UNIVERSITI PUTRA MALAYSIA

INFERENCE FOR AUTOREGRESSIVE AND MOVING AVERAGE MODELS WITH EXTREME VALUE DISTRIBUTION VIA SIMULATION STUDY

BAKO SUNDAY SAMUEL

INFERENCE FOR AUTOREGRESSIVE AND MOVING AVERAGE MODELS WITH EXTREME VALUE DISTRIBUTION VIA SIMULATION STUDY

By

BAKO SUNDAY SAMUEL

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

May 2015

DEDICATIONS

To the memory of my late dad, God bless his soul.
To my mum, for her steadfastness in prayers.
To my wife, Lovinda, for her Love and understanding while I was away.
To my daughter, Shiayet: Daddy will make up for the lost time!

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in ful lment of the requirement for the degree of Master of Science

INFERENCE FOR AUTOREGRESSIVE AND MOVING AVERAGE MODELS WITH EXTREME VALUE DISTRIBUTION VIA SIMULATION STUDY

By

BAKO SUNDAY SAMUEL

May 2015

Chairman: Mohd Bakri Adam, Ph.D
Faculty: Science

Time series analysis has emerged as one of the most important statistical discipline and it has been applied in di erent elds over the years. Literature reviews show that independent identical distributed Gaussian random variables is not suitable for modelling extreme events. We evaluate the impact of dependence on the parameter estimates of Autoregressive (AR) and Moving Average (MA) processes with Gumbel distributed innovation. The performance of the parameter estimates of the Gumbel-generalised Pareto distribution tted to the autoregressive and moving average processes and their respective cluster maxima is also assess.

The extension of time series to extreme value theory can be achieved by inducing time dependence in the underlying state of an extreme value process. Extreme values occur in clusters in the presence of dependence. Gumbel distribution, a member of the family of the generalised extreme value distribution is the possible limit for the entire range of tail behaviour between polynomial decrease and essentially a nite endpoint and it is known to t well in many situations. It is important to make general statements that characterises time series extreme models over a range of sample sizes with varying degree of dependence. Such general characterisation for a given model is useful for the extremal behaviour of physical processes.

To achieve our objectives, a stationary autoregressive and moving average models with Gumbel distributed innovation is proposed and we characterise the short-term dependence among maxima, arising from light-tailed Gumbel distribution over a range of sample sizes with varying degrees of dependence. Dependence is induced through a

linear lter operation. The linear lter operation takes a weighted sum of past innovations. The estimate of the maximum likelihood of the parameters of the Gumbel autoregressive and Gumbel moving average processes and their respective residuals are evaluated. Gumbel-AR(1) and Gumbel-MA(1) was tted to the Gumbel-generalised Pareto distribution and we evaluate the performance of the parameter estimates tted to the cluster maxima and the original series. Ignoring the e ect of dependence leads to overestimation of the location parameter of the Gumbel-AR(1) and Gumbel-MA(1) processes respectively. The estimate of the location parameter of the autoregressive process using the residuals gives a better estimate. The estimate of the scale parameter perform marginally better for the original series than the residual estimate. The degree of clustering increases as dependence is enhance for both the AR and MA processes. The Gumbel-AR(1) and Gumbel-MA(1) are tted to the Gumbel-generalised Pareto distribution show that the estimates of the scale and shape parameters tted to the cluster maxima perform better as sample size increases, however, ignoring the e ect of dependence leads to an underestimation of the parameter estimates of the scale parameter. The shape parameter of the original series gives a superior estimate compare to the threshold excesses tted to the Gumbel-generalised Pareto distribution.

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk ijazah Master Sains

INFERENS BAGI AUTOREGRESIF DAN PURATA-TERGERAK MODEL DENGAN TABURAN NLAI-NILAI EKSTRIM MELALUI KAJIAN SIMULASI

Oleh

BAKO SUNDAY SAMUEL

Mei 2015

Pengerusi: Mohd Bakri Adam, Ph.D
Fakulti: Sains

Analisis siri masa telah muncul sebagai salah satu disiplin statistik yang paling penting dan ia telah digunakan didalam pelbagai bidang setelah sekian lama. Hasil tinjauan daripada kajian-kajian yang lepas, ia menunjukkan bahawa pembolehubah rawak Gaussian yang bertaburan bebas-sama adalah tidak sesuai untuk menghasilkan model peristiwa ekstrim. Kami menilai kesan pergantungan terhadap anggaran parameter daripada proses Autoregresif (AR) dan Purata-Tergerak (MA) dengan inovasi taburan Gumbel. Prestasi anggaran parameter daripada taburan Pareto Gumbel-umum yang sepadan dengan proses autoregrasi dan purata-tergerak serta kelompok maksima masing-masing juga dinilai.

Perlanjutan siri masa kepada teori nilai ekstrim boleh di capai dengan mendorong pergantungan masa dalam keadaan mendasari proses nilai yang ekstrim. Nilai ekstrem berlaku di dalam kelompok di mana pergantungan hadir. Taburan Gumbel, merupakan salah satu daripada taburan nilai ekstrim umum adalah had sempadan yang mungkin bagi keseluruhan rangkaian tingkah laku hujung ekor antara penurunan polinomial dan asasnya adalah titik akhir yang terbatas dan ia dikenali sebagai pemadansuai yang baik dalam pelbagai situasi. Ia adalah penting untuk membuat kenyataan umum yang menyifatkan model ekrtrim siri masa lebih pelbagai saiz sampel dengan pelbagai tahap kebergantungan. Pencirian umum untuk model yang diberikan adalah berguna bagi ciri-ciri ekstrim dengan proses zikal.

Bagi mencapai objektif-objektif kajian, model autoregresi pegun dan purata-tergerak dengan taburan Gumbel inovasi dicadangkan dan kami mencirikan pergantungan jangka pendek antara maksima yang terhasil daripada taburan Gumbel yang berekor nipis yang lebih pelbagai saiz sampel dengan pelbagai tahap kebergantungan. Kebergantungan didorong melalui operasi penapis linear. Operasi penapis linear mengambil

kira jumlah wajaran inovasi yang lalu. Anggaran kemungkinan parameter maksimum bagi proses autoregresi Gumbel dan Gumbel purata-tergerak beserta residual mereka masing-masing dinilai. Gumbel-AR(1) dan Gumbel-MA(1) dilengkapi dengan taburan Pareto Gumbel-umum dan kami menilai prestasi anggaran parameter yang sepadan dengan kelompok maksima dan siri-siri asal. Mengabaikan kesan pergantungan boleh membawa kepada perlebihan anggaran daripada parameter lokasi dan kekurangan anggaran daripada parameter skala daripada proses Gumbel-AR(1) dan Gumbel-MA(1) masing-masing. Anggaran parameter lokasi daripada proses autoregresif menggunakan residual memberikan anggaran yang lebih baik. Anggaran parameter skala memberikan sedikit hasil yang lebih baik untuk siri-siri asal berbanding anggaran residual. Tahap kelompokan meningkat apabila pergantungan meningkat untuk kedua-dua proses AR dan MA. Gumbel-AR(1) dan Gumbel-MA(1) sepadan pada taburan Pareto Gumbel-umum menunjukkan bahawa anggaran parameter skala dan bentuk sepadan pada kelompok maksima memberi hasil yang lebih baik apabila saiz sampel bertambah, bagaimanapun, mengabaikan kesan pergantungan membawa kepada kekurangan terhadap anggaran parameter skala dan parameter bentuk taburan Pareto Gumbel-umum masing-masing.

ACKNOWLEDGEMENTS

I would like to express my profound gratitude to Associate Prof. Mohd Bakri Adam who introduced me to the world of extreme values by giving me the chance to work with him whilst guiding my rst attempt to in depth scienti c research. I am also grateful to Dr. Anwar Fitrianto, a member of my supervisory committee, for his useful comments. The constructive criticisms of the members of the Thesis examination committee towards the improvement of this research work is higly appreciated. I also acknowledge the sta of the department of Mathematics for providing a friendly working environment.

I am one of the recipients of Tertiary Education Trust Fund (TETFund) schorlaship in Nigeria, and it is gratefully acknowledge. I must also acknowledge my employer, Kaduna State University Nigeria, without whom the TETFund schorlaship would not have been possible.

I am grateful to my fellow students under the supervision of Associate Prof. Mohd Bakri Adam. I appreciate my at mates (Nicodemus, Jaaf and MD) and my fellow students whom we share same study room without any incident! I am also grateful to my friends, Capt. GD Bako, Lt. GJ Aluku, Bayei Mantwang Zachiam, Ados Peter, Sarki Philips, Peter Waziri, Peter Adamu, Snow Sini, Usman, Lakin, Mainangwa, Emmanuel Nok, Mr Alphosus, Mr Paul Dalatu and Arzuka Ibrahim. I also acknowledge my colleagues in Kaduna State University.

Special thanks to my mum, Mrs Esther Bako, and my siblings (Martha, Daniel, Rhoda, Joshua, Amos, and Emmanuel) for their prayers and support have kept me during my study. Also to my nephews, niece and in-laws especially my parent in-law, Mr and Mrs Bayero Peter, I take a bow!

Most of all, I wish to thank my wife, Lovinda, and my daugther, Shiayet, for without their support and patience this work would mean nothing!

Finally, I owe whatever I have achieve to the author and nisher of my faith for without God I could accomplish nothing!

This thesis was submitted to the Senate of Universiti Putra Malaysia and has been accepted as ful lment of the requirement for the degree of Master of Science.

The members of the Supervisory Committee were as follows:

Mohd Bakri Adam, Ph.D
Associate Professor
Faculty of Science
Universiti Putra Malaysia
(Chairperson)

Anwar Fitrianto, Ph.D
Senior Lecturer
Faculty of Science
Universiti Putra Malaysia
(Member)

BUJANG KIM HUAT, Ph.D.
Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date:

Declaration by Members of Supervisory Committee

This is to con rm that:

- the research conducted and the writing of this thesis was under the 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: ____________________

Signature: ____________________
Name of
Member of
Supervisory
Committee: ____________________

TABLE OF CONTENTS

	Page
ABSTRACT	i
ABSTRAK	iii
ACKNOWLEDGMENTS	v
APPROVAL	vi
DECLARATION	viii
LIST OF TABLES	xii
LIST OF FIGURES	xiv
LIST OF ABBREVIATIONS	xvi
CHAPTER	
1 INTRODUCTION	1
1.1 Background	1
1.1.1 Time Series Processes	2
1.1.2 Extreme Value Theory	6
1.1.3 Threshold Exceedance and the Generalised Pareto Distribution	9
1.1.4 Dependence in Extreme Value Distribution	11
1.2 Motivation of the Study	14
1.3 Objectives of the Thesis	14
1.4 Expected Outcome	15
1.5 Organisation of the Thesis	15
2 LITERATURE REVIEW	16
2.1 Clustering in Dependent Sequence	16
2.2 Extreme Value Analysis of Dependent Sequences	19
2.3 Bootstrapping Dependent Series	24
2.4 Resampling Scheme for Stationary Bootstrap	24
2.5 Residual Bootstrap Resampling Scheme	26
2.6 Summary	27
3 AR(1) MODELLING WITH EXTREME VALUE INNOVATIONS	28
3.1 Introduction	28
3.2 Methods	29
3.2.1 Generalized Extreme Value Distribution	29
3.2.2 Model Speci cation for Gumbel-AR(1)	31

3.2.3 Gumbel Distribution and the Maximum Likelihood Estimator 33
3.2.4 Declustering and the Extremal Index Estimation 35
3.3 Results and Discussion 36
3.3.1 Simulation Study 36
3.3.2 Real Data Example 49
3.4 Conclusion 53

4 MA(1) MODELLING WITH EXTREME VALUE INNOVATIONS 55
4.1 Introduction 55
4.2 Methods 55
4.2.1 Generalized Extreme Value Distribution 55
4.2.2 Model Speci cation for Gumbel-MA(1) 57
4.3 Results and Discussion 59
4.4 Conclusion 71

5 SUMMARY, CONCLUSION AND FUTURE WORK 72
5.1 Summary 72
5.2 Conclusion 72
5.3 Recommendations for Future Research 73

BIBLIOGRAPHY 74
APPENDICES 79
BIODATA OF STUDENT 79
LIST OF PUBLICATIONS 80

LIST OF TABLES

Table		Page
3.1	Probability values (p-values) of Augmented Dickey Fuller test for stationarity for sample size, n, with degree of dependence.	38
3.2	Gumbel-AR(1) lag 1 autocorrelation.	38
3.3	Biases of the estimate of the location parameter for Gumbel-AR(1).	38
3.4	Biases of the estimate of the location parameter for the residuals of Gumbel-AR(1).	39
3.5	Biases for the estimate of the scale parameter for Gumbel-AR(1).	41
3.6	Biases of the estimate of the scale parameter for the residuals of Gumbel-AR(1).	42
3.7	Threshold use to decluster Gumbel-AR(1) for n and .	42
3.8	Gumbel-AR(1) estimates of the extremal index for n and .	44
3.9	Declustered Gumbel-AR(1) for n and .	45
3.10	Biases for estimate of GPD_{gumbel} tted to GPD_{gumbel}-AR(1) and Gumbel-AR(1) for n and .	46
3.11	Biases for estimate of tted to GPD_{gumbel}-AR(1) and Gumbel-AR(1) for n and .	49
4.1	Probability values (p-values) of Augmented Dickey Fuller test for stationarity for sample size, n, with degree of dependence.	60
4.2	Gumbel-MA(1) lag 1 autocorrelation.	60
4.3	Biases of the estimate of the location parameter for Gumble-MA(1).	61
4.4	Biases of the estimate of the location parameter for the residuals of Gumbel-MA(1).	62
4.5	Biases of the estimate of the scale parameter for Gumbel-MA(1).	63
4.6	Biases of the estimate of the scale parameter for the residuals of Gumbel-MA(1).	64

4.7 Threshold use to decluster Gumbel-MA(1) for n and . 66

4.8 Gumbel-MA(1) estimates of the extremal index for n and 66

4.9 Declustered Gumbel-MA(1) for n and . 67

4.10 Biases for estimate of GPD_{gumbel} tted to GPD_{gumbel}-MA(1) and Gumbel-MA(1) for n and . 68

4.11 Biases for estimate of tted to GPD_{gumbel}-MA(1) and Gumbel-MA(1) for n and . 68

LIST OF FIGURES

Figure		Page
3.1	Biases of ^ for Gumbel-AR(1) with $\in (0; 1)$.	39
3.2	Biases of ^ for the residuals of Gumbel-AR(1) with $\in (0; 1)$.	40
3.3	Biases of ^ for Gumbel-AR(1) with $\in (0; 1)$.	40
3.4	Biases of ^ for the residuals of Gumbel-AR(1) with $\in (0; 1)$.	41
3.5	Quantile-Quantile plot for Gumbel-AR(1) for n = 3000; = 0:1.	43
3.6	Threshold use to decluster Gumbel-AR(1) with $\in (0; 1)$.	44
3.7	Extremal index of Gumbel-AR(1) with $\in (0; 1)$.	45
3.8	Number of independent observations after declustering Gumbel-AR(1) with $\in (0; 1)$.	46
3.9	Biases of $\hat{}_{GPDgumbel}$ for GPD_{gumbel}-AR(1) with $\in (0; 1)$.	47
3.10	Biases of $\hat{}_{GPDgumbel}$ for Gumbel-AR(1) with $\in (0; 1)$.	47
3.11	Biases of ^ for GPD_{gumbel}-AR(1) with $\in (0; 1)$.	48
3.12	Biases of ^ for Gumbel-AR(1) with $\in (0; 1)$.	48
3.13	Scatter plot of daily maximum wind speeds.	50
3.14	Goodness of t test for wind speed dataset to the Gumbel distribution.	51
3.15	Sample ACF of daily maximum wind speed series.	52
3.16	Sample PACF of daily maximum wind speed series.	52
3.17	Quantile-Quantile plot for wind speed data set.	53
4.1	Biases of ^ for Gumbel-MA(1) with $\in (0; 1)$.	61
4.2	Biases of ^ for the residuals of Gumbel-MA(1) with $\in (0; 1)$.	62
4.3	Biases of ^ for Gumbel-MA(1) with $\in (0; 1)$.	63

4.4 Quantile-Quantile plot for Gumbel-MA(1) for n = 5000; = 0:2: 64

4.5 Biases of ^ for the residuals of Gumbel-MA(1) with $\in (0; 1)$. 65

4.6 Threshold use to decluster Gumbel-MA(1) with $\in (0; 1)$. 65

4.7 Extremal index of Gumbel-MA(1) with $\in (0; 1)$. 66

4.8 Number of independent observations after declustering Gumbel-MA(1) with $\in (0; 1)$. 67

4.9 Biases of $\hat{}_{GPDgumbel}$ for GPD_{gumbel}-MA(1) with $\in (0; 1)$. 69

4.10 Biases of $\hat{}_{GPDgumbel}$ for Gumbel-MA(1) with $\in (0; 1)$. 69

4.11 Biases of ^ for GPD_{gumbel}-MA(1) with $\in (0; 1)$. 70

4.12 Biases of ^ for Gumbel-MA(1) with $\in (0; 1)$. 70

LIST OF ABBREVIATIONS

ADF	Augmented Dickey Fuller
AR	Autoregressive
ARCH	Autoregressive Conditional Heteroskedasticity
EVT	Extreme Value Theory
GEV	Generalised Extreme Value
GPD	Generalised Pareto Distribution
GPD_{gumbel}	Gumbel-Generalised Pareto Distribution
MA	Moving Average
ML	Maximum Likelihood
MLE	Maximum Likelihood Estimate
MSE	Mean Squared Error
POT	Peak Over Threshold

CHAPTER 1

INTRODUCTION

This chapter presents a general introduction to the theory of extreme values. Basic concept of time series, extreme value theory, and dependence in extreme value distribution are highlighted. The objectives and motivation of the thesis, expected outcome and organisation of the thesis are all discussed.

1.1 Background

Extreme Value Theory (EVT) is one of the most important statistical disciplines and has been applied in various elds, from nance to hydrology to atmospheric chemistry to environmental sciences to nancial econometrics. An important feature of extreme value analysis is the objective of assessing the extremal behaviour of random variables and quantifying the stochastic behaviour of maxima or minima of independent and identically distributed random variables. The distributional properties of extreme observations (maximum and minimum) and exceedances over high (or low) thresholds are determined by the upper and lower tails of the underlying distribution. Extreme value analysis necessitate the estimation of the probability of events that are more extreme than any that have already been observed, Coles (2001). The theory underlying EVT methods arises from studying block maxima or threshold exceedances.

Let $X_1; \ldots; X_n$ be a random samples from a given probability distribution, to understand extreme behavior, one would intuitively inspect the value of the maximum of the sample. If there are repeated samples of random variables, and we look at the maxima of each of these samples, then one might begin to gain an understanding of the behavior found in the tail of the distribution. The modelling of extreme value has attracted tremendous interest in many di erent elds. Speci cally, the Generalised Pareto distribution is a very popular extreme value model, which can give a good model for the upper tail providing reliable extrapolation for quantiles beyond a reasonably high threshold. An important challenge in application of such extreme value models is the threshold choice beyond which an observation can be considered as extreme that is, the asymptotically justi ed extreme value models will give su cient approximation to the tail of the distribution.

To analyse the statistical properties of maxima or minima, previous studies usually focus on independent and identically distributed random variables, Coles (2001). However, due to temporal dependence, short-range dependence of extremes which leads to cluster of observations usually arises in extreme value theory. The dependence in extreme values leads to breakdown of the independence assumption in EVT and an appropriate approach is thus needed to analyse and characterise the extremes of dependent series.

The extension of time series to extreme values can be achieved by inducing time-dependence in the underlying state of the extreme value process. One approach to capture dependence in EVT is to consider a time series process for the model parameters of the extreme value distribution using a state space representation, Huerta and Sanso (2007).

We study and characterise the impact of dependence on the parameter estimates of stationary Gumbel autoregressive and Gumbel moving average processes of an extreme value process with varying degree of dependence induce through the linear ltering of Gumbel distributed innovation or error term. The innovations of the time series process basically make up the variability observed in part of the system when it moves from one time period to another.

The Gumbel distribution is a member of the Generalised Extreme Value (GEV) distribution. The GEV distribution has three parameters: a location parameter, ; a scale parameter, ; and a shape parameters,. Generalised extreme value distribution commonly use for the analysis of maxima of some large set of random variables has cumulative distribution function given as

$$G(x) = \exp\left\{-\left[1 + \left(\frac{x -}{}\right)\right]^{-\frac{1}{\xi}}\right\} \qquad (1.1)$$

where $> 0, \in R, \in R$ and $1 + (\frac{x-}{}) > 0$:
When $\to 0$ in Eq. (3.1) the GEV distribution lead to the Gumbel family of distribution with light-tailed normalized maxima given as

$$G(x) = \exp\left\{-\exp\left[-\left(\frac{x -}{}\right)\right]\right\}; -\infty < x < \infty; \qquad (1.2)$$

Coles (2001).

1.1.1 Time Series Processes

A time series is a sequence of observations observed sequentially in time, Box et al. (2008). The ordering is usually through time particularly in terms of some equally spaced time intervals. The time interval could be every second, every minute, hourly, daily, weekly, monthly, quarterly and annually. An important feature of a time series is that successive observations are usually dependent. The behaviour of this dependence among successive observations is of practical importance. Indeed, it is this dependent from one time period to another that is exploited in making reliable forecast. It is imperative to distinguish between a time series process and a time series realisation. The observed time series is an actual realisation of an underlying time series process. By a realisation we mean a sequence of observed data points and not just a single observation, while a time series process refers to the underlying mechanism giving rise to an observed time series data.

Stationarity, autocorrelation and autocovariance are important concepts in the study of time series. The autocorrelation function (ACF) assess the level of dependence in a time series process and is use to choose a model for the observations that reflects this dependence, Brockwell and Davis (2002). Let $\{Y_t\}$ be realised values of a stationary time series process, the estimate of the autocorrelation function will provide an estimate of the ACF of $\{Y_t\}$ which may suggest which of the various stationary time series model is suitable in explaining the dependence in the observations. For any stationary process $\{Y_t\}$, the mean, $E\{Y_t\} = \mu$, and variance written as

$$\begin{aligned} var\{Y_t\} &= E\{(Y_t - \mu)^2\} \\ &= E\{(y_t - \mu)(y_t - \mu)\} \end{aligned}$$

are constant for all t and the covariance, denoted by $cov(y_t, y_s)$ are functions of only the time difference $|t - s|$.

In order to make statistical inference about the structure of a stochastic process on the basis of a finite record of the process, an important assumption is that of stationarity. If the mean and variance of a time series process do not change with time and if strictly periodic variations have been removed, then, the time series process is said to be stationary and the process is assumed to be in statistical equilibrium, Chatfield (2003). Stationarity requirement may seem too restrictive, however, most non-stationary time series process that are encountered in practice can be transformed into a stationary time series using relatively simple operations, Pankratz (1983). If a process is not stationary, we cannot get useful estimates of the parameters of the process. Non-stationarity implies that the mean is different each time period, and the variance is also not constant through time. A model that violates the stationarity restrictions will produce forecasts whose variance increases without limit, an undesirable result. An examination of times series plot for wandering behaviour could suggest non-stationarity. Also, examining the sample autocorrelation function and observing that the autocorrelations do not die out quickly could suggest non-stationarity. A more formal test for testing stationarity is the Augmented Dickey Fuller test. It is important to note that the non-stationary components such as trend could be of more interest than the stationary residuals depending on the objective of a researcher. We shall distinguish between two forms of stationarity namely, strictly stationary and weakly stationary processes.

- Strictly stationary process

  A time series process $\{Y_t\}, \; t = 0, \pm 1, \ldots,$ is said to be strictly stationary if it is defined by the condition that $(Y_1, Y_2, \ldots, Y_n)$ and $(Y_{1+k}, Y_{2+k}, \ldots, Y_{n+k})$ have the same joint distributions for all integers k and $n > 0$, Brockwell and Davis (2002).

- Weakly stationary process

  A times series process $\{Y_t\}, \; t = 0, \pm 1, \ldots, \pm n$ is said to be weakly (or second-order or covariance) stationary if the mean and variance are constant over time. In other words, $E(Y_t) = \mu$ and $E\{(Y_t - \mu)^2\} = \sigma^2$ for all t and the autocovariance, $\text{cov}(Y_t; Y_{t+k})$, depend only on the distance k and not on the time period, Box et al. (2008). If $\{Y_t\}$ is a strictly stationary time series and $E(Y_t^2) < \infty$ for all t, Brockwell and Davis (2002) state that $\{Y_t\}$ is also covariance stationary. Since it is often relatively easier to check for the first two moments, we shall work with weakly stationary time series in this thesis.

One important example of a stationary process is the so-called White Noise (WN) process, which is defined as a sequence of independent identically distributed random variable. This is often time called strict white noise and the phrase uncorrelated white noise is used when successive values are merely uncorrelated, rather than independent, Chatfield (2003). The importance of a white noise process is derive not from the fact that it is an interesting process itself but from the fact that many processes can be constructed from the white noise process. A white noise process is distributed with mean zero and finite variance, σ^2, Brockwell and Davis (2002).

Time series models have different forms and represent varying stochastic processes. There are three broad classes of practical importance in modelling variations in the level of a process and they are the Autoregressive (AR) models, the Integrated (I) models, and the Moving Average (MA) models. These three models are each linearly dependent on past observations. A combination of these models gives rise to different models. There are also other types of non-linear time series models, such as Autoregressive Conditional Heteroskedasticity (ARCH) that represent fluctuations of variance over time (heteroskedasticity). According to Brockwell and Davis (2002), autoregressive moving average models give a basic structure for studying stationary processes, in this thesis, we therefore study autoregressive process and moving average process of order one respectively.

Autoregressive Process

An autoregressive model of order p is obtain by regressing the series Y_t on its lagged values $Y_{t-1}; Y_{t-2}; \ldots; Y_{t-p}$. It can be represented by equation of the form

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \phi_3 Y_{t-3} + \ldots + \phi_p Y_{t-p} + \varepsilon_t \qquad (1.3)$$

and can be written in a more compact form as $\phi(B)Y_t = \varepsilon_t$ where B is the back-shift operator, ε_t denotes white noise (random shock) and

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p.$$

The positive integer p is the order of the model and is determined from the data using a plot of the autocorrelation and partial autocorrelation functions respectively. The sequence $\phi_1; \phi_2; \ldots$ may be finite or infinite. If this sequence is finite or infinite but absolutely summable in the sense that $\sum_{j=0}^{\infty} |\phi_j| < \infty$, a linear filter generated time series is said to be stable. The condition $\sum_{j=0}^{\infty} |\phi_j| < \infty$ implies $\sum_{j=0}^{\infty} \phi_j^2 < \infty$ and therefore ensure the series converges in mean square, Brockwell and Davis (2002).

A time series $\{Y_t\}$ is said to be governed by a rst-order AR process if the current value of the series Y_t can be expressed as a linear function of the previous values of the series plus the random shock "$_t$. The AR(1) process can be written as

$$Y_t = {}_1 Y_{t-1} + "_t \qquad (1.4)$$

and this can be rewritten as $(B)Y_t = "_t$ where $(B) = (1 - {}_1 B)$: For the AR(1) to be stationary, the root of $(1 - {}_1 B) = 0$ must be greater than 1 in absolute value, Pankratz (1983).The AR(1) process is sometime called the Markov process because the value of Y_t is completely determined by the knowledge of Y_{t-1}. The autocovariance of Y_t at lag k denoted by $_k$, is a covariance between Y_t and Y_{t-k} and therefore de ned as

$${}_k = cov(Y_t; Y_{t-k}):$$

The ACF of AR(1) decays exponentially while the Partial Autocorrelation Function (PACF) cuts o after lag one. For stationarity assumption to hold for AR(1) process, we must have that $|{}_1| < 1$. Since $|{}_1| < 1$, the autocorrelation function is an exponentially decreasing curve as the lag k increases. It follows therefore that if $0 < {}_1 < 1$, autocorrelation function decays exponentially to zero, and if $-1 < {}_1 < 0$ the autocorrelation function also decays exponentially to zero but oscillates in sign, Box et al. (2008) .

Moving Average Process

The moving average process expresses the current value of $\{Y_t\}$ as a linear function of the current and previous errors, "$_t$ and "$_{t-j}$; $j = 1; 2; :::; q$:A moving average process of order q, MA(q), is of the form

$$Y_t = "_t - {}_1"_{t-1} - {}_2"_{t-2} - ::: - {}_q"_{t-q} \qquad (1.5)$$
$$Y_t = (B)"_t:$$

As with an autoregressive process, the random shocks in a moving average process are assumed to be normally and independently distributed with mean zero and constant variance $^2_"$: Also, q is a positive integer indicating the order of the model and the order can be determined by plotting the autocorrelation and partial autocorrelation functions respectively. As with the AR process, we shall limit the scope of this study to a rst order moving average process, MA(1).

The rst order moving average, MA(1), process is of the form

$$y_t = "_t - "_{t-1} \qquad (1.6)$$
$$= (1 - B)"_t$$

where "$_t$ is a zero-mean white noise process with constant variance ^{2}: Contrary to the behaviour of the autocorrelation and partial autocorrelation functions plots of the AR(1), the ACF of MA(1) cuts o after lag one while the PACF tails o exponentially in one of two forms depending on the signs of

It is important to note that while an AR process is require to satisfy the stationarity conditions, a MA process is require to satisfy the invertibility condition, Pankratz (1983). For the MA process to be invertible, must lie in the range$| \ |<1$. However, the process is ofcourse stationary for all values of A non-invertible MA model implies that the weights placed on past realisations of the time series do not decline as we move backward into past observations; but rational thinking suggest that larger weights should be attached to more recent observations than to past observations, Brockwell and Davis (2002). Invertibility condition ensures that larger weights are placed on recent observations than past observations. Invertibility also ensures a unique association between processes and theoretical autocorrelation functions.

1.1.2 Extreme Value Theory

The objective of extreme value theory is to formulate probabilistic results that allows characterisation in the tail behaviour of any probability distribution without requiring knowledge of the form of the underlying probability distribution because, the cumulative distribution function of extremes of any distribution approaches a known distribution asymptotically as the sample size increases, Makkonen (2008). Modelling observed data extremes and making generalisation about the probable recurrence of these events is the objective of an extreme value analysis.

Let $Y_1; Y_2; :::; Y_n$ be a sequence of independent identically distributed random variables with distribution function F. The theory of classical extreme value is concern with the behaviour of $M_n = \max\{Y_1; Y_2; :::; Y_n\}$: As $n \to \infty$; the characteristics of M_n is of particular importance. Let

$$
\begin{aligned}
\Pr\{M_n \le z\} &= \Pr\{Y_1 \le z; Y_2 \le z; ::::; Y_n \le z\} \\
&= \prod_{j=1}^{n} \Pr\{Y_j \le z\} \\
&= \{P(Y_j \le z)\}^n \\
&= \{F(z)\}^n;
\end{aligned}
$$

for all values of n. In practice, the distribution function of F is unknown, Makkonen (2008). One approach use in determining F is by employing standard statistical techniques to the observed data and this lead to an estimate for the distribution of M_n. This approach could however give rise to small discrepancies in the estimate of F, which could give rise to large variations for F^n. Alternatively, extreme data could be use in estimating an approximate families of models for the distribution function F^n. However, M_n has a degenerate limiting distribution because it converges with probability 1 to the upper endpoint of F, Coles (2001). The implication of this is that as $n \to \infty$, to consider the behaviour of F^n would not be enough. This challenge is avoided by considering a linear renormalisation of M_n:

$$M_n^* = \frac{M_n - b_n}{a_n};$$

where $\{a_n > 0\}$ and $\{b_n\}$ are sequences of constants. Choosing $\{a_n > 0\}$ and $\{b_n\}$ appropriately stabilise both the location and scale parameters of M_n^* as n increases and this avoid the di culties that arise with M_n. Therefore, with appropriate choices of $\{a_n > 0\}$ and $\{b_n\}$, we seek limit distributions for M_n^* rather than M_n. The proposition below by Embrechts et al. (1997) gives the continuity conditions on F which ensures that the limit of $\Pr(M_n \leq u_n)$ as $n \to \infty$ for the independent and identically distributed random variable exist for appropriate constants $u_n = a_n + b_n$

Proposition 1.1 :
For given $\in [0;\infty]$ *and a sequence* (u_n) *of real numbers the following are equivalent*

$$n\overline{F}(U_n) \to \ ;$$

$$P(M_n \leq u_n) \to e^{-}$$

where $\overline{F} = 1 - F$ *is the tail of the distribution function* F *and it guarantee the existence of a limit distribution for a heavy-tailed case.*

Embrechts et al. (1997) also state that a better understanding of the order of magnitude of maxima is given by weak convergence results for centered and normalised maxima. The extremal types theorem give the entire range of the possible limit distributions for M_n and an aspect of this result were proved by Fisher and Tippett (1928). The extremal types theorem by Coles (2001) state:

Theorem 1.1 :(Extremal Types Theorem)
If there exist sequences of constants $\{a_n > 0\}$ *and* $\{b_n\}$ *such that*

$$\Pr\left\{\frac{M_n - b_n}{a_n} \leq z\right\} \to G(z) \text{ as } n \to \infty; \tag{1.7}$$

where G *is a non-degenerate distribution function, then* G *belongs to one of the following families:*

$$G(z) = \exp\left\{-\exp\left[-\left(\frac{z - }{}\right)\right]\right\}; -\infty < z < \infty \tag{1.8}$$

$$G(z) = \begin{cases} 0; & z \leq \ ; \\ \exp\left[-\left(\frac{z-}{}\right)^{-}\right]; & z > \ ; \end{cases} \tag{1.9}$$

$$G(z) = \begin{cases} \exp\left\{-\left[-\left(\frac{z-}{}\right)\right]\right\}; & z < \ ; \\ 0; & z \geq \ ; \end{cases} \tag{1.10}$$

for parameters > 0, b *and, in the case of families of Eq. (1.8) and Eq. (1.9),* > 0.

The extremal types theorem in other words states that, if $\{a_n\}$ and $\{b_n\}$ are suitable sequences that stabilise M_n, then M_n^* converge in distribution to one of the three families of distribution stated in Eq. (1.7), Eq. (1.8) and Eq. (1.9) respectively. This three distributions are known as the *Gumbel*, *Frechet* and *Weibull* families respectively and collectively they are termed the *extreme value distributions*. Each of the three families has a location parameter; „ and a scale parameter; , respectively while the Frechet and Weibull families have an additional shape parameter; One remarkable feature of this result is that the Gumbel, Frechet and Weibull families of extreme value distributions are the only possible limits for the distributions of M_n^*, irrespective of the distribution F of the population therefore providing an extreme value counterpart of the central limit theorem.

Generalized Extreme Value Distribution

The Generalised Extreme Value (GEV) distribution is use traditionally for the analysis of extreme values, Huerta and Sanso (2007). Theorem (1.1) give rise to three types of limits distribution with distinct forms of tail behaviour for the distribution function F of X_i. The upper end-point, z_+, of the Weibull distribution is nite, while the upper end-point, z_+, for both Frechet and Gumbel distributions are in nite. It is important to note that the density of G(z) decays exponentially for the Gumbel distribution but polynomially for the Frechet distribution corresponding to di erent rate of decay behaviour in the tail of F. Hence in practice, this correspond to different representation of extreme behaviour. Since Gumbel distribution ts well in many practical situations, Kotz and Nadarajah (2000) and Cordeiro et al. (2012), this thesis therefore focus on Gumbel family of the GEV distribution. According to Coles (2001), the uncertainty involve in choosing which of the three families is most suitable for a particular data can be avoided by a reformulation of Theorem (1.1) to a modi ed form

Theorem 1.2 :(Modified Extremal Types Theorem)
If there exist sequences of constants $\{a_n > 0\}$ *and* $\{b_n\}$ *such that*

$$\Pr\left\{\frac{M_n - b_n}{a_n} \leq z\right\} \to G(z) \text{ as } n \to \infty; \qquad (1.11)$$

for a non-degenerate distribution function G, *then* G *is a member of the GEV family*

$$G(z) = \exp\left\{-\left[1 + \quad \left(\frac{z - \quad}{\quad}\right)\right]^{-\frac{1}{\xi}}\right\} \qquad (1.12)$$

defined on $\{z : 1 + \frac{(z- \ )}{\ } > 0\}$, *where* $-\infty < \ < \infty$, $\ > 0$ *and* $-\infty < \ < \infty$:

The subset of Eq. (1.12) with $\ = 0$, which is interpreted as the limit of G(z) as $\ \to 0$, is known as the Gumbel distribution with normalized maxima from GEV with light-tailed given as in Eq. (1.8).

The Frechet and Weibull families of extreme value correspond to the case > 0 and < 0 respectively. Assuming Eq. (1.11), then

$$\Pr\left\{\frac{M_n - b_n}{a_n} \leq z\right\} \approx G(z)$$

for large n. Similarly,

$$\begin{aligned}\Pr\{M_n \leq z\} &\approx G\left\{\frac{(z - b_n)}{a_n}\right\}\\ &= G^*(z);\end{aligned}$$

where G^* is also a member of the Generalised Extreme Value family, hence resolving the challenge that the normalising constants $\{a_n > 0\}$ and b_n will be unknown in practice, Coles (2001).

1.1.3 Threshold Exceedance and the Generalised Pareto Distribution

The Generalised Pareto Distribution (GPD) is a dual parameter family of distribution use in modelling exceedances over a given threshold u, The GPD has scale parameter; and shape parameter; . Maximum Likelihood Estimates (MLE) of the parameters are asymptotically normal and asymptotically e cient in many cases and are therefore preferred in the estimation of both and , Grimshaw (1993). We proceed by stating the generalised Pareto distribution and show how generalised extreme value distribution and Gumbel distribution are related to the GPD and the Gumbel distributed GPD (GPD_{gumbel}) respectively. We derive the MLE for GPD_{gumbel} in Section (3.2.3). Coles (2001) give the main result of GPD as follows:
Let $X_1; X_2; \cdots; X_n$ be a sequence of independent and identically distributed random variables with common distribution function, F, and let $M_n = \max\{X_1; \cdots; X_n\}$ so that

$$Pr\{M_n\} \approx G(z)$$

where

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

for some , > 0 and . Then, for any reasonable large threshold u, the distribution function of $(X - u)$, conditional on $X > u$, is approximately

$$H(y) = 1 - \left(1 + \frac{y}{\sim}\right)^{-\frac{1}{\xi}} \qquad (1.13)$$

de ned on $\{y : y > 0\}$ and $(1 + \frac{y}{\sim}) > 0$, where $\sim = \quad + \quad (u - \quad)$.

For the Gumbel distributed generalised Pareto distribution, GPD_{gumbel}; we take the limit of Eq. (1.13) as $\to 0$; leading to

$$H(y) = 1 - \exp(-y=\sim); \; y > 0 \tag{1.14}$$

which correspond to the exponential distribution with parameter 1=

The family of distribution de ned by Eq. (1.13) is the generalised Pareto family of distributions while Eq. (1.14) is the Gumbel-distributed generalised Pareto distribution. Since Gumbel distribution is a member of the family of the GEV distribution, to show the transformation from Gumbel distribution to the GPD, we commence with the the GEV. Lets denote an arbitrary term in the sequence $X_1; X_2; \cdots; X_n$ by X. Coles (2001) describe the behaviour of extreme events as

$$\begin{aligned} \Pr\{X > u + z | X > u\} &= \frac{Pr(X > u + z; X > u)}{Pr(X > u)} \\ &= \frac{Pr(X > u + z)}{Pr(X > u)} \\ &= \frac{1 - F(u + z)}{1 - F(u)} \end{aligned} \tag{1.15}$$

but we know that

$$\begin{aligned} \Pr\{M_n\} &= F^n(X) \approx G(z) \\ &= \exp\left\{ - \left[1 + \left(\frac{X -}{} \right) \right]^{-\frac{1}{\xi}} \right\}; \end{aligned}$$

with parameters , and : Therefore,

$$n \log F(X) \approx - \left[1 + \left(\frac{X -}{} \right) \right]^{-\frac{1}{\xi}} : \tag{1.16}$$

For reasonable large values of X, a Taylor series expansion implies that

$$\log F(X) \approx -\{1 - F(X)\}$$

and substituting into Eq. (1.16) gives

$$\begin{aligned} n(-\{1 - F(u)\}) &\approx - \left[1 + \left(\frac{u -}{} \right) \right]^{-\frac{1}{\xi}} \\ 1 - F(u) &\approx \frac{1}{n} \left[1 + \left(\frac{u -}{} \right) \right]^{-\frac{1}{\xi}} \end{aligned} \tag{1.17}$$

for reasonable large u: Also, for $z > 0$;

$$1 - F(u + z) \approx \frac{1}{n}\left[1 + \quad \left(\frac{u + z -}{\quad}\right)\right]^{-\frac{1}{\xi}} \tag{1.18}$$

substituting Eq. (1.17) and Eq. (1.18) into Eq. (1.15) then gives

$$\begin{aligned}
Pr\{X > u + z | X > u\} &= \frac{1 - F(u + z)}{1 - F(u)} \\
&\approx \frac{\frac{1}{n}\left[1 + \quad \left(\frac{u+z-}{\quad}\right)\right]^{-\frac{1}{\xi}}}{\frac{1}{n}\left[1 + \quad \left(\frac{u-}{\quad}\right)\right]^{-\frac{1}{\xi}}} \\
&= \left[\frac{1 + \quad \left(\frac{u-}{\quad}\right) + \quad \frac{z}{\quad}}{1 + \quad \left(\frac{u-}{\quad}\right)}\right]^{-\frac{1}{\xi}} \\
&= \left[\frac{1 + \quad \left(\frac{u-}{\quad}\right)}{1 + \quad \left(\frac{u-}{\quad}\right)} + \frac{\frac{z}{\quad}}{1 + \quad \left(\frac{u-}{\quad}\right)}\right]^{-\frac{1}{\xi}} \\
&= \left[1 + \frac{z}{\quad + \quad (u - \quad)}\right]^{-\frac{1}{\xi}} \\
&= \left[1 + \frac{z}{\sim}\right]^{-\frac{1}{\xi}}
\end{aligned}$$

hence, the GPD_{gumbel} as $\quad \to 0$ is,

$$H(y) = 1 - \exp(-y = \sim); \; y > 0$$

where $\sim = \quad + \quad (u - \quad)$ and z as use here is the excess above a given threshold.

The derivation of the GPD_{gumbel} from the Gumbel family of distribution makes the dependence of the Gumbel distributed generalised Pareto distribution parameters on the threshold, u, obvious. The parameters of the GPD_{gumbel} of threshold excesses are determined by those of the corresponding Gumbel distribution and in particular, the shape parameter of the GPD_{gumbel} is equal to that of the associated Gumbel distribution.

1.1.4 Dependence in Extreme Value Distribution

So far, we have focused on independent identically distributed random variables to consider the statistical properties of their maxima. However, in dependent sequence, extremal events in practice often occur in clusters, Markovich (2014). The basic assumption governing an extreme value analysis is that the sequence of random variables are independent and identically distributed and the maxima of the samples converge in distribution to a random variable which have a non-degenerate extreme value distribution G, Coles (2001). These basic assumptions are not met in time series data because time series data are usually dependent. As a result of temporal dependence, short-range dependence of extremes leading to cluster of observations usually arises in extreme value theory and this leads to the breakdown of the independence assumption, Chavez-Demoulin and Davison (2012). In order to analyse the extreme of stationary series, the dependence in the series must be consider, Eastoe and Tawn (2012).

Dependence takes many di erent forms in stationary series and it is not unusual to assume a condition that limits the extent of long range dependence at an extreme levels. This will ensure that for a stationary time series $X_1; X_2; \ldots; X_n$, two events $X_i > u$ and $X_j > u$ are approximately independent, provided the threshold u, is su ciently large and time points i and j are reasonable far apart, Coles (2001). Many stationary time series processes satisfy this property. Thus, one can focus on the e ect of short-range dependence by eliminating long-range dependence at extreme levels. For a stationary time series model, exceedances often times occur in clusters. Instead of trying to develop methods which will specify the joint probability distribution of exceedances, the data are more often declustered, Ferro and Segers (2003) and Ledford and Tawn (2003). The declustering schemes involve setting a threshold such that the resulting sequence of exceedances contains approximately independent observations.

Extremal Index

An important parameter in dependent series that measures the degree of clustering of extreme values is the extremal index, Reiss and Thomas (2007). Extremal index make it possible to characterises changes in the sequences of an extremal distribution as a result of serial dependence, Embrechts et al. (1997). For dependent random variables, Proposition (1.1) is violated because it require observations to be independent identically distributed, and therefore, the distribution of the maxima M_n is not determined by F alone, but rather from the complete distribution of the dependent series. A comparative approximate relationship can be derived as

$$P(M_n \leq z) \approx F^{n} (z) \geq F^{n}(z)$$

for large n, where $\in [0;1]$ is called the *extremal index*. Recall that for an independent case the following holds

$$n\overline{F}(U_n) \to \;\; ; 0 < \;\; < \infty$$

if and only if

$$P(M_n^* \leq U_n) \to e^{-} \; :$$

A formal de nition of extremal index is given by Embrechts et al. (1997) as:

Definition 1.1 :(Extremal Index)
Let $(Y_j); -\infty < j < \infty$; *be a strictly stationary sequence and* $\in (0;1)$ *a non-negative number. Assume that for every* > 0 *there exists a sequence* (u_n) *such that*

$$\lim_{n\to 0} n\overline{F}(u_n) = \qquad (1.19)$$

and

$$\lim_{n\to\infty} \Pr(M_n \leq u_n) = e^{-} : \qquad (1.20)$$

Then is called the extremal index *of the sequence* (Y_j).

Unless the extremal index, , is equal to one, the limit distributions for stationary and independent sequences are not the same. Provided > 0 exist, then, for any sequence of real numbers $\in 0;\infty$, and u_n, the relations in the extremal index de nition and $\Pr(M_n^* \leq u_n) \to e^{-}$ are equivalent. Thus from the above de nition, the relationship between the distribution of maximum and the exceedances probability is

$$\begin{aligned} \Pr(M_n \leq u_n) &\approx e^{-} \\ &\approx e^{- n\overline{F}(u_n)} \\ &= \left(e^{-\overline{F}(u_n)}\right)^n \\ &\approx \left(1 - \overline{F}(u_n)\right)^n \\ &= F^n(u_n); \end{aligned} \qquad (1.21)$$

when u_n is large, then $\overline{F}(u_n) \approx 0$: If we ignore the extremal index, , in the data, we risk underestimating the quantile of F, Chavez-Demoulin and Davison (2012). Independent identically distributed variables has extremal index = 1, since Y_j here are independent. The case = 0 entails that sample maxima of a dependent process are of smaller order than sample maxima of the independent sequnce, see Beirlant et al. (2006) for a detail proof of extremal index.

Modelling Extremes in Stationary Processes

There are two popular approach use in statistical modeling of extremes. The Block Maxima and the Peaks over Threshold methods. The block maxima method entail observing the maximum values of a series over a number of blocks. The observed block maxima are modelled using an extreme value distribution with distribution function given in Eq. (1.12). The use of the extreme value distributions is motivated by the fact that they are the only limit of linearly normalised maxima and they are max-stable, Rootzæn and Tajvidi (2006). Max-stability implies that a change in the size of the block only leads to corresponding change in both the location and scale parameters respectively in the distribution. However, according to Coles (2001), modelling only block maxima is an imprudent way of modelling extreme values provided other observations on extremes are available. Therefore, if a complete time series is available, avoiding the procedure of blocking will lead to better use of the data. Typically, the identification of independent clusters and the declustering scheme used has a significant effect on the parameter estimates of cluster characteristics when inference is to be made for clusters of extreme in a time series process, Ferro and Segers (2003).

The POT approach uses all observed values which are greater than a chosen threshold. These extreme observations are then assumed to be from the GPD given in Eq. (1.13). Rootzæn and Tajvidi (2006) state that the choice of the generalised Pareto distribution in modeling threshold exceedances is motivated by the following:

- The asymptotic threshold (as the threshold approaches the endpoint of the distribution) of the distribution of a scale normalised exceedances approaches a generalised Pareto distribution if and only if the block maxima distribution converges to an extreme value distribution.
- The generalised Pareto distribution is the only distribution for which the conditional distribution of exceedances is a scale transformation of the original distribution.

Rootzæn and Tajvidi (2006) state that the peak over threshold method often times give better estimation precision than the block maxima approach since it uses more of the data.

1.2 Motivation of the Study

The motivation for this study arises from an empirical observation that dependent sequence of observations from finance, environmental sciences, hydrology, etc. occur in cluster in the presence of an extremal event. Specifically, it has been identified in recent years that independent identical distributed Gaussian random variables is not suitable for modelling extreme returns of stocks observed during financial crisis, Robert (2008). However, many real life applications of extreme values provide one with dependent and stationary data rather than the conventional independent identically distributed data use in analysing extreme values. Therefore, it is important to study and understand the behaviour of the parameter estimates of the extreme value process under the stationarity assumption of the observed sequence in the presence of different level of weak dependence. The need to make general statements that characterise dependence in extreme value analysis over a range of sample sizes with varying degree of dependence motivates this study. Such general characterisation is useful in understanding the underlying mechanism giving rise to an extreme event in physical processes.

1.3 Objectives of the Thesis

We propose stationary Gumbel-AR(1) and Gumbel-MA(1) models and generate sample sizes with varying degrees of dependence. The objective of the thesis therefore are to:

1. Evaluate the impact of dependence on the parameter estimates of Gumbel-AR(1) and Gumbel-MA(1) models.

2. Evaluate the impact of dependence on the degree of clustering of the Gumbel-AR(1) and Gumbel-MA(1) models .

3. Fit Gumbel-AR(1) and Gumbel-MA(1) models to $\text{GPD}_{\text{gumbel}}$ and then evaluate the performance of the parameter estimates tted to the threshold exceedances and the original series.

1.4 Expected Outcome

The general expectation of this thesis is to evaluate the performance of the parameter estimates of an autoregressive and moving average processes with Gumbel distributed innovation arising from a linear lter operation for di erent sample sizes with varying degrees of dependence. We t the threshold exceedances of the AR and MA processes to $\text{GPD}_{\text{gumbel}}$ and to the declustered series and then assess the performance of the parameter estimates of the Gumbel generalised Pareto distribution. Results obtain will give a general characterisation of extreme values parameter estimates of AR and MA processes with Gumbel distributed innovation. The parameter estimates of Gumbel distributed generalised Pareto distribution tted to threshold exceedances before and after declustering the AR and MA processes for the di erent sample sizes over a varying degree of dependence will also be evaluated.

1.5 Organisation of the Thesis

This thesis is divided into ve chapters. Chapter 1 give the theoretical background by discussing the basic concept of time series, extreme value theory and dependence in extreme value distributions. Chapter 2 reviews relevant literatures on dependence in extreme value distributions. Chapter 3 presents simulation study, results and discussion of the AR(1) model with extreme value innovations. In Chapter 4, we presents the simulation study, results and discussion for the MA(1) model with extreme value innovations. In Chapter 5, we presents the general summary of results, conclusion and recommendations for future research.

BIBLIOGRAPHY

Ancona-Navarrete, M. A. and Tawn, J. A. 2000. A Comparison of Methods for Estimating the Extremal Index. *Extremes* 3 (1): 5{38.

Beirlant, J., Goegebeur, Y. and Segers, J. 2006. *Statistics of extremes: theory and applications*. Chichester: Wiley.

Berkowitz, J. and Kilian, L. 2000. Recent Developments in Bootstrapping Time Series. *Econometric Reviews* 19 (1): 1{48.

Bhattacharya, B. 2008. The Extremal Index and the Maximum of a Dependent Stationary Pulse Load Process Observed above a High Threshold. *Structural Safety* 30 (1): 34{48.

Box, G. E., Jenkins, G. M. and Reinsel, G. C. 2008. *Time Series Analysis: Forecasting and Control*. Wiley series in probability and statistics, NJ: John Wiley & Sons.

Brockwell, P. J. and Davis, R. A. 2002. *Introduction to Time Series and Forecasting*. New York, NY: Springer New York.

Buhlmann, P. 2002. Bootstraps for Time Series. *Statistical Science* 17 (1): 52{72.

Cechet, R. and Sanabria, L. 2010. Extreme-value time-series analysis of Australian Region A gust wind speeds to examine instrument bias. In *IOP Conference Series: Earth and Environmental Science*, 2016.

Chat eld, C. 2003. *The Analysis of Time Series: An Introduction*. 6th edn. Boca Raton, FL: Chapman & Hall/CRC.

Chavez-Demoulin, V. and Davison, A. 2012. Modelling Time Series Extremes. *REVSTAT-Statistical Journal* 10 (EPFL-ARTICLE-180506): 109{133.

Chernick, M. R., Hsing, T. and McCormick, W. P. 1991. Calculating the extremal index for a class of stationary sequences. *Advances in applied probability* 23 (4): 835{850.

Coles, S. 2001. *An Introduction to Statistical Modeling of Extreme Values*. , vol. 208. London, Britain: Springer.

Cordeiro, C. and Neves, M. 2006. The Bootstrap Methodology in Time Series Forecasting. *Proceedings of CompStat2006(J. Black and A. White, Eds.), Springer Verlag* 1067{1073.

Cordeiro, G. M., Nadarajah, S. and Ortega, E. M. 2012. The Kumaraswamy Gumbel Distribution. *Statistical Methods & Applications* 21 (2): 139{168.

Davis, R. and Resnick, S. 1985. Limit theory for moving averages of random variables with regularly varying tail probabilities. *The Annals of Probability* 179{195.

Davis, R. A. and Mikosch, T. 2009, In Handbook of Financial Time Series, In *Handbook of Financial Time Series*, 187{200, New York, NY: Springer, 187{200.

Dombry, C. 2013. Maximum Likelihood Estimators for the Extreme Value Index Based on the Block Maxima Method. *arXiv preprint arXiv:1301.5611* .

Eastoe, E. F. and Tawn, J. A. 2012. Modelling the distribution of the cluster maxima of exceedances of subasymptotic thresholds. *Biometrika* 99 (1): 43{55.

Efron, B. 1979. Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics* 7 (1): 1{26.

Embrechts, P., Kluppelberg, C. and Mikosch, T. 1997. *Modelling Extremal Events: for Insurance and Finance*. , vol. 33. New York: Springer.

Fasen, V., Kluppelberg, C. and Schlather, M. 2010. High-level Dependence in Time Series Models. *Extremes* 13 (1): 1{33.

Fawcett, L. and Walshaw, D. 2006a. A Hierarchical Model for Extreme Wind Speeds. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 55 (5): 631{646.

Fawcett, L. and Walshaw, D. 2006b. Markov Chain Models for Extreme Wind Speeds. *Environmetrics* 17 (8): 795{809.

Fawcett, L. and Walshaw, D. 2007. Improved Estimation for Temporally Clustered Extremes. *Environmetrics* 18 (2): 173{188.

Felici, M., Lucarini, V., Speranza, A. and Vitolo, R. 2006. Extreme Value Statistics of the Total Energy in an Intermediate Complexity Model of the Mid-latitude Atmospheric Jet. Part I: stationary case. *ArXiv Preprint Physics/0601081* .

Ferro, C. A. and Segers, J. 2003. Inference for Clusters of Extreme Values. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 65 (2): 545{556.

Fisher, R. A. and Tippett, L. H. C. 1928. Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Mathematical Proceedings of the Cambridge Philosophical Society* 24: 180{190.

Fuentes, C. and Casella, G. 2009. Testing for the Existence of Clusters. *SORT-Statistics and Operations Research Transactions* 33 (2): 115{158.

Grimshaw, S. D. 1993. Computing Maximum Likelihood Estimates for the Generalized Pareto Distribution. *Technometrics* 35 (2): 185{191.

Hall, P., Horowitz, J. L. and Jing, B.-Y. 1995. On blocking rules for the bootstrap with dependent data. *Biometrika* 82 (3): 561{574.

Hardle, W., Horowitz, J. and Kreiss, J.-P. 2003. Bootstrap Methods for Time Series. *International Statistical Review* 71 (2): 435{459.

Huerta, G. and Sanso, B. 2007. Time-varying Models for Extreme Values. *Environmental and Ecological Statistics* 14 (3): 285{299.

Hughes, G. L., Rao, S. S. and Rao, T. S. 2007. Statistical Analysis and Time-series Models for Minimum/Maximum Temperatures in the Antarctic Peninsula. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science* 463 (2077): 241{259.

Ji, C. 2013. Analysis of Temporally Dependent Extremes for the Gumbel Distribution. Master's thesis, University of Calgary. Canada.

Kluppelberg, C. and Lindner, A. 2005. Extreme value theory for moving average processes with light-tailed innovations. *Bernoulli* 11 (3): 381{410.

Kotz, S. and Nadarajah, S. 2000. *Extreme Value Distributions*. , vol. 31. London: Imperial College Press.

Kreiss, J.-P. and Lahiri, S. N. 2012. Bootstrap methods for time series. *Handbook of Statistics: Time Series Analysis: Methods and Applications* 30: 3{23.

Kunsch, H. R. 1989. The jackknife and the Bootstrap for General Stationary Observations. *The Annals of Statistics* 1217{1241.

Laurini, F. and Tawn, J. A. 2003. New Estimators for the Extremal Index and other Cluster Characteristics. *Extremes* 6 (3): 189{211.

Leadbetter, M. 1983. Extremes and Local Dependence in Stationary Sequences. *Probability Theory and Related Fields* 65 (2): 291{306.

Ledford, A. W. and Tawn, J. A. 2003. Diagnostics for Dependence within Time Series Extremes. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 65 (2): 521{543.

Liu, R. Y. and Singh, K. 1992. Moving Blocks Jackknife and Bootstrap Capture Weak Dependence. *Exploring the Limits of Bootstrap* 225{248.

Loynes, R. 1965. Extreme Values in Uniformly Mixing Stationary Stochastic Processes. *The Annals of Mathematical Statistics* 36 (3): 993{999.

Makkonen, L. 2008. Problems in the Extreme Value Analysis. *Structural Safety* 30 (5): 405{419.

Markovich, N. M. 2014. Modeling Clusters of Extreme Values. *Extremes* 17 (1): 97{125.

Montgomery, D. C., Jennings, C. and Kulahci, M. 2008. *Introduction to Time Series Analysis and Forecasting*. New Jersey: John Wiley & Sons.

Nakajima, J., Kunihama, T., Omori, Y. and Fruhwirth-Schnatter, S. 2012. Generalized Extreme Value Distribution with Time-dependence using the AR and MA Models in State Space Form. *Computational Statistics & Data Analysis* 56 (11): 3241{3259.

O'Brien, G. 1974. The Maximum Term of Uniformly Mixing Stationary Processes. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* 30 (1): 57{63.

O'Brien, G. L. 1987. Extreme Values for Stationary and Markov Sequences. *The Annals of Probability* 15 (1): 281{291.

Pankratz, A. 1983. *Forecasting with Univariate Box-Jenkins Models*. New York: Wiley.

Politis, D. N. 2003. The Impact of Bootstrap Methods on Time Series Analysis. *Statistical Science* 18 (2): 219{230.

Politis, D. N. and Romano, J. P. 1994. The Stationary Bootstrap. *Journal of the American Statistical Association* 89 (428): 1303{1313.

Prado, R. and Lopes, H. F. 2013. Sequential parameter learning and ltering in structured autoregressive state-space models. *Statistics and Computing* 23 (1): 43{57.

Reich, B. J., Shaby, B. A. and Cooley, D. 2014. A Hierarchical Model for Serially-dependent Extremes: A study of Heat Waves in the Western US. *Journal of Agricultural, Biological, and Environmental Statistics* 19 (1): 119{135.

Reiss, R.-D. and Thomas, M. 2007. *Statistical Analysis of Extreme Values: with Applications to Insurance, Finance, Hydrology and other Fields*. Basel: Birkhauser.

Robert, C. Y. 2008. Estimating the multivariate extremal index function. *Bernoulli* 14 (4): 1027{1064.

Robert, C. Y. 2013. Automatic Declustering of Rare Events. *Biometrika* 100 (3): 587{606.

Rootzen, H. 1986. Extreme Value Theory for Moving Average Processes. *The Annals of Probability* 14 (2): 612{652.

Rootzen, H. and Tajvidi, N. 2006. The multivariate Generalized Pareto Distributions. *Bernoulli* 5 (5): 917{930.

Scotto, M. G., Alonso, A. M. and Barbosa, S. M. 2009. Clustering Time Series of Sea Levels: Extreme Value Approach. *Journal of Waterway, Port, Coastal, and Ocean Engineering* 136 (4): 215{225.

Singh, K. 1981. On the Asymptotic Accuracy of Efron's Bootstrap. *The Annals of Statistics* 9 (6): 1187{1195.

Smith, R. L. 2003. Statistics of extremes, with applications in environment, insurance and nance, chap 1. *Statistical analysis of extreme values: with applications to insurance, finance, hydrology, and other fields. Birkhuser, Basel* .

Toulemonde, G., Guillou, A., Naveau, P., Vrac, M. and Chevallier, F. 2010. Autoregressive Models for Maxima and their Applications to CH4 and N2O *Environmetrics* 21 (2): 189{207.

Verhoeven, P. and McAleer, M. 2000. Modelling Outliers and Extreme Observations for ARMA-GARCH Processes. In *Econometric Society World Congress 2000 Contributed Papers*. Econometric Society.

Young, P. 1996. Jackknife and Bootstrap Resampling Methods in Statistical Analysis to Correct for Bias. *Statistical Science* 11: 189{228.