Hacettepe Journal of Mathematics & Statistics

Hacet. J. Math. Stat.
Volume 51 (6) (2022), 1723 – 1735
DOI : 10.15672/hujms.682730

Research Article

Bayesian estimation of bivariate Pickands dependence function

Alireza Ahmadabadi[1], Gholamhossien Gholami[2], Burcu Hudaverdi*[1]

[1] *Dokuz Eylul University, Department of Statistics, Izmir*
[2] *Urmia University, Department of Mathematics, Orumiyeh*

Abstract

In the present study, Bayesian method of estimating the Pickands dependence function of bivariate extreme-value copulas is proposed. Initially, cubic B-spline regression is used to model the dependence function. Then, the estimator of Pickands dependence function is obtained by the Bayesian approach. Through the estimation process, the prior and the posterior distributions of the parameter vectors are provided. The posterior sampling algorithm is presented in order to approximate the posterior distribution. We give a simulation study to measure and compare the performance of the proposed Bayesian estimator of the Pickands dependence function. A real data example is also illustrated.

Mathematics Subject Classification (2020). 62F15, 62F40, 60G70

Keywords. Bivariate extreme value copula, Pickands dependence function, Bayesian method, reversible jump MCMC

1. Introduction

There has been a growing interest for multivariate modeling of extreme values especially in the fields of environmental and financial studies. The bivariate extreme-value distributions are related to the limiting distributions of componentwise maxima of a given sample. Dependence structure of bivariate rare events like in asset pricing and portfolio choice can be modelled using extreme value distributions. In this context, inference methods for modeling the extreme-value dependence structure are getting improved. The stable tail dependence function and the spectral measure are used to model the extreme-value dependence structure through extreme-value copulas. Consider the random pair (X, Y) with joint distribution function H and continuous marginal distributions F and G, respectively, then following [17], the pair (X, Y) has extreme value dependence if, and only if, its copula C can be expressed, for all u, $v \in$ (0,1), in the form

$$C(u, v) = P(F(X) \leq u, G(Y) \leq v) = \exp\left(\ln(uv) A\left[\frac{\ln v}{\ln(uv)}\right]\right), \tag{1.1}$$

where A belongs to the class $\mathcal{A}$ which is defined by $\mathcal{A} = \{A : [0, 1] \to \mathcal{R} | A$ is convex and $\max(t, 1-t) \leq A(t) \leq 1$ for all $t \in [0, 1]\}$.

*Corresponding Author.
Email addresses: alireza.ahmadabadi@ogr.deu.edu.tr (A. Ahmadabadi), gh.gholami@urmia.ac.ir (G. Gholami), burcu.hudaverdi@deu.edu.tr (B. Hudaverdi)
Received: 31.01.2020; Accepted: 01.07.2022

Estimation of the Pickands dependence function is the main step in modeling the multivariate extremes. Pickands [17] and Deheuvels [8] had early works in estimating the extreme-value dependence function. Capéraà et al. [6], Hall and Tajvidi [15], Segers [18] suggested some alternative estimators for Pickands dependence function and they investigated some of its properties. Genest and Segers [9] proposed a rank-based estimator of [17] and [6]. Bücher et al. [5] considered minimum-distance principle for the estimator and provided an infinite class estimators. Zhang et al. [20] and Gudendorf and Segers [12] studied multivariate case assuming knowledge of marginal distributions and also they discussed the asymptotic properties of the estimator. Berghaus et al. [3] constructed an alternative class of estimator for Pickands dependence function for generalized case. They used the minimum distance approach. Marcon et al. [16] proposed a nonparametric estimator by using Bernstein polynomial approximation. They discussed the properties of the proposed estimation method. Cormiér et al. [7] proposed a visual tool for detecting the presence of extreme-value dependence and they used B-splines method for their estimator. There are also some other Bayesian estimations are available in the literature, [13] has considered a certain functional form for the Pickands function. In this study, a Bayesian approach which is based on specifying a probability model for the data is suggested to estimate Pickands dependence function. Our method is different by two ways. First, we do not consider any functional form for the Pickands function except its constraints, second, the method of estimating the function is cubic B-splines, which is not used in this framework. We choose Bayesian approach in estimation of Pickands dependence function as another approach compared to conventional ones since it is shown that in many statistical problems, the Bayesian methods are more efficient. See, [4].

The remainder of the paper is organized as follows. In Section 2, we define cubic B-spline regression function for the Pickands dependence function and the Bayesian model which consists of prior and posterior distributions for the unknown vector of parameters. We present posterior sampling algorithm using reversible jump MCMC method. In subsection 2.4, we discuss about constraints which are imposed to the Pickands dependance function. In Section 3, a simulation study is carried out in order to compare our new estimator and two well-known CFG and Pickands estimators in terms of accuracy and efficiency and a real-data example is also illustrated in Section 4. Section 5 is devoted to the conclusion.

2. The statistical model

2.1. Cubic B-spline regression model

Consider the extreme value copula given in Equation (1.1). Let $t = \frac{\ln v}{\ln(uv)}$ and then $A(t) = \frac{\ln C(u,v)}{\ln(uv)}$. The transformation from (u, v) to $(t, A(t))$ reduce to a convex curve if and only if, the copula C is the form by Equation (1.1). For more details, see [7].

Let $\{(X_i, Y_i)\}_{i=1}^n$ be a random sample of size n from an extreme value copula C and marginals F and G. Consider $\hat{F}_n$ and $\hat{G}_n$ be the empirical distributions of F and G which are defined respectively as

$$\hat{F}_n(z) = \frac{1}{n}\sum_{i=1}^{n} I(X_i \leq z), \quad \hat{G}_n(z) = \frac{1}{n}\sum_{i=1}^{n} I(Y_i \leq z). \tag{2.1}$$

Now, For all i in $\{1, ..., n\}$, we set $\hat{U}_i = \hat{F}_n(X_i)$ and $\hat{V}_i = \hat{G}_n(Y_i)$. Let $\hat{C}_n$ be empirical copula of $\{(\hat{U}_i, \hat{V}_i)\}_{i=1}^n$. Define $T_i = \frac{\ln(\hat{V}_i)}{\ln(\hat{U}_i\hat{V}_i)}$ and $Q_i = \frac{\ln \hat{C}_n(\hat{U}_i,\hat{V}_i)}{\ln(\hat{U}_i\hat{V}_i)}$. In accordance with the property of the Pickands dependence function, that is, $max(t, 1-t) \leq A(t) \leq 1$, division by n in Equation (2.1) ensures that for $i = 1, ..., n$, $Q_i \geq max(T_i, 1 - T_i)$. See, [7].

Suppose that the Pickands dependence function A is to be fitted through a set of points $(t_1, q_1), (t_2, q_2), ..., (t_n, q_n)$. The regression model can be expressed as

$$q_i = A(t_i) + \sigma^2 \varepsilon_i, \quad i = 1, 2, \ldots, n$$

where $\varepsilon_1, ..., \varepsilon_n$ are assumed to be standard normal random variables. As we do not consider any specific functional form on A, we model $\{Q_i\}_{i=1}^n$ as

$$Q_i = g(t_i; \boldsymbol{\beta}) + \sigma^2 \varepsilon_i, \quad i = 1, 2, \ldots, n, \quad t \in [0, 1],$$

where $g(t_i; \boldsymbol{\beta})$ is an arbitrary smooth function and $\boldsymbol{\beta}$ is the coefficients vector of the model. Suppose that m interior knots are $k_1, k_2, \ldots, k_m$ and g is represented by the cubic B-spline basis $S_{j,3}(t, \boldsymbol{k})$, as follows:

$$g(t; \boldsymbol{\beta}) = \sum_{j=1}^{m+4} \beta_j S_{j,3}(t, \boldsymbol{k}).$$

Matrix form of the above model can be written as

$$\boldsymbol{Q} = \chi_{m,\boldsymbol{k}} \boldsymbol{\beta} + \sigma^2 \boldsymbol{\varepsilon},$$

where

$$\boldsymbol{Q} = \begin{bmatrix} Q_1 \\ Q_2 \\ \vdots \\ Q_n \end{bmatrix}_{(n\times 1)}, \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_{m+4} \end{bmatrix}_{((m+4)\times 1)}, \quad \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}_{(n\times 1)}$$

$$\chi_{m,\boldsymbol{k}} = \begin{bmatrix} S_{1,3}(t_1, \boldsymbol{k}) & S_{2,3}(t_1, \boldsymbol{k}) & \ldots & S_{m+4,3}(t_1, \boldsymbol{k}) \\ S_{1,3}(t_2, \boldsymbol{k}) & S_{2,3}(t_2, \boldsymbol{k}) & \ldots & S_{m+4,3}(t_2, \boldsymbol{k}) \\ \vdots & \vdots & & \vdots \\ S_{1,3}(t_n, \boldsymbol{k}) & S_{2,3}(t_n, \boldsymbol{k}) & \ldots & S_{m+4,3}(t_n, \boldsymbol{k}) \end{bmatrix}_{(n\times(m+4))}.$$

2.2. The Bayesian model

2.2.1. The priors. To determine the prior distributions for the parameters, we should check the number of unknown parameters of the model. The model parameters are: The number of interior knots, m, knots locations, $\boldsymbol{k}$, model coefficients, $\boldsymbol{\beta}$, and the variance of the errors, σ^2.

If one fixes the number and the locations of the knots, the only unknown parameters in the model are $\boldsymbol{\beta}$ and variance σ^2. The Normal-Inverse-Gamma (NIG) with $(m+4)$-dimensional mean vector $\boldsymbol{\mu}$ and $(m+4) \times (m+4)$-demensional covariance matrix $\boldsymbol{V}$ is the popular *conjugate* choice for the prior of $\boldsymbol{\beta}$ and σ^2, which is usually denoted by

$$\pi(\boldsymbol{\beta}, \sigma^2) = NIG(\boldsymbol{\mu}, \boldsymbol{V}, a, b), \quad a, b > 0,$$

that is

$$\pi(\boldsymbol{\beta}, \sigma^2) = \pi(\boldsymbol{\beta}|\sigma^2) \times \pi(\sigma^2) = \mathcal{N}_{(m+4)}(\boldsymbol{\mu}, \sigma^2 \boldsymbol{V}) \times \mathcal{IG}(a, b),$$

where $\mathcal{IG}(a, b)$ denotes the Inverse-Gamma distribution with parameters a and b. Then, we obtain the prior distribution as

$$\begin{aligned} \pi(\boldsymbol{\beta}, \sigma^2) &= \frac{1}{(2\pi)^{\frac{m+4}{2}} |\boldsymbol{V}|^{\frac{1}{2}} (\sigma^2)^{\frac{m+4}{2}}} \exp\left[-\frac{1}{2\sigma^2} (\boldsymbol{\beta} - \boldsymbol{\mu})' \boldsymbol{V}^{-1} (\boldsymbol{\beta} - \boldsymbol{\mu})\right] \\ &\times \frac{b^a}{\Gamma(a)} \left(\frac{1}{\sigma^2}\right)^{a+1} \exp - \frac{b}{\sigma^2} \\ &\propto \left(\frac{1}{\sigma^2}\right)^{a+\frac{m+4}{2}+1} \exp\left[-\frac{1}{\sigma^2}\left(b + \frac{1}{2}(\boldsymbol{\beta} - \boldsymbol{\mu})' \boldsymbol{V}^{-1} (\boldsymbol{\beta} - \boldsymbol{\mu})\right)\right]. \end{aligned}$$

If we let $a = b = 0$ and $\boldsymbol{V} = n(\chi_{m,\boldsymbol{k}}^{'}\chi_{m,\boldsymbol{k}})^{-1}$, the resulting prior is Zellner's G-prior [19]. Following [10], for a given number of knots, m, we consider a discrete uniform distribution on the knots locations as follows:

$$\pi(\boldsymbol{k}|m) = \binom{K}{m}^{-1}, \quad m = 1, 2, \ldots, M$$

where K is the size of candidate set of knot locations and M is the maximum number of knot sequence allowed. By using the above prior one ensures that all models with dimension m have equal weights. Finally, to determine a prior to m, we consider uniform distribution over all possible values, i.e.

$$\pi(m) = \frac{1}{M}, \quad m \in \{1, 2, \ldots, M\}.$$

Then the prior for the parameter vector is

$$\pi(\boldsymbol{\theta}) = \pi(\boldsymbol{\beta}, \sigma^2, \boldsymbol{k}, m) = \pi(\boldsymbol{\beta}|\sigma^2, \boldsymbol{k}, m)\pi(\sigma^2)\pi(\boldsymbol{k}|m)\pi(m).$$

For the choice of prior, we use well known Zellners G-prior. The priors on the other parameters of the model (number of knots, location of knots) serve mainly as computation tools rather than subjective believes. They are used as a key for exploring the parameters' space. See also, [13].

2.2.2. The likelihood. The likelihood function for the model is defined as the joint probability of the observed data with the given parameters. Since $\chi_{m,\boldsymbol{k}}$ is fixed, the likelihood function is obtained as

$$f(\boldsymbol{q}|\boldsymbol{\theta}) = f(\boldsymbol{q}|\boldsymbol{\beta}, \sigma^2, \boldsymbol{k}, m) = \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{n}{2}} \exp\left[-\frac{1}{2\sigma^2}(\boldsymbol{q} - \chi_{m,\boldsymbol{k}}\boldsymbol{\beta})^{'}(\boldsymbol{q} - \chi_{m,\boldsymbol{k}}\boldsymbol{\beta})\right].$$

2.2.3. The posterior. The first step of the inference procedure is to obtain the joint posterior distribution of $(\boldsymbol{\beta}, \sigma^2)$,

$$\pi(\boldsymbol{\beta}, \sigma^2|\boldsymbol{q}) \propto \pi(\boldsymbol{\beta}, \sigma^2)f(\boldsymbol{q}|\boldsymbol{\beta}, \sigma^2),$$

that is

$$\pi(\boldsymbol{\beta}, \sigma^2|\boldsymbol{q}) \propto \left(\frac{1}{\sigma^2}\right)^{a^*+\frac{m+4}{2}+1} \exp\left[-\frac{1}{\sigma^2}\left(b^* + \frac{1}{2}(\boldsymbol{\beta} - \boldsymbol{\xi})^{'}\boldsymbol{\Psi}^{-1}(\boldsymbol{\beta} - \boldsymbol{\xi})\right)\right],$$

where

$$\boldsymbol{\xi} = (\boldsymbol{V}^{-1}\chi_{m,\boldsymbol{k}}^{'}\chi_{m,\boldsymbol{k}})^{-1}(\boldsymbol{V}^{-1}\boldsymbol{\mu} + \chi_{m,\boldsymbol{k}}^{'}\boldsymbol{q})$$
$$\boldsymbol{\Psi} = (\boldsymbol{V}^{-1}\chi_{m,\boldsymbol{k}}^{'}\chi_{m,\boldsymbol{k}})^{-1}$$
$$a^* = a + \frac{n}{2}$$
$$b^* = b + \frac{1}{2}[\boldsymbol{\mu}^{'}\boldsymbol{V}^{-1}\boldsymbol{\mu} + \boldsymbol{q}^{'}\boldsymbol{q} + \boldsymbol{\xi}^{'}\boldsymbol{\Psi}^{-1}\boldsymbol{\xi}].$$

Note that the marginal posterior distribution is Inverse-Gamma distribution with parameters a^* and b^*. The marginal posterior of $\boldsymbol{\beta}$, $\pi(\boldsymbol{\beta}|\boldsymbol{q})$, results from integrating out σ^2 from the above joint posterior distribution as

$$\int \frac{b^{*a^*}}{(2\pi)^{\frac{m+4}{2}}|\boldsymbol{\Psi}|^{\frac{1}{2}}\Gamma(a^*)} \left(\frac{1}{\sigma^2}\right)^{a^*+\frac{m+4}{2}+1} \exp\left[-\frac{1}{\sigma^2}\left(b^* + \frac{1}{2}(\boldsymbol{\beta} - \boldsymbol{\xi})^{'}\boldsymbol{\Psi}^{-1}(\boldsymbol{\beta} - \boldsymbol{\xi})\right)\right] d\sigma^2.$$

Let $\varphi = b^* + \frac{1}{2}(\boldsymbol{\beta} - \boldsymbol{\xi})^{'}\boldsymbol{\Psi}^{-1}(\boldsymbol{\beta} - \boldsymbol{\xi})$, then,

$$\pi(\boldsymbol{\beta}|\boldsymbol{q}) = \frac{b^{*a^*}}{(2\pi)^{\frac{m+4}{2}}|\boldsymbol{\Psi}|^{\frac{1}{2}}\Gamma(a^*)} \int \left(\frac{1}{\sigma^2}\right)^{a^*+\frac{m+4}{2}+1} \exp(-\frac{\varphi}{\sigma^2})d\sigma^2,$$

by using the properties of Gamma function we have

$$\pi(\boldsymbol{\beta}|\boldsymbol{q}) = \frac{b^{*a^*}\Gamma(a^* + \frac{m+4}{2})}{(2\pi)^{\frac{m+4}{2}}|\boldsymbol{\Psi}|^{\frac{1}{2}}\Gamma(a^*)}\varphi^{-(a^* + \frac{m+4}{2})}.$$

By simplifying the equation, we see that $\pi(\boldsymbol{\beta}|\boldsymbol{q})$ is a multivariate t density with n degrees of freedom. i.e.

$$\pi(\boldsymbol{\beta}|\boldsymbol{q}) = \frac{\Gamma(\frac{n+p}{2})}{\Gamma(\frac{n}{2})\pi^{\frac{m+4}{2}}|n\Sigma|^{\frac{1}{2}}}\left[1 + \frac{(\boldsymbol{\beta}-\boldsymbol{\xi})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\beta}-\boldsymbol{\xi})}{n}\right]^{(-\frac{n+(m+4)}{2})},$$

where $n = 2a^*$, $\boldsymbol{\Sigma} = (\frac{b^*}{a^*})\boldsymbol{\Psi}$.

Since we choose the Zellner's G-prior with $\mu = 0$ as a prior distribution of $(\boldsymbol{\beta}, \sigma^2)$, we should make some changes in the obtained posterior distribution and also marginal posterior distributions. Therefore, the posterior of the parameter vector is, $\pi(\boldsymbol{\theta}|\boldsymbol{q}) \propto f(\boldsymbol{q}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})$, that is

$$\begin{aligned}\pi(\boldsymbol{\theta}|\boldsymbol{q}) &\propto (\sigma^2)^{-\frac{n}{2}}\exp\left[-\frac{1}{2\sigma^2}(\boldsymbol{q} - \chi_{m,\boldsymbol{k}}\boldsymbol{\beta})'(\boldsymbol{q} - \chi_{m,\boldsymbol{k}}\boldsymbol{\beta})\right] \\ &\times (2\pi)^{-\frac{m+4}{2}}|\sigma^2 n(\chi_{m,\boldsymbol{k}}{}'\chi_{m,\boldsymbol{k}})^{-1}|^{-\frac{1}{2}}\exp\left[-\frac{1}{2\sigma^2}\boldsymbol{\beta}'(n^{-1}\chi_{m,\boldsymbol{k}}{}'\chi_{m,\boldsymbol{k}})^{-1}\boldsymbol{\beta}\right] \\ &\times \frac{1}{\sigma^2}\times\frac{1}{\binom{K}{m}}\times\frac{1}{M}.\end{aligned}$$

The marginal posterior distribution of $\boldsymbol{\beta}$ is $(m+4)$-variate Student's t-distribution with n degrees of freedom, i.e.

$$\boldsymbol{\beta}|\boldsymbol{q}, m, \boldsymbol{k} \sim t_{m+4}(n; \boldsymbol{\xi}, \Sigma),$$

where

$$\boldsymbol{\xi} = \frac{n}{n+1}(\chi'_{m,\boldsymbol{k}}\chi_{m,\boldsymbol{k}})^{-1}\chi'_{m,\boldsymbol{k}}\boldsymbol{q},$$

$$\Sigma = \frac{\gamma}{n+1}(\chi'_{m,\boldsymbol{k}}\chi_{m,\boldsymbol{k}})^{-1},$$

$$\gamma = \boldsymbol{q}'(I_n - \frac{n}{n+1}\chi_{m,\boldsymbol{k}}(\chi'_{m,\boldsymbol{k}}\chi_{m,\boldsymbol{k}})^{-1}\chi_{m,\boldsymbol{k}})\boldsymbol{q},$$

then,

$$\mathbb{E}[\boldsymbol{\beta}|\boldsymbol{q}, m, \boldsymbol{k}] = \frac{n}{n+1}(\chi'_{m,\boldsymbol{k}}\chi_{m,\boldsymbol{k}})^{-1}\chi'_{m,\boldsymbol{k}}\boldsymbol{Q}$$

and also the marginal posterior distribution of σ^2 given m and $\boldsymbol{k}$ is $IG(\frac{n}{2}, \frac{\gamma}{2})$. Then we obtain

$$\mathbb{E}[\sigma^2|\boldsymbol{q}, m, \boldsymbol{k}] = \frac{\gamma}{n-2}.$$

2.3. Posterior sampling

The posterior distribution, $\pi(\boldsymbol{\theta}|\boldsymbol{q})$, does not have an analytical closed form. Because of the complex form of this posterior, it does not admit an analytical solution, and should be approximated. In order to approximate the posterior distribution, we generate random numbers from posterior distribution by using *Gibbs sampling* method. To complete the sampling process by Gibbs method, we need to generate random numbers from the conditional distribution of $(\boldsymbol{k}, m)$. Note that due to the varying dimension of this marginal posterior, we are not able to use the standard sampling methods. So we use the *Reversible Jump* algorithm [11]. The marginal posterior distribution of the $(\boldsymbol{k}, m)$ is

$$\pi(\boldsymbol{k}, m) \propto f(\boldsymbol{q}|\boldsymbol{k}, m)\pi(\boldsymbol{k}|m)\pi(m) \propto (n+1)^{-(m+4)/2}\gamma^{-n/2}\binom{\boldsymbol{k}}{m}^{-1}.$$

The Bayesian inference is then carried out based on the assumption that the *true* model is unknown but comes from the class of models $\mathcal{M}_1, \mathcal{M}_2, \ldots$ where $\mathcal{M}_m$ denotes the model with exactly m knots. The overall subspace is $\mathbf{X} = \bigcup_{m=1}^{\infty} \mathbf{X}_m$ where $\mathbf{X}_m$ is the m dimensional space corresponding to model $\mathcal{M}_m$.

Due to the varying dimensionality of our problem, we should design move types between the subspaces of $\mathbf{X}_m$. This process allows the sampler to freely explore the combined parameter space if we use the Reversible Jump algorithm. We use three move types, ***birth***, ***death*** and ***relocation***. ***Birth*** and ***death*** types propose moves between different dimensions while ***relocation*** type proposes moves within a dimension. Throughout the procedure, we assume that the current model is of dimension m.

Birth. With probability $p_{b,m}$, propose to add a new knot at a chosen data point randomly from those which do not currently have a knot.

Death. With probability $p_{d,m}$, propose to remove a randomly chosen knot which is present in the model.

Relocation. With probability $p_{r,m}$, propose to alter a randomly chosen knot, say k_j, by swapping for a randomly chosen knot which is not present in the model.

We choose the proposal probabilities $p_{p,m} = p_{d,m} = p_{r,m} = \frac{1}{3}$ for $m = 2, \ldots, M-1$; $p_{d,1} = 0, p_{b,1} = \frac{1}{3}, p_{r,1} = \frac{2}{3}$ and $p_{b,M} = 0, p_{d,M} = \frac{1}{3}, p_{r,M} = \frac{2}{3}$. The acceptance probability for a Reversible Jump algorithm [11] for a proposal move from model $\boldsymbol{k}$ (of dimension m) to $\boldsymbol{k}'$ (of dimension m') is given by

$$\rho_{\mathcal{M}_m \to \mathcal{M}_{m'}} = \min\left\{1, LR_{\mathcal{M}_m \to \mathcal{M}_{m'}} \times PR_{\mathcal{M}_m \to \mathcal{M}_{m'}} |J|\right\},$$

where the ratio of marginal likelihood ($LR_{\mathcal{M}_m \to \mathcal{M}_{m'}}$), prior and proposal ratio ($PR_{\mathcal{M}_m \to \mathcal{M}_{m'}}$) together with a Jacobian term $|J|$, that accounts for the change in scale when moving between models are of potentially different dimensions. In our problem, the marginal likelihood ratio is

$$\begin{aligned} LR_{\mathcal{M}_m \to \mathcal{M}_{m'}} &= \frac{p(\boldsymbol{q}|\boldsymbol{k}')}{p(\boldsymbol{q}|\boldsymbol{k})} = \frac{p(\boldsymbol{q}|\boldsymbol{k}'_{m'}, m')}{p(\boldsymbol{q}|\boldsymbol{k}_m, m)} = \frac{1}{\sqrt{n+1}} \left(\frac{\gamma_{m',\boldsymbol{k}'}}{\gamma_{m,\boldsymbol{k}}}\right)^{\frac{n}{2}}, \\ PR_{\mathcal{M}_m \to \mathcal{M}_{m'}} &= \frac{\pi(m')}{\pi(m)} \frac{\pi(\boldsymbol{k}'_{m'})}{\pi(\boldsymbol{k}_m)} \frac{\pi(\mathcal{M}_m)}{\pi(\mathcal{M}_{m'})} \frac{\pi(\boldsymbol{k}_m|\boldsymbol{k}'_{m'})}{\pi(\boldsymbol{k}'_{m'}|\boldsymbol{k}_m)}. \end{aligned}$$

Note that because of the use of linear transformations (Birth, Death and Relocation), the Jacobian term is $|J| = 1$.

The prior and the proposal ratios for birth, death and relocation moves are $\frac{p_{d,m+1}}{p_{b,m}}$, $\frac{p_{b,m-1}}{p_{d,m}}$ and 1, respectively. Then the last move is just a standard update in a *Metropolis-Hasting* sampler. The rest of the algorithm consists of two usual Gibbs steps: updating both the coefficients and the variance which can be done easily. In nonparametric settings, the most important parameter to estimate is the functional form of the regression function $A(t) = \boldsymbol{\eta}$ namely $\boldsymbol{\eta} = \chi\boldsymbol{\beta}$. The posterior distribution of $\boldsymbol{\eta}$ is a mixture of multivariate t-Student distributions

$$\pi(\boldsymbol{\eta}|\boldsymbol{q}) = \sum_{m,\boldsymbol{k}} \pi(m, \boldsymbol{k}|\boldsymbol{q}) \pi(\boldsymbol{\eta}_{m,\boldsymbol{k}}|m, \boldsymbol{k}, \boldsymbol{q}),$$

where

$$\boldsymbol{\eta}_{m,\boldsymbol{k}}|m, \boldsymbol{k}, \boldsymbol{q} \sim t_{m+4}\left(n, \chi_{m,\boldsymbol{k}}\,\boldsymbol{\xi}_m, \frac{\gamma_{m,\boldsymbol{k}}}{n}\chi_{m,\boldsymbol{k}}\,\boldsymbol{\Sigma}_{m,\boldsymbol{k}}\,\chi'_{m,\boldsymbol{k}}\right),$$

and

$$\boldsymbol{\xi}_m = \frac{n}{n+1}\left(\chi'_{m,\boldsymbol{k}}\,\chi_{m,\boldsymbol{k}}\right)^{-1}\chi'_{m,\boldsymbol{k}}\,\boldsymbol{q}, \qquad \boldsymbol{\Sigma}_{m,\boldsymbol{k}} = \frac{n}{n+1}\left(\chi'_{m,\boldsymbol{k}}\,\chi_{m,\boldsymbol{k}}\right)^{-1}.$$

The overall mean and variance of this mixture distribution can be estimated by using the sample produced from the algorithm.

2.4. Constraints on the Pickands dapendence function

In order to ensure that the new estimator $\hat{A}_B$ which is obtained by Bayesian approach satisfies the properties of the Pickands dependence function, after constructing the estimator $\hat{A}_B$, we use *Convhull* function in R which is based on the *QHull* algorithm [2] to reconstruct a bounded convex estimator. Note that this algorithm has been used with a slight change so that we consider $A(0) = 1$ as the start point and $A(1) = 1$ as the end point of the algorithm. The rate of convergence in this method is preserved. For more details, see [1].

3. Simulation study

In this section, we consider 4 families of extreme value distributions; Asymmetric logistic distribution with $\alpha = 0.35$, $\beta = 0.75$ and $\tau = 0.25$, Galambos distribution with $\tau = 0.5$, Logistic distribution with $\tau = 0.5$ and Extreme value t-student with 4 degrees of freedom and $\tau = 0.5$. In each family, we generate a random sample of size 300 and estimate the Pickands dependence function using our new Bayesian estimator by running 10000 iterations of algorithm following burn-ins of 5000, intrinsic versions of the rank-based Pickands and Caperaa-Fougeres-Genest (CFG) estimators which are denoted by $\hat{A}_B$, $\hat{A}_{PIC}$ and $\hat{A}_{CFG}$, respectively. Figure 1 shows the estimated Bayesian Pickands dependence function (blue dashed line) for Asymmetric logistic distribution, Galambos distribution, Logistic distribution, and Extreme value t-student, respectively. Pink and green lines present Pickands and CFG estimators, respectively and red line is the true Pickands dependence function. Comparing with the other estimators, Bayesian Pickands estimator is closer to the true Pickands dependence function for all types of extreme-value copulas. For all the models, Bayesian estimator fits best even for the asymmetric model.

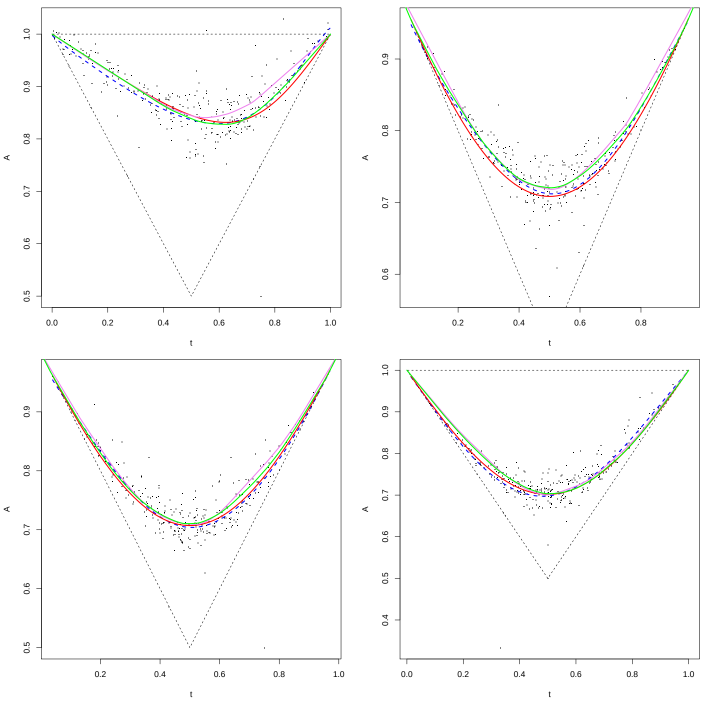

Figure 1. Plots with true Pickands dependence function A (red) and three estimates based on a random sample of size 300 from the Asymmetric Logistic ($\alpha = 0.35$, $\beta = 0.75$, $\tau = 0.25$) (top left), Galambos ($\tau = 0.5$) (top right), Logistic ($\tau = 0.5$) (bottom left), Extreme-value t distributions ($\tau = 0.5$)(bottom right). $\hat{A}_{CFG}$ (green), $\hat{A}_{PIC}$ (pink), $\hat{A}_B$ (blue dashed line).

Figure 2 shows the sequence of m's and the corresponding approximated probability densities. The MAP estimation of m is 2. MAP estimation is the value of the parameter which maximize the posterior distribution, see [10]. The upper plot gives the sequence of interior knots and the lower plot gives the corresponding probability densities for asymmetric logistic distribution, Galambos distribution, Logistic distribution and Extreme value t-student distribution, respectively. Figure 3 shows the mixing behavior of the σ^2 with corresponding running mean (red line) and its histogram. Figure 4 shows the sequence of MSE's against iterations. It is seen when the iteration number getting increased then MSE's are getting smaller as we expected.

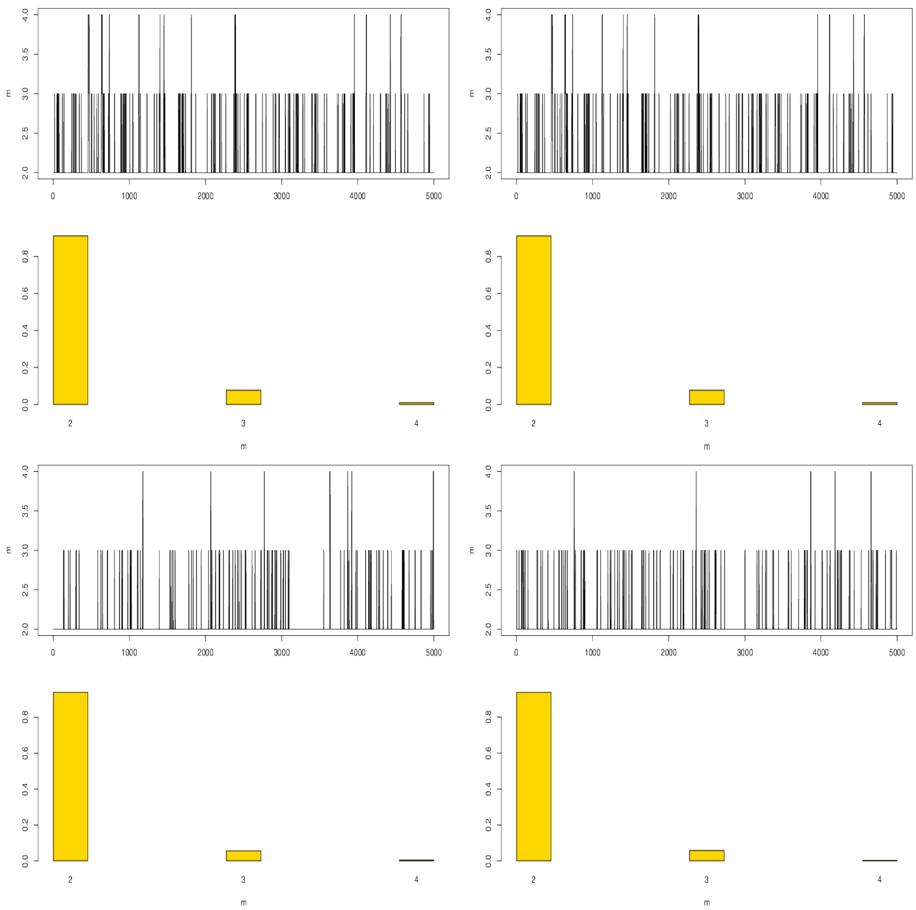

Figure 2. Sequence of m's of $\hat{A}_B$ and the corresponding approximated probability densities based on a random sample of size 300 from Asymmetric Logistic ($\alpha = 0.35$, $\beta = 0.75$, $\tau = 0.25$) (top left), Galambos ($\tau = 0.5$) (top right), Logistic ($\tau = 0.5$) (bottom left), Extreme-value t distributions ($\tau = 0.5$)(bottom right).

Finally, a simulation study is carried out to investigate the behavior of the new estimator. Comparisons between three estimators are obtained in terms of MSE. We draw 500 samples of size 300 from two well known extreme value distribution; Logistic and Negative Logistic distributions. We consider three dependence levels as measured by Kendall's correlation coefficent $\tau = 0.25, 0.50, 0.75$. For each sample, we approximate $\hat{A}_B$ by running 2000 iterations of algorithm following burn-ins of 1000, $\hat{A}_{PIC}$ and $\hat{A}_{CFG}$ and also compute the true Pickands dependence function. The results of these comparisons are summarized in Tables 1-2, and Figures 5-6. One can conclude from these results that $\hat{A}_B$ has the best performance at all levels of dependence.

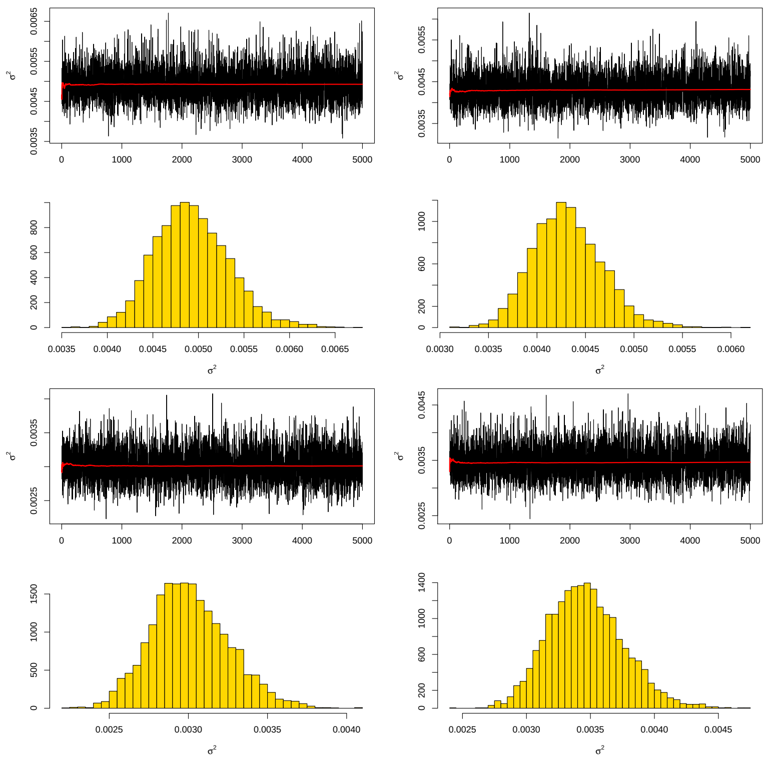

Figure 3. Estimated Variance of $\hat{A}_B$ with running mean (red line) and corresponding histogram based on a random sample of size 300 from Asymmetric Logistic ($\alpha = 0.35$, $\beta = 0.75$, $\tau = 0.25$) (top left), Galambos ($\tau = 0.5$) (top right), Logistic ($\tau = 0.5$) (bottom left), Extreme-value t distributions ($\tau = 0.5$)(bottom right).

Table 1. Mean and Median of MSE of $\hat{A}$ for Bayesian, Pickands and CFG estimators based on 500 samples size of 300 from Negative Logistic or Galambos model when $\tau = 0.25, \tau = 0.5$ and $\tau = 0.75$.

		MSE		
Neg. Logistic		$\hat{A}_B$	$\hat{A}_{PIC}$	$\hat{A}_{CFG}$
$\tau = 0.25$	Median	0.0004016	0.0006648	0.0014030
	Mean	0.0006550	0.0008944	0.0014450
$\tau = 0.50$	Median	0.0003228	0.0006447	0.0007887
	Mean	0.0005192	0.0008564	0.0011850
$\tau = 0.75$	Median	0.0001250	0.0003021	0.0002973
	Mean	0.0002368	0.0003825	0.0005012

Table 2. Mean and Median of MSE of $\hat{A}$ for Bayesian, Pickands and CFG estimators based on 500 samples size of 300 from Logistic or Gumbel model when $\tau = 0.25, \tau = 0.5$ and $\tau = 0.75$.

		MSE		
Logistic		$\hat{A}_B$	$\hat{A}_{PIC}$	$\hat{A}_{CFG}$
$\tau = 0.25$	Median	0.0004766	0.0007120	0.0013680
	Mean	0.0007081	0.0009970	0.0013910
$\tau = 0.50$	Median	0.0003026	0.0005753	0.0008571
	Mean	0.0005170	0.0007923	0.0011840
$\tau = 0.75$	Median	0.0001352	0.0003039	0.0002876
	Mean	0.0002530	0.0004215	0.0004596

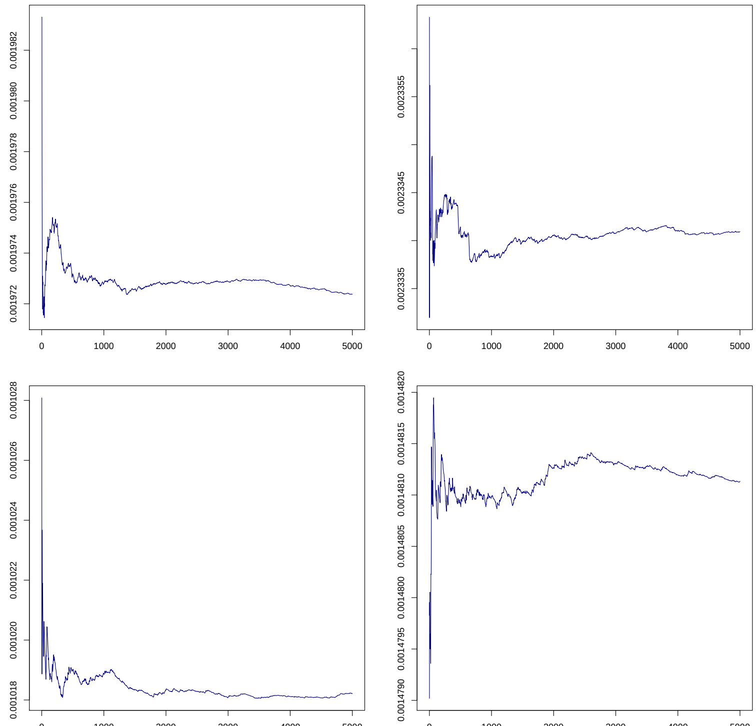

Figure 4. Sequence of MSE's of $\hat{A}_B$ against iterations based on a random sample of size 300 from Asymmetric Logistic ($\alpha = 0.35$, $\beta = 0.75$, $\tau = 0.25$) (top left), Galambos ($\tau = 0.5$) (top right), Logistic ($\tau = 0.5$) (bottom left), Extreme-value t distributions ($\tau = 0.5$)(bottom right).

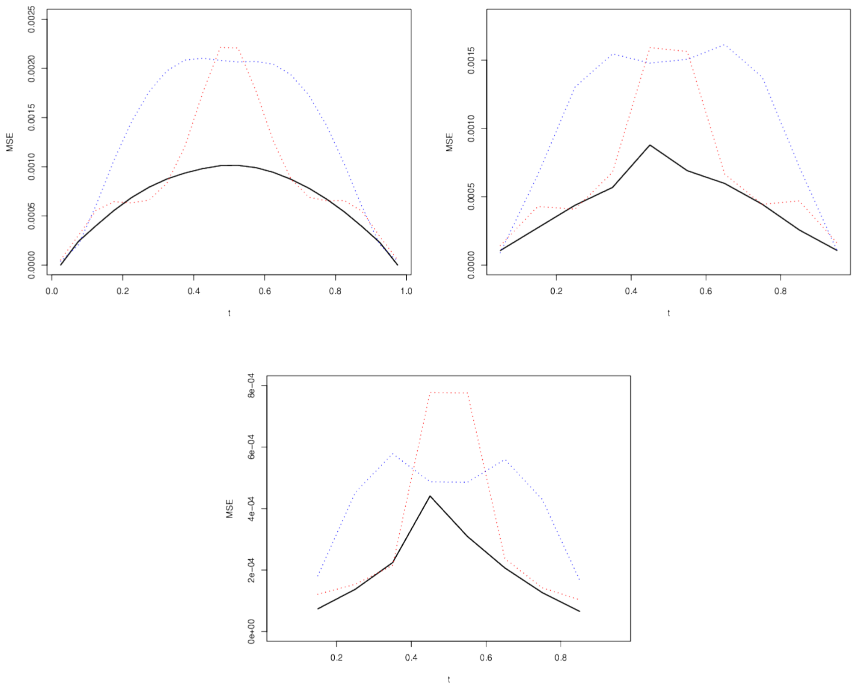

Figure 5. MSE of $\hat{A}$ for Bayesian (black line), Pickands (red dashed line) and CFG (blue dashed line) estimators based on 500 sample size of 300 from Logistic or Gumbel model when $\tau = 0.25$ (first column), $\tau = 0.5$ (second column), and $\tau = 0.75$ (third column).

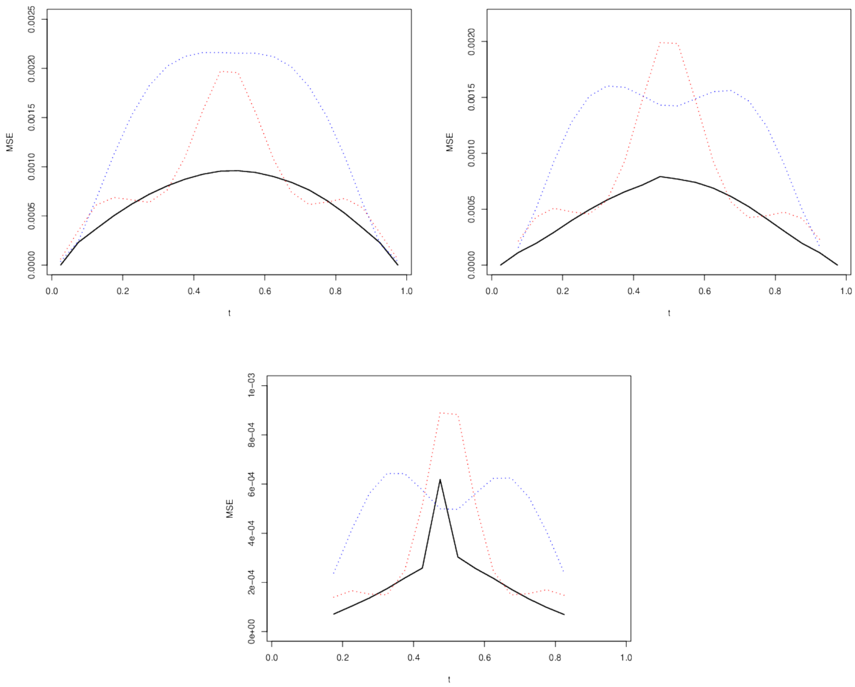

Figure 6. MSE of $\hat{A}$ for Bayesian (black line), Pickands (red dashed line) and CFG (blue dashed line) estimators based on 500 sample size of 300 from Negative Logistic or Galambos model when $\tau = 0.25$ (first column), $\tau = 0.5$ (second column), and $\tau = 0.75$ (third column).

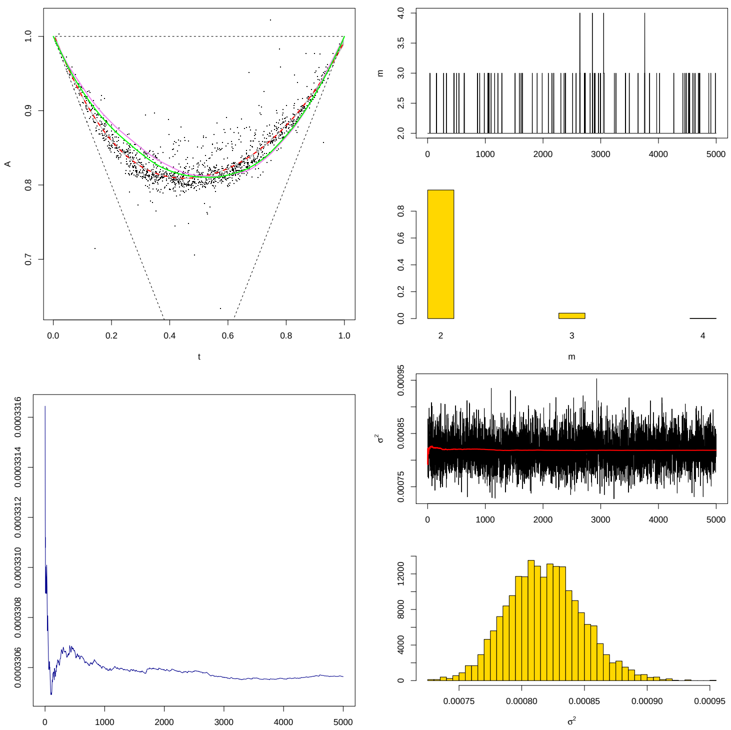

Figure 7. Real data analysis: Plots with three estimates of Pickands dependence function A (red); $\hat{A}_{CFG}$ (green), $\hat{A}_{PIC}$ (pink), $\hat{A}_B$ (red dashed line) (top left), Sequence of m's of $\hat{A}_B$ and the corresponding approximated probability densities (top right), Estimated Variance of $\hat{A}_B$ with running mean (red line) and corresponding histogram (bottom left), Sequence of MSE's of $\hat{A}_B$ against iterations (bottom right).

4. Real Data Analysis

The paper considered a real-data example which is well-known LOSS/ALAE data. The data comprise 1500 general liability claims randomly chosen from late settlement lags and were provided by Insurance Services Office. Each claim consists of an indemnity payment (LOSS) and an allocated loss adjustment expense (ALAE). For more details, see [14]. We estimate Pickands dependence function for LOSS/ALAE data by Bayesian, Pickands and CFG methods. The plots are given in Figure 7. To compare our estimator with the alternative estimators, we use MSE. The results are reported in Table 3. As seen from the values obtained for MSE's, Bayesian estimate of Pickands dependence function has the least value among other estimators.

Table 3. MSE of $\hat{A}$ for LOSS/ALAE data of $\hat{A}_B$, $\hat{A}_{PIC}$ and $\hat{A}_{CFG}$.

MSE		
$\hat{A}_B$	$\hat{A}_{PIC}$	$\hat{A}_{CFG}$
0.0003302	0.0006948	0.0007949

5. Conclusion

The limiting distributions of componentwise maxima of a given sample are related with the bivariate extreme-value distributions. These extreme-value distributions are used for modelling the rare events especially in the fields of financial and enviromental studies. The extreme-value dependence structures of the rare data can be modelled using the extreme-value copulas. The Pickands dependence function characterizes the extreme-value copulas. In this study, we use cubic B-spline regression approach for modeling the Pickands dependence function and then a Bayesian approach which consists of prior and posterior distributions for the unknown vector of parameters is used to estimate the Pickands dependence function. We obtain the prior for the parameter vector and use Zellner's G-prior for the choice of the prior. We also give a posterior sampling algorithm using a reversible jump MCMC method.

We measure the performance of the Bayesian Pickands estimator by a simulation study. In the simulation study, we consider four families of the extreme-value distributions. For each family of the distributions, the Bayesian Pickands estimator shows a better performance when compared with its alternatives.

Also, a real data analysis consisting the well-known LOSS/ALAE data is presented. We estimate the Pickands dependence function by Bayesian, Pickands and CFG methods. It can also be concluded that the Bayesian Pickands estimator has the best performance.

Acknowledgments
We thank the anonymous referees and the asscoiated editor for their helpful suggestions which improved the presentation of the paper.

References

[1] A. Ahmadabadi and B. Hudaverdi Ucer, *Bivariate nonparametric estimation of the Pickands dependence function using Bernstein copula with kernel regression approach*, Comput. Statist. **32** (4), 1515–1532, 2017.

[2] C.B. Barber, D.P. Dobkin and H.T. Huhdanpaa, *The Quickhull algorithm for convex hulls*, ACM Trans. Math. Software **22** (4), 469–483, 1996.

[3] B. Bergahus, A. Bücher and H. Dette, *Minimum distance estimation of Pickands dependence function for multivariate distributions*, Working Paper, 2012.

[4] S.P. Brooks, *Bayesian computation: a statistical revolution*, Philos. Trans. Royal Soc. A **361** (1813), 2681–2697, 2003.

[5] A. Bücher, H. Dette and S. Volgushev, *New estimators of the Pickands dependence function and a test for extreme-value dependence*, Ann. Statist. **39** (4), 1963–2006, 2011.

[6] P. Capéraà, A.-L.Fougeres and C. Genest, *A nonparametric estimation procedure for bivariate extreme value copulas*, Biometrika **84** (4), 567–577, 1997.

[7] E. Cormiér, C. Genest and J.G. Neslehova, *Using B-splines for nonparametric inference on bivariate extreme-value copulas*, Extremes **17** (4), 633–659, 2014.

[8] P. Deheuvels, *On the limiting behavior of the Pickands estimator for bivariate extreme-value distributions*, Statist. Probab. Lett. **12** (5), 429–439, 1991.

[9] C. Genest and J.Segers, *Rank-based inference for bivariate extreme-value copulas*, Ann. Statist. **37** (5B), 2990–3022, 2009.

[10] G. Gholami, *On the Bayesian change-point problem in regression analysis*, J. Stat. Theory Appl. **9** (1), 9–27, 2010.

[11] P. Green, *Reversible jump Markov chain Monte Carlo computation and Bayesian model determination*, Biometrika **82** (4), 711–732, 1995.

[12] G. Gudendorf and J. Segers, *Nonparametric estimation of an extreme-value copula in arbitrary dimensions*, J. Multivariate Anal. **102** (1), 37–47, 2011.

[13] S. Guillotte and F. Perron, *A Bayesian estimator for the dependence function of a bivariate extremevalue distribution*, Canad. J. Statist. **36** (3), 83-396, 2008.

[14] E.W. Frees and E.A. Valdez, *Understanding relationships using copulas*, N. Am. Actuar. J. **2** (1), 1–25, 1998.

[15] P. Hall and N. Tajvidi, *Distribution and dependence-function estimation for bivariate extreme-value distributions*, Bernoulli **6** (5), 835–844, 2000.

[16] G. Marcon, S.A. Padoan, P. Naveau, P. Muliere and J. Segers, *Multivariate nonparametric estimation of the Pickands dependence function using Bernstein polynomials*, J. Statist. Plann. Inference **183**, 1-17, 2017.

[17] J. Pickands, *Multivariate extreme value distribution*, in: Proceedings of the 43rd Session of the International Statistical Institute, Buenos Aires, Brazil, 859–878, 1981.

[18] J. Segers, *Non-parametric inference for bivariate extreme-value copulas*, in: M. Ahsanullah and S. Kirmani (ed.) Extreme Value Distributions, Nova Science Publishers, 181–203, 1985.

[19] A. Zellner, *On assessing prior distributions and Bayesian regression analysis with g prior distributions*, in: P. Goel and A. Zellner (ed.) Bayesian Inference and Decision Techniques: Essays in Honor of Bruno de Finetti. Studies in Bayesian Econometrics. Elsevier, New York, 233-243, 1986.

[20] D. Zhang, M.T Wells and L. Peng, *Nonparametric estimation of the dependence function for a multivariate extreme value distribution*, J. Multivariate Anal. **99**, 577–588, 2008.