A Tapered Pareto-Poisson Model for Extreme Pyroclastic Flows: Application to the Quantification of Volcano Hazards

by

Fan Dai

Program in Statistical and Economic Modeling
Duke University

Date: _____

Approved:

_____
Robert Wolpert, Supervisor

_____
Surya Tokdar

_____
Kent Kimbrough

Thesis submitted in partial fulfillment of the requirements for the degree of
Master of Science in Program in Statistical and Economic Modeling
in the Graduate School of Duke University
2015

Abstract

A Tapered Pareto-Poisson Model for Extreme Pyroclastic Flows: Application to the Quantification of Volcano Hazards

by

Fan Dai

Program in Statistical and Economic Modeling
Duke University

Date: _____

Approved:

_____
Robert Wolpert, Supervisor

_____
Surya Tokdar

_____
Kent Kimbrough

An abstract of a Thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Program in Statistical and Economic Modeling in the Graduate School of Duke University
2015

Abstract

This paper intends to discuss the problems of parameter estimation in a proposed tapered Pareto-Poisson model for the assessment of large pyroclastic flows, which are essential in quantifying the size and risk of volcanic hazards. In dealing with the tapered Pareto distribution, the paper applies both maximum likelihood estimation and a Bayesian framework with objective priors and Metropolis algorithm. The techniques are further illustrated by an example of modeling extreme flow volumes at Soufriere Hills Volcano, and their simulation results are addressed.

Contents

List of Tables

List of Figures

List of Abbreviations and Symbols

Symbols

X	A random variable X.
Y	A random variable Y.
Z	A random variable Z.
z	A random variable z.
Ga	Gamma distribution.
Ex	Exponential distribution.
Poi	Poisson distribution.
ϵ	Default threshold ϵ.
β	Shape parameter β.
η	Rate parameter η.
θ	Location parameter θ.
δ_1	Proposal standard deviation δ_1.
δ_2	Proposal standard deviation δ_2.
r	Accepted ratio r.
V_j	Extreme volume V_j.
τ_j	Time point τ_j.
$\ell(\lambda, \beta, \eta)$	Log-likelihood function $\ell(\lambda, \beta, \eta)$.
$\mathrm{H}(\beta, \eta)$	Hessian Matrix $\mathrm{H}(\beta, \eta)$.
$\pi(\beta, \eta)$	Prior distribution $\pi(\beta, \eta)$.

Abbreviations

SNF	National Science Foundation.
MLEs	Maximum Likelihood Estimations.
SD	Standard Deviation.

Acknowledgements

First of all, I give my particular thanks to Professor Alan Gelfand, for introducing me to the field of applying statistical techniques to Environmental Science, and Professor Robert Wolpert, my supervisor, for enlightening me with his special intelligence and giving me confidence and knowledge on the research road.

Second, I wish to express my sincere appreciation to my committee members who were more than generous with their expertise and precious time. A special feeling of gratitude to Professor Mike West, our Director of Graduate Studies, and Professor Surya Tokdar, our Master's Program Director for their countless hours of instructing, encouraging, and most of all patience throughout my entire master's studies. Thank you Professor Kent Kimbrough for agreeing to serve on my committee.

Finally, I wish to thank my family and all of my friends, especially my parents for never leaving me side, my cousin brother's family for always supporting and encouraging me, my friends Sheng Jiang who have shared his invaluable time, insights to help me deal with problems and provide me useful advice, and Shuxian Chen, when time gets tough, we know we have each other.

1

Introduction

This paper focuses on modeling the size of geophysical mass flows, such as debris flows, volcanic avalanches and landslides (Lee Siebert, 1984), which are caused by earthquakes or volcanic activities. These severe environmental disturbances are supposed to be important index for measuring the risk of seismic and volcanic hazards, i.e. the volume of the pyroclastic flows exceeding an upper point for a certain time period would suggest an intensification of volcanic activities, which are becoming focuses in a growing amount of statistical applications to environmental data.

My work is mainly inspired by Professor Robert L. Wolpert, who has been one of the statisticians in the NSF collaborative research: Prediction and Risk of Extreme Events Utilizing Mathematical Computer Models of Geophysical Processes, in which a Pareto-Poisson model was proposed for quantifying the pyroclastic mass flows at Soufriere Hills, providing comprehensive material and a solid foundation for the subject of this paper.

In this paper, we determine the size of volcanic activities as the daily average level of the pyroclastic flow volumes recorded in the Soufriere Hills data set (The data set is compiled largely from the Montserrat Volcano Observatory daily reports). As the

volume's magnitude exceeding an observational threshold increases, the frequency of occurrences follows approximately a power-law decrease (Figure 4.1), which is hence typically fitted with a heavy tailed Pareto distribution.

However, in real-life applications, observations of extreme events are limited by physical constraints, i.e. The volcanic flow volumes could reach more than a thousand cubic kilometers, while none on this large scale would appear for hundred thousand years. But a Pareto model still permits higher density for those rarely occurred cases. Additionally, with respect to the volcanic data for Soufriere Hills, the sample tail (Figure 4.1) also shows a faster decay than the fitted Pareto line, indicating evidences for possible improvements to the upper tail part of the pure Pareto distribution.

Two common methods used to modify the Pareto distribution are truncated Pareto and tapered Pareto. Compared with the truncated Pareto that introduces a fixed upper threshold to control the distribution tail, the tapered Pareto distribution has an exponential taper to zero and provides much more richness and flexibilities when fitted to data by including a shape parameter and a location parameter for the tapering effect. Therefore, the tapered Pareto distribution has gained a wide use in quantifying the size of earthquakes and wildfires (Kagan et al., 2001). We extend the applications of the tapered Pareto to model volcanic mass flows, and introduce a new Bayesian framework to estimate model parameters.

Next chapter introduces the tapered Pareto distribution and implements it into a Poisson model, in order to describe the stochastic properties of large pyroclastic flows. Chapter 3 investigates the problems of estimating model parameters via maximum likelihood estimations (MLEs) and Bayesian methods. Chapter 4 presents model fitting results based on the Soufriere Hills data set. The final part discusses the problem of model selection and possible extensions for our proposed model.

2

Modeling Extreme Pyroclastic Flows

In this chapter, we explain basic properties of tapered Pareto distribution and construct a tapered Pareto-Poisson model for extreme pyroclastic flows.

2.1 Tapered Pareto Distribution

A tapered Pareto distribution modifies the Pareto distribution by adding an exponential tapering effect to the survival function, hence it is closely related to both Pareto and Exponential distributions. Define random variable $X > \epsilon$ to be Pareto distributed with shape parameter β, and $Y - \epsilon$ to be Exponential with rate parameter η, where X and Y are independent. Then, a tapered Pareto distributed variable Z can be obtained by $Z = \min(X, Y)$ (David Harte, 2010). Since:

$$P(X > z) = (\frac{z}{\epsilon})^{-\beta}, z > \epsilon > 0, \beta > 0 \tag{2.1}$$

And,

$$P(Y > z) = e^{-\eta(z-\epsilon)}, z > \epsilon > 0, \eta > 0 \tag{2.2}$$

So,

$$P(Z > z) = P(X > z, Y > z) = (\frac{z}{\epsilon})^{-\beta} e^{-\eta(z-\epsilon)}, z > \epsilon > 0, \beta, \eta > 0 \tag{2.3}$$

Which is the survival function of the tapered Pareto distribution, indicating a relatively lower frequency of extreme occurrences (Figure 2.1).

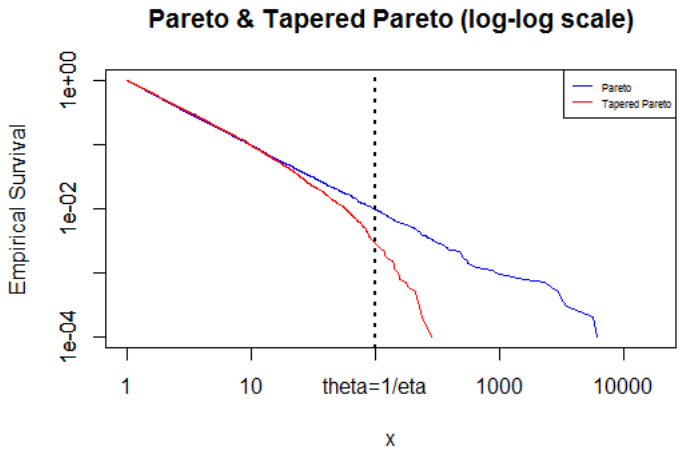

FIGURE 2.1: Survival Function: Pareto vs Tapered Pareto

The cumulative distribution function for the tapered Pareto distribution is:

$$F(z) = 1 - (\frac{z}{\epsilon})^{-\beta} e^{-\eta(z-\epsilon)}, z > \epsilon > 0, \beta, \eta > 0 \tag{2.4}$$

And the density:

$$f(z) = (\frac{\beta}{z} + \eta)(\frac{z}{\epsilon})^{-\beta} e^{-\eta(z-\epsilon)}, z > \epsilon > 0, \beta, \eta > 0 \tag{2.5}$$

Where ϵ represents a lower threshold which is known, β is the shape parameter that controls the slope of the power-law decrease of frequency versus magnitude. For the new rate parameter η, $\theta = 1/\eta$ determines the location of the exponential taper to zero (Tapering effect can become negligible if the magnitude of θ is far more than that of $z - \epsilon$).

2.2 Tapered Pareto-Poisson Model

We use the tapered Pareto distribution to model the pyroclastic flow volume $V_j > v$, at time $\tau_j \geqslant 0$:

$$P(V_j > v) \sim (\frac{v}{\epsilon})^{-\beta} e^{-\eta(v-\epsilon)}, v > \epsilon > 0, \beta, \eta > 0 \tag{2.6}$$

Independence is assumed among these tapered Pareto-distributed volumes.

Then, we employ a marked Poisson process for the initial mass flows whose volumes exceed a fixed lower point $\epsilon > 0$. The tapered Pareto-Poisson model can be constructed as follows:

1. The number, N, of flow volumes exceeding the level ϵ in time interval T has a Poisson distribution with mean λT.

2. Conditionally on $N \geqslant 1$, the extreme volumes $V_1, ..., V_j, ..., V_N$ are independent and identically distributed from the tapered Pareto distribution.

The proposed tapered Pareto-Poisson model follows an assumption of stationary over a given time period (Bayarri et al., 2009), which means that the constant Poisson rate λ is only associated with the observational threshold ϵ.

The likelihood of tapered Pareto-Poisson model is:

$$\begin{aligned} P(V_1, ..., V_j, ..., V_N, N = n) =& \text{Poi}(T\lambda) \prod_j f(V_j) \\ \propto& e^{-T\lambda} \lambda^n \prod_j (\beta + \eta V_j)(\frac{V_j}{\epsilon})^{-\beta} e^{-\eta(V_j - \epsilon)} \end{aligned} \tag{2.7}$$

And the log-likelihood:

$$\ell(\lambda, \beta, \eta) \propto n \log \lambda - T\lambda + \sum_{j=1}^{n} \log(\beta + \eta V_j) - \beta \sum_{j=1}^{n} (\log V_j - \log \epsilon) - \eta \sum_{j=1}^{n} (V_j - \epsilon) \tag{2.8}$$

Where $\beta, \eta > 0, V_j > \epsilon > 0, j = 1, ..., n$

3

Parameter Estimation

Based on the proposed model in Chapter 2, the maximum likelihood estimates for parameters λ, β and η can be derived from the gradients of the log-likelihood function. As an alternative, the Bayesian estimates, which include parameter uncertainty in the probability model, can be simulated by Random Walk Metropolis algorithm.

3.1 Maximum Likelihood Estimations

The three gradients of $\ell(\lambda, \beta, \eta)$ are:

$$\begin{cases} \dfrac{\partial \ell(\lambda,\beta,\eta)}{\partial \lambda} = \dfrac{n}{\lambda} - T \\ \dfrac{\partial \ell(\lambda,\beta,\eta)}{\partial \beta} = \displaystyle\sum_{j=1}^{n} \frac{1}{\beta + \eta V_j} - \sum_{j=1}^{n} \log V_j + n \log \epsilon \\ \dfrac{\partial \ell(\lambda,\beta,\eta)}{\partial \eta} = \displaystyle\sum_{j=1}^{n} \frac{V_j}{\beta + \eta V_j} - \sum_{j=1}^{n} V_j + n\epsilon \end{cases} \tag{3.1}$$

Set the partial derivatives into zero, we have:

MLE for $\lambda : \hat{\lambda} = n/T$;

MLEs for β and η satisfy:

$$\begin{cases} \sum_{j=1}^{n} \frac{1}{\beta + \eta V_j} = \sum_{j=1}^{n} \log V_j - n \log \epsilon \\ \sum_{j=1}^{n} \frac{V_j}{\beta + \eta V_j} = \sum_{j=1}^{n} V_j - n\epsilon \end{cases} \tag{3.2}$$

With Hessian matrix $\mathrm{H}(\beta, \eta)$:

$$\begin{bmatrix} -\sum_{j=1}^{n} \frac{1}{(\beta+\eta V_j)^2} & -\sum_{j=1}^{n} \frac{V_j}{(\beta+\eta V_j)^2} \\ -\sum_{j=1}^{n} \frac{V_j}{(\beta+\eta V_j)^2} & -\sum_{j=1}^{n} \frac{V_j^2}{(\beta+\eta V_j)^2} \end{bmatrix}$$

It is easy to compute MLEs in R by numerical fitting routines, i.e. R function "nlm" or "optim". However, one primary difficulty of applying MLEs in practice is that the observations of extreme events which essentially dominate the estimation of taper Pareto parameters are very rare, possible to generate inaccurate estimations, especially for η. An alternative is to use Bayesian approaches, which allow stochastic uncertainty for model parameters.

3.2 Bayesian Inference

The following parts present a Bayesian framework where parameters are estimated with objective priors and Random Walk Metropolis Algorithm.

3.2.1 Parameter's Priors

Since there is little prior information available concerning λ, β and η, we chiefly propose objective priors. The popular conjugate prior for Poisson rate λ is the gamma distribution $\mathrm{Ga}(a, b)$, which is assumed independent of β and η.

Considering the relationships of tapered Pareto with Pareto and Exponential (Section 2.1), we propose a gamma prior $\mathrm{Ga}(c, d)$ for β, and an exponential prior

$\text{Ex}(\gamma)$ for η. Both are commonly used priors for the shape parameter in Pareto distribution and the rate parameter in Exponential distribution. More importantly, they can guarantee proper posterior distributions in the tapered Pareto model.

Thus, for β, η, we have:

$$\pi(\beta, \eta) \sim \beta^{c-1} e^{-d\beta} e^{-\gamma\eta} \tag{3.3}$$

The choice of the hyper-parameter γ should help prevent trivial tapering effect.

Parameter's Posteriors

Based on above priors, the posterior distributions are derived from the likelihood function as follows:

For λ,

$$p(\lambda|\beta, \eta, V_1, ..., V_n) \propto \lambda^{n+a-1} e^{-\lambda(T+b)} \tag{3.4}$$

For β and η,

$$p(\beta, \eta|\lambda, V_1, ..., V_n) \propto \prod_j (\beta + \eta V_j)(\frac{V_j}{\epsilon})^{-\beta} e^{-\eta(V_j - \epsilon)} \beta^{c-1} e^{-d\beta} e^{-\gamma\eta} \tag{3.5}$$

3.2.2 Random Walk Metropolis Algorithm

The conjugated gamma posterior for λ is approximated with the Gibbs sampler. The non-conjugate posterior for β, η can be simulated via Random Walk Metropolis algorithm, which allows normal proposal distributions for the log transformations of $\beta, \eta > 0$ as follows:

$$\tilde{\beta} = \log \beta, \tilde{\eta} = \log \eta \tag{3.6}$$

Symmetric proposal distributions:

$$q(\tilde{\beta}^*|\tilde{\beta}) \sim \text{Normal}(\tilde{\beta}, \delta_1^2), q(\tilde{\eta}^*|\tilde{\eta}) \sim \text{Normal}(\tilde{\eta}, \delta_2^2), \delta_1, \delta_2 > 0 \tag{3.7}$$

So,

$$\begin{aligned}\tilde{\beta}^* &= \tilde{\beta} + \delta_1 z_1, z_1 \sim \text{Normal}(0,1), \delta_1 > 0 \\ \tilde{\eta}^* &= \tilde{\eta} + \delta_2 z_2, z_2 \sim \text{Normal}(0,1), \delta_2 > 0\end{aligned} \tag{3.8}$$

Then, for β, η,

$$\begin{aligned}\beta^* &= \beta e^{\delta_1 z_1}, z_1 \sim \text{Normal}(0,1), \delta_1 > 0 \\ \eta^* &= \eta e^{\delta_2 z_2}, z_2 \sim \text{Normal}(0,1), \delta_2 > 0\end{aligned} \tag{3.9}$$

By function transformation of random variables,

$$\begin{aligned}q(\beta^*|\beta) &= q(\log\beta^*|\log\beta) \times \frac{1}{\beta^*} \\ q(\eta^*|\eta) &= q(\log\eta^*|\log\eta) \times \frac{1}{\eta^*}\end{aligned} \tag{3.10}$$

$$\frac{q(\beta|\beta^*)}{q(\beta^*|\beta)} = \frac{\beta^*}{\beta}, \frac{q(\eta|\eta^*)}{q(\eta^*|\eta)} = \frac{\eta^*}{\eta} \tag{3.11}$$

Therefore, the acceptance ratio r can be computed as:

$$r = \frac{p(\beta^*, \eta^*|\lambda, V_1, ..., V_n) q(\beta|\beta^*) q(\eta|\eta^*)}{p(\beta, \eta|\lambda, V_1, ..., V_n) q(\beta^*|\beta) q(\eta^*|\eta)} = \frac{p(\beta^*, \eta^*|\lambda, V_1, ..., V_n)\beta^*\eta^*}{p(\beta, \eta|\lambda, V_1, ..., V_n)\beta\eta} \tag{3.12}$$

For the tapered Pareto-Poisson model, the $\log r$ is:

$$\begin{aligned}\log r &= \ell(\lambda, \beta^*, \eta^*) - \ell(\lambda, \beta, \eta) + \log\frac{\pi(\beta^*, \eta^*)}{\pi(\beta, \eta)} + \log\frac{\beta^*\eta^*}{\beta\eta} \\ &= \sum_{j=1}^{n} \log\frac{\beta^* + \eta^* V_j}{\beta + \eta V_j} - (\beta^* - \beta)\sum_{j=1}^{n}\log\frac{V_j}{\epsilon} - (\eta^* - \eta)\sum_{j=1}^{n}(V_j - \epsilon) \\ &+ c\log\frac{\beta^*}{\beta} - d(\beta^* - \beta) - \gamma(\eta^* - \eta) + \log\frac{\eta^*}{\eta}\end{aligned} \tag{3.13}$$

Under Random Walk Metropolis, convergences of Markov chains are sensitive to scale parameters. Therefore, to promote the proficiency of our algorithm, we employ

Component-wise Metropolis in order to find proper δ_1, δ_2 that keep the rejected rates within $[0.15, 0.5]$ (The default interval was addressed by Robert et al., 1997). The simulation algorithm is shown as follows:

For $s = 1, ..., S$:

1. Generate λ^s from $\text{Ga}(n + a, T + b)$;
2. Generate random variable z_1 from $\text{Normal}(0, 1)$ and compute $\beta^* = \beta^{s-1} e^{\delta_1 z_1}$;
3. Compute $\log r_1$ based on $\lambda^s, (\beta^*, \eta^{s-1}), (\beta^{s-1}, \eta^{s-1})$ and $V_1, ..., V_n$;
4. Generate random variable ε_1 from $\text{Ex}(1)$, set (β^s, η^{s-1}) to (β^*, η^{s-1}) if $\varepsilon_1 + \log r_1 \geqslant 0$, or to $(\beta^{s-1}, \eta^{s-1})$ if $\varepsilon_1 + \log r_1 < 0$;
5. Generate random variable z_2 from $\text{Normal}(0, 1)$ and compute $\eta^* = \eta^{s-1} e^{\delta_2 z_2}$;
6. Compute $\log r_2$ based on $\lambda^s, (\beta^s, \eta^*), (\beta^s, \eta^{s-1})$ and $V_1, ..., V_n$;
7. Generate random variable ε_2 from $\text{Ex}(1)$, set (β^s, η^s) to (β^s, η^*) if $\varepsilon_2 + \log r_2 \geqslant 0$, or to (β^s, η^{s-1}) if $\varepsilon_2 + \log r_2 < 0$.

4

Data Analysis: Extreme Pyroclastic Flows at Soufriere Hills

The data set of Soufriere Hills Volcano contains 900 daily records starting from March 27, 1996 to December 31, 2008 since the eruption in 1995, in which there are 899 non-negative and non-missing daily measurements of the pyroclastic flows. We are interested in modeling the daily average level between the new minimum and maximum volumes ($\times 10^6 m^3$).

4.1 Threshold ϵ and Extreme Volumes

Empirical survival function of the average volumes is plotted on log-log scale in Figure 4.1, from which we can see that after a lower truncated point $\epsilon = 0.5 \times 10^6 m^3$, the trajectory indicates linear properties and thus a power-law decrease of the frequency of extreme volumes. However, we also find that the sample upper tail decays faster than the fitted line, suggesting a possible tapering effect.

Based on above evidence, we construct a tapered Pareto-Poisson model for the average extreme flows of $V_1, ..., V_n$, where $V_j > \epsilon = 0.5 \times 10^6 m^3, n = 138, T = 899$ (Here the λ represents the daily rate of extreme occurrences). Also, In the Bayesian

inference, we choose $\gamma = 10^2$ in η's exponential prior which is comparable to the magnitude of $V_j - \epsilon$.

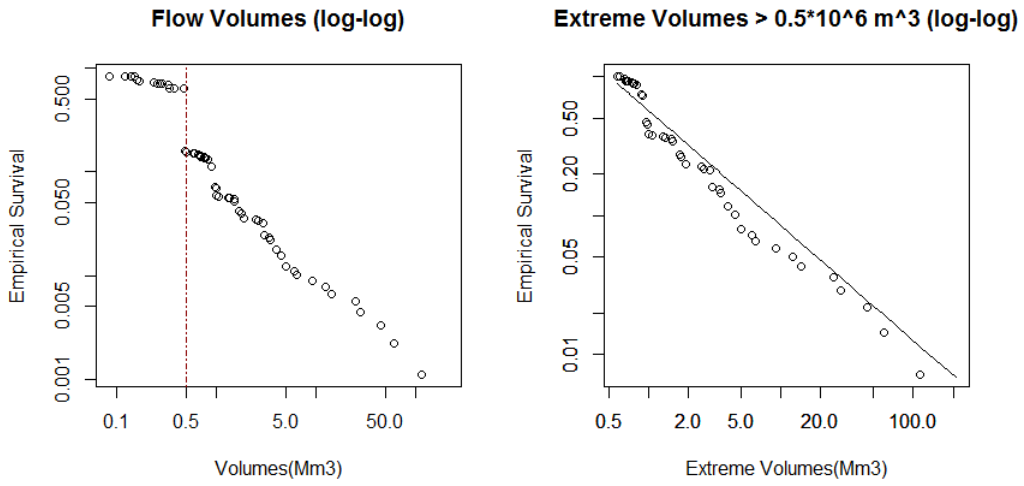

FIGURE 4.1: Volumes Plot

4.2 Parameter Estimations

The maximum likelihood estimations are given as: $\hat{\lambda} = 0.154, \hat{\beta} = 0.863$ and $\hat{\eta} = 0.003$. Summaries of the Bayesian estimates are shown in table 4.1. The posterior means of λ, β, η are much closed to MLEs, while Figure 4.2 illustrates the posterior models influenced by parameter uncertainty, mainly generated from η, which was simulated with a large proposal variance.

Table 4.1: Bayesian Estimation

Parameter	Mean	CI ($\alpha = 5\%$)	Rejected rate and proposal SD
λ	0.154	(0.130, 0.181)	–
β	0.855	(0.715, 1.006)	$33.2\%, \delta_1 = 0.1$
η	0.005	(0.0002, 0.159)	$33.7\%, \delta_2 = 1$

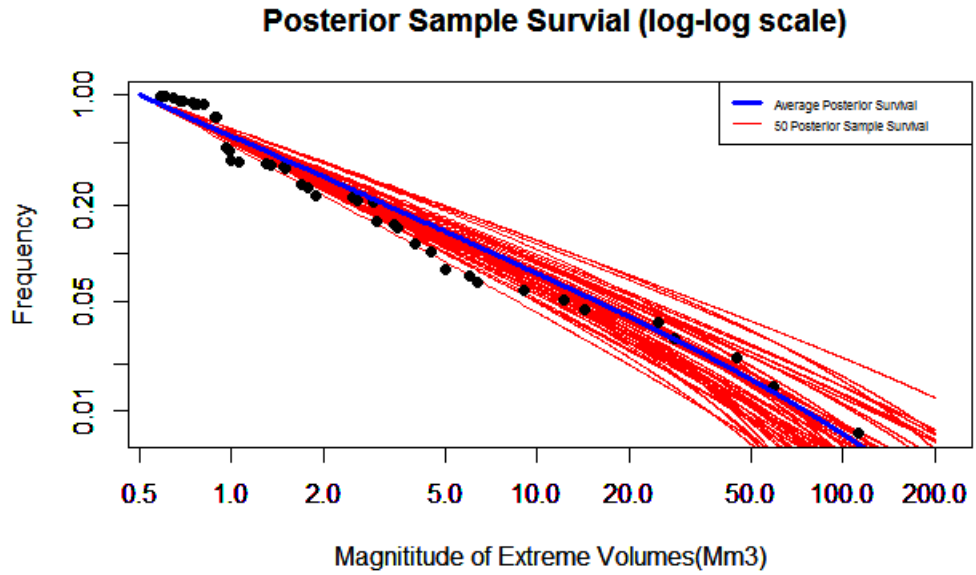

FIGURE 4.2: Posterior sample Survival Function

Figure 4.3 and Figure 4.4 show good convergences of the three posterior samples and that the Markov chains of β and η are weakly negative-correlated.

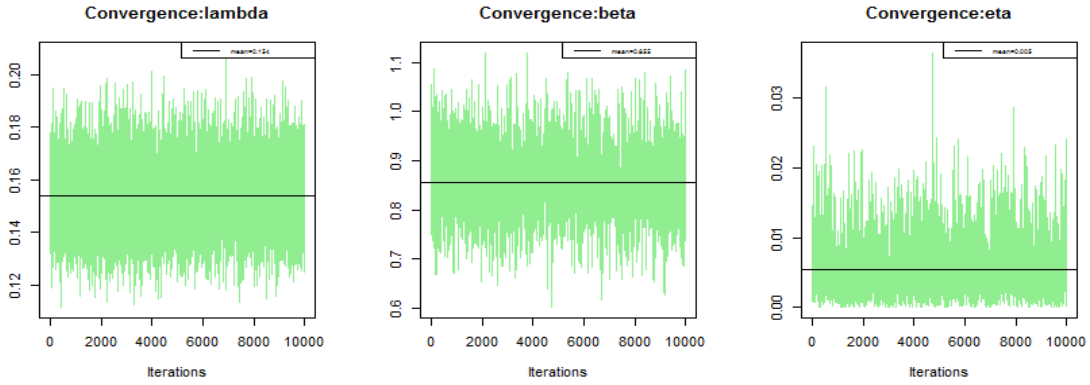

FIGURE 4.3: Convergence of Parameter Samples

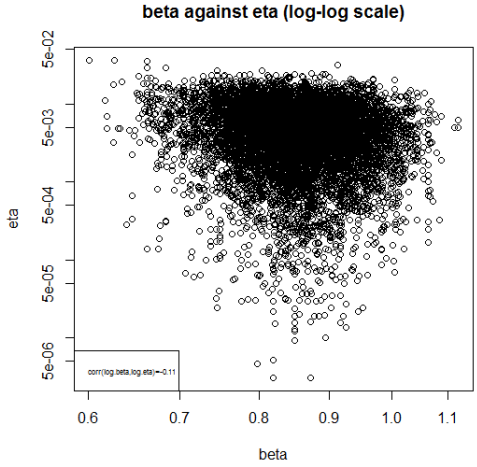

FIGURE 4.4: Correlation between β and η

4.3 Goodness of Fit

We compare the data's empirical survival function with the proposed ones in order to evaluate the local goodness of fit. For the tapered Pareto with maximum likelihood estimations $\hat{\beta}, \hat{\eta}$, we have:

$$P_{\text{mle}}(V_j > v) \sim (\frac{v}{\epsilon})^{-\hat{\beta}} e^{-\hat{\eta}(v-\epsilon)}, v > \epsilon > 0, \beta, \eta > 0 \tag{4.1}$$

For the tapered Pareto with Bayesian estimations, we have:

$$P_{\text{mcmc}}(V_j > v) \sim \frac{1}{M} \sum_{m=1}^{M} (\frac{v}{\epsilon})^{-\beta_m} e^{-\eta_m(v-\epsilon)}, v > \epsilon > 0, \beta, \eta > 0 \tag{4.2}$$

Where $\beta_1, ..., \beta_m, ..., \beta_M$ and $\eta_1, ..., \eta_m, ..., \eta_M$ are MCMC samples with size of M.

These two proposed survival function were plotted with the empirical one and the fitted linear function (Pareto distribution) in Figure 4.5, which indicates that the data tends to support the tapered Pareto model in the upper tail, although the differences are relatively small.

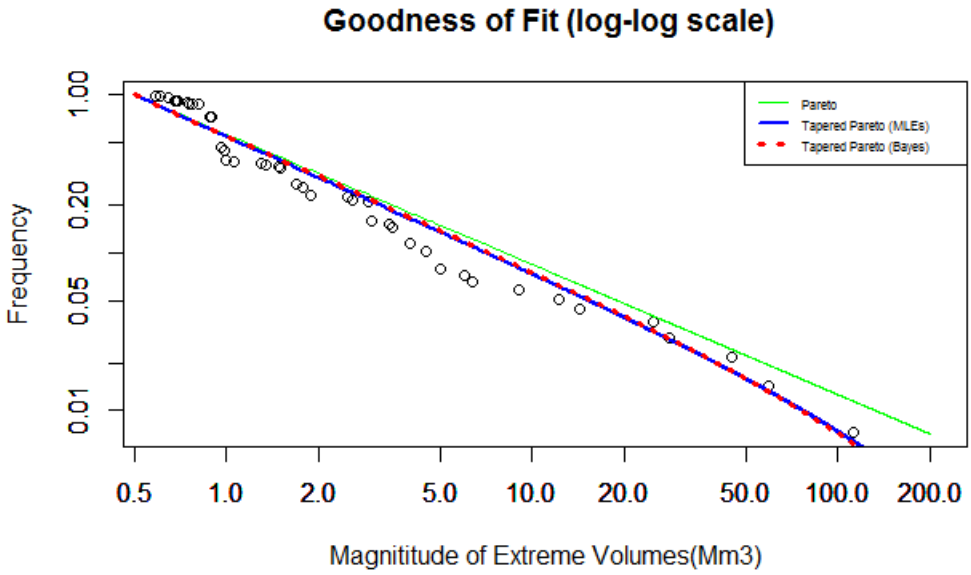

FIGURE 4.5: Goodness of Fit

5

Discussion

In this chapter, the model fitting results are further addressed. Then, several potential improvements are proposed to extend our current model.

5.1 Model Selection

From Table 4.1, we can see that the estimated rate parameter $\eta > 0$ has a very small value either in MLEs or Bayesian inference. Consequently, the estimated $\theta = 1/\eta$ appears extremely large compared with $V_j - \epsilon$, indicating a weak tapering effect fitted to the data. Additionally, the 95% confidence interval of η ([0.0002, 0.159]) also suggests strong evidence that favors of a negligible exponential taper.

However, the fitted results in Figure 6 still shows superior fit of the tapered Pareto distribution to the Pareto distribution in the upper tail. We could expect that provided sufficient observations of extreme pyroclastic flows, the tapered Pareto distribution, which can be easily combined with point processes with a not complicated structure, would gain more preference.

5.2 Proposed Extentions

The proposed model assumes that the tapered Pareto-distributed volumes $V_1, ..., V_j$ are independent of each other, but it is possible that they could correlated because they are often initiated by similar volcanic activities. Furthermore, The Poisson process may have non-stationary properties during a certain time period given that the occurrence rate would change over time, which means we also need to consider the Poisson rate as time series, $\lambda_1, ..., \lambda_t, ...$

On the other hand, additional to quantify the average of the flow volumes, a more appealing choice would be modeling the accumulated volumes $X_t = \sum_{j=1}^{n} V_j$, which was once explored by Bayarri et al., 2009 via a skewed α-stable distribution. The new derived variable X_t may introduce possible correlations of the Poisson rate λ with the tapered Pareto parameters β and η.

Appendix A

R Code

```
# MLE for lambda
lambda.mle=length(vols)/899
# neg-loglikelihood
nllh = function(par, vols, eps) {
bet = exp(par[1]); nu = exp(par[2]);
rv = bet*sum(log(vols/eps)) + nu*sum(vols-eps) - sum(log(bet/vols+nu));
return(rv);
}

# MLEs for beta, eta
fit = nlm(nllh, p=c(log(1),log(1e-2)),vols=vols,eps=eps,hessian=TRUE)
beta.mle=exp(fit$estimate[1])
eta.mle=exp(fit$estimate[2])
para.mle=c(lambda.mle, beta.mle, eta.mle)
```

```
# Bayesian Inference
rw.single.Metropolis = function(sigma, it, burnin, params, vols){
# lambda prior: gamma(1,2)
# beta prior: gamma(1,1)
# eta prior: exponential(100)
# it & burnin: size of interation and burn-in
# params=c(beta, eta): initial params for beta  eta
# sigma=c(sigma1, sigma2): sd for proposal distribution
# vols: vector of extreme volumes
params.samp = matrix(nrow=it, ncol=3)
accepted = c(0,0)
params.final = NULL
# iteration
for (i in 1:it){
# sample lambda
lambda = rgamma(1,length(vols)+1,899+2)
# Metropolis-Histings Algorithm
# update beta
beta=params[1];eta=params[2]
beta.star=beta*exp(sigma[1]*rnorm(1))
log.r1 = sum(log((beta.star+eta*vols)/(beta+eta*vols)))-
(beta.star-beta)*sum(log(vols/0.5))+log(beta.star/beta)-(beta.star-beta)
if(rexp(1)+log.r1>0){
beta=beta.star
params=c(beta,eta)
accepted[1]=accepted[1]+1
}
```

```
# update eta
beta=params[1];eta=params[2]
eta.star=eta*exp(sigma[2]*rnorm(1))
log.r2 = sum(log((beta+eta.star*vols)/(beta+eta*vols)))-
(eta.star-eta)*sum(vols-0.5)-100*(eta.star-eta)+log(eta.star/eta)
if(rexp(1)+log.r2>0){
eta=eta.star
params=c(beta,eta);
accepted[2]=accepted[2]+1
}
params.samp[i,]=c(lambda,params)
}
params.final = params.samp[(burnin+1):it, ]
return(list(params.final=params.final, r=1-accepted/it))
}
```

Bibliography

Bayarria, M. J., Berger, J. O., Calderc, E. S., Dalbeyd, K., Lunagomeze, S., Patraf, A. K., Pitmang, E. B., Spillerh, E. T., and Wolpert, R. L. (2009), "Using statistical and computer models to quantify volcanic hazards," *Technometrics*, 16, 402–413.

Burgman, M., Franklin, J., Hayes, K., Hosack, G., Peters, G., and Sisson, S. (2012), "Modeling extreme risks in ecology," *Risk Analysis*, 32, 1956–1966.

Cai, W., Santoso, A., Wang, G., Weller, E., Wu, L., Ashok, K., Masumoto, Y., and Yamagata, T. (2014), "Increased frequency of extreme Indian Ocean Dipole events due to greenhouse warming," *Nature*, 501, 254–258.

Consul, P. C. and Jain, G. C. (1973), "On some interesting properties of the generalized Poisson distribution," *Biometrische Zeitschrift*, 15, 495–500.

Falk, M., Hüsler, J., and Reiss, R.-D. (2006), "Laws of Small Numbers: Extremes and Rare Events," *Metrika*, 64, 127–129.

Harte, D. (2010), "PtProcess: An R Package for Modelling Marked Point Processes Indexed by Time," *Journal of Statistical Software*, 35.

Jaramillo, E., Dugan, J. E., Hubbard, D. M., Melnick, D., Manzano, M., Duarte, C., Campos, C., and Sanchez, R. (2012), "Ecological Implications of Extreme Events: Footprints of the 2010 Earthquake along the Chilean Coast," *PLoS ONE*, 7.

Kagan and Y., Y. (2010), "Earthquake size distribution: Power-law with exponent $\beta = 1/2$," *Tectonophysics*, 490, 103–114.

Kagan, Y. Y. and Schoenberg, F. (2001), "Estimation of the Upper Cutoff Parameter for the Tapered Pareto Distribution," *Journal of Applied Probability*, 38, 158–175.

Malevergne, Y., Pisarenko, V., and Sornette, D. (2006), "On the power of generalized extreme value (GEV) and generalized Pareto distribution (GPD) estimators for empirical distributions of stock returns," *Applied Financial Economics*, 3, 213–289.

Pisarenko, V., Sornette, A., Sornette, D., and Rodkin, M. (2014), "Characterization of the Tail of the Distribution of Earthquake Magnitudes by Combining the GEV and GPD Descriptions of Extreme Value Theory," *Pure and Applied Geophysics*, 171, 1599–1624.

Roberts, G. O., Gelman, A., and Gilks, W. R. (1997), "Weak convergence and optimal scaling of random walk Metropolis algorithms," *The Annals of Applied Probability*, 7, 110–120.

Rossi, F., Fiorentino, M., and Versace, P. (1984), "Two-Component Extreme Value Distribution for Flood Frequency Analysis," *Water Resources Research*, 20, 847–856.

Schoenberg, F. and Patel, R. (2012), "Comparison of Pareto and tapered Pareto distributions for environmental phenomena," *The European Physical Journal Special Topics*, 205, 159–166.

Schoenberg, F. P., Barr, C., and Seo, J. (2009), "The distribution of Voronoi cells generated by Southern California earthquake epicenters," *Environmetrics*, 20, 159–171.

Sobradelo, R., Martí, J., Mendoza-Rosas, A. T., and Gómez, G. (2011), "Volcanic hazard assessment for the Canary Islands (Spain) using extreme value theory," *Natural Hazards and Earth System Science*, 11, 2741–2753.

Sornette, D. and Ouillon, G. (2012), "Dragon-kings: mechanisms, statistical methods and empirical evidence," *The European Physical Journal Special Topics*, 205, 1–26.

Zwiers, F. W., Alexander, L. V., Hegerl, G. C., Knutson, T. R., Kossin, J. P., Naveau, P., Nicholls, N., Schär, C., Seneviratne, S. I., and Zhang, X. (2013), *Climate Extremes: Challenges in Estimating and Understanding Recent Changes in the Frequency and Intensity of Extreme Climate and Weather Events*, pp. 339–389, Springer.

Zöllera, G., Holschneidera, M., Hainzlb, S., and Zhuang, J. (2014), "The Largest Expected Earthquake Magnitudes in Japan: The Statistical Perspective," *Seismological Society of America*, 104, 769–779.