mathematics

MDPI

Article

Using Negative Power Transformation to Model Block Minima

Thanawan Prahadchai [1,2], Piyapatr Busababodhin [3,4], Taeyong Kwon [5] and Sanghoo Yoon [2,5,*]

[1] Department of Statistics, Khon Kaen University, Khon Kaen 40002, Thailand; thanpra@kku.ac.th
[2] Department of Mathematics and Statistics, Chonnam National University, Gwangju 61186, Republic of Korea
[3] Digital Innovation Research Cluster for Integrated Disaster Management in the Watershed, Mahasarakham University, Maha Sarakham 44150, Thailand; piyapatr.b@msu.ac.th
[4] Department of Mathematics, Mahasarakham University, Maha Sarakham 44150, Thailand
[5] G-LAMP, Chonnam National University, Gwangju 61186, Republic of Korea; tdstat91@jnu.ac.kr
* Correspondence: statstar77@jnu.ac.kr

Abstract

This study proposes a novel transformation method for analyzing block minima using the generalized extreme value distribution (GEVD). The negative power transformation (NPT), which includes a tunable hyperparameter and reduces to the reciprocal transformation (RT) when set to 1, improves the accuracy and robustness in estimating long-term return levels (RL). Compared to traditional methods, the NPT-GEVD demonstrates lower bias, standard errors, and root mean square errors in Monte Carlo simulations. Furthermore, the NPT-GEVD provides consistent RL estimates with improved robustness across varying parameterizations and sample sizes, mainly when using L-moments for small datasets. The application of the NPT-GEVD to rainfall data from South Korea revealed that the RLs for detecting hourly cumulative rainfall threshold levels varied from 30 min to over 4 h, depending on the location and threshold. This research underscores the value of advanced transformation techniques in environmental risk management, offering critical insights for flood prediction and mitigation strategies in climate change.

Keywords: inter-amount time; negative power transformation; reciprocal transformation; L-moment estimation; flood

MSC: 62P12; 62G32

1. Introduction

1.1. A Brief Literature Survey

Several transformation-based approaches have been proposed to enhance the flexibility of the generalized extreme value distribution (GEVD). Early work introduced the conventional transformation (CT), beginning with [1] and later applied by [2], which enabled the unified modeling of block maxima and minima. Subsequent developments sought to improve the estimation stability for environmental extremes, such as the refinement of the transformation accuracy for hydrometeorological applications [3,4] and inverse transformation methods incorporating the extremal index [5]. Additional contributions, including the inverse Pareto formulations [6,7] and Weibull-based inverse distributions [8,9], demonstrated improved adaptability for bounded or highly skewed environmental variables.

Further transformation families, such as the alpha-power transformation [10] and the Box–Cox class [11,12], have been widely applied in environmental and financial extremes [13,14]. However, these transformation methods were fundamentally developed for block maxima or for normalizing the main body of a distribution, rather than

Check for updates

Academic Editor: Indranil Ghosh

Received: 6 March 2026
Revised: 5 April 2026
Accepted: 8 April 2026
Published: 20 April 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

preserving the tail asymptotics required in extreme value theory [2,3]. Furthermore, recent studies in extreme hydrological modeling emphasize that applying arbitrary power-based approaches like the Box–Cox transformation to the GEVD framework is highly problematic for physically bounded data [8,15]. It disrupts the max-stability postulate, alters the theoretical domain of attraction, and requires transforming the GEVD through the logarithm of its cumulative distribution function (CDF), which eliminates the distribution's analytical closed-form structure. Such generic transformations do not naturally map the true physical lower bounds (zero) of hydrological data, such as inter-event durations, to the domain boundaries of the GEVD. Consequently, methods that cannot theoretically guarantee strictly positive support yield unrealistic predictive limits, revealing a persistent methodological gap that requires a targeted analysis of non-negative block minima. For these theoretical reasons, the Box–Cox transformation is not considered as a competing method in this study.

1.2. Research Motivation

Extreme value theory, especially the GEVD, is widely used to model block maxima such as annual maximum rainfall or wind speeds. However, in many hydrological and meteorological applications, block minima such as the minimum time between heavy rainfall events are equally important for flood risk assessments. A common approach to modeling minima is to multiply the data by -1, converting the problem into a maxima framework. This method, known as the conventional transformation GEVD (CT-GEVD) [1–3], enables the use of existing tools and theory.

However, the CT-GEVD method inherently possesses a theoretical flaw that can yield unrealistic negative results. The fundamental cause stems from its simple linear mapping ($Y = -X$). While this negative sign reversal mathematically aligns the data with a block maxima framework, it fails to encode the physical boundary constraint ($X > 0$) into the GEVD support. Specifically, when the shape parameter $\xi < 0$ (the Weibull domain), the transformed GEVD possesses a finite upper bound [2]. During parameter estimation with high data variability, this theoretical bound can easily cross the zero threshold when transformed back to the original scale. Consequently, the model occasionally produces negative return level estimates, rendering the predictions physically impossible. This phenomenon is clearly corroborated by the comprehensive empirical comparison provided in the Supplementary Information (SI); notably, Supplementary Information Figure S5 illustrates that the CT-GEVD model yields physically impossible negative RLs when modeling the minimum inter-amount time (IAT) at the 110 mm threshold in South Korea (also summarized in Table 1).

To theoretically resolve this, it is necessary to employ nonlinear transformations that can asymptotically map the zero boundary to infinity, strictly enforcing the non-negative characteristics of the block minima [9]. Two such methods are the reciprocal transformation (RT) and the proposed negative power transformation (NPT). While the RT provides a fixed mapping (equivalent to NPT with $\alpha = 1$), estimating the hyperparameter α in the NPT provides flexible, data-driven scaling. In practice, this allows the transformation to dynamically adapt to varying degrees of skewness and specific tail behaviors near the zero boundary across different environmental datasets. Both of these approaches prevent negative outputs by structurally bounding the inverse variables. These approaches preserve the extreme tail properties without violating the GEVD assumptions, offering more robust and physically meaningful estimates for extreme hydrological minima.

Therefore, comparing the proposed NPT against the CT and RT models provides the most logically relevant evaluation framework for this study. The CT serves as the traditional linear baseline, highlighting the pitfalls of boundary violations, while the RT represents the simplest fixed-boundary alternative. Evaluating these three models systematically

isolates the specific theoretical benefits of both structural bound preservation and flexible shape adaptation.

Table 1. The results of fitting the GEVD for transformed minima values using the traditional transformation method, including return level estimation for IAT data in Korea with a 110 mm threshold at various stations.

ID	Location	$\hat{\mu}$ (SE)	$\hat{\sigma}$ (SE)	$\hat{\xi}$ (SE)	CvM (p-Value)	RL25 (L, U)	RL100 (L, U)
112	Incheon	−22.86 (7.34)	30.96 (6.68)	−1.31 (0.22)	0.19 (0.29)	−0.42 (−3.46, 2.63)	−0.72 (−4.59, 3.15)
133	Daejeon	−24.41 (8.83)	38.62 (10.62)	−1.55 (0.23)	0.20 (0.27)	−0.37 (−3.68, 2.93)	−0.53 (−4.40, 3.34)
235	Boryeong	−30.35 (9.80)	49.20 (13.97)	−1.60 (0.26)	0.14 (0.44)	−0.19 (−3.87, 3.5)	−0.35 (−4.73, 4.03)
247	Namwon	−22.73 (9.94)	39.58 (16.59)	−1.96 (0.21)	0.27 (0.16)	2.56 (−0.06, 5.19)	2.53 (−0.46, 5.52)
277	Yeongdeok	−42.10 (12.77)	59.45 (13.84)	−1.39 (0.21)	0.10 (0.57)	−0.17 (−5.65, 5.31)	−0.60 (−7.20, 6.01)
285	Hapcheon	−19.08 (6.68)	27.52 (8.02)	−1.58 (0.21)	0.17 (0.33)	1.80 (−0.40, 4.00)	1.70 (−0.77, 4.17)

Note: CvM refers to the Cramér–von Mises goodness-of-fit test, while RL25 and RL100 represent the return levels for 25- and 100-year return periods, respectively. SE denotes the standard error, and L and U refer to the lower and upper bounds of the RL.

1.3. Objectives of This Research

This study aims to evaluate the performance of GEVD models for block minima using three transformation approaches: the CT, NPT, and RT. Model accuracy is assessed using bias, the standard error (SE), and the root mean square error (RMSE). Parameter estimation is carried out using both maximum likelihood estimation (MLE) and L-moments (LM), with model adequacy verified through the Cramér–von Mises (CvM) goodness-of-fit test. Return levels for 20-, 50-, and 100-year periods are calculated based on the estimated parameters. Additionally, we apply our proposed technique to real-world data, namely minimum IAT observations in South Korea, to support flood risk mitigation.

Section 2 explains the GEVD for block minima, including transformation methods, and the parameter and return level estimation methodologies. Section 3 discusses the simulation scenarios and evaluates each distribution. Section 4 presents the results of fitting the GEVD to IAT data. Section 5 contains the discussion, followed by the conclusions in Section 6. The SI contains technical details, tables, and figures.

2. Methodology

2.1. Generalized Extreme Value Distribution (GEVD)

In this study, the GEVD is used to analyze the data; it is asymptotically appropriate and is widely applied to analyze univariate extreme values [1,2]. Assume that $X = \max(X_1, \cdots, X_n)$. The CDF of the GEVD is given by

$$F_X(x) = P(X \le x) = \begin{cases} \exp\left(-\left(1+\xi\frac{x-\mu}{\sigma}\right)^{-1/\xi}\right), & \text{if } \xi \neq 0 \\ \exp\left(-\exp\left(-\frac{x-\mu}{\sigma}\right)\right), & \text{if } \xi = 0 \end{cases} \tag{1}$$

where $1+\xi\frac{x-\mu}{\sigma} > 0$, and μ, σ, and ξ represent the location, scale, and shape parameters, respectively. When $\xi = 0$, the GEV distribution simplifies to the Gumbel distribution (GD). The corresponding PDF is expressed as

$$f(x) = \begin{cases} \frac{1}{\sigma}\left(1+\xi\frac{x-\mu}{\sigma}\right)^{-1/\xi-1}\exp\left(-\left(1+\xi\frac{x-\mu}{\sigma}\right)^{-1/\xi}\right), & \text{for } \xi \neq 0 \\ \frac{1}{\sigma}\exp\left(\frac{x-\mu}{\sigma}\right)\exp\left(-\exp\left(\frac{x-\mu}{\sigma}\right)\right), & \text{for } \xi = 0 \end{cases} \tag{2}$$

2.2. GEVD for Block Minima

This study applies the GEVD to model block minima, focusing on extremely small observations. The data are transformed to align with the GEVD framework, enabling an exploration of the distribution's heavy-tailed properties. The block minima approach is analogous to the block maxima method; however, the minimum value is selected within each block instead of the maximum. It is essential to verify the form and constraints of the distribution following the transformation to ensure compatibility with the GEVD model.

Let $X \sim GEV(\mu, \sigma, \xi)$. To investigate the effects of different transformation schemes on the distributional form and parameter behavior, three monotonic transformations are considered.

2.2.1. Conventional Transformation

The conventional transformation introduces a negative sign to the original minimum data, as suggested by [1–3,16]. Let $Y_i = -X_i$ for $i = 1, ..., n$. If $\tilde{M}_n = \min(X_1, ..., X_n)$ and $M_n = \max(Y_1, ..., Y_n)$, then $\tilde{M}_n = -M_n$. Thus, for large n, we have

$$\begin{aligned} P(\tilde{M}_n \leq y) &= P(-M_n \leq y) \\ &= P(M_n \geq y) \\ &= 1 - P(M_n \leq -y) \\ &= 1 - F_X(-y) \end{aligned} \tag{3}$$

Thus, the CDF of transformed data Y under the conventional transformation of CT-GEVD is

$$F_Y^{CT}(y) = \begin{cases} 1 - \exp\left\{ -\left[1 - \xi'\left(\frac{y-\tilde{\mu}'}{\sigma'}\right)\right]^{-1/\xi'} \right\}, & \text{if } \xi' \neq 0, \\ 1 - \exp\left\{ -\exp\left(\frac{y-\tilde{\mu}'}{\sigma'}\right) \right\}, & \text{if } \xi' = 0 \end{cases} \tag{4}$$

for large n. The corresponding CT-PDF of Y is obtained by taking the derivative of F_Y^{CT}, dF_Y^{CT}/dy, and is expressed as follows:

$$f_Y^{CT}(y) = \begin{cases} \frac{1}{\sigma'}\left[1 - \xi'\left(\frac{y-\tilde{\mu}'}{\sigma'}\right)\right]^{-1/\xi'-1} \exp\left\{ -\left[1 - \xi\left(\frac{y-\tilde{\mu}'}{\sigma'}\right)\right]^{-1/\xi'} \right\}, & \text{if } \xi' \neq 0, \\ \frac{1}{\sigma'}\exp\left(\frac{y-\tilde{\mu}'}{\sigma'}\right) \exp\left\{ -\exp\left(\frac{y-\tilde{\mu}'}{\sigma'}\right) \right\}, & \text{if } \xi' = 0. \end{cases} \tag{5}$$

This formulation leads to the operationally transformed parameters $(\tilde{\mu}', \sigma', \xi')$. To avoid any ambiguity in the parameter definitions, we explicitly define the correspondence between the underlying distribution of the block maxima and the transformed parameters. Specifically, the negative sign reversal $Y = -X$ shifts the location parameter such that $\tilde{\mu}' = -\mu'$, where μ' represents the location parameter of the underlying block maxima. The scale and shape parameters, σ' and ξ', remain structurally consistent with the GEVD form. The cumulative probability $F_Y^{CT}(y)$ is strictly defined on the domain $[y : 1 - \xi'(y - \tilde{\mu}')/\sigma' > 0]$.

2.2.2. Negative Power Transformation

The NPT-GEVD model is proposed to address the issue of negative values arising in high quantile estimates from the CT-GEVD model. This is achieved by shifting the data to a positive lower bound using a hyperparameter (α). Let $Y_i = (X_i)^{-\alpha}$, where X_i is a random variable for $i = 1, ..., n$ and $\alpha > 0$.

To ensure that the transformed variable Y falls within a valid support for GEVD modeling, we require that all original values X_i be strictly positive. When the original data include values close to zero, a small positive constant (e.g., $\epsilon = 0.01$) is added to

avoid extreme behavior or undefined transformation outcomes. This adjustment helps to maintain a proper domain for the transformed distribution and ensures that the shape parameter ξ remains interpretable under GEVD assumptions.

If $\tilde{M}_n = \min\{(X_1)^{-\alpha}, ..., (X_n)^{-\alpha}\}$ and $M_n = \max(Y_1, ..., Y_n)$, then $\tilde{M}_n = (M_n)^{-\alpha}$. Thus, for large n, we have

$$\begin{aligned} P(\tilde{M}_n \leq y) &= P((M_n)^{-\alpha} \leq y) \\ &= P(M_n \geq y^{-\frac{1}{\alpha}}) \\ &= 1 - P(M_n \leq y^{-\frac{1}{\alpha}}) \\ &= 1 - F_X(y^{-\frac{1}{\alpha}}) \end{aligned} \tag{6}$$

Thus, the CDF of transformed data Y under the NPT-GEVD is given by

$$F_Y^{NPT}(y) = \begin{cases} 1 - \exp\left\{-\left[1 + \xi' w_y^{(\alpha)}\right]^{-1/\xi'}\right\}, & \text{if } \xi' \neq 0, \\ 1 - \exp\left\{-\exp\left(-w_y^{(\alpha)}\right)\right\}, & \text{if } \xi' = 0 \end{cases} \tag{7}$$

leading to transformed parameters (μ', σ', ξ'), where $1 + \xi' w_y^{(\alpha)} > 0$ and $w_y^{(\alpha)} = \frac{y^{-\frac{1}{\alpha}} - \mu'}{\sigma'}$. The corresponding NPT-PDF is obtained by differentiating Y obtained by taking the derivative of F_Y^{NPT}, dF_Y^{NPT}/dy, and is expressed as follows:

$$f_Y^{NPT}(y) = \begin{cases} \frac{y^{-\frac{1}{\alpha}-1}}{\alpha\sigma'}\left[1 + \xi' w_y^{(\alpha)}\right]^{-1/\xi'-1} \exp\left\{-\left[1 + \xi' w_y^{(\alpha)}\right]^{-1/\xi'}\right\}, & \text{if } \xi' \neq 0, \\ \frac{y^{-\frac{1}{\alpha}-1}}{\alpha\sigma'} \exp\left(-w_y^{(\alpha)}\right) \exp\left\{-\exp\left(-w_y^{(\alpha)}\right)\right\}, & \text{if } \xi' = 0. \end{cases} \tag{8}$$

To ensure the validity of the GEVD after transformation, we impose $\mu' > 0$, $\sigma' > 0$, and $-\infty < \xi' < \infty$. These constraints maintain the positivity of the transformed variable Y, which is especially important when modeling block minima. Additionally, we monitor the transformed shape parameter ξ' to ensure that it remains within a range consistent with theoretical expectations for the GEVD. Cases where $\xi' \geq 0$ are interpreted carefully, considering the extended upper support of $Y = X^{-\alpha}$ when X approaches zero. The following evidence demonstrates that the CDF of the NPT-GEVD for minima is as presented:

$$\tilde{H}(y) = 1 - \exp\left\{-\left[1 + \xi' w_y^{(\alpha)}\right]^{-1/\xi'}\right\} \tag{9}$$

defined on $1 + \xi' w_y^{(\alpha)} > 0$. There exist sequences of constants $\{a_n > 0\}$ and $\{b_n\}$ such that

$$P\left(\frac{\tilde{M}_n - b_n}{a_n} \leq y\right) \to \tilde{H}(y) \quad \text{as } n \to \infty \tag{10}$$

distributions for minima limit distributions as $n \to \infty$.

As a visual verification of this result, Figure 1 compares the theoretical CDF (solid lines) of the NPT-GEVD with the empirical CDF (points) obtained from randomly generated samples of size n = 100 under four parameter configurations ($\xi' = 0.25, -0.25$; $\alpha = 1, 2$). The close agreement across all four panels confirms the validity of the derived CDF in Equation (7) and supports the asymptotic convergence stated in Equation (10).

2.2.3. Reciprocal Transformation

For the reciprocal transformation (RT) method, the distribution can be obtained by setting the hyperparameter to $\alpha = 1$ in the NPT-GEVD. Consequently, the CDF of the RT-GEVD can be expressed in the terms of the NPT-GEVD from Equation (7) as follows:

$$F_Y^{RT}(y) = \begin{cases} 1 - \exp\left\{-\left[1 + \xi' w_y^{(1)}\right]^{-1/\xi'}\right\}, & \text{if } \xi' \neq 0, \\ 1 - \exp\left\{-\exp\left(-w_y^{(1)}\right)\right\}, & \text{if } \xi' = 0 \end{cases} \tag{11}$$

leading to transformed parameters (μ', σ', ξ'), where $1 + \xi' w_y^{(1)} > 0$ and $w_y^{(1)} = \frac{y^{-1} - \mu'}{\sigma'}$.

Here, the parameters satisfy $\mu' > 0, \sigma' > 0$, and $-\infty < \xi' < \infty$. Since $F_Y^{RT}(y)$ ranges from 0 to a value strictly less than 1 and does not remain constant over any interval, it indicates that $F_Y^{RT}(y)$ does not collapse to a single point or stay constant over any non-trivial interval as y increases. Thus, the distribution of Y is non-degenerate for any finite values of (μ', σ', ξ'). This confirms that the RT-GEVD belongs to the GEV distribution family for the minima limit distribution as $n \to \infty$. The PDF of Y can be obtained by differentiating the CDF, F_Y^{RT}:

$$f_Y^{RT}(y) = \begin{cases} \frac{1}{\sigma' y^2}\left[1 + \xi' w_y^{(1)}\right]^{-1/\xi' - 1} \exp\left\{-\left[1 + \xi' w_y^{(1)}\right]^{-1/\xi'}\right\}, & \text{if } \xi' \neq 0, \\ \frac{1}{\sigma' y^2} \exp\left(-w_y^{(1)}\right) \exp\left\{-\exp\left(-w_y^{(1)}\right)\right\}, & \text{if } \xi' = 0. \end{cases} \tag{12}$$

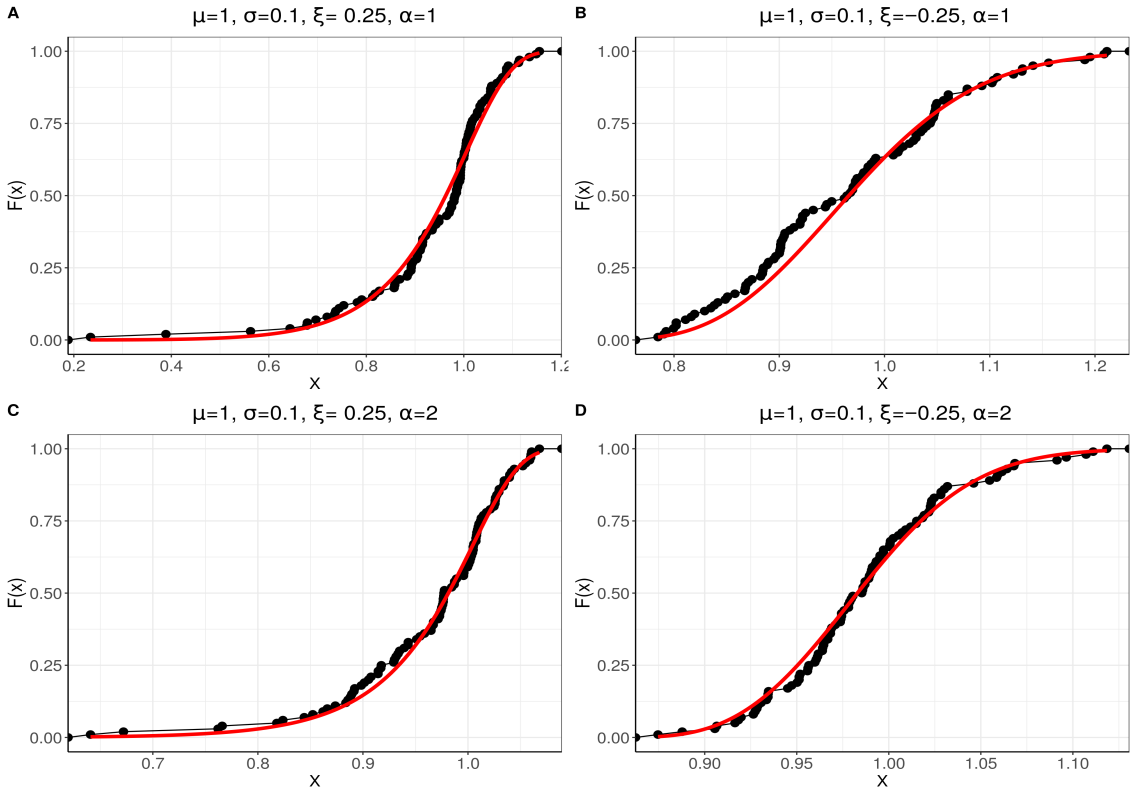


Figure 1. Cumulative density function plots for the NPT-GEVD with randomly generated samples. Parameters used: $\mu' = 1$, $\sigma' = 0.1$, $\xi' = 0.25$ (**A**,**C**), $\xi' = -0.25$ (**B**,**D**), and $\alpha = 1$ (top) and $\alpha = 2$ (below), with $n = 100$. The plots compare the theoretical (solid line) and empirical (points) CDFs.

2.3. Parameter Estimation

2.3.1. Method of Maximum Likelihood

The MLE is obtained by maximizing the function $L(\theta|y_i) = \prod_{i=1}^{n} f(y_i;\theta)$, where $\theta = \{\mu', \sigma', \xi'\}$. Directly maximizing the likelihood function can be challenging due to a product term. To simplify, the log-likelihood function is used, as it converts the product into a sum, and it is expressed as $l(\theta|y_i) = \sum_{i=1}^{n} \log(f(y_i;\theta))$.

For a sample of data y_i from a CT-GEVD where $\xi \neq 0$, the log-likelihood function is given by

$$\begin{aligned} l^{CT}(\theta|y_i) = & -m\log(\sigma') - \left(\frac{1}{\xi'}+1\right)\sum_{i=1}^{m}\log\left[1-\xi'\left(\frac{y_i-\tilde{\mu}'}{\sigma'}\right)\right] \\ & -\sum_{i=1}^{m}\left[1-\xi'\left(\frac{y_i-\tilde{\mu}'}{\sigma'}\right)\right]^{-1/\xi'} \end{aligned} \tag{13}$$

subject to $1-\xi'\left(\frac{y_i-\tilde{\mu}'}{\sigma'}\right) > 0$ for $i = 1, \ldots, n$ and $\sigma' > 0$.

For a sample of data y_i corresponding to the NPT-GEVD where $\xi' \neq 0$, the log-likelihood function is expressed as

$$\begin{aligned} l^{NPT}(\theta|y_i) = & -n\log(\alpha\sigma') - \left(\frac{1}{\alpha}+1\right)\sum_{i=1}^{n}\log(y_i) \\ & -\left(\frac{1}{\xi'}+1\right)\sum_{i=1}^{n}\log\left(1+\xi'\left(\frac{y_i^{-\frac{1}{\alpha}}-\mu'}{\sigma'}\right)\right) \\ & -\sum_{i=1}^{n}\left[1+\xi'\left(\frac{y_i^{-\frac{1}{\alpha}}-\mu'}{\sigma'}\right)\right]^{-1/\xi'}, \end{aligned} \tag{14}$$

subject to $1+\xi'\left(\frac{y_i^{-\frac{1}{\alpha}}-\mu'}{\sigma'}\right) > 0$ for $i = 1, \ldots, n, \sigma' > 0$, and the hyperparameter $\alpha > 0$.

When $\alpha = 1$, Equation (14) reduces to the log-likelihood for the RT-GEVD:

$$\begin{aligned} l^{RT}(\theta|y_i) = & -n\log(\sigma') - 2\sum_{i=1}^{n}\log(y_i) \\ & -\left(\frac{1}{\xi'}+1\right)\sum_{i=1}^{n}\log\left(1+\xi'\left(\frac{y_i^{-1}-\mu'}{\sigma'}\right)\right) \\ & -\sum_{i=1}^{n}\left[1+\xi'\left(\frac{y_i^{-1}-\mu'}{\sigma'}\right)\right]^{-1/\xi'}, \end{aligned} \tag{15}$$

subject to $1+\xi'\left(\frac{y_i^{-1}-\mu'}{\sigma'}\right) > 0$ for $i = 1, \ldots, n$ and $\sigma' > 0$.

To maximize Equations (13)–(15), the partial derivatives are set to zero:

$$\frac{\partial l(\theta|y_i)}{\partial \mu'} = 0, \quad \frac{\partial l(\theta|y_i)}{\partial \sigma'} = 0, \quad \frac{\partial l(\theta|y_i)}{\partial \xi'} = 0. \tag{16}$$

The hyperparameter is fixed to a positive value, and the 'optim' function in the R program is used for numerical optimization, supported by the 'stats' and 'ismev' packages to perform the numerical optimization for this function [2,17].

2.3.2. Method of L-Moments

L-moments (LM) are analogous to conventional moments but represent linear combinations of probability-weighted moments (PWMs) [18]. The PWM can be defined using the CDF of a random variable z [19]:

$$\beta_r = E\left\{z\left[F_Z(z)\right]^r\right\} = \int_0^1 z(F)F^r dF, \qquad r = 0, 1, \ldots \tag{17}$$

where β_r is the r-th-order PWM and $z(F)$ represents the inverse CDF (quantile function). The first four L-moments (λ_r) are related to PWMs as follows:

$$\begin{aligned} \lambda_1 &= \beta_0, \\ \lambda_2 &= 2\beta_1 - \beta_0, \\ \lambda_3 &= 6\beta_2 - 6\beta_1 + \beta_0, \\ \lambda_4 &= 20\beta_3 - 30\beta_2 + 12\beta_1 - \beta_0. \end{aligned} \tag{18}$$

Similar to the log-likelihood formulation, the L-moments can be explicitly derived for the specific generalized extreme value distributions. To derive the L-moments, we first determine the inverse distribution function $Y(F)$ for the respective distributions and substitute it into the PWM integral $\beta_r = \int_0^1 Y(F)F^r dF$.

Based on the PWMs, the first four L-moments for the CT-GEVD ($\xi \neq 0$) are given by

$$\begin{aligned} \lambda_1 &= \tilde{\mu} + \sigma[1 - \Gamma(1-\xi)], \\ \lambda_2 &= \frac{\sigma}{\xi}\Gamma(1-\xi)(2^{\xi-1} - 1), \\ \lambda_3 &= \frac{\sigma}{\xi}\Gamma(1-\xi)[3(2^{\xi-1}) - 3^{\xi-1} - 1], \\ \lambda_4 &= \frac{\sigma}{\xi}\Gamma(1-\xi)[4^{\xi-1} - 6(3^{\xi-1}) + 4(2^{\xi-1}) - 1]. \end{aligned} \tag{19}$$

For the NPT-GEVD, the inverse distribution function $Y^{NPT}(F)$ can be written as

$$Y^{NPT}(F) = \left[\mu' + \frac{\sigma'}{\xi'}\left((-\log(1-F))^{-\xi'} - 1\right)\right]^{-\alpha}. \tag{20}$$

Therefore, the r-th order of the PWMs is evaluated as $\beta_r = \int_0^1 Y^{NPT}(F)F^r dF$. By applying the substitution $u = -\log(1-F)$, where $F = 1 - e^{-u}$ and $dF = e^{-u}du$, and expanding the binomial series $(1-e^{-u})^r = \sum_{k=0}^r \binom{r}{k}(-1)^k e^{-ku}$, the integral can be reformulated as

$$\beta_r = \sum_{k=0}^{r} \binom{r}{k}(-1)^k \int_0^\infty \left[\mu' + \frac{\sigma'}{\xi'}(u^{-\xi'} - 1)\right]^{-\alpha} e^{-(k+1)u} du. \tag{21}$$

For the approximation representing extreme behavior ($u \to 0$), the integrand is dominated by the term $(\frac{\xi'}{\sigma'})^\alpha u^{\alpha\xi'}$. Applying the Gamma function $\Gamma(\cdot)$ definition $\int_0^\infty u^{\alpha\xi'} e^{-(k+1)u} du = \frac{\Gamma(\alpha\xi'+1)}{(k+1)^{\alpha\xi'+1}}$ to this approximation, we obtain

$$\beta_r \approx \left(\frac{\xi'}{\sigma'}\right)^\alpha \Gamma(\alpha\xi' + 1) \sum_{k=0}^{r} \binom{r}{k} \frac{(-1)^k}{(k+1)^{\alpha\xi'+1}}. \tag{22}$$

By substituting these evaluated β_r values into the linear combinations defined in Equation (18), the exact L-moments can be theoretically constructed. Under these necessary asymptotic approximations, the first four L-moments for the NPT-GEVD can be expressed as

$$\begin{aligned}
\lambda_1 &= \left(\frac{\xi}{\sigma}\right)^{\alpha}\Gamma(\alpha\xi+1),\\
\lambda_2 &= \left(\frac{\xi}{\sigma}\right)^{\alpha}\Gamma(\alpha\xi+1)\left(1-\frac{1}{2^{\alpha\xi}}\right),\\
\lambda_3 &= \left(\frac{\xi}{\sigma}\right)^{\alpha}\Gamma(\alpha\xi+1)\left(\frac{5}{2}-\frac{2}{3^{\alpha\xi}}\right),\\
\lambda_4 &= \left(\frac{\xi}{\sigma}\right)^{\alpha}\Gamma(\alpha\xi+1)\left(1-\frac{12}{2^{\alpha\xi+1}}+\frac{30}{3^{\alpha\xi+1}}-\frac{20}{4^{\alpha\xi+1}}\right).
\end{aligned} \tag{23}$$

The RT-GEVD is a special case of the NPT-GEVD where the hyperparameter $\alpha = 1$. Consequently, its L-moments simplify to

$$\begin{aligned}
\lambda_1 &= \frac{\xi}{\sigma}\Gamma(\xi+1),\\
\lambda_2 &= \frac{\xi}{\sigma}\Gamma(\xi+1)\left(1-\frac{1}{2^{\xi}}\right),\\
\lambda_3 &= \frac{\xi}{\sigma}\Gamma(\xi+1)\left(\frac{5}{2}-\frac{2}{3^{\xi}}\right),\\
\lambda_4 &= \frac{\xi}{\sigma}\Gamma(\xi+1)\left(1-\frac{12}{2^{\xi+1}}+\frac{30}{3^{\xi+1}}-\frac{20}{4^{\xi+1}}\right).
\end{aligned} \tag{24}$$

Ref. [20] defined the unbiased sample estimator of β_r as

$$b_r = \frac{1}{n}\binom{n-1}{r}^{-1}\sum_{j=1}^{n}\binom{j-1}{r}z_j, \tag{25}$$

where z_j is the ordered sample $z_1 \le z_2 \le \cdots \le z_n$. The sample L-moment ratios are then defined as $\tau_r = \lambda_r/\lambda_2$ for $r \ge 3$. For GEVD parameter estimation, the 'optim' function in R, along with the 'stats' package, can be used to address numerical estimation issues by matching theoretical L-moments with sample estimates [17].

2.4. Return Level

Setting $F(z) = 1 - p = 1 - 1/T$ and solving for z yields the quantile value $(1 - p)$, also known as the return level (RL). This value represents the threshold expected to be exceeded once over a specified return period T. For block maxima in hydrology, such as rainfall or temperature data, the RL for the GEVD [2] is

$$RL(T) = \begin{cases} \mu - \dfrac{\sigma}{\xi}\left\{1-\left[-\log\left(1-\dfrac{1}{T}\right)\right]^{-\xi}\right\} & \text{for } \xi \neq 0,\\ \mu - \sigma\log\left[-\log\left(1-\dfrac{1}{T}\right)\right] & \text{for } \xi = 0. \end{cases} \tag{26}$$

For block minima, transformations are applied (Section 2.2), yielding the following RL forms:

$$RL(T)^{CT} = \tilde{\mu}' + \frac{\sigma'}{\xi'}\left\{1-\left[-\log\left(1-\frac{1}{T}\right)\right]^{-\xi'}\right\}, \tag{27}$$

$$RL(T)^{NPT} = \left\{\mu' - \frac{\sigma'}{\xi'}\left[1-\left(-\log\left(1-\frac{1}{T}\right)\right)^{-\xi'}\right]\right\}^{-\alpha}, \tag{28}$$

$$RL(T)^{RT} = \left\{ \mu' - \frac{\sigma'}{\xi'} \left[1 - \left(-\log\left(1 - \frac{1}{T}\right)\right)^{-\xi'} \right] \right\}^{-1}. \tag{29}$$

These expressions are discussed in [1–3,16]. The Delta and parametric bootstrap methods are used to approximate the variance of the RL based on the parameter variance $\hat{\theta}$ from estimation methods.

2.5. Goodness-of-Fit Test

The Cramér–von Mises (CvM) test evaluates the goodness of fit of transformed data before parameter estimation. Given a sorted sample $z_1, z_2, \ldots, z_n$, the CvM statistic W^2 is defined as

$$W^2 = n \int_{-\infty}^{+\infty} [F_n(z) - F(z)]^2 dF(z), \tag{30}$$

where $F_n(z)$ is empirical distribution function (EDF) of the sample. The test statistic W^2 measures the squared difference between the EDF and the theoretical CDF, integrated across the full range of the distribution. A smaller value of W^2 indicates a closer fit between the empirical and theoretical distributions.

2.6. Estimation of Hyperparameter (α)

The complexity of likelihood functions in the newly proposed distribution makes MLE computationally intensive. It often requires advanced optimization techniques, which can overestimate GEV parameters and exhibit sensitivity to initial estimates, leading to inaccuracies. We utilize parameter estimation methods, including MLE and LM, alongside goodness-of-fit (GOF) statistic tests to address these challenges and evaluate the model fit.

For estimating the hyperparameter α in the NPT-GEVD for minima, GOF statistics are minimized using the Cramér–von Mises (CvM) test. From a statistical theory perspective, the CvM test is fundamentally more suitable for this model than the Anderson–Darling (AD) test. The AD statistic is designed to assign heavier weights to the tails of the distribution [21]. In the context of extreme block minima with small sample sizes, empirical tails are sparse and exhibit high variance. Consequently, optimizing a global transformation hyperparameter based on the AD test causes the objective function to become numerically unstable and highly sensitive to outliers [22].

Conversely, the CvM statistic applies uniform weighting across the empirical distribution function, measuring the global integrated squared error. This property ensures a more robust, theoretically stable optimization landscape for determining the structural hyperparameter α before the MLE or LM methods are used to estimate the main distribution parameters. Using the provisionally estimated parameters, the CDF of the GEVD for minima is computed, and the robust CvM statistic is minimized to determine α. In practice, this CvM-based optimization allows α to be selected dynamically, yielding a custom transformation curve that optimally fits the empirical distribution of the given sample, rather than relying on the rigid assumption of $\alpha = 1$.

When sensitivity to initial estimates produced unrealistic parameter values, a resampling technique was applied with a reduced sample size. The 50 most optimal hyperparameter estimates were identified using the median. The complete algorithm is detailed in Algorithm 1.

Algorithm 1 Resampling Algorithm for Hyperparameter Estimation of NPT-GEVD.

Input:

- The true value of hyperparameter α.
- A generated random sample $y \sim$ NPT-GEVD of size n with the true α.
- Number of bootstrap samples $B = 50$.
- GOF optimization functions.

Output:

- Final robust estimate of hyperparameter α.

Steps:

1: **Initialize:** Create an empty list $\alpha^{[b]}$ to store the estimates of α.
2: **for** for $b = 1$ to B **do**
   a. Generate a random sample y_b^* from y with a smaller sample size $n^* = n \times 0.8$ (sampling without replacement).
   b. Perform GOF optimization as follows (refer to Section 2.6 for detailed steps):
      i. Estimate the parameters μ_b, σ_b, and ξ_b using MLE and LM methods on y_b^*.
      ii. Calculate the CvM statistic using the empirical data (y_b^*) and the theoretical CDF of (y).
      iii. Refine α using GOF optimization on y_b^*, initializing α^* with α.
   c. Store the refined value of α_b in the list $\alpha^{[b]}$.
3: **end for**
4: Construct a distribution of the estimated hyperparameter $\alpha^{[b]}$ where $b = 1, \ldots, B$, and compute the median of this distribution.
5: **Return** the final robust estimate of α.

3. Monte Carlo Simulation

3.1. Simulation Framework

Our experiments revealed that the performance of the NPT method deteriorates when the parameters μ' and σ' of the transformed GEVD are large. While conventional standardization (a mean of 0 and a variance of 1) is common practice, it poses significant challenges for heavy-tailed generalized extreme value distributions. Transformations involving negative powers (e.g., $y^{-\alpha}$) become numerically unstable and exhibit exploding gradients as $y \to 0^+$ [2].

To ensure strict positivity and numerical stability during parameter optimization, the transformed variables must remain well separated from zero [23]. Furthermore, scaling the data to a mean of 1 before calculating L-moments prevents artificial variance inflation, enabling highly stable parameter estimation [18,24,25].

Therefore, we empirically rescaled the data to have a mean of approximately 1 and a small standard deviation of about 0.1. This targeted normalization provides a tight, numerically stable domain for the GEVD, keeping all values safely away from the singularity at zero and significantly enhancing the reliability of the transformed distribution [26].

To evaluate the performance of three transformed GEVD(s) for block minima such as conventional, reciprocal, and negative power, we conducted a Monte Carlo (MC) simulation study. We generated 1000 random samples from the NPT-GEVD using known quantiles, fixing the hyperparameter at values of 0.5, 1, 1.5, and 2. The simulation included different scenarios focusing on the shape parameter $\xi' = -0.5, -0.25, -0.05, 0.05, 0.25, 0.5$ to cover fundamental tail behaviors, with sample sizes (n) of 30, 100, and 500. The location and scale parameters were deliberately fixed at $\mu' = 1$ and $\sigma' = 0.1$ to ensure strict positivity and numerical stability, as justified previously. While the main text highlights the key findings from these extensive parameter variations, the complete and exhaustive simulation results across all evaluated ranges of ξ' and α are provided in the SI. Parameter estimation was

performed using the MLE and LM methods, and return levels for 25, 50, and 100 years were calculated.

3.2. Model Evaluation

The performance of each method was evaluated using the bias, SE, and RMSE:

$$\text{Bias} = \bar{\hat{RL}}(T) - RL(T), \tag{31}$$

$$\text{SE} = \sqrt{Var(\hat{RL}(T))}, \tag{32}$$

$$\text{RMSE} = \sqrt{[Bias^2 + SE^2]}. \tag{33}$$

Here, $\hat{RL}(T)$ and $RL(T)$ represent the estimated and true RLs, respectively, while $\bar{\hat{RL}}(T) = \frac{1}{N}\sum_{i=1}^{N}\hat{RL}(T,i)$ denotes the average estimated RL from N successful convergences across 1000 trials. Lower RMSE and SE values and bias values closer to zero indicate better performance.

3.3. Simulation Results

Tables 2–5 summarize the MC simulation results for $\mu' = 1$, $\sigma' = 0.1$, and varying ξ', with $\alpha = 0.5$ (in Tables 2 and 3) and $\alpha = 1$ (in Tables 4 and 5). These simulations assessed return levels (RL25, RL50, RL100) for block minima using three GEVDs: CT-GEVD, NPT-GEVD, and RT-GEVD. MLE-based methods (CT.MLE, NPT.MLE, RT.MLE) and LM-based methods (CT.LM, NPT.LM, RT.LM) were compared. For larger sample sizes ($n = 500$), the NPT.MLE method outperformed the others, particularly for longer return periods (e.g., RL25 to RL100). Conversely, for smaller sample sizes ($n = 30$), the NPT.LM method exhibited superior performance in most cases, yielding lower RMSE values. However, the NPT methods underperformed in specific scenarios, such as when $\xi' = -0.5$ or $\xi' = 0.05$, or for shorter return periods.

Table 2. Results of Monte Carlo simulation for the case where $\mu' = 1, \sigma' = 0.1$, and ξ' varies as -0.5, -0.25, -0.05, 0.05, 0.25, and 0.5. The evaluation metric used is the *RMSE*. A hyperparameter $\alpha = 0.5$ is employed using the NPT-GEVD method to estimate the 25-, 50-, and 100-year return levels for each transformation approach. The sample size is $n = 500$.

T	Method	ξ'					
		−0.50	**−0.25**	**−0.05**	**0.05**	**0.25**	**0.50**
RL25	CT.MLE	**0.00158**	0.00301	0.00507	**0.00649**	0.01143	0.02065
	RT.MLE	0.00570	0.00575	**0.00505**	0.00660	0.01144	0.01984
	NPT.MLE	0.00179	**0.00275**	0.00507	0.00664	0.01141	**0.01905**
	CT.LM	0.00202	0.00324	0.00531	0.00677	**0.01108**	0.02005
	RT.LM	0.00181	0.00313	0.00519	0.00667	0.01130	0.02001
	NPT.LM	0.00177	0.00309	0.00515	0.00664	0.01152	0.02092
RL50	CT.MLE	**0.00135**	0.00332	0.00640	**0.00862**	0.01619	0.04034
	RT.MLE	0.00409	0.00509	**0.00639**	0.00865	0.01575	0.02910
	NPT.MLE	0.00152	**0.00322**	0.00640	0.00865	0.01546	**0.02553**
	CT.LM	0.00271	0.00425	0.00695	0.00906	**0.01520**	0.02602
	RT.LM	0.00213	0.00385	0.00665	0.00879	0.01550	0.02911
	NPT.LM	0.00205	0.00377	0.00654	0.00869	0.01588	0.03129
RL100	CT.MLE	**0.00127**	0.00385	0.00804	0.01116	0.02268	0.07924
	RT.MLE	0.00284	0.00479	0.00792	0.01100	0.02070	0.03963
	NPT.MLE	0.00142	**0.00378**	**0.00789**	**0.01091**	**0.01991**	**0.03151**
	CT.LM	0.00344	0.00553	0.00896	0.01180	0.02020	0.03286
	RT.LM	0.00246	0.00467	0.00832	0.01122	0.02020	0.03755
	NPT.LM	0.00235	0.00454	0.00812	0.01102	0.02078	0.04203

Note: Bold values indicate the smallest RMSE among the compared methods.

Table 3. Same as Table 2 but for $n = 30$.

T	Method	ξ'					
		−0.50	**−0.25**	**−0.05**	**0.05**	**0.25**	**0.50**
RL25	CT.MLE	**0.0066**	0.0130	0.0219	0.0289	0.0537	0.1045
	RT.MLE	0.0067	0.0134	0.0220	0.0291	0.0509	0.0883
	NPT.MLE	0.0067	0.0133	0.0221	0.0293	0.0501	0.0835
	CT.LM	0.0075	0.0135	0.0212	0.0260	0.0417	**0.0714**
	RT.LM	0.0073	0.0131	0.0207	0.0253	0.0412	0.0715
	NPT.LM	0.0072	**0.0129**	**0.0204**	**0.0251**	**0.0412**	0.0721
RL50	CT.MLE	0.0074	0.0160	0.0287	0.0409	0.0890	0.1954
	RT.MLE	**0.0074**	0.0161	0.0287	0.0401	0.0724	0.1226
	NPT.MLE	0.0075	0.0162	0.0288	0.0397	0.0690	0.1109
	CT.LM	0.0090	0.0167	0.0270	0.0340	0.0571	0.1005
	RT.LM	0.0087	0.0161	0.0259	0.0326	0.0553	**0.0995**
	NPT.LM	0.0086	**0.0158**	**0.0255**	**0.0320**	**0.0551**	0.1008
RL100	CT.MLE	0.0084	0.0192	0.0371	0.0566	0.1501	0.3640
	RT.MLE	**0.0084**	0.0193	0.0364	0.0528	0.0955	0.1537
	NPT.MLE	0.0084	0.0195	0.0363	0.0514	0.0885	0.1356
	CT.LM	0.0106	0.0203	0.0337	0.0437	0.0754	0.1342
	RT.LM	0.0102	0.0194	0.0318	0.0408	0.0709	**0.1288**
	NPT.LM	0.0101	**0.0191**	**0.0311**	**0.0398**	**0.0704**	0.1315

Note: Bold values indicate the smallest RMSE among the compared methods.

Table 4. Same as Table 2 but for $\alpha = 1$ and $n = 500$.

T	Method	ξ'					
		−0.50	**−0.25**	**−0.05**	**0.05**	**0.25**	**0.50**
RL25	CT.MLE	0.0036	0.0061	0.0090	**0.0110**	0.0172	0.0293
	RT.MLE	0.0034	**0.0050**	0.0089	0.0112	0.0173	0.0268
	NPT.MLE	0.0034	0.0050	**0.0089**	0.0112	0.0173	**0.0268**
	CT.LM	0.0041	0.0061	0.0094	0.0114	**0.0168**	0.0310
	RT.LM	**0.0033**	0.0056	0.0091	0.0111	0.0177	0.0321
	NPT.LM	0.0033	0.0056	0.0091	0.0111	0.0177	0.0322
RL50	CT.MLE	0.0029	0.0063	0.0110	0.0141	0.0235	0.0596
	RT.MLE	**0.0029**	0.0057	0.0109	0.0141	0.0222	0.0323
	NPT.MLE	0.0029	**0.0057**	**0.0109**	0.0141	**0.0221**	**0.0323**
	CT.LM	0.0059	0.0084	0.0123	0.0149	0.0226	0.0336
	RT.LM	0.0038	0.0068	0.0113	**0.0140**	0.0232	0.0418
	NPT.LM	0.0038	0.0068	0.0113	0.0140	0.0232	0.0419
RL100	CT.MLE	**0.0025**	0.0069	0.0136	0.0178	0.0331	0.1270
	RT.MLE	0.0027	0.0067	0.0132	0.0172	0.0269	0.0355
	NPT.MLE	0.0027	**0.0067**	**0.0132**	0.0172	**0.0269**	**0.0354**
	CT.LM	0.0076	0.0113	0.0159	0.0191	0.0309	0.0558
	RT.LM	0.0043	0.0081	0.0138	0.0171	0.0287	0.0487
	NPT.LM	0.0043	0.0081	0.0138	**0.0171**	0.0287	0.0488

Note: Bold values indicate the smallest RMSE among the compared methods.

To interpret these numerical patterns, the distinct statistical behaviors of the estimation methods must be considered. MLE-based methods (e.g., NPT.MLE) perform optimally for larger datasets ($n = 500$) due to their asymptotic efficiency and consistency; with sufficient data, the likelihood surface is well behaved, allowing the model to fully utilize the tail information [2]. Conversely, for smaller samples ($n = 30$), maximum likelihood estimators can be highly sensitive to a few extreme outliers, leading to inflated variance in shape parameter estimation. In such cases, LM-based methods (e.g., NPT.LM) demonstrate superior stability because they rely on linear combinations of order statistics, which inherently downweight the influence of extreme fluctuations and provide robust parameter bounds [18]. This is consistent with established findings that L-moments are highly preferable to MLE for small-sample GEVD applications ($n \leq 50$) [27]. Furthermore, the advantage of the NPT approach over CT is most noticeable for longer return periods (RL50, RL100) and heavily

bounded or skewed shape parameters (ξ'). Under these extreme extrapolations, the linear mapping of the traditional CT-GEVD often approaches or violates the theoretical zero boundary, causing high bias and standard errors. The NPT structurally enforces a strictly positive domain through its nonlinear scaling, effectively stabilizing the tail estimation and providing highly reliable, physically possible long-term return levels.

Table 5. Same as Table 2 but for $\alpha = 1$ and $n = 30$.

T	Method	ξ'					
		−0.50	**−0.25**	**−0.05**	**0.05**	**0.25**	**0.50**
RL25	CT.MLE	**0.0124**	0.0227	0.0391	0.0514	0.0837	0.1536
	RT.MLE	0.0127	0.0231	0.0397	0.0530	0.0803	0.1214
	NPT.MLE	0.0127	0.0231	0.0397	0.0530	0.0802	0.1198
	CT.LM	0.0145	0.0239	0.0384	0.0493	0.0734	0.1050
	RT.LM	0.0136	0.0227	0.0366	0.0468	**0.0712**	**0.1021**
	NPT.LM	0.0136	**0.0227**	**0.0366**	**0.0468**	0.0713	0.1021
RL50	CT.MLE	0.0141	0.0275	0.0508	0.0689	0.1252	0.2803
	RT.MLE	**0.0140**	0.0276	0.0505	0.0686	0.1032	0.1424
	NPT.MLE	0.0141	0.0276	0.0505	0.0686	0.1029	0.1410
	CT.LM	0.0174	0.0292	0.0489	0.0641	0.1003	0.1522
	RT.LM	0.0161	0.0271	0.0453	0.0585	**0.0909**	**0.1279**
	NPT.LM	0.0160	**0.0271**	**0.0453**	**0.0585**	0.0910	0.1281
RL100	CT.MLE	0.0160	0.0331	0.0642	0.0896	0.1844	0.4988
	RT.MLE	**0.0159**	0.0331	0.0623	0.0849	0.1241	0.1535
	NPT.MLE	0.0159	0.0332	0.0623	0.0848	0.1235	0.1524
	CT.LM	0.0202	0.0351	0.0608	0.0812	0.1340	0.2223
	RT.LM	0.0188	0.0324	0.0548	0.0710	**0.1105**	**0.1495**
	NPT.LM	0.0188	**0.0324**	**0.0548**	**0.0710**	0.1106	0.1498

Note: Bold values indicate the smallest RMSE among the compared methods.

Figures 2 and 3 present the results of a simulation corresponding to the outcomes in Tables 2–5, using the evaluated metrics of SE (Figure 2a,c) and bias (Figure 2b,d). The figures compare the performance of the methods (CT.MLE, RT.MLE, NPT.MLE, CT.LM, RT.LM, and NPT.LM) under varying parameter values. Figure 2 corresponds to the hyperparameter $\alpha = 0.5$, while Figure 3 represents $\alpha = 1$ (RT-GEVD). For $n = 500$, NPT.MLE consistently outperforms the other methods, exhibiting lower SEs and minimal bias across most values of ξ' and for increasing return periods (RL25 to RL100).

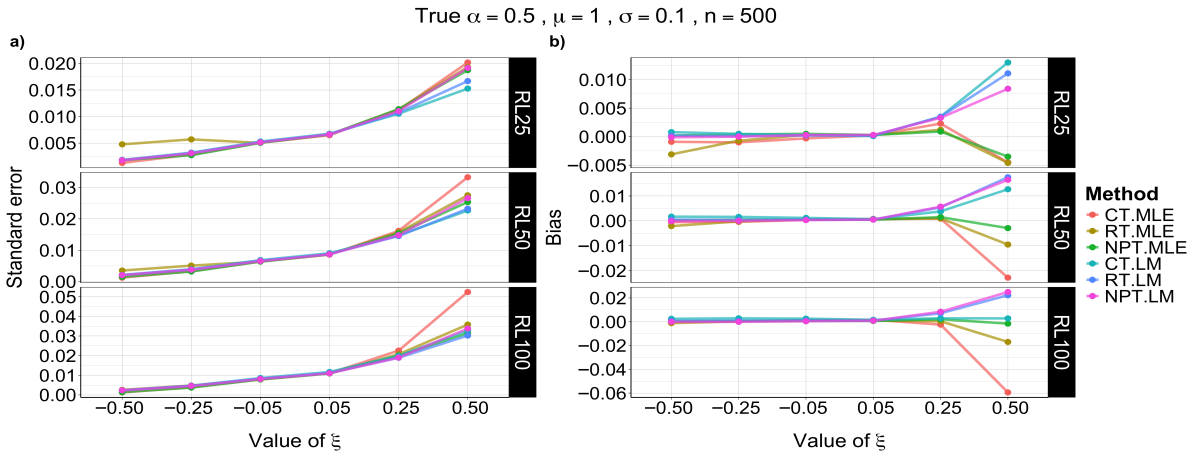


Figure 2. *Cont.*

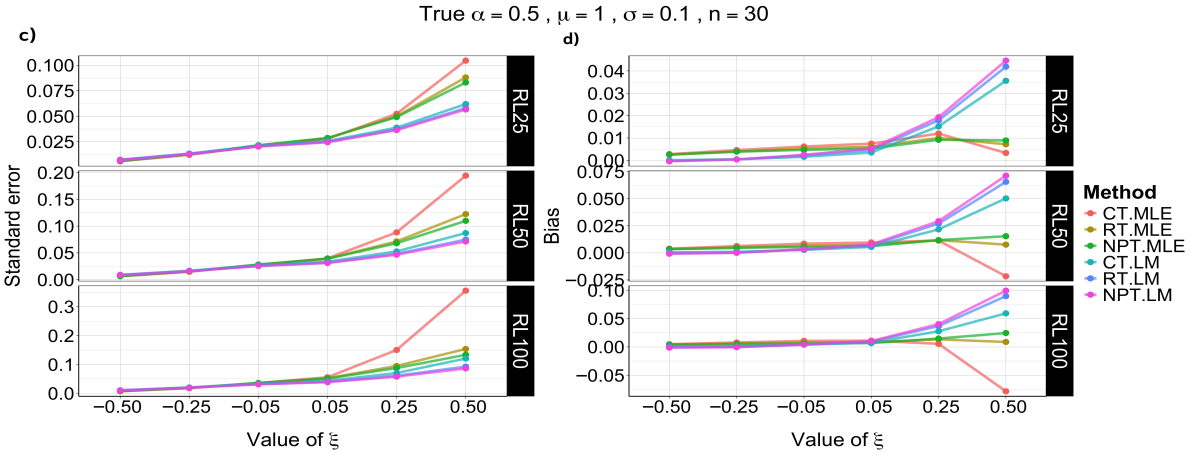


Figure 2. Results of Monte Carlo simulation for the cases with $[\mu', \sigma']$ = [1, 0.1] and varying $\xi' = -0.5, -0.25, -0.05, 0.05, 0.25$, and 0.5. The evaluation measures include the standard error (**a**,**c**) and bias (**b**,**d**), displayed on the vertical axis, evaluated across the specific shape parameter scales (ξ') on the horizontal axis. A hyperparameter $\alpha = 0.5$ is used in the NPT-GEVD method to estimate the 25, 50, and 100-year return levels derived from each transformation method. The legend indicates the specific estimation method (CT.MLE, NPT.MLE, etc.), corresponding to different line types and colors. The sample sizes are $n = 500$ (**top**) and $n = 30$ (**bottom**).

In contrast, for $n = 30$, NPT.LM performs better, particularly for lower values of ξ' and larger return periods, highlighting its robustness with smaller sample sizes. Across all configurations, the SE and bias generally increase as ξ' increases, particularly for higher return levels. RT.LM and NPT.LM demonstrates more stable performance compared to MLE-based methods under these conditions.

A simulation-based evaluation of the statistical consistency under varying true values of ξ' is provided in Section 3. The results demonstrate that parameter estimation remains stable across a range of tail behaviors, particularly when using L-moment estimation in small-sample scenarios. This supports the robustness of the NPT-GEVD framework for practical applications.

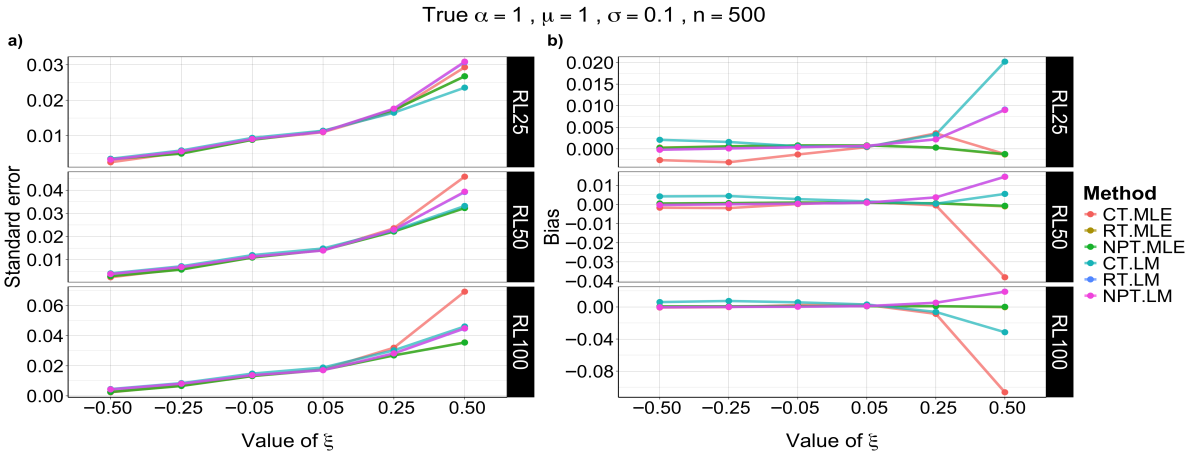


Figure 3. *Cont.*

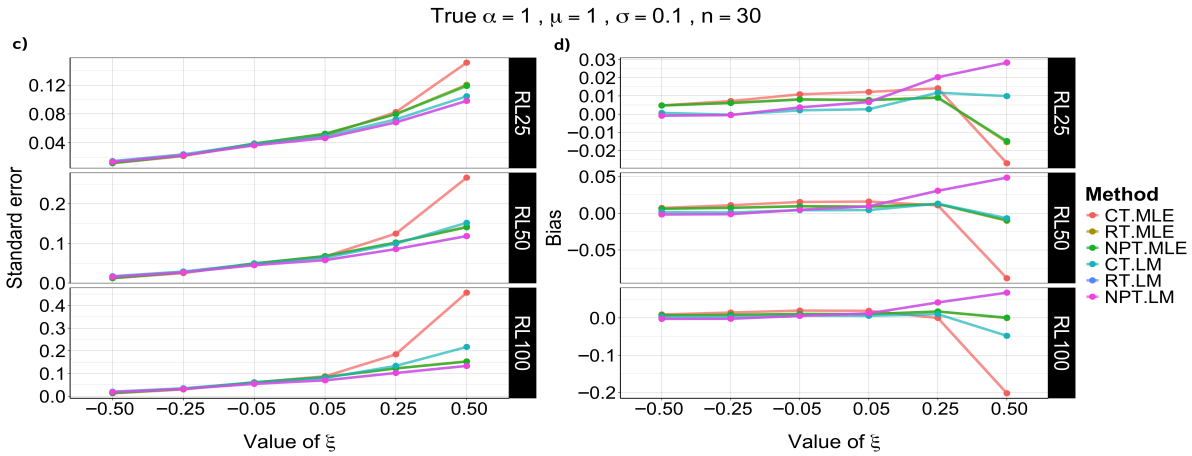


Figure 3. Same as Figure 2, but for the hyperparameter $\alpha = 1$ set to 1 using the NPT-GEVD (which becomes RT-GEVD when $\alpha = 1$). The sample sizes are $n = 500$ (**top**) and $n = 30$ (**bottom**).

4. Application: Inter-Amount Time of Hourly Rainfall Data

4.1. Climate and Data

The IAT refers to the interval between two consecutive occurrences of a specific event, such as rainfall exceeding a predefined threshold [28]. It is commonly used in stochastic processes, time series analysis, and extreme value theory. Let $\{t_1, t_2, t_3, \dots\}$ represent a sequence of times when rainfall events occur, where each t_i denotes the time of the i-th event. The IAT τ_i between two consecutive events t_i and t_{i+1} is given by

$$\tau_i = t_{i+1} - t_i, \quad \forall\, i \in \mathbb{N} \tag{34}$$

where τ_i represents the time difference between the i-th and $(i+1)$-th events.

When applying a threshold u (e.g., $u = 60$ mm), the IAT is calculated only for rainfall events t_i and t_{i+1} where the rainfall amount R_i exceeds the threshold u. Let $\mathcal{R}(k)$ denote the rainfall amount at time k, and define the set of times t as

$$t = \{k : \mathcal{R}(k) \geq u\} \tag{35}$$

where t contains the times when the rainfall amount $\mathcal{R}(k)$ meets or exceeds the threshold u. The threshold-based IAT τ_i is then defined as in Equation (34), but for $t_i, t_{i+1} \in t$, where t_i and t_{i+1} are consecutive times when the rainfall amount reaches the threshold u.

Hourly rainfall data from 63 stations across Korea were used to calculate the IAT. These stations were strategically selected to represent a wide array of distinct microclimatic and topographical zones across the peninsula, ranging from typhoon-exposed southern coastal areas to rain-shadowed inland basins and continental northwestern regions [29–31]. The annual minimum IAT was selected to fit the NPT-GEVD model. The data were recorded over a long-term time scale from 1986 to 2022 [32] (additional details can be found in the SI file). The CvM statistical test was subsequently applied to assess the fit of the estimated parameters to the transformed IAT data.

4.2. Results of Modeling

Given the relatively small sample size, we apply the methodology outlined in Section 2 to estimate the hyperparameters and utilize the LM method for parameter estimation in the NPT-GEVD. The CvM statistic is employed to assess the GOF, which is particularly well suited for the LM method. Thresholds 60, 90, and 110 mm are used to define the IAT of cumulative hourly rainfall reaching these levels. Analyzing the IAT across these

progressively higher thresholds allows us to robustly validate the stability of the NPT-GEVD model under varying degrees of extreme conditions. These thresholds are commonly used for issuing heavy rainfall advisories and warnings in Korea, as determined by the Korea Metrological Administration [32].

Parameter estimates for the NPT-GEVD, RT-GEVD, and CT-GEVD models are derived by fitting the distributions to the minimum data, which include estimates for $(\hat{\mu}', \hat{\sigma}', \hat{\xi}', \hat{\alpha})$ along with their corresponding SEs. A quantitative comparison of these models reveals that the NPT-GEVD provides a structurally superior fit for the actual minimum data. Specifically, across the analyzed stations, the NPT-GEVD generally yielded lower Cram'er–von Mises (CvM) statistics compared to the CT-GEVD, indicating improved overall goodness of fit. Furthermore, when analyzing high thresholds (e.g., 110 mm), the traditional CT-GEVD frequently produced physically impossible negative return levels (as detailed in Table 1), suffering from theoretical boundary violations. In contrast, the NPT-GEVD completely resolved this issue, generating stable and strictly non-negative return levels (RL25, RL50, RL100) across all stations. Full comprehensive results, including individual GOF statistics and SEs for all competing models, are provided in the SI file.

Figures 4–6 present the estimated RLs (in hours) for all 63 stations in Korea, corresponding to the parameters estimated in the SI. The results show consistent spatial patterns of RL estimates across all three thresholds. Overall, shorter RLs, which indicate the rapid accumulation of extreme rainfall, are concentrated along the southern coast, Jeju Island, and the northwestern region (Seoul and Gyeonggi-do province). This spatial distribution is strongly governed by the regional meteorology and complex topography. The southern coastal areas and Jeju Island are highly vulnerable to intense rainfall due to their proximity to the paths of summer typhoons and the East Asian summer monsoon (Changma) front, compounded by the orographic lifting of moist oceanic air masses along coastal mountain ranges [29,30].

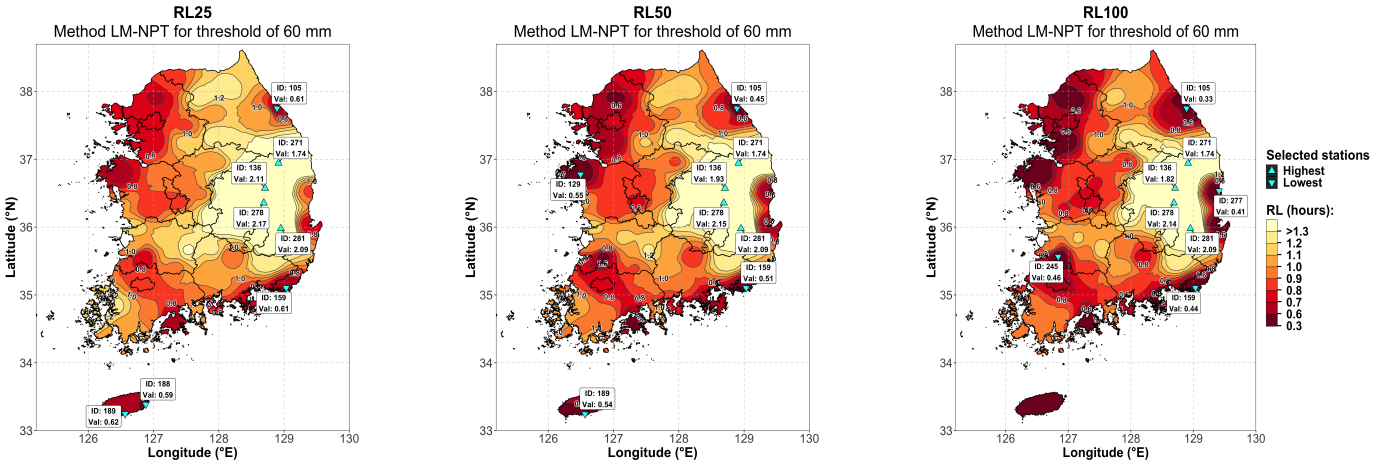


Figure 4. Map of the 25-, 50-, and 100-year return levels for minimum IAT (in hours) using the NPT-GEVD model with LM estimation and a threshold of 60 mm. The color gradient represents the magnitude of return levels (hours) as indicated in the legend. Longitude and latitude are expressed in degrees (°E and °N). Selected stations with the highest and lowest return levels are marked by upward and downward triangles, respectively, with their corresponding values. The geographic coordinates and estimated return levels for all stations are provided in the SI.

For the 60 mm threshold, the shortest RLs are often less than 0.6 h (approximately 36 min) in these highly vulnerable regions. For the 90 mm threshold, the shortest RLs are approximately 1.0 to 1.5 h. When the threshold increases to 110 mm, the RL(s) approach 1 to 2 h, especially in

the northwestern region (Gyeonggi-do province), central western region (Chungcheongnam-do), and southern coastal regions (Jeollanam-do and Gyeongsangnam-do).

In contrast, stations in the inland basin of Gyeongsangbuk-do require the most time to reach the threshold value across all return periods. This clearly demonstrates the rain shadow effect created by the Taebaek and Sobaek mountain ranges, which block the influx of moisture and consequently attenuate the intensity of extreme rainfall in this leeward region [31,33].

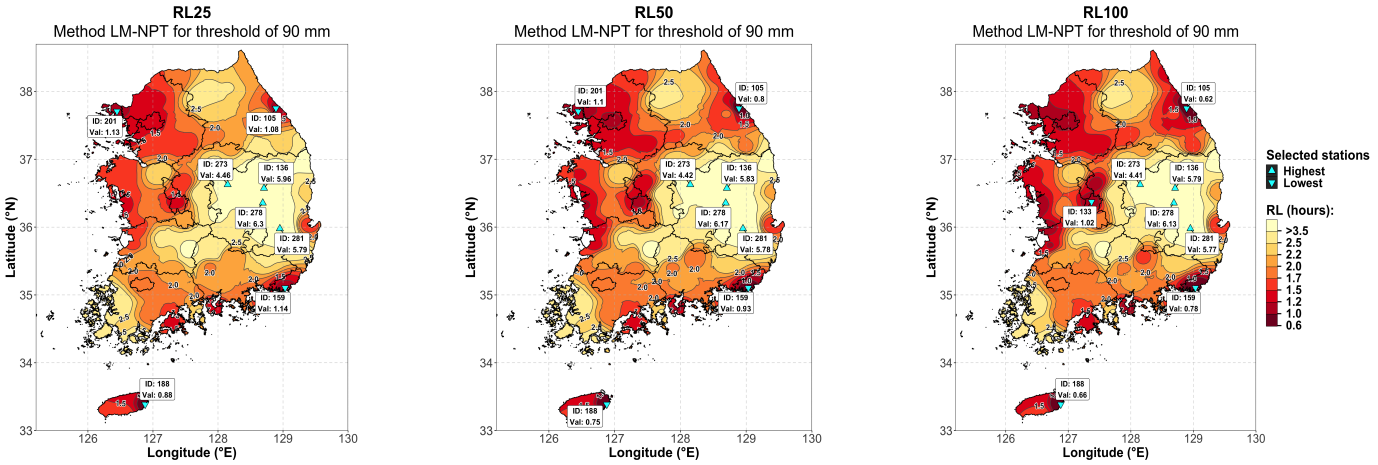


Figure 5. Same as Figure 4 but for a threshold level of 90 mm. Longitude and latitude are expressed in degrees (°E and °N). The color gradient represents return levels (hours), and selected stations with the highest and lowest return levels are marked by upward and downward triangles, respectively.

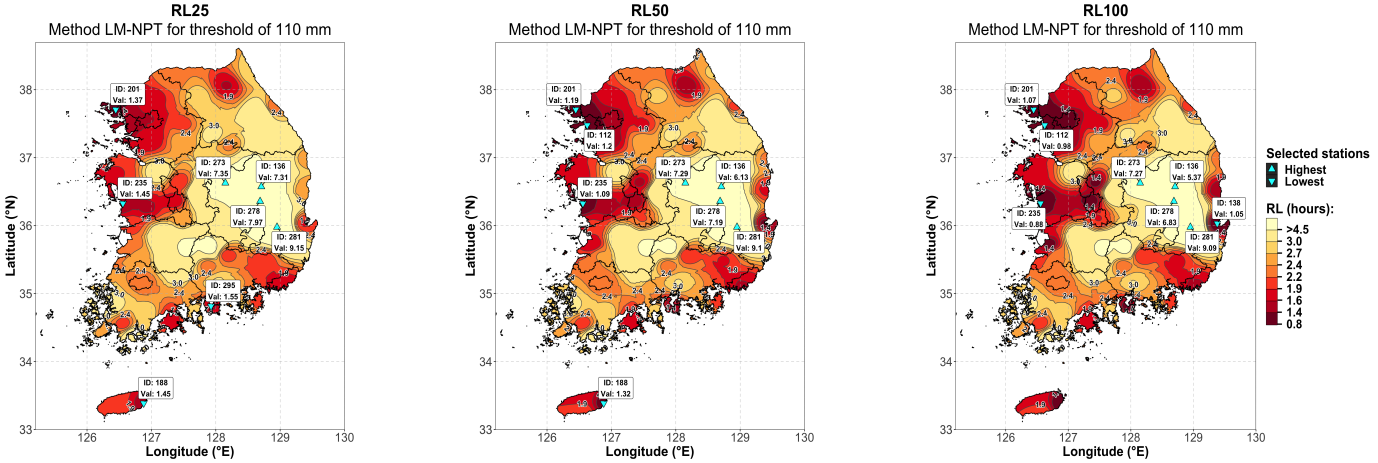


Figure 6. Same as Figure 4 but for a threshold level of 110 mm. Longitude and latitude are expressed in degrees (°E and °N). The color gradient represents return levels (hours), and selected stations with the highest and lowest return levels are marked by upward and downward triangles, respectively.

5. Discussion

This study focuses on the development of statistical models for environmental risk management. Transformation techniques, which are crucial for converting minimum data into a form suitable for the GEVD, require rigorous theoretical validation to ensure compliance with the extreme value framework. We propose an alternative approach to establish the validity of the NPT-GEVD for minimum values, demonstrating its effectiveness through our experimental analysis.

Moreover, robust parameter estimation methods, such as MLE and LM, are vital. In this study, we have explicitly derived and presented both the likelihood functions for MLE

and the direct theoretical formulation of L-moments for the NPT-GEVD applied to block minima. Our experiments reveal that LM provide more stable parameter estimates than MLE for smaller datasets, such as IAT data, which are commonly encountered in real-world applications. Future research will focus on enhancing the hyperparameter estimation techniques and applying this method to real-time warning systems, emphasizing the prediction of the timing of extreme events rather than sparse rainfall data.

6. Conclusions

This study proposes the NPT-GEVD as a robust, theoretically sound alternative for modeling block minima in environmental risk management. Through comprehensive Monte Carlo simulations, we demonstrated that the NPT-GEVD systematically outperforms traditional methods (such as the CT and RT models) by preventing the estimation of physically impossible negative boundaries. Particularly for small sample sizes, typical of hydro-meteorological extremes, the L-moments method proved highly stable and reliable.

The empirical application to South Korean IAT data across multiple heavy rainfall thresholds (60, 90, and 110 mm) confirmed the practical utility of the proposed model. The NPT-GEVD successfully captured the spatial variability of extreme rainfall rapidly accumulating along typhoon-exposed coastal regions while strictly maintaining physically meaningful, non-negative return levels—a theoretical requirement where the traditional CT-GEVD failed. Overall, the NPT transformation offers an effective mechanism for preserving the positivity and interpretability of block minima, serving as a reliable alternative for flood risk assessment.

Despite these advantages, the current model has limitations that suggest clear avenues for future research. First, the present framework assumes stationarity. Given the increasing volatility of extreme weather, driven by climate change, future studies should extend the NPT-GEVD to a non-stationary framework by incorporating time-varying covariates or global climate indices into the parameter estimation. Second, the current model is limited to univariate extremes. Since severe hydro-meteorological disasters are often driven by interacting factors, extending the NPT transformation to multivariate extreme value models (e.g., using copula-based frameworks) will be crucial for assessing compound hazards, such as the joint probability of minimum IAT and maximum storm surges.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/math14081383/s1. References [34–36] are cited in Supplementary Materials.

Author Contributions: Conceptualization, S.Y.; Methodology, T.P. and T.K.; Software, T.P.; Validation, T.P., P.B. and T.K.; Formal analysis, T.P. and T.K.; Data curation, P.B.; Writing—original draft, T.P.; Writing—review and editing, P.B., T.K. and S.Y.; Visualization, T.P.; Supervision, T.K.; Project administration, S.Y.; Funding acquisition, S.Y. All authors have read and agreed to the published version of the manuscript.

Funding: Thanawan's research was financially supported by Khon Kaen University. Piyapatr's research was financially supported by Mahasarakham University. This research was supported by the Global Learning & Academic Research Institution for Master's, PhD students, and Postdocs (LAMP) Program of the National Research Foundation of Korea (NRF), through a grant funded by the Ministry of Education (No. RS-2024-00442775).

Data Availability Statement: Code: https://github.com/TanawanP/NPT-GEVD.git (accessed on 27 February 2025). Data: Observational data in Korea are provided by the Korea Meteorological Administration (KMA) at https://www.weather.go.kr/w/index.do (accessed on 3 March 2024) and can be obtained upon request to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

GEVD	Generalized extreme value distribution
CT	Conventional transformation (sign reversal for minima)
CT-GEVD	Conventional transformation generalized extreme value distribution (model)
RT	Reciprocal transformation
RT-GEVD	GEVD under the reciprocal transformation (model)
NPT	Negative power transformation
NPT-GEVD	GEVD under the negative power transformation (model)
CvM	Cramér–von Mises (goodness-of-fit statistic)
AD	Anderson–Darling (goodness-of-fit statistic)
BFGS	Broyden–Fletcher–Goldfarb–Shanno (quasi-Newton optimizer)
CDF	Cumulative distribution function
PDF	Probability density function
EDF	Empirical distribution function
GD	Gumbel distribution (the $\xi = 0$ case of GEVD)
GOF	Goodness of fit
IAT	Inter-amount time (between threshold-exceeding rainfall events)
KMA	Korea Meteorological Administration
LM	L-moments
MLE	Maximum likelihood estimation
MC	Monte Carlo (simulation)
PWM	Probability-weighted moment
RL	Return level
RL_{25}	Return level for the 25-year return period
RL_{50}	Return level for the 50-year return period
RL_{100}	Return level for the 100-year return period
RMSE	Root mean square error
SD	Standard deviation
SE	Standard error
SI	Supplementary Information

References

1. Jenkinson, A.F. The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Q. J. R. Meteorol. Soc.* **1955**, *81*, 158–171. [CrossRef]
2. Coles, S. *An Introduction to Statistical Modeling of Extreme Values*; Springer: London, UK, 2001.
3. Haan, L.; Ferreira, A. *Extreme Value Theory: An Introduction*; Springer: New York, NY, USA, 2006.
4. Fawad, M.; Yan, T.; Chen, L.; Huang, K.; Singh, V.P. Multiparameter probability distributions for at-site frequency analysis of annual maximum wind speed with L-moments for parameter estimation. *Energy* **2019**, *181*, 724–737. [CrossRef]
5. Smith, R.L.; Weissman, I. Estimating the Extremal Index. *J. R. Stat. Soc. Ser. B (Stat. Methodol.)* **1994**, *56*, 515–528. [CrossRef]
6. Dankunprasert, S.; Jaroengeratikun, U.; Talangtam, T. *The Properties of Inverse Pareto Distribution and Its Application to Extreme Events*; Thailand Statistician: Bangkok, Thailand, 2021; pp. 1–13.
7. Özdemir, Ö; Çavuş, M. Performance of the inverse transformation method for extreme value distributions. In Proceedings of the Xth International Statistics Days Conference (ISDC'2016), Giresun, Turkey, 7–9 October 2016.
8. Bardsley, E. A class of probability distributions for application to non-negative annual maxima. *Hydrol. Earth Syst. Sci. Discuss.* **2017**, *2017*, 1–6
9. Bardsley, E. The Weibull distribution as an extreme value model for transformed annual maxima. *J. Hydrol.* **2019**, *58*, 123–131.
10. Mudholkar, G.S.; Srivastava, D.K. Exponentiated Weibull family for analyzing bathtub failure-rate data. *IEEE Trans. Reliab.* **1993**, *42*, 299–302. [CrossRef]
11. Box, G.E.; Cox, D.R. An analysis of transformations. *J. R. Stat. Soc. Ser. B (Stat. Methodol.)* **1964**, *26*, 211–243. [CrossRef]
12. Asar, O.; Ilk, O.; Dag, O. Estimating Box-Cox power transformation parameter via goodness-of-fit tests. *Commun. Stat. Simul. Comput.* **2017**, *46*, 91–105. [CrossRef]
13. Bali, T.G. The generalized extreme value distribution. *Econ. Lett.* **2003**, *79*, 423–427. [CrossRef]
14. Hong, H.P. Application of the Box-Cox power transformation in extreme value analysis of wind speed. In Proceedings of the 12th International Conference on Applications of Statistics and Probability in Civil Engineering (ICASP12), Vancouver, BC, Canada, 12–15 July 2015.

15. Naveau, P.; Huser, R.; Ribereau, P.; Hannart, A. Modeling jointly low, moderate, and heavy rainfall intensities without a threshold selection. *Water Resour. Res.* **2016**, *52*, 2753–2769. [CrossRef]
16. Kotz, S.; Nadarajah, S. *Extreme Value Distributions: Theory and Applications*; Imperial College Press: London, UK, 2000.
17. R Core Team. *'stats' Package: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2023. Available online: https://www.R-project.org/ (accessed on 3 March 2024).
18. Hosking, J.R. L-moments: Analysis and estimation of distributions using linear combinations of order statistics. *J. R. Stat. Soc. Ser. B (Stat. Methodol.)* **1990**, *52*, 105–124. [CrossRef]
19. Greenwood, J.A.; Landwehr, J.M.; Matalas, N.C.; Wallis, J.R. Probability weighted moments: Definition and relation to parameters of several distributions expressable in inverse form. *Water Resour. Res.* **1979**, *15*, 1049–1054. [CrossRef]
20. Hosking, J.R.M.; Wallis, J.R. *Regional Frequency Analysis*; Cambridge University Press: Cambridge, UK, 1997.
21. Stephens, M.A. EDF statistics for goodness of fit and some comparisons. *J. Am. Stat. Assoc.* **1974**, *69*, 730–737. [CrossRef]
22. Choulakian, V.; Stephens, M.A. Goodness-of-fit tests for the generalized Pareto distribution. *Technometrics* **2001**, *43*, 478–484. [CrossRef]
23. Nocedal, J.; Wright, S.J. *Numerical Optimization*; Springer: New York, NY, USA, 2006.
24. Vogel, R.M.; Fennessey, N.M. L moment diagrams should replace product moment diagrams. *Water Resour. Res.* **1993**, *29*, 1745–1752. [CrossRef]
25. Malekinezhad, H.; Zare-Garizi, A. Regional frequency analysis of daily rainfall extremes using L-moments approach. *Atmósfera* **2014**, *27*, 411–427. [CrossRef]
26. Gelman, A.; Jakulin, A.; Pittau, M.G.; Su, Y.S. A weakly informative default prior distribution for logistic and other regression models. *Ann. Appl. Stat.* **2008**, *2*, 1360–1383. [CrossRef]
27. Hosking, J.R.M.; Wallis, J.R.; Wood, E.F. Estimation of the generalized extreme-value distribution by the method of probability-weighted moments. *Technometrics* **1985**, *27*, 251–261. [CrossRef]
28. Schleiss, M.; Smith, J.A. Two simple metrics for quantifying rainfall intermittency: The burstiness and memory of interamount times. *J. Hydrometeorol.* **2016**, *17*, 421–436. [CrossRef]
29. Jung, Y.; Shin, J.-Y.; Ahn, H.; Heo, J.-H. The Spatial and Temporal Structure of Extreme Rainfall Trends in South Korea. *Water* **2017**, *9*, 809. [CrossRef]
30. Jung, H.-S.; Choi, Y.; Oh, J.-H.; Lim, G.-H. Recent trends in temperature and precipitation over South Korea. *Int. J. Climatol.* **2002**, *22*, 1327–1337. [CrossRef]
31. Kim, I.-K.; Ahmed, R. Long-Term Trend of Summer Rainfall at Selected Stations in the Republic of Korea. In *The Geographical Bulletin*; Gamma Theta Upsilon: Bowling Green, OH, USA, 2004.
32. Korea Meteorological Administration. Observational Climate Data, 2024. Available online: https://www.weather.go.kr/w/index.do (accessed on 20 November 2024).
33. Park, J.; Jang, D.-H. Application of MK-PRISM for interpolation of wind speed and comparison with co-kriging in South Korea. *GIScience Remote Sens.* **2016**, *53*, 421–443. [CrossRef]
34. IPCC. *AR6 Synthesis Report: Climate Change 2023*; Contribution ofWorking Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change; IPCC AR6 Synthesis Report Website; IPCC: Interlaken, Switzerland, 2023. Available online: https://www.ipcc.ch/report/sixth-assessment-report-cycle/ (accessed on 20 November 2024).
35. World Meteorological Organization. Climate Change Impacts Increase in Asia. 2024. Available online: https://wmo.int/news/media-centre/climate-change-impacts-increase-asia (accessed on 19 August 2024).
36. National Geographic Information Institute (NGII). National Atlas of Korea: Administrative Districts. 2016. Available online: http://nationalatlas.ngii.go.kr/ (accessed on 4 April 2026).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.