Hacettepe Journal of Mathematics and Statistics
Volume 47 (6) (2018), 1690 – 1714

Slice sampler algorithm for generalized Pareto distribution

Mohammad Rostami*†, Mohd Bakri Adam Yahya‡, Mohamed Hisham Yahya§, Noor Akma Ibrahim¶

Abstract

In this paper, we developed the slice sampler algorithm for the generalized Pareto distribution (GPD) model. Two simulation studies have shown the performance of the peaks over given threshold (POT) and GPD density function on various simulated data sets. The results were compared with another commonly used Markov chain Monte Carlo (MCMC) technique called Metropolis-Hastings algorithm. Based on the results, the slice sampler algorithm provides closer posterior mean values and shorter 95% quantile based credible intervals compared to the Metropolis-Hastings algorithm. Moreover, the slice sampler algorithm presents a higher level of stationarity in terms of the scale and shape parameters compared with the Metropolis-Hastings algorithm. Finally, the slice sampler algorithm was employed to estimate the return and risk values of investment in Malaysian gold market.

Keywords: Extreme value theory, Markov chain Monte Carlo, slice sampler, Metropolis-Hastings algorithm, Bayesian analysis, gold price.

Mathematics Subject Classification (2010): 62C12, 62F15, 60G70, 62G32, 97K50, 65C40.

Received : 05.10.2016 *Accepted* : 02.02.2017 *Doi* : 10.15672/HJMS.2017.441

*Institute for Mathematical Research, Universiti Putra Malaysia, Email: rostamy.mohammad@gmail.com

†Corresponding Author.

‡Institute for Mathematical Research, Universiti Putra Malaysia, Email: pmbakri@gmail.com

§Department of Accounting and Finance, Faculty of Economics And Management, Universiti Putra, Malaysia, Email: datohisham@gmail.com

¶Department of Mathematics, Faculty of Science, Universiti Putra, Malaysia, Email: nakma@upm.edu.my

1. Introduction

The field of extreme value theory (EVT) goes back to 1927, when Fréchet [18] formulated the functional equation of stability for maxima, which later was solved with some restrictions by Fisher and Tippett [16] and finally by Gnedenko [22] and De Haan [12]. There are two main approaches in the modelling of extreme values. First, under certain conditions, the asymptotic distribution of a series of maxima (minima) can be properly approximated by Gumbel, Weibull and Frechet distributions which have been unified in a generalized form named generalized extreme value (GEV) distribution [6]. The second approach is related to a model associated with observation over (below) a given threshold. EVT indicates that such approximated model represents a generalized Pareto distribution (GPD) [11, 26]. The GPD has the benefit of using more sample information for tail estimation, as compared to the generalized extreme value GEV which considers the block maxima (BM). It therefore can minimize the problem of being wasteful of extreme information for gathering more extreme data compared to the GEV [6]. The BM approach group the data into epochs (months, years, etc.) or events (storms) and use the maxima (minima) as representative of each epoch or event. This leads to the fact that every epoch or event contains one representative no matter the size of the associated extreme value and all values which are not extremes of epochs or events are discarded. Thus, some information seems to be lost. The larger the epoch the larger the loss of data [4]. A more natural way to define extremes in a given sample is to set a high threshold u and to consider as extreme any observation exceeding u [42]. The threshold approach is the analogue of the GEV distribution for the block maxima, but it leads to a distribution called the GPD which is proven to be more flexible than the block maxima [45, 44, 23]. This approach allows in principle for a more parsimonious use of data and hinges on theoretical foundations as solid as those of the BM method, as shown by the results of [2] and [38]. One advantage of the threshold methods over the block maxima is that they can deal with asymmetries in the tails of distributions [30].

1.1. Generalized Pareto Distribution. Let the random variable X follow a GPD model and indicate the excess above the selected threshold u. The distribution function of X is in form

$$F_X(x|u,\sigma,\xi) = \begin{cases} 1 - \left[1 + \xi\left(\frac{x-u}{\sigma}\right)\right]^{-1/\xi} & \xi \neq 0, \\ 1 - \exp\left(-\frac{x-u}{\sigma}\right) & \xi = 0, \end{cases} \tag{1.1}$$

where the probability density function is given by

$$f_X(x|u,\sigma,\xi) = \begin{cases} \frac{1}{\sigma}\left[1 + \xi\left(\frac{x-u}{\sigma}\right)\right]^{(1+1/\xi)} & \xi \neq 0, \\ \frac{1}{\sigma}\exp\left(-\frac{x-u}{\sigma}\right) & \xi = 0, \end{cases} \tag{1.2}$$

with

$$x \geq u, \qquad \sigma > 0, \qquad 1 + \xi\left(\frac{x-\mu}{\sigma}\right) > 0.$$

where, σ is the scale parameter, ξ is the shape parameter and u is the threshold. There are three type of tail distributions associated with GPD regarding to the shape parameter value. The excesses distribution has an upper bound of the distribution if $\xi < 0$. A exponential decayed type tail correspond to $\xi = 0$, considered in the limit $\xi \to 0$. The excesses above the threshold has a slowly decaying tail and no upper bound if $\xi > 0$. Therefore, the shape parameter of GPD is dominant in determining the qualitative behaviour of the tail.

1.1.1. *Threshold Selection.* The threshold selection of the GPD can be problematic in applications. The threshold selection has to satisfy the balance of validity of the asymptotic and variance of the estimators. The threshold must be sufficiently high to ensure the threshold excesses such that the asymptotically motivated GPD provides a reliable approximation to avoid bias. The threshold cannot be too high otherwise there is potential little sample information leading high variance on estimates. The idea of threshold selection is to pick as low a threshold as possible subject to the limit model providing a reasonable approximation [6]. Traditionally, two methods are available for this: the first method is an exploratory technique carried out prior to model estimation, e.g. using mean residual life (MRL) plots, also referred to as mean excess plots in statistical literature; the second method is an assessment of the stability of parameter estimates based on the fitting of models across a range of different thresholds. Above a level u_0, the asymptotic motivation for GPD is valid and the estimates of shape and scale parameters should be approximately constant [6].

In this study, we use MRL plot as a commonly used tool for threshold selection, see [6, 8, 21, 27, 19, 49].

1.1.2. *Mean Residual Life Plot.* Suppose a sequence of independent and identically distributed measurements $x_1 \cdots, x_n$ and let $x_1, \cdots, x_k$ represent the subset of data points that exceed a particular threshold, u, where $x_1, \cdots, x_k$ consist of the k observations that exceed u. Define threshold excesses by:

$$e_n(u) = \frac{1}{k}\sum_{i=1}^{k}(x_i - u), \qquad u < x_{\max}, \tag{1.3}$$

where $x_{\max}$ is the largest of the X_i. The expected value of excess over threshold, for $(X-u) \sim \mathrm{GPD}(\sigma, \xi)$, is

$$E(X-u) = \frac{\sigma}{1-\xi},$$

provided $\xi < 1$. If the GPD is valid for excesses of a threshold u_0, it should be valid for all $u > u_0$ subject to the change of scale parameter σ_u. Hence, for $u > u_0$

$$\begin{aligned} E(X-u \mid X > u) &= \frac{\sigma_u}{1-\xi} \\ &= \frac{\sigma_{u_0} + \xi_u}{1-\xi}. \end{aligned} \tag{1.4}$$

Furthermore, $E(X-u \mid X > u)$ is simply the mean of the excesses of the threshold u, for which the sample mean of the threshold excesses of u provides an empirical estimate [8]. According to (1.4), these estimates are expected to change linearly with u, at levels of u for which the GPD is appropriate. This leads to the following procedure. The locus of points $\{(u, e_u) : u < x_{\max}\}$ is termed the MRL plot. Above a threshold u_0 at which the GPD provides a valid approximation to the excess distribution, the MRL plot should be approximately linear in u. Confidence intervals can be added to the plot based on the approximate normality of sample means.

1.2. Slice Sampler. The idea of introducing auxiliary variables into the process of conditional simulation was introduced in statistical physics by Trotter and Tukey [50], who proposed a powerful technique called conditional Monte Carlo, which has been later generalized by Hammersley and Morton [24]. In addition, Edwards et al. [14], Besag and Green [3] suggest additional algorithms of using auxiliary variables in MCMC simulation for the purpose of improving the efficiency of the simulation process. Furthermore, Mira and Tierney [31] present a sufficient condition for the uniform ergodicity of the auxiliary variable algorithm and an upper bound for the rate of convergence to stationarity. These

important results guarantee that the Markov chain converges efficiently to the target distribution.

As a special case of the auxiliary variable method, slice sampler algorithm also requires introducing additional variable(s), to accompany the variable of interest, in order to generate realizations from a target probability density function. The history of the slice sampler can be traced back to Swendsen and Wang [46], who introduced the idea of using auxiliary variables for improving the efficiency of the MCMC sampling technique for a statistical physics model. Later, Swendsen and Wang's notion was further enhanced by the introduction of the slice sampler which has been studied in recent years by many researchers. For example, Besag and Green [3] apply a similar algorithm in agricultural field experiments. Higdon [25] introduces an improved auxiliary variable method for MCMC techniques based on the Swendsen and Wang algorithm called partial decoupling with applications in Bayesian image analysis. Other developments and detailed descriptions of several applications of slice sampler can be found in [14, 17, 41, 33]. Damlen and Wakefield [10] demonstrate the use of latent variables for Bayesian non conjugate, nonlinear, and generalized linear mixed models. The purpose of their paper is to provide a possible sampling algorithm other than rejection based methods or other sampling methods similar to the Metropolis-Hastings algorithm. They point out that with the aid of latent variables the process of constructing a Markov chain is more efficient than a Metropolis-Hastings independence chain. Damlen and Wakefield [10] propose samplers using multiple auxiliary variables for Bayesian inference problems. They factor the probability density of $f(x)$ into the product of k parts, i.e., $f(x) \propto f_1(x), f_2(x), \cdots, f_k(x)$, then introduce k auxiliary variables, $y_1, \cdots, y_k$, one for each factor. The joint density function for x and y_j's is proportional to the product of the indicator functions: $f(x, y_1, \cdots, y_k) \propto \prod_i I\{0 < y_i < f_i(x)\}$. The main idea of [10] is to make all the conditional distributions for the auxiliary variables and the components of x easy to sample from by such factorization. They also compare the auxiliary variable method with the independence Metropolis-Hastings algorithm and conclude that the former is more efficient.

Two years later, Damien and Walker [9] provide a "black-box" algorithm for sampling from truncated probability density functions based on the introduction of latent variables. They show that by introducing a single auxiliary variable, the process of sampling truncated density functions, such as truncated normal, beta, and gamma distributions can be reduced to the sampling of a couple of uniform random variables. However, their discussions are mainly focused on the one dimensional case.

Thereafter, Neal [35] introduces a single auxiliary variable slice sampler method which can be used for univariate and multivariate distribution sampling. He summarizes the single variable slice sampler method in three steps. First, uniformly sample from$(0, f(x_0))$, where x_0 is the current state, and define a current horizontal slice "S" as $S = \{x : y < f(x)\}$. Then, around the current state x_0, find an interval $I = (L, R)$ containing much, if not all, of the current slice "S ". Third, uniformly sample from the part of the slice within interval I to get the new state x_1. Neal [35] proves that the resulting Markov chain from this algorithm converges to the target distribution. He points out that the sampling efficiency can be improved by suppressing random walks. This can be done for univariate slice sampling by "overrelaxation," and for multivariate slice sampling by "reflection" from the edges of the slice. To improve the efficiency of the slice sampler, Neal [35] also introduce the idea of "shrinkage", that is, to shrink the hyper rectangle sampling region when rejection happens. By doing this, the number of rejections decreases dramatically. As pointed out by Chen and Schmeiser [5], for single variable slice sampler, the sampling process proposed by Neal [35] operates analogously to the Gibbs sampling method in the sense that given the current state x_0, to obtain the next state

x_1, first an auxiliary variable y is generated from the conditional distribution $[y \mid x_0]$; and then x_1 is sampled from the conditional distribution $[x \mid y]$. The reason for using the auxiliary variable is that directly sampling from $[x \mid y]$ is not possible since the closed form of the support of $[x \mid y]$ is not available. By introducing the auxiliary variable, this problem is solved by sampling from two uniform distributions: $[y \mid x_0]$ and $[x \mid y]$. Agarwal and Gelfand [1] illustrate the application of the auxiliary variable method in a simulation based fitting strategy of Bayesian models in the context of fitting stationary spatial models for geo-referenced or point source data.

The properties of the slice sampler have been discussed by several researchers. Mira and Tierney [32] prove that the slice sampler algorithm performs better than the corresponding independence Metropolis-Hastings algorithm in terms of asymptotic variance in the central limit theorem. Based on the findings, they suggest that given any independence Metropolis-Hastings algorithm a corresponding slice sampler that has a smaller asymptotic variance of the sample path averages for every function obeying the central limit theorem can be constructed. Subsequently, Roberts and Rosenthal [41] prove that the simple slice sampler is stochastically monotone under an appropriate ordering on its state space. Based on this property, they derive useful rigorous quantitative bounds on the convergence of slice samplers for certain classes of probability distributions. Roberts and Rosenthal [41] show that the simple slice sampler is nearly always geometrically ergodic and very few other MCMC algorithms exhibit comparably robust properties. Their paper discusses the theoretical properties of slice samplers, especially the convergence of slice sampler Markov chains and shows that the algorithm has desirable convergence properties. Roberts and Rosenthal [41] prove the geometric ergodicity of all simple slice samplers on probability density functions with asymptotically polynomial tails, which indicates that this algorithm has extremely robust geometric ergodicity properties. They derive analytic bounds on the total variation distance from the stationarity of the algorithm by using the Foster-Lyapunov drift condition methodology. Mira and Tierney [31] show that slice samplers are uniformly ergodic under certain conditions. Furthermore they provide upper bounds for the rates of convergence to stationarity for such samplers. In addition, Walker [51] proposes an axillary variable technique based on slice sampler algorithm to sample from well known mixture of Dirichlet process model. The key to the algorithm detailed in his paper, which also keeps the random distribution functions, is the introduction of a latent variable which allows a finite number, which is known, of objects to be sampled within each iteration of a Gibbs sampler.

Murray et al. [34] present a MCMC algorithm called elliptical slice sampler for performing inference in models with multivariate Gaussian priors. Its key properties are: 1) it has simple, generic code applicable to many models, 2) it has no free parameters, 3) it works well for a variety of Gaussian process based models. These properties make their method ideal for use while model building, removing the need to spend time deriving and tuning updates for more complex algorithms. It performs similarly to the best possible performance of a related Metropolis-Hastings scheme, and could be applied to a wide variety of applications in both low and high dimensions. Furthermore, Merrill and Jingjing [29] suggest two variations of a multivariate normal slice sampler method that uses multiple auxiliary variables to perform multivariate updating. Their method is flexible enough to allow for truncation to a rectangular region and/or exclusion of any n-dimensional hyper-quadrant. Merrill and Jingjing [29] compare efficiency and accuracy of their proposed method and existing state-of-the-art slice sampler. According to the results, by using this method, one can generate approximately independent and identically distributed samples at a rate that is more efficient than other methods that update all dimensions at once. Additionally, Tibbits et al. [48] propose an approach to multivariate slice sampler that naturally lends itself to a parallel implementation. They

study approaches for constructing a multivariate slice sampler, and they show how parallel computing can be useful for making MCMC algorithms computationally efficient. Tibbits et al. [48] examine various implementations of their algorithm in the context of real and simulated data. Moreover, Kalli et al. [28] present a more efficient version of the slice sampler for Dirichlet process mixture models described by Walker [51]. Their proposed sampler allows for the fitting of infinite mixture models with a wide-range of prior specifications. Two applications are considered: density estimation using mixture models and hazard function estimation. Kalli et al. [28] show how the slice efficient sampler can be applied to make inference in the models. In the mixture case, two sub models are studied in detail. The first one assumes that the positive random variables are Gamma distributed and the second one is assumed to be inverse Normal distributed. Both priors have two hyper parameters and they consider their effect on the prior distribution of the number of occupied clusters in a sample. Extensive computational comparisons with alternative conditional simulation techniques for mixture models using the standard Dirichlet process prior and their new priors are made. According to the findings, concerning performance of slice-efficient and retrospective samplers, both samplers are approximately the same in terms of efficiency and performance. However, the savings are in the prerunning work where setting up a slice sampler is far easier than setting up a retrospective sampler. In addition, Favaro and Walker [15] consider the problem of slice sampler mixture models for a large class of mixing measures generalizing the celebrated Dirichlet process. Such a class of measures, known in the literature as σ-stable Poisson-Kingman models, includes as special cases most of the discrete priors currently known in Bayesian nonparametrics, e.g., the two parameter Poisson-Dirichlet process and the normalized generalized Gamma process. Favaro and Walker [15] show how to slice sample a class of mixture models which includes all of the popular choices of mixing measures. Based on the results, with standard stick-breaking models the stick-breaking variables are independent, even as they appear in the full conditional distribution sampled in the posterior MCMC algorithm. They show how to sample this joint distribution and hence implement a valid MCMC algorithm. Moreover, Nieto et al. [36] introduce an automatic method for rail inspection that detects rail flaws using computer vision algorithms. The proposed technique is based on the novel combination of simple but effective laser-camera calibration procedure with the application of an MCMC framework. They evaluate the performance of the proposed strategy applying various sampling techniques including sequential importance resampling, Metropolis-Hastings, slice sampler and overrelaxated slice sampler. The results show that the overrelaxed slice sampler is capable of more efficiently representing a probability density function using the slice method, which allows faster computation without sacrifying accuracy. In addition, Nishihara et al. [37] present a parallelizable MCMC algorithm for efficiently sampling from continuous probability distributions that can take advantage of hundreds of cores. This method shares information between parallel Markov chains to build a scale-location mixture of Gaussians approximation to the density function of the target distribution. Nishihara et al. [37] combine this approximation with elliptical slice sampler algorithm presented by Murray et al. Murray et al. [34] to create a Markov chain with no step-size parameters that can mix rapidly without requiring gradient or curvature computations. Nishihara et al. [37] compare their algorithm to several other parallel MCMC algorithms in a variety of settings. They find that generalized elliptical slice sampler mixes more rapidly than the other algorithms on a variety of distributions, and they find evidence that the performance of generalized elliptical slice sampler can scale super-linearly in the number of available cores. Further, Dubois et al. [13] introduce an approximate slice sampler that uses only small mini-batches of data in every iteration. They show that their proposed algorithm can significantly improve (measured in terms of the risk) the traditional slice

sampler of Neal [35] when faced with large datasets. Dubois et al. [13] evaluate the ability of the proposed approximate slice sampler by using three experiments including regularized linear regression, multinomial regression and logistic regression. The findings indicate that the performance for the multinomial regression model is most significant compared to regularized linear regression and logistic regression.

Recently, Tibbits et al. [47] describe a two-pronged approach for constructing efficient, automated MCMC algorithms: (1) they propose the factor slice sampler, a generalization of the univariate slice sampler where they treat the selection of a coordinate basis (factors) as an additional tuning parameter, and (2) they develop an approach for automatically selecting tuning parameters to construct an efficient factor slice sampler. Tibbits et al. [47] examine the performance of the standard and factor univariate slice samplers within the context of several examples, including two challenging examples from spatial data analysis. They show how the algorithm can be fully automated, which makes it very useful for routine application by modelers who are not experts in MCMC. Furthermore, according to the findings, the automated and parallelized factor slice sampler provides an efficient technique that has broad application to statistical sampling problems.

It can be seen from the above review that the slice sampler algorithm has been extensively studied, yet the literature on extreme value case is sparse. Nowadays, many researchers are dealing with extreme data sets and computational challenges. Hence, there are increasingly more requirements for efficient sampling algorithms of extreme value distributions. In this paper, slice sampler algorithm is developed for the GPD model.

2. Metropolis-Hastings algorithm

Metropolis-Hastings algorithm simulates samples from a probability distribution by making use of the full joint density function and (independent) proposal distributions for each of the variables of interest [52]. The idea behind the Metropolis-Hastings algorithm is to start with an (almost) arbitrary transition density q. This density will not give the correct asymptotic distribution f, but we could try to repair this by rejecting some of the moves it proposes, see [43].

In this study, candidate generator is assumed to have a normal distribution, so-called "random walk Metropolis-algorithm with normal increments". Metropolis-Hastings algorithm used in this study is based on simulation of a random walk chain. Random walk chain is a more natural approach for the practical construction of a Metropolis-Hastings proposal is thus to take into account the value previously simulated to generate the following value; that is, to consider a local exploration of the neighbourhood of the current value of the Markov chain [40]. Note that each step of the two stage Gibbs sampler amounts to an infinity steps of a special slice sampler. Moreover, the Gibbs sampling method is equivalent to the composition of p Metropolis-Hastings algorithm, with acceptance probabilities uniformly equal to 1 [39].

3. Bayesian inference of the GPD using slice sampler algorithm

Although, the central idea of the threshold approach is to avoid the loss of information produced by the BM approaches [4], one of the common issues with extreme value modeling is the lack of observations [7]. There are a number of reasons why a Bayesian analysis of extreme value data might be useful. First and foremost, owing to scarcity of data, the facility to include other sources of information through a prior distribution has obvious appeal. Second, the output of a Bayesian analysis - the posterior distribution - provides a more complete inference than the corresponding maximum likelihood analysis

[6]. Therefore, in the current study, Bayesian inference is used for fitting the GPD posterior model as we can potentially take advantage of any expert prior information, which can be important in tail estimation due to the inherent sparsity of extremal data, and to account for all uncertainties in the estimation. Although, in this study, we deliberately apply vague priors, to show our little prior information about the GPD parameters. The slice sampler algorithm was developed to obtain the posterior distribution of the GPD.

3.1. Likelihood function. Given a value of threshold u, and the original sample $X_1, \cdots, X_n$, the extremes over u compose a new sample $X_1, \cdots, X_k$, $k < n$. The likelihood function for a GPD sample is given by

$$L(\mathbf{x} \mid \sigma, \xi) = \frac{1}{\sigma^k} \prod_{i=1}^{k} \left[1 + \xi \left(\frac{x_i - u}{\sigma}\right)\right]^{-(1+1/\xi)}, \tag{3.1}$$

where

$$1 + \xi \left(\frac{x_i - u}{\sigma}\right) > 0, \qquad \xi \neq 0, \qquad i = 1, 2, \cdots, k.$$

3.2. Prior distribution. We have defined the gamma and normal with large variances denoted by $\mathrm{Ga}(\alpha, \lambda)$ and $\mathrm{N}(\beta, \eta^2)$ as prior distributions on the scale (σ) and shape (ξ) parameters, respectively. The probability densities are:

$$\begin{aligned} \pi(\sigma) &= \frac{1}{\lambda^\alpha \Gamma(\alpha)} \exp(-\frac{\sigma}{\lambda}) \sigma^{\alpha-1} \qquad , \qquad \sigma > 0, \quad \alpha > 0, \quad \lambda > 0 \\ \pi(\xi) &= \frac{1}{\eta\sqrt{2\pi}} \exp\left[-\frac{(\xi - \beta)^2}{2\eta^2}\right] \qquad , \qquad \xi \in \mathbb{R}, \quad \beta \in \mathbb{R}, \quad \eta > 0 \end{aligned} \tag{3.2}$$

3.3. Posterior distribution. By multiplying the likelihood function in Equation (3.1) and the prior distributions in Equation (3.2), the posterior distribution becomes:

$$\begin{aligned} \pi(\sigma, \xi|\mathbf{x}) &\propto L(\mathbf{x} \mid \sigma, \xi) \times \pi(\sigma) \times \pi(\xi) \\ &\propto \frac{1}{\sigma^k} \prod_{i=1}^{k} \left[1 + \xi \left(\frac{x_i - u}{\sigma}\right)\right]^{-(1+1/\xi)} \\ &\quad \times \exp\left[-\frac{(\xi - \beta)^2}{2\eta^2} - \frac{\sigma}{\lambda}\right] \sigma^{\alpha-1}. \end{aligned} \tag{3.3}$$

4. Simulation study

Two simulation studies consider the performance of the slice sampler algorithm in the GPD model and compare the results with the Metropolis-Hastings algorithm. One parameter sets from the GEV model, including $\mathrm{GEV}(10, 5, 0.3)$ and three parameter sets from the GPD density function, containing $\mathrm{GPD}(5, 0.3)$, $\mathrm{GPD}(5, 10^{-9})$ and $\mathrm{GPD}(5, -0.3)$ were chosen to simulate data from POT and GPD density function with different tail behaviours, respectively. The different shape parameter (ξ) values used for these three sets are to identify the model features for heavier tails. In the POT approach, estimation of the shape parameter is problematic due to requiring a large size of data to obtain reliable estimate [6].

Simulation study I represents a population with the POT model obtained from the $\mathrm{GEV}(10, 5, 0.3)$. Simulation study II represents a GPD density function for the different values of the shape parameter. The sample sizes for the first and second simulation studies are $n = 155$ and $n = 150$, respectively. The prior for the simulation sample is as: $\sigma \sim \mathrm{Ga}(2, 1000)$ and $\xi \sim \mathrm{N}(0, 1000)$. The values of M_p, SD_p, SE_p and 95% quantile based credible intervals for the scale σ and shape ξ parameters of the GPD model are shown in Table 1 to Table 4.

We wrote all the computer code with advanced statistics package R.

4.1. Simulation study I. A sample was simulated from the GEV$(10, 5, 0.3)$ with a size 200 with 155 observations over the threshold u. The threshold value was obtained through MRL plot as shown in Figure 1. The Figure displays curvature until around $u = 8$, after which there is reasonable linearity. Moreover, Figure 2 reveals the n=155 excesses over the threshold $u = 8$. Figure 3 supplies the diagnostic plots of the exceedence model fitting on the simulated data. As can be seen from Figure 3, the quantile plot and density plot show a good fit of GPD to the simulated data. By using the slice sampler and Metropolis-Hastings algorithms with 5000 iterations, M_p, SD_p, SE_p and 95% quantile based credible intervals of the scale (σ) and shape (ξ) parameters of the GPD posterior density were calculated as shown in Table 1. Additionally, Figures 4 and 5 present the slice sampler series, the Metropolis-Hastings series and the posterior densities of the algorithms. The first 1000 iterations are discarded as burn-in.

4.1.1. *Results from the slice sampler and Metropolis-Hastings algorithms for the observations above the given threshold.* Table 1 illustrates the statistical results along with 95% quantile based credible intervals of the posterior distribution of the POT model by specifying the gamma and normal with large variances as prior distributions on the scale σ and shape ξ parameters, respectively. As can be seen from the Table, for the slice sampler algorithm, the posterior means of σ and ξ are 23.054 and -1.272, respectively. On the other hand, the posterior means of σ and ξ for the Metropolis-Hastings algorithm are 19.928 and -0.316, respectively. In this simulation, since the true parameter values of the GPD model were unknown, the posterior mean values of the Metropolis-Hastings and slice sampler algorithms are relatively far from each other.

Moreover, Table 1 reveals the SE_p for both Metropolis-Hastings and slice sampler algorithms are less than 0.025 which shows a low level of error. Further, based on Table 1, the lower and upper bounds of the 95% quantile based credible intervals of σ and ξ for the slice sampler algorithm are equal to $(22.223, 24.432)$ and $(-1.291, -1.219)$, respectively. In addition, the lower and upper bounds of σ and ξ for the Metropolis-Hastings algorithm are $(16.870, 23.114)$ and $(-0.384, -0.233)$, respectively. Overall, both the slice sampler and Metropolis-Hastings algorithm show a low level of error ($SE_p < 0.03$), however, the slice sampler presents closer posterior means and shorter credible intervals for σ and ξ compared with Metropolis-Hastings algorithm.

In addition, in this research, the maximum likelihood estimations of the GPD parameters are calculated. According to the results, the maximum likelihood estimations of σ and ξ are equal to 20.573 and -0.137, respectively, which are close to the posterior mean values of the slice sampler technique. These results show that there is a high level of consistency between maximum likelihood estimations and and posterior means of slice sampler algorithm.

Figures 4a and 4b show convergency in both the Metropolis-Hastings and slice sampler series. In this work, in order to determine the algorithm which shows a higher level of stationarity, we compare the running time to converge on a satisfactory model for Metropolis-Hastings and slice sampler algorithms. According to the findings, the Metropolis-Hastings algorithm requires around 9 seconds to perform 5000 iterations, while the running time to converge on a satisfactory model for slice sampler technique is around 3 seconds. Therefore, the slice sampler algorithm reduces the time required to perform a number of iterations. Consequently, the slice sampler shows a higher level of stationarity compared to the Metropolis-Hastings algorithm.

Additionally, density plots of the posteriors in the Metropolis-Hastings and slice sampler algorithms are shown in Figures 5a and 5b, respectively.

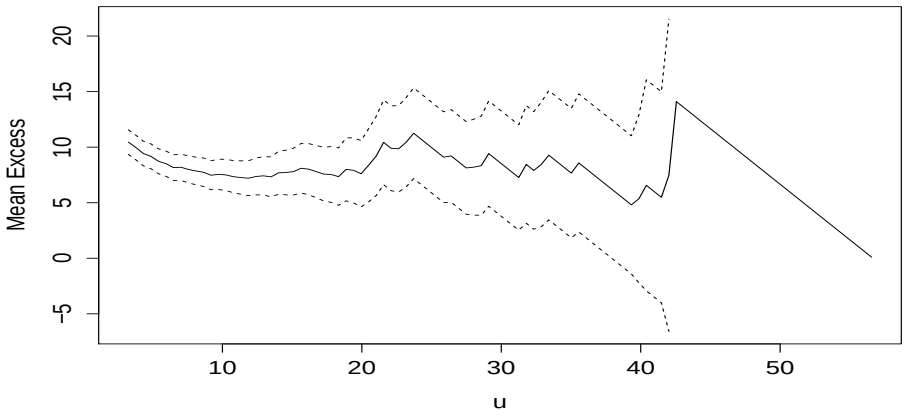


Figure 1. MRL plot (with the 95% confidence intervals) of simulated data. The Figure shows curvature until around $u = 8$, after which there is reasonable linearity.

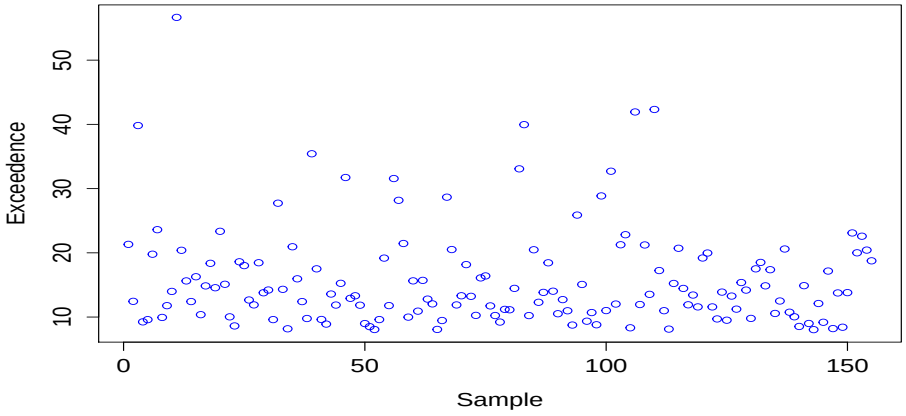
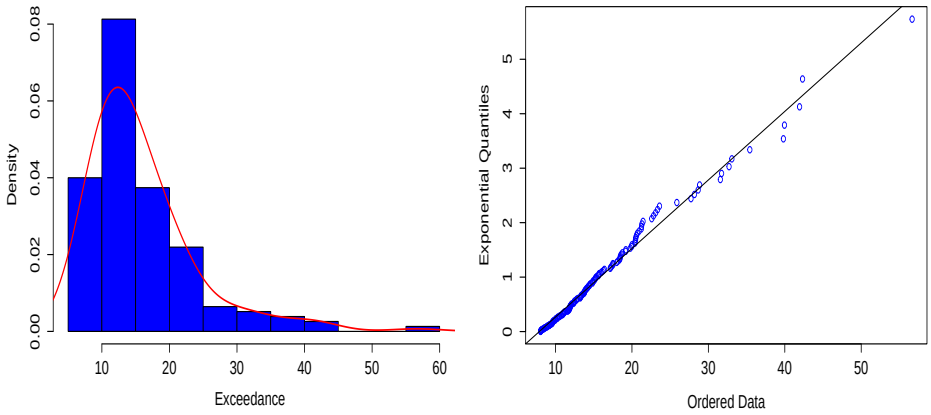


Figure 2. Exceedence over a high threshold of the simulated data. Figure shows the n=155 excesses over the threshold $u = 8$.

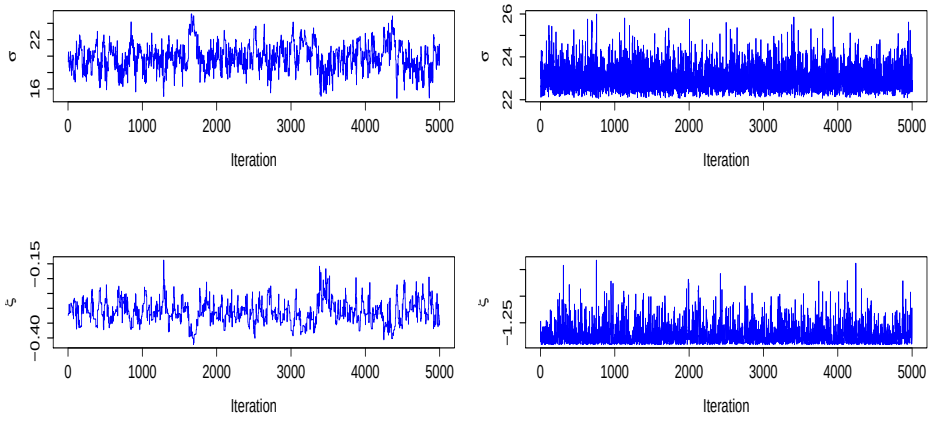


(a) Density plot **(b)** Quantile plot

Figure 3. Diagnostic plot of GPD fitting for the exceedence over a high threshold of the simulated data. The figures show that the peaks over the threshold value $u = 8$ is well approximated by the GPD.

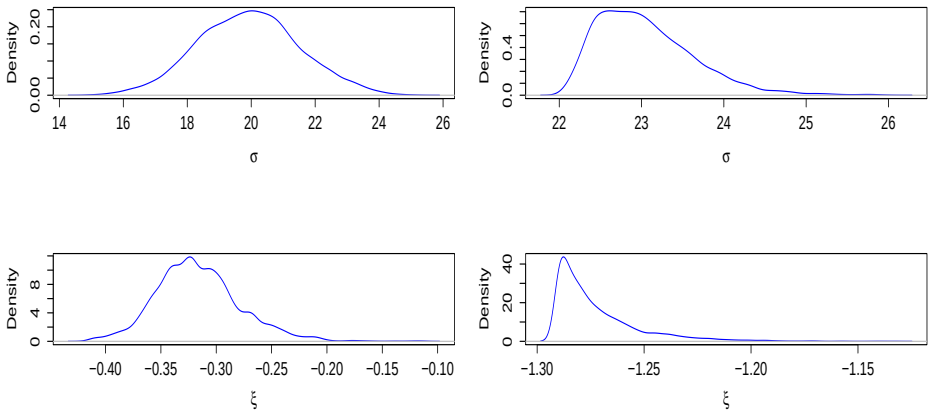
Table 1. Summary of properties of the slice sampler and Metropolis-Hastings algorithms of the POT model parameters for single simulated dataset from the GEV$(10, 5, 0.3)$ with 155 observations over the threshold u. Values of the M_p, SD_p and SE_p along with the 95% quantile based credible interval across the 5000 iterations.

	MH		SS	
	σ	ξ	σ	ξ
M_p	19.928	-0.316	23.054	-1.272
SD_p	1.591	0.038	0.593	< 0.020
SE_p	0.022	< 0.001	0.008	0.001
LCI	16.870	-0.384	22.223	-1.291
UCI	23.114	-0.233	24.432	-1.219



(a) Metropolis-Hastings **(b)** Slice sampler

Figure 4. Series of the scale σ and shape ξ parameters of the POT model with $n = 155$ sample size and 5000 iterations.



(a) Metropolis-Hastings **(b)** Slice sampler

Figure 5. Posterior distributions of σ and ξ with sample size $n = 155$ and 5000 iterations by defining the gamma and normal with large variances as prior distributions on the scale and shape parameters, respectively.

4.2. Simulation study II. We consider the performance of the slice sampler algorithm for the GPD density function and compare it with the Metropolis-Hastings algorithm. Three different parameter sets of GPD(σ, ξ) containing GPD$(5, 0.3)$, GPD$(5, 10^{-9})$ and

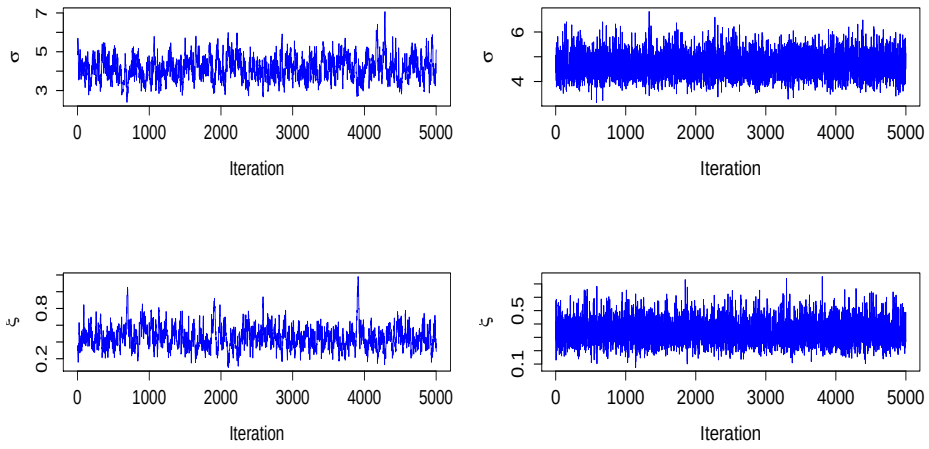
GPD$(5, -0.3)$ were examined, with 5000 iterations and a sample size of $n = 15$. There are three type of tail distributions associated with GPD regarding to the shape parameter value. The excesses distribution has an upper bound of the distribution if $\xi < 0$. A exponential decayed type tail correspond to $\xi = 0$, considered in the limit $\xi \to 0$. The excesses above the threshold has a slowly decaying tail and no upper bound if $\xi > 0$. Hence, in this study the different shape parameter values applied for this simulation study are to identify the model characteristic for heavier tails. The statistical results are presented in Table 2 to Table 4. In addition, Figure 6 to Figure 11 display the slice sampler series, the Metropolis-Hastings series and the posterior densities of the three sample groups. The first 1000 iterations are discarded as burn-in.

4.2.1. *Values of the M_p, SD_p, SE_p and 95% based credible intervals of the scale σ and shape ξ parameters of GPD$(5, 0.3)$ for the slice sampler and Metropolis-Hastings algorithms.* Table 2 demonstrates the M_p, SD_p, SE_p and 95% quantile based credible intervals of the posterior distribution of GPD$(5, 0.3)$ by specifying the gamma and normal with large variances as prior distributions on the scale and shape parameters. The results, as shown in Table 2, indicate that the posterior means of σ and ξ for the slice sampler algorithm are close to the true parameter values, equal to 4.680 and 0.345,respectively, while the posterior means of σ and ξ for the Metropolis-Hastings algorithm are relatively far from the true parameter values, equal to 4.150 and 0.453, respectively. Further, Table 2 reveals the SE_p for both Metropolis-Hastings and slice sampler algorithms are less than 0.01 which shows a low level of error. Moreover, As the Table shows, the lower and upper bounds of the 95% quantile based credible intervals of σ and ξ for the slice sampler algorithm are $(3.790, 5.770)$ and $(0.184, 0.555)$, respectively. On the other hand, the lower and upper bounds of σ and ξ for the Metropolis-Hastings algorithm are $(3.030, 5.412)$ and $(0.221, 0.740)$, respectively. These results indicate that the slice sampler algorithm provides closer posterior mean values and shorter credible intervals compared to the Metropolis-Hastings algorithm.

Moreover, in this paper, the maximum likelihood estimations of the GPD parameters are calculated. According to the results, the maximum likelihood estimations of σ and ξ are equal to 4.603 and 0.290, respectively, which are very close to the posterior mean values of the slice sampler technique. These results indicate that there is a high level of consistency between maximum likelihood estimations and posterior means of slice sampler algorithm.

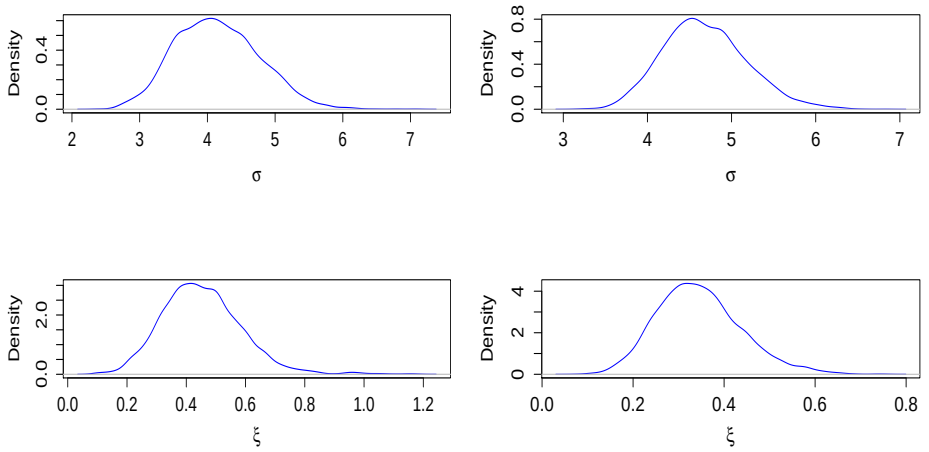
Figures 6a and 6b show iteration series of σ and ξ in both the Metropolis-Hastings and slice sampler algorithms.

Based on the results, the Metropolis-Hastings algorithm requires around 11 seconds to perform 5000 iterations, while the running time to converge on a satisfactory model for slice sampler technique is around 3 seconds. Therefore, the slice sampler algorithm reduces the time required to perform a number of iterations. Consequently, the slice sampler shows a higher level of stationarity compared to the Metropolis-Hastings algorithm.



(a) Metropolis-Hastings **(b)** Slics sampler

Figure 6. Series of the scale σ and shape ξ parameters of the GPD posterior model with sample size $n = 150$ and 5000 iterations. Note: the true parameter values for σ and ξ are 5 and 0.3, respectively



(a) Metropolis-Hastings **(b)** Slice sampler

Figure 7. Posterior distributions of σ and ξ with 5000 iterations by specifying gamma and normal with large variances as prior distributions on the scale and shape parameters of the GPD density function. Note: the true parameter values for σ and ξ are 5 and 0.3, respectively

Table 2. Summary of statistical results of the scale σ and shape ξ parameters of the GPD posterior model for both the slice sampler and Metropolis-Hastings algorithms with $n = 150$ sample size and 5000 iterations. Note: the true parameter values of σ and ξ are 5 and 0.3, respectively.

	MH		SS	
	σ	ξ	σ	ξ
True Value	5	0.3	5	0.3
M_p	4.150	0.453	4.680	0.345
SD_p	0.621	0.135	0.499	0.093
SE_p	0.009	0.002	0.007	0.001
LCI	3.030	0.221	3.790	0.184
UCI	5.412	0.740	5.770	0.555

4.2.2. *Values of the M_p, SD_p, SE_p and 95% quantile based credible intervals of the scale σ and shape ξ parameters of GPD$(5, 10^{-9})$ for the slice sampler and Metropolis-Hastings algorithms.* Table 3 presents the M_p, SD_p, SE_p and 95% quantile based credible intervals of the posterior distribution of GPD$(5, 10^{-9})$ by defining the gamma and normal with large variances as prior distributions on the scale and shape parameters. From Table 3, it can be seen that the posterior means of σ and ξ for the slice sampler algorithm are close to the true parameter values, equal to 4.840 and 0.032, respectively, while the posterior means of σ and ξ for the Metropolis-Hastings algorithm are relatively far from the true parameter values, equal to 4.293 and 0.127, respectively. In addition, Table 3 reveals the SE_p for both Metropolis-Hastings and slice sampler algorithms are less than 0.01 which shows a low level of error. In addition, From Table 3, we can see that the lower and upper bounds of the 95% quantile based credible intervals of σ and ξ for the slice sampler algorithm are $(4.120, 5.700)$ and $(-0.077, 0.180)$, respectively. By contrast, the lower and upper bounds of σ and ξ for the Metropolis-Hastings algorithm are $(3.259, 5.484)$ and $(-0.055, 0.375)$, respectively. Taken together, these results suggest that the slice sampler algorithm provides closer posterior means and shorter credible intervals compared with the Metropolis-Hastings algorithm.

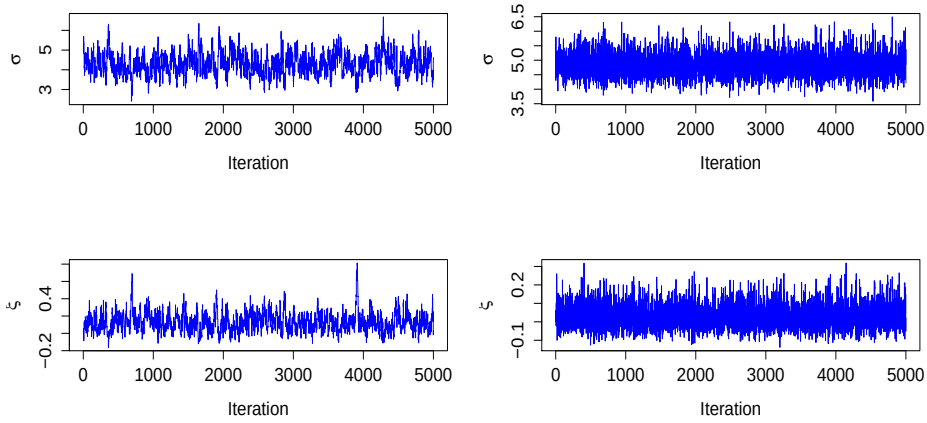
Moreover,the maximum likelihood estimations of the GPD parameters are calculated. According to the findings, the maximum likelihood estimations of σ and ξ are equal to 4.852 and 0.022, respectively, which are very close to the posterior mean values of the slice sampler technique. These results indicate that there is a high level of consistency between maximum likelihood estimations and posterior means of slice sampler algorithm.

Figures 8a and 8b provide iteration series of σ and ξ in both the Metropolis-Hastings and slice sampler algorithms. According to the findings, the Metropolis-Hastings algorithm requires approximately 9 seconds to perform 5000 iterations, while the running time to converge on a satisfactory model for slice sampler technique is about 4 seconds. Therefore, the slice sampler algorithm reduces the time required to perform a number of iterations. Consequently, the slice sampler shows a higher level of stationarity compared to the Metropolis-Hastings algorithm.

Additionally, the posterior densities in the Metropolis-Hastings and slice sampler algorithms are revealed in Figures 9a and 9b respectively.

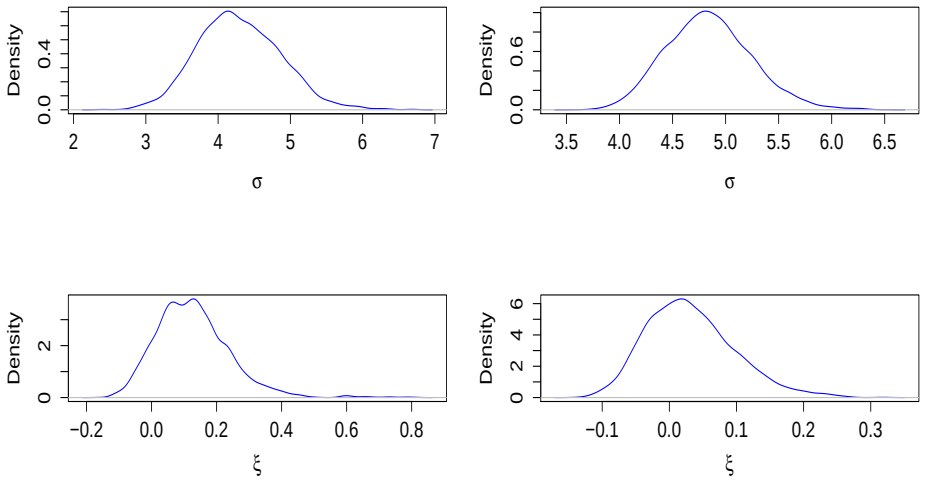
Table 3. Results from the slice sampler and Metropolis-Hastings algorithms for single simulated dataset with sample size $n = 150$ from the GPD posterior model. Note: the true parameter values for σ and ξ are 5 and 10^{-9}, respectively

	MH		SS	
	σ	ξ	σ	ξ
True Value	5	10^{-9}	5	10^{-9}
M_p	4.293	0.127	4.840	0.032
SD_p	0.574	0.114	0.405	0.066
SE_p	0.008	0.002	0.006	0.001
LCI	3.259	-0.055	4.120	-0.077
UCI	5.484	0.375	5.700	0.180



(a) Metropolis-Hastings **(b)** Slice sampler

Figure 8. Series of σ and ξ as scale and shape parameters of the GPD posterior model with sample size $n = 150$ and 5000 iterations. Note: the true parameter values for σ and ξ are 5 and 10^{-9}, respectively



(a) Metropolis-Hastings **(b)** Slice sampler

Figure 9. Posterior distributions of σ and ξ with 5000 iterations by defining the gamma and normal with large variances as prior distributions on the scale and shape parameters of the GPD density function. Note: the true parameter values for σ and ξ are 5 and 10^{-9}, respectively

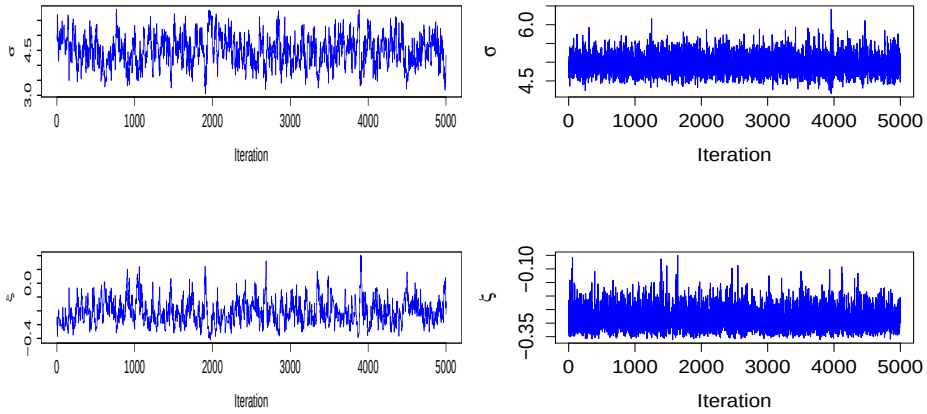
4.2.3. *Values of the M_p, SD_p, SE_p and 95 % quantile based credible intervals of the scale σ and shape ξ parameters of GPD$(5, -0.3)$ for the slice sampler and Metropolis-Hastings algorithms.* Table 4 demonstrates the posterior mean, SD_p, SE_p and 95% quantile based credible intervals of the posterior distribution of GPD$(5, -0.3)$ by specifying the gamma and normal with large variances as prior distributions on the scale and shape parameters. As shown in Table 4, the posterior means of σ and ξ for the slice sampler algorithm are close to the true parameter values, equal to 4.960 and -0.266, whereas the posterior means of σ and ξ for the Metropolis-Hastings algorithm are relatively far from the true parameter values, equal to 4.488 and -0.209, respectively. Additionally, Table 4 reveals the SE_p for both Metropolis-Hastings and slice sampler algorithms are less than 0.01 which shows a low level of error.

Table 4 shows that the lower and upper bounds of the 95% quantile based credible intervals of σ and ξ for the slice sampler algorithm are $(4.480, 5.530)$ and $(-0.346, -0.186)$, respectively. In contrast, the lower and upper bounds of σ and ξ for the Metropolis-Hastings algorithm are $(3.533, 5.458)$ and $(-0.350, 0.002)$, respectively. In summary, these results show that the slice sampler algorithm presents relatively closer posterior means and shorter credible intervals compared to the Metropolis-Hastings algorithm. Further, in the current study, the maximum likelihood estimations of the GPD parameters are calculated. Based on the findings, the maximum likelihood estimations of σ and ξ are equal to 4.632 and -0.309, respectively, which are very close to the posterior mean values of the slice sampler technique. These results show that there is a high level of consistency between maximum likelihood estimations and posterior means of slice sampler algorithm.

Figures 10a and 10b illustrate iteration series of σ and ξ in both the Metropolis-Hastings and slice sampler algorithms.

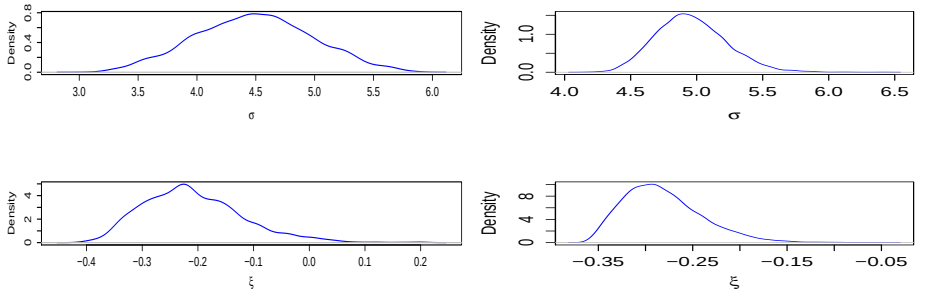
According to the findings, the Metropolis-Hastings algorithm requires around 13 seconds to perform 5000 iterations, while the running time to converge on a satisfactory model for slice sampler technique is around 5 seconds. Therefore, the slice sampler algorithm reduces the time required to perform a number of iterations. Consequently, the slice sampler shows a higher level of stationarity compared to the Metropolis-Hastings algorithm.

In addition, the posteriors' densities in the Metropolis-Hastings and slice sampler algorithms are shown in Figures 11a and 11b, respectively.



(a) Metropolis-Hastings **(b)** Slice sampler

Figure 10. Series of the scale σ and shape parameters of the GPD posterior model with sample size $n = 150$ and 5000 iterations. Note: the true parameter values of σ and ξ are 5 and -0.3, respectively



(a) Metropolis-Hastings **(b)** Slice sampler

Figure 11. Posterior densities of σ and ξ with sample size $n = 150$ and 5000 iterations by defining gamma and normal with large variances as prior distributions on the scale and shape parameters, respectively. Note: the true parameter values of σ and ξ are 5 and -0.3, respectively

Table 4. Summary of properties of the slice sampler and Metropolis-Hastings algorithms of the GPD posterior model parameters for single simulated dataset with sample size $n = 150$. Note: the true parameter values of σ and ξ are 5 and -0.3, respectively.

	MH		SS	
	σ	ξ	σ	ξ
True Value	5	-0.3	5	-0.3
M_p	4.488	-0.209	4.960	-0.282
SD_p	0.049	0.090	0.266	0.042
SE_p	0.004	0.001	0.004	0.001
LCI	3.533	-0.350	4.480	-0.346
UCI	5.458	0.002	5.530	-0.186

5. Real data analysis

In this study, the slice sampler algorithm is applied to the threshold exceedences of Malaysian daily gold returns. The data that are used in this research is generated specifically from daily reports of Malaysian gold price, which have been downloaded from www.kitco.com. The period is from January 5, 2004 to December 18, 2015, leaving a total of 3121 days. The daily returns are defined by $r_t = \log(p_t/p_{t-1})$ where p_t denotes the price of the gold at day t, see [20]. The threshold value was obtained using the MRL plot in Figure 12. The Figure shows curvature until approximately $u = -0.07$, after which there is reasonable linearity. With this threshold, the 12 years of available daily gold return data, generate 1900 threshold exceedances where the peaks over the threshold is well approximated by the GPD as shown in Figure 14 which provides the diagnostic plots of the exceedence model fitting on the Malaysian gold returns. In Figure 14, the quantile plot and density plot display a good fit of GPD to the returns. The description of the excesses above $u = -0.07$ is summarized in Table 5. Based on the results, the mean value of the returns above the threshold is (M= 0.310, SD= 0.355). Further, the median value is 0.219 indicating 50% of the threshold exceedance returns are less than 0.219. Additionally, as the Table shows, the minimum and maximum returns are -0.070 and 4.654, respectively. Moreover, the skewness and kurtosis are 2.774 and 18.498, respectively providing evidence of fat-tailless in the return series. To consider the fat-tailless of the exceedance returns, a sample histogram together with the best fitted density (red curve) and best fitted normal distribution (green curve) was used as shown in Figure 14a. Clearly, the Figure exhibits a heavy upper tail. When ξ is positive, the GPD distribution is a fat-tailed one since the support of the density in this case is $[0, \infty]$ [20]. Moreover, the quantile plot shown in Figure 14b confirms that the peaks over the threshold $u = -0.07$ approximately lie on the line $y = x$ and hence the GPD well fits the set of the exceedance returns.

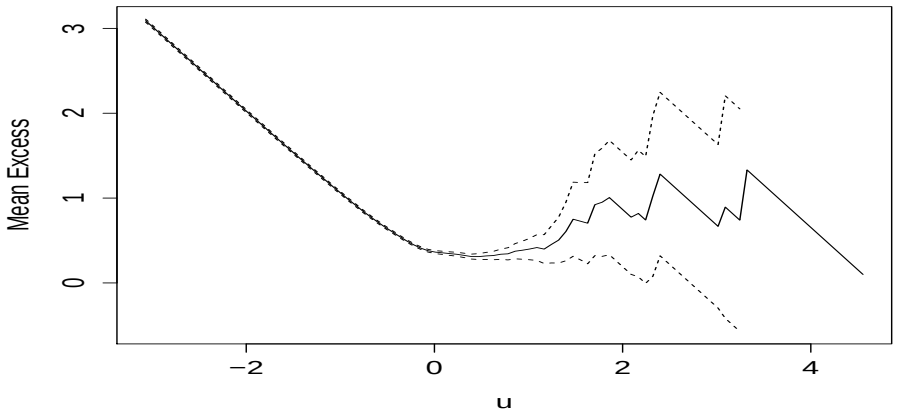


Figure 12. MRL plot (with the 95% confidence intervals) of Malaysian gold returns. The Figure reveals curvature until approximately $u = -0.07$, after which there is reasonable linearity.

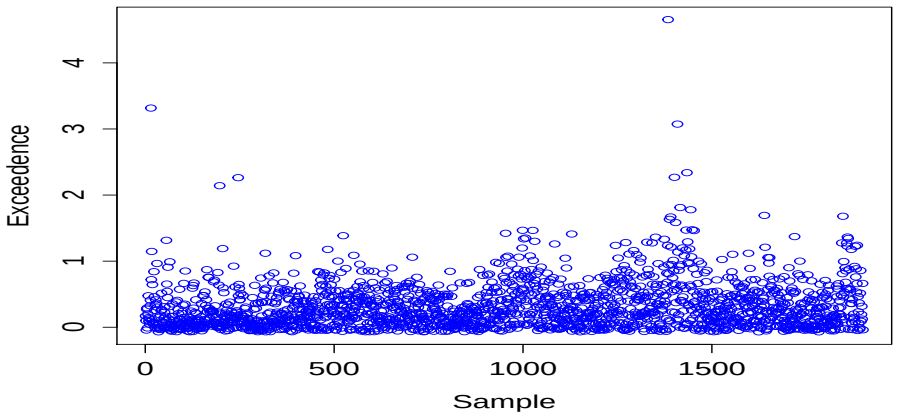
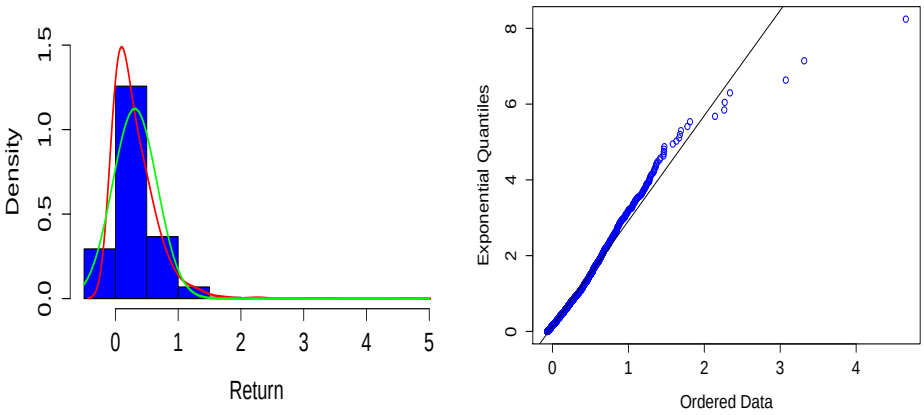


Figure 13. Exceedences over a high threshold of Malaysian gold return



(a) Density plot **(b)** Quantile plot

Figure 14. Diagnostic plot of GPD fitting for the exceedences over a high threshold of Malaysian gold return. The plot show that the peaks over the threshold value $u = -0.07$ is well approximated by the GPD.

Table 5. Descriptive statistics of threshold exceedances of Malaysian gold returns from 2004 to 2015

Gold	n	Mean	SD	Median	Min	Max	Range	Skew	Kurtosis	SE
Value	1900	0.310	0.355	0.219	-0.070	4.654	4.724	2.774	18.498	0.008

5.1. The slice sampler algorithm for the threshold exceedances of Malaysian gold return. We applied the slice sampler algorithm for the threshold exceedances of Malaysian gold return by defining the gamma and normal with large variances on the scale (σ) and shape (ξ) parameters of the GPD model. The typical problem related to the extreme value model is the lack of the tail information and the GPD has the advantage in the availability of the extreme observations compare to the GEV. It is therefore more preferred for the applications which is hard to collect large size of sample and used more generally relatively, see [6].

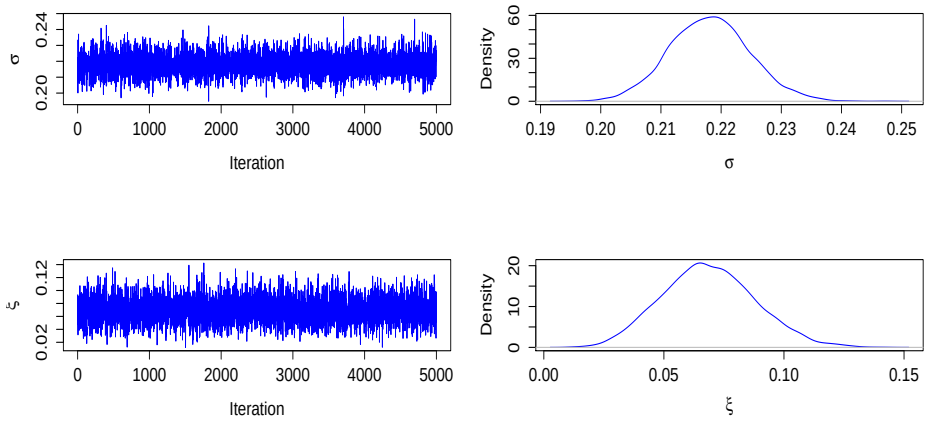
Table 6 illustrates the posterior mean, SD_p and SE_p values along with the 95% quantile based credible interval across 5000 iteration series of the slice sampler algorithm. As can be observed from Table 6, the posterior mean and SD_p values of σ and ξ are ($\text{M}_p = 0.218$, $\text{SD}_p = 0.007$) and ($\text{M}_p = 0.069$, $\text{SD}_p = 0.019$), respectively. Additionally, Table 6 indicates that the posterior standard error (SE_p) for the scale (σ) and shape ξ parameters of the threshold exceedances of Malaysian gold return are less than 0.001 which shows a low level of error. When $\xi > 0$, exceedances over the threshold takes the form of the ordinary Pareto distribution. This particular case is the most relevant for financial time series analysis since it is a heavy tailed one [20]. Hosking [26] introduces a formula for calculation of the mean and standard deviation of the peaks over threshold (POT) observations when $\xi > 0$. In the current study, after replacing the unknown parameters

including σ and ξ by their posterior means σ_{PM} and ξ_{PM} respectively, the mean ($\mathrm{M}_{\mathrm{POT}}$) and standard deviation ($\mathrm{SD}_{\mathrm{POT}}$) of the threshold exceedance returns are calculated as:

$$\begin{aligned}
\mathrm{M}_{\mathrm{POT}} &\simeq u + \frac{\sigma_{\mathrm{PM}}}{1-\xi_{\mathrm{PM}}} \\
&= -0.07 + \frac{0.218}{1-0.069} \\
&\simeq 0.164 \\
\mathrm{SD}_{\mathrm{POT}} &\simeq \frac{\sigma_{\mathrm{PM}}}{1-\xi_{\mathrm{PM}}}\sqrt{\frac{1}{1-2\xi}} \\
&\simeq \frac{0.218}{1-0.069}\sqrt{\frac{1}{1-2\times 0.069}} \\
&\simeq 0.272
\end{aligned}$$

Therefore, the average of the threshold exceedances of Malaysian gold return is 0.164. In other words, according to the threshold exceedance model, the benefit to the investor resulting from an investment in Malaysian gold market is around 0.164%. Additionally, based on the standard deviation ($\mathrm{SD}_{\mathrm{POT}}$), the risk of investment in Malaysian gold market is approximately 0.272% which is almost twice the size of the return.

The iteration series of the scale σ and shape ξ parameters of the threshold exceedance model are plotted in Figure 15a. These plots show stationarity in both aspects of mean and variance. Furthermore, Figure 15b displays the posterior distributions of the scale and shape parameters.



(a) Iteration series **(b)** Posterior densities

Figure 15. Iteration series and posterior distributions of the scale σ and shape ξ parameters of the threshold exceedances of Malaysian gold return with 5000 iterations.

Table 6. Results from the slice sampler algorithm for the threshold exceedances of Malaysian gold return by specifying the gamma and normal with large variances as prior distributions on the scale (σ) and shape (ξ) parameters of the GPD posterior model with 5000 iterations.

	σ	ξ
M_p	0.218	0.069
SD_p	0.007	0.019
SE_p	< 0.001	< 0.001
LCI	0.205	0.033
UCI	0.232	0.108

6. Conclusion

It is always challenging to justify the form of extreme models and to estimate parameters due to the inherent sparsity of tail information, see [7]. This research therefore develops the slice sampler algorithm for the peaks over a threshold (POT) and generalized Pareto distribution (GPD), since this model can minimize the problem of being wasteful of extreme information for collecting more extreme observations compared to maxima over blocks as used for the GEV. A simulation study has been used to show the performance of the slice sampler approach for fitting the POT and GPD posterior models, and is shown to perform well when compared to the Metropolis-Hastings algorithm. In fact, the slice sampler algorithm provides similar performance to the Metropolis-Hastings algorithm, but has an important benefit in that it reveals a high level of stationarity, thus making posterior estimation more straightforward. The slice sampler was also applied to threshold exceedances of Malaysia gold returns and was shown to provide posterior means with low errors for the parameters, which indicate the general applicability of the algorithm.

References

[1] Agarwal, D.K. and Gelfand, A.E. *Slice sampling for simulation based fitting of spatial data models*, Statistics and Computing **15**, 61-69, 2005.

[2] Balkema, A.A. and De Haan, L. *Residual Life Time at Great Age*, The Annals of Probability **2**, 792-804, 1974.

[3] Besag, J. and Green, P. *Spatial statistics and Bayesian computation*, Journal of the Royal Statistical Society. Series B (Methodological), 25-37, 1993.

[4] Castillo, E. *Extremes in engineering applications*, Extreme Value Theory and Applications, Springer, Dordrecht, Netherlands, 15-42, 1994.

[5] Chen, M.H. and Schmeiser, B.W. *[Slice Sampling]: Discussion*, Annals of Statistics, 42-743, 2003.

[6] Coles, S. *An introduction to statistical modeling of extreme values*, Springer-Verlag, 2001.

[7] Coles, S. and Powell, E. *Bayesian methods in extreme value modelling: a review and new developments*, International Statistical Review/Revue Internationale de Statistique, 119-136, 1996.

[8] Coles, S. , Pericchi, L.R. and S. Sisson, *A fully probabilistic approach to extreme rainfall modeling*, Journal of Hydrology **273**, 35-50, 2003.

[9] Damien, P. and Walker, S.G. *Sampling truncated normal, beta, and gamma densities*, Journal of Computational and Graphical Statistics **10**, 206-215, 2001.

[10] Damlen, P. , Wakefield, J. and Walker, S. *Gibbs sampling for Bayesian non-conjugate and hierarchical models by using auxiliary variables*, Journal of the Royal Statistical Society: Series B (Statistical Methodology) **61**, 331-344, 1999.

[11] Davison, A.C. and Smith, R.L. *Models for exceedances over high thresholds*, Journal of the Royal Statistical Society. Series B (Methodological), 393-442, 1990.

[12] De Haan, L. *On regular variation and its application to the weak convergence of sample extremes*, Mathematical Centre tracts Mathematisch Centrum, 1970.

[13] DuBois, C. , Korattikara, A. , Welling, M. and P. Smyth, *Approximate slice sampling for bayesian posterior inference*, Artificial Intelligence and Statistics, 2014.

[14] Edwards, R.G. and Sokal, A.D. *Dynamic critical behavior of Wolff's collective-mode Monte Carlo algorithm for the two-dimensional O (n) nonlinear a model*, Phys. Rev. D **40**, 1374-1377, 1989.

[15] Favaro, S. and Walker, S.G. *Slice sampling σ-stable Poisson-Kingman mixture models*, Journal of Computational and Graphical Statistics **22**, 830–847 2013.

[16] Fisher, R. and Tippett, L. *Limiting forms of the frequency distribution of the largest or smallest member of a sample*, Mathematical Proceedings of the Cambridge Philosophical Society **24**, 180-190, 1928.

[17] Fishman, G.S. *An analysis of Swendsen–Wang and related sampling methods*, Journal of the Royal Statistical Society: Series B (Statistical Methodology) **61** ,623-641, 1999.

[18] Fréchet, M. *Sur la loi de probabilité de l'écart maximum*, Annales de la societe Polonaise de Mathematique **6**, 93-116, 1927.

[19] Garavaglia, F. , Gailhard, J. , Paquet, E. , Lang, M. , Garçon, R. and Bernardara, P. *Introducing a rainfall compound distribution model based on weather patterns sub-sampling*, Hydrology and Earth System Sciences Discussions **14**, 951, 2010.

[20] Gençay, R. , Selçuk, F. and A. Ulugülyağci, *High volatility, thick tails and extreme value theory in value-at-risk estimation*, Insurance: Mathematics and Economics **33**, 337-356, 2003.

[21] Gilli, M. and Kellezi, E. *An application of extreme value theory for measuring financial risk*, Computational Economics **27** , 207-228, 2006.

[22] Gnedenko, B. *Stir La Distribution Limite Du Terme Maximum D'une Sarie Alaatoire*, Annales de Mathématiques , 423-453, 1943.

[23] Goldstein, J. , Mirza, M. , Etkin, D. , and Milton, J. *J2. 6 hydrologic assessment: Application of extreme value theory for climate extremes scenarios construction*, 14th Symposium on Global Change and Climate Variations, American Meteorological Society 83rd Annual Meeting, 2003.

[24] Hammersley, J. and Morton, K. *A new Monte Carlo technique: antithetic variates*, Mathematical Proceedings of the Cambridge Philosophical Society **52**, 449-475, 1956.

[25] Higdon, D. *Auxiliary variable methods for Markov chain Monte Carlo with applications*, Journal of the American Statistical Association **93**, 585-595, 1998.

[26] Hosking, J.R. and Wallis, J.R. *Parameter and quantile estimation for the generalized Pareto distribution*, Technometrics **29** , 339-349, 1987.

[27] Jagger, T.H. and Elsner, J.B. *Climatology models for extreme hurricane winds near the United States*, Journal of Climate **19**, 3220-3236, 2006.

[28] Kalli, M. , Griffin, J.E. and Walker, S.G. *Slice sampling mixture models*, Statistics and Computing **21**, 93-105, 2011.

[29] Liechty, M.W. and Lu, J. *Multivariate Normal Slice Sampling*, Journal of Computational and Graphical Statistics **19**, 281-294, 2010.

[30] McNeil, A.J. and Frey, R. *Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach*, Journal of Empirical Finance **7** , pp. 271-300, 2000.

[31] Mira, A. and Tierney, L. *Efficiency and convergence properties of slice samplers*, Scandinavian Journal of Statistics **29**, 1-12, 2002.

[32] Mira, A. and Tierney, L. *On the use of auxiliary variables in Markov chain Monte Carlo sampling*, Technical Report, University of Minnesota, School of Statistics, 1997.

[33] Mira, A. , Møller, J. and Roberts, G.O. *Perfect slice samplers*, Journal of the Royal Statistical Society: Series B (Statistical Methodology) **63**, 593-606, 2001.

[34] Murray,I. , Adams,R. and MacKay, D. *Elliptical slice sampling*, arXiv preprint arXiv:1001.0175, 2009.
[35] Neal, R. *Slice sampling*, Annals of Statistics, 705–741, 2003.
[36] Nieto, M. , Cortés, A. , Barandiaran, J. , Otaegui, O. and Etxabe, I. *Single Camera Railways Track Profile Inspection Using an Slice Sampling-Based Particle Filter*, Computer Vision, Imaging and Computer Graphics. Theory and Application, Springer, 326-339, 2013.
[37] Nishihara, R. , Murray, I. and Adams, R.P. *Parallel MCMC with generalized elliptical slice sampling*, The Journal of Machine Learning Research **15**2087-2112, 2014.
[38] Pickands III, J. *Statistical inference using extreme order statistics*, the Annals of Statistics, 119-131, 1975.
[39] Robert, C. and Casella, G. *Monte Carlo Statistical Methods Springer*, Citeseer **319**, 2004.
[40] Robert, C. and Casella, G. *Introducing Monte Carlo Methods with R*, Springer Science & Business Media, 2009.
[41] Roberts, G.O. and Rosenthal, J.S. *Convergence of slice sampler Markov chains*, Journal of the Royal Statistical Society: Series B (Statistical Methodology) **61**,643-660, 1999.
[42] Rocco, M. *Extreme value theory in finance: A survey*, Journal of Economic Surveys **28**, 82-108, 2014.
[43] Skold, M. *Computer intensive statistical methods*, Center for Mathematical Sciences, Lund University, 2005.
[44] Smith, R. *Extreme value statistics in meteorology and the environment*, Environmental Statistics **8**, 300-357, 2001.
[45] Smith, R. L. and Shively, T. S. *Point process approach to modeling trends in tropospheric ozone based on exceedances of a high threshold*, Atmospheric Environment **29**, pp. 3489–3499, 1995.
[46] Swendsen, R. and Wang, J. *Nonuniversal critical dynamics in Monte Carlo simulations*, Physical Review Letters **58**, 86-88, 1987.
[47] Tibbits, M.M. , Groendyke, C. , Haran, M. and Liechty, J.C. *Automated Factor Slice Sampling*, Journal of Computational and Graphical Statistics 23, 543-563, 2014.
[48] Tibbits, M.M. , Haran, M. and J.C. Liechty, *Parallel multivariate slice sampling*, Statistics and Computing **21**, 415-430, 2011.
[49] Trepanier, J.C. and Scheitlin, K.N. *Hurricane wind risk in Louisiana*, Natural hazards **70**, 1181-1195, 2014.
[50] Trotter, H.F. and Tukey, J.W. *Conditional Monte Carlo for Normal-Samples*, Symposium on Monte Carlo Methods, HA Meyer, ed.(New York: John Wiley, 1956) , 64, 1954.
[51] Walker, S.G. *Sampling the Dirichlet mixture model with slices*, Communications in Statistics-Simulation and Computation **36**, 45-54, 2007.
[52] Yildirim, I. *Bayesian Inference: Metropolis-Hastings Sampling*, Dept. of Brain and Cognitive Sciences, Univ. of Rochester, Rochester, NY, 2012.