UNIVERSITI PUTRA MALAYSIA

SLICE SAMPLER AND METROPOLIS HASTINGS APPROACHES FOR BAYESIAN ANALYSIS OF EXTREME DATA

MOHAMMAD ROSTAMI

IPM 2016 10

SLICE SAMPLER AND METROPOLIS HASTINGS APPROACHES FOR BAYESIAN ANALYSIS OF EXTREME DATA

By

MOHAMMAD ROSTAMI

Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfilment of the Requirements for the Degree of Doctor of Philosophy

February 2016

DEDICATIONS

To my dad & mum
for their love, endless support and encouragement.

Abstract of thesis presented to the Senate of Universiti Putra Malaysian in fulfilment of the requirement for the degree of Doctor of Philosophy

SLICE SAMPLER AND METROPOLIS HASTINGS APPROACHES FOR BAYESIAN ANALYSIS OF EXTREME DATA

By

MOHAMMAD ROSTAMI

February 2016

Chairman: Mohd Bakri Adam, PhD
Faculty : Institute For Mathematical Research

Modelling the tails of distributions is important in many areas of research where the risk of unusually small or large events are of interest. In this research, application of extreme value theory within a Bayesian framework using the Metropolis Hastings algorithm and the slice sampler algorithm as an alternative approach, has been introduced.

Selection of prior distributions are very crucial in Bayesian analysis. Here, we have exhaustedly studied all the possible priors for location and scale parameters and come out with a few suggestions for the prior selection of a Gumbel model.

The slice sampler method can adaptively change the scale of changes made, which makes it easier to tune than Metropolis Hastings algorithm. Another important benefit of the slice sampler algorithm is that it provides posterior means with low errors for the shape parameters of the monthly maximum and threshold exceedances models. The slice sampler algorithm has been extended for more complex bivariate extreme value model with logistic dependence structure and exponential margins. A simulation study shows that the slice sampler algorithm provides posterior means with low errors for the parameters along with a high level of stationarity in iteration series. Furthermore, the slice sampler algorithm has been successfully applied to Malaysian gold returns which has been calculated using Malaysian daily gold prices from 2000 to 2011. By using a Bivariate extreme model and the slice sampler algorithm, the relationship between the gold and American dollar returns in Malaysian market has been considered.

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk ijazah Doktor Falsafah

KAEDAH PENSAMPEL IRISAN DAN *METROPOLIS HASTINGS* BAGI ANALISIS BAYES UNTUK DATA EKSTRIM

Oleh

MOHAMMAD ROSTAMI

Februari 2016

Pengerusi: Mohd Bakri Adam, PhD
Fakulti : Institut Penyelidikan Matematik

Permodelan di hujung taburan sangat penting di kebanyakan bidang di mana risiko luarbiasa yang melibatkan kepentingan terhadap kejadian peristiwa kecil atau besar. Di dalam penyelidikan ini, applikasi teori nilai ekstrim dalam kerangka Bayes menggunakan algoritma *Metropolis Hastings* dan algoritma pensampel irisan sebagai kaedah alternatif, telah diperkenalkan dalam kes univariat ekstrim mudah kepada yang lebih kompleks iaitu kes bivariat ekstrim.

Pilihan taburan prior sangat kritikal dalam analisis Bayes. Di sini, kajian mendalam terhadap semua kemungkinan prior bagi parameter lokasi dan skala di jalankan dan beberapa cadangan pilihan taburan prior terhadap model Gumbel telah diberikan.

Algoritma pensampel irisan boleh menghindari daripada mensampel taburan tidak piawai dengain lebih berkesan. Tambahan lagi, kaedah pensampel irisan boleh menyesuaikan perubahan terhadap perubahan parameter skala yang dibuat, di mana kaedah pensampel irisan lebih mudah ditala berbanding algoritma *Metropolis Hastings*. Kebaikan utama lain bagi algoritma pensampel irisan ialah ianya memberikan posterior purata dengan ralat lebih rendah bagi parameter bentuk bagi model maksimum bulanan dan batas kekangan. Algoritma pensampel irisan ini telah juga digunakan terhadap model nilai bivariat nilai ekstrim dengan struktur logistik bersandar dan juga margin eksponen. Kajian simulasi menunjukkan algoritma pensampel irisan memberikan posterior purata dengan ralat yang rendah untuk semua parameters dengan aras pegun yang tinggi dalam siri iterasi. Selanjutnya, agoritma pensampel irisan in telah berjaya diaplikasikan terhadap data pulangan harga emas di Malaysia dan juga perhubungan di antara harga emas dengan pulangan nilai dolar Amerika di pasaran Malaysia.

ACKNOWLEDGEMENTS

I would like to show my deepest gratitude to Dr Mohd Bakri Adam, Chairman of my supervisory committee for his advice, generous help, expertise, and patience throughout the preparation of this thesis. I am indebted to my supervisory committee members, Professor Noor Akma Ibrahim and Dr Mohamed Hisham Yahya for their invaluable assistance and constructive suggestions.

Words are not enough to express my gratitude to my parents for their love and support throughout my life. The thesis would not be possible without their understanding and encouragement. Thank you for always being there for me. My brothers and sisters deserve my wholehearted thanks as well. Their love gave me forces to make this work.

Many thanks also go to all my friends, who share a lot of happiness and dreams with me.

I certify that a Thesis Examination Committee has met on 2 February 2016 to conduct the final examination of Mohammad Rostami on his thesis entitled "Slice Sampler and Metropolis Hastings Approaches for Bayesian Analysis of Extreme Data" in accordance with the Universities and University Colleges Act 1971 and the Constitution of the Universiti Putra Malaysia [P.U.(A) 106] 15 March 1998. The Committee recommends that the student be awarded the Doctor of Philosophy.

Members of the Thesis Examination Committee were as follows:

Leong Wah June, PhD
Associate Professor
Faculty of Science
Universiti Putra Malaysia
(Chairman)

Jayanthi a/p Arasan, PhD
Associate Professor
Faculty of Science
Universiti Putra Malaysia
(Internal Examiner)

Kamarulzaman Ibrahim, PhD
Professor
Universiti Kebangsaan Malaysia
Malaysia
(External Examiner)

David Nott, PhD
Associate Professor
National University of Singapore
Singapore
(External Examiner)

ZULKARNAIN ZAINAL, PhD
Professor and Deputy Dean
School of Graduate Studies
Universiti Putra Malaysia

Date: 24 March 2016

This thesis was submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfilment of the requirement for the degree of Doctor of Philosophy.

The members of the Supervisory Committee were as follows:

Mohd Bakri Adam, PhD
Associate Professor
Institute for Mathematical Research
Universiti Putra Malaysia
(Chairperson)

Noor Akma Ibrahim, PhD
Professor
Faculty of Science
Universiti Putra Malaysia
(Member)

Mohammad Hisham Yahya, PhD
Senior Lecturer
Faculty of Economics and Management
Universiti Putra Malaysia
(Member)

BUJANG KIM HUAT, PhD
Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date:

Declaration by graduate student

I hereby confirm that:

- this thesis is my original work;
- quotations, illustrations and citations have been duly referenced;
- this thesis has not been submitted previously or concurrently for any other degree at any other institutions;
- intellectual property from the thesis and copyright of thesis are fully-owned by Universiti Putra Malaysia, as according to the Universiti Putra Malaysia (Research) Rules 2012;
- written permission must be obtained from supervisor and the office of Deputy Vice-Chancellor (Research and Innovation) before thesis is published (in the form of written, printed or in electronic form) including books, journals, modules, proceedings, popular writings, seminar papers, manuscripts, posters, reports, lecture notes, learning modules or any other materials as stated in the Universiti Putra Malaysia (Research) Rules 2012;
- there is no plagiarism or data falsification/fabrication in the thesis, and scholarly integrity is upheld as according to the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) and the Universiti Putra Malaysia (Research) Rules 2012. The thesis has undergone plagiarism detection software.

Signature:__________________________ Date:__________________________

Name and Matric No: Mohammad Rostami, GS31629

Declaration by Members of Supervisory Committee

This is to confirm that:

- the research conducted and the writing of this thesis was under our supervision;
- supervision responsibilities as stated in the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) are adhered to.

Signature: ______________________________
Name of
Chairman of
Supervisory
Committee: Assoc. Prof. Dr. Mohd Bakri Adam

Signature: ______________________________
Name of
Member of
Supervisory
Committee: Prof. Dr. Noor Akma Ibrahim

Signature: ______________________________
Name of
Member of
Supervisory
Committee: Dr. Mohammad Hisham Yahya

TABLE OF CONTENTS

	Page
ABSTRACT	i
ABSTRAK	ii
ACKNOWLEDGEMENTS	iii
APPROVAL	iv
LIST OF TABLES	xii
LIST OF FIGURES	xiv
LIST OF ABBREVIATIONS	xvi

CHAPTER

1 **INTRODUCTION**	1
1.1 Introduction	1
1.2 Extreme Value Modelling	1
1.2.1 Generalized Extreme Value Distribution	1
1.2.2 Generalized Pareto Distribution	3
1.3 Relationship Between the GEV and GPD	3
1.4 Bayesian Theory	5
1.5 Prior Distribution	6
1.5.1 Informative Priors	6
1.5.2 Weakly Informative Priors	7
1.5.3 Least Informative Priors	7
1.5.4 Uninformative Priors	8
1.6 Credible Interval	8
1.7 Markov Chain Monte Carlo	9
1.7.1 Monte Carlo Sampling	9
1.7.2 Metropolis Hastings Algorithm and Slice Sampler Technique	11
1.8 Problem Statement	11
1.9 Data	13
1.9.1 Simulation Using Inversion Technique	13
1.9.2 Real Data Collection	14
1.10 Scope of the Report	15
1.11 Research Objectives	16
1.12 Summary	17
2 **LITERATURE REVIEW**	18
2.1 Extreme Value Theory	18
2.2 Bayesian Analysis of Extremes	19
2.3 Slice Sampler	20
2.4 Summary	24

3 **PRIOR SELECTION OF THE GEV$_{\xi=0}$ USING METROPOLIS HASTINGS ALGORITHM** 25
3.1 Introduction 25
3.2 Likelihood Function of the GEV$_{\xi=0}$ 25
3.3 Prior Distribution 26
3.3.1 Normal and Gamma 26
3.4 Simulation Study 33
3.4.1 Posterior mean (M_p), posterior standard deviation (SD_p), posterior standard error (SE_p) and the 95% quantile based credible intervals of μ and σ by defining the various pairs of priors on the location and scale parameters of $G(10,5)$ 33
3.4.2 Posterior mean (M_p), posterior standard deviation (SD_p), posterior standard error (SE_p) and the 95% quantile based credible intervals of μ and σ by specifying the various pairs of priors on the location and scale parameters of $G(10,10)$ 38
3.4.3 Posterior mean (M_p), posterior standard deviation (SD_p), posterior standard error (SE_p) and the 95% quantile based credible intervals of μ and σ by defining the various pairs of priors on the location and scale parameters of $G(100,5)$ 42
3.4.4 Posterior mean (M_p), posterior standard deviation (SD_p), posterior standard error (SE_p) and the 95% quantile based credible intervals of μ and σ by defining the various pairs of priors on the location and scale parameters of $G(100,10)$ 46
3.5 Summary 50

4 **SLICE SAMPLER ALGORITHM FOR GENERALIZED EXTREME VALUE DISTRIBUTION** 51
4.1 Introduction 51
4.2 Generalized Extreme Value Distribution 51
4.3 Bayesian Inference of the GEV Using Slice Sampler Algorithm 52
4.3.1 Likelihood Function of the GEV 52
4.3.2 Prior Distribution 52
4.3.3 Posterior Distribution 53
4.4 Slice Sampler 53
4.5 Simulation Study 56
4.5.1 Posterior mean (M_p), posterior standard deviation (SD_p), posterior standard error (SE_p) and the 95% quantile based credible intervals of μ, σ and ξ as location, scale and shape parameters of $GEV(10,5,0.3)$ for the slice sampler and Metropolis Hastings algorithms 56
4.5.2 Posterior mean (M_p), posterior standard deviation SD_p, posterior standard error (SE_p) and the 95% quantile based credible intervals of μ, σ and ξ as location, scale and shape parameters of $GEV(10,5,10^{-9})$ for the slice sampler and Metropolis Hastings algorithms 59

4.5.3 Posterior mean (M_p), posterior standard deviation (SD_p), posterior standard error (SE_p) and the 95% quantile based credible intervals of μ, σ and ξ as location, scale and shape parameters of GEV$(10, 5, -0.3)$ for the slice sampler and Metropolis Hastings algorithms 62
4.6 Real Data Analysis 64
4.6.1 The slice Sampler Algorithm For Malaysian Monthly Maxima Gold Returns 68
4.7 Summary 70

5 **SLICE SAMPLER ALGORITHM FOR GENERALIZED PARETO DISTRIBUTION** 71
5.1 Threshold Selection 71
5.1.1 Mean Residual Life Plot 72
5.2 Bayesian Inference of the GPD Using Slice Sampler Algorithm 73
5.2.1 Likelihood Function 73
5.2.2 Prior Distribution 73
5.2.3 Posterior Distribution 73
5.3 Simulation Study 74
5.3.1 Simulation Study I 74
5.3.2 Simulation Study II 79
5.4 Real Data Analysis 86
5.4.1 The Slice Sampler Algorithm For The Threshold Exceedances Of Malaysian Gold Return 89
5.5 Summary 91

6 **SLICE SAMPLER ALGORITHM FOR BIVARIATE EXTREME VALUE DISTRIBUTION** 92
6.1 Bivariate Extreme Value Distributions 92
6.1.1 Parametric Methods 94
6.2 Bayesian Analysis of the BEVD Using Slice Sampler Algorithm 95
6.2.1 Likelihood Function 95
6.2.2 Prior Distribution 96
6.2.3 Posterior Distribution 96
6.3 Simulation Study 96
6.3.1 Posterior mean (M_p), posterior standard error (SD_p), posterior standard error (SE_p) and the 95% quantile based credible intervals of μ_x, σ_x, ξ_x, μ_y, σ_y, ξ_y and ω as location, scale, shape and dependence parameters of BEVD$(\mu_x = 10,\ \sigma_x = 5,\ \xi_x = 0.1,\ \mu_y = 10,\ \sigma_y = 5,\ \xi_y = 0.1,\ \omega = 0.4)$ for the slice sampler algorithm 97

6.3.2 Posterior mean (M_p), posterior standard error (SE_p), posterior standard deviation (SD_p) and the 95% quantile based credible intervals of μ_x, σ_x, ξ_x, μ_y, σ_y, ξ_y and ω as location, scale, shape and dependence parameters of BEVD($\mu_x = 10,\ \sigma_x = 5,\ \xi_x = 0.4,\ \mu_y = 10,\ \sigma_y = 5,\ \xi_y = 0.4,\ \omega = 0.4$) for the slice sampler algorithm 100
6.3.3 Posterior mean (M_p), posterior standard error (SD_p), posterior standard error (SE_p) and the 95% quantile based credible intervals of μ_x, σ_x, ξ_x, μ_y, σ_y, ξ_y and ω as location, scale, shape and dependence parameters of BEVD($\mu_x = 10,\ \sigma_x = 5,\ \xi_x = 0.7,\ \mu_y = 10,\ \sigma_y = 5,\ \xi_y = 0.7,\ \omega = 0.4$) for the slice sampler algorithm 103
6.4 Real Data Analysis 106
6.5 Summary 109

7 **CONCLUSIONS AND RECOMMENDATIONS** 110
7.1 Conclusion 110
7.2 Suggestions for Future Research 111

REFERENCES 112
APPENDICES 123
BIODATA OF STUDENT 131
LIST OF PUBLICATIONS 132

LIST OF TABLES

Table		Page
3.1	Posterior densities for different pairs of priors defined on the location (μ) and scale (σ) parameters of the $\text{GEV}_{\xi=0}$	30
3.2	Marginal posterior densities for various pairs of priors specified on the location (μ) and scale (σ) parameters of the $\text{GEV}_{\xi=0}$	31
3.3	Acceptance probability for different pairs of priors defined on the location (μ) and scale (σ) parameters of the $\text{GEV}_{\xi=0}$	32
3.4	Results from Bayesian inference for single simulated dataset with sample size $n = 150$ and 5000 iterations from the $\text{GEV}_{\xi=0}$ model. The posterior mean, SD_p and SE_p of the Metropolis Hastings samples within the 95% quantile based credible interval for the different pair of priors are revealed.	35
3.5	Comparison of Bayesian inference results for the $\text{GEV}_{\xi=0}$ applied to various prior distributions (calculated using the Metropolis Hastings algorithm) with $n = 150$ observations and 5000 iterations. The posterior mean, SD_p and SE_p of the Metropolis Hastings samples within the the 95% quantile based credible interval are shown.	39
3.6	Results from Bayesian inference for single simulated dataset with sample size $n = 150$ and 5000 iterations from the $\text{GEV}_{\xi=0}$ model. The posterior mean, SD_p and SE_p of the Metropolis Hastings samples within the 95% quantile based credible interval for the different pair of priors are shown.	43
3.7	Comparison of Bayesian inference results for the $\text{GEV}_{\xi=0}$ applied to various prior distributions (calculated using the Metropolis Hastings algorithm) with $n = 150$ observations and 5000 iterations. The posterior mean, SD_p and SE_p of the Metropolis Hastings samples within the the 95% quantile based credible interval are demonstrated.	47
4.1	Results from the slice sampler and Metropolis Hastings algorithms for single simulated dataset with sample size $n = 150$ from the GEV posterior model. The posterior mean, SD_p and SE_p values along with the 95% quantile based credible interval across 5000 iterations.	59
4.2	Summary of properties of the slice sampler and Metropolis Hastings algorithms of the GEV posterior model parameters for single simulated dataset with sample size $n = 150$. The posterior mean, SD_p and SE_p values along with the 95% quantile based credible interval across 5000 iterations.	61
4.3	Results from the slice sampler and Metropolis Hastings algorithms for single simulated dataset with sample size $n = 150$ from the GEV posterior model. The posterior mean, SD_p and SE_p values along with the 95% quantile based credible interval across 5000 iterations.	63
4.4	Descriptive statistics of Malaysian daily and maxima monthly gold data from 2000 to 2011.	67

4.5 Results from the slice sampler algorithm for Malaysian monthly maxima gold returns by defining the normal, gamma and normal with large variances as prior distributions on the location (μ), scale (σ) and shape (ξ) parameters of the GEV posterior model with 5000 iterations. 69

5.1 Summary of properties of the slice sampler and Metropolis Hastings algorithms of the POT model parameters for single simulated dataset with sample size $n = 155$. The posterior mean, SD_p and SE_p values along with the 95% quantile based credible interval across the 5000 iterations. 77
5.2 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. 81
5.3 Results from the slice sampler and Metropolis Hastings algorithms for single simulated dataset with sample size $n = 150$ from the GPD posterior model. The posterior mean, SD_p and SE_p values along with the 95% quantile based credible interval across 5000 iterations. 83
5.4 Summary of properties of the slice sampler and Metropolis Hastings algorithms of the GPD posterior model parameters for single simulated dataset with sample size $n = 150$. The posterior mean, SD_p and SE_p values along with the 95% quantile based credible interval across 5000 iterations. Note: the true parameter values of σ and ξ are 5 and -0.3, respectively. 86
5.5 Descriptive statistics of threshold exceedances of Malaysian gold returns from 2000 to 2011 89
5.6 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. 91

6.1 Results from the slice sampler algorithm for bivariate simulated dataset with sample size $n = 150$ from the BEVD posterior model. The posterior mean and SE_p values along with the 95% quantile based credible interval across 5000 iterations. 98
6.2 Results from the slice sampler algorithm for bivariate simulated dataset with sample size $n = 150$ from the BEVD posterior model. The posterior mean and SE_p values along with the 95% quantile based credible interval across 5000 iterations. 101
6.3 Results from the slice sampler algorithm for bivariate simulated dataset with sample size $n = 150$ from the BEVD posterior model. The posterior mean and SE_p values along with the 95% quantile based credible interval across 5000 iterations. 104
6.4 Results from the slice sampler algorithm for the bivaiate model of Malaysian monthly maxima gold and USD returns with 5000 iterations. 107

LIST OF FIGURES

Figure		Page
3.1	Series of the location μ and scale σ parameters of the $\text{GEV}_{\xi=0}$ posterior model with sample size $n = 150$ and 5000 iterations.	36
3.2	Posterior distributions of μ and σ with sample size $n = 150$ and 5000 iterations.	37
3.3	Series of the location μ and scale σ parameters of the $\text{GEV}_{\xi=0}$ posterior model with sample size $n = 150$ and 5000 iterations.	40
3.4	Posterior distributions of μ and σ with sample size $n = 150$ and 5000 iterations.	41
3.5	Series of the location μ and scale σ parameters of the $\text{GEV}_{\xi=0}$ posterior model with sample size $n = 150$ and 5000 iterations.	44
3.6	Posterior distributions of μ and σ with sample size $n = 150$ and 5000 iterations.	45
3.7	Series of the location μ and scale σ parameters of the $\text{GEV}_{\xi=0}$ posterior model with sample size $n = 150$ and 5000 iterations.	48
3.8	Posterior distributions of μ and σ with sample size $n = 150$ and 5000 iterations.	49
4.1	Series of μ, σ and ξ as location, scale and shape parameters of the GEV posterior model with sample size $n = 150$ and 5000 iterations. In this simulation study, the true parameter values of μ, σ and ξ are 10, 5 and 0.3, respectively.	58
4.2	Posterior distributions of μ, σ and ξ with sample size $n = 150$ and 5000 iterations by defining normal, gamma and normal with large variances as prior distributions on the location, scale and shape parameters, respectively.	58
4.3	Series of μ, σ and ξ as location, scale and shape parameters of the GEV posterior model with sample size $n = 150$ and 5000 iterations.	60
4.4	Posterior densities of μ, σ and ξ with sample size $n = 150$ and 5000 iterations by defining normal, gamma and normal with large variances as prior distributions on the location, scale and shape parameters, respectively.	61
4.5	Series of the location μ, scale σ and shape ξ parameters of the GEV posterior model with sample size $n = 150$ and 5000 iterations. Note: the true parameter values of μ, σ and ξ are 10, 5 and -0.3, respectively.	63
4.6	Posterior distributions of μ, σ and ξ with sample size $n = 150$ and 5000 iterations by defining the normal, gamma and normal with large variances as prior distributions on the location, scale and shape parameters, respectively.	64
4.7	Empirical distribution of Malaysian gold monthly maximum return (1 month maximum). The red curve indicates the best fitted density and the green curve represents the best fitted normal distribution. This implies that there is a higher probability of extreme values than for a normal distribution.	66

4.8 Malaysian gold series from January 3, 2000 to December 19, 2011. **Top:** Daily gold price. **Middle:** Daily gold return. **Bottom:** Monthly maxima gold return. 66
4.9 Iteration series and posterior distributions of the location μ, scale σ and shape ξ parameters of the monthly maxima gold return. 69

5.1 MRL plot (with the 95% confidence intervals) of simulated data 76
5.2 Exceedence over a high threshold of the simulated data 76
5.3 Diagnostic plot of GPD fitting for the exceedence over a high threshold of the simulated data. They show that the peaks over the threshold value $u = 8$ is well approximated by the GPD. 77
5.4 Series of the scale σ and shape ξ parameters of the POT model with $n = 155$ sample size and 5000 iterations. 78
5.5 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. 78
5.6 Series of the scale σ and shape ξ parameters of the GPD posterior model with sample size $n = 150$ and 5000 iterations. 80
5.7 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. 81
5.8 Series of σ and ξ as scale and shape parameters of the GPD posterior model with sample size $n = 150$ and 5000 iterations. 83
5.9 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. 84
5.10 Series of the scale σ and shape parameters of the GPD posterior model with sample size $n = 150$ and 5000 iterations. 85
5.11 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. 86
5.12 MRL plot (with the 95% confidence intervals) of Malaysian gold returns 87
5.13 Exceedences over a high threshold of Malaysian gold return 88
5.14 Diagnostic plot of GPD fitting for the exceedences over a high threshold of Malaysian gold return. They show that the peaks over the threshold value $u = -0.07$ is well approximated by the GPD. 88
5.15 Iteration series and posterior distributions of the scale σ and shape ξ parameters of the threshold exceedances of Malaysian gold return with 5000 iterations. 90

6.1 The slice sampler series of μ_x, σ_x, ξ_x, ω, μ_y, σ_y and ξ_y as location, scale, shape and dependence parameters of the BEVD posterior model with sample size $n = 150$ and 5000 iterations. 98
6.2 Posterior distributions of μ_x, σ_x, ξ_x, ω, μ_y, σ_y and ξ_y with sample size $n = 150$ and 5000 iteration series by defining normal, gamma and normal with large variances as prior distributions on the location, scale and shape parameters, respectively and a uniform prior on the dependence parameter. 99

6.3 The slice sampler series of μ_x, σ_x, ξ_x, ω, μ_y, σ_y and ξ_y as location, scale, shape and dependence parameters of the BEVD posterior model with sample size $n = 150$ and 5000 iterations. 101
6.4 Posterior distributions of μ_x, σ_x, ξ_x, ω, μ_y, σ_y and ξ_y with sample size $n = 150$ and 5000 iteration series by specifying normal, gamma and normal with large variances as prior distributions on the location, scale and shape parameters, respectively and a uniform prior on the dependence parameter. 102
6.5 The slice sampler series of μ_x, σ_x, ξ_x, ω, μ_y, σ_y and ξ_y as location, scale, shape and dependence parameters of the BEVD posterior model with sample size $n = 150$ and 5000 iterations. 104
6.6 Posterior distributions of μ_x, σ_x, ξ_x, ω, μ_y, σ_y and ξ_y with sample size $n = 150$ and 5000 iteration series by defining normal, gamma and normal with large variances as prior distributions on the location, scale and shape parameters, respectively and a uniform prior on the dependence parameter. 105
6.7 The slice sampler series of the location (μ_x, μ_y), scale (σ_x, σ_y), shape (ξ_x, ξ_x) and dependence (ω) parameters of the gold and USD returns with 5000 iterations. 107
6.8 Posterior distributions of the bivariate monthly maxima model of gold and USD returns in the slice sampler algorithm. 108
6.9 Scatter Plot of the gold and American dollar returns in Malaysian market. 109

LIST OF ABBREVIATIONS

GEV	Generalized Extreme Value
GPD	Generalized Pareto Distribution
Ga	Gamma
MRL	Mean Residual Life
G	Gumbel
Lo	Logistic
W	Weibull
Exp	Exponential
IN	Inverse Normal
N	Normal
LN	Log Normal
R	Rayleigh
MCMC	Markov Chain Monte Carlo
MH	Metropolis Hastings
SS	Slice Sampler
EVT	Extreme Value Theory
BM	Block Maxima
POT	Peaks Over Threshold
SD	Standard Deviation
M_p	Posterior Mean
SE_p	Posterior Standard Error
CI	Credible Interval
LCI	Lower Bound of Credible Interval
UCI	Upper Bound of Credible Interval
BEVD	Bivariate Extreme Value Distribution

CHAPTER 1

INTRODUCTION

1.1 Introduction

During recent 50 years, extreme value theory (EVT) has been broadly applied in many areas of interest such as hydrology, financial studies and actuarial analysis (Coles (2001) and Tancredi et al. (2006)). Rare events can have catastrophic consequences for human activities, through their impact on the constructed and natural environments. The recent development of a sophisticated methodology for the prediction and estimation of functionals of rare events has contributed to saving endangered natural resources and to modelling earthquakes, climate and other environmental phenomena, like temperature, floods and precipitation, situations where we have to deal with large risks or with very low probabilities of overpassing a high (low) level (Gomes and Guillou, 2014). The distinguishing feature of an extreme value analysis is that it assesses a data generating processes of rare events- in other words, tail behaviour. The extrapolation of tail behaviour is accomplished by the asymptotic EVT. This theory supplies the asymptotic motivated approximate distributions in describing extremes, providing flexible and simple parametric models for fitting tail-related distributions. However, application of extreme value models is not always straightforward and there are some issues and difficulties in practice. Inherent sparsity of the extremal observations is a typical problem in modelling of extreme values that can reduce accuracy of the model (Coles and Powell, 1996).

1.2 Extreme Value Modelling

The field of EVT goes back to 1927, when Frechet (1927) formulated the functional equation of stability for maxima, which later was solved with some restrictions by Fisher and Tippett (1928) and finally by Gnedenko (1943) and De Haan (1970). 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 (Coles, 2001). 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) (Davison and Smith (1990) and Hosking and Wallis (1987)). Both extreme value distributions can be used to reliably estimate very high quantiles, permitting examination of the tail behaviour even out of the scope of observations (Danielsson and de Vries, 1997).

1.2.1 Generalized Extreme Value Distribution

Generalized extreme value (GEV) distribution is explained by the following theorem from Coles (2001).

Theorem 1.1 *Assume $X_1, \cdots, X_n$ be a sequence of random variables with common distribution function F, and let a_n and b_n be two sequences of constants. By considering $M_n = \max\{X_1, \cdots, X_n\}$:*

$$P_r\{(M_n - b_n)/a_n \leq z\} \rightarrow G(z) \quad as \quad n \rightarrow \infty.$$

If G be a non-degenerate distribution function, G belongs to the GEV family

$$G(z) = \exp\left\{-\left[1+\xi\left(\frac{z-\mu}{\sigma}\right)\right]^{-1/\xi}\right\},$$

for $\{z : 1+\xi(z-\mu)/\sigma > 0\}$, where $\mu \in \mathbb{R}$, $\sigma > 0$ and $\xi \in \mathbb{R}$.

The GEV distribution has three parameters: the location μ, scale σ and shape ξ. The GEV model can represent three types of tail behaviour determined by the value of shape parameter ξ: Weibull type (upper bounded) tail ($\xi < 0$), Frechet type (slowly decaying) tail ($\xi > 0$) and GEV $_{\xi=0}$ or Gumbel type (exponentially decay, like upper tail of normal distribution) which is defined in the limit as the shape parameter $\xi \rightarrow 0$. The EVT proves that, if a limiting distribution for the maxima (minima) of a sequence of independent and identically distributed random variables exists, then it must fall into one of these three types.

1.2.1.1 GEV$_{\xi=0}$ Model

The GEV$_{\xi=0}$ model is named after Emil Julius Gumbel (1891-1966), based on his original papers describing the model (Gumbel (1935) and Gumbel et al. (1941)). The GEV$_{\xi=0}$ model is a particular case of the GEV distribution (also known as the Fisher-Tippett model). This model might be applied to represent the distribution of the maximum level of a river in a particular year if there was a list of maximum values for the past ten years. The potential applicability of the GEV$_{\xi=0}$ model to represent the distribution of maxima relates to EVT which indicates that it is likely to be useful if the distribution of the underlying sample data is of the normal or exponential type. In real application, therefore, the GEV$_{\xi=0}$ model is applied to analyze such variables as monthly and annual maximum daily values (Oosterbaan and Ritzema, 1994). The book by Kotz and Nadarajah (2000) which describes the GEV$_{\xi=0}$ model, presents some of its application areas in engineering include network engineering, flood frequency analysis, , offshore engineering, space engineering, nuclear engineering, structural engineering, wind engineering and software reliability engineering. See also Nadarajah (2006), Prescott and Walden (1980) and Cooray (2010) for some generalizations of the GEV$_{\xi=0}$ model.

Let $X_1, \cdots, X_n$ be a sequence of independent and identically distributed sample which follow GEV$_{\xi=0}$ distribution denoted as G(μ, σ). The cumulative function is

$$F(x \mid \mu, \sigma) = \exp\left\{-\exp\left[-\left(\frac{x-\mu}{\sigma}\right)\right]\right\} \qquad x \in \mathbb{R}, \quad \mu \in \mathbb{R}, \quad \sigma > 0, \qquad (1.1)$$

where μ and σ are the location and scale parameters, respectively. The probability density function is given by

$$f(x) = \frac{1}{\sigma} \exp\left[-\left(\frac{x-\mu}{\sigma}\right)\right] F(x \mid \mu, \sigma). \tag{1.2}$$

1.2.2 Generalized Pareto Distribution

Davison and Smith (1990) introduce an asymptotic extreme value model to represent distribution of the excess over (below) a given threshold u, called GPD which has advantage of containing more samples for estimation of the parameters compared with GEV distribution. Hence, GPD has potential power to reduce problem of the information wasting.

Let X be an independent and identically distributed random variable of a GPD model and indicates the excess above the selected threshold u. The distribution function of X is in form

$$F_X(x|u,\sigma,\xi) = \begin{cases} 1 - \left[1 + \xi\left(\frac{x-u}{\sigma}\right)\right]^{-1/\xi} & \xi \neq 0, \\ 1 - \exp\left(-\frac{x-u}{\sigma}\right) & \xi = 0, \end{cases} \tag{1.3}$$

where the probability density function is given by

$$f_X(x|u,\sigma,\xi) = \begin{cases} \frac{1}{\sigma}\left[1 + \xi\left(\frac{x-u}{\sigma}\right)\right]^{(1+1/\xi)} & \xi \neq 0, \\ \frac{1}{\sigma}\exp\left(-\frac{x-u}{\sigma}\right) & \xi = 0, \end{cases} \tag{1.4}$$

with

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

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.

1.3 Relationship Between the GEV and GPD

The following theorem (see Coles (2001)) gives the relationship between the GEV and GPD models.

Theorem 1.2 *By considering assumption of Theorem 1.1, for a large size of n and large enough value of u, the distribution function of* $(X - u)$*, conditional on* $X > u$*, is approximated by*

$$H(y) = 1 - \left(1 + \frac{\xi y}{\tilde{\sigma}}\right), \tag{1.5}$$

for $\{y : y > 0$ *and* $(1+\xi y/\tilde{\sigma}) > 0\}$, *where*

$$\tilde{\sigma} = \sigma + \xi(u-\mu). \tag{1.6}$$

Proof:
Let X as an arbitrary term in the X_i with the distribution function F. For a large size of n

$$F^n(z) \approx \exp\left\{-\left[1+\xi\left(\frac{z-\mu}{\sigma}\right)\right]^{-1/\xi}\right\},$$

for $\mu \in \mathbb{R},\ \sigma > 0$ and $\xi \in \mathbb{R}$. Consequently,

$$n\log F(z) \approx -\left[1+\xi\left(\frac{z-\mu}{\sigma}\right)\right]^{-1/\xi}, \tag{1.7}$$

by using a Taylor expansion

$$\log F(z) \approx -\{1-F(z)\},$$

substitution into (1.7), gives

$$1-F(u) \approx \frac{1}{n}\left[1+\xi\left(\frac{u-\mu}{\sigma}\right)\right]^{-1/\xi},$$

for large enough u.

In a similar way, for $y > 0$,

$$1-F(u+y) \approx \frac{1}{n}\left[1+\xi\left(\frac{u+y-\mu}{\sigma}\right)\right]^{-1/\xi}. \tag{1.8}$$

Therefore,

$$\begin{aligned} P_r\{X > u+y \mid X > u\} &\approx \frac{\frac{1}{n}\left[1+\xi\left(\frac{u+y-\mu}{\sigma}\right)\right]^{-1/\xi}}{\frac{1}{n}\left[1+\xi\left(\frac{u-\mu}{\sigma}\right)\right]^{-1/\xi}} \\ &= \left[1+\frac{1+\xi\left(\frac{u+y-\mu}{\sigma}\right)}{1+\xi\left(\frac{u-\mu}{\sigma}\right)}\right]^{-1/\xi} \qquad (1.9) \\ &= \left[1+\frac{\xi y}{\tilde{\sigma}}\right]^{-1/\xi}, \end{aligned}$$

where

$$\tilde{\sigma} = \sigma + \xi(u-\mu).$$

1.4 Bayesian Theory

Term of Bayesian theory is named after Thomas Bayes, who introduced a special case of what is now called as the Bayes' theorem or Bayes' rule, see (Bayes, 1764). Bayes (1764) presents idea of revising and updating parameters of a density function based on the new information. Practically, by using Bayes' theorem, we obtain a posterior distribution through combination of marginal distribution and conditional probability distributions.

Theorem 1.3 (Bayes' theorem) *Let A and B represent two events. The probability of A given B is in form:*

$$P_r(A \mid B) = \frac{P_r(B \mid A)P_r(A)}{P_r(B)},$$

where $P_r(A)$ represents the marginal probability or prior probability of A and $P_r(B \mid A)$ denotes the conditional probability of B given A or called likelihood of B. The $P_r(B)$ represents the marginal probability of B.

Bayes' theorem provides a procedure for calculation of the posterior distribution by combination of likelihood function and prior distribution as follows:

$$\begin{aligned} \pi(A \mid B) &= \frac{L(B \mid A) \times \pi(A)}{P_r(B)} \\ &\propto L(B \mid A) \times \pi(A), \end{aligned} \tag{1.10}$$

where $\pi(A \mid B)$ is the posterior distribution, $\pi(A)$ is the prior probability and $L(B \mid A)$ is the likelihood function. The likelihood function, likelihood, or $L(\mathbf{x} \mid \Theta)$, contains the available information provided by the sample. The likelihood function is defined as

$$L(\mathbf{x} \mid \Theta) = \prod_{i=1}^{n} f(\mathbf{x_i} \mid \Theta).$$

The data $\mathbf{x}$ affects the posterior distribution $f(\Theta \mid \mathbf{x})$ only through the likelihood function $L(\mathbf{x} \mid \Theta)$. In this way, Bayesian inference obeys the likelihood principle, which states that for a given sample of data, any two probability models $f(\mathbf{x} \mid \Theta)$ that have the same likelihood function yield the same inference for Θ, see Koch (1990).

Bayes' theorem has been broadly used in many disciplines such as operational research and economical analysis. Cyert and DeGroot (1987) develop Bayes' theorem for both continuous and discrete random variables. Cyert and DeGroot (1987) point out if θ has a prior distribution denoted as $\pi(\theta)$, for any $X = x$, the posterior distribution denoted as $\pi(\theta \mid x)$ is obtained through following formula derived from Bayes' theorem:

$$\begin{aligned} \pi(\theta \mid x) &= \frac{f(x \mid \theta)\ \pi(\theta)}{\int_{\theta} f(x \mid \theta)\ \pi(\theta)\ d\theta} \\ &\propto f(x \mid \theta) \times \pi(\theta). \end{aligned} \tag{1.11}$$

Equation (1.11) is widely applied in data analysis. Now, assume x denotes the data or given information which comes from a probability distribution $f(. \mid \theta)$. We would like to revise or update the probability distribution of θ given x. If we have information about the behaviour of the parameter and know the prior distribution, then the posterior distribution is obtained through Formula (1.11). Otherwise, when there is no information about the parameter, we can use uninformative prior such as conjugate and Jeffrey's priors, see Kass and Wasserman (1996).

1.5 Prior Distribution

In Bayesian inference, a prior probability distribution, often called simply the prior, of an uncertain parameter θ or latent variable is a probability distribution that expresses uncertainty about θ before the data are taken into account. The parameters of a prior distribution are called hyper parameters, to distinguish them from the parameters (Θ) of the model. When applying Bayes' theorem, the prior is multiplied by the likelihood function and then normalized to estimate the posterior probability distribution, which is the conditional distribution of Θ given the data. Moreover, the prior distribution affects the posterior distribution. Prior probability distributions have traditionally belonged to one of two categories: informative priors and uninformative priors. Here, four categories of priors are presented according to information and the goal in the use of the prior. The four categories are informative, weakly informative, least informative, and uninformative.

1.5.1 Informative Priors

When prior information is available about θ, it should be included in the prior distribution of θ. For example, if the present model form is similar to a previous model form, and the present model is intended to be an updated version based on more current data, then the posterior distribution of θ from the previous model may be used as the prior distribution of θ for the present model. In this way, each version of a model is not starting from scratch, based only on the present data, but the cumulative effects of all data, past and present, can be taken into account. To ensure the current data do not overwhelm the prior, Ibrahim and Chen (2000) introduce the power prior. The power prior is a class of informative prior distribution that takes previous data and results into account. If the present data is very similar to the previous data, then the precision of the posterior distribution increases when including more and more information from previous models. If the present data differs considerably, then the posterior distribution of θ may be in the tails of the prior distribution for θ, so the prior distribution contributes less density in its tails. Hierarchical Bayes is also a popular way to combine data sets. Sometimes informative prior information is not simply ready to be used, such as when it resides in another person, as in an expert. In this case, their personal beliefs about the probability of the event must be elicited into the form of a proper probability density function. This process is called prior elicitation.

1.5.2 Weakly Informative Priors

Weakly informative prior (WIP) distributions use prior information for regularization and stabilization, providing enough prior information to prevent results that contradict our knowledge or problems such as an algorithmic failure to explore the state-space. Another goal is for WIPs to use less prior information than is actually available. A WIP should provide some of the benefit of prior information while avoiding some of the risk from using information that doesn't exist. WIPs are the most common priors in practice, and are favoured by subjective Bayesians. Selecting a WIP can be tricky. WIP distributions should change with the sample size, because a model should have enough prior information to learn from the data, but the prior information must also be weak enough to learn from the data. After updating a model in which WIPs exist, the user should examine the posterior to see if the posterior contradicts knowledge. If the posterior contradicts knowledge, then the WIP must be revised by including information that will make the posterior consistent with knowledge, see Gelman et al. (2014).

1.5.2.1 Vague Priors

The first formal move from vague to weakly informative priors is Lambert et al. (2005). Typically a vague prior, also called a diffuse prior, is a conjugate prior with a large scale parameter. After conjugate priors were introduced (Raiffa, 1974), most applied Bayesian modelling has used vague priors, parametrized to approximate the concept of uninformative priors (better considered as least informative priors). Often researchers want the data to dominate when there is no prior information and thus attempt to use vague prior distributions (Lambert et al., 2005). A vague prior is defined here as usually being a conjugate prior that is intended to approximate an uninformative prior (or actually, a least informative prior).

1.5.3 Least Informative Priors

The term 'Least Informative Priors', or LIPs, is used here to describe a class of prior in which the goal is to minimize the amount of subjective information content, and to use a prior that is determined solely by the model and observed data.

1.5.3.1 Flat Priors

The flat prior was historically the first attempt at an uninformative prior. The unbounded, uniform distribution, often called flat prior, is

$$\theta \sim u(-\infty, +\infty),$$

where θ is uniformly-distributed from negative infinity to positive infinity. Although this seems to allow the posterior distribution to be affected solely by the data with no impact from prior information, this should generally be avoided because this probability distribution is improper, meaning it will not integrate to one since the integral of the assumed $p(\theta)$ is infinity (which violates the assumption that the probabilities sum to one). This may cause the posterior to be improper, which invalidates the model. Thomas Bayes $(1701-1761)$ was the first to use inverse probability and Bayes and Price (1763) use flat prior for his billiard example so that all possible values of θ are equally likely a priori (Gelman et al., 2014). Pierre-Simon Laplace $(1749-1827)$ also use the flat prior to estimate the proportion of female births in a population, and for all estimation problems present or justify as a reasonable expression of ignorance. Laplace's use of this prior distribution was later referred to as the 'principle of indifference' or 'principle of insufficient reason', and is now called the at prior Gelman et al. (2014). Laplace was aware that it was not truly uninformative, and used it as a LIP. Another problem with the flat prior is that it is not invariant to transformation. For example, a flat prior on a standard deviation parameter is not also at for its variance or precision.

1.5.4 Uninformative Priors

Traditionally, most of the above descriptions of prior distributions were categorized as uninformative priors. However, uninformative priors do not truly exist and all priors are informative in some way (Irony and Singpurwalla, 1997). Traditionally, there have been many names associated with uninformative priors, including diffuse, minimal, non-informative, objective, reference, uniform, vague, and perhaps weakly informative.

Prior Distributions For This Study

When analysing data from a Bayesian perspective it is necessary to specify prior distributions for all unknown parameters. This can be a potential advantage, but in many situations there is a desire for the data to dominate when no prior information is available (or when MCMC methods are being used for computational convenience and the researcher does not want to include prior information), which has led to the use of vague or reference priors (Kass and Wasserman, 1996). In this study, we use some informative priors for the location and scale parameters of the $\text{GEV}_{\xi=0}$ model. Here, Metropolis Hastings algorithm initiates the parameter values from the parameter space. Multiple starting value under the support of parameter values are used in testing the convergency of the chain and the sensitivity to the initial value (Gelman et al., 2003). The simulation results show the chain can converge (in probability) to the true value quickly even with the initial value which is relatively far from the true parameter value. In contrast, we define vague priors on the location, scale and shape parameters of the GEV, GPD and bivariate extreme value distribution (BEVD) along with a flat prior defined on the dependence parameter of the BEVD, to show our little prior information allowing the data to speak for themselves which demonstrates the worst case for estimation performance (Coles and Tawn (1996), Coles et al. (2003), Coles and Tawn (2005), Reis Jr and Stedinger (2005) and Yoon et al. (2010)).

1.6 Credible Interval

In Bayesian inference, a credible interval is a probabilistic region around a posterior moment, and is similar in use to a frequentest confidence interval. In fact, the idea of a credibility interval is to give an analogue of a confidence interval in classical statistics. The reasoning is that point estimates give no measure of accuracy, so it is preferable to give a interval within which is likely that the parameter lies. In frequentist approach, confidence intervals have the interpretation that if the sampling were repeated, there is a specified probability that the interval so obtained would contain the parameter- it is the interval which is random and not the parameter. There is no such difficulty in the Bayesian approach because parameters are treated as random. A Bayesian analysis can provide credible intervals for parameters or any function of the parameters which are more easily interpreted than the concept of confidence interval in classical statistics (Congdon, 2007). A credible interval, incorporates information from the prior distribution into the estimate, while confidence intervals are based solely on the data. An interval C is a $100(1-\alpha)\%$ credible interval for θ if

$$\int_C \pi(\theta \mid \mathbf{X})d\theta = 1-\alpha. \tag{1.12}$$

That is, there is a probability of $1-\alpha$, based on the posterior distribution, that θ lies in C. Note that, if the parameter space is discrete, a sum replaces the integral in equation (1.12). There are a variety of Bayesian probability intervals. For example, when generalized to multivariate forms, it is called a probability region (or credible region). Aside from whether it is univariate or multivariate, there are quantile based probability intervals, uni-modal Highest Posterior Density (HPD) intervals, multimodal HPD intervals, and the Lowest Posterior Loss (LPL) interval, among others.

In this study, we use quantile based credible interval to obtain a probabilistic region around the posterior means. If θ_L^* be the $\alpha/2$ posterior quantile for θ, and θ_U^* be the $1-\alpha/2$ posterior quantile for θ, then (θ_L^*, θ_U^*) is a $100(1-\alpha)\%$ credible interval for θ.

$$\begin{aligned} P\{\theta \in (\theta_L^*, \theta_U^*) \mid \mathbf{X}\} &= 1 - P\{\theta \notin (\theta_L^*, \theta_U^*) \mid \mathbf{X}\} \\ &= 1 - [P(\theta < \theta_L^* \mid \mathbf{X}) + P(\theta > \theta_U^* \mid \mathbf{X})] \\ &= 1-\alpha. \end{aligned}$$

1.7 Markov Chain Monte Carlo

Practically, we sometimes face some obstacles in analytical calculation of the posterior distributions. One effective method to overcome this problem is using a Monte Carlo simulation technique. By performing this method, we can estimate some features of the posterior distribution, such as marginal distribution, covariance and posterior mean.

1.7.1 Monte Carlo Sampling

Suppose we need to calculate an expectation that cannot be expressed in a closed form. An acceptable approach would be to use random sampling to evaluate the analytical or numerical integration. For example, we can collect a large sample from a population and consider the corresponding sample mean as an approximation of the population mean. Based on the law of large numbers, we know that when the sample size is large enough, the estimate may be acceptably accurate. So if we want to estimate a posterior mean, we attempt to generate independent and identically distributed observations from the posterior distribution and consider the sample mean as an approximation of the posterior mean. However, one of the usual drawbacks of this approach is that often we encounter posterior distributions which are non-standard and are difficult to sample from. In such event, the notion of importance sampling introduced by Metropolis et al. (1953), provides a new algorithm for sampling points from a given probability density function. When the original density function is difficult or computationally burdensome to sample from directly, the importance sampling approach suggests finding a probability density function that is very close to the original density function and easier to sample from. In other words, importance sampling involves choosing a "good" distribution from which to simulate the random variables of interest. As a result, it yields the expectation of a quantity that varies less than the original integral over the region of integration. This approach is one of the basic Monte Carlo sampling methodologies. Since it is believed that certain values of an input random variable in a simulation have more impact on the estimation than others, and by emphasizing these "important" values more frequently, the variance of the estimator can be reduced. Hence, importance sampling can also be viewed as a variance reduction technique.

One drawback of the traditional Monte Carlo sampling or Monte Carlo importance sampling is that the functional form of the posterior density function needs to be specified. Otherwise, this sampling technique would be difficult to implement for cases where the posterior distributions are handled indirectly or incompletely. Such cases are not rare, especially for Bayesian hierarchical models that involve the joint posterior distribution of the parameter set specified in combination of conditional and marginal distributions. This is due to the fact that while the joint posterior distribution is difficult to specify directly, the conditional posteriors, given the relevant parameter values at different hierarchy levels, are easier to derive.

Usually, the generation of random vectors is not an easy task. Devroye (1986) argues that most of the rejection based algorithms are practically limited to relatively small dimensions (up to at most 10). Also, there exist a large number of distributions which are difficult to sample from in even as few as three or four dimensions. Under these circumstances, constructing a Markov chain, which has the desired fixed multivariate distribution as its unique stationary distribution, is a better choice. The basic idea of this method is to use a Markov chain to simulate from random vectors. This technique has attracted considerable research attention recently.

To overcome the problem of evaluating multidimensional integrals, necessary in Bayesian statistics, Markov chain Monte Carlo (MCMC) techniques have been used frequently over the last three decades. The hierarchical prior structure in Bayesian computations can lead to analytically tractable conditional posteriors, which makes it possible to adopt the MCMC procedure for obtaining random draws from the target joint posterior distribution. Ghosh et al. (2007) state that these iterative Monte Carlo procedures typically generate a random sequence with the Markov property such that this Markov chain is ergodic with the limiting distribution being the target posterior distribution.

1.7.2 Metropolis Hastings Algorithm and Slice Sampler Technique

Among the Monte Carlo sampling techniques, the well-known Markov chain Monte Carlo (MCMC) method is widely used for handling complex computational problems, especially in multivariate cases. Given an initial vector $X_0 \in \mathbb{R}^d$ and a conditional distribution $K(X_{t+1} \mid X_t)$ that depends only on the current state vector X_t, we can generate a sequence of random vectors $X_0, X_1, \cdots, X_n$ which form a Markov chain with the transition kernel $K(X_{t+1} \mid X_t)$. When the transition kernel does not depend on t, this Markov chain is called a time homogeneous Markov chain. One question that arises here involves the effect of the starting vector X_0 on the distribution of X_t, denoted by $K^t(X_t \mid X_0)$. In the Markov chain sampling process, we require $K_t(X_t \mid X_0)$ to converge to a unique stationary distribution, which neither depends on t nor on X_0. These Markov chain based generators produce dependent random variables and the first vectors of these sequences do not replicate the target distribution and, thus, need to be discarded. This is called the ”burn-in” phase of such a Markov chain sampling procedure. The length of the burn-in period should be long enough to guarantee the convergence of the Markov chain. Theoretically speaking, when applying Monte Carlo sampling based approaches, it is necessary to wait until the Markov chain converges to the invariant target distribution, and then sampling from the resulting distribution. It is a good idea to start a large number of chains beginning with different starting points, and pick the draws after allowing these chains to run over a sufficiently long period of time. In other words, it may be necessary to use many different chains to ensure that convergence occurs and we need to discard the samples obtained during the burn-in phase. Nevertheless, the law of large numbers for dependent chains implies that one could just use a single Markov chain as long as this chain is ”long” enough.

1.8 Problem Statement

In recent years, there have been increasing number of researches that propose adopting the Bayesian approach by application of Monte Carlo simulation methods to achieve optimal making of decisions. For instance, Jensen (2001) introduces an estimator based on Markov chain Monte Carlo (MCMC) technique for a long-memory stochastic volatility in a financial model.

EVT based models are based upon an asymptotic approximation for the tail distributions, which are very flexible in terms of the allowable tail shape behaviour. The attraction of the EVT based techniques is that they can provide statistically and mathematically justifiable parametric model for the tails of distribution which can give reliable extrapolations beyond the range of the observed data. The typical problem with fitting extreme value models is the inherent lack of observations (Coles and Powell, 1996). Moreover, a large sample size is required to estimate parameter of the shape (ξ), see Coles (2001). Therefore in EVT, other information of the distribution can be useful if the estimation method can take account these information such as expert knowledge. Hence, MCMC techniques have been frequently used in the EVT. For example, Coles and Powell (1996) apply a Bayesian framework for the GEV model in three different cases by defining various prior distributions including gamma as informative prior, a flat distribution as non informative prior and multivariate normal as empirical prior. Coles and Powell (1996) use Gibbs sampler algorithm to estimate the posterior distributions. According to the findings, in the situation of both informative and empirical priors, the estimations are at least as trustworthy as MLE, but remarkably more informative and flexible.

A key issue in the Bayesian analysis is subjectivity in choosing the prior distribution. In fact, there is no standard procedure for specifying the prior distribution on the parameters and it can be completely subjective. According to Kass and Wasserman (1996) and Skold (2005), prior selection can be different from researcher to researcher and this result in subjectivity in choice of prior distribution for a parameter. Consequently, one of the greatest challenges of applying Bayesian approach for EVT is that such subjectivity in prior selection necessarily leads to some errors in the final conclusion. Another obstacle for Bayesian analysis of extreme data is related to the computation of the posterior distribution. Since in most cases, posterior distributions are high dimensional and complex, it is difficult to calculate the parameter estimation through analytical methods. In such situations, MCMC technique provides some procedures to overcome this problem. Thus, there are high level of interest to use this technique in Bayesian computation by researchers. For example, Bray (2002) presents a usage of MCMC techniques for projecting mortality and cancer incidence. Also, San Martini et al. (2006) apply MCMC techniques to inference aerosol observations and forecast gas phase concentrations. Monte Carlo sampling techniques contain the well-known Gibbs sampler introduced by Geman and Geman (1984) and Gelfand and Smith (1990), Metropolis-Hastings algorithm proposed by Metropolis et al. (1953) and generalized by Hastings (1970), and hit-and-run sampler, for which a detailed discussion may be obtained in Smith (1984), Bélisle et al. (1993), Schmeiser and Chen (1991) and Chen and Schmeiser (1993).

However, in analysis of the features of extreme value posterior distributions, Gibbs sampler and Metropolis Hastings algorithms are popular, it is often problematic for researchers to gain satisfactory simulation results. Metropolis Hastings algorithm requires a good candidate density functions for efficient sampling. But, sometimes, it is difficult to find a proper proposal density function even in the one dimensional states. Further, in Gibbs sampler somebody may face problem to sample from a complex non-standard conditional distribution. Therefore, it seems that researchers need to introduce a more efficient algorithm for collecting sample from non-standard extreme value posterior distributions.

Despite many studies conducted on Bayesian analysis of extreme data using MCMC methods such as Metropolis Hastings and Gibbs sampler algorithms, the Bayesian analysis of extreme data using the slice sampler approach has not been undertaken by the time of this study. Therefore, the current research attempts to fill the gap in the body of literature on Bayesian analysis of extreme data. The objectives of this research are to consider the prior selection and develop the slice sampler algorithm for extreme value models.

1.9 Data

Two data sets are used in the current study, including simulated data and real data. The simulated data are generated by using the inversion technique for various extreme models such as GEV, $\text{GEV}_{\xi=0}$, GPD and bivariate extreme value. For real data analysis, we adopt Malaysian daily gold returns from January 3, 2000 to December 19, 2011, taking a total of 3120 days.

1.9.1 Simulation Using Inversion Technique

The inversion method transforms a single uniform random variable into the random variable. This method is based on (inverses of) cumulative distribution functions F. Inversion method is the only truly universal method: If all we can do is compute distribution function $F(x)$ for all x, and we have enough (*i.e.*, infinite) time on our hands, then we can generate random variate with distribution function F (Devroye, 1986).

Let $F(x)$, $x \in \mathbb{R}$, denote any cumulative distribution function (cdf) (continuous or not). Recall that $F : \mathbb{R} \to [0,1]$ is thus a non-negative and non-decreasing (monotone) function that is continuous from the right and has left hand limits, with values in $[0,1]$; moreover $F(\infty) = 1$ and $F(-\infty) = 0$. Our objective is to generate (simulate) random variable X distributed as F; that is, we want to simulate a random variable X such that $P(X \leq x) = F(x)$, $x \in \mathbb{R}$. Define the generalized inverse of F, $F^{-1} : [0,1] \to \mathbb{R}$, via

$$F^{-1}(y) = \min\{x : F(x) \geq y\}, \quad y \in [0,1]. \tag{1.13}$$

If F is continuous, then F is invertible (since it is thus continuous and strictly increasing) in which case $F^{-1}(y) = \min\{x : F(x) \geq y\}$, the ordinary inverse function and thus $F(F^{-1}(y)) = y$ and $F^{-1}(F(x)) = x$. In general it should that $F^{-1}(F(x)) \leq x$ and $F(F^{-1}(y)) \geq y$. $F^{-1}(y)1$ is a non-decreasing (monotone) function in y. This simple fact yields a simple method for simulating a random variable X distributed as F:

Theorem 1.4 *Let $F(x)$, $x \in \mathbb{R}$, denote any cumulative distribution function (cdf) (continuous or not). Let $F^{-1}(y)$, $y \in [0,1]$ denote the inverse function defined in (1.13). Define $X = F^{-1}(U)$, where U has the continuous uniform distribution over the interval* $(0,1)$. *Then X is distributed as F, that is, $P(X \leq x) = F(x), x \in \mathbb{R}$.*

Proof:
We must show that $P(F^{-1}(U) \leq x) = F(x)$, $x \in \mathbb{R}$. First suppose that F is continuous. Then we will show that (equality of events) $\{F^{-1}(U) \leq x\} = \{U \leq F(x)\}$, so that by taking probabilities (and letting $a = F(x)$ in $P(U \leq a) = a$) yields the result: $P(F^{-1}(U) \leq x) = P(U \leq F(x)) = F(x)$. To this end: $F(F^{-1}(y)) = y$ and so (by monotonicity of F) if $F^{-1}(U) \leq x$, then $F(F^{-1}(U)) \leq F(x)$. Similarly $F^{-1}(F(x)) = x$ and so if $U \leq F(x)$, then $F^{-1}(U) \leq x$. We conclude equality of the two events as was to be shown (Devroye, 1986).

1.9.2 Real Data Collection

Data collection is referred to the sources that are used in the study. There are mainly two forms of data collections, primary and secondary data. Primary data is data that has been observed, experienced or recorded directly from immediate experience and that has not been analysed. Secondary data, on the other hand, is gathered data established by a third party (Walliman, 2010). Saunders et al. (2011) claim that secondary data is not always intentionally published for the research subject. Journals, newspapers, policy documents, annual reports and all other type of data that already exist are some example of secondary data

The documents and information that are gathered for the current study are mainly based on secondary data. The secondary data that were are in this study is generated specifically from daily reports of Malaysian gold returns from January 3, 2000 to December 19, 2011, which have been downloaded from www.kitco.com. Over the past two decades, many countries have been turning their attention to the issue of economic stability more and more because of the occurrence of several economic crises repeatedly, such as the Asian financial crisis in 1970, the sub-prime mortgage crisis in the United States in 2007, the European sovereign-debt crisis in 2009, and, lastly, the Cyprus crisis. These crises introduced adverse economic impacts that spread widely from country to country and, subsequently, reached global levels. All nations around the world have suffered from this economic impact in one way or the other, in terms of international trade, foreign investment, international financial market, foreign capital movement, stock index, and foreign exchange rate market. These impacts affected each and every country's economic goals, and, ultimately, it is bound to influence the Gross Domestic Production (GDP) growth. For centuries, investors have been found to protect their capital by investing in assets that offer safer stores of value (World Gold Council, 2008).

Gold is widely regarded as representative of precious metals. In the history of international currencies, gold has not only contributed significantly to the stabilization of the international money market, but has also served as an important financial asset in international currency reserves. For instance, most countries around the world hold a certain proportion of their foreign exchange reserves in gold. Apart from its superior industry characteristics, gold has also served as a medium of exchange for several thousand years. According to a report prepared by the World Gold Council, there are around $166,000$ tons of gold in the world, and the growth rate of the stock of gold on earth increases by only 2% per year. In 2011, the global demand for gold was 4067.1 tonnes, which was 6.5% higher than the previous years' level of 3818.2 tonnes. However, the total supply of gold was 3994 tonnes in 2011, which was 8% below the previous years' level of 4163.9 tonnes. When the smaller global gold reserves are in even shorter supply, the cost of gold production reduces the shortfall in supply increases, and vice versa. Since the global financial crisis began in late 2007, global investors have not aggressively sought complicated portfolios of financial assets, but have rather resorted to simpler financial and risk management. Investing in gold can provide a store of value and hedge risks (Kolluri (1981), Moore (1990), Dooley et al. (1995), Taylor (1998), Capie et al. (2005) and Hammoudeh et al. (2010)). Even though gold prices will change due to fluctuations in market prices, gold will not lose its value over time. In order to avoid currency depreciation and low interest rates, particularly during times of recession, investors prefer to provide a store of value for their wealth by holding gold (Koutsoyiannis (1983), Mishra and Rahman (2005) and Hammoudeh et al. (2011)). An increase in the demand for gold can increase gold prices, which can subsequently lead to a high value of the gold investment. Consequently, this can push up gold prices even further. When all factors are considered, there is still optimism that the gold market will remain bullish. In fact, the price and production behaviour of gold differs from most other mineral commodities. Gold is an important asset that provides stability to international money markets and international currency reserves (Chang et al., 2013). In the late 2007 financial crisis, the gold price increased by 6% while many key mineral prices fell and other equities dropped by around 40%.

Given the importance of gold price in the global market, for the application part of the current study, we attempt to conduct Bayesian analysis of Malaysian extreme gold returns by using the slice sampler algorithm. EVT has been shown to be a very useful tool in estimating and predicting the extremal behaviour of actuarial and financial products, such as predicting the largest claim in insurance and the Value-at-Risk (VaR), see Embrechts et al. (1997). Financial data is well known to exhibit relatively heavy (heavier than normal) tails, and typically show clusters of observations in the tails, often termed volatility clustering which creates challenges when applying extreme value models, since classical EVT assume independence of the underlying process. However, the assumption of independence can be easily dropped and the theoretical results follow through (McNeil, 1997). Moreover, under certain conditions, Beirlant et al. (2004) show that a stationary process with short range dependence between the observations can also lead to the same extreme distribution family.

1.10 Scope of the Report

This report is divided into 7 chapters. Chapter 1 looks at the general overview of the EVT and some issues that exist in Bayesian analysis of extreme values in terms of prior selection and computation of the posterior distribution. It also looks into some background studies of Markov chain Monte Carlo (MCMC) techniques and the objectives of this research.

Chapter 2 reviews some background studies and literatures in the fields of extreme value modelling, Bayesian analysis and slice sampler algorithm in order to support the current work.

Chapter 3 focuses on the prior selection of $\text{GEV}_{\xi=0}$ model using the Metropolis Hastings algorithm. The posterior distributions are computed by multiplying the likelihood function and various pairs of prior distributions. A simulation study is conducted to compare the performance of the priors defined on the location (μ) and scale (σ) parameters.

Chapter 4 concentrates on the development of slice sampler algorithm for block maxima (BM) approach and GEV distribution. Both the simulation study and real data application are used to show the performance of the slice sampler algorithm in the computation of the GEV posterior distribution and to model the tail behaviour of Malaysian monthly maximum gold returns.

Chapter 5 focuses on the development of 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 the BM approach as used for the GEV. A simulation study is carried out to demonstrate the performance of the slice sampler algorithm for fitting the POT and GPD posterior models. Moreover, the slice sampler technique is employed to estimate the return and risk values of investment in Malaysian gold market.

Chapter 6, concentrates on the development of slice sampler algorithm for the bivariate extreme value distribution (BEVD) with logistic dependence structure and exponential margins. This is an extension of the model discussed in chapter 4. By conducting a simulation study, the performance of the slice sampler algorithm in the computation of the posterior distributions is considered. Further, by using the slice sampler algorithm, the relationship between returns of the gold and united states dollar in Malaysian market is examined.

Finally, chapter 7 offers some concluding remarks and suggestions for future work.

1.11 Research Objectives

It is clear that to date the slice sampler algorithm for Bayesian analysis of extreme data have not been developed. More specifically, some studies have focused on the Bayesian analysis of extreme values by using other MCMC algorithm such as Metropolis Hastings and Gibbs sampler. Thus, in this study, the slice sampler algorithm contributes to Bayesian analysis of GEV, GPD and bivariate extreme value distribution (BEVD). Moreover, since there is no standard approach for defining the prior distribution on the parameters of extreme value models and it should vary according to application, this research also provides an analysis of prior selection of the $GEV_{\xi=0}$. In addition, this study attempts to perform Bayesian analysis of Malaysian extreme gold returns using the slice sampler algorithm. Briefly, this research embark on the following objectives:

1. Analysis for prior selection for the $GEV_{\xi=0}$ model by using Metropolis-Hastings algorithm.
2. To develop the slice sampler algorithm for the GEV model.
3. To develop the slice sampler algorithm for the GPD model.
4. To develop the slice sampler algorithm for the bivariate extreme with logistic dependence structure.
5. Applying the slice sampler algorithm for Malaysian extreme gold returns from 2000 to 2011.

1.12 Summary

The EVT makes it possible to model those phenomena which happen rarely such as earthquake, flood, financial risk and insurance losses. The typical problem in extreme value modelling is the inherent sparsity of extreme observations, for which Bayesian framework supplies a procedure to take advantage of prior information such as expert knowledge. In this study, we will develop the slice sampler algorithm for extreme value models such as GEV, GPD and bivariate extremes. For application part, we will concentrate on the Bayesian inference of Malaysian gold returns using the slice sampler algorithm.

REFERENCES

Achcar, J., Bolfarine, H., and Pericchi, L. R. (1987). Transformation of survival data to an extreme value distribution. *The Statistician*, pages 229–234.

Agarwal, D. K. and Gelfand, A. E. (2005). Slice sampling for simulation based fitting of spatial data models. *Statistics and Computing*, 15(1):61–69.

Albeverio, S., Jentsch, V., and Kantz, H. (2006). *Extreme events in nature and society*. Springer-Verlag Berlin Heidelberg.

Altmann, E. G., Hallerberg, S., and Kantz, H. (2006). Reactions to extreme events: Moving threshold model. *Physica A: Statistical Mechanics and its Applications*, 364:435–444.

Balakrishnan, N. and Lai, C. D. (2009). *Continuous bivariate distributions*. Springer Science & Business Media, New York.

Balkema, A. A. and de Haan, L. (1974). Residual life time at great age. *The Annals of Probability*, 2(5):792–804.

Bayes (1764). An essay toward solving a problem in the doctrine of chances. *Philosophical Transactions of the Royal Society of London: 53*, pages 370–418.

Bayes, M. and Price, M. (1763). An essay towards solving a problem in the doctrine of chances. by the late rev. mr. bayes, frs communicated by mr. price, in a letter to john canton, amfrs. *Philosophical Transactions (1683-1775)*, pages 370–418.

Behrens, C. N., Lopes, H. F., and Gamerman, D. (2004). Bayesian analysis of extreme events with threshold estimation. *Statistical Modelling*, 4(3):227–244.

Beirlant, J., Goegebeur, Y., Teugels, J., and Segers, J. (2004). *Statistics of Extremes: Theory and Applications*. Wiley Online Library.

Bélisle, C. J., Romeijn, H. E., and Smith, R. L. (1993). Hit-and-run algorithms for generating multivariate distributions. *Mathematics of Operations Research*, 18(2):255–266.

Bermudez, P. Z. and Turkman, M. (2003). Bayesian approach to parameter estimation of the generalized pareto distribution. *TEST: An Official Journal of the Spanish Society of Statistics and Operations Research*, 12(1):259–277.

Bernstein, W. (2000). *The Intelligent Asset Allocator: How to Build Your Portfolio to Maximize Returns and Minimize Risk*. McGraw Hill Professional.

Besag, J. and Green, P. (1993). Spatial statistics and bayesian computation. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 25–37.

Bottolo, L., Consonni, G., Dellaportas, P., and Lijoi, A. (2003). Bayesian analysis of extreme values by mixture modeling. *Extremes*, 6(1):25–47.

Bray, I. (2002). Application of markov chain monte carlo methods to projecting cancer incidence and mortality. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 51(2):151–164.

Brodin, E. and Klüppelberg, C. (2008). Extreme value theory in finance. *Encyclopedia of Quantitative Risk Analysis and Assessment*.

Burton, P. W. (1979). Seismic risk in southern europe through to india examined using gumbel's third distribution of extreme values. *Geophysical Journal of the Royal Astronomical Society*, 59(2):249–280.

Capie, F., Mills, T. C., and Wood, G. (2005). Gold as a hedge against the dollar. *Journal of International Financial Markets, Institutions and Money*, 15(4):343–352.

Castillo, E. (1994). Extremes in engineering applications. In *Extreme value theory and applications*, pages 15–42. Springer, Dordrecht, netherlands.

Castillo, E., Hadi, A. S., Balakrishnan, N., and Sarabia, J.-M. (2005). *Extreme value and related models with applications in engineering and science*. Wiley Hoboken, NJ.

Chang, C.-L., Della Chang, J.-C., and Huang, Y.-W. (2013). Dynamic price integration in the global gold market. *The North American Journal of Economics and Finance*, 26:227–235.

Chen, M.-H. and Schmeiser, B. (1993). Performance of the gibbs, hit-and-run, and metropolis samplers. *Journal of Computational and Graphical Statistics*, 2(3):251–272.

Chen, M.-H. and Schmeiser, B. W. (2003). [slice sampling]: Discussion. *Annals of Statistics*, pages 742–743.

Coles, S. (2001). *An introduction to statistical modeling of extreme values*. Springer-Verlag.

Coles, S. (2003). The use and misuse of extreme value models in practice. *Finkenstädt and Rootzén (2003)*, pages 79–100.

Coles, S., Heffernan, J., and Tawn, J. (1999). Dependence measures for extreme value analyses. *Extremes*, 2(4):339–365.

Coles, S. and Pericchi, L. (2003). Anticipating catastrophes through extreme value modelling. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 52(4):405–416.

Coles, S., Pericchi, L. R., and Sisson, S. (2003). A fully probabilistic approach to extreme rainfall modeling. *Journal of Hydrology*, 273(1):35–50.

Coles, S. and Powell, E. (1996). Bayesian methods in extreme value modelling: a review and new developments. *International Statistical Review/Revue Internationale de Statistique*, pages 119–136.

Coles, S. and Tawn, J. (2005). Bayesian modelling of extreme surges on the uk east coast. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 363(1831):1387–1406.

Coles, S. G. and Tawn, J. A. (1991). Modelling extreme multivariate events. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 377–392.

Coles, S. G. and Tawn, J. A. (1996). A bayesian analysis of extreme rainfall data. *Applied Statistics*, 45(4):463–478.

Congdon, P. (2007). *Bayesian statistical modelling*, volume 704. John Wiley & Sons, Hoboken, NJ.

Cooray, K. (2010). Generalized gumbel distribution. *Journal of Applied Statistics*, 37(1):171–179.

Cornell, C. A. (1968). Engineering seismic risk analysis. *Bulletin of the Seismological Society of America*, 58(5):1583–1606.

Corsini, G., Gini, F., and Verrazzani, L. (1995). Cramer-rao bounds and estimation of the parameters of the gumbel distribution. *Aerospace and Electronic Systems, IEEE Transactions on*, 31(3):1202–1204.

Cruz, M. G. (2002). *Modeling, measuring and hedging operational risk*. John Wiley & Sons, New York.

Cyert, R. M. and DeGroot, M. H. (1987). *Bayesian analysis and uncertainty in economic theory*. Springer, USA.

Damien, P. and Walker, S. G. (2001). Sampling truncated normal, beta, and gamma densities. *Journal of Computational and Graphical Statistics*, 10(2):206–215.

Damlen, P., Wakefield, J., and Walker, S. (1999). Gibbs sampling for bayesian non-conjugate and hierarchical models by using auxiliary variables. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 61(2):331–344.

Danielsson, J. and de Vries, C. G. (1997). Tail index and quantile estimation with very high frequency data. *Journal of Empirical Finance*, 4(2):241–257.

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

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

de Haan, L. and Resnick, S. I. (1977). Limit theory for multivariate sample extremes. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 40(4):317–337.

de Zea Bermudez, P. and Kotz, S. (2010). Parameter estimation of the generalized pareto distributionpart i. *Journal of Statistical Planning and Inference*, 140(6):1353 – 1373.

de Zea Bermudez, P., Turkman, M. A., and Turkman, K. (2001). A predictive approach to tail probability estimation. *Extremes*, 4(4):295–314.

Deheuvels, P. (1983). Point processes and multivariate extreme values. *Journal of Multivariate Analysis*, 13(2):257–272.

Devroye, L. (1986). *Non-uniform random variate generation*. Springer-Verlag.

Diebolt, J., El-Aroui, M.-A., Garrido, M., and Girard, S. (2005). Quasi-conjugate bayes estimates for gpd parameters and application to heavy tails modelling. *Extremes*, 8(1-2):57–78.

Dimson, E., Marsh, P., and Staunton, M. (2009). *Triumph of the optimists: 101 years of global investment returns*. Princeton University Press.

Dooley, M. P., Isard, P., and Taylor, M. P. (1995). Exchange rates, country-specific shocks, and gold. *Applied financial economics*, 5(3):121–129.

DuBois, C., Korattikara, A., Welling, M., and Smyth, P. (2014). Approximate slice sampling for bayesian posterior inference. In *Artificial Intelligence and Statistics*.

Edwards, R. G. and Sokal, A. D. (1989). 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.

Embrechts, P. (1999). Extreme value theory in finance and insurance. *Manuscript, Department of Mathematics, ETH, Swiss Federal Technical University*.

Embrechts, P., Klüppelberg, C., and Mikosch, T. (1997). *Modelling extremal events: for insurance and finance*, volume 33. Springer Verlag.

Embrechts, P., Kluppelberg, C., and Mikosch, T. (1999). Modelling extremal events. *British Actuarial Journal*, 5(2):465–465.

Engeland, K., Hisdal, H., and Frigessi, A. (2004). Practical extreme value modelling of hydrological floods and droughts: a case study. *Extremes*, 7(1):5–30.

Engelund, S. and Rackwitz, R. (1992). On predictive distribution functions for the three asymptotic extreme value distributions. *Structural Safety*, 11(3):255–258.

Evans, M., Hastings, N., and Peacock, B. (2001). Statistical distributions.

Everitt, B. (2002). *The Cambridge dictionary of statistics/BS Everitt*. Cambridge University Press, Cambridge, UK New York:.

Favaro, S. and Walker, S. G. (2013). Slice sampling σ-stable poisson-kingman mixture models. *Journal of Computational and Graphical Statistics*, 22(4):830–847.

Felici, M., Lucarini, V., Speranza, A., and Vitolo, R. (2007a). Extreme value statistics of the total energy in an intermediate-complexity model of the midlatitude atmospheric jet. Part I: Stationary case. *Journal of the Atmospheric Sciences*, 64(7):2137–2158.

Felici, M., Lucarini, V., Speranza, A., and Vitolo, R. (2007b). Extreme value statistics of the total energy in an intermediate-complexity model of the midlatitude atmospheric jet. part ii: Trend detection and assessment. *Journal of the Atmospheric Sciences*, 64(7).

Fiorentino, M. and Gabriele, S. (1984). A correction for the bias of maximum-likelihood estimators of gumbel parameters. *Journal of Hydrology*, 73(1-2):39–49.

Fisher, R. and Tippett, L. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 24, pages 180–190. Cambridge Univ Press.

Fishman, G. S. (1999). An analysis of swendsen–wang and related sampling methods. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 61(3):623–641.

Frechet, M. (1927). Sur la loi de probabilité de lécart maximum. In *Annales de la societe Polonaise de Mathematique*, volume 6, pages 93–116. Bibliothèque des Sciences Humaines, Editions Gallimard.

Frey, B. (1997). Continuous sigmoidal belief networks trained using slice sampling. *Advances in neural information processing systems*, pages 452–458.

Friederichs, P. and Hense, A. (2007). Statistical downscaling of extreme precipitation events using censored quantile regression. *Monthly weather review*, 135(6).

Galambos, J. (1978). *The asymptotic theory of extreme order statistics*, volume 352. Wiley, New York.

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

Gelfand, A. and Smith, A. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American statistical association*, 85(410):398–409.

Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (2003). *Bayesian Data Analysis, Second Edition (Chapman & Hall/CRC Texts in Statistical Science)*. Chapman and Hall/CRC, 2 edition.

Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (2014). *Bayesian data analysis*, volume 2. Taylor & Francis.

Geman, S. and Geman, D. (1984). Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, (6):721–741.

Gençay, R., Selçuk, F., and Ulugülyaci, A. (2003). High volatility, thick tails and extreme value theory in value-at-risk estimation. *Insurance: Mathematics and Economics*, 33(2):337–356.

Ghosh, J. K., Delampady, M., and Samanta, T. (2007). *An introduction to Bayesian analysis: theory and methods*. Springer.

Gilli, M. et al. (2006). An application of extreme value theory for measuring financial risk. *Computational Economics*, 27(2-3):207–228.

Gnedenko, B. (1943). Stir la distribution limite du terme maximum d'une sarie alaatoire.

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

Gomes, M. I. and Guillou, A. (2014). Extreme value theory and statistics of univariate extremes: a review. *International Statistical Review*.

Gumbel, E. (1935). Les valeurs extrêmes des distributions statistiques. In *Annales de l'institut Henri Poincaré*, volume 5, pages 115–158. Presses universitaires de France.

Gumbel, E. J. et al. (1941). The return period of flood flows. *The Annals of Mathematical Statistics*, 12(2):163–190.

Hallerberg, S. and Kantz, H. (2008). Influence of the event magnitude on the predictability of an extreme event. *Physical Review E*, 77(1):011108.

Hammersley, J. and Morton, K. (1956). A new monte carlo technique: antithetic variates. In *Mathematical proceedings of the Cambridge philosophical society*, volume 52, pages 449–475. Cambridge Univ Press.

Hammoudeh, S., Malik, F., and McAleer, M. (2011). Risk management of precious metals. *The Quarterly Review of Economics and Finance*, 51(4):435–441.

Hammoudeh, S. M., Yuan, Y., McAleer, M., and Thompson, M. A. (2010). Precious metals–exchange rate volatility transmissions and hedging strategies. *International Review of Economics & Finance*, 19(4):633–647.

Hastings, W. (1970). Monte carlo sampling methods using markov chains and their applications. *Biometrika*, 57(1):97–109.

Higdon, D. (1998). Auxiliary variable methods for markov chain monte carlo with applications. *Journal of the American Statistical Association*, 93(442):585–595.

Hosking, J. (1985). Algorithm as 215: Maximum-likelihood estimation of the parameters of the generalized extreme-value distribution. *Applied Statistics*, pages 301–310.

Hosking, J., Wallis, J. R., and Wood, E. F. (1985). Estimation of the generalized extreme-value distribution by the method of probability-weighted moments. *Technometrics*, 27(3):251–261.

Hosking, J. R. (1984). Testing whether the shape parameter is zero in the generalized extreme-value distribution. *Biometrika*, 71(2):367–374.

Hosking, J. R. and Wallis, J. R. (1987). Parameter and quantile estimation for the generalized pareto distribution. *Technometrics*, 29(3):339–349.

Ibrahim, J. G. and Chen, M.-H. (2000). Power prior distributions for regression models. *Statistical Science*, pages 46–60.

Irony, T. and Singpurwalla, N. (1997). Noninformative priors do not exist: A discussion with jose m. bernardo. *Journal of Statistical Inference and Planning*, 65:159–189.

Jagger, T. H. and Elsner, J. B. (2006). Climatology models for extreme hurricane winds near the united states. *Journal of Climate*, 19(13):3220–3236.

Jenkinson, A. F. (1955). The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Quarterly Journal of the Royal Meteorological Society*, 81(348):158–171.

Jensen, M. J. (2001). Bayesian inference of long-memory stochastic volatility via wavelets. Technical report, Missouri Economics Working Paper.

Kaewkheaw, M., Leeahtam, P., and Chaiboosri, C. (2014). An analysis of relationship between gold price and us dollar index by using bivariate extreme value copulas. In *Modeling Dependence in Econometrics*, pages 455–462. Springer.

Kalli, M., Griffin, J. E., and Walker, S. G. (2011). Slice sampling mixture models. *Statistics and computing*, 21(1):93–105.

Kass, R. E. and Wasserman, L. (1996). The selection of prior distributions by formal rules. *Journal of the American Statistical Association*, 91(435):1343–1370.

Katz, R. (1999). Extreme value theory for precipitation: Sensitivity analysis for climate change. *Advances in Water Resources*, 23(2):133–139.

Katz, R. W. and Brown, B. G. (1992). Extreme events in a changing climate: variability is more important than averages. *Climatic change*, 21(3):289–302.

Katz, R. W., Brush, G. S., and Parlange, M. B. (2005). Statistics of extremes: Modeling ecological disturbances. *Ecology*, 86(5):1124–1134.

Koch, K.-R. (1990). *Bayes Theorem*. Springer.

Kolluri, B. R. (1981). Gold as a hedge against inflation-an empirical-investigation. *Quarterly Review of Economics and Business*, 21(4):13–24.

Kotz, S. and Nadarajah, S. (2000). *Extreme value distributions*, volume 31. Imperial College Press.

Koutsoyiannis, A. (1983). A short-run pricing model for a speculative asset, tested with data from the gold bullion market. *Applied Economics*, 15(5):563–581.

Lambert, P. C., Sutton, A. J., Burton, P. R., Abrams, K. R., and Jones, D. R. (2005). How vague is vague? a simulation study of the impact of the use of vague prior distributions in mcmc using winbugs. *Statistics in medicine*, 24(15):2401–2428.

Landwehr, J. M., Matalas, N., and Wallis, J. (1979). Probability weighted moments compared with some traditional techniques in estimating gumbel parameters and quantiles. *Water Resources Research*, 15(5):1055–1064.

Leese, M. N. (1973). Use of censored data in the estimation of gumbel distribution parameters for annual maximum flood series. *Water Resources Research*, 9(6):1534–1542.

Liechty, M. W. and Lu, J. (2010). Multivariate normal slice sampling. *Journal of Computational and Graphical Statistics*, 19(2).

Lingappaiah, G. (1984). Bayesian prediction regions for the extreme order statistics. *Biometrical journal*, 26(1):49–56.

Longin, F. M. (2000). From value at risk to stress testing: The extreme value approach. *Journal of Banking & Finance*, 24(7):1097–1130.

Lye, L., Hapuarachchi, K., and Ryan, S. (1993). Bayes estimation of the extreme-value reliability function. *Reliability, IEEE Transactions on*, 42(4):641–644.

McNeil, A. J. (1997). Estimating the tails of loss severity distributions using extreme value theory. *Astin Bulletin*, 27(01):117–137.

McNeil, A. J. and Frey, R. (2000). Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. *Journal of empirical finance*, 7(3):271–300.

Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A., and Teller, E. (1953). Equation of state calculations by fast computing machines. *The journal of chemical physics*, 21:1087.

Mira, A., Møller, J., and Roberts, G. O. (2001). Perfect slice samplers. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(3):593–606.

Mira, A. and Tierney, L. (1997). On the use of auxiliary variables in markov chain monte carlo sampling. *Scandinavian Journal of Statistics*.

Mira, A. and Tierney, L. (2002). Efficiency and convergence properties of slice samplers. *Scandinavian Journal of Statistics*, 29(1):1–12.

Mishra, B. and Rahman, M. (2005). The dynamics of bombay stock, us stock and london gold markets. *Indian Journal of Economics and Business*, 4(1):151–160.

Moore, G. H. (1990). Analysis: Gold prices and a leading index of inflation. *Challenge*, pages 52–56.

Mousa, M. A., Jaheen, Z., and Ahmad, A. (2002). Bayesian estimation, prediction and characterization for the gumbel model based on records. *Statistics: A Journal of Theoretical and Applied Statistics*, 36(1):65–74.

Murray, I., Adams, R., and MacKay, D. (2009). Elliptical slice sampling. *arXiv preprint arXiv:1001.0175*.

Nadarajah, S. (2006). The exponentiated gumbel distribution with climate application. *Environmetrics*, 17(1):13–23.

Neal, R. (2003). Slice sampling. *Annals of statistics*, pages 705–741.

Nicholis, N. (1997). Clivar and ipcc interests in extreme events. In *Workshop Proceedings on Indices and Indicators for Climate Extremes, Asheville, NC. Sponsors, CLIVAR, GCOS and WMO*.

Nieto, M., Cortés, A., Barandiaran, J., Otaegui, O., and Etxabe, I. (2013). Single camera railways track profile inspection using an slice sampling-based particle filter. In *Computer Vision, Imaging and Computer Graphics. Theory and Application*, pages 326–339. Springer.

Nishihara, R., Murray, I., and Adams, R. P. (2014). Parallel mcmc with generalized elliptical slice sampling. *The Journal of Machine Learning Research*, 15(1):2087–2112.

Oosterbaan, R. J. and Ritzema, H. (1994). Frequency and regression analysis. *Drainage principles and applications.*, (Ed. 2):175–223.

Phien, H. N. (1987). A review of methods of parameter estimation for the extreme value type-1 distribution. *Journal of Hydrology*, 90(3):251–268.

Pickands, J. (1981). Multivariate extreme value distributions. In *Proceedings 43rd Session International Statistical Institute*, volume 2, pages 859–878.

Pickands III, J. (1975). Statistical inference using extreme order statistics. *the Annals of Statistics*, pages 119–131.

Prescott, P. and Walden, A. (1980). Maximum likelihood estimation of the parameters of the generalized extreme-value distribution. *Biometrika*, 67(3):723–724.

Raftery, A. E. (1996). Approximate bayes factors and accounting for model uncertainty in generalised linear models. *Biometrika*, 83(2):251–266.

Raiffa, H. (1974). *Applied statistical decision theory*. Div. of Research, Graduate School of Business Administration, Harvard Univ.

Rakonczaia, P. and Tajvidia, N. (2010). On prediction of bivariate extremes. 2:174–192.

Reis Jr, D. S. and Stedinger, J. R. (2005). Bayesian mcmc flood frequency analysis with historical information. *Journal of Hydrology*, 313(1):97–116.

Reiss, R.-D. and Thomas, M. (2007). *Statistical analysis of extreme values: with applications to insurance, finance, hydrology and other fields*. Springer.

Resnick, S. (1987). *Extreme values, point processes and regular variation*. Springer-Verlag, New York.

Robert, C. and Casella, G. (2009). *Introducing Monte Carlo Methods with R*. Springer Science & Business Media.

Roberts, G. O. and Rosenthal, J. S. (1999). Convergence of slice sampler markov chains. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 61(3):643–660.

Rocco, M. (2014). Extreme value theory in finance: A survey. *Journal of Economic Surveys*, 28(1):82–108.

Ruggiero, P., Komar, P. D., and Allan, J. C. (2010). Increasing wave heights and extreme value projections: The wave climate of the us pacific northwest. *Coastal Engineering*, 57(5):539–552.

San Martini, F., Dunlea, E., Volkamer, R., Onasch, T., Jayne, J., Canagaratna, M., Worsnop, D., Kolb, C., Shorter, J., Herndon, S., et al. (2006). Implementation of a markov chain monte carlo method to inorganic aerosol modeling of observations from the mcma-2003 campaign–part ii: Model application to the cenica, pedregal and santa ana sites. *Atmospheric Chemistry and Physics*, 6(12):4889–4904.

Saunders, M. N., Saunders, M., Lewis, P., and Thornhill, A. (2011). *Research methods for business students, 5/e*. Pearson Education India.

Schmeiser, B. and Chen, M.-H. (1991). *On Hit-and-Run Monte Carlo sampling for evaluating multidimensional integrals*. Purdue University, Department of Statistics.

Sibuya, M. (1959). Bivariate extreme statistics, i. *Annals of the Institute of Statistical Mathematics*, 11(2):195–210.

Skold, M. (2005). Computer intensive statistical methods. *Center for Mathematical Sciences, Lund University*.

Smith, R. (2001). Extreme value statistics in meteorology and the environment. *Environmental statistics*, 8:300–357.

Smith, R. L. (1984). Efficient monte carlo procedures for generating points uniformly distributed over bounded regions. *Operations Research*, 32(6):1296–1308.

Smith, R. L. (1987). Estimating tails of probability distributions. *The annals of Statistics*, pages 1174–1207.

Smith, R. L. (1989). Extreme value analysis of environmental time series: an application to trend detection in ground-level ozone. *Statistical Science*, 4(4):367–377.

Smith, R. L. and Goodman, D. (2000). Bayesian risk analysis. *Extremes and integrated risk management*, pages 235–251.

Smith, R. L. and Shively, T. S. (1995). Point process approach to modeling trends in tropospheric ozone based on exceedances of a high threshold. *Atmospheric Environment*, 29(23):3489–3499.

Sornette, D., Knopoff, L., Kagan, Y., and Vanneste, C. (1996). Rank-ordering statistics of extreme events: Application to the distribution of large earthquakes. *Journal of Geophysical Research: Solid Earth (1978–2012)*, 101(B6):13883–13893.

Stephenson, A. (2004). A users guide to the evd package (version 2.1). *Department of Statistics. Macquarie University. Australia.*

Stephenson, A. and Tawn, J. (2004). Bayesian inference for extremes: Accounting for the three extremal types. *Extremes*, 7(4):291–307.

Sveinsson, O. G., Salas, J. D., and Boes, D. C. (2002). Regional frequency analysis of extreme precipitation in northeastern colorado and fort collins flood of 1997. *Journal of Hydrologic Engineering*, 7(1):49–63.

Swendsen, R. and Wang, J. (1987). Nonuniversal critical dynamics in monte carlo simulations. *Physical review letters*, 58(2):86–88.

Tancredi, A., Anderson, C., and OHagan, A. (2006). Accounting for threshold uncertainty in extreme value estimation. *Extremes*, 9(2):87–106.

Tawn, J. A. (1988). Bivariate extreme value theory: models and estimation. *Biometrika*, 75(3):397–415.

Taylor, N. J. (1998). Precious metals and inflation. *Applied Financial Economics*, 8(2):201–210.

Tiago de Oliveira, J. (1962). *Structure theory of bivariate extremes extensions*. Est. Mat., Estat. e. Econ.

Tiago de Oliveira, J. (1980). Bivariate extremes: Foundations and statistics. *Multivariate Analysis V*, pages 349–366.

Tibbits, M. M., Groendyke, C., Haran, M., and Liechty, J. C. (2014). Automated factor slice sampling. *Journal of Computational and Graphical Statistics*, 23(2):543–563.

Tibbits, M. M., Haran, M., and Liechty, J. C. (2011). Parallel multivariate slice sampling. *Statistics and Computing*, 21(3):415–430.

Trepanier, J. C. and Scheitlin, K. N. (2014). Hurricane wind risk in louisiana. *Natural hazards*, 70(2):1181–1195.

Trotter, H. (1954). Tukeyˆ jw," conditional monte carlo for normal-samples. In *Symposium on Monte Carlo Methods, HA Meyer, ed.(New York: John Wiley, 1956)*, page 64.

Vitolo, R., Ruti, P., Dell'Aquila, A., Felici, M., Lucarini, V., and Speranza, A. (2009). Accessing extremes of mid-latitudinal wave activity: methodology and application. *Tellus A*, 61(1):35–49.

Vitols, S. and Engelhardt, L. (2005). National institutions and high tech industries: A varieties of capitalism perspective on the failure of germany's neuer markt. *WZB, Markets and Political Economy Working Paper No. SP II*, 3.

Walker, S. G. (2007). Sampling the dirichlet mixture model with slices. *Communications in StatisticsSimulation and Computation®*, 36(1):45–54.

Walliman, N. (2010). *Research methods: The basics*. Routledge.

Walshaw, D. (2000). Modelling extreme wind speeds in regions prone to hurricanes. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 49(1):51–62.

Wang, Z. (2012). *The relationships between silver price, gold price and US dollar index before and after the subprime crisis*. Halifax, NS: Saint Mary's University.

Yoon, S., Cho, W., Heo, J.-H., and Kim, C. E. (2010). A full bayesian approach to generalized maximum likelihood estimation of generalized extreme value distribution. *Stochastic Environmental Research and Risk Assessment*, 24(5):761–770.