

# A semiparametric generalized exponential regression model with a principled distance-based prior

Arijit Dey<sup>1</sup> and Arnab Hazra<sup>2\*</sup>

<sup>1</sup>Department of Statistical Science, Trinity College of Arts and Sciences,  
Duke University, Durham, NC, 27708-0251, United States.

<sup>2\*</sup>Department of Mathematics and Statistics, Indian Institute of Technology  
Kanpur, Kanpur, 208016, India.

\*Corresponding author(s). E-mail(s): [ahazra@iitk.ac.in](mailto:ahazra@iitk.ac.in);  
Contributing authors: [arijit.dey@duke.edu](mailto:arijit.dey@duke.edu);

## Abstract

The generalized exponential distribution is a well-known probability model in lifetime data analysis and several other research areas, including precipitation modeling. Despite having broad applications for independently and identically distributed observations, its uses as a generalized linear model for non-identically distributed data are limited. This paper introduces a semiparametric Bayesian generalized exponential (GE) regression model. Our proposed approach involves modeling the GE rate parameter within a generalized additive model framework. An important feature is the integration of a principled distance-based prior for the GE shape parameter; this allows the model to shrink to an exponential regression model that retains the advantages of the exponential family. We draw inferences using the Markov chain Monte Carlo algorithm. Extensive simulations demonstrate that the proposed model outperforms simpler alternatives. The Western Ghats mountain range holds critical importance in regulating monsoon rainfall across Southern India, profoundly impacting regional agriculture. Here, we analyze daily wet-day rainfall data for the monsoon months between 1901–2022 for the Northern, Middle, and Southern Western Ghats regions. Applying the proposed model to analyze the rainfall data over 122 years provides insights into model parameters, short-term temporal patterns, and the impact of climate change. We observe a significant decreasing trend in wet-day rainfall for the Southern Western Ghats region.

**Keywords:** Climate change; Generalized exponential distribution; Markov chain Monte Carlo; Penalized complexity prior; Semiparametric Bayesian regression; Western Ghats; Wet-day precipitation modeling

# 1 Introduction

The Western Ghats, a prominent mountain range along the western coast of India, plays a crucial role in shaping the climatic patterns and hydrological dynamics of Southern India. Known for its exceptional biodiversity, lush forests, and vital water resources, the Western Ghats has long captured the attention of researchers and environmentalists [1, 2]. Among the various climatic parameters that influence this ecologically significant region, rainfall is a crucial driver of its diverse ecosystems, water availability, and overall environmental health. The Western Ghats, characterized by its rugged terrain and proximity to the Arabian Sea, experiences a unique and intricate rainfall pattern heavily influenced by monsoon dynamics [3]. Over the last century, this region has experienced notable climatic shifts due to global climate change and local human activities [4, 5]. Analyzing wet-day rainfall in this region during monsoon months over an extended period of 122 years using a flexible statistical model offers a valuable opportunity to gain insights into short and long-term trends, variability, and potential shifts in the monsoonal regime. While our methodology is motivated by these goals in the context of rainfall data analysis for the Western Ghats mountain range, it is also broadly applicable to analyzing rainfall and other types of datasets, including lifetime data, financial data, etc., for other parts of the globe.

Researchers have widely employed the exponential distribution to model rainfall data [6, 7] in literature; its simplicity integrates seamlessly into hydrological and climatological frameworks. However, contemporary research increasingly recognizes the need for innovative probability distribution models to encompass complex real-world data patterns better. This realization has prompted the introduction of novel probability classes with far-reaching implications across diverse research domains. [8] provide an excellent overview of newly developed distributions. A notable collection of models is generalized distributions, which have gained attention from both practical and theoretical statisticians for their adaptability to various datasets. Some examples include the Marshall-Olkin generalized exponential distribution [9], generalized inverse Gaussian distribution [10], generalized Rayleigh distribution [11], etc. These distributions are comprehensively examined in [12].

[13] introduced another crucial generalized distribution called the generalized exponential (GE) distribution, which emerges as a specific case within the three-parameter exponentiated-Weibull model. The GE distribution has two parameters- a shape and a rate (or scale, defined as the inverse of rate) parameter. This distribution boils down to an exponential distribution when the shape parameter is one. Thus, an additional shape parameter expands the capabilities of the exponential distribution, making it more adaptable to various datasets. Since its introduction, many researchers have integrated substantial advancements in exploring different properties, estimation strategies, extensions, and applications of this distribution. For instance, [14] found the efficacy of GE distribution compared to gamma or Weibull distributions, whereas [15] discussed different methods of estimating the parameters of GE distribution. [16], [17], and [18] explored Bayesian estimation and prediction methods

in this context. [19] reviewed the existing results and discussed some new estimation methods and their characteristics. Numerous researchers have modeled the experimental data using GE distribution across several disciplines like meteorological studies [20, 21]; flood frequency analysis [22]; reliability analysis [23]; lifetime analysis [24]; risk analysis [25]. However, as of our knowledge, it is most likely that GE regression-type models have never been proposed in the literature; searching using the phrase “Generalized exponential regression” in Google Scholar does not lead to any paper relevant to our proposed methodology. In this study, taking a step beyond exponential regression, we employ the GE regression framework to explore rainfall patterns.

Rainfall data collected over a century are inherently nonstationary. Here, modeling the temporal trend using traditional parametric regression would struggle to capture the intricate and evolving short-term temporal patterns. In this context, a semiparametric regression setup emerges as a promising approach. In the existing literature, many researchers have delved into applying semiparametric regression techniques for analyzing rainfall patterns [26, 27]. While a generalized linear model (GLM) assumes the link function to be a linear combination of the covariates, the more flexible generalized additive models [GAM, 28] allow the link function to be a sum of nonlinear smooth functional forms of the underlying covariates. We generally model each smooth function in GAMs as a linear combination of basis functions like cubic B-splines. Instead of estimating the entirely unknown function, following a finite truncation (hence semiparametric) of the number of B-splines, we draw inferences based on basis function coefficients [29]. Henceforth, instead of GAM, we use the term ‘semiparametric regression’, which is common in Bayesian nonparametrics. The rate parameter of the GE distribution is always positive, and hence, it would be reasonable to model the log-rate in a semiparametric regression framework.

Within the Bayesian methodology, priors hold a pivotal role in inference, and the literature provides a diverse spectrum of prior distributions utilized for regression coefficients in semiparametric regression frameworks. For instance, Gaussian prior was employed by [30], while [31] opted for Laplace prior. [32] utilized Zellner’s  $g$ -prior, while [33] considered flat priors, and [34] used the Normal-Gamma prior. On the other hand, the gamma distribution has consistently been considered the most natural prior choice for the shape parameter of the GE distribution; authors who introduced the GE distribution chose a gamma prior for the shape parameter in [18] as well. Besides [17] and [35] also employed gamma prior for the shape parameter. However, the literature demonstrates that a handful of alternative prior choices have also been utilized. For example, [36] employed a Jeffrey’s prior, indicating their preference for an objective prior, and [37] opted for a non-informative prior in their study.

The Penalized Complexity (PC) prior, introduced by [38], has emerged in recent literature, which mitigates the model complexity through penalization. In cases where a model extends from a simpler foundational model by incorporating an additional parameter, this type of prior becomes applicable; it penalizes the escalation in model complexity that arises when favoring the extended model over its more straightforward counterpart. Existing literature encompasses instances of this approach across various models [39]. [40] developed

PC priors for estimating the effective degrees of freedom in Bayesian penalized splines (P-splines), while [41] discussed a PC prior for the skewness parameter of the power links family, and [42] proposed interpretable and comprehensive PC priors for the coefficients of a stationary autoregressive process.

In this paper, along with proposing a semiparametric Bayesian GE regression model where we build the rate parameter in a generalized additive model framework (in a log-scale), we employ the PC prior for the GE shape parameter, which allows the GE regression to shrink towards an exponential regression. Thus, the exponential distribution is considered to be the base model for the GE distribution. In several practical examples [7, 43], the exponential distribution is found to be a reasonable model and enjoys several benefits of being a member of the exponential family, thus shrinking the GE distribution to its base model through shrinking the shape parameter to one is justified. On the other hand, we opt for the independent Gaussian priors for the regression coefficients. We draw inferences using the Markov chain Monte Carlo (MCMC) algorithm; here, conjugate priors are not available for the model parameters, and thus, we update them using Metropolis-Hastings steps. We conduct a thorough simulation study by simulating 1000 datasets from each combination of the model generating and model fitting scenarios, and we compare the performances of parametric and semiparametric Bayesian GE regression models under the conventional gamma prior choices for the GE shape parameter along with our proposed one. We study the coverage probabilities for the shape parameter and the rate functions and compare these two models using widely applicable information criterion [WAIC, 44]. We implement the proposed methodology to the daily wet-day precipitation spanning from 1901 to 2022 in different regions of the Western Ghats mountain range, using the year as a covariate and wet-day precipitation as a response. We study the convergence and mixing of the MCMC chains and compare different model fits in terms of WAIC.

The paper is structured as follows. Section 2 delves into the necessary background about the GE distribution, thoroughly examining its properties. In Section 3, we introduce the GE regression model. Proceeding to Section 4, we concentrate on delineating the prior specifications for the regression model, including introducing a principled distance-based prior for the shape parameter of the GE distribution. Bayesian inference under small and large sample scenarios are addressed in Section 5. Section 6 presents the outcomes of the simulation study, while Section 7 discusses an exploratory data analysis that justifies our semiparametric GE model assumption for the wet-day precipitation data, followed by the results obtained from our proposed model and some simpler alternatives. Finally, Section 8 summarizes our findings and contributions.

## 2 Background: Generalized Exponential (GE) Distribution

We say a random variable  $Y$  follows GE distribution if its cumulative distribution function (CDF) is given by

$$F(y; \alpha, \lambda) = (1 - e^{-\lambda y})^\alpha; \quad y, \alpha, \lambda > 0,$$

where  $\alpha$  is the shape parameter and  $\lambda$  is the rate parameter. The corresponding probability density function (PDF) is given by

$$f(y; \alpha, \lambda) = \alpha \lambda (1 - e^{-\lambda y})^{\alpha-1} e^{-\lambda y}; \quad y, \alpha, \lambda > 0. \quad (1)$$

The GE distribution is a more complex model than the exponential distribution, as it incorporates an extra shape parameter. Both models coincide when  $\alpha = 1$ .

### 2.1 Properties of GE

The hazard function of the GE distribution is given by

$$h(y; \alpha, \lambda) = \frac{f(y; \alpha, \lambda)}{1 - F(y; \alpha, \lambda)} = \frac{\alpha \lambda (1 - e^{-\lambda y})^{\alpha-1} e^{-\lambda y}}{1 - (1 - e^{-\lambda y})^\alpha}; \quad y > 0.$$

The GE distribution has an increasing or decreasing hazard rate depending on the value of the shape parameter. The hazard function decreases for  $\alpha < 1$ , remains constant for  $\alpha = 1$ , and increases for  $\alpha > 1$ . The moment generating function (MGF) of the GE distribution is given by

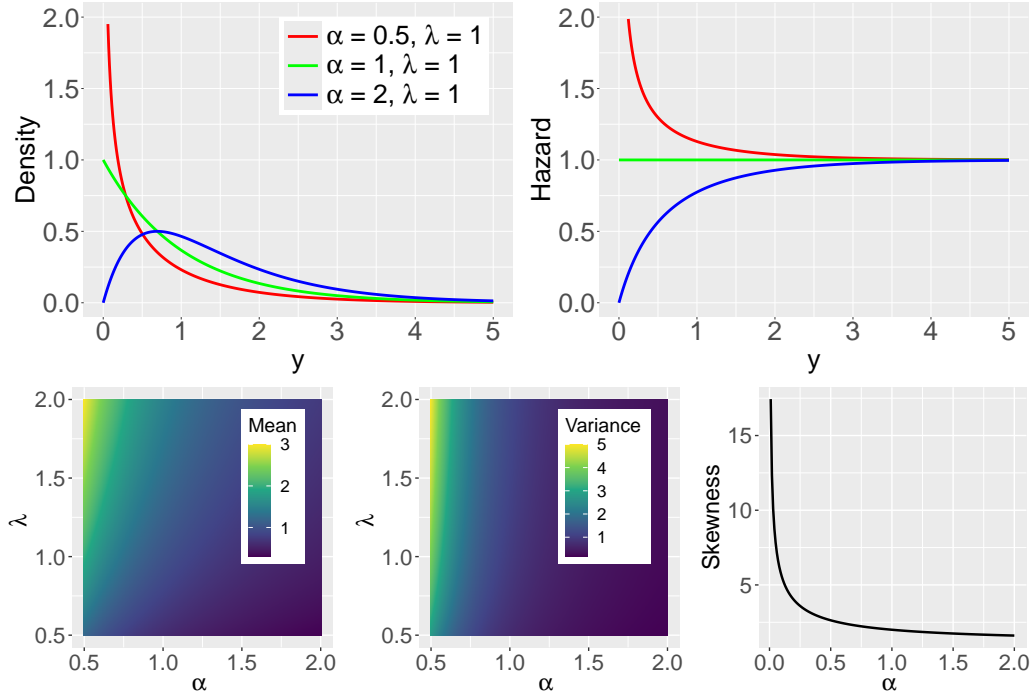
$$M_Y(t) = \frac{\Gamma(\alpha + 1)\Gamma(1 - t/\lambda)}{\Gamma(1 + \alpha - t/\lambda)}; \quad 0 \leq t < \lambda,$$

and differentiating the log of the MGF with respect to  $t$  repeatedly and then setting  $t = 0$ , we get the expectation, variance, and skewness of GE distribution as

$$\begin{aligned} E(Y) &= \lambda^{-1} [\psi(\alpha + 1) - \psi(1)], \\ V(Y) &= \lambda^{-2} [\psi^{(1)}(1) - \psi^{(1)}(\alpha + 1)], \\ \text{Skewness}(Y) &= \left[ \psi^{(2)}(\alpha + 1) - \psi^{(2)}(1) \right] / \left[ \psi^{(1)}(1) - \psi^{(1)}(\alpha + 1) \right]^{3/2}, \end{aligned}$$

where  $\psi^{(m)}(z) = \frac{\partial^m}{\partial z^m} \psi(z) = \frac{\partial^{m+1}}{\partial z^{m+1}} \ln \Gamma(z)$  is the polygamma function of order  $m$ ; for  $m = 0$ , it denotes the digamma function.

Figure 1 sheds light on different aspects of the GE distribution, e.g., PDF, hazard function, mean, variance, and skewness. The top-left panel of Figure 1 shows that for  $\alpha < 1$ , the curve depicting the PDF of the GE distribution has an asymptote at the Y-axis and then decreases exponentially and monotonically as we move across the positive real line. With  $\alpha = 1$ , GE coincides with the exponential distribution, thus having mode at zero (with value



**Fig. 1** Generalized Exponential probability density function (top-left), hazard function (top-right), mean (bottom-left), variance (bottom-middle), and skewness (bottom-right) functions. Both top panels share the same legend.

$= \lambda$ ) and gradually decreasing similarly as the previous case. When  $\alpha > 1$ , the curve initiates at zero, then increases over a range of values, and eventually decreases monotonically, having a unique mode at  $\log(\alpha)/\lambda$ . As mentioned earlier, the top-right panel of Figure 1 shows that the hazard function is monotonically increasing when  $\alpha < 0$ , monotonically decreasing when  $\alpha > 0$ , and constant (the value being  $\lambda = 1$ ) when  $\alpha = 1$ . The mean and variance of GE behave somewhat similarly. From the bottom-left and the bottom-middle panel of Figure 1, we see for a fixed value of  $\alpha$ , both mean and variance decrease with increasing  $\lambda$  and for a fixed value of  $\lambda$ , both increase as  $\alpha$  increases. On the other hand, the skewness of the GE distribution depends only on the shape parameter and decreases exponentially with increasing  $\alpha$  (bottom-right panel of Figure 1).

### 3 Generalized Exponential (GE) Regression

#### 3.1 Parametric GE Regression

The observations often do not satisfy the assumption of being independently and identically distributed. For example, the rainfall data observed across 122 years in the Western Ghats mountain range are unlikely to be identically distributed due to several potential short-term and long-term factors, such as El Nino and global warming. However, different rainfall events across the years can be safely assumed to be independent, which is a common assumption

in the environmental statistics literature. We thus can consider a regression model where the response variable  $Y$  follows a GE distribution, and the relationship between  $Y$  and the covariates  $\mathbf{X} = (X_1, \dots, X_P)'$  is represented through a linear predictor  $\eta$ , i.e., a linear combination of the covariates with associated regression coefficients, given by

$$\eta(\mathbf{X}) = \phi_0 + \phi_1 X_1 + \phi_2 X_2 + \dots + \phi_P X_P, \quad (2)$$

where  $\phi = (\phi_0, \phi_1, \dots, \phi_P)'$  is the vector of regression coefficients.

Here, the shape parameter  $\alpha$  is considered to be an inherent property of the distribution that characterizes the shape and asymmetry of the distribution, allowing for a more flexible modeling approach compared to a standard linear regression with a Gaussian error component. On the other hand, the rate parameter  $\lambda$  is the parameter of interest in the regression model, which captures the association between the covariates and the response variable. Moreover, given that the rate parameter of the GE distribution is positive, we relate it to the linear predictor from (2) using a link function designed to ensure the rate parameter always stays positive. Thus, for  $n$  observed responses and the corresponding covariate vectors  $\{Y_i, \mathbf{X}_i; i = 1, \dots, n\}$ , we conceptualize the GE regression model as

$$Y_i | \mathbf{X}_i = \mathbf{x}_i \stackrel{\text{Indep}}{\sim} \text{GE}(\alpha, \lambda_i), \quad i = 1, \dots, n,$$

where  $g(\lambda_i) = \eta(\mathbf{x}_i)$ , with  $g(\cdot)$  representing an appropriate link function and  $\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{iP})'$ . A natural choice for  $g(\cdot)$  is  $g(\lambda) = \log(\lambda)$ , which ensures  $\lambda_i > 0$  for all  $i$ .

### 3.2 Semiparametric GE Regression

The parametric GE regression model introduced in Section 3.1 aims to capture the linear relationship between the response (suitably transformed) and the covariates. However, such an assumption may be practically unreliable; for example, assuming the mean or median rainfall to vary linearly across 122 years for the Western Ghats range is unlikely, and several short-term decadal patterns can be present in the data. In contrast, a nonparametric regression model assumes no specific form for the relationship between the response and covariates, allowing flexibility based on data-derived information. While this approach allows for more flexible modeling, it is computationally intensive, less interpretable, and can be affected by the *curse of dimensionality*. Semiparametric regression integrates the above two approaches, allowing us to have the best of both regimes; it incorporates the interpretability of the parametric setup and flexibility of the nonparametric setup.

In linear models or generalized linear models (GLM) that fall under the parametric setup, we assume the conditional mean of the distribution of the response variable is linked with the linear predictor through a linear combination of the covariates or their functions. The generalized additive model [28] setup extends this domain of regression models by introducing the nonparametric component in the linear predictors. In this setup, the most

general formulation of the linear predictor can be given as

$$\eta(\mathbf{x}) = \sum_{p=1}^P f_p(x_p), \quad (3)$$

where  $f_j$ 's are smoothing functions of continuous covariates and  $\mathbf{x} = (x_1, x_2, \dots, x_P)'$ . Under a purely nonparametric scenario, each of  $f_j$ 's allows an infinite basis function expansion, while most semiparametric methods assume they can be expressed as a linear combination of finite basis functions, often denoted as

$$f_p(z) = \sum_{k=1}^{K_p} \beta_{p,k} B_{p,k}(z); \quad p = 1, 2, \dots, P, \quad (4)$$

where  $B_{p,k}(\cdot)$ 's are known basis functions and  $\beta_{p,k}$  are unknown basis function coefficients that determine the shape of the smoothing function  $f_p(z)$ . A basis expansion of  $M$ -many terms can match the true curve  $f_p(\cdot)$  at any  $M$  points  $X_1, \dots, X_M$  in the range of covariates. Hence, increasing  $M$  gives us an arbitrarily flexible model, and cross-validation or information criterion-based choice of  $M$  is necessary for practical purposes.

In this study, we employ a semiparametric model akin to (3) for the rate parameter of the GE distribution. With a covariate vector comprising  $P$  components and the appropriate logarithmic link function, the regression model takes the form:

$$Y_i | \mathbf{X}_i = \mathbf{x}_i \stackrel{\text{Indep}}{\sim} \text{GE}(\alpha, \lambda(\mathbf{x}_i)) \text{ with } \log \{\lambda(\mathbf{x}_i)\} = \sum_{p=1}^P \sum_{k=1}^{K_p} \beta_{p,k} B_{p,k}(x_{ip}), \quad (5)$$

where  $B_{p,k}(\cdot)$ 's are cubic B-splines and  $\beta_{p,k}$ 's are the spline coefficients representing the weights assigned to the corresponding spline functions. Here, a cubic B-spline is a piecewise-defined polynomial cubic function that is defined on a set of knots or control points, taking the form  $B_s(x) = (x - v_s)_+^3$  where  $v_s$ 's are the fixed knots that span the range of  $x$  and '+' denotes the positive part. In our model, we choose equidistant knots. Denoting  $\boldsymbol{\beta}_p = (\beta_{p,1}, \dots, \beta_{p,K_p})'$ ,  $p = 1, \dots, P$  and combining them, we have  $\boldsymbol{\beta} = (\boldsymbol{\beta}_1', \dots, \boldsymbol{\beta}_P')'$ . Further, for the B-spline components, we denote  $\mathbf{B}_p(\mathbf{x}_i) = [B_{p,1}(x_{i,p}), \dots, B_{p,K_p}(x_{i,p})]'$ ,  $p = 1, \dots, P$  and  $\mathbf{B}(\mathbf{x}_i) = [\mathbf{B}_1(\mathbf{x}_i)', \dots, \mathbf{B}_P(\mathbf{x}_i)']'$ . By an abuse of notation, for  $K = \sum_{p=1}^P K_p$ , we henceforth denote  $\mathbf{B}(\mathbf{x}_i) = (b_{i,1}, \dots, b_{i,K})'$  and  $\boldsymbol{\beta} = (\beta_1, \dots, \beta_K)'$  for  $i = 1, \dots, n$ .

## 4 Prior Specification

We first discuss the prior specification for the rate parameter of the GE distribution. If the observations are independently and identically distributed with the same rate parameter  $\lambda$ , we can formulate an explicit prior for  $\lambda$ . Alternatively, in the case of a parametric GE regression model, we can choose independent priors for the elements of  $\boldsymbol{\phi}$  in (2). In the case of a generalized additive model setup, as in (3), without any finite basis function expansion, we



need to choose priors for  $f_p$ 's; a standard prior choice in the literature is a Gaussian process [29]. In semiparametric Bayesian regression, independent prior distributions are explicitly specified for the spline coefficients  $\beta = (\beta_1, \beta_2, \dots, \beta_K)'$ . This paper considers independent weakly-informative Gaussian priors for  $\beta_k$ 's; specifically  $\beta_k \stackrel{\text{IID}}{\sim} \text{N}(0, 10^2)$ .

We employ a newly developed class of priors for the shape parameter of the GE distribution. In cases where a model is constructed based on a simpler base model, the chosen prior should accurately reflect the characteristics of the model considered and capture its departure from the base model. This type of prior construction is founded upon the work of [38], who introduced the Penalized Complexity (PC) prior. The PC prior is a model-based prior that imposes a penalty on the deviation of the model of consideration from its simpler base version at a logarithmic constant rate. The following subsection discusses the PC prior for  $\alpha$ .

#### 4.1 Penalized Complexity (PC) prior

The PC prior is an informative proper prior that exhibits robustness properties of high quality and invariance under reparameterization. It aims to penalize the complexity that arises when we move from a simpler base model to a more complex one, thereby preventing overfitting and adhering to *Occam's razor principle* [38]. By using the PC prior, we uphold the *principle of parsimony*, which suggests a preference for simpler models until sufficient evidence supports more complex alternatives.

The PC prior is established based on the statistical difference between the proposed complex model and its base model. We quantify this distance using the Kullback-Leibler divergence (KLD) [45], which essentially measures the information loss when we substitute a complex model with PDF  $f$  with its simpler version with PDF  $g$ . The exponential distribution is a natural choice as the appropriate base model for the GE distribution. Hence, for our purposes, we consider  $f$  and  $g$  as the GE and exponential densities, respectively.

For two continuous distributions with PDF  $f$  and  $g$  defined over the same support, KLD is defined as

$$\text{KLD}(f \parallel g) = \int_{-\infty}^{\infty} f(y) \log \left( \frac{f(y)}{g(y)} \right) dy, \quad (6)$$

where we define the distance between the two models by the 'unidirectional' distance function  $d(f \parallel g) = \sqrt{2\text{KLD}(f \parallel g)}$  [38]. The absence of symmetry in KLD is not a concern within this context. Our main focus is on quantifying the additional complexity that arises from employing the intricate model rather than the other way around.

The main idea of the PC prior involves assigning priors to the distance between two models rather than directly on the model parameters, and then by employing a change-of-variables approach, one can obtain a prior distribution for the parameter of interest. In our context, while constructing the PC prior for the shape parameter  $\alpha$ , we take this distance as a function of  $\alpha$ , i.e.,  $d(\alpha) = \sqrt{2\text{KLD}(\alpha)} \equiv \sqrt{2\text{KLD}(f \parallel g)}$ .

To incorporate the fact that the prior should have a decaying nature as a function of the distance between the two models, we take the constant rate penalization assumption and construct the PC prior by assigning an exponential prior to the distance, i.e.,  $d(\alpha) \sim \text{Exp}(\alpha_0)$

with  $\alpha_0 > 0$ ; this gives us the PC prior for  $\alpha$  as

$$\pi(\alpha) = \alpha_0 \exp[-\alpha_0 d(\alpha)] \left| \frac{\partial d(\alpha)}{\partial \alpha} \right|, \quad (7)$$

where  $\alpha_0$  is a user-defined quantity that controls the prior mass at the tail. We select this quantity in such a way that this characterizes how informative the PC prior we want. We achieve this by imposing a condition  $\Pr[d(\alpha) > U] = \xi$ , where  $U$  is the upper bound of the tail-event and  $\xi$  is the weight of the event [38].

**Theorem 4.1.** *The KLD between the GE density function in (1) and the density function of the exponential distribution with rate  $\lambda$  is given by*

$$KLD(\alpha) = \log(\alpha) + 1/\alpha - 1.$$

*Proof.* Here,  $KLD(\alpha) = KLD(f \parallel g)$ , with  $f$  being the GE density function in (1) and  $g(y) = \lambda \exp[-\lambda y]$ ,  $y > 0$ . Thus,  $KLD(\alpha)$  is obtained as

$$\begin{aligned} KLD(\alpha) &= \int_0^\infty \log \left( \frac{f(y)}{g(y)} \right) f(y) dy \\ &= \log(\alpha) \int_0^\infty f(y) dy + (\alpha - 1) \int_0^\infty \log[1 - \exp(-\lambda y)] f(y) dy \\ &= \log(\alpha) + (\alpha - 1) \int_0^\infty \log(1 - \exp[-\lambda y]) \alpha \lambda [1 - \exp(-\lambda y)]^{\alpha-1} \exp(-\lambda y) dy \\ &= \log(\alpha) - (\alpha - 1) \int_0^\infty \alpha x \cdot \exp[-x(\alpha - 1)] \exp(-x) dx \quad [\text{replacing } \exp(-x) = 1 - \exp(-\lambda y)] \\ &= \log(\alpha) + \frac{(1 - \alpha)}{\alpha}. \end{aligned}$$

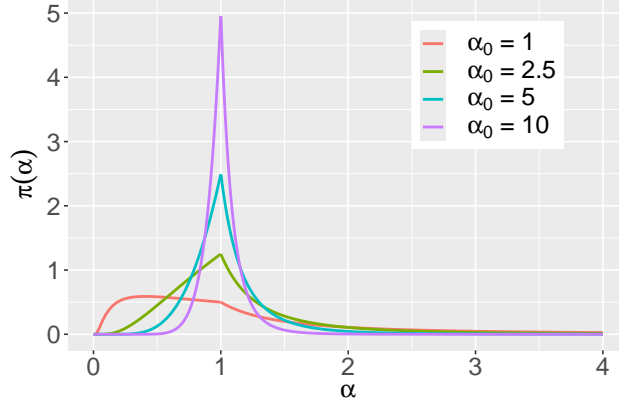
□

**Theorem 4.2.** *The PC prior, with hyperparameter  $\alpha_0$ , for the shape parameter  $\alpha$  is*

$$\pi(\alpha) = \frac{\alpha_0}{2} \exp \left( -\alpha_0 \sqrt{2 \log(\alpha) + \frac{2(1 - \alpha)}{\alpha}} \right) \left( 2 \log(\alpha) + \frac{2(1 - \alpha)}{\alpha} \right)^{-1/2} \left| \frac{1}{\alpha} - \frac{1}{\alpha^2} \right|, \quad \alpha, \alpha_0 > 0.$$

*Proof.* The result follows directly from (7) and the expression  $d(\alpha) = \sqrt{2KLD(\alpha)}$ , where  $KLD(\alpha)$  is given by Theorem 4.1. Moreover, it includes a scaling factor that ensures  $\int_0^\infty \pi(\alpha) d\alpha = 1$ . □

Figure 2 illustrates  $\pi(\alpha)$  for different hyperparameter specifications  $\alpha_0$ . We notice a proportional relationship between the value of  $\alpha_0$  and the extent of contraction to the base model. As the value of  $\alpha_0$  decreases, the tails become heavier, resulting in reduced contraction towards the base model. We also observe that for  $\alpha_0 \leq 4/3$ , the mode of the density function occurs at a value of  $\alpha$  less than one, but for  $\alpha_0 \geq 4/3$ , the mode is at  $\alpha = 1$ . While one might expect the prior would have mode at  $\alpha = 1$  irrespective of the value of  $\alpha_0$ ,



**Fig. 2** The PC prior for the shape parameter of the GE distribution for different choices of the hyperparameter  $\alpha_0$ .

we do not necessarily need the mode at  $\alpha = 1$ , and rather, we should rely on a prior that is consistent with the principles of PC prior.

## 5 Bayesian Inference

This paper employs a Bayesian estimation method to infer and quantify the uncertainty surrounding the parameters of interest. In this context, the likelihood function based on  $n$  observations from the GE distribution under the regression setting from (5) is given by

$$L(\alpha, \beta | \mathbf{y}) = \prod_{i=1}^n f(y_i; \alpha, \lambda(\mathbf{x}_i)), \quad (8)$$

where  $\mathbf{y} = (y_1, y_2, \dots, y_n)'$  is the observed data vector,  $f(y; \alpha, \lambda)$  is the GE density function in (1) and  $\lambda(\mathbf{x}_i)$  taking form as given in (5). Also, let  $\pi(\alpha)$  and  $\pi(\beta)$  denote the specified mutually independent priors for the parameters  $\alpha$  and  $\beta$ .

Combining the priors for  $\alpha$  and  $\beta$ , and the likelihood function as given in (8), we obtain the joint posterior distribution as  $\pi(\alpha, \beta | \mathbf{y}) \propto L(\alpha, \beta | \mathbf{y}) \times \pi(\alpha) \times \pi(\beta)$  from which Bayesian inference is facilitated. However, the explicit form of the marginal posterior density of the parameters is not analytically tractable, leading to employing simulation-based techniques such as Markov chain Monte Carlo (MCMC) methods or numerical approximation methods like Integrated Nested Laplace Approximations (INLA), introduced by [46]. This paper employs MCMC techniques for parameter inference, specifically utilizing the adaptive Metropolis-Hastings algorithm within Gibbs sampling. We iteratively adjust the variance of the proposal distribution within the burn-in phase so that the acceptance rate remains between 0.3 and 0.5. We initiate the MCMC chains with an initial value of 1 for  $\alpha$  and the maximum likelihood estimate for  $\beta$  as calculated under  $\alpha = 1$ . In our implementation, we update the model parameters one at a time within Gibbs sampling.

Regarding computational resources, conducting the simulations outlined in Section 6, involving 144 different configurations, each with 1000 datasets, and employing an MCMC output of 10000 iterations, required approximately 8 hours. We performed this computational task on a desktop system with an AMD Ryzen 9 5900x 12-core 24-threaded processor and 64GB of RAM. On average, the computation time for a single MCMC chain was approximately 3 seconds for our proposed semiparametric model, while it is approximately 1 second for the parametric counterpart using a simple linear form for the GE rate parameter.

Furthermore, describing the asymptotic (as  $n \uparrow \infty$  and keeping  $K$  fixed) distribution of the posterior estimates for the parameters  $\alpha$  and  $\beta$  is feasible using the Bernstein-von Mises theorem. To gauge the level of uncertainty linked with these parameter estimations, we investigate the asymptotic variance of the parameters, which is encapsulated by the inverse of the information matrix.

## 5.1 Bayesian Asymptotics

Doob's theorem [47] assures the asymptotic behavior of the posterior density; according to this theorem, assuming the sampling model  $P_{\theta}$  with  $\theta \in \Theta$  to be identifiable in the sense that  $\theta \neq \theta'$  implies  $P_{\theta} \neq P_{\theta'}$ , there exists  $\Theta_* \subseteq \Theta$  with  $\Pi(\Theta_*) = 1$  such that for each  $\theta_* \in \Theta_*$ , if  $\mathbf{Z}_n = (Z_1, \dots, Z_n)$  are iid  $P_{\theta_*}$ , then for all  $\epsilon > 0$ , we have

$$\lim_{n \rightarrow \infty} \mathbb{P}(\theta \in \mathcal{N}_{\epsilon}(\theta_*) | \mathbf{Z}_n) = 1,$$

where  $\mathcal{N}_{\epsilon}(a) = \{\theta \in \Theta : d(\theta, a) < \epsilon\}$ ,  $d$  being a metric on  $\Theta$ . The theorem implies as long as the set of possible parameter values under consideration has a positive probability assigned to it by the prior distribution, the posterior distribution obtained after incorporating new information will tend to concentrate around the true value of the parameter.

In this context, the Bernstein-von Mises theorem addresses the query concerning the form or attributes of the posterior distribution as it tends to its limiting behavior. Suppose  $\mathbf{Z}_n = (Z_1, \dots, Z_n)$  denotes the data,  $\mathbb{P}(\mathbf{Z}_n | \theta)$  is the likelihood and  $\hat{\theta}_n = \arg \max_{\theta} \mathbb{P}(\mathbf{Z}_n | \theta)$  be the MLE of  $\theta$ . Then, Fisher information matrix is defined as

$$\mathcal{I}_n(s) = \mathbb{E} \left[ \left( \frac{\partial}{\partial \theta} \log \mathbb{P}(\mathbf{Z}_n | \theta) \right)^2 \Big|_{\theta=s} \right] = -\mathbb{E} \left[ \left( \frac{\partial^2}{\partial \theta^2} \log \mathbb{P}(\mathbf{Z}_n | \theta) \right) \Big|_{\theta=s} \right],$$

where the expectation is calculated with respect to the data. In this context, Bernstein-von Mises theorem says that in the limit, the posterior distribution of  $\sqrt{n}(\theta - \hat{\theta}_n)$  approaches a Gaussian distribution. However, in the Bayesian approach, we consider conditioning on the data  $\mathbf{Z}_n$ , so that the MLE remains constant, resulting in  $\sqrt{n}(\theta - \hat{\theta}_n)$  being random solely due to the variability of the parameter  $\theta$ . Thus under sufficient regularity conditions [Section 2.2, 48], we have

$$\sqrt{n}(\theta - \hat{\theta}_n) | \mathbf{Z}_n \xrightarrow{d} N_K \left( 0, \tilde{\mathcal{I}}(\theta_*)^{-1} \right), \quad (9)$$

where  $\tilde{\mathcal{I}}(\boldsymbol{\theta}) = \lim_{n \uparrow \infty} n^{-1} \mathcal{I}_n(\boldsymbol{\theta})$ ,  $\boldsymbol{\theta}_*$  is the true parameter value, and  $K$  denotes the length of  $\boldsymbol{\theta}$ . Further, [49] discuss the corresponding results within a linear model framework, and hence, under certain regularity conditions, the above results also hold in our setup.

Following the notations in Section 3.2, we have  $Y_i | \mathbf{X}_i = \mathbf{x}_i \stackrel{\text{Indep}}{\sim} \text{GE}(\alpha, \lambda(\mathbf{x}_i))$  with  $\log[\lambda(\mathbf{x}_i)] = \mathbf{B}(\mathbf{x}_i)' \boldsymbol{\beta}$  for  $i = 1, \dots, n$ , where  $\mathbf{B}(\mathbf{x}_i) = (b_{i1}, \dots, b_{iK})'$  and  $\boldsymbol{\beta} = (\beta_1, \dots, \beta_K)'$ . Hence, the likelihood function in (8) can be rewritten as  $L(\alpha, \boldsymbol{\beta} | \mathbf{y}) = \prod_{i=1}^n f(y_i; \alpha, \exp[\mathbf{B}(\mathbf{x}_i)' \boldsymbol{\beta}])$ . Hence, the explicit form of the log-likelihood is

$$l(\alpha, \boldsymbol{\beta} | \mathbf{y}) = n \log(\alpha) + \sum_{i=1}^n S_i + (\alpha - 1) \sum_{i=1}^n \log(1 - E_i) - \sum_{i=1}^n y_i \exp(S_i),$$

with  $S_i = \mathbf{B}(\mathbf{x}_i)' \boldsymbol{\beta}$  and  $E_i = \exp\{-y_i \exp(S_i)\}$  for  $i = 1, \dots, n$ . Then, the first derivatives of the log-likelihood with respect to the parameters are computed as follows:

$$\begin{aligned} \frac{\partial l(\alpha, \boldsymbol{\beta} | \mathbf{y})}{\partial \alpha} &= \frac{n}{\alpha} + \sum_{i=1}^n \log(1 - E_i), \\ \frac{\partial l(\alpha, \boldsymbol{\beta} | \mathbf{y})}{\partial \beta_k} &= \sum_{i=1}^n b_{ik} - \sum_{i=1}^n y_i \exp(S_i) b_{ik} + (\alpha - 1) \sum_{i=1}^n \left[ \frac{y_i \exp(S_i) b_{ik}}{E_i^{-1} - 1} \right], \quad k = 1, \dots, K. \end{aligned}$$

Further, the second derivatives are computed as follows:

$$\begin{aligned} \frac{\partial^2 l(\alpha, \boldsymbol{\beta} | \mathbf{y})}{\partial \alpha^2} &= -\frac{n}{\alpha^2} \\ \frac{\partial^2 l(\alpha, \boldsymbol{\beta} | \mathbf{y})}{\partial \alpha \partial \beta_k} &= \sum_{i=1}^n \left[ \frac{y_i \exp(S_i) b_{ik}}{E_i^{-1} - 1} \right], \quad k = 1, \dots, K \\ \frac{\partial^2 l(\alpha, \boldsymbol{\beta} | \mathbf{y})}{\partial \beta_{k'} \partial \beta_k} &= -\sum_{i=1}^n [y_i b_{ik} b_{ik'} \exp(S_i)] + (\alpha - 1) \sum_{i=1}^n \left[ y_i b_{ik} \cdot \frac{\partial}{\partial \beta_{k'}} \left\{ \frac{\exp(S_i)}{E_i^{-1} - 1} \right\} \right], \\ &\quad \text{where } \frac{\partial}{\partial \beta_{k'}} \left\{ \frac{\exp(S_i)}{E_i^{-1} - 1} \right\} = \frac{\exp(S_i) b_{ik'}}{(E_i^{-1} - 1)^2} \cdot [E_i^{-1} \{1 - y_i \exp(S_i)\} - 1], \end{aligned}$$

and the results hold for all  $k, k' = 1, \dots, K$ .

We calculate the Fisher information matrix under the regression setup by taking expectations of the terms obtained above and it is given by

$$\mathcal{I}_n(\alpha, \boldsymbol{\beta}) = \begin{pmatrix} I_{\alpha, \alpha} & I_{\alpha, 1} & \cdots & I_{\alpha, K} \\ I_{1, \alpha} & I_{1, 1} & \cdots & I_{1, K} \\ \vdots & \vdots & \ddots & \vdots \\ I_{K, \alpha} & I_{K, 1} & \cdots & I_{K, K} \end{pmatrix} = - \begin{pmatrix} \mathbb{E}(J_{\alpha, \alpha}) & \mathbb{E}(J_{\alpha, 1}) & \cdots & \mathbb{E}(J_{\alpha, K}) \\ \mathbb{E}(J_{1, \alpha}) & \mathbb{E}(J_{1, 1}) & \cdots & \mathbb{E}(J_{1, K}) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbb{E}(J_{K, \alpha}) & \mathbb{E}(J_{K, 1}) & \cdots & \mathbb{E}(J_{K, K}) \end{pmatrix}, \quad (10)$$

with  $J_{\alpha, \alpha} = \frac{\partial^2 l(\alpha, \boldsymbol{\beta} | \mathbf{y})}{\partial \alpha^2}$ ,  $J_{\alpha, k} = J_{k, \alpha} = \frac{\partial^2 l(\alpha, \boldsymbol{\beta} | \mathbf{y})}{\partial \alpha \partial \beta_k}$ , and  $J_{k, k'} = \frac{\partial^2 l(\alpha, \boldsymbol{\beta} | \mathbf{y})}{\partial \beta_k \partial \beta_{k'}}; k, k' = 1, 2, \dots, K$ . For calculating the limiting posterior uncertainty of  $\sqrt{n}(\alpha, \boldsymbol{\beta}')'$ , we calculate

$\tilde{\mathcal{I}}(\alpha, \beta) = \lim_{n \uparrow \infty} n^{-1} \mathcal{I}_n(\alpha, \beta)$ . Considering the form of the double derivatives as specified in  $\mathcal{I}_n(\alpha, \beta)$ , it is evident that the integrals involved in evaluating their expectation are challenging to compute analytically. We can resort to sampling-based techniques such as Monte Carlo methods to obtain the ultimate form of the expected Fisher information matrix.

## 6 Simulation Study

We conduct an extensive simulation study to demonstrate (i) the effectiveness of the PC prior described in Section 4.1 over a conventional choice of a gamma prior and (ii) the effectiveness of the proposed semiparametric model over a parametric model, where the GE rate parameter (in log scale) is modeled as a linear or nonlinear function of the covariate(s). We design two separate simulation setups for this purpose. For the sake of simplicity and also to match our data application scenario, we assume  $P = 1$  in (5) under all the simulation settings. We consider two cases for the number of samples,  $n = 24$  and  $n = 99$ , to gain insights into scenarios with small and large sample sizes, respectively. We utilize two different covariate sequences depending on the sample size. When the sample size is  $n = 24$ , we assume that the realized values of the covariate are  $\mathcal{X} = (0.04, 0.08, \dots, 0.96)'$ , while for a larger sample size of  $n = 99$ , we assume that the realized values are  $\mathcal{X} = (0.01, 0.02, \dots, 0.99)'$ .

All the simulations use three different values of  $\alpha$ ; they are 0.5, 1, and 2. The choice of these values allows us to explore various scenarios;  $\alpha = 1$  represents the exponential distribution scenario, and in contrast, the values of  $\alpha$  being 0.5 and 2 indicate deviations from the exponential-like behavior. When  $\alpha = 1$ , the proposed PC prior would try to shrink the fitted model towards the true data-generating mechanism. However, when  $\alpha = 0.5$  or  $\alpha = 2$ , the true-generating mechanism is not an exponential regression setting, and hence, the PC prior would shrink the fitted model towards a wrong model. Thus, our goal remains to judge the advantages of the PC prior when  $\alpha = 1$  and whether the prior provides a robust estimate when  $\alpha \neq 1$ .

When the true data generating scheme is linear, we simulate data following  $Y_i|X_i = x_i \stackrel{\text{Indep}}{\sim} \text{GE}(\alpha, \lambda_i = \exp[5 + 10x_i])$ ,  $i = 1, \dots, n$ , where  $x_i$  are elements of  $\mathcal{X}$ . Here, we construct a design matrix with two columns, one for the intercept and the other for  $\mathcal{X}$ . When fitting the parametric model, we use the same design matrix employed during the linear data generation process. Alternatively, when the true data generating scheme is nonlinear, data are generated from  $Y_i|X_i = x_i \stackrel{\text{Indep}}{\sim} \text{GE}(\alpha, \lambda_i = \exp[5 + 10 \sin(2\pi x_i)])$ ,  $i = 1, \dots, n$ . In practice, we do not know the true nonlinear function, and to fit a semiparametric model, we introduce basis splines with ten basis functions for the semiparametric setup.

In the first simulation setup, we compare the PC prior with the conventional gamma prior for the shape parameter of the GE distribution under four different scenarios as follows:

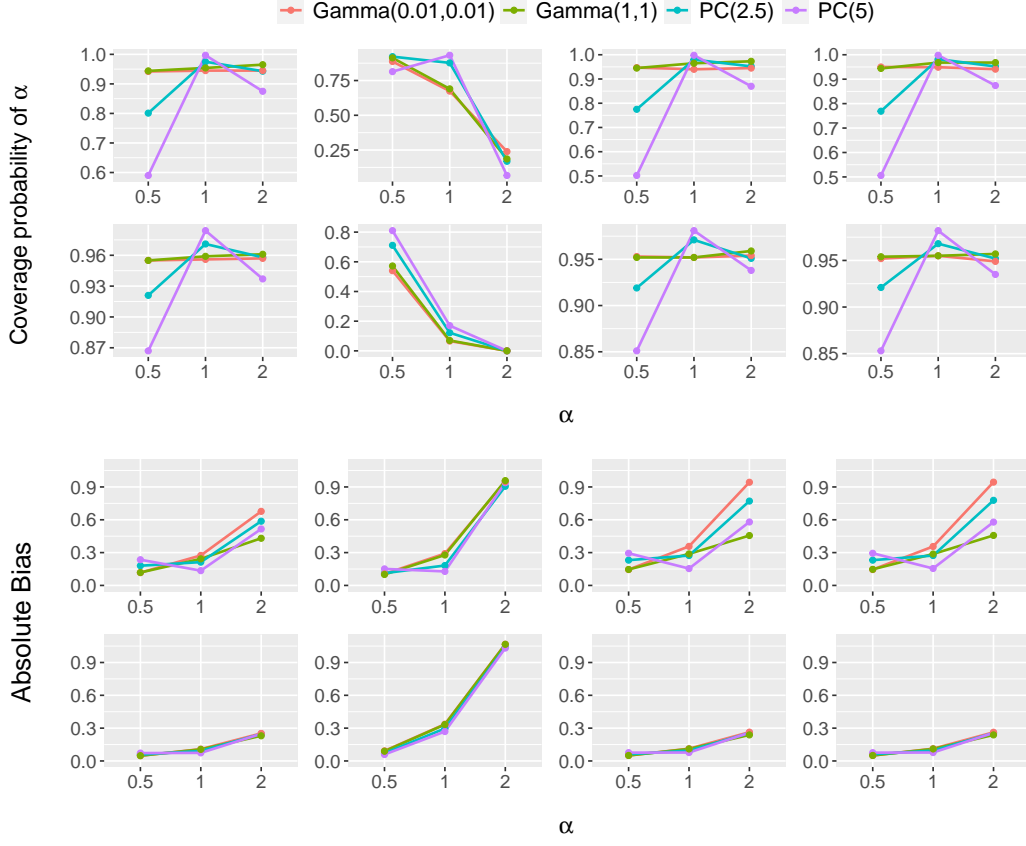
- Setting 1: Generating data from a linear setup and fitting a parametric model,
- Setting 2: Generating data from a nonlinear setup and fitting a parametric model,
- Setting 3: Generating data from a linear setup and fitting a semiparametric model,
- Setting 4: Generating data from a nonlinear setup and fitting a semiparametric model.

In the second simulation setup, we compare the efficiency of the semiparametric model with that of the parametric model. We consider four settings for this comparison as follows:

- Setting 5: Generating data from a linear setup and using a gamma prior for  $\alpha$ ,
- Setting 6: Generating data from a linear setup and employing a PC prior for  $\alpha$ ,
- Setting 7: Generating data from a nonlinear setup and using a gamma prior for  $\alpha$ ,
- Setting 8: Generating data from a nonlinear setup and employing a PC prior for  $\alpha$ .

We use four different prior specifications for the first four settings: the PC prior with parameters 2.5 and 5 and the gamma prior with parameters (0.01, 0.01) and (1, 1); we compare their effectiveness in estimation. In each case of this simulation setup, we calculate the coverage probability of  $\alpha$  based on 95% credible intervals and the absolute bias in estimating  $\alpha$  for comparison. We fit the parametric and semiparametric models for each of the last four settings and compare their goodness of fit. We also examine the variations in hyperparameters for each case. Specifically, in Settings 5 and 7, we consider the gamma prior with parameters (0.01, 0.01) and (1, 1), respectively. In Settings 6 and 8, we employ the PC prior with parameters 2.5 and 5, respectively. In every case within this simulation setup, we compute the mean absolute fitting error and the WAIC as a part of the comparison process. Here, the first metric assesses model fitting, while the latter assesses prediction performances as an alternative to cross-validation. In total, we generate 1,000 datasets from the GE distribution. We employ the Markov chain Monte Carlo (MCMC) algorithm to draw samples from the posterior distributions. We draw 20,000 MCMC samples for each dataset and discard the first 10,000 samples as burn-in. Further, we thin the chains, keeping one of five consecutive samples, and draw inferences based on the remaining 2,000 MCMC samples. We monitor the convergence and mixing of the MCMC chains using trace plots and summaries like effective sample sizes, and we obtain the desired results; we do not show the MCMC chain-related diagnostics corresponding to the simulation studies for brevity and focus on the metrics that show the advantages of the proposed methodology.

Figure 3 corresponds to the first simulation setup. The coverage probabilities are the proportion of times the true value of  $\alpha$  falls under the 95% posterior credible intervals of  $\alpha$  obtained from MCMC. The top illustration presents the coverage probabilities of different simulation setups. Four columns represent Settings 1, 2, 3, and 4, with two rows representing sample sizes of 24 and 99, respectively. Each panel showcases four prior specifications represented by different lines, while the X-axis represents the different  $\alpha$  values considered. The illustration shows that when the true value of  $\alpha$  approximates one, the PC prior exhibits superior coverage probability compared to the conventional gamma prior. This pattern holds across all four configurations in the initial simulation setup, except Setting 2, where both the gamma and PC priors yield undesirable outcomes due to attempts to fit a parametric linear model to highly nonlinear data. In the bottom illustration of Figure 3, the identical simulation setups are depicted, but instead of coverage probabilities, it focuses on illustrating the absolute bias in the estimation of  $\alpha$ ; this illustration highlights a reduction in estimation bias when  $\alpha$  equals one, aligning with the inherent characteristic of the

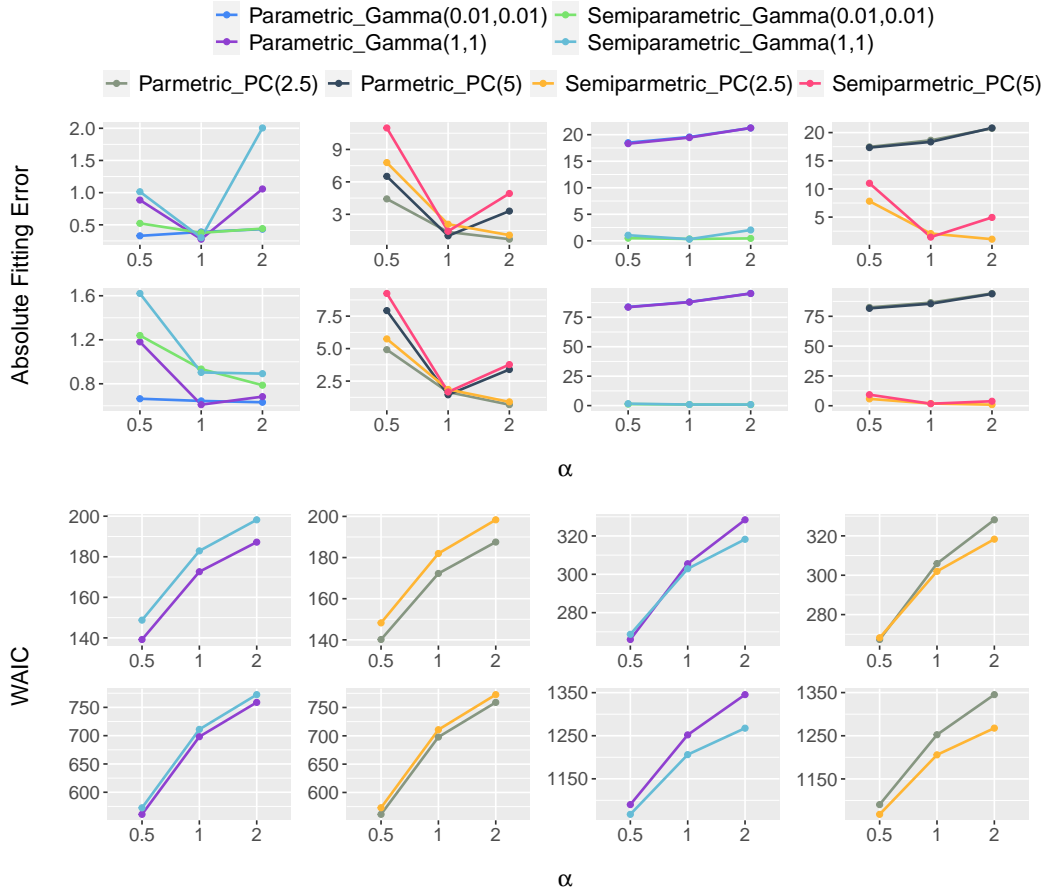


**Fig. 3** Coverage probabilities (top illustration) and absolute bias values (bottom illustration) computed based on imposing a PC prior. For each illustration, from the left-most to the right-most columns: (i) Setting 1 with  $n = 24, 99$ , (ii) Setting 2 with  $n = 24, 99$ , (iii) Setting 3 with  $n = 24, 99$ , (iv) Setting 4 with  $n = 24, 99$ . A higher coverage probability in the top illustration and a lower value of absolute bias in the bottom illustration indicates a better model fit.

PC prior to shrink the estimate towards the base model. Additionally, as the sample size increases, the influence of the prior gradually diminishes; this pattern is evident as the lines representing absolute bias nearly overlap, regardless of the chosen prior.

Figure 4 focuses on the second simulation setup. In the top illustration, we compare the goodness of fit between the parametric and semiparametric models using the absolute fitting error. The four columns represent Settings 5, 6, 7, and 8, with two rows representing sample sizes of 24 and 99. Each panel displays specifications of different values of  $\alpha$  at the X-axis, and lines correspond to either the parametric setup or the semiparametric setup, with varying hyperparameter specifications. The top illustration highlights a substantial increase in the absolute fitting error as we transition from a dataset demonstrating a linear trend to one illustrating a nonlinear trend. However, the plots associated with settings involving the generation of linear data and fitting a parametric model (Settings 5 and 6) exhibit minimal differences and occasionally yield superior results. When examining the settings





**Fig. 4** Absolute fitting error (top illustration) and WAIC values (bottom illustration) based on fitting a semiparametric GE regression model. For each illustration, from the left-most to the right-most columns: (i) Setting 5 with  $n = 24, 99$ , (ii) Setting 6 with  $n = 24, 99$ , (iii) Setting 7 with  $n = 24, 99$ , (iv) Setting 8 with  $n = 24, 99$ . (For the bottom illustration, the lines representing the gamma prior with parameters (0.01, 0.01) and the PC prior with  $\theta = 5$  are excluded since they closely resemble the demonstrated prior specification). A lower value of absolute fitting error indicates a better model fit, while a lower value of WAIC indicates a better out-of-sample prediction performance.

where data originates from a nonlinear context (Settings 7 and 8), a distinct and considerable gap emerges between the lines representing the parametric and semiparametric models, where the latter consistently performs less favorably than the former, occupying a lower position on the graph. The bottom illustration provides a similar comparison but focuses on the WAIC; a smaller WAIC indicates a better out-of-sample prediction performance. Despite Settings 5 and 6 pertaining to data generated from a linear setup, we observe that the parametric model consistently exhibits lower WAIC scores. In Settings 7 and 8, this ordering of WAIC scores is entirely reversed, with the lines representing the semiparametric model falling consistently below those of the parametric model. This reversal underscores the clear advantage of employing semiparametric modeling, particularly in high nonlinearity scenarios.

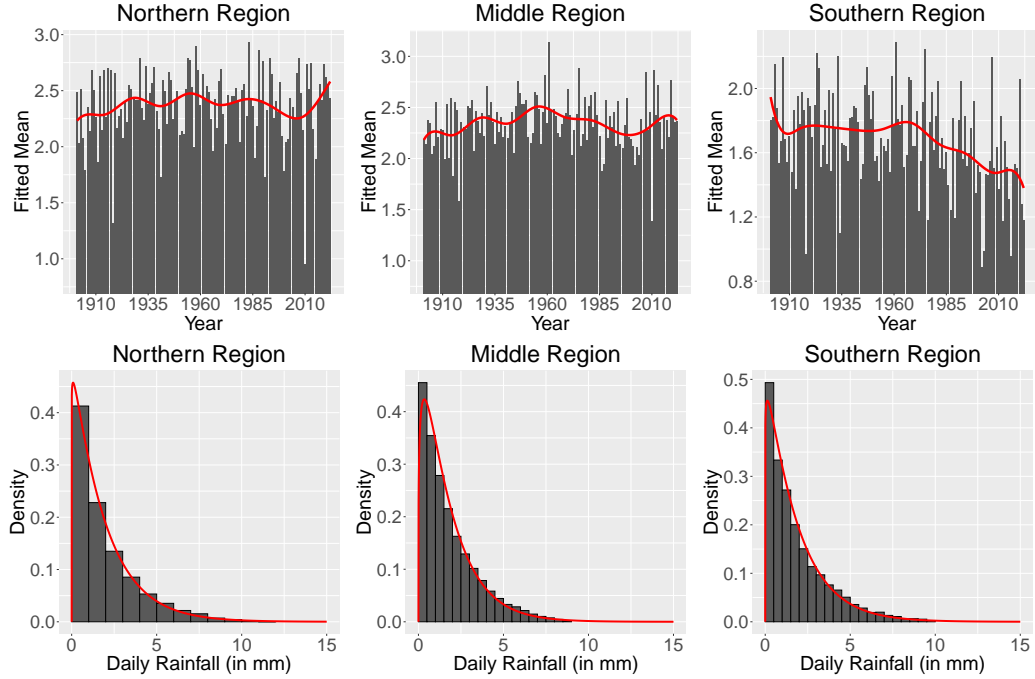
## 7 Data Application

We obtain daily gridded rainfall (in mm) data over the Western Ghats of India region with a spatial resolution of  $1.0^\circ \times 1.0^\circ$ , covering the period from 1901–2022. The dataset is publicly available at the official website of the Indian Meteorological Department (IMD), Pune ([https://www.imdpune.gov.in/cmpg/Griddata/Rainfall\\_1\\_NetCDF.html](https://www.imdpune.gov.in/cmpg/Griddata/Rainfall_1_NetCDF.html)). The gridded data product was prepared by IMD through spatial interpolation of ground-station data following the procedure described in [50]. We extract the daily rainfall information for the monsoon months of June, July, August, and September (JJAS) throughout 1901–2022. Additionally, we exclude days within the JJAS months where recorded rainfall amounts were zero. Out of the pixels representing the Western Ghats area, we group them into three distinct significant regions: the Northern, Middle, and Southern regions (the regions are shown in the supplementary material). We compute the daily rainfall values for each region by calculating the average of the corresponding pixel values within that region. Afterward, we conduct a separate analysis for each of these regions.

### 7.1 Exploratory Data Analysis

Given our dataset (after preprocessing) spans over a century, our initial focus involves performing necessary analyses to address any potential trends within the data. In the top panels of Figure 5, we present a bar diagram depicting the average yearly rainfall for each year. No clear long-term linear trend is observable for any of the three regions. However, several short-term upward and downward patterns are noticeable. We use a basis spline regression approach to explore such short-term trends, which treats daily rainfall values as response variables and corresponding years as covariates. Considering the residuals from this regression, we can effectively eliminate any potential trends embedded in the data. We overlap the estimated means with the bar diagrams in the top panels of Figure 5. Firstly, the estimated mean curve aligns well with the visualized bar diagram. Moreover, both components highlight the presence of a nonstationary rainfall pattern. This pattern, in turn, underscores the suitability of employing a semiparametric regression model, which can effectively accommodate and incorporate these nonstationary patterns within the data.

Subsequently, after removing outliers via the popular adjusted-boxplot method developed by [51], we present two critical visualizations in the bottom panels of Figure 5, where the panels correspond to three regions of interest. We first showcase histograms illustrating the distribution of the detrended residuals obtained by exponentiating the residuals obtained by fitting a semiparametric regression curve to the log-transformed wet-day rainfall observations, which aligns with the standard link function formulations for generalized additive models. Additionally, a red line denotes the fitted density of the GE distribution, with parameters estimated from the detrended residuals. We observe a strong alignment between the estimated density and the associated histograms, indicating a favorable fit. This visual representation significantly supports the rationale behind the semiparametric GE regression model proposed in this paper. Additionally, the plots highlight a marked resemblance of the



**Fig. 5** Bar diagrams of the annual average wet-day rainfall during June through September along with fitted mean curves based on twelve cubic B-splines with equidistant knots, for Northern, Middle, and Southern Western Ghats regions (top panels). Histograms of the detrended residuals from the daily rainfall overlapped with the fitted GE densities (bottom panels).

GE distribution towards its foundational model, the exponential distribution, which has a density with mode at zero and then decays exponentially. This observation also reinforces our second consideration of using a novel distance-based prior for the shape parameter of the GE distribution.

## 7.2 Model Specification

In our case, except for the rainfall measures and the corresponding year information, we do not have access to additional covariates. Hence, we consider the year  $T$  as a covariate in our analysis and assume it is defined over the continuous interval  $[1901, 2022]$ , i.e. if  $Y_t$  denotes a rainfall measure for year  $t$ , our model is given by  $Y_t|T = t \stackrel{\text{Indep}}{\sim} \text{GE}(\alpha, \lambda(t))$ ,  $t = 1901, \dots, 2022$ . We conduct the analysis using two distinct models based on two different choices of  $\lambda(t)$ . The first model employs a parametric approach to model the rate parameter, while the second model is our proposed semiparametric formulation. We employ a simple linear regression model for the rate parameter in the parametric setting, given as  $\lambda_{(L)}(t)$  in (11). On the other hand, for the semiparametric regression, we adopt the basis spline

regression form presented in (4) and choose  $\lambda_{\text{NL}}(t)$  in (11) as

$$\lambda_{\text{L}}(t) = \exp(\beta_0 + \beta_1 t), \quad \lambda_{\text{NL}}(t) = \exp \left[ \sum_{k=1}^K \beta_k B_k(t) \right]. \quad (11)$$

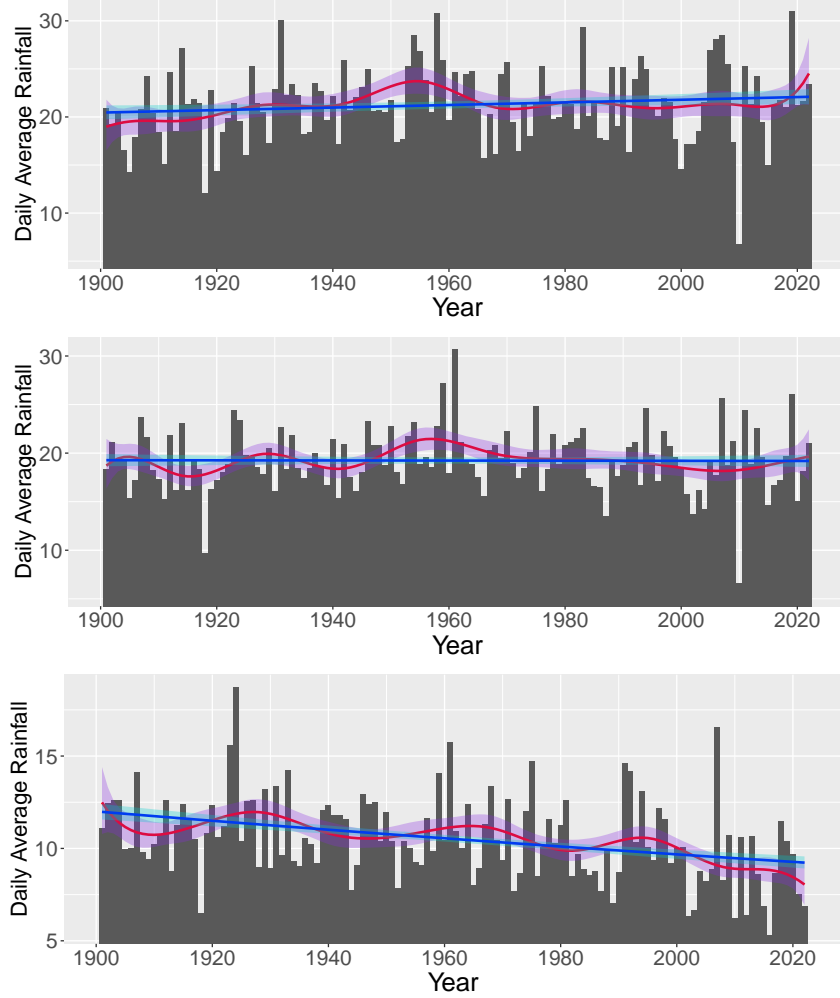
We employ Bayesian methods to estimate the model parameters. As the exploratory analysis shows a strong resemblance of the considered GE distribution towards its base model, i.e., the exponential distribution, we use the proposed PC prior for the shape parameter  $\alpha$  and choose independent weakly-informative Gaussian priors with mean zero and variance 100 for the regression parameters. As discussed in Section 5, we use MCMC techniques to draw inferences about the model parameters.

For our analysis, we use  $K = 12$  cubic B-spline basis functions with equidistant knots. With data available across 122 years, adopting 12 splines enables us to effectively capture decadal patterns using each spline [52]. Here, the hyperparameter  $\alpha_0$  for the PC prior being a tuning parameter, we compute WAIC values across a range of  $\alpha_0$  values, spanning within the interval  $\{0.5, 1, \dots, 5\}$ , and achieve the most precise fit for our semiparametric regression. After individually examining the northern, middle, and southern regions, we identify the optimal values for  $\alpha_0$  that yield the lowest WAIC values, and the optimal choices for these regions are  $\alpha_0 = 4.5$ ,  $\alpha_0 = 3.5$ , and  $\alpha_0 = 1.5$ , respectively. As a result, these optimal  $\alpha_0$  values are employed for their respective regions during the final model fitting stage.

For all three regions, we fit the two competing models in (11). For each of the model fits, we generate 50,000 MCMC samples for each model parameter. The initial 15,000 samples are removed as burn-in and subsequently excluded from the analysis. Additionally, we employ a thinning interval of 5, and finally, we have 7,000 posterior samples for drawing inference. To evaluate the convergence and mixing of the chains derived from the MCMC process, we visualize the trace plots of the parameters associated with both model fits in the supplementary materials. Specifically, we present the trace plots of the shape parameter  $\alpha$  for each region. The regression parameters also exhibit similar satisfactory mixing and convergence.

### 7.3 Model Comparison

In this section, we compare the results based on two models mentioned in (11), and we present the estimated mean daily rainfall on wet days for each year between 1901 and 2022 for the Northern, Middle, and Southern Western Ghats regions and two competing models in Figure 6. Across all three regions, a noticeable trend emerges: the semiparametric models exhibit a notably superior fit. This distinction becomes evident as we observe multiple abrupt fluctuations in the bars representing the annual averages of wet-day precipitation. Remarkably, the semiparametric model effectively captures these fluctuations. Particularly noteworthy is the ability of the semiparametric model to accurately capture the nonstationarity present in the precipitation patterns. This heightened ability to encapsulate the dynamic variations in precipitation is a notable strength of the semiparametric model fitting. We also present the pointwise 95% credible intervals for the trajectories estimated from



**Fig. 6** Estimated mean of daily wet-day rainfall (in mm) with semiparametric (red line) and parametric (blue line) models in (11), along with corresponding pointwise 95% credible intervals (ribbons). The top, middle, and bottom panels show the results for the Northern, Middle, and Southern Western Ghats regions.

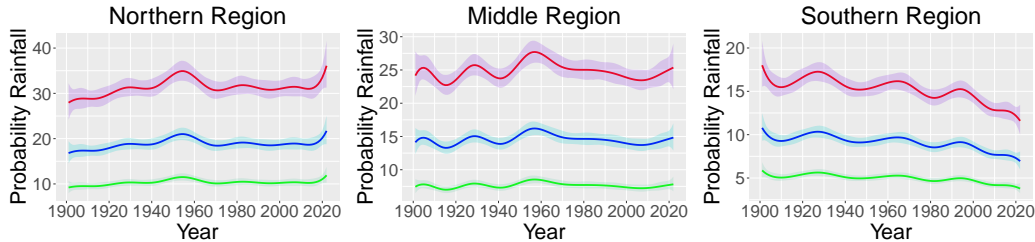
the MCMC samples. The credible bands based on the semiparametric model are generally wider than the ones based on the parametric GE regression model; semiparametric models generally provide robust estimates but have higher uncertainty due to a bias-variance trade-off, and we observe the same here as well.

#### 7.4 Inferences about Western Ghats Rainfall

In this subsection, we discuss results based on fitting the proposed GE regression model with a PC prior for the GE shape parameter. The estimated GE shape parameters (posterior means) for the Northern, Middle, and Southern Western Ghats regions are 0.859, 0.949, and 0.873, respectively, and the corresponding posterior standard deviations are 0.096, 0.100,

and 0.097, respectively. These shape parameter values indicate a pronounced alignment with the exponential distribution of wet-day rainfall. Consistent with the fluctuating pattern in the annual average of daily wet-day rainfall, the fitted mean lines for each region also demonstrate short-term fluctuations. Besides, a consistent and stable mean rainfall trend is noticeable across the Northern and Middle Western Ghats regions. However, in the Southern Western Ghats region, the fitted parametric and semiparametric models distinctly reveal a decaying pattern in the annual averages of daily wet-day rainfall.

We present two significant insights into the rainfall patterns within these regions: the overarching decade-long shifts and individual region-specific probability rainfall plots. The calculation of the decadal change involves determining the overall rainfall shift and dividing it by the number of decades, resulting in  $\{\mu(2022) - \mu(1901)\}/12.1$ , where  $\mu(t) = \lambda(t)^{-1} [\psi(\alpha + 1) - \psi(1)]$ , and  $t$  representing the corresponding year; here,  $\psi(\cdot)$  denotes the digamma function. Subsequently, the estimated decadal shifts in rainfall amount are 0.458 mm, 0.078 mm, and -0.367 mm for the northern, middle, and southern regions, respectively. In Figure 7, we display the probability rainfall graphs for three distinct probabilities: 0.3 (red line), 0.5 (blue line), and 0.7 (green line); in agrometeorology,  $100p\%$  probability rainfall means the  $(1 - p)^{th}$  quantile of the probability distribution of rainfall. These plots hold significant implications in agriculture as they empower farmers to formulate their harvesting strategies based on the anticipated likelihood of rainfall, allowing them to align their plans with the rainfall patterns to fulfill their specific requirements. Figure 7 showcases the estimated probability rainfall graphs, along with pointwise 95% credible intervals for the estimated rainfall. We derive these intervals from the MCMC samples; they illustrate the uncertainty associated with the estimation process. For 70% probability rainfall, the pointwise credible bands are wider than the other two probability levels.



**Fig. 7** The 30% (red line), 50% (blue line), and 70% (green line) probability rainfall (in mm) with corresponding point-wise 95% credible intervals (ribbons).

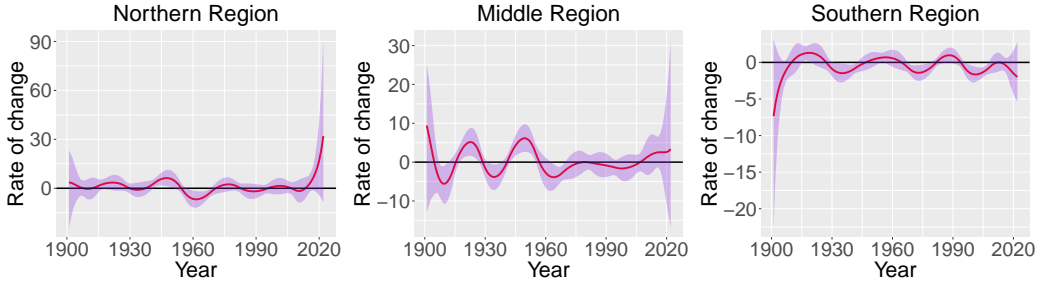
As a crucial component of our comprehensive analysis, we further discuss the dynamic nature of annual average rainfall for the Western Ghats range over the past century by exploring the plot of its rate of change. Interpreting this quantity unveils insights into trends, variations, and shifts in mean values over time, offering glimpses into rainfall behavior. A higher absolute magnitude implies fast changes, while a lower one indicates gradual shifts. A positive or negative rate means an increasing or decreasing mean rainfall over time,

potentially signaling rising or reducing annual average rainfall. We compute this quantity by taking the derivative of the fitted mean from our semiparametric model with respect to the time component from  $\lambda_{NL}(t)$  in (11), given by

$$\frac{\partial\mu(t)}{\partial t} = -\frac{\psi(\alpha+1) - \psi(1)}{[\lambda(t)]^2} \sum_{k=1}^K \beta_k \frac{\partial B_k(t)}{\partial t}, \quad (12)$$

where we compute the derivatives of the cubic B-splines using `fda` package [53] in R.

Figure 8 illustrates varying trends in the rate of change in mean annual rainfall over the years across the three regions of the Western Ghats range. Initially, the Northern and Middle regions exhibit more pronounced fluctuations in the rate-of-change graphs than the Southern region. This pattern suggests more rapid variations in rainfall trends in the Northern and Middle regions, while a more stable rainfall pattern is visible for the Southern area. Moreover, the small-scale positive and negative rate-of-change instances are well-balanced for the Northern and Middle regions. This pattern implies that over the past century, changes in rainfall have been relatively symmetric in terms of increase and decrease, with no significant alterations in long-term patterns. In contrast, the Southern region displays a substantial portion of years with graphs below the zero line, signifying a prevalent decreasing trend in rainfall. The rate-of-change in mean for the last 30 years shows consistent negative values in the Southern sector, indicating the declining rainfall trend, while the graphs for the other two regions consistently exhibit positive values, indicating an increasing trend in rainfall over the past three decades in those areas. The pointwise 95% credible intervals include the zero line for the Northern and Middle regions; hence, the positive values for the last years are not significant. On the other hand, while the posterior mean rate-of-change remains negative for the Southern region in general, the credible intervals indicate that the negative values of rate-of-change are significant for several timestamps; however, the positive values are generally insignificant.



**Fig. 8** Rate of change in the annual average of daily wet-day rainfall in the monsoon months across the year 1901–2022, given by  $\frac{\partial\mu(t)}{\partial t}$  in (12) (red line), and the corresponding pointwise 95% credible intervals (ribbons). The black line represents the zero value.

## 8 Discussions and Conclusions

With its shape and rate parameters, the generalized exponential (GE) distribution facilitates more rigorous skewness attributes than several other distributions, specifically the exponential distribution. Thus, it is a better choice as a potential flexible model to incorporate high positive skewness in the data. Additionally, depending on the shape parameter, the varying nature of the hazard function provides GE distribution more compatibility in fitting into complex data structures. In the regression arena, semiparametric regression is a powerful statistical method that combines the flexibility of the nonparametric models with the interpretability and efficiency of the parametric models. The superiority of our proposed model in capturing nonlinearity compared to the corresponding parametric model is depicted in Sections 6 and 7. On the other hand, PC prior is a principled distance-based prior that penalizes departure from a base model. It is used for specifying priors on parameters that are difficult to elicit directly from expert knowledge. This paper introduces a PC prior for the GE shape parameter, with the motivation of driving the GE distribution closer to the characteristics of the exponential distribution, a well-known probability distribution model for classical rainfall modeling.

The proposed semiparametric GE regression model reasonably fits the wet-day rainfall data for the Northern, Middle, and Southern regions of the Western Ghats mountain range of India. We observe a consistent overall trend with periodic fluctuations in the Northern and Middle Western Ghats regions. However, a declining trend is prominent in the Southern Western Ghats region. This observation is further supported by the decadal analysis of rainfall changes in these three regions, where only the Southern region exhibits a clear and significant negative value indicating the effects of climate change. This research not only enhances our comprehension of the intricate climatic dynamics within the Western Ghats but also emphasizes the critical role of precise predictive models in anticipating seasonal rainfall variations.

There are several directions for extending this research. In addition to modeling the rate parameter, we can consider treating the shape parameter as a time-dependent variable. Instead of utilizing splines for the rate parameter, an alternative approach could involve employing a Gaussian process prior. Moreover, to ensure the applicability of our comparisons to large datasets, we may explore various approximation techniques like Gaussian Markov random fields [54]. While this paper has primarily focused on the temporal analysis of rainfall data, further enhancements can be made by incorporating spatial components [55]. This extension involves investigating the variability in rainfall patterns across diverse geographical regions or watersheds [56]. Additionally, there is potential for developing a real-time rainfall prediction system, offering timely information for tasks such as flood forecasting, reservoir management, and emergency response, based on the foundation provided by this model. For the high-dimensional spatial problems, our model can be implemented as a two-stage model where the GE parameters can be estimated at each spatial location, ignoring the spatial structure, and those estimates can be smoothed using a Gaussian process [57].



## Supplementary information

Supplementary materials include the trace plots of the MCMC chains of the shape parameter of the generalized exponential regression model, under the proposed semiparametric model and the competing parametric model. We further provide a map identifying the Northern, Middle, and Southern Western Ghats regions on the map of India. Subsequently, we provide some details about selecting the hyperparameter of the proposed penalized complexity prior for the shape parameter of the generalized exponential regression model. Codes (written in R) for obtaining MCMC samples from the posterior distribution of the parameters of the proposed semiparametric GE regression model and the processed dataset analyzed in this paper are also provided.

## Disclosure statement

No potential conflict of interest was reported by the authors.

## References

- [1] Mathew, M.M., Sreelash, K., Mathew, M., Arulbalaji, P., Padmalal, D.: Spatiotemporal variability of rainfall and its effect on hydrological regime in a tropical monsoon-dominated domain of Western Ghats, India. *Journal of Hydrology: Regional Studies* **36**, 100861 (2021)
- [2] Venkatesh, B., Nayak, P., Thomas, T., Jain, S.K., Tyagi, J.: Spatio-temporal analysis of rainfall pattern in the Western Ghats region of India. *Meteorology and Atmospheric Physics* **133**, 1089–1109 (2021)
- [3] Varikoden, H., Revadekar, J., Kuttippurath, J., Babu, C.: Contrasting trends in south-west monsoon rainfall over the Western Ghats region of India. *Climate Dynamics* **52**, 4557–4566 (2019)
- [4] Abe, G., James, E.: Impacts of anthropogenic regulation on streamflow in the humid tropics of Western Ghat regions of Kerala state. *International Journal of Advances in Engineering & Technology* **6**(4), 1895 (2013)
- [5] Veerabhadranavar SA, V.B.: Assessment of Impact of Climate Change in the Western Ghats Region, India. *Indian Journal of Science and Technology*, 1466–1472 (2022)
- [6] Todorovic, P., Woolhiser, D.: Stochastic model of daily rainfall. *Miscellaneous Publication, US Department of Agriculture* **1275**, 232–246 (1974)
- [7] Hazra, A., Bhattacharya, S., Banik, P.: A Bayesian zero-inflated exponential distribution model for the analysis of weekly rainfall of the Eastern Plateau Region of India. *Mausam* **69**(1), 19–28 (2018)

- [8] Ahmad, Z., Hamedani, G.G., Butt, N.S.: Recent developments in distribution theory: a brief survey and some new generalized classes of distributions. *Pakistan Journal of Statistics and Operation Research*, 87–110 (2019)
- [9] Marshall, A.W., Olkin, I.: A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families. *Biometrika* **84**(3), 641–652 (1997)
- [10] Jorgensen, B.: Statistical properties of the generalized inverse Gaussian distribution. *Lecture Notes in Statistics* **9** (1982)
- [11] Kundu, D., Raqab, M.Z.: Generalized Rayleigh distribution: different methods of estimations. *Computational Statistics & Data Analysis* **49**(1), 187–200 (2005)
- [12] Tahir, M.H., Cordeiro, G.M.: Compounding of distributions: a survey and new generalized classes. *Journal of Statistical Distributions and Applications* **3**, 1–35 (2016)
- [13] Gupta, R.D., Kundu, D.: Theory & methods: Generalized exponential distributions. *Australian & New Zealand Journal of Statistics* **41**(2), 173–188 (1999)
- [14] Gupta, R.D., Kundu, D.: Exponentiated exponential family: an alternative to gamma and Weibull distributions. *Biometrical Journal: Journal of Mathematical Methods in Biosciences* **43**(1), 117–130 (2001)
- [15] Gupta, R.D., Kundu, D.: Generalized exponential distribution: different method of estimations. *Journal of Statistical Computation and Simulation* **69**(4), 315–337 (2001)
- [16] Jaheen, Z.F.: Empirical Bayes inference for generalized exponential distribution based on records. *Communications in Statistics-Theory and Methods* **33**(8), 1851–1861 (2004)
- [17] Raqab, M.Z., Madi, M.T.: Bayesian inference for the generalized exponential distribution. *Journal of Statistical Computation and Simulation* **75**(10), 841–852 (2005)
- [18] Kundu, D., Gupta, R.D.: Generalized exponential distribution: Bayesian estimations. *Computational Statistics & Data Analysis* **52**(4), 1873–1883 (2008)
- [19] Gupta, R.D., Kundu, D.: Generalized exponential distribution: Existing results and some recent developments. *Journal of Statistical Planning and Inference* **137**(11), 3537–3547 (2007)
- [20] Madi, M.T., Raqab, M.Z.: Bayesian prediction of rainfall records using the generalized exponential distribution. *Environmetrics* **18**(5), 541–549 (2007)
- [21] Hazra, A.: Minimum density power divergence estimation for the generalized exponential distribution. *Communications in Statistics-Theory and Methods*, 1–21 (2024)

- [22] Markiewicz, I., Strupczewski, W.G., Bogdanowicz, E., Kochanek, K.: Generalized exponential distribution in flood frequency analysis for Polish rivers. *PloS One* **10**(12), 1–26 (2015)
- [23] Aslam, M., Shahbaz, M.Q.: Economic reliability test plans using the generalized exponential distribution. *Journal of Statistics* **14**(1), 53–60 (2007)
- [24] Cota-Felix, J.E., Rivas-Davalos, F., Maximov, S.: An alternative method for estimating mean life of power system equipment with limited end-of-life failure data. In: 2009 IEEE Bucharest PowerTech, pp. 1–4 (2009). IEEE
- [25] Sarhan, A.M.: Analysis of incomplete, censored data in competing risks models with generalized exponential distributions. *IEEE Transactions on Reliability* **56**(1), 132–138 (2007)
- [26] Wigena, A.H., Djuraidah, A., Rizki, A.: Semiparametric modeling in statistical down-scaling to predict rainfall. *Applied Mathematical Sciences* **9**(88), 4371–4382 (2015)
- [27] Nguyen-Huy, T., Deo, R.C., Mushtaq, S., Khan, S.: Probabilistic seasonal rainfall forecasts using semiparametric  $d$ -vine copula-based quantile regression. In: *Handbook of Probabilistic Models*, pp. 203–227. Elsevier, Oxford, UK (2020)
- [28] Hastie, T.J., Tibshirani, R.J.: *Generalized Additive Models*. CRC press, Boca Raton, USA (1990)
- [29] Ghosal, S., Vaart, A.: *Fundamentals of Nonparametric Bayesian Inference* vol. 44. Cambridge University Press, Cambridge (2017)
- [30] Gelfand, A.E., Kottas, A.: Bayesian semiparametric regression for median residual life. *Scandinavian Journal of Statistics* **30**(4), 651–665 (2003)
- [31] Lee, J., Sison-Mangus, M.: A Bayesian semiparametric regression model for joint analysis of microbiome data. *Frontiers in Microbiology* **9**, 522 (2018)
- [32] Li, L., Hanson, T.E.: A Bayesian semiparametric regression model for reliability data using effective age. *Computational Statistics & Data Analysis* **73**, 177–188 (2014)
- [33] Fahrmeir, L., Lang, S.: Bayesian semiparametric regression analysis of multicategorical time-space data. *Annals of the Institute of Statistical Mathematics* **53**, 11–30 (2001)
- [34] Koop, G., Poirier, D.J.: Bayesian variants of some classical semiparametric regression techniques. *Journal of Econometrics* **123**(2), 259–282 (2004)
- [35] Kim, C., Song, S.: Bayesian estimation of the parameters of the generalized exponential distribution from doubly censored samples. *Statistical Papers* **51**, 583–597 (2010)

- [36] Naqash, S., Ahmad, S., Ahmed, A.: Bayesian analysis of generalized exponential distribution. *Journal of Modern Applied Statistical Methods* **15**(2), 38 (2016)
- [37] Dey, S.: Bayesian estimation of the shape parameter of the generalised exponential distribution under different loss functions. *Pakistan Journal of Statistics and Operation Research*, 163–174 (2010)
- [38] Simpson, D., Rue, H., Riebler, A., Martins, T.G., Sørbye, S.H.: Penalising model component complexity: A principled, practical approach to constructing priors. *Statistical Science* **32**(1), 1–28 (2017)
- [39] Van Niekerk, J., Bakka, H., Rue, H.: A principled distance-based prior for the shape of the Weibull model. *Statistics & Probability Letters* **174**, 109098 (2021)
- [40] Ventrucci, M., Rue, H.: Penalized complexity priors for degrees of freedom in Bayesian p-splines. *Statistical Modelling* **16**(6), 429–453 (2016)
- [41] Ordoñez, J.A., Prates, M.O., Bazán, J.L., Lachos, V.H.: Penalized complexity priors for the skewness parameter of power links. *Canadian Journal of Statistics* (Published online) (2023)
- [42] Sørbye, S.H., Rue, H.: Penalised complexity priors for stationary autoregressive processes. *Journal of Time Series Analysis* **38**(6), 923–935 (2017)
- [43] Hazra, A., Ghosh, A.: Robust statistical modeling of monthly rainfall: The minimum density power divergence approach. *Sankhya B* **86**(1), 241–279 (2024)
- [44] Watanabe, S., Opper, M.: Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. *Journal of Machine Learning Research* **11**, 3571–3594 (2010)
- [45] Kullback, S., Leibler, R.A.: On information and sufficiency. *The Annals of Mathematical Statistics* **22**(1), 79–86 (1951)
- [46] Rue, H., Martino, S., Chopin, N.: Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **71**(2), 319–392 (2009)
- [47] Doob, J.L.: Application of the theory of martingales. *Le calcul des probabilités et ses applications*, 23–27 (1949)
- [48] Patriota, A.G.: A q-exponential regression model. *Sankhya B* **74**, 149–170 (2012)
- [49] Ghosal, S.: Asymptotic normality of posterior distributions in high-dimensional linear models. *Bernoulli* **5**(11), 315–331 (1999)

- [50] Rajeevan, M., Bhate, J., Jaswal, A.K.: Analysis of variability and trends of extreme rainfall events over India using 104 years of gridded daily rainfall data. *Geophysical Research Letters* **35**(18) (2008)
- [51] Hubert, M., Vandervieren, E.: An adjusted boxplot for skewed distributions. *Computational Statistics & Data Analysis* **52**(12), 5186–5201 (2008)
- [52] Hazra, A., Reich, B.J., Staicu, A.-M.: A multivariate spatial skew-t process for joint modeling of extreme precipitation indexes. *Environmetrics* **31**(3), 2602 (2020)
- [53] Ramsay, J.: *fda: Functional Data Analysis*. (2023). R package version 6.1.4. <https://CRAN.R-project.org/package=fda>
- [54] Rue, H., Held, L.: *Gaussian Markov Random Fields: Theory and Applications*. CRC press, New York, USA (2005)
- [55] Gotway, C.A., Stroup, W.W.: A generalized linear model approach to spatial data analysis and prediction. *Journal of Agricultural, Biological, and Environmental Statistics*, 157–178 (1997)
- [56] Yang, C., Chandler, R., Isham, V., Wheeler, H.: Spatial-temporal rainfall simulation using generalized linear models. *Water Resources Research* **41**(11) (2005)
- [57] Hazra, A., Huser, R., Jóhannesson, Á.V.: In: Hrafnkelsson, B. (ed.) *Bayesian latent Gaussian models for high-dimensional spatial extremes*, pp. 219–251. Springer, Cham (2023)