UNIVERSITI PUTRA MALAYSIA

EXTREME AIR POLLUTANT DATA ANALYSIS USING CLASSICAL AND BAYESIAN APPROACHES

NOR AZRITA BT MOHD AMIN

IPM 2015 15

EXTREME AIR POLLUTANT DATA ANALYSIS USING CLASSICAL AND BAYESIAN APPROACHES

By

NOR AZRITA BT MOHD AMIN

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

December 2015

DEDICATIONS

My deepest wish to my husband for his great support, understanding and being a strength for my PhD journey. To my lovely son and daughter, thank you for your love. My family and friends that always encourage and support me. Thank you so much. …

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfillment of the requirement for the degree of Doctor of Philosophy

EXTREME AIR POLLUTANT DATA ANALYSIS USING CLASSICAL AND BAYESIAN APPROACHES

By

NOR AZRITA BT MOHD AMIN

December 2015

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

Extreme value (EV) theory has raised researcher intention for modeling and forecasting of catastrophic or higher risk events. The concept of EV theory affords attention to the tails of distribution where standard models are proved unreliable. Generalized extreme value (GEV) distribution and generalized Pareto (GP) distribution are two main models in EV theory based on block maxima and threshold exceedances approaches. These two models are obviously different in terms of the sampling routine used in the formation of the extreme series. However, decisions on block sizes and threshold selection should be made by taking into consideration the limiting distribution properties.

Inferences on the extremes of environmental events are essential as guidelines in designing structures in order to survive under the utmost extreme conditions. Extreme air pollutants caused various effects associated to human health and material damages. In many cases, the pollutants are responsible for huge impacts on economic performances. The EV theory is applied to model the extreme PM_{10} pollutant for three air monitoring stations in Johor. This study started with the analysis of extreme PM_{10} data based on maximum likelihood estimation technique. Several block sizes were chosen to compare the model fit and hence estimate the return level. Using threshold exceedances technique, the selection of threshold value was made using mean residual life plot and threshold choice plot. Comparable estimates are found when the numbers of samples for both techniques are almost similar.

Alternatively, Bayesian framework is implemented to allow priors or additional information concerning the data into the analysis which expectantly improve the model fit. Bayesian inference in the context of EV theory obviously overcomes the scarcity of

extreme observations. The applications of Bayesian techniques have become practical through the development of simulation based techniques such as Markov chain Monte Carlo (MCMC). Two MCMC techniques are considered for the inferences namely Metropolis-Hastings (MH) algorithm and the Multiple-try Metropolis (MTM) algorithm. MTM algorithm is an extension of MH algorithm, designed to improve the convergence of MH algorithm by performing parallel computation. In general, both methods are performing well for analyzing extreme model but numerical results show that MTM method performs slightly better than MH method in terms of efficiency and convergency to the stationary distribution.

The univariate and bivariate extreme processes have been considered extensively using a frequentist perspective and recently there has been an increasing interest in the application of Bayesian methods to EV problems. Generally the univariate extreme inference has been considered commonly in Bayesian perspective. Bayesian techniques for bivariate model have not yet received much attention due to the hitches in dealing with much more parameters. Literature on Bayesian extremes based on MCMC techniques are dealing with either Gibbs sampling method or MH method, or the combination of both methods. This research implemented the MTM method as an alternative for modeling of univariate and bivariate extremes with non-informative priors. Bayesian technique for bivariate monthly maxima data from each pair of sites were employed to analyze the dependencies between two stations.

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

ANALISIS DATA PENCEMARAN UDARA EKSTRIM MENGGUNAKAN KAEDAH KLASIKAL DAN BAYESIAN

Oleh

NOR AZRITA BT MOHD AMIN

Disember 2015

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

Teori nilai ekstrim (NE) menarik perhatian penyelidik dalam pemodelan dan ramalan terhadap bencana alam atau kejadian-kejadian yang berisiko tinggi. Teori NE menyediakan kefahaman terhadap penghujung taburan di mana model-model lain telah dibuktikan tidak benar. Taburan nilai ekstrim teritlak (NET) dan taburan Pareto teritlak (PT) merupakan dua model utama bagi teori NE berasaskan kaedah maksima blok dan lebihan ambangan. Kedua-dua model tersebut sangat berbeza dari segi cara persampelan siri ekstrim. Walaubagaimanapun, keputusan bagi menentukan saiz blok dan pemilihan nilai ambangan perlu dilakukan dengan mengambil kira ciri-ciri khas had taburan.

Kesimpulan bagi peristiwa ekstrim terhadap alam sekitar adalah perlu sebagai tanda aras dalam rekaan struktur binaan supaya selamat walaupun ketika keadaan yang sangat ekstrim. Pencemaran udara yang ekstrim menyebabkan pelbagai kesan terhadap kesihatan seseorang dan kerosakan harta benda. Dalam pelbagai kes, bahan pencemaran tersebut adalah bertanggungjawab terhadap kesan yang besar bagi prestasi ekonomi. Teori NE diaplikasikan terhadap model ekstrim pencemaran PM_{10} untuk tiga stesen kawalan udara di Johor. Kajian dimulakan dengan analisis terhadap data ekstrim PM_{10} menggunakan kaedah panganggar kebarangkalian maksimum. Beberapa saiz blok dipilih bagi perbandingan model dan seterusnya menganggar tahap pulangan. Dengan menggunakan kaedah lebihan ambangan, pemilihan nilai ambangan adalah dengan menggunakan kaedah plot purata baki kehidupan dan plot pilihan ambangan. Nilai anggaran yang hampir sama diperolehi apabila bilangan sampel bagi kedua-dua teknik yang digunakan hampir sama.

Kaedah alternatif adalah dengan menggunakan kaedah Bayesian dengan membenarkankan keutamaan atau maklumat tambahan berkenaan data diambil kira dalam analisis yang diharapkan dapat memperbaiki pemodelan data. Aplikasi Bayesian semakin praktikal dengan adanya kemajuan teknik simulasi seperti teknik rantaian Markov Monte Carlo (RMMC). Dua kaedah RMMC digunakan dalam membuat kesimpulan iaitu algoritma Metropolis-Hastings (MH) dan algoritma Pelbagai-percubaan Metropolis (PPM). Algoritma PPM merupakan lanjutan bagi algoritma MH bagi memperbaiki penumpuan algoritma MH dengan menggunakan pengiraan selari. Umumnya, kedua-dua kaedah melaksanakan analisis model ekstrim dengan baik tetapi keputusan berangka menunjukkan PPM sedikit lebih baik dari MH dari segi kecekapan dan penumpuan ke taburan pegun.

Proses ekstrim bagi univariat dan bivariat telah dipertimbangkan dengan meluas menggunakan perspektif kekerapan. Keadaan pada masa kini mendapati terdapat peningkatan minat dalam kaedah Bayesian bagi aplikasi masalah NE. Secara umumnya, kesimpulan proses ekstrim bagi univariat telah dipertimbangkan dengan meluas menggunakan perspektif Bayesian. Teknik Bayesian bagi model bivariat masih tidak lagi mendapat lebih perhatian disebabkan kesukaran dalam menguruskan lebih banyak parameter. Sastera terdahulu terhadap bidang ekstrim Bayesian adalah terhad kepada kaedah RMMC berpandukan pesampelan Gibbs atau kaedah MH atau gabungan kedua-duanya. Kajian ini membangunkan kaedah PPM sebagai alternatif terhadap pemodelan ekstrim bagi univariat dan bivariat dengan menggunakan keutamaan tidak bermaklumat.

ACKNOWLEDGEMENTS

Thank to Allah for the opportunity and grace to have accomplished this research. My sincere appreciation to my very helpful supervisor, PM Dr Mohd Bakri Adam for his tireless supervision and motivation. He is not only a supervisor but most of all a guide from both professional and human sense. I will always be in debt with him for being so understanding even in my gloomiest hour.

I also wish to thank Professor Dr Noor Akma Ibrahim and PM Dr Ahmad Zaharin Aris as members of my committee for their guidance and their valuable suggestion along the way to the accomplishment of this thesis.

I would like to thank Malaysian Higher Education Ministry for the financial support for this study under Universiti Malaysia Perlis. I gratefully acknowledge the Universiti Putra Malaysia for the financial support supplied to my work under the ERGS grant project, Bayesian Extreme in Finance and Environment.

I certify that a Thesis Examination Committee has met on 28 August 2015 to conduct the final examination of Nor Azrita Bt Mohd Amin on his thesis entitled ”Extreme Air Pollutant Data Analysis using Classical and Bayesian Approach” 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:

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

Maman Abdurachman Djauhari, PhD
Professor
Faculty of Science
Universiti Putra Malaysia
(Internal Examiner)

Mohd Rizam b Abu Bakar, PhD
Professor
School of Mathematical Sciences
Universiti Sains Malaysia
(External Examiner)

Fabrizio Ruggeri, PhD
Professor
Laboratoire de Mathématiques Nicolas Oresme
Université de Caen Basse Normandie
France
(External Examiner)

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

Date:

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
Assosiate Professor
Faculty of Science
Universiti Putra Malaysia
(Chairperson)

Noor Akma Ibrahim, PhD
Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Member)

Ahmad Zaharin Aris, PhD
Assosiate Professor
Faculty of Science
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: Nor Azrita binti Mohd Amin, GS30752

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: Mohd Bakri Adam, PhD

Signature: ______________________________
Name of
Member of
Supervisory
Committee: Noor Akma Ibrahim, PhD

Signature: ______________________________
Name of
Member of
Supervisory
Committee: Ahmad Zaharin Aris, PhD

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

EV	extreme value
iid	independent and identically distributed
GEV	generalized extreme value
GP	generalized Pareto
MCMC	Markov chain Monte Carlo
MH	Metropolis-Hastings
MTM	Multiple-try Metropolis
API	air pollutant index
PM_{10}	particulate matters
SO_2	sulphur dioxide
NO_2	nitrogen dioxide
CO	carbon monoxide
DoE	Department of Environment, Malaysia

CHAPTER 1

INTRODUCTION

1.1 Introduction

Dealing with extreme observation events is critical in order to decide the appropriate actions for future extreme circumstances. Extreme studies on environmental events are among the most important areas to be explored extensively due to its worse impacts to humankind and materials. To achieve this purpose, a suitable statistical analysis must be adopted efficiently. Extreme value (EV) theory is one of the statistical methodologies that handle extreme situations. This chapter introduces the basic ideas of the research together with the motivation and the list of objectives.

1.2 Motivation

The occurrence of extreme events such as atmospheric pollutions, high rainfalls, floods and windstorms and many others are due to physical processes and also human activities. The impacts of these extreme phenomena have caused serious injuries, material damages as well as affecting the economic developments of a country. The relations of these catastrophic events with the statistical analysis of EV theory have been developed some decades earlier. EV theory is unlike other statistical approaches since its focus is on the tail of distribution either on maxima or minima values. The scope of EV theory has been widely explored in various fields. Recently, it has become an area vigorously researched due to its significance in many applications. Literatures on EV theory among others are by Coles (2001) and Haan and Ferreira (2006) that provide from basic EV theory to the application of EVT in various fields. Behrens et al. (2004) investigate the alternative, threshold approach based on Bayesian idea.

Statistical modelling of extreme air pollutants has a very practical motivation since these events have major effects on serious health threats. This motivates the need to estimate the most terrible air pollutant level that will be occur over a certain benchmark value in the future. In air quality control, particulate matter (PM_{10}) is recognized as the most influencing atmospheric pollutant for air quality index in a majority of cities in Malaysia. This situation is particularly due to the haze and biomass burning as well as industrial and vehicle emissions which usually contribute to high PM_{10} levels. This situation of unease has been an annual problem across Malaysia. The main concerns of this study are on the extreme levels of PM_{10} concentrations at three air quality monitoring stations in Johor, Malaysia. These three stations are located in different districts which have diverse roles. Therefore, we expect different patterns of statistical modelling for the extreme PM_{10} data from each station. The aim is to approximate the possible extreme PM_{10} levels in the future and provide important information to facilitate the proper procedures for combating these problems.

Fitting the model to the extreme data required the use of estimation methods for the unknown parameter θ. Among the very common methods are by using maximum likelihood estimation, method of moments and probability weighted moments. Undoubtedly, the most distinguished method is by using the maximum likelihood estimation method. However, data for rare events are often scarce because such events are necessarily unusual. Therefore, careful and sophisticated modelling is desirable to extract the fullest information from the data and to provide more accurate forecasts and associated measures of uncertainty. Bayesian framework offers an alternative to deal with small sample size and has managed to estimate models that are difficult to estimate using standard statistical approaches. Thus, combining the extreme analysis with Bayesian framework gives an advantage in management of the scarcity of the extreme data. One of the objectives for Bayesian extreme analysis is to elicit prior information for extreme PM_{10} in such a way that when combined with data through a Bayesian analysis, the posterior analysis obtained would provide a rational basis for extrapolation.

This study focuses on the estimation of the model parameters by using Metropolis-Hastings (MH) algorithm and new methods in Bayesian extreme models which is based on the Multiple-try Metropolis (MTM) algorithm. MTM extends the MH by promising larger areas for exploration in order to find the right points positions. The MTM results are consistent with those obtained by the MH method. The advantage in MTM is that it takes shorter iterations to meet the stationary distribution although the same initial values are used for both methods. MTM performs parallel computation depending on the number of proposals for which the rates of convergence accelerate as the number of proposals in MTM increase. The apparent drawback of MTM is the additional backward computation that makes the programming much more complex.

In order to implement the Markov chain Monte Carlo methods, some procedures have to be considered. The issues of initial values, the length of chains and the burn-in periods are also discussed as these matters has always been problematic in Bayesian modeling practices. Convergence diagnostic provides better understanding in terms of assessing the performance of MCMC algorithms. Researches on MCMC techniques are widely applied in numerous fields but only a few studies worked on the practical use of diagnosing the convergence of the algorithms. MCMC techniques do not give a clear indication on whether the iterations have been converged. The fundamental theory of MCMC only guarantees that the distribution of the output will converge to the posterior distribution as the number of iterations increases to infinity. However, it is not guaranteed that the chain will converge after a certain number of iterations.

1.3 Problem Statement

Maximum likelihood is a very well-known method for estimation of parameters while Bayesian framework gives an alternative thought on data modeling with the facility of prior information. This countenance the additional knowledge of the process based on expert knowledge before exploring the behavior of the data. Currently there is a wide variety of Bayesian techniques developed and practiced for statistical modelling. But it is important to understand that each idea developed based on Bayesian framework has its own distinct advantages and drawbacks. This thesis considers the MCMC techniques which are MH and MTM methods for the analysis of extreme data. MTM algorithm modifies the MH algorithm by expanding the proposal region to improve convergence performance by generating a larger number of candidates, k and therefore improving exploration of the chain near current value, $x^{(t)}$. It is expected that the higher number of candidates, k give better convergence of the draws to the preferred stationary distribution.

Using Bayesian inference on extremes of environmental problems would allow any additional information about the processes to be incorporated as prior information. Due to the lack of data, the benefits of using any information available are likely to be great. In this thesis, the priors are constructed by assuming there is no information available about the process apart from the data. It is to be expected that posterior means would be close to the maximum likelihood estimates, since the priors were almost flat and added very little information to the likelihood but do tend to be slightly higher.

The main anxieties in environmental management are on extreme phenomena (catastrophic) instead of common events. However, most statistical approaches are concerned primarily with the center of a distribution or on the average value rather than the tail of the distribution which contains all the high observations. EV theory offers a strong statistical tool for analyzing rare events and predicts the maximum concentration in a certain return period for air quality management purposes. The adverse effects of PM_{10} to human health and material damages are the main reasons for extensive explorations on the behavior of this pollutant especially on the extreme level. High level of PM_{10} has strong effects on mortality and morbidity among population with high health risk. Environmental risk management is more concerned on the occurrence of extreme pollutant level than normal level due to the serious impacts of the pollutant on individuals, organizations and also to the developments of a country. Thus is the significance to give more attention to the modeling and analysis of these extreme events.

1.4 Objectives of the Thesis

1. To evaluate the extreme PM_{10} concentration model based on block maxima and threshold exceedances approaches.

2. To investigate the statistical inferences of extreme PM_{10} data using maximum likelihood estimation method.

3. To propose a new approach in Bayesian extreme studies that is the MTM technique for analysing the statistical inferences of extreme PM_{10} data.

4. To analyse the efficiency for estimating location, μ and scale, σ parameters of Gumbel distribution simulated data, with different number of proposals and the influence of initial values using MH and MTM approach.

5. To examine the convergence of the MTM and MH for the inferences of EV distributions.

6. To evaluate the dependencies of extreme PM_{10} data between two air monitoring stations in Johor.

1.5 Thesis Outline

The thesis is structured as follows. Chapter 1 describes the motivation of the research area and sets up the problem statement as well as the objectives of the research. Chapter 2 provides the methodologies of EV theory together with some important literatures. The concepts of block maxima and threshold exceedances approaches are discussed with the corresponding return levels for both approaches.

In Chapter 3, preliminary studies on univariate EV analysis are introduced to investigate the behavior of high PM_{10} data for different block sizes. Air quality data for three air quality monitoring stations in Johor are analyzed separately. The statistical analysis is performed using maximum likelihood method. An alternative threshold approach to analyse the similar data discussed in Chapter 3 is implemented in Chapter 4. The threshold exceedances approach considers the extreme data exceeding an appropriate chosen threshold. Some techniques for threshold selection are also presented. At the end of the chapter, we compare and discuss the analysis of block maxima and threshold exceedances approaches for extreme PM_{10} data.

Bayesian framework with focus on Markov chain Monte Carlo (MCMC) techniques are introduced in Chapter 5. The general ideas of Bayesian modeling and MCMC techniques applied throughout the thesis are presented. MH and MTM are Bayesian methods developed based on MCMC idea for the analysis of the posterior distribution. The simulation study for Gumbel distribution is covered in this chapter. This provides an illustration of the implementation of Bayesian techniques for EV distributions. Convergence diagnostic tests are introduced to investigate the performance of MH algorithms. In Chapter 6, the Bayesian modeling for extreme PM_{10} data are executed.

Besides working on the univariate extreme modeling, some extent of bivariate extreme models will be considered by comparing the PM_{10} data from two different stations. Chapter 7 applies the component wise block maxima for the analysis of bivariate extreme data and the estimation of parametric models which are computed using maximum likelihood and Bayesian methods. The dependencies of extreme PM_{10} between two air monitoring stations are analyzed. Finally Chapter 8 concludes the overall thesis and provides some recommendations.

BIBLIOGRAPHY

Abatzoglou, G., Chaloulakou, A., Assimacopoulos, D., and Lekkas, T. (1996). Prediction of air pollution episodes: Extreme value theory applied in athens. *Environmental Technology*, 17:349–359.

Abidin, N. Z., Adam, M. B., and Midi, H. (2012). The goodness-of-fit test for gumbel distribution: A comparative study. *MATEMATIKA*, 28(1):35–48.

Ali, N., Adam, M. B., Ibrahim, N. A., and Daud, I. (2012). Statistical analysis of extreme ozone data. *Journal of Statistical Modeling and Analytics*, 3:11–18.

Aryal, S. K., Bates, B. C., Campbell, E. P., Li, Y., Palmer, M. J., and Viney, N. R. (2009). Characterizing and modeling temporal and spatial trends in rainfall extremes. *Journal of Hydrometeor*, 10:241–253.

Balakrishnan, N., Davies, K. F., Keating, J. P., and Mason, R. L. (2011). Correlation-type goodness-of-fit test for extreme value distribution based on simultaneous closeness. *Communications in Statistics-Simulation and Computation*, 40:1074–1095.

Bedard, M., Douc, R., and Moulines, E. (2012). Scaling analysis of multiple-try mcmc methods. *Stochastic Processes and their Applications*, 122(3):758–786.

Behrens, C. N., Lopes, H. F., and Gamerman, D. (2004). Bayesian Analysis of Extreme Events with Threshold Estimation. *Statistical Modelling*, 4:227–244.

Beirlant, J., Geogebeur, Y., Teugels, J., and Segers, J. (2004). *Statistics of Extremes: Theory and Applications*. John Wiley and Sons, Ltd, England.

Bermudez, P. Z. and Kotz, S. (2010). Parameter estimation of the generalized pareto distribution part ii. *Journal of Statistical Planning and Inference*, 140:1374–1388.

Brodin, E. and Rootzen, H. (2009). Univariate and bivariate gpd methods for predicting extreme wind storm losses. *Insurance: Mathematics and Economics*, 44:345–356.

Brooks, S. P. (1998). Markov chain monte carlo method and its application. *Journal of the Royal Statistical Society. Series D*, 47(1):69–100.

Brooks, S. P. and Roberts, G. O. (1998). Convergence assessment of markov chain monte carlo algorithms. *Statistics and Computing*, 8:319–335.

Buishand, T. A. (1984). Bivariate extreme value data and the station year method. *Journal of Hydrology*, 69:77–95.

Cabras, S., Castellanos, M. E., and Gamerman, D. (2011). A default bayesian approach for regression on extremes. *Statistical Modelling*, 11(6):557–580.

Carlin, B. P. and Louis, T. A. (2009). *Bayesian Method for Data Analysis*. Chapman and Hall/CRC, United States of America, 3 edition.

Casarin, R., Craiu, R., and Leisen, F. (2013). Interacting multiple-try algorithms with different proposal distributions. *Statistical Computation*, 23:185–200.

Castellanos, M. E. and Cabras, S. (2007). A default bayesian procedure for the generalized pareto distribution. *Journal of Statistical Planning and Inference*, 137:473–483.

Chaloulakou, A., Assimacopoulos, D., and Lekkas, T. (1996). Forecasting dailymaximum ozone concentrations in the athens basin. *Environmental Monitoring and Assessment*, 56:97–112.

Chib, S. and Greenberg, E. (1995). Understanding the Metropolis-Hastings Algorithm. *The American Statistician*, 49(4):327–335.

Coles, S. G. (2001). *An Introduction to Statistical Modeling of Extreme Values*. Springer-Verlag, London.

Coles, S. G. and Powell, E. A. (1996). Bayesian Methods in Extreme Value Modelling: A Review and New Developments. *International Statistical Review*, 64:119–136.

Coles, S. G. and Tawn, J. A. (1991). Modelling extreme multivariate events. *Journal of the Royal Statistical Society. Series B (Methodological)*, 53:377–392.

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

Cooley, D. S. (2005). *Statistical Analysis of Extremes Motivated by Weather and Climate Studies: Applied and Theoretical Advances*. PhD thesis, Department of Applied Mathematics, University of Colorado.

Cowles, M. K. and Carlin, B. P. (1996). Markov chain Monte Carlo Convergence Diagnostics: a Comparative Review. *Journal of the American Statistical Association*, 91(434):883–904.

Davison, A. C. and Smith, R. L. (1990). Models for exceedances over high thresholds. *Journalm of Royal Statistical Society B*, 52(3):393–442.

Deidda, R. (2010). A multiple threshold method for fitting the generalized pareto distribution to rainfall time series. *Hydrology and Earth System Sciences*, 14:2559–2575.

DOE (2010). Malaysia environmental quality report, ministry of natural resources and environment malaysia.

Dominick, D., Juahir, H., Latif, M. T., Zain, S. M., and Aris, A. Z. (2012). Spatial Assessment of Air Quality Patterns in Malaysia using Multivariate Analysis. *Atmospheric Environment*, 60:172–181.

Eastoe, E. F. and Tawn, J. A. (2009). Modelling non-stationary extremes with application to surface level ozone. *Applied Statistics*, 58:22–45.

Eli, A., Shaffie, M., and Zin, W. Z. W. (2012). Preliminary Study on Bayesian Extreme Rainfall Analysis: A Case Study of Alor Setar, Kedah, Malaysia . *Sains Malaysiana*, 41:1403–1410.

Ercelebi, S. G. and Huseyin, T. (2009). Extreme value analysis of istanbul air pollution data. *CLEAN - Soil, Air, Water*, 37(2):122–131.

Fawcett, L. and Walshaw, D. (2006). Markov chain models for extreme wind speeds. *Environmetrics*, 17:795–809.

Finkenstadt, B. and Rootzen, H. (2004). *Extreme Values in Finance, Telecommunications, and the Environment*. Chapman and Hall / CRC., United States of America.

Fisher, R. A. and Tippett, L. H. C. (1928). The Frequency Distribution of the Largest and Smallest Member of a Sample. *Proceeding Cambridge Philosophy Society*, 24:180–190.

Gaioni, E., Dey, D., and Ruggeri, F. (2010). Bayesian modeling of flash floods using generalized extreme value distribution with prior elicitation. *Chilean Journal of Statistics*, 1(1):75–90.

Galambos, J., Lechner, J., and Simiu, E. (1994). *Extreme Value Theory and Applications*, pages 1–14. Kluwer Academic Publishers.

Gamerman, D. and Lopes, H. F. (2006). *Markov chain Monte Carlo Stochastic Simulation for Bayesian Inference*. Chapman and Hall / CRC, United States of America, 2 edition.

Gelfand, A. E. and Smith, A. F. M. (1990). Sampling-based approaches to calculate marginal densities. *Journal of the American Statistical Association*, 85(410):398–409.

Gelman, A. and Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, 7(4):57–472.

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

Geweke, J. (1992). *Bayesian Statistics 4*, pages 169–193. Oxford University Press, Oxford.

Geweke, J. and Tanizaki, H. (2003). Note on the sampling distribution for the metropolis-hastings algorithm. *Communications in Statistics - Theory and Methods*, 32:775–789.

Gilli, M. and Kellezi, E. (2006). An application of extreme value theory for measuring financial risk. *Computational Economics*, 1:1–23.

Givens, G. H. and Hoeting, J. A. (2012). *Computational Statistics*. John Wiley and Sons, United States, 2 edition.

Gumbel, E. J. and Goldstein, N. (1964). Analysis of empirical bivariate extremal distributions. *Journal of the American Statistical Association*, 59:794–816.

Haan, D. and Ferreira, L. (2006). *Extreme Value Theory An Introduction*. Springer, New York.

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

Heffernan, J. E. and Tawn, J. A. (2004). A conditional approach for multivariate extreme values. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 66(3):497–546.

Hirose, H. (1996). Maximum likelihood estimation in the 3-parameter weibull distribution. a look through the generalized extreme-value distribution. *IEEE Transactions on Dielectrics and Electrical Insulation*, 3(1):43–55.

Hurairah, A., Ibrahim, N. A., Daud, I., and Haron, K. (2005). An application of a new extreme value distribution to air pollution data. *Management of Environmental Quality: An International Journal*, 16(1):17–25.

Jenkinson, A. F. (1955). The Frequency Distribution of Annual Maxima and Minima of Meteorological Elements. *Quarterly Journal of the Royal Meteorological Society*, 81:158–171.

Joe, H. (1990). Families of min-stable multivariate exponential and multivariate extreme value distributions. *Statistics and Probability Letters*, 9:75–81.

Joe, H., Smith, R. L., and Weissman, I. (1992). Bivariate threshold methods for extremes. *Journal of the Royal Statistical Society. Series B (Methodological)*, 54:171–183.

Johansen, S. S. (2004). *Bivariate Frequency Analysis of Flood Characteristics in Glomma and Gudbrandsdalslagen*. PhD thesis, Department of Geosciences, University of Oslo.

Juneng, L., Latif, M. T., Tangang, F. T., and Mansor, H. (2009). Spatio-temporal characteristics of pm_{10} concentration across malaysia. *Atmospheric Environment*, 43:4584–4594.

Katz, R. W. (2010). Statistics of extremes in climate change. *Climatic Change*, 100:71–76.

Kochenhoff, H. and Thamerus, M. (1996). Extreme value analysis of munich air pollution data. *Environmental and Ecological Statistics*, 3:127–141.

Kotz, S. and Nadarajah, S. (2000). *Extreme Value Distributions Theory and Applications*. Imperial College Press, London.

Liu, J. S., Liang, F., and Wong, W. H. (2000). The multiple-try method and local optimization in metropolis sampling. *Journal of the American Statistical Association*, 95(449):121–134.

Lu, H. S. (2004). Estimating the emission source reduction of pm_{10} in central taiwan. *Chemosphere*, 54(7):805–814.

Martino, L. and Read, J. (2013). On the flexibility of the design of multiple-try metropolis schemes. *Computational Statistics*, 28(6):2797–2823.

Mengersen, K. and Robert, C. P. (1999). *Bayesian Statistics 6*, pages 415–440. Oxford University Press, Oxford.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953). Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087–1092.

Muller, P. (1991). A generic approach to posterior integration and gibbs sampling. Technical report, Technical Report 91-09, Department Statistics, Purdue University.

Nadarajah, S. and Eljabri, S. (2014). On chen et al.'s extreme value distribution. *Journal of Data Science*, 12:87–106.

O'Hagan, A. (2008). *The Bayesian Approach to Statistics*. SAGE Publications, UK.

Pandolfi, S., Bartolucci, F., and Friel, N. (2010). A generalization of the multiple-try metropolis algorithm for bayesian estimation and model selection. *Journal of Machine Learning Research*, 9:581–588.

Pandolfi, S., Bartolucci, F., and Friel, N. (2014). A generalized multiple-try version of the reversible jump algorithm. *Computational Statistics and Data Analysis*, 72:298–314.

Paulino, P. R., Humberto, V. H., and Jose, A. V. (2009). A goodness-of-fit test for the gumbel distribution based on kullbackleibler information. *Communications in Statistics - Theory and Methods*, 38(6):842–855.

Payus, C., Abdullah, N., and Sulaiman, N. (2013). Airborne particulate matter and meteorological interactions during the haze period in malaysia. *International Journal of Environmental Science and Development*, 4(4):398402.

Pickands, J. (1975). Statistical inference using extreme order statistics. *Annals Statistics*, 3:119–131.

Raftery, A. E. and Lewis, S. M. (1992). *Bayesian Statistics 4*, pages 763–773. Oxford University Press, Oxford.

Rakonczai, P. (2012). Asymmetric dependence models for bivariate threshold exceedance models. *Forum Statisticum Slovacum*, 1:25–32.

Rakonczai, P. and Tajvidi, N. (2010). On prediction of bivariate extremes. *The Thailand Econometrics Society*, 2:174–192.

Rakonczai, P. and Zempleni, A. (2011). Bivariate generalized pareto distribution in practice: Models and estimation. *Environmetrics*, 23:219–227.

Robert, C. and Casella, G. (2010). *Introducing Monte Carlo Methods with R.* Springer Science + Business Media, London.

Rocco, M. (2012). Extreme value theory in finance: a survey. *Journal of Economic Surveys*, 28:82–108.

Rootzen, H. and Tajvidi, N. (2006). Multivariate generalized pareto distributions. *Bernouli*, 12:917–930.

Rostami, M. and Adam, M. B. (2013). Analyses of prior selections for gumbel distribution. *MATEMATIKA*, 29(1):95–107.

Roth, M., Buishand, T. A., Jongbloed, G., A.M.G.KleinTank, and Zanten, J. H. (2014). Projections of precipitation extremes based on a regional, non-stationary peaks over threshold approach: a case study for the netherlands and north-western germany. *Weather and Climate Extremes*, 4:1–10.

Sandoval, C. E. (2007). Application of bivariate extreme value distribution to flood frequency analysis: a case study of northwestern mexico. *Nat Hazards*, 42:37–46.

Sani, S. (1999). *Beyond Environmental Legislation: Environmental education in Malaysia*. Shiga University, United States of America.

Shabri, A., Daud, Z. M., and Ariff, N. M. (2011). Regional analysis of annual maximum rainfall using TL-moments method. *Theoretical and Applied Climatology*, 104(1):561–570.

Sharma, P., Avinash, C., Kaushik, S. C., Sharma, P., and Suresh, J. (2012). Predicting violations of national ambient air quality standards using extreme value theory for delhi city. *Atmospheric Pollution Research*, 3:170–179.

Shinyie, W. L. and Ismail, N. (2012). Analysis of T-Year Return Level for Partial Duration Rainfall Series. *Sains Malaysiana*, 41(11):1389–1401.

Sinharay, S. (2003). Assessing convergence of the markov chain monte carlo algorithms: a review. Technical report, Educational Testing Service, Princeton NJ.

Sinnadura, J. (2006). Clearing the air "about the haze". *Medical Journal Malaysia*, 61(1):117–121.

Smith, A. F. M. and Roberts, G. O. (1993). Bayesian computation via the gibbs sampler and related markov chain monte carlo methods. *Journal of the Royal Statistical Society. Series B*, 55(1):3–23.

Smith, R. L. (1986). Extreme value statistics and reliability applications. *Reliability Engineering*, 15:161–170.

Smith, R. L. (1987). Estimating tails of probability distribution. *The Annals of Statistics*, 15(3):1174–1207.

Smith, R. L. (1994). *Statistical Extremes and Applications*, pages 621–638. Springer Science + Business Media B.V., Springer Netherlands.

Smith, R. L. and Naylor, J. C. (1987). A Comparison of Maximum Likelihood and Bayesian Estimators for the Three Parameter Weibull Distribution. *Applied Statistics*, 36:358–396.

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

Thibaud, E., Mutzner, R., and Davison, A. C. (2013). Threshold modeling of extreme rainfall. *Water Resources Research*, 49:4633–4644.

Wadsworth, J. L. and Tawn, J. A. (2012). Likelihood-based procedures for threshold diagnostics and uncertainty in extreme value modelling. *Journal of the Royal Statistical Society*, 74:543–567.

Yue, S. (2001). The Gumbel Logistic Model for Representing a Multivariate Storm Event. *Advances in Water Resources*, 24:179–185.

Yue, S., Ouardaa, T. B. M. J., Bobeea, B., Legendre, P., and Bruneau, P. (1999). The gumbel mixed model for flood frequency analysis. *Journal of Hydrology*, 226:88–100.

Yue, S. and Wang, C. Y. (2004). A Comparison of Two Bivariate Extreme Value Distributions. *Stochastic Environmental Research*, 18:61–66.

Yuguo, D., Bingyan, C., and Zhihong, J. (2008). A newly-discovered gpd-gev relationship together with comparing their models of extreme precipitation in summer. *Advances in Atmospheric Sciences*, 25(3):507–516.

Yusof, N. F. F. M., Ramli, N. A., and Yahaya, A. S. (2011). Extreme value distribution for prediction of future pm_{10} exceedences. *International Journal of Environmental Protection*, 1(4):28–36.

Zalina, M. D., Desa, M. N. M., Nguyen, V. T. V., and Kassim, A. H. M. (2002). Selecting a probability distribution for extreme rainfall series in malaysia. *Water Science and Technology*, 45(2):63–68.

Zin, W. Z. W. (2009). A comparative study of extreme rainfall in peninsular malaysia: with reference to partial duration and annual extreme series. *Sains Malaysiana*, 38(5):751–760.

Zin, W. Z. W., Jemain, A. A., and Ibrahim, K. (2012). Bayesian changepoint analysis of the extreme rainfall events. *Journal of Mathematics and Statistics*, 8(1):85–91.