MODELLING ELECTRICITY DEMAND IN SOUTH AFRICA

by

Caston SIGAUKE

THESIS

Submitted in fulfillment of the requirements for the degree of

PHILOSOPHIAE DOCTOR

in

STATISTICS

in the

FACULTY OF NATURAL AND AGRICULTURAL SCIENCES
DEPARTMENT OF MATHEMATICAL STATISTICS AND ACTUARIAL SCIENCE

at the

UNIVERSITY OF THE FREE STATE

BLOEMFONTEIN: JANUARY 2014

PROMOTER: DR. DELSON CHIKOBVU

Declaration

I declare that the thesis which is hereby submitted for the qualification of Doctor of Philosophy in Statistics at the University of the Free State, is my own independent work and has not been handed in before for a qualification at/in another University/Faculty. I further declare that all sources cited or quoted are indicated and acknowledged by means of a comprehensive list of references. I further cede copyright of the thesis to the University of the Free State.

Signature.....

Date: 20 January 2014

Dedication

To my wife Sabelo and children,
Tinovimbanashe, Tinevimbo and Thandiwe

Acknowledgments

Firstly I would like to thank my academic promoter, Dr Delson Chikobvu for his patience, inspiration and guidance through out my PhD studies, for introducing me to extreme value theory and Bayesian statistics and co-authoring papers based on this thesis. I am greatly indebted to Dr. Andrehette Verster of the Department of Mathematical Statistics and Actuarial Science at the University of the Free State in South Africa for co-authoring five papers and one peer-reviewed conference proceeding.

I would like to express my sincere gratitude to Professor Maseka Lesaoana for her support during my PhD studies. I am deeply indebted to the late Professor Nkatini for his encouragement, inspiration and direction during the early stages of identifying a research area and proposal writing. Thank you Dr. Zindoga Mukandavire for helpful comments and suggestions throughout the writing of this thesis. I would like to thank Professor James Cochran for helpful comments. I am grateful to Brian T Darikwa, Rahab Makwela, Professor Abebe Tessera and Paul Moleko for their encouragement and support.

The success of this thesis would not have been realized had it not been for Eskom, South Africa's power utility company which provided both electricity and temperature data which is used in this thesis. I would like to thank my mother Mrs Celiah Sigauke for her support throughout my doctoral studies. To the almighty God I am grateful for the gift of life and wisdom. To my wonderful wife and lovely children, I thank you for your patience and support and finally Dad is coming home.

Contents

List of Figures

List of Tables

List of Special Symbols

$G_\gamma(x)$	extreme value distribution function for maxima
$W_\xi(x)$	generalized Pareto distribution function
$W_\eta(x)$	generalized single Pareto distribution function
x_p	quantile function
x_t	daily peak electricity demand at time t
$\pi(\theta)$	prior distribution
$\pi(x\|\theta)$	likelihood function
$\pi(\theta\|x)$	posterior distribution
$\pi(x_{n+1}\|x)$	posterior predictive density of a future observation x_{n+1}
τ	a sufficiently high threshold
log	natural logarithm
ε_t	white noise process
X'	transpose of X
γ	Extreme Value Index (EVI) for the generalized extreme value distribution
ξ	EVI for the generalized Pareto distribution
η	EVI for the generalized single Pareto distribution
$\mathcal{D}(G_\gamma(x))$	domain of attraction of $G_\gamma(x)$
$F^{[\tau]}$	empirical exceedance distribution function of τ
$\alpha(F)$	left end point of distribution function F
$\omega(F)$	right end point of distribution function F

Distributions

Generalized extreme value family of max-stable distributions

Gumbel distribution

Distribution function

$$G(x) = \mathbf{exp}\left\{-\mathbf{exp}\left[-\frac{x-\mu}{\sigma}\right]\right\}, \gamma = 0$$

Fr$\acute{e}$chet distribution

Distribution function

$$G_\gamma(x) = \mathbf{exp}\left\{-\left[1+\gamma\left(\frac{x-\mu}{\sigma}\right)\right]^{-\frac{1}{\gamma}}\right\}, \gamma > 0$$

Weibull distribution

Distribution function

$$G_\gamma(x) = \mathbf{exp}\left\{-\left[1+\gamma\left(\frac{x-\mu}{\sigma}\right)\right]^{-\frac{1}{\gamma}}\right\}, \gamma < 0$$

Generalized Extreme Value Distribution, $\gamma \neq 0$

Notation	$X \sim \text{GEVD}(\gamma, \mu, \sigma)$
Distribution function	$G_\gamma(x) = \exp\left\{-\left[1+\gamma\left(\frac{x-\mu}{\sigma}\right)\right]^{-\frac{1}{\gamma}}\right\}$
Density function	$g_\gamma(x) = \frac{1}{\sigma}\left[1+\gamma\left(\frac{x-\mu}{\sigma}\right)\right]^{-\frac{1}{\gamma}-1}\exp\left\{-\left[1+\gamma\left(\frac{x-\mu}{\sigma}\right)\right]^{-\frac{1}{\gamma}}\right\}$
Range	$\left\{x \vert 1+\gamma\left(\frac{x-\mu}{\sigma}\right) > 0\right\}$
Parameters	$\gamma \neq 0, -\infty < \mu < \infty, \sigma > 0$
EVI	γ

Generalized Extreme Value Distribution, $\gamma = 0$

Notation	$X \sim \text{GEVD}(0, \mu, \sigma)$
Distribution function	$G_\gamma(x) = \exp\left\{-\exp\left(-\frac{x-\mu}{\sigma}\right)\right\}$
Density function	$g_\gamma(x) = \frac{1}{\sigma}\exp\left(-\frac{x-\mu}{\sigma}\right)\exp\left\{-\exp\left(-\frac{x-\mu}{\sigma}\right)\right\}$
Range	$-\infty < x < \infty$
Parameters	$-\infty < \mu < \infty, \sigma > 0$
EVI	0

Generalized Pareto Distribution, $\xi \neq 0$

Notation	$X \sim \text{GPD}(\xi, \sigma)$
Distribution function	$W_\xi(x) = 1 - \left(1 + \frac{\xi(x-\tau)}{\sigma}\right)^{-\frac{1}{\xi}}$
Density function	$w_\xi(x) = \frac{1}{\sigma}\left(1 + \frac{\xi(x-\tau)}{\sigma}\right)^{-(1+\frac{1}{\xi})}$
Range	$x > \tau$, for $\xi > 0$
	$\tau < x < \tau - \frac{\sigma}{\xi}$, for $\xi < 0$
Parameters	$\xi \neq 0, \sigma > 0$
EVI	ξ

Generalized Single Pareto Distribution

Notation	$X \sim \text{GSPD}(\eta)$
Distribution function	$W_\eta(x) = \left\{1 + \frac{\eta}{1+\tau\eta}(x-\tau)\right\}^{-\frac{1}{\eta}}, \eta \neq 0$
Density function	$w_\eta(x) = \frac{1}{1+\tau\eta}\left[1 + \frac{\eta(x-\tau)}{1+\tau\eta}\right]^{-\frac{1}{\eta}-1}$
Range	$x > \tau$
Parameters	$\eta \neq 0$
EVI	η

Generalized Gamma Distribution

$$f(y) = \frac{\alpha^k}{\beta\Gamma(k)} e^{-\alpha y^{\frac{1}{\beta}}} y^{\frac{k}{\beta}-1}, y > 0$$

Generalized Gamma Distribution for $\beta = k = 1$

When $\beta = k = 1$ we have the exponential distribution

$$f(y) = \alpha e^{-\alpha y}, y > 0$$

Generalized Gamma Distribution for $\beta = 1$

If $\beta = 1$ we have the gamma distribution

$$f(y) = \frac{\alpha^k}{\Gamma(k)} e^{-\alpha y} y^{k-1}, y > 0$$

Generalized Gamma Distribution for $k = 1$

For $k = 1$ we have the Weibull distribution

$$f(y) = \frac{\alpha}{\beta} e^{-\alpha y^{\frac{1}{\beta}}} y^{\frac{1}{\beta}-1}, y > 0$$

List of Abbreviations and Acronyms

ADT	Average Daily Temperature
APED	Annual Peak Electricity Demand
AIC	Akaike Information Criterion
ARIMA	Autoregreesive Integrated Moving Average
CDD	Cooling Degree Days
CDF	Cumulative Distribution Function
DLP	Daily Load Profile
DoE	Department of Energy
DPED	Daily Peak Electricity Demand
DSM	Demand Side Management
EDPED	Extreme Daily Peak Electricity Demand
EE	Energy Efficiency
e.g.	for example
EGARCH	Exponential Generalized Autoregreesive Conditional Heteroskedasticity
Eskom	South Africa's power utility company
EVT	Extreme Value Theory
GARCH	Generalized Autoregressive Conditional Heteroskedasticity

GBG	Generalized Burr Gamma
GCV	Generalized Cross Validation
GEVD	Generalized Extreme Value Distribution
GDP	Gross Domestic Product
GPD	Generalized Pareto Distribution
GSPD	Generalized Single Pareto Distribution
HDD	Heating Degree Days
IPP	Independent Power Producers
LTF	Long Term Forecast
MAE	Mean Absolute Error
MARS	Multivariate Adaptive Regression Splines
MAPE	Mean Absolute Percentage Error
MTF	Medium Term Forecast
MDI	Maximal Data Information
MLE	Maximum Likelihood Estimate
MLP	Monthly Load Profile
MCMC	Markov Chain Monte Carlo
NESO	Nett Energy Sent Out
PDF	Probability Density Function
PLR	Piecewise Linear Regression
POE	Probability Of Exceedance
QQ	Quantile to Quantile
RMSE	Root Mean Square Error
SARIMA	Seasonal Autoregressive Integrated Moving Average
STLF	Short Term Load Forecast
WLP	Weekly Load Profile

Research Outputs

A list of research outputs from this thesis is given below.

Peer Reviewed Journal Publications

1. Sigauke C., Verster A. and Chikobvu D. (2013). Extreme daily increases in peak electricity demand: tail-quantile estimation. Energy Policy Journal, Vol. 53, pp. 90-96.

2. Chikobvu D. and Sigauke C. (2013). Modelling influence of temperature on daily peak electricity in South Africa. Journal of Energy in Southern Africa, Vol. 24 (4), pp. 63-70.

3. Verster A., Chikobvu D. and Sigauke C. (2013). Analysis of the same day of the week increases in peak electricity demand in South Africa. ORiON Journal, Vol. 29 (2), pp. 125-136.

4. Sigauke C., Chikobvu D. and Verster A. (2012). Modelling daily increases in peak electricity demand using a generalized Pareto distribution. South African Statistical Journal: Peer-reviewed Proceedings of the 54th Annual Conference of the South African Statistical Association, pp. 58-66.

5. Chikobvu D., Sigauke C. and Verster A. (2012). Winter peak electricity load forecasting in South Africa using extreme value theory. South African Statistical Journal, Vol. 46, pp. 377-394.

6. Sigauke C., Verster A. and Chikobvu D. (2012). Tail quantile estimation of heteroskedastic intraday increases in peak electricity demand, Open Journal of Statistics, Vol. 2(4), pp. 435-442.

7. Sigauke C. and Chikobvu D. (2012). Short-term daily winter peak electricity demand in South Africa. African Journal of Business Management, Vol. 6(32), pp. 9243-9249.

8. Chikobvu D. and Sigauke C. (2012). Regression-SARIMA modelling of daily peak electricity demand in South Africa. Journal of Energy in Southern Africa, Vol. 23 (3), pp. 23-30.

9. Chikobvu D. and Sigauke C. (2012). A frequentist and Bayesian regression analysis to daily peak electricity demand in South Africa. African Journal of Business Management, Vol. 6 (40), pp. 10524-10533.

10. Chikobvu D., Sigauke C. and Verster A. (2012). Winter peak electricity load forecasting in South Africa using extreme value theory with a Bayesian flavour. Journal of Business and Economics, Vol. 3(5), pp. 380-389.

11. Sigauke C. and Chikobvu D. (2011). Prediction of daily peak electricity demand in South Africa using volatility forecasting models. Energy Economics Journal, Vol. 33(5), pp. 882-888.

12. Sigauke C. and Chikobvu D. (2010). Daily peak electricity load forecasting in South Africa using a multivariate non-parametric regression approach. ORiON Journal, Vol. 26(2), pp. 97-111.

International Conferences

1. Chikobvu D., Sigauke C. and Verster A., Winter peak electricity load forecasting in South Africa using extreme value theory with a Bayesian

flavour. International Society for Bayesian Analysis. Poster presented at the International Society for Bayesian Analysis (ISBA) June 2012 Conference, Kyoto, Japan. http://bayesian.org/node/2593

2. Sigauke C. and Chikobvu D., Modelling daily peak electricity load forecasting in South Africa using a multivariate non-parametric regression approach. 19th Triennial Conference of the International Federation of Operations Research Societies, Melbourne, Australia, July10-15, 2011. Chaired the invited session FC-1: OR, Energy, and Africa in stream OR Applications in Energy. http://www.euro-online.org/conf/admin/tmp/chair-ifors2011.idx

Other Conferences

1. Sigauke C. and Chikobvu D. A comparative analysis of dynamic regression models to daily peak load forecasting in South Africa. 55th Annual Conference of the South African Statistical Association, 4-8 November 2013, University of Limpopo, Polokwane, South Africa. Chaired a session: General/Applied Statistics.

2. Sigauke C. and Chikobvu D., Stochastic modelling of daily peak electricity demand. Second ICCSSA Convention, 19-21 June 2013, Eskom Academy of Learning Midrand.

3. Sigauke C., Chikobvu D. and Verster A., Modelling daily increases in peak electricity demand using the generalized Pareto distribution. 54th Annual Conference of the South African Statistical Association, 5-9 November 2012, Nelson Mandela Metropolitan University, Port Elizabeth, South Africa. Chaired a session: Analytics Competition.

4. Sigauke C., Verster A. and Chikobvu D., Tail quantile estimation of heteroskedastic intraday increases in peak electricity demand. 54th Annual

Conference of the South African Statistical Association, 5-9 November 2012, Nelson Mandela Metropolitan University, Port Elizabeth, South Africa.

5. Sigauke C., Verster A., Chikobvu D., Extreme daily increases in peak electricity demand: tail-quantile estimation. 41st Annual Conference of the Operations Research Society of South Africa, 16-19 September 2012, Aloe Ridge Hotel, Cradle of Humankind, South Africa. Chaired a session: Analytics.

6. Sigauke C. and Chikobvu D., The influence of temperature on daily electricity load forecasting in South Africa using Bayesian parameter estimates to a MARS model. Annual Conference of the South African Statistical Association, 31 October to 4 November, 2011, CSIR, Pretoria, South Africa.

7. Sigauke C. and Chikobvu D., Daily peak electricity load forecasting in South Africa using a multivariate non-parametric regression approach. 2nd Annual Forecasting Eskom Conference, 1-2 December 2010, Midrand, South Africa.

8. Chikobvu D. and Sigauke C., Modelling daily peak electricity demand in South Africa using SARIMA and RegSARIMA models. Second Isibalo Young African Statisticians Conference, 1-3 December 2010, Pretoria, South Africa.

9. Sigauke C. and Chikobvu D., Daily peak electricity load forecasting in South Africa using a multivariate non-parametric regression approach. 39th Annual Conference of the Operations Research Society of South Africa, 26-29 September 2010, University of Limpopo, South Africa.

Chapter 1

Introduction

1.1 Background

Electricity demand forecasting involves planning under uncertainty and has to cope with operational, tactical and strategic planning. Planning under uncertainty in the electricity sector involves determining appropriate location of a generating plant, plant size, transmission and distribution (load flow analysis, analysis of the frequency and occurrence of extreme peak loads and scheduling of electricity). Uncertainties in electricity demand forecasting may arise due to increased technology making use of electricity, population growth, general randomness in individual usage of electricity, prevailing economic and weather conditions. Uncertainty is involved in variables such as electricity demand, raw material supply cost (coal, gas, amount of water needed for hydroelectricity generation, etc), electricity generation and generation cost.

1.2 Statement of the Problem

Electricity demand forecasting is carried out for various time frames, from short term to long term forecasting. Short Term Load Forecasting (STLF) is important for scheduling of electricity, for daily system operations and for supply and demand balance of electricity (Taylor, 2008). Medium Term (MT) forecasts are generated for up to a few years ahead which are important for risk assessment and maintenance planning. On the other hand long term electricity demand is important for strategic planning which involves capacity expansion, power system planning, power security and supply reliability (Ismail *et al*., 2009). Long term peak electricity demand forecasting ranges from several months to several years ahead (Hyndman and Fan, 2008). In this thesis we concentrate on short term daily peak electricity demand forecasting.

Short term peak electricity demand in South Africa is not only influenced by means and averages but also by the tails of probability distributions. The problem considered in this thesis consists of developing accurate forecasting models for short term daily peak electricity demand and developing a modelling framework for assessing Extreme Daily Peak Electricity Demand (EDPED) uncertainty in South Africa.

1.2.1 Peak electricity demand

Short term daily peak electricity demand forecasting is important for load management (i.e. load flow analysis, scheduling of electricity, etc.). Modelling of extreme daily peaks helps in the analysis of the frequency of occurrence of extreme daily peak load forecasts. This helps in the quantification of electricity which can be shifted to off peak hours which is important for grid stability by a power utility company. This also helps in determining the number of critical peak days. A critical peak day is defined as one in which demand exceeds

normal demand which may cause severe stress on the power system. A critical peak price (a tariff higher than the off peak tariff) is then used on all those days which have been declared critical peak days. Critical peak days are determined by various factors which include among others, extreme weather conditions, special calendar events and generating capacity constraints. The quantification of the number of critical peak days can also be done by modelling of extreme daily peak loads.

Predictions of short term peak electricity demand can be achieved by including in forecasting models important drivers of electricity demand (Hyndman and Fan, 2010; Fan and Hyndman, 2012; among others). Several factors which affect peak electricity demand are population growth, weather variables, calendar effects, economic variables, changing technology as well as the general randomness in individual usage (Hyndman and Fan, 2010). Peak load forecasting is also affected by other factors such as transmission and generation constraints, interruption of supply which include power outages, tariffs including demand side management also known as load management. Weather variables include temperature, humidity, wind speed, cloud cover and precipitation; economic factors are gross domestic product, producer price index, population growth; calendar effects include day of the week effect, holiday effect (Munoz et al., 2010).

Peak electricity demand is an energy policy concern for all economies throughout the world, causing blackouts, power outages and increased costs of electricity (Strengers, 2012). As a result many countries around the world are designing energy efficient and demand side management (load management) strategies to either redistribute or reduce electricity demand during peak periods. In this thesis we propose a modelling framework which deals with short term predictions of Daily Peak Electricity Demand (DPED) and use of extreme

value theory to quantify the amount of the predicted DPED to be shifted to off peak periods. The modelling framework discussed in this thesis can also be used to assess the frequency and level of occurrence of extreme DPED which is important in guiding planning in the optimal scheduling of electricity.

1.2.2 South Africa

Efficient use of electricity in South Africa is a national priority and a necessity for the future economic development of the country. Forecasting electricity demand constitutes a vital part of energy policy of a country, especially for a developing country like South Africa whose electricity demand has grown rapidly over the past twenty years since democracy in 1994.

The data used in this thesis is from Eskom, South Africa's power utility company. The data considered is Net Energy Sent Out (NESO) in response to demand for electricity. NESO (measured in megawatts) is defined as the rate at which electrical energy is delivered to customers. In this thesis, NESO is used as a proxy of electricity demand after adjusting for energy losses. The data are for the period 2000 to 2010. This definition of electricity demand has its weaknesses. Electricity demand is bounded by the power plants' capacity to provide supply at any time of the day, including the need for reserve capacity. Demand cannot exceed supply. If demand were to exceed supply, intervention takes place in the form of, for example, load shedding. This NESO definition excludes the demand from people, companies, who are willing (or unwilling) and able (or unable) to pay for electricity, but currently do not have access to electrical power. Despite the weakness in the NESO definition of electricity demand, it is still a good and measurable proxy for electricity demand. The NESO considered in this thesis is Daily Peak Electricity Demand (DPED) which is the maximum hourly demand in a 24-hour day period. Aggregated DPED data are used for the agricultural, commercial, domestic and industrial sectors of South

Africa.

In line with our definition of NESO is the definition given by Cottet and Smith (2003). Cottet and Smith (2003) define electricity load as electricity demand if electricity generated is equal to electricity drawn from the grid in the absence of blackouts and load-shedding.

Daily peak electricity demand in South Africa has been increasing steadily over the years 2000 to 2008 and levelled off after 2008 as shown by Figure 1.1. After 2009 electricity demand started increasing again. The decrease in

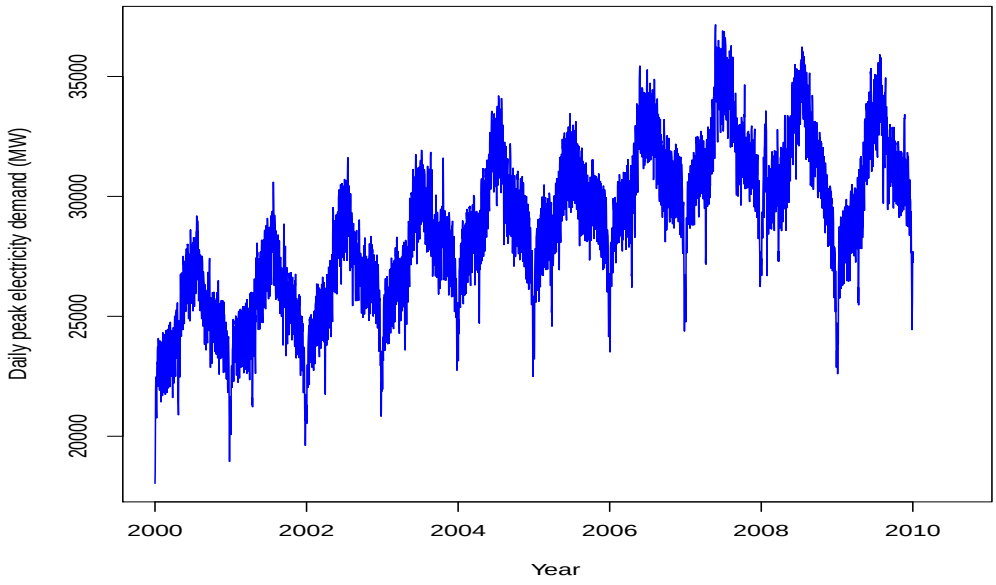

Figure 1.1: Daily peak electricity demand in megawatts for the years 2000 to 2010.

electricity demand over the years 2008 to 2009 is probably due to blackouts and energy saving campaigns. This was happening during the period of world recession which led to load shedding. The data in Figure 1.1 exhibits strong seasonality and is not stationary. In this thesis data will be made stationary by: (1) detrending and the detrended data will be used for modelling winter

peak electricity demand, (2) taking first difference which will give rise to day to day changes in peak electricity demand and (3) seventh difference which will give rise to same day of the week changes in peak electricity demand.

Historical data on temperature from 36 meteorological stations from all the provinces of South Africa are also used in the modelling of daily peak demand forecasting. A plot of the average daily electricity demand against average daily temperature is given in Figure 1.2. A visual inspection of Figure 1.2 suggests that the relationship between temperature and electricity demand is non-linear and that electricity demand is more sensitive to cold temperatures than hot temperatures in South Africa. Some of the challenges faced by Eskom is the

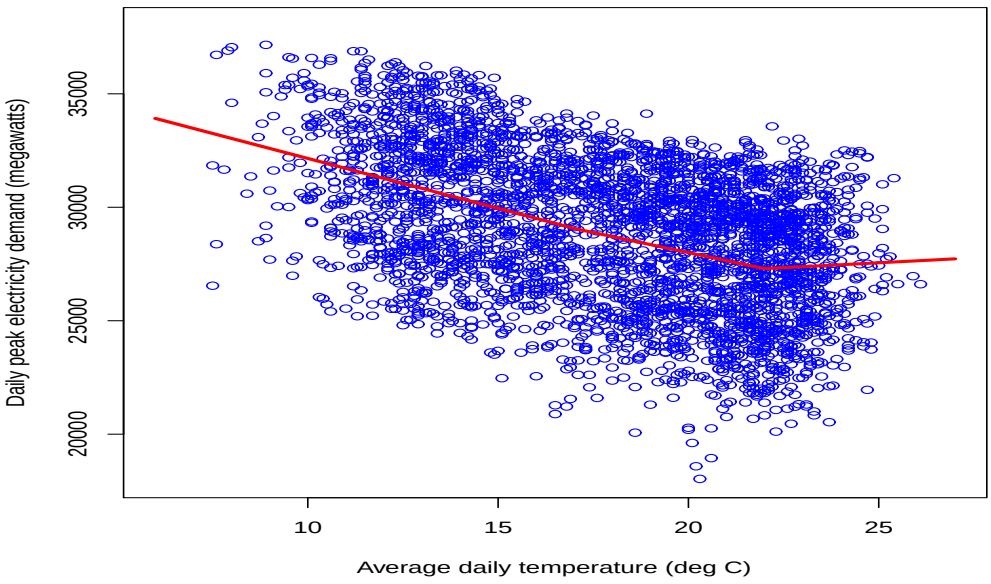

Figure 1.2: Scatter plot of daily peak electricity demand (in megawatts) against average daily temperature (in degrees Celsius).

description of the influence of temperature on electricity demand and that of accurate predictions and modelling of peak electricity load forecasts. Accurate peak load forecasts are important as they assist decision makers in improving capital expenditures which include securing of adequate capacities for gener-

ation, transmission and distribution. Daily, weekly, monthly and annual load profiles help in determining the time and level of peak loads. A typical daily load profile for South Africa is given in Figure 1.3. The daily load profile shown in Figure 1.3 is bimodal with two peaks, one in the morning around 09:00hrs and the other in the evening around 20:00hrs. DPED is around 20:00hrs. A

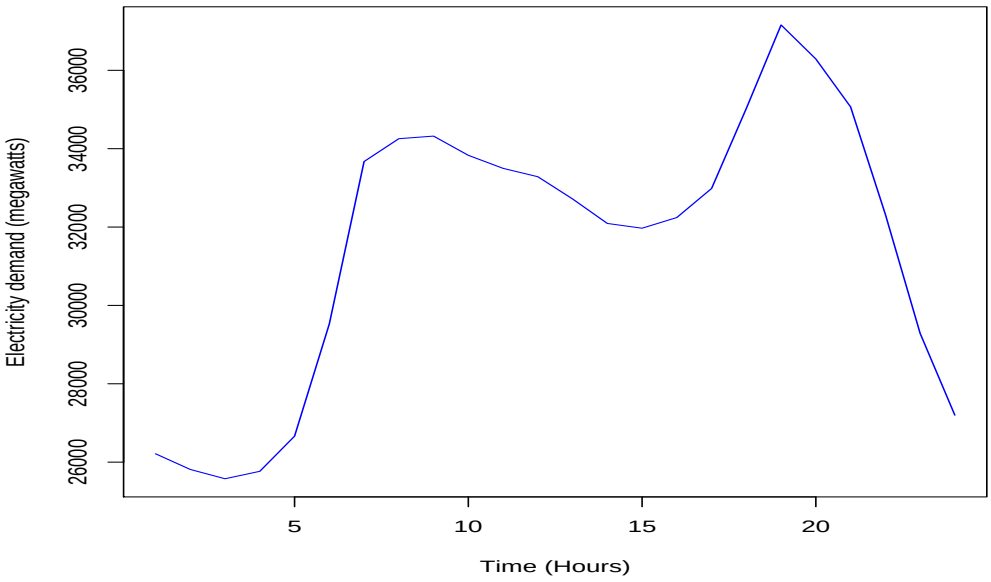

Figure 1.3: Daily load profile. There are two peaks, one in the morning and the other peak in the evening.

typical yearly load profile for the year 2007 given in Figure 1.4 shows that there are two peaks, one in winter and the other in summer. The winter period in the Southern Hemisphere is between May and August of each year. In South Africa demand for electricity increases significantly in winter than in summer as shown in Figure 1.4.

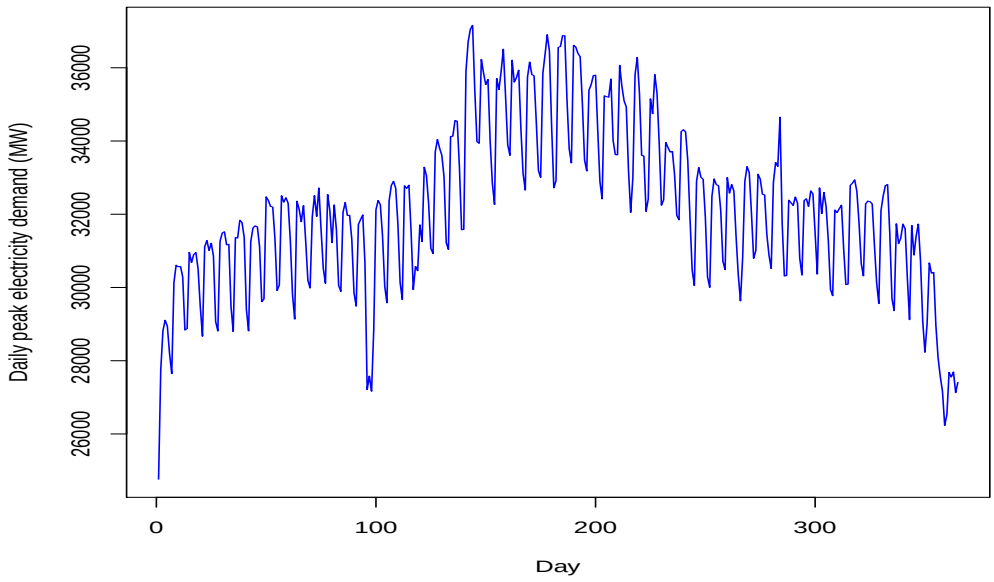

Figure 1.4: Daily peak electricity demand data in megawatts for the year 2007.

The weekly load profile given in Figure 1.5 shows that demand for electricity decreases significantly on Friday through to Sunday. For the rest of the week days the load variations are small.

1.3 Study aim and objectives

The main aim of this study is to develop a modelling framework which can be used in the electricity sector for carrying out accurate assessments of the frequency and level of occurrence of mean and extreme daily peak electricity demand. The objectives are to

1. identify the major drivers of electricity demand in South Africa,
2. develop forecasting models for short term peak electricity demand,
3. model the tail behaviour of extreme daily peak electricity demand,
4. improve and extend some of the forecasting models,

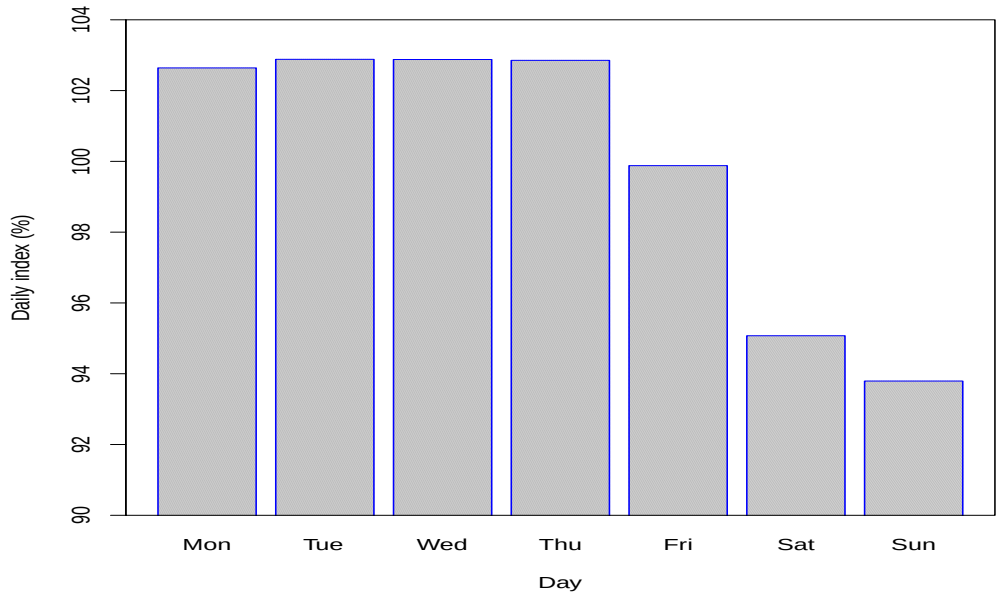

Figure 1.5: Weekly seasonal index plot of daily peak electricity demand.

5. suggest areas for further study.

1.4 Modelling framework: Interplay of Frequentist and Bayesian analysis

In light of the problem statement the thesis concentrates on the modelling of daily peak electricity demand in the energy sector using a frequentist and Bayesian analysis. For large sample sizes the Bayesian approach with non informative priors give results for parametric models which are similar to results from frequentist methods. However the Bayesian approach is more attractive than the frequentist approach as it combines prior information with data, enables analysis with small samples, the estimation is robust (Ardia and Hoogerheide, 2010). The Bayesian approach does not rely on asymptotic approximation (Coles, 2001; Bayarri and Berger, 2004; Beirlant *et al.*, 2004). The Bayesian modelling framework obeys the likelihood principle which states that

for a given sample of data, any two probability models that have the same likelihood function yield the same inference (Gelman *et al.*, 2004). Bayesian analysis provides a coherent modelling framework for solving problems under conditions of uncertainty (Marin *et al.*, 2003).

Interplay of Bayesian and frequentist analysis is discussed in Bayarri and Berger (2004). Bayarri and Berger (2004) identify the most common areas of important and useful connections between the Bayesians and frequentists as those scenarios when no external information, except the data and the model itself, is to be introduced into the analysis. Bayarri and Berger (2004) indicate that there are many areas of frequentist methodology that can be replaced by the Bayesian methodology that provide superior answers.

1.5 Significance of the study

Energy security is very important for the economic growth of any country. It is therefore important to have a reliable supply of electricity to meet demand at any given time. Accurate load forecasts will enable effective load shifting between transmission substations, scheduling of start-up times of peak stations, load flow analysis and power system security studies. Modelling of extreme daily peak electricity demand improves the reliability and stability of a power network.

1.6 Contributions

The major contribution of this thesis is in applying statistical techniques in modelling mean and extreme daily peak electricity demand in South Africa. The contributions are:

1. modelling the influence of temperature on electricity demand in South

Africa,

2. quantification of the amount of electricity which can be shifted to off peak hours. This is achieved through accurate assessment of the frequency and level of future extreme load forecasts. This modelling approach provides a policy framework for load curtailment and determination of the number of critical peak days for power utility companies. This has not been done for electricity demand modelling in the context of South Africa to the best of our knowledge,

3. application of extreme value theory in modelling extreme daily electricity demand winter peaks in South Africa,

4. application of extreme value theory statistical distributions in modelling extreme daily increases and same day of the week increases in peak electricity demand,

5. extension of the general regression-seasonal autoregressive integrated moving average modelling approach to a regression-seasonal autoregressive integrated moving average-generalized autoregressive conditional heteroskedasticity model,

6. development of a piecewise linear regression model using both classical and Bayesian statistical approaches, to explain the influence of temperature on daily peak electricity demand in South Africa,

7. estimating the extreme value index of the generalized single Pareto distribution (by Verster and De Waal, 2011; they only used the Bayesian approach) using the maximum likelihood estimation method.

8. extension of an autoregressive moving average-exponential generalized autoregressive conditional heteroskedasticity model to an autoregressive moving average-exponential generalized autoregressive conditional

heteroskedasticity-generalized single Pareto distribution. The developed hybrid model can be used for modelling the risk of under or over demand forecasts and also helps in determining the number of critical peak days which are important for dynamic pricing programmes such as critical peak pricing.

1.7 Thesis structure

The rest of the thesis is organized as follows. **Chapter 2** reviews literature on daily peak electricity demand forecasting. The thesis is divided into two broad sections. **Chapters 3-6** deal with modelling the mean of DPED. Peak electricity demand is influenced by the tails of probability distributions as well as by means or averages. At times there is a need to depart from the average thinking and exploit information provided by the extremes (tails of distributions). Modelling of tail behaviour of DPED through the Extreme Value Theory (EVT) approach is then covered in **Chapters 7-11**.

Modelling the influence of temperature on daily peak electricity demand using multivariate adaptive regression splines models then follows in **Chapter 3**[1]. It is shown in **Chapter 3** that electricity demand in South Africa is highly sensitive to winter periods than summer periods. An extension of the analysis of the effect of temperature on daily peak electricity demand discussed in **Chapter 3** is presented in **Chapter 4** in which we model the effect of temperature on daily peak electricity demand in South Africa using a piecewise linear regression model. The use of the Bayesian parameter estimates to a piecewise linear regression model in explaining the influence of temperature on daily peak electricity demand is discussed in **Chapter 5**[2]. The developed

[1]Daily peak electricity load forecasting in South Africa using a multivariate non-parametric regression approach. ORiON Journal, Vol. 26(2), 2010, pp. 97-111.

[2]A frequentist and Bayesian regression analysis to daily peak electricity demand in South Africa. African Journal of Business Management, Vol. 6 (40), 2012, pp. 10524-10533.

model captures a wide variety of electricity demand drivers such as temperature, seasonal, lagged demand and calendar effects. A Bayesian modelling framework is introduced in the analysis so as to take into account uncertainty in the estimation of the piecewise linear regression parameters.

In **Chapter 6**[1] we develop some hybrid models for daily peak electricity demand forecasting in South Africa. We develop a regression with seasonal autoregressive integrated moving average errors (Reg-SARIMA) model to predict daily peak electricity demand. Regression-SARIMA modelling which we call Reg-SARIMA modelling framework captures important drivers of electricity demand. The Reg-SARIMA model is then extended to a SARIMA model with generalized autoregressive conditional heteroskedastic (SARIMA-GARCH) errors and to a Reg-SARIMA-GARCH (Reg-SARIMA-GARCH). The Reg-SARIMA-GARCH modelling methodology is introduced to accommodate the possibility of serial correlation in volatility since the daily peak demand data exhibits non-constant mean and variance, and multiple seasonality corresponding to weekly and monthly periodicity.

Chapters 3 to 6 discussed the modelling of the means of daily peak electricity demand. Electricity demand is also influenced by the tails of the distributions. **Chapter 7**[2] discusses an analysis of the effect of temperature on daily electricity demand in South Africa using extreme value theory. An analysis of extreme low temperatures is important for load forecasters in the electricity sector. An analysis of extreme daily and same day of the week increases in peak electricity demand using a Generalized Pareto-type (GP-type) distribution which has

[1]Regression-SARIMA modelling of daily peak electricity demand in South Africa. Journal of Energy in Southern Africa, Vol. 23 (3), 2012, pp. 23-30; Prediction of daily peak electricity demand in South Africa using volatility forecasting models. Energy Economics Journal, Vol. 33(5), 2011, pp. 882-888.

[2]Modelling influence of temperature on daily peak electricity in South Africa. Journal of Energy in Southern Africa, Vol. 24 (4), pp. 63-70.

one parameter to estimate is discussed in **Chapter 8**[1]. The GP-type distribution belongs to the domain of attraction of the Generalized Pareto Distribution (GPD) and as such it is referred to as the Generalized Single Pareto Distribution (GSPD). A comparative analysis is then done with a GPD.

Chapters 3 to 7 established that electricity demand in South Africa is highly sensitive to cold temperatures. A detailed discussion of winter peak electricity demand modelling using the generalized single Pareto distribution in South Africa is presented in **Chapter 9**[2]. A comparative analysis is then done with a Generalized Extreme Value Distribution (GEVD).

An assessment of daily peak electricity demand uncertainty using an Autoregressive Moving Average-Exponential Generalized Autoregressive Conditional Heteroskedastic-Generalized Single Pareto Distribution (ARMA-EGARCH-GSPD) approach is discussed in **Chapter 10**[3]. The advantage of this modelling approach lies in its ability to capture conditional heteroskedasticity in the data through the EGARCH framework, while at the same time estimating the extreme tail quantiles through the GSPD modelling framework. The developed model is then used for extreme tail quantile estimation using daily peak electricity demand data and modelling the risk of under or over demand predictions. **Chapters 3-9** discussed short term modelling of daily peak electricity demand which is important for operational planning. Strategic decision making which involves planning for the construction of generating power plants

[1]Modelling daily increases in peak electricity demand using a generalized Pareto distribution. South African Statistical Journal: Peer-reviewed Proceedings of the 54th Annual Conference of the South African Statistical Association, 2012, pp. 58-66; Extreme daily increases in peak electricity demand: tail-quantile estimation. Energy Policy Journal, Vol. 53, 2013, pp. 90-96; Analysis of the same day of the week increases in peak electricity demand in South Africa. ORiON Journal, Vol. 29 (2), 2013, pp. 125-136.

[2]Winter peak electricity load forecasting in South Africa using extreme value theory. South African Statistical Journal, Vol. 46, 2012, pp. 377-394.

[3]Tail quantile estimation of heteroskedastic interday increases in peak electricity demand, Open Journal of Statistics, Vol. 2(4), 2012, pp. 435-442.

requires long term peak electricity demand forecasts. **Chapter 11** presents a brief discussion of the use of the generalized gamma distribution in forecasting annual peak electricity demand. A summary of research findings, contributions, suggested areas for further study and concluding remarks are presented in **Chapter 12**.

The following statistical packages were used in this thesis: R, Eviews and the Salford Predictive Modeller. Some of the R codes are given in an appendix at the end of the thesis.

Chapter 2

Literature Review

2.1 Introduction

Prediction of peak load demand is very important for decision making processes in the electricity sector. Decision making in this sector involves planning under uncertainty. This involves for example strategic planning for generating capacity expansion and even for the day to day operation of power plants which includes among others optimal scheduling of electricity. It is important therefore to produce very accurate forecasts as the consequences of underestimation or overestimation can be costly. Underestimation has a serious negative impact on the national electricity supply system of a country. Overestimation results in wastage of resources due to excess production (Hyndman and Fan, 2010). This may even lead to further loss of business as restoration of a power plant or construction of new generating plants take a long time before generation can start.

Load demand forecasting has been studied extensively for over three decades.

Hahn *et al.* (2009) give an overview of some of the methods used in demand load forecasting. The methods are classified into regression based, time series, state space and Kalman-filtering. The other group of methods suggested are the artificial intelligence and computational intelligence. Neural networks and support vector regression methods fall into this class. Munoz *et al.* (2010) discuss the importance of short term load forecasting as one which is at the core of power system planning, scheduling and control of electricity. Factors such as multiple seasonalities, temperature and the calendar effects have significant effects on short term electricity demand (Munoz *et al.*, 2010). A review of energy demand models using classical methods such as time series, regression, econometric, decomposition, cointegration, SARIMA, artificial (expert systems and artificial neural networks) and computational intelligence, including new techniques such as support vector regression, ant colony, particle swarm optimization and bottom up models such as Long-range Energy Alternatives Planning system (LEAP) and MARKet ALlocation (MARKAL) is given in (Suganthi and Samuel, 2012). The current trend is to develop hybrid models as they are seen to be more robust.

The Chapter is organized as follows. In Section 2.2 we review articles on modelling mean (average) daily peak electricity demand. Articles on modelling of extreme peak electricity demand using extreme value theory are discussed in Section 2.3. A review of literature on electricity demand forecasting in South Africa is discussed in Section 2.4 and concluding remarks are given in Section 2.5.

2.2 Modelling mean peak electricity demand

2.2.1 Weather variables

One of the most weather-sensitive sectors of any economy is the electricity sector. In this sector accurate prediction of daily peak electricity demand is very important. It provides short term forecasts which are required for dispatching and economic grid management of electric energy (Amaral *et al*., 2008; Amin-Naseri and Soroush, 2008; Chen, 1997; Franco and Sanstad, 2008; Mirasgedis *et al*., 2006; Psiloglou *et al*., 2009; Ismail *et al*., 2009; Ramanathan *et al*., 1997; among others). The most important weather factors which affect DPED is temperature. Changing weather conditions represent the major source of variation in peak demand forecasting and the inclusion of temperature has a significant effect due to the fact that in winter heating systems are used, whilst in summer air conditioning appliances are used (De Gooijer and Ray, 2003; Ghosh, 2008; Goia *et al*., 2010; Hekkenberg *et al*., 2009; Makridakis *et al*., 1998; Munoz and Felicisimo, 2004; Pilli-Sihvola *et al*., 2010; among others).
The predictive power of electricity demand models is improved if temperature is included in the models. This helps system operators and load forecasters to understand better factors that have a greater influence on electricity demand. Weather variables such as temperature, solar radiation, humidity, wind speed, cloudiness and rainfall are often used as explanatory variables in regression based load forecasting models (Munoz *et al*., 2010). Most authors however use temperature as the main driver (Munoz *et al*., 2010). Other weather factors include relative humidity, wind speed, cloud cover, precipitation and thunder showers. Electricity demand forecasting has received extensive attention in the literature using various techniques ranging from classical time series methods and neural networks to regression methods.

The effect of temperature on daily electricity demand has been studied exten-

sively in the energy sector. Taylor and Buizza (2003) use weather ensemble predictions in electricity demand forecasting for lead times of up to ten days ahead. Fifty one scenarios for weather-related components of electricity demand are used and the demand forecast uncertainty is then estimated using the distribution of the demand scenarios (Taylor and Buizza, 2003). The results from this study are then compared with univariate volatility forecasting methods and are found to be comparable (Taylor and Buizza, 2003). Regression models with autoregressive structures to reduce serial correlation are developed by Mirasgedis *et al.* (2006) to predict medium term electricity demand of up to 12 months ahead. Meteorological variables such as relative humidity and derived variables, heating and cooling degree-days are included in the models. The relationship between average daily temperature and daily electricity demand is known to be nonlinear. Mirasgedis *et al.* (2006) use a cubic regression curve to determine the reference temperature which separates winter temperatures from summer temperatures. However Mirasgedis *et al.* (2006) did not determine temperature values which separate the weather neutral period from the winter and summer sensitive periods. Moral-Carcedo and Vicéns-Otero (2005) analyse the effect of temperature on Spanish daily electricity demand using a Smooth Transition (STR), Threshold Regression (TR) and Switching Regression (SR) models. The models discussed determine appropriate threshold temperatures which separate the weather neutral period from the winter and summer sensitive periods. Results from this study show that the Logistic STR (LSTR) model captures adequately the non-linearity between electricity demand and temperature (Moral-Carcedo and Vicéns-Otero, 2005). The impact of temperature, holidays and other variables on daily peak electricity demand in Malaysia is investigated in Ismail *et al.* (2008). A time series regression model with autoregressive errors is used for out of sample predictions giving a MAPE of 1.71% (Ismail *et al.*, 2008). Psiloglou *et al.* (2009) present a comparison of characteristics of electricity demand and explore the relation-

ship with both climatic and non-climatic factors. Empirical results show that temperature plays the most important role in controlling the electricity load demand (Psiloglou *et al.*, 2009). The relationship between electricity demand and temperature is non-linear (Mirasgedis, 2006; Franco and Sanstad, 2008; Saini, 2008; Psilogu *et al*., 2009; Munoz *et al*., 2010; Pilli-Sihvola *et al*., 2010; among others). This non-linear relationship is modelled in literature using Heating Degree Days (HDD) and Cooling Degree Days (CDD). HDD and CDD are calculated using the following functions

$$\text{HDD} = \max(T_r - T_t, 0) \tag{2.1}$$

$$\text{CDD} = \max(T_t - T_r, 0) \tag{2.2}$$

where T_t is the average daily temperature at time t and T_r is the reference temperature. If $T_t < T_r$ there are no CDD and similarly if $T_t > T_r$ the are no HDD (Munoz *et al.*, 2010). It is important that T_r is selected in such a way that it correctly separates the winter sensitive temperatures from the summer sensitive temperatures of the load-temperature relationship (Munoz *et al.*, 2010). In addition to the use of HDD and CDD we also explore in this thesis, the use of the Multivariate Adaptive Regression Splines (MARS) method to determine temperature values which separate the weather neutral period from the summer and winter sensitive periods in the context of South Africa. MARS method is a non-parametric multivariate regression modelling approach which was developed by Friedman (1991) and has been used to solve high-dimensional problems with complex model structures, such as nonlinearities, interactions, multicollinearity and missing values (Chen, 1995; De Gooijer and Ray, 2003; Munoz and Felicisimo, 2004; Tsekouras *et al*.; among others). The method does not make any assumptions about the functional relationship between the response variable and the predictor variables (Friedman, 1991).

Hyndman and Fan (2010) propose a semi-parametric additive model in the regression framework with serially correlated errors to forecast long term peak electricity demand. The methodology proposed by Hyndman and Fan (2010) captures the nonlinear effect of temperature including other drivers of electricity demand such as calendar effects, price changes and economic growth. Fan and Hyndman (2012) extend the modelling approach by Ramanathan *et al.*, (1997) to a semi-parametric additive modelling framework to forecast short-term load demand. Half-hourly Australian electricity demand data is used. Each half-hourly period is modelled separately (Fan and Hyndman, 2012). Fan and Hyndman (2012) use regression splines to model temperature and lagged demand effects. The model developed by Fan and Hyndman (2012) is given in equation (2.3).

$$\log(y_{t,p}) \quad = \quad h_p(t) + f_p(\mathbf{w}_{1,t}, \mathbf{w}_{2,t}) + a_p(\mathbf{y}_{t-p}) + n_t \tag{2.3}$$

where $y_{t,p}$ denotes the demand at time t (measured in half-hourly intervals) during period $p(p = 1, ..., 48)$; $h_p(t)$ models all calendar effects; $f_p(\mathbf{w}_{1,t}, \mathbf{w}_{2,t})$ models all temperature effects where $\mathbf{w}_{1,t}$ is a vector of recent temperatures at Richmond (a suburb in Melbourne city in Australia) and $\mathbf{w}_{2,t}$ is a vector of recent temperatures at Melbourne airport; $a_p(\mathbf{y}_{t-p})$ models the effects of recent demands; n_t denotes the model error at time t. The model developed by Fan and Hyndman (2012) allows for nonlinear and nonparametric terms which can capture the non-linear relationship between electricity demand and its drivers.

Huang *et al.* (2012) develop a forecasting model for daily peak load in summer. The developed model captures meteorological factors such as season, temperature, humidity, precipitation, continuous sunny days, persistent drought, average temperature difference, atmospheric conditions, human factors and other factors (Huang *et al.*, 2012). Mestekemper *et al.* (2013) use a semi-parametric regression methodology to estimate the mean demand of electricity using both

calendar and meteorological variables. Results from this study show that the dynamic factor model provides more accurate forecasts over horizons of up to one week (Mestekemper *et al*., 2013). In contrast to these works we use a frequentist and Bayesian regression modelling framework to a piecewise linear regression model to explore the influence of temperature on electricity demand using South African data.

2.2.2 Statistical time series and regression models

Time series statistical models are pioneered by Box-Jenkins (1976). The SARIMA model developed by Box-Jenkins (1976) is the most commonly used model in short-term load forecasting (Munoz *et al*. 2010). The SARIMA model can be classified as either univariate or multivariate. Most of the univariate SARIMA models are generally used as naive (benchmark) models for comparative purposes with other models (see for example Taylor, 2008; Taylor and McSharry, 2008; Soares and Medeiros, 2008; Munoz *et al.*, 2010; among others). Seasonal ARIMA (SARIMA) and regression-SARIMA (also known as ARIMAX) models have been used extensively in electricity demand forecasting. These models are generally used for short term forecasting. The general SARIMA model can be written as

$$\phi_p(L)\Phi_P(L^s)\nabla^d\nabla_s^D x_t \;=\; \theta_q(L)\Theta_Q(L^s)\varepsilon_t \tag{2.4}$$

where x_t represents the load at time t, ε_t is the white noise error term at time t, s is the seasonal length, L is a lag operator, ∇^d and ∇_s^D are the nonseasonal and seasonal difference operators of order d and D respectively, where $\nabla^d = (1-L)^d$ and $\nabla_s^D = (1-L^s)^D$, and $\phi_p, \Phi_P, \theta_q, \Theta_Q$ and are polynomial functions of orders p, P, q and Q, respectively. Equation (2.4) can also be written as a single seasonal ARMA (SARMA) model (Taylor, 2010). The model is given in equation

(2.5).

$$\phi_p(L)\Phi_P(L^s)(x_t - a - bt) = \theta_q(L)\Theta_Q(L^s)\varepsilon_t \quad (2.5)$$

where a is a constant term and b is the coefficient of a linear deterministic trend term and the other terms are as defined in equation (2.1). The standard single seasonal ARMA model given in equation (2.2) can be extended to multiple seasonality models (Box *et al.*, 1994). A SARMA model for capturing both intraday and intraweek seasonal cycles in electricity data is proposed by Taylor (2008, 2010). The proposed model can be written as

$$\phi_p(L)\Phi_{P_1}(L^{s_1})\Omega_{P_2}(L^{s_2})(x_t - a - bt) = \theta_q(L)\Theta_{Q_1}(L^{s_1})\Psi_{Q_2}(L^{s_2})\varepsilon_t \quad (2.6)$$

The polynomial functions Φ_{P_1} and Θ_{Q_1} enable a SARMA modelling of the intraday seasonality while Ω_{P_2} and Ψ_{Q_2} are for modelling intraweek seasonality (Taylor, 2010). In order to model the intrayear seasonality in electricity data, Taylor (2010) extended the double seasonal ARMA (DSARMA) model given in equation (2.6) to a triple seasonal ARMA (TSARMA) model which is given in equation (2.7).

$$\phi_p(L)\Phi_{P_1}(L^{s_1})\Omega_{P_2}(L^{s_2})\Gamma_{P_3}(L^{s_3})(x_t - a - bt)$$

$$= \theta_q(L)\Theta_{Q_1}(L^{s_1})\Psi_{Q_2}(L^{s_2})\Lambda_{Q_3}(L^{s_3})\varepsilon_t \quad (2.7)$$

The terms Γ_{P_3} and Λ_{Q_3} represent lag polynomial functions for modelling the intrayear seasonality (Taylor, 2010).

The multiplicative double seasonal ARMA model in equation (2.3) is used by Taylor (2008) to model minute-by-minute British electricity demand data. Taylor, (2008) note that short term lead times are important for scheduling of elec-

tricity generation. ARIMA models, an adaptation of the Holt-Winters' exponential smoothing model and an exponential smoothing method that focuses on the evolution of intraday cycle are used (Taylor, 2008). Out of these methods the double seasonal adaptation of the Holt- Winters' exponential smoothing model gives the best results and this is consistent with results from previous studies (Taylor, 2008). Taylor (2010) extends the DSARMA model to modelling of British and French electricity data using TSARMA model given in equation (2.4). Empirical results from this study show that the TSARMA model outperforms the DSARMA and also a univariate neural network model for predictions of up to a day-ahead (Taylor, 2010).

Amjay (2001) propose a time series model for short-term hourly and daily peak load forecasting using Iranian data. Amjay (2001) compares the Box-Jenkins models with artificial neural networks. The proposed model outperformed the benchmark models (Amjay, 2001). McSharry *et al.* (2005) present a regression model which predicts probabilistic forecasts of future peak electricity demand. McSharry *et al.* (2005) argue that such forecasts can improve risk management and assessment of uncertainty in the forecasts. Aguirre *et al.* (2008) decompose load forecasting into dynamic prediction and pattern mapping. Real load data and surrogate data analysis are used as evidence in understanding fundamental issues underlying load forecasting problems (Aguirre *et al.*, 2008). Amin-Naseri and Soroush (2008) develop a hybrid neural network model for daily electrical Peak Load Forecasting (PLF). A novel approach for clustering data by using a self-organizing map is proposed by Amin-Naseri and Soroush (2008). It is concluded that the proposed hybrid model produces superior forecasts than those of the linear regression. The modelling and short-term forecast of daily power demand in the state of Victoria, Australia is studied by Truong *et al.* (2008). A two-dimensional wavelet based state dependent (SDP) modelling approach is adopted to formulate a compact mathematical model that is used to

forecast daily peak power demand. With a MAPE of 1.9%, the model is found to be effective. A peak load forecasting model for prediction of peak loads of up to 24 hours ahead using a neuro-evolutionary technique known as the Cartesian Genetic Algorithm evolved Artificial Neural Network (GGPANN) is developed in Khan *et al.* (2011).

Taylor (2008) states that accurate short-term forecasts are needed by both generators and retailers of electricity, particularly during periods of abnormal peak load demand. Accurate forecasts will enable effective load shifting between transmission substations. In order to improve forecast accuracy it is important to combine statistical forecasting methods together with judgmental techniques. Through experience, the judgmental experts develop intuitive relationships between electrical load and weather parameters, time of day, day of week, season and time lag of response, (Ismail *et al.*, 2009). On the other hand statistical techniques provide a scientific approach for producing consistent and accurate forecasts.

Amaral *et al.* (2008) develop a Smooth Transition Periodic Autoregressive (STPAR) model to forecast electricity load series data from Australia. STPAR is compared with other alternative load forecasting models. STPAR proves to be a useful tool when forecasting the electricity load (Amaral *et al.*, 2008). In their paper Sumer *et al*. (2009) develop ARIMA, SARIMA (seasonal ARIMA) and regression model with seasonal latent variable in forecasting electricity demand from Kayseri and Vicinity Electricity Joint-Stock Company. Their results show that the regression model with seasonal latent variable is more efficient than ARIMA and SARIMA. Soares and Medeiros (2008) consider a two-level method for hourly electricity load. A Two-Level Seasonal Autoregressive (TLSAR) model is developed and compared with a modified version of a SARIMA model called Dummy-Adjusted SARIMA (DASARIMA) (Soares and

Medeiros, 2008). A specific class of seasonal ARIMA models (the benchmark model) and the generalized long memory (GLM) model discussed by Soares and Souza (2006) are better than DASARIMA. A possible extension of this methodology to combining forecasts, interval forecasts and forecast density evaluation is suggested by Soares and Medeiros, (2008).

Regression-based methods have been used extensively in load demand forecasting (Ramanathan *et al*., 1997; Chou *et al*., 2004; Ghosh, 2008; Ismail *et al*., 2008; among others). These methods range from simple linear to multivariate linear regression models and work very well when the relationship between the dependent variable and the predictor variables is linear. They are usually easy to implement and give relatively robust solutions. Ismail *et al*. (2009) use a rule-based forecasting approach for forecasting peak electricity demand. The authors conclude that rule-based forecasting increases the forecast accuracy when compared to the traditional SARIMA model and that improvement depends on the conditions of the data, knowledge development and validation. Ramanathan *et al*. (1997) develop simple and flexible multiple regression models for hourly load forecasting. The regression modelling framework is in such a way that each one of the models is used for each hour of the day (Ramanathan *et al*., 1997) . Each of these regression models has a dynamic error structure as well as adaptive adjustments to correct for forecast errors of previous hours (Ramanathan *et al*., 1997). The structure of the multiple regression model developed by Ramanathan *et al*. (1997) is given as

$$y_{ht} = x_{ht}\beta_h + u_{ht} \tag{2.8}$$

where y_{ht} is the load on hour h on day t, x_{ht} represents the variables in the regression equation. The regression equation (2.8) captures important drivers of electricity demand which are temperature, lagged load variables including variables for the day of the week, month of the year and yearly trends. The

error term u_{ht} follows an autoregression structure given in equation (2.9)

$$u_{ht} = \sum_{i=1}^{5} \rho_{hi} u_{ht-i} + v_{ht} \tag{2.9}$$

The models developed perform well against a wide range of alternative models (Ramanathan *et al*., 1997). The major shortcoming of this modelling approach lies in treating special holidays as long weekends or the first of January (Munoz *et al.*, 2010). The treatment of special holidays such as Christmas, Easter or public holidays is usually done using dummy variables (Munoz *et al.*, 2010). Hong (2010) proposes an integrated systematic approach that incorporates interactions in multiple linear regression models for short term load forecasting. The proposed modelling framework produces accurate forecasts.

Cho *et al.* (2013) propose a hybrid linear regression model with both curve response and curve regressors for modelling and forecasting of daily electricity load curves using French data for the period 1996 to 2009. Empirical results show that the hybrid model is more adaptive to daily electricity consumption variations (Cho *et al.*, 2013).

2.2.3 Volatility forecasting

In conventional SARIMA models, the variance of the disturbance term is assumed to be constant. Electricity data normally exhibits non-constant mean and variance, and multiple seasonality corresponding to daily, weekly and monthly periodicity. Hence the assumption of homoskedasticity (constant variance) seems inappropriate. The GARCH modelling methodology is used to accommodate the possibility of serial correlation in volatility. Models for volatility forecasting were first developed by Engle (1982). The models, known as the Autoregressive Conditional Heteroskedasticity (ARCH) models were developed to capture the non-constant variance. ARCH models were later extended to

Generalized ARCH (GARCH) models by Bollerslev (1986). Electricity demand using volatility forecasting models are discussed in literature. Taylor (2006) argue that estimating an ARMA model jointly with an appropriate GARCH model is more efficient than estimating the ARMA model with an incorrect homoskedastic error assumption. Taylor (2006) investigates the methods for Net Imbalance Volume density forecasting using British electricity data. Taylor (2006) decomposes the problem into point forecasting and volatility forecasting. A seasonal ARMA model and a periodic AR model with simplistic volatility forecasting produce accurate forecasting results. An ARMA-ARCH model is used by Chen *et al.* (2006) to forecast electricity load giving results which compare favourably with those from ARIMA models. Hor *et al.* (2006) use GARCH models to predict daily load data and model the residuals under the assumption of the student-t distribution. The maximum load demand which is likely to occur in the short term is estimated. The GARCH-t model outperforms the GARCH model. ARMAX and GARCH models are used to assess the electricity demand pattern and the contribution of tourism to electricity consumption in Bakhat and Rosselo (2011). Chen *et al.*, (2011) model the volatility in short term load forecasting by combining regime-switching models with GARCH models, two types of regime-switching GARCH models, Threshold Auto-Regressive GARCH (TAR-GARCH) and Logistic Smooth Transition Auto-Regressive GARCH (LSTAR-GARCH) models. Empirical results show that the LSTAR-GARCH model produces more accurate forecasts compared to the other models. The LSTAR-GARCH model proposed by Chen *et al.*, (2011)

is given as

Conditional mean equation

$$y_t = E(y_t|\Psi_{t-1}) + \varepsilon_t \tag{2.10}$$

Conditional variance equation

$$h_t = \alpha_0 + \sum_{j=1}^{p} \beta_j h_{t-j} + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2$$

$$-\lambda \varepsilon_{t-1}^2 F(\varepsilon_{t-1} - \Delta) \tag{2.11}$$

where y_t is the load series data, ε_t denotes the error term, h_t represents the variance equation, Δ is the threshold and $F(\varepsilon_{t-1} - \Delta)$ is the LSTAR (Logistic Smooth Transition Autoregressive) written as

$$F(\varepsilon_{t-1} - \Delta) = \frac{1}{\exp\left(-\gamma(\varepsilon_{t-1} - \Delta)\right)} \tag{2.12}$$

Xu *et al*. (2012) uses ARCH type models to forecast and model long and short term load time series. The authors analyse volatility clustering, asymmetric impact and long memory volatility on the load series data. Empirical results from this study show that the developed models captures very well the volatility of the load time series. Kesavabhotla *et al.* (2012) use ARMA and GARCH models to predict day ahead electricity demand.

We extend the existing ARMA-GARCH and regression-SARIMA models to a regression-SARIMA-GARCH (REG-SARIMA-GARCH) model in Chapter 5. Taylor (2006) develop a SARIMA-GARCH model which captures seasonal cycles, that is daily, weekly and yearly seasonalities. In contrast our Reg-SARIMA-GARCH model captures important drivers of electricity demand such as calendar effects (day of the week, holidays) and temperature effects including multiple seasonality (daily, weekly, monthly and yearly seasonalities).

2.2.4 Bayesian forecasting of peak electricity demand

In Bayesian analysis the first step is to take into account information about the parameters including having access to expert information based on past experience or related studies (Marin *et al.*, 2003). This is modelled through a prior distribution $\pi(\boldsymbol{\theta})$, where $\boldsymbol{\theta}$ denotes the parameter space, that is $\boldsymbol{\theta} = \{\theta_1, ..., \theta_p\}$. If expert information is not available, non-informative priors are used. The second step involves collecting data and form the likelihood function, $\pi(\boldsymbol{x}|\boldsymbol{\theta})$ where $\boldsymbol{x} = \{x_1, ..., x_n\}$. The likelihood describes how the data $\boldsymbol{x}$ depends on the parameter $\boldsymbol{\theta}$ (Marin *et al.*, 2003). The next step combines the prior distribution with the likelihood function to get the posterior distribution, $\pi(\boldsymbol{\theta}|\boldsymbol{x})$, which is the distribution of the parameter $\boldsymbol{\theta}$ given the data $\boldsymbol{x}$. The posterior distribution is given as

$$\pi(\boldsymbol{\theta}|\boldsymbol{x}) \quad = \quad \frac{\pi(\boldsymbol{x}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\int_{\boldsymbol{\theta}} \pi(\boldsymbol{x}|\boldsymbol{\theta})\pi(\boldsymbol{\theta}) d\boldsymbol{\theta}} \tag{2.13}$$

which is often written as

$$\pi(\boldsymbol{\theta}|\boldsymbol{x}) \quad \propto \quad \pi(\boldsymbol{x}|\boldsymbol{\theta})\pi(\boldsymbol{\theta}) \tag{2.14}$$

that is, the posterior is proportional to the likelihood function multiplied by the prior distribution. The denominator in equation (2.13) is a normalizing constant. Computing the normalizing constant is usually done using Markov Chain Monte Carlo (MCMC) techniques. All inferences will then be based on the posterior distribution since it summarizes all available information about the parameters (Marin *et al.*, 2003). In point estimation, for example, the mean $E(\boldsymbol{\theta}|\boldsymbol{x}) = \int_{\boldsymbol{\theta}} \boldsymbol{\theta}\pi(\boldsymbol{\theta}|\boldsymbol{x}) d\boldsymbol{\theta}$ can serve as an estimate of the parameter $\boldsymbol{\theta}$. Future values of $\boldsymbol{x}$ can also be predicted by using the posterior predictive distribution

$\pi(x_{n+1}|\boldsymbol{x})$ which is given by

$$\pi(x_{n+1}|\boldsymbol{x}) = \int_{\boldsymbol{\theta}} \pi(\boldsymbol{x}|\boldsymbol{\theta})\pi(\boldsymbol{\theta}|\boldsymbol{x})d\boldsymbol{\theta} \tag{2.15}$$

Most papers in literature concentrate on frequentist point forecasting only. One major drawback of frequentist point forecasting only is that it does not take into account uncertainty in the estimation of the parameters. One way of overcoming this, is the use of Bayesian analysis and density forecasting. For electricity load forecasting the Bayesian approach provides a full range of sample inferences which includes the predictive distribution of electricity demand and the distribution of the time and level of daily peak demand (Cottet and Smith, 2003). This is not easy using classical statistics. Cottet and Smith (2003) applied the Bayesian approach to New South Wales (Australia) total system load data and were able to decompose the load into meteorological, seasonal, trend and day-type effects.

Bayesian analysis does not depend on regularity conditions required for maximum likelihood estimation of parameters of extreme value distributions (Coles, 2001; Beirlant *et al.*, 2004). For a detailed discussion on the advantages and disadvantages of using Bayesian analysis see Berger (1985); Berger and Wolpert (1988); Bernado and Smith (1994); Carlin and Louis (2000); Wade (2000); Robert (2001); Wasserman (2004); Hugo (2012); among others).

Chaturvedi (1996) discusses the use of robust Bayesian analysis. The author concludes that the use of robust Bayesian analysis overcomes the major drawback of coming up with a correct elicitation of a prior distribution. Kiartzis *et al.* (1997) use a Bayesian combination method to predict short-term electric loads using data from Greece. The hybrid model developed by Kiartzis *et al.* (1997) is a weighted sum of artificial neural network predictor and two linear regression predictors. Ohtsuka *et al.* (2010) use a Bayesian spatial autoregres-

sive ARMA approach to forecast electricity demand in Japan. The developed model captures spatial heterogeneity and spatial correlation simultaneously (Ohtsuka *et al.*, 2010). Ohtsuka *et al.* (2010) conclude that spatial correlation play an important role in forecasting electricity demand in Japan. A Bayesian neural network learned by a hybrid Monte Carlo algorithm for short-term load forecasting is discussed in Niu *et al.* (2012). In contrast our work proposes the use of the generalized gamma distribution in predicting annual hourly peak electricity demand using South African data. We model the annual peak load as a function of gross domestic product using the Bayesian framework.

2.3 Modelling extreme peak electricity demand using extreme value theory

Extreme Value Theory (EVT) is a powerful and fairly robust framework for modelling the tail behaviour of a wide class of distributions (Gencay and Selcuk, 2004). In this thesis we model the tail behaviour of extreme daily peak electricity demand using the Generalized Extreme Value Distribution (GEVD), the Generalized Pareto Distribution (GPD) and the Generalized Single Pareto Distribution (GSPD). For an overview of EVT see Coles (2001); Bierlant *et al*. (2004); de Haan and Ferreira (2006); Reiss and Thomas (2007); among others. The EVT distributions are presented in the following sections.

2.3.1 Generalized Extreme Value Distribution (GEVD)

GEVD is a family of parametric distributions which are the Gumbel, Frechet and Weibull distributions. The distributions are also known as the type I, II and III extreme value distributions respectively (Hor *et al.*, 2008). The asymptotic distribution of maxima belongs to one of the three distributions of the GEVD irrespective of the distribution of the observed data (Fisher and Tippett,

1928). This is a convenient way of forecasting electricity demand uncertainty as the distribution of demand can change over time due to climatic changes, air-conditioning demand and societal changes (Hor *et al.*, 2008). The generalized extreme value family of distributions (von Mises, 1936) is given by

$$G_\gamma(x) = \exp\left\{-\left[1+\gamma\left(\frac{x-\mu}{\sigma}\right)\right]^{-\frac{1}{\gamma}}\right\},\ \text{if}\ 1+\gamma\left(\frac{x-\mu}{\sigma}\right) > 0, \gamma \neq 0 \tag{2.16}$$

where γ is the extreme value index (EVI). For $\gamma = 0$ (the distribution function is interpreted as a limit as $\gamma \to 0$) equation (2.16) is the Gumbel class of distributions which is given by

$$G_\gamma(x) = \exp\left\{-\exp\left[-\frac{x-\mu}{\sigma}\right]\right\}, \gamma = 0 \tag{2.17}$$

When $\gamma > 0$ we have the Frechet class of distributions and when $\gamma < 0$ we have the Weibull class of distributions. GEVD models block maxima and this makes it an appropriate distribution for modelling daily peak electricity demand. In this thesis hourly electricity data is put into blocks of 24 hours each and from each block the maximum is taken. EVT has been applied in various fields such as flood frequency analysis (Katz *et al.*, 2002), financial time series forecasting (Gencay and Secluk, 2004) including weather and climate studies (Cooley, 2005) among others.

Application of EVT in the electricity sector is discussed in literature. Chan and Gray (2006) propose a model that accommodates autoregression and weekly seasonalities in both the conditional mean and conditional volatility of daily electricity spot price returns. The tails of the distribution are then modelled using the EVT approach. The developed EVT-based model performs well in forecasting out-of-sample value-at-risk (VAR). VAR is a measure of the risk of a portfolio. Electricity returns are highly volatile and display seasonalities in both their mean and as well as volatility, exhibit leverage effects and cluster-

ing in volatility, and feature extreme levels of skewness and kurtosis (Chan and Gray, 2006).

The use of extreme value distributions requires that the assumptions of independent and identical distributed observations are met (McNeil and Frey, 2000). These assumptions provide obstacles to the straight forward application of extreme value to both financial market returns and electricity return series (McNeil and Frey, 2000; Bystrom, 2005). Using a two stage approach, McNeil and Frey (2000) estimate a GARCH model in stage one with a view to filtering the return series to get nearly independent and identical distributed residuals. In stage two, the EVT framework is then applied to the standardized residuals. An application of the McNeil and Frey (2000) modelling approach to electricity demand forecasting is discussed in literature. Bystrom (2005) applies a generalized Pareto Distribution (GPD) to an autoregressive GARCH filtered price change series. Empirical results from this study show that a peaks-over-threshold method provides accurate results in modelling tails of hourly electricity price changes (Bystrom, 2005). However Leadbetter (1974) shows that the assumptions of independent and identical distributed observations hold for a stationary series. This is given in the following theorem:

Theorem 2.1. : *(Leadbetter, 1974)*
Let $\{X_n\}$ *be a stationary sequence for which there exist sequences of constants* $a_n > 0$ *and* b_n *and a non-degenerate distribution function* G *such that*

$$P\left(\frac{M_n - b_n}{a_n} \leq x\right) \xrightarrow{D} G(x)$$

as $n \to \infty$. *If* $D(u_n)$ *condition holds with* $u_n = a_n x + b_n$ *for each* x *such that* $G(x) > 0$*, then* G *is an extreme value distribution*

The $D(u_n)$ condition, Leadbetter (1974), states that for any two events of the form $\{M(I_1) \leq u_n\}$ and $\{M(I_2) \leq u_n\}$ can become approximately independent

as $n \to \infty$ when the index sets $I_i \subset \{1, ..., n\}$ are separated by a relatively short distance $s_n = o(n)$ (Beirlant *et al.*, 2004). The proof of Theorem 2.1 is given in Leadbetter *et al.* (1983). For a detailed discussion of the requirement of independent and identical distributed observations see Leadbetter (1983); Cooley (2005); among others.

2.3.2 Generalized Pareto Distribution (GPD)

In extreme value theory there are times when we are not only interested in modelling maxima or minima but with observations which are above a high threshold. Extreme electricity demand is one such example. Balkema and de Haan (1974) and Pickands (1975) show that the distribution function of the excesses above a high threshold converges to a Generalized Pareto Distribution (GPD) as the threshold tends to the right endpoint. The GPD is a Peaks Over Threshold (POT) distribution whose distribution function is given as

$$W_\xi(x) = 1 - \left(1 + \frac{\xi(x-\tau)}{\sigma}\right)^{-\frac{1}{\xi}} \text{ if } \xi \neq 0 \tag{2.18}$$

Smith (1987) shows that estimation of GPD parameters with the maximum likelihood method is a non regular problem when $\xi < -\frac{1}{2}$, where ξ is the shape parameter which is also known as the Extreme Value Index (EVI). This means that certain regularity conditions have to be met. It is therefore more attractive to use the Bayesian approach which does not depend on these regularity conditions (Beirlant *et al.,* 2004; Coles, 2004). Since the peaks over threshold (POT) methodology involves tail estimation based on small data sets with little information available, the Bayes approach can be used to capture and take into account all the available information including additional information through prior elicitation (Beirlant *et al.*, 2004). However Hosking and Wallis (1987) show that the regularity conditions by Smith (1987) are satisfied for $\xi > -\frac{1}{2}$ and that the MLEs are asymptotically normally distributed.

Fitting a GPD to exceedances over a sufficiently large threshold has been discussed in literature. Castillo and Hadi (1997) discuss the fitting of the GPD to a given set of data. The estimation of the two parameters of the GPD which are the shape and scale parameters is usually not easy. In their paper Castillo and Hadi (1997) propose a method for estimating the parameters and quantiles of the GPD. Their proposed method works well over a wide range of parameter values. A Bayesian analysis of extreme events considering uncertainty about the threshold is discussed in Behrens *et al.* (2004). Behrens *et al.*, (2004) consider the threshold as a parameter in each of the models developed and prior information is allowed to be incorporated into the analysis. The proposed models are applied to a financial market index, Nasdaq 100. Empirical results show that the proposed gamma GPD model outperforms the other distributions (Behrens *et al.*, 2004).

Threshold selection

Modelling extreme value data using POT distributions such as a GPD requires estimation of a sufficiently high threshold. The threshold estimation requires a balance between asymptotic tail approximation bias and quantification of parameter estimation uncertainty (Scarrott and MacDonald, 2012). Threshold estimation methods are classified as classical fixed threshold approaches, resampling based methods, tail fraction estimation and mixture models (Scarrott and MacDonald, 2012). Classical fixed threshold methods use graphical diagnostics to select a threshold (Coles, 2001; Katz *et al.*, 2002; Beirlant *et al.*, 2004; among others). Several methods are discussed in literature which use tail fraction estimation together with graphical diagnostics. For details see Hill (1975); Drees *et al.* (2000); Beirlant *et al.* (2004); Goegebeur *et al.* (2008); among others. One of the drawbacks is that these tail fraction estimation methods use asymptotic optimality-based arguments under various population distribution assumptions (Scarrott and MacDonald, 2012). To overcome this drawback re-

sampling based methods which require weaker assumptions are used (Scarrott and MacDonald, 2012). For details on resampling methods see Hall (1990); Danielsson *et al.* (2001); Ferreira *et al.* (2003); Beirlant *et al.* (1996); Beirlant *et al.* (2004); among others. One of the drawbacks of using classical fixed threshold approaches, resampling based and tail fraction estimation methods is that the uncertainty associated with the threshold selected is not taken onto account (Scarrott and MacDonald, 2012). To overcome this drawback mixture models are proposed in literature. In mixture models the entire distribution function is approximated in such a way that the distribution below the threshold is not influenced by the tail fit (Scarrott and MacDonald, 2012). Mixture methods are discussed in Behrens and Castellanos (2010); MacDonald (2012); Scarrott and MacDonald (2012); among others. A detailed review of threshold selection methods is found in (De Zea Bermudez and Kotz, 2010; Scarrott and MacDonald, 2012).

2.3.3 Generalized Single Pareto Distribution (GSPD)

The Generalized Single Pareto Distribution (GSPD) discussed in Verster and De Waal (2011), is an approximation to the GPD which is a POT distribution used in extreme value theory to model observations above a sufficiently high threshold. The GSPD has one parameter, which is the EVI. The distribution function of the GSPD distribution is given as

$$W_\eta(x) = \left\{1 + \frac{\eta}{1+\tau\eta}\left(x-\tau\right)\right\}^{-\frac{1}{\eta}}, \eta \neq 0, x > \tau \tag{2.19}$$

where η is the EVI. Verster and De Waal (2011) show that the tail of a Generalized Burr-Gamma (GBG) distribution can be approximated by a Generalized Pareto-type (GP-type) distribution. The GBG class of distributions is popular in extreme value theory when the whole data set is modelled (Verster and De Waal, 2011). The GP-type distribution developed by Verster and De Waal

(2011) belong to the domain of attraction of the GPD. This modelling approach is convenient since the GP-type distribution has only one parameter to estimate given that the threshold is known, compared to the two parameters of the usual GPD. As such it is referred to as the GSPD. Verster and De Waal (2011) use the Bayesian approach to estimate the EVI.

Modelling of extreme peak loads using statistical extreme value theory is discussed in literature. The extreme value distribution (EVD), in conjunction with Monte Carlo simulations, is used to analyse sources of uncertainty in forecasting annual peak power loads (Bezler, 1993). An application of McNeil and Frey's (2000) modelling approach to electricity demand forecasting is discussed in literature. Bystrom (2005) applies a generalized Pareto distribution (GPD) to an autoregressive GARCH filtered price change series. Empirical results from this study show that a peaks-over-threshold method provides accurate results in modelling tails of hourly electricity price changes (Bystrom, 2005). Mori (2005) uses the POT approach to evaluate the risk which exceeds the upper bound of generation capacity. The proposed method is successfully applied to real data of one-step ahead daily maximum load forecasting. Hor *et al.* (2008) use the generalized extreme value (GEV) theory and block maxima approach to estimate the maximum load forecast errors in order to assess risk in long-term electricity load forecasting. Constable *et al.* (2012) develop a planning tool for predicting extreme peak loads using the generalized extreme value distributions. Results from this study show that a given extreme peak load which occurs once every 20 years is likely to occur annually by around 2050. The authors conclude that such information is important for future electrical power planning needs. Chidodo and Lauria (2012) propose a Bayesian modelling framework for estimating the frequency of occurrence of peak loads. The procedure is based on the assumption that peak power follows a Poisson process (Chidodo and Lauria, 2012). Results from this study show that the pro-

posed estimation methodology is highly efficient (Chidodo and Lauria, 2012). The model proposed by Chidodo and Lauria (2012) is given as

$$P_M(k,t) \equiv P\left[M(t)=k\right] = e^{-\phi q t} \times \frac{(\phi q t)^k}{k!} \tag{2.20}$$

where ϕ is the expected number of peak load occurrence over a threshold b, $q = 1 - F(b) = P(W_j > b)$ is the exceedance probability of the value b by any single peak load amplitude W_j, $M(t)$ is a Poisson process representing the number k of extreme peak load amplitudes occurring in time interval $(0, t)$.

Our work is related to that of Hor *et al.* (2008) who model load forecast uncertainty using extreme value theory and Chidodo and Lauria (2012) who estimate the frequency of occurrence of peak loads. We use the modelling approach of Hor *et al.* (2008) to model electricity demand uncertainty and carry out a frequency analysis of the days when extreme daily increases in peak electricity demand occur in South Africa. We then take these days as critical peak days in which the power utility company is expected to charge appropriate cost reflective tariffs. We define a critical peak day as one in which the system demand for electricity in a given day exceeds normal demand or when total outages on a given day are higher than normal. It should be noted that critical peak days can also be determined by considering generating capacity constraints of a power utility company during abnormal electricity peak demand which may be caused by for instance extreme weather conditions.

2.4 Electricity demand in South Africa

2.4.1 Modelling mean daily peak electricity demand

Since 2007 South Africa experienced a lack of capacity generation. Tassel (2003) explains that the gap between installed generating capacity and the

total electricity demand is narrowing. The author warned that South Africa will be in throes of power crisis as early as 2007. There were blackouts which resulted in load shedding in the first quarter of 2008. This affected severely sectors of the South African economy (Inglesi, 2010). In 2008 Eskom designed intervention strategies needed to minimize the risk of load shedding and a possible collapse of the national electricity supply system. A detailed discussion of these intervention strategies is found in National Response to South Africa's Electricity Shortage (2008). Prior to the 2008 power crisis there were few energy research centres and institutions in South Africa. Amongst these are Energy research Centre based at the university of Cape Town and Eskom's research division (Energy Markets: Economics, Statistics and Decision Sciences (n.d.)).

The power crisis of 2007-2008 was worsened by the global recession of 2008-2009. After the power crisis of 2007-2008 and global recession of 2008-2009, additional energy demand forecasting centres were established. The Centre for Energy Research at Nelson Mandela Metropolitan University now has a unit on energy demand forecasting and modelling. The newly established energy research centre at University of Pretoria focuses on energy production, energy distribution, energy optimization, policy, economics and society among others. The university of the Free State concentrates on energy risk. In an effort to reduce electricity demand especially during peak periods, Eskom has embarked on a Critical Peak Day Pricing Project (CPDPP) (Eskom CPDPP Website (n.d.)). The South African Department of Energy's division of Energy Efficiency and Environment (EEE) is currently working on some projects on energy efficiency designed to save energy in all sectors of the economy (EEE website (n.d.)).

Electricity demand forecasting in South Africa is discussed in literature. Amusa *et al*. (2009) apply the bounds testing approach to cointegration within an au-

toregressive distributed lag framework to examine the aggregate demand for electricity in South Africa during the period 1960-2007. Their findings are that, in the long run income will be the main determinant of electricity demand. Ziramba (2008) examines the residential demand for electricity in South Africa as a function of real gross domestic product per capita and the price elasticity during the period 1978-2005. The bounds testing approach to cointegration within an autoregressive distributed framework is used. The linear double-logarithmic form is used with income and price as independent variables. Findings of this study are that in the long run, income would be the main determinant of electricity demand and that electricity price is insignificant. Debba *et al.* (2010) present forecasts for electricity demand in South Africa for the period 2010-2035 using the CSIR sectoral regression model.

In an effort to reduce electricity demand, Eskom's demand side management (DSM) division has had projects running since 1994 (Eskom DSM website (n.d.)). DSM often called integrated demand management (IDM) in Eskom includes load management, peak clipping and energy efficiency (Eskom IDM website (n.d.)) in all sectors of the South African economy, i.e. residential, commercial and industrial. Research on load management is ongoing. Boake (2003) develops an electrical load management (ELM) strategy in the South African mining industry. Results from this study show that the developed ELM can be used to reduce input costs which will ultimately reduce electrical energy costs. Some of the work done on daily peak demand forecasting in South Africa is not published as it is done mainly by private consulting firms. Accurate peak load forecasts are important as they assist decision makers in improving capital expenditures which include securing of adequate capacities for generation, transmission and distribution.

2.4.2 Modelling extreme daily peak electricity demand in South Africa

Use of extreme value theory in energy research in South Africa is discussed in literature. Micali (2007) use EVT to assess the occurrences of excessive unexpected electricity losses experienced by electrical generating units in South Africa. Micali (2007) use both the GEVD and the GPD to model the tail of the gigawatt-hour losses through the Bayesian approach. Empirical results from this study show that the GPD fits very well the gigawatt-hour losses above a sufficiently high threshold (Micali, 2007).

Our modelling approach differs from all these in that we use EVT to help in the quantification of the amount of electricity which should be shifted from the grid to off peak hours. Frequency analysis of the occurrence of extreme daily increases and same day of the week increases in peak electricity demand is carried out. This helps in identifying the days and months when those abnormal large increases will occur. This modelling approach provides a policy framework for load curtailment and determining the number of critical peak days for a power utility company. Verster and De Waal (2011) use the Bayesian approach to estimate the EVI. We extend this to estimating EVI using the maximum likelihood estimation method. Appropriate software is then developed using R.

Our work concentrates on modelling the highest demand of electricity (hourly peak demand) recorded in a 24-hour period (which is a day). We also explore the use of the Generalized Gamma (GG) distribution in predicting annual hourly peak electricity loads in South Africa for the period 1996 to 2010. We model the annual peak load as a function of gross domestic product using the Bayesian framework.

2.5 Concluding remarks

Modelling the mean and extreme peak electricity demand has been reviewed in this Chapter. Conventional electricity demand forecasting models discussed in literature consider mainly mean changes. At times there is need to depart from the average thinking and exploit information provided by the extremes. This helps to understand how extremes affect forecasts and in assessing daily peak electricity demand uncertainty. This knowledge is important to decision makers in power utility planning and reduction of operational costs. Modelling of extreme daily peak electricity demand helps to improve the reliability of a power network if an accurate assessment of the level and frequency of future extreme load forecasts is carried out (Hor *et al.*, 2008). Bayesian EVT modelling helps to understand how extreme peak electricity demand can be modelled incorporating uncertainties in the parameters.

Chapter 3

Modelling the influence of temperature on daily peak electricity demand in South Africa using a multivariate non-parametric regression approach

3.1 Introduction

Changing weather conditions are a major contribution to variation in Daily Peak Electricity Demand (DPED) (Munoz *et al.*, 2010). Temperature is the main variable that is usually included in load forecasting models (Hippert *et*

al. 2001a). This chapter presents empirical results on the influence of temperature on daily peak electricity demand in South Africa using Multivariate Adaptive Regression Splines (MARS) models.

One of the challenges for South African data is to determine thresholds for separating the winter-sensitive period (T_w) from the weather-neutral period and also the summer-sensitive period (T_s) from the weather-neutral period. The MARS modelling approach is used in this thesis to determine these two reference temperatures, T_w and T_s respectively. The MARS method is discussed in Section 3.2.

3.2 The Multivariate Adaptive Regression Splines (MARS) model

MARS is a non-parametric multivariate regression method which was developed by Friedman (1991). It is used to solve problems with complex model structures, such as multicollinearity, interactions including nonlinearities (Craven and Wahba, 1979; Chen, 1997; Munoz and Felicisimo, 2004; Tsekouras *et al.,* 2007; among others). In MARS modelling no assumptions about the functional relationship between the dependent variable and the regressor (independent) variables are made (Friedman, 1991). Modelling using MARS involves splitting the modelling space into sub-regions (sub-spaces) and fitting simple linear regression models in each of the sub-regions identified. The MARS model building process involves the construction of a large number of basis functions using a forward stepwise algorithm which over-fits the data (Friedman, 1991). In the backward stepwise step basis functions are deleted in order of least contribution using the generalized cross validation (GCV) criterion (Craven and Wahba, 1979). The general MARS model developed by Friedman (1991) will be

fitted to South African data and is given as

$$f(x) = \omega_0 + \sum_{m=1}^{M} \omega_m \prod_{k=1}^{K_m} \left[s_{km}(x_{v(k,m)} - t_{km}) \right] \tag{3.1}$$

where ω_0 and ω_m are parameters, M is the number of basis functions, K_m is the number of knots, s_{km} takes on values of either 1 or -1 indicating the right or left sense of the associated step function, $v(k,m)$ is the label of the independent variable and t_{km} represents the knot location. The basis function $B_m(x)$ (Friedman, 1991) is given as

$$B_m(x) = \prod_{k=1}^{K_m} \left[s_{km}(x_{v(k,m)} - t_{km}) \right] \tag{3.2}$$

The GCV criterion is a measure of the goodness of fit which takes into account the residual error and the model complexity (Craven and Wahba, 1979). In its simplest form the GCV criterion (Craven and Wahba, 1979) may be written as

$$\text{GCV}(M) = \frac{\frac{1}{N} \sum_{i=1}^{N} \left[y_i - \hat{f}_M(x_i) \right]^2}{\left[1 - \frac{C(M)}{N} \right]^2} \tag{3.3}$$

where N is the sample size, and $C(M)$ is the cost-penalty measure of a model containing M basis functions. The numerator measures the lack of fit on the M basis function model $\hat{f}_M(x_i)$ and the denominator represents the penalty for the model complexity $C(M)$. The complexity cost function (Friedman, 1991) may be written as

$$C(M) = \text{trace}(B(B^T B)^{-1} B^T) + 1 \tag{3.4}$$

where B is the $M \times N$ data matrix of the M (non constant) basis functions $(B_{ij} = B_i(x_j))$. The best model is one with the lowest GCV criterion value. The MARS model developed for DPED and T_t as the predictor variable is given in

equation (3.5).

$$\text{DPED} = r_0 + r_1 \text{max}\left\{0, T_t - T_w\right\} + r_2 \text{max}\left\{0, T_s - T_t\right\} + \varepsilon_{t(1)} \tag{3.5}$$

where r_0, r_1 and r_2 are constants, and the parameters T_w and T_s have meanings as declared in Section (3.1) and and $\varepsilon_{t(1)}$ is the error term.

Using Average Daily Electricity Demand (ADED) and average daily temperature (T_t) the MARS model for assessing the influence of temperature on electricity demand is given in equation (3.6).

$$\text{ADED} \quad = \quad k_0 + k_1 \text{max}(0, T_s - T_t) + k_2 \text{max}(0, T_t - T_w) + \varepsilon_{t(2)} \tag{3.6}$$

where k_0, k_1 and k_2 are constants, T_s and T_w are as defined in Section (3.1) and $\varepsilon_{t(2)}$ is the error term.

3.3 Load and temperature data

3.3.1 Load data

Electricity load is equal to electricity demand in the absence of blackouts and load-shedding (Cottet and Smith, 2003). Aggregated DPED data from all sectors of the South African economy for the period 2000 to 2010 is used. The plot of DPED (top left panel) in Figure 3.1 shows that DPED is not stationary, exhibits a positive upward trend and strong seasonality. A spectral analysis is carried out to investigate the periodicity in the data. The spectral density in Figure 3.2 shows a seven day periodicity.

The data is made stationary by using the following methods: Figure 3.3 shows DPED after taking a seventh difference (which is seasonal differencing since data exhibits a seven day periodicity (see Figure 3.2)). The data is now sta-

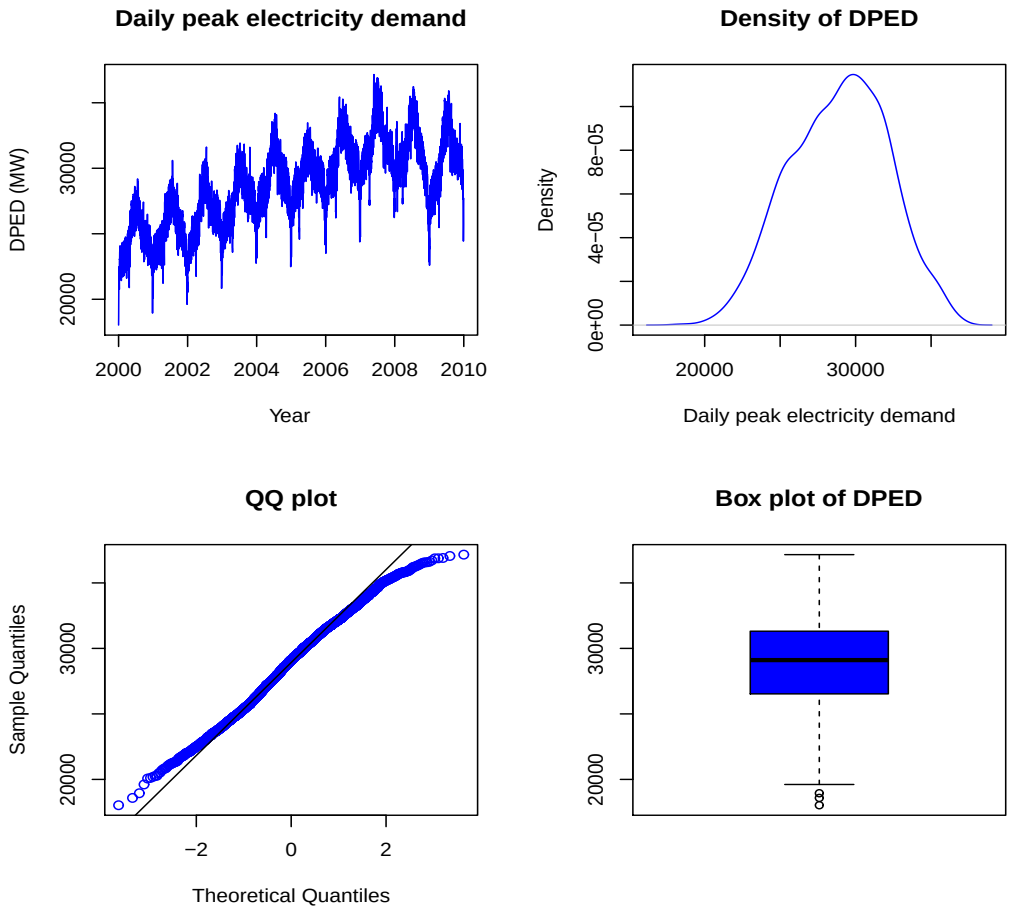

Figure 3.1: (a) Top left panel: Plot of DPED (b) Top right panel: Probability density function of DPED. The distribution is left skewed. (c) Bottom left panel: Normal QQ plot of DPED (d) Bottom right panel: Box plot of DPED.

tionary. The bottom panel of Figure 3.3 shows the graphs of the Autocorrelation Function (ACF) of the residuals and the Partial Autocorrelation Function (PACF) of the residuals after taking the seventh difference.

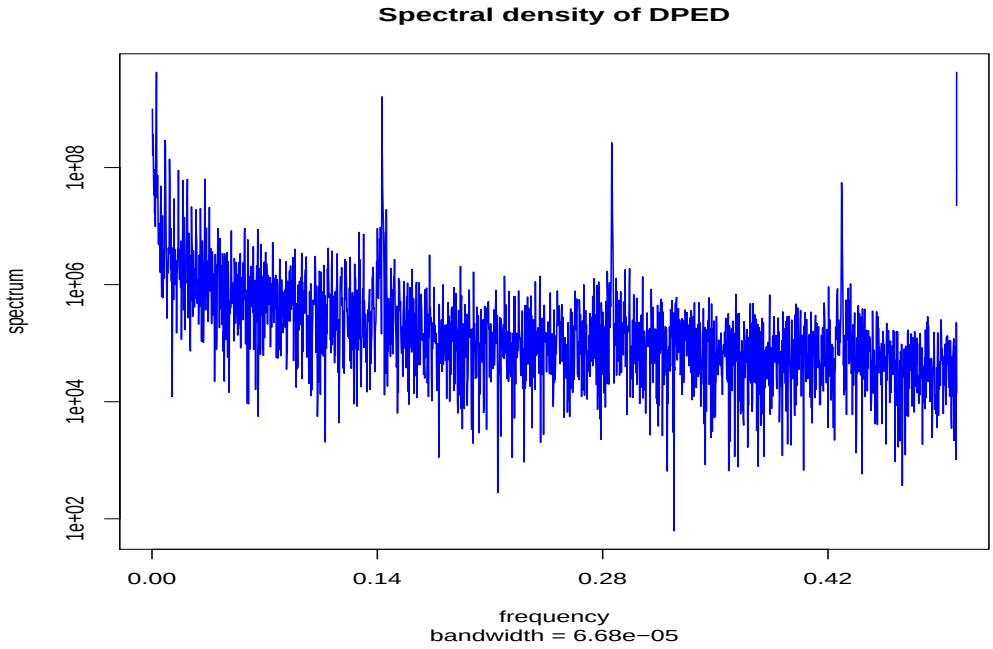

Figure 3.2: Spectral density of DPED showing a seven day periodicity.

A visual inspection of the probability density function of DPED (for the whole data set in the historic period, 2000-2010) which is estimated using kernel density estimation (Silverman, 1986) and is plotted in Figure 3.1 shows that the data is not normally distributed. The normal QQ and box plots also given in Figure 3.1 confirms the non normality of the DPED data. Graphs of similar probability densities for DPED given in Figures 3.12 and 3.13 respectively show the gradual changes in the shape of the densities throughout the years 2000 to 2009. These changes are likely due to economic growth. Figures 3.14, 3.15 and 3.16 (given in appendix 3.3 at the end of the chapter) show a gradual change in the probability densities for each of the hours of the day. This is possibly due to day to day variations in electricity usage.

We then carry out a frequency analysis of the occurrence of DPED in different hours of the day. The frequency of DPED occurring in different hours of the day are given in Table 3.1 and shown in Figure 3.4. Daily peak electricity demand occur at 20:00 on 46% of the days in the data set. This is followed

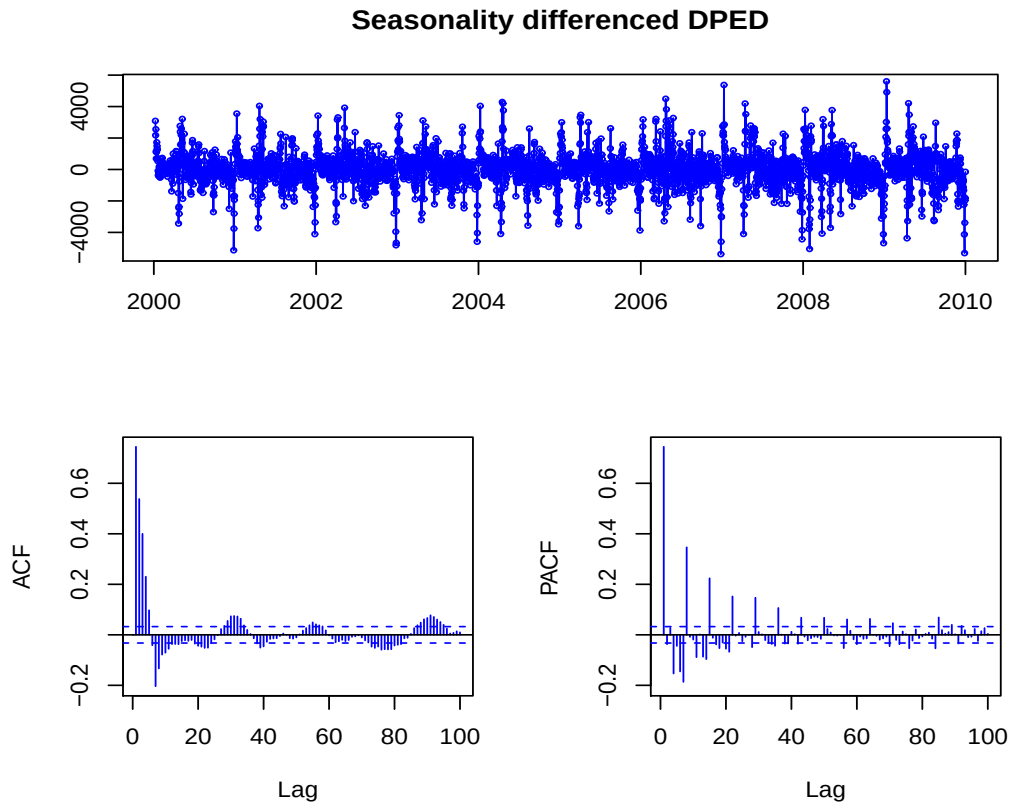

Figure 3.3: (a) Upper panel: Seasonality differenced DPED data (b) Bottom left panel: Autocorrelation function of the residuals and (c) Bottom right panel: Partial autocorrelation function of the residuals.

by an occurrence of DPED on 41% at 19:00. On some occasions DPED is also experienced around 21:00 for 5% of the time. In the historic data set electricity demand also peaks in the morning around 09:00 for 3% of the time and 10:00 for 5% of the time.

3.3.2 Temperature data

Historical hourly data on temperature from 36 meteorological stations from all the provinces of South Africa is used. The hourly data provided which was already in aggregated form from the 36 meteorological stations is for the period 2000 to 2010. The meteorological stations are: Upington, Calvinia, De Aar, Kimberley, Beaufort West, Cape Town, Langebaanweg, George, East Lon-

Table 3.1: Frequency of occurrence of DPED in different hours of the day.

Hour	1	2	3	4	5	6	7	8	9	10	11	12
Frequency	0	0	0	0	0	0	0	5	124	232	0	0
Hour	13	14	15	16	17	18	19	20	21	22	23	24
Frequency	0	0	0	0	1	3	1734	1968	204	0	0	0

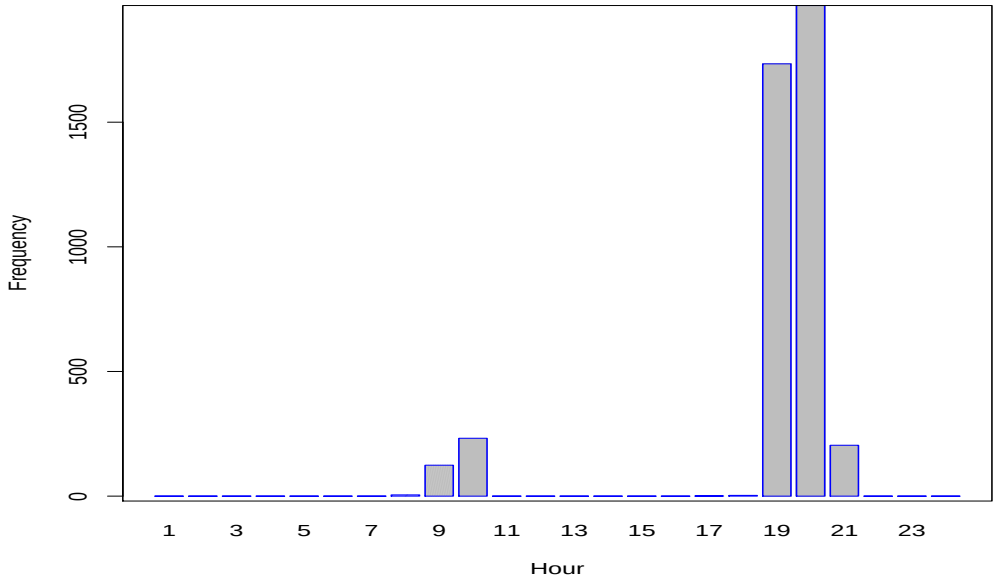

Figure 3.4: Frequency of daily peak demands occurring in different hours of the day for the years 2000 to 2010.

don, Port Elizabeth, Tsitsikama, Umtata, Bethlehem, Bloemfontein, Welkom, Vereeniging, Ladysmith, Pietermaritzburg, Durban, Richards Bay, Airport, Belfast, Nelspruit, Witbank, Johannesburg, Pretoria , Eendracht, Irene, Pilanesberg, Rustenburg, Polokwane, Thohoyandou, Ellisras, Mafikeng, Margate and Newcastle. Table 3.2 gives a summary of the descriptive statistics of temperature.

Table 3.2: Descriptive statistics for average temperature.

	Mean	Median	Max	Min	Std Dev	Skew	Kurtosis
Maximum Temp	23.9	24.2	32.2	10.8	3.5	-0.355	2.66
Minimum Temp	13.0	13.9	21.4	-0.2	4.5	-0.397	1.99

A plot of the minimum, average and maximum daily temperatures together with their respective densities are given in Figure 3.5.

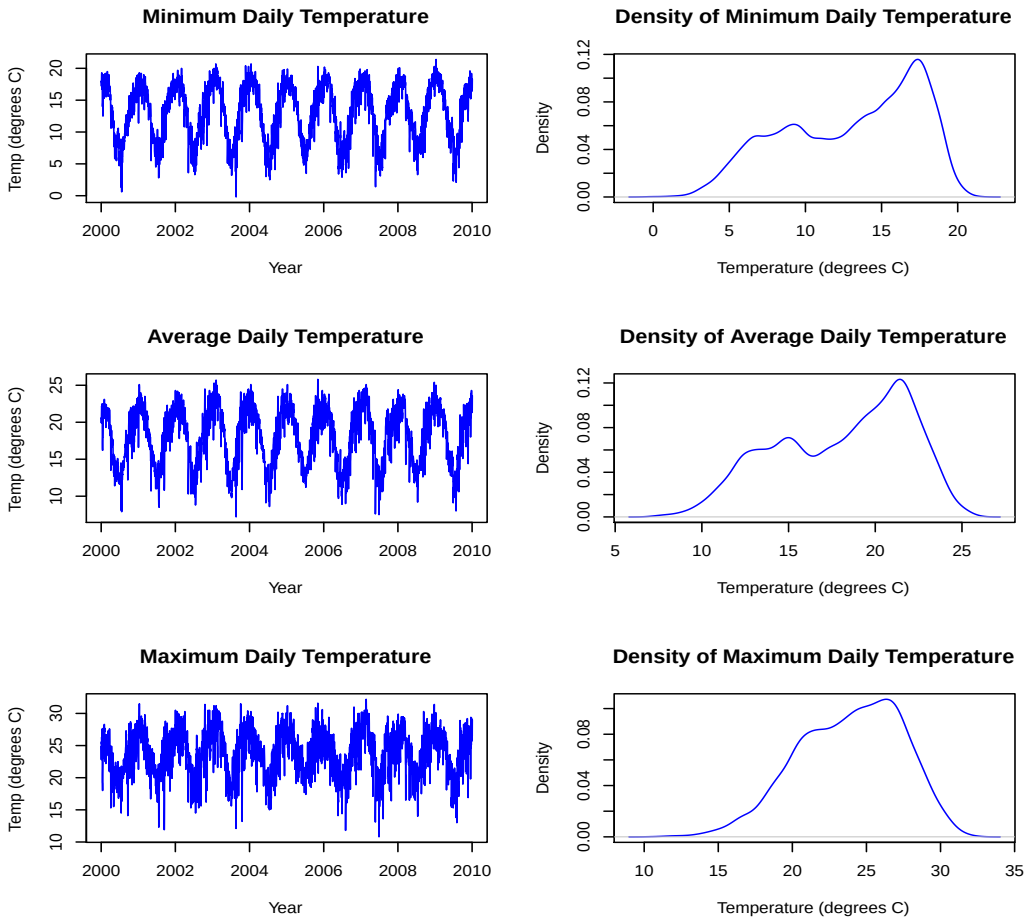

Figure 3.5: (a) Top left panel: Plot of the minimum (b) Middle left panel: Average and (c) Bottom left panel: maximum daily temperatures together with their probability density functions on the right panels of each for the years 2000 to 2010.

The frequencies of the maximum and minimum temperature occurring in different hours of the day are presented in Figure 3.6. Figures 3.14, 3.15 and 3.16 in appendix 3.2 at the end of the chapter show plots of electricity demand against temperature at different hours when minimum and maximum temperatures were recorded. In order to analyse the influence of temperature on

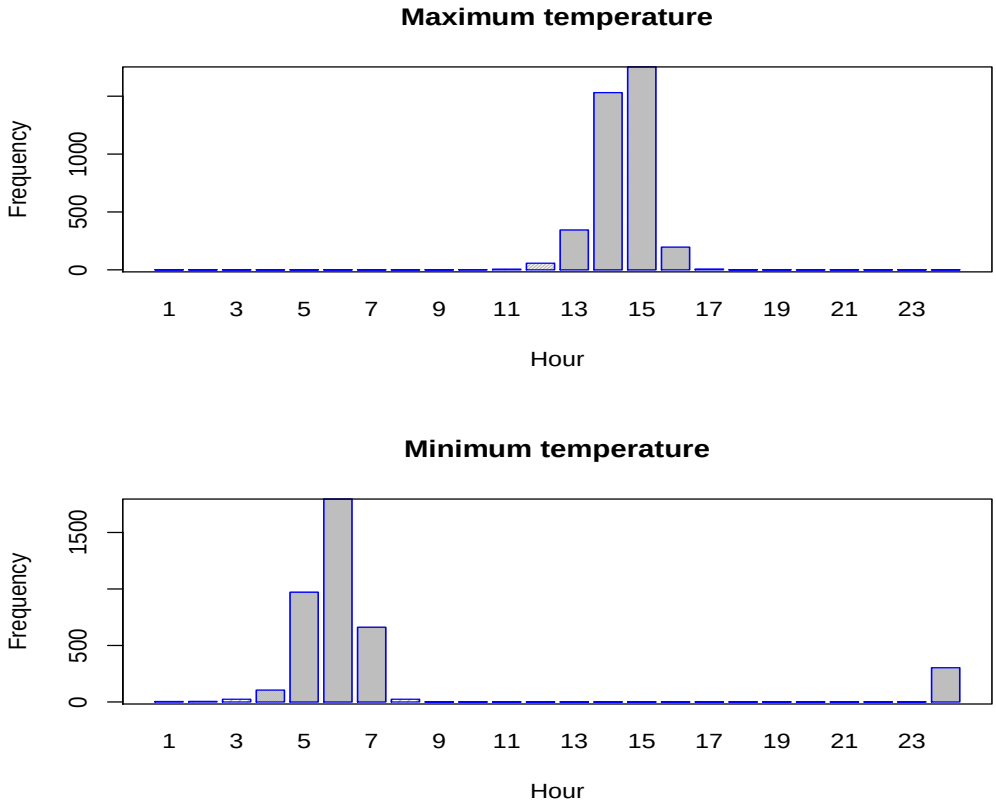

Figure 3.6: (a) Top panel: Frequency of occurrence of maximum and (b) Bottom panel: minimum daily temperatures in different hours of the day.

electricity demand we initially remove the trend and seasonality in the DPED data which is not linked to temperature. We initially fit a trend line to DPED, that is, we regress DPED (x_t) against time t, for $t = 1, .., n$. The trend line is given by the following equation

$$x_t = \beta_0 + \beta_1 t + \varepsilon_t$$

We then detrend by subtracting the trend component. This is then followed by deseasonalizing the detrended data. This is done by initially calculating the seasonal indices followed by dividing the DPED data by the seasonal index for the given season. There are other several methods of filtering data (i.e. removing both the trend and the calendar effects) which are discussed in literature (see Moral-Carcedo and Vicéns-Otero, 2005; Munoz *et al,*, 2010; among others). Figure 3.7 shows that the relationship between filtered DPED and temperature is nonlinear. Plots of electricity demand against average minimum temperature and those of electricity demand against average maximum temperature at different hours of the day are shown in Figures 3.10 and 3.11 respectively. These are given in appendix 3.2.

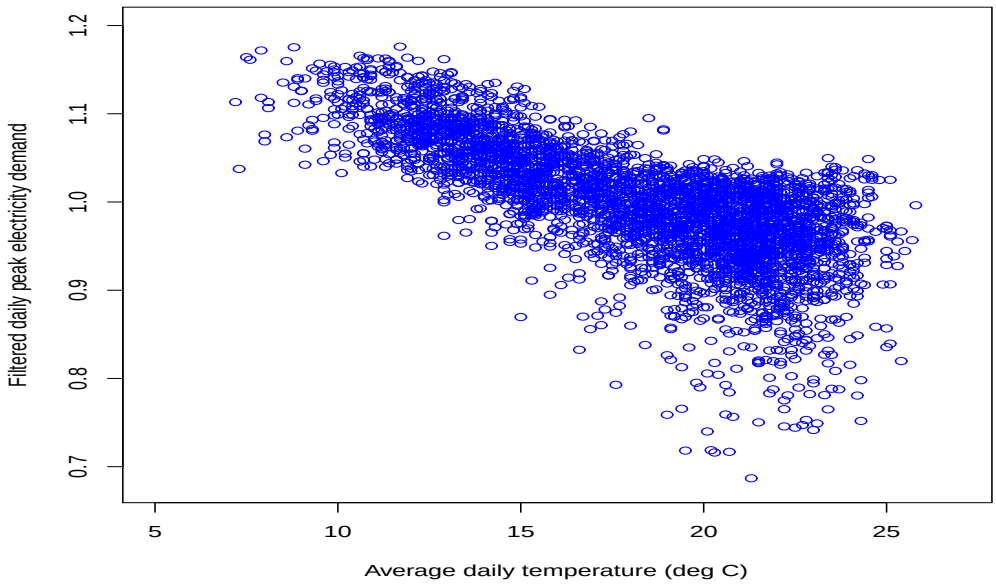

Figure 3.7: Scatter plot of filtered DPED (in MW) against average daily temperature (degree C).

The demand for electricity is highly sensitive to temperature fluctuations in winter and less sensitive in summer. The nonlinear relationship between load and temperature is modelled in literature using cooling degree-days and heating degree-days (Mirasgedis *et al.*, 2006; Munoz *et al.*, 2010; among others).

In the presence of only one reference point, Cooling Degree-Days (CDD_t) and Heating Degree-Days (HDD_t) are estimated on the basis of the following two linear functions defined as

$$\text{CDD}_t = \max\{T_t - T_{ref}, 0\}$$

and

$$\text{HDD}_t = \max\{T_{ref} - T_t, 0\}$$

where T_{ref} represents the temperature which separates the winter and summer periods of average daily electricity demand and temperature relationship and T_t represents average daily temperature (mean outdoor air temperature) on day t. If $T_t > T_{\text{ref}}$ there are no HDD_t and similarly when $T_t < T_{\text{ref}}$ there are no CDD_t. The reference temperature should be selected adequately so as to separate correctly the winter (cold temperature) and summer (hot temperature) periods of average daily electricity demand and temperature relationship (Munoz *et al.*, 2010). The challenge for South Africa is on how to determine the reference temperature T_{ref}.

3.4 Empirical results

3.4.1 The multivariate adaptive regression splines models

The MARS models discussed in Section 3.2 are used to identify thresholds for separating the winter-sensitive period from the weather-neutral period and also the summer-sensitive period from the weather-neutral period. The MARS model with DPED as the dependent variable and average daily temperature

(T_t) as the predictor variable is given as

$$DPED = r_0 + r_1 \textbf{max}\{0, T_t - T_w\} + r_2 \textbf{max}\{0, T_s - T_t\} + \varepsilon_{t(1)} \tag{3.7}$$

The parameters, r_0, r_1 and r_2, and knots, T_w and T_s respectively are estimated by using the Salford Predictive Modeller version 7.0. The resulting model is

$$\hat{\textbf{DPED}} = 26870.80 + 440.16 \textbf{max}\{0, 22 - T_t\} + 86.46 \textbf{max}\{0, T_t - 17.1\} \tag{3.8}$$

If the average daily temperature is less than or equal to 17.1^0C we use equation (3.9).

$$\hat{\textbf{DPED}} = 26870.80 + 440.16 \textbf{max}\{0, 22 - T_t\} \tag{3.9}$$

That is if temperature decreases by 1^0C (*e.g.* from 17.1^0C to 16.1^0C) DPED increases by about 440.16MW. Similarly if average daily temperature is greater than or equal to 22^0C we use equation (3.10)

$$\hat{\textbf{DPED}} = 26870.80 + 86.46 \textbf{max}\{0, T_t - 17.1\} \tag{3.10}$$

If temperature increases by 1^0C (*e.g.* from 22^0C to 23^0C) DPED increases by about 86.46MW. For an average daily temperature between 17.1^0C and 22^0C) we use the full model given in equation (3.8). A scatter plot of DPED against T_t is shown in Figure 3.8.

Using Average Daily Electricity Demand (ADED) as the response variable and T_t as the predictor variable we get the following MARS model

$$ADED = k_0 + k_1 \textbf{max}\{0, T_t - T_w\} + k_2 \textbf{max}\{0, T_s - T_t\} + \varepsilon_{t(2)} \tag{3.11}$$

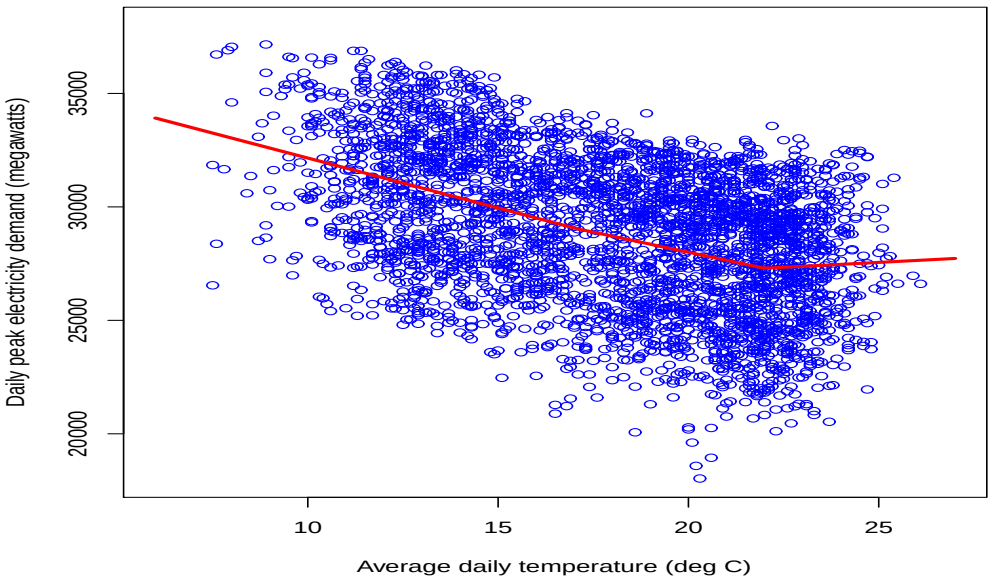

Figure 3.8: Scatter plot of DPED (in MW) against average daily temperature (in degree C).

After estimating the parameters and knots we get

$$\hat{\text{ADED}} = 23555.1 + 303.66\text{max}\left\{0, 22 - T_t\right\} + 151,52\text{max}\left\{0, T_t - 16\right\} \qquad (3.12)$$

If temperature increases by 1^0C above 22^0C (e.g. from 22^0C to 23^0C) ADED increases by about 151.52MW. Similarly if temperature decreases by 1^0C below 16^0C (*e.g.* from 16^0C to 15^0C) ADED increases by about 303.66MW. Figure 3.9 shows the relationship between average daily electricity demand against average daily temperature. Comparing Figure 3.9 with Figure 3.8 it is seen that DPED increases more sharply as temperature decreases when compared to ADED. The models given in equations 3.7 and 3.11 only use temperature as an explanatory variable. Lagged error terms and trend are ignored in these models. These models cannot be used for forecasting. They are only used for analyzing the relationship between electricity demand with temperature. A summary of the MARS models discussed in this section are given in appendix

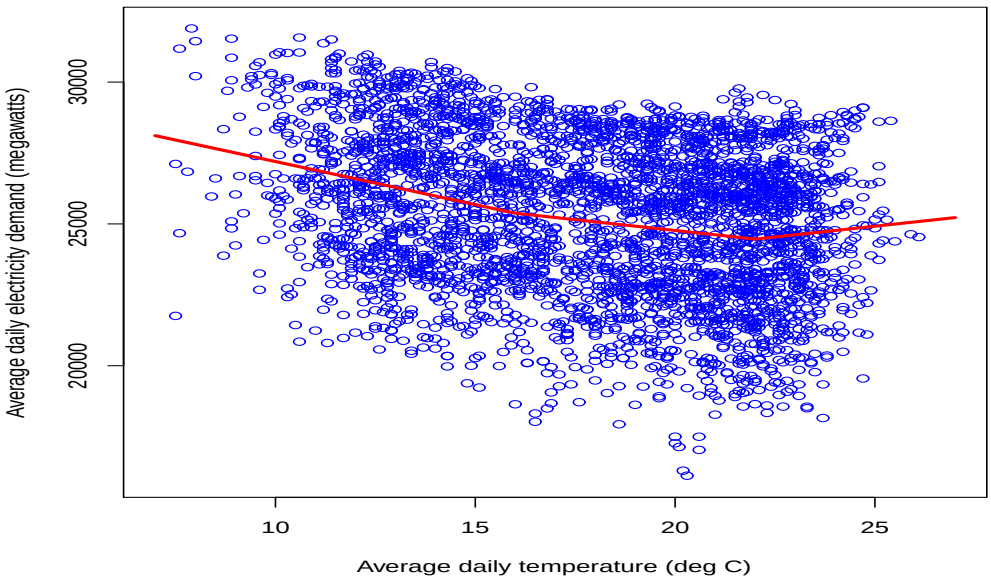

Figure 3.9: ADED (in MW) against average daily temperature (in degree C).

3.1 at the end of the chapter.

The threshold temperatures determined by the MARS algorithm are found to be 16^0C and 22^0C using ADED with average daily temperature, and 17.1^0C and 22^0C using DPED with average daily temperature. We round up 17.1^0C to 18^0C. From now onwards we will be using 18^0C for separating the winter-sensitive period from the weather-neutral period and 22^0C for separating the summer-sensitive period from the weather-neutral period. It should be noted that the reference temperatures differ from country to country, e.g. for Spain they are 15^0C and 17^0C (Munoz *et al.*, 2010). For both winter and summer, Sailor and Munoz (1997) use a reference temperature of 18.3^0C; Moral-Carcedoa and Vicéns-Otero (2005) use a threshold (reference) temperature of 18^0C and for the Australian electricity data, Hyndman and Fan (2008) used a reference temperature of 18.5^0C.

3.5 Concluding remarks

Modelling the influence of temperature on daily peak electricity demand using Multivariate Adaptive Regression Splines (MARS) is discussed. Using the MARS algorithm the reference temperatures which separate the weather neutral period from the winter and summer sensitive periods are determined. The developed MARS models can only be used for analyzing the relationship between electricity demand with temperature and should not be used for forecasting. The chapter establishes temperature as an important variable in explaining electricity demand. Empirical evidence from this study shows that for temperature values below 18^0C demand for electricity in South Africa increases significantly while for temperature values above 22^0C demand increases slightly. Within the weather neutral period residents would neither use a heater nor a cooling system.

In Chapter 4 we discuss the use of piecewise linear regression models in modelling the influence of temperature on electricity demand.

APPENDIX 3.1: MARS MODELS

1. **Daily Peak Electricity Demand (DPED) with Average Daily Temperature (ADT)**

   ==========

   Basis Functions

   ==========

   BF2 = max(0, 22 - ADT);

   BF3 = max(0, ADT - 17.1);

   DPED = 26870.8 + 440.156 * BF2 + 86.4596 * BF3;

   MODEL DPED = BF2 BF3.

2. **Average Daily Electricity Demand with Average Daily Temperature (ADT)**

   ==========

   Basis Functions

   ==========

   BF2 = max(0, 22 - ADT);

   BF3 = max(0, ADT - 16);

   ADED = 23555.1 + 303.663 * BF2 + 151.519 * BF3;

   MODEL ADED = BF2 BF3.

APPENDIX 3.2: PLOTS OF DEMAND AGAINST TEMPERATURE

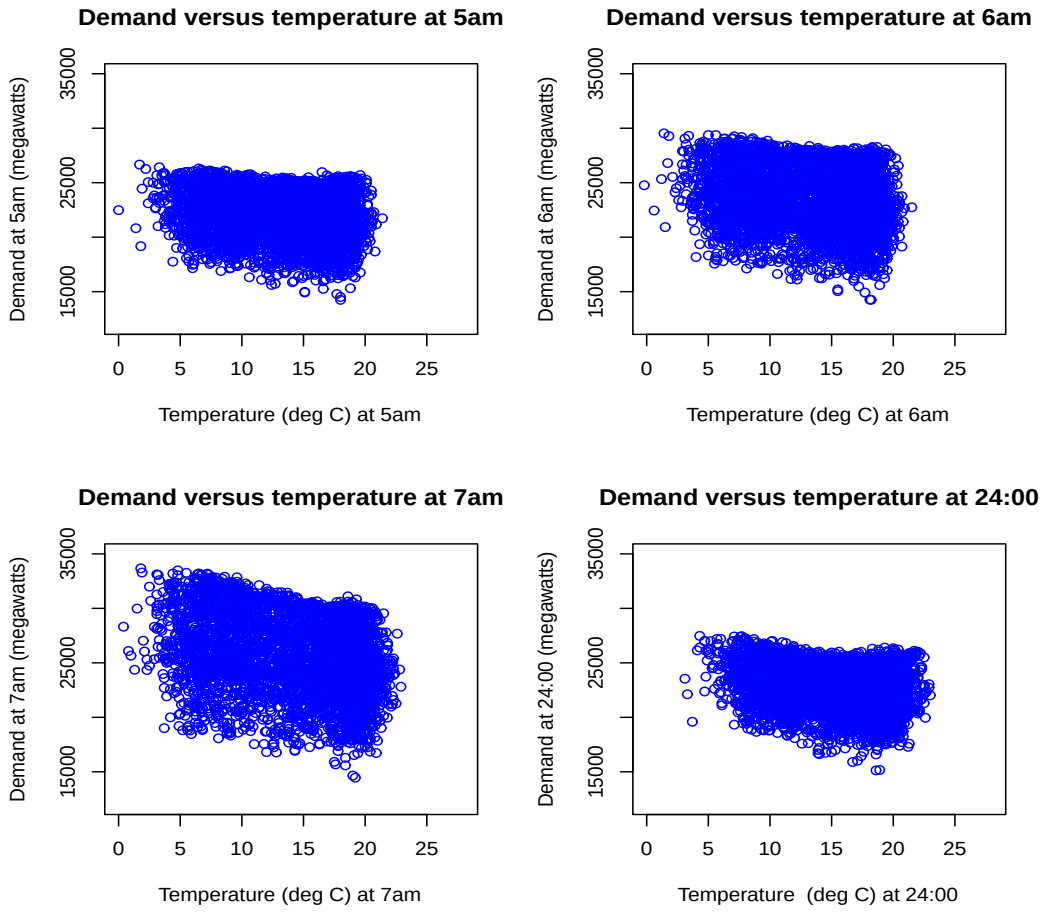

Figure 3.10: Plot of electricity demand against average minimum temperature at different hours of the day.

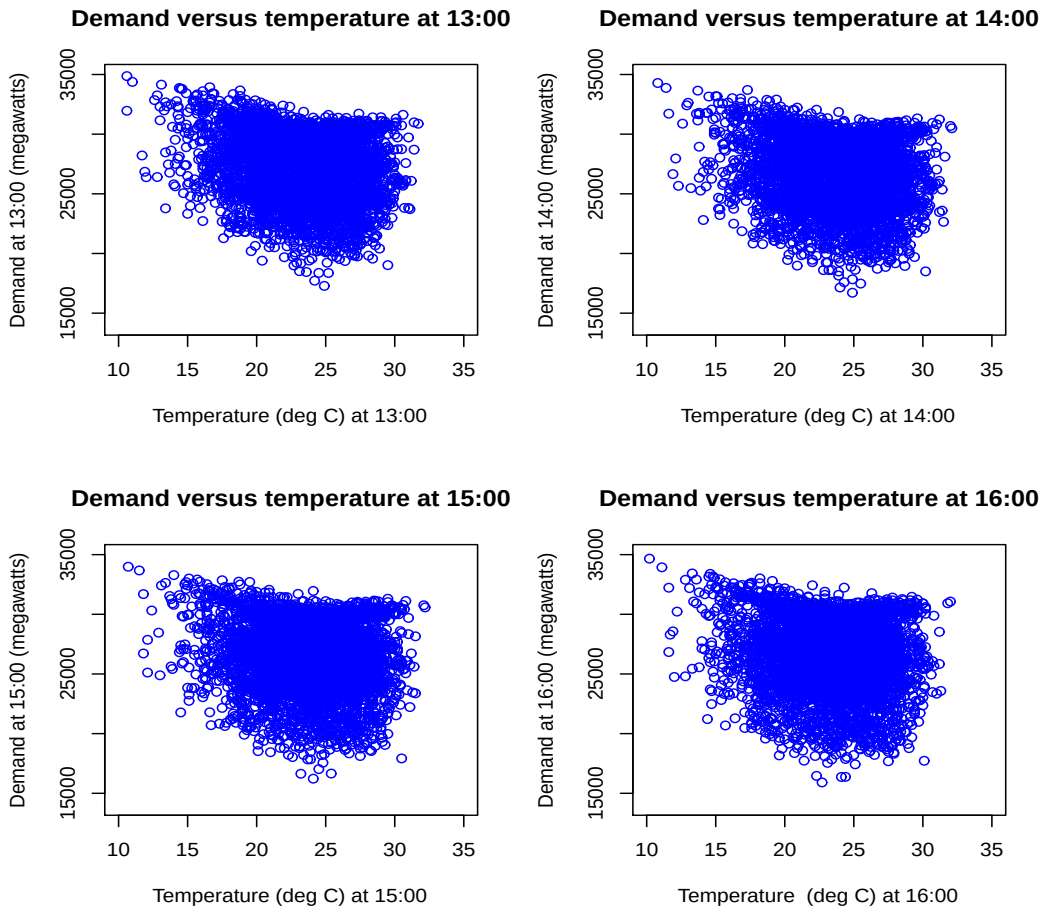

Figure 3.11: Plot of electricity demand against average maximum temperature at different hours of the day.

APPENDIX 3.3: PROBABILITY DENSITY FUNCTIONS OF DPED: YEARS 2000 TO 2009 AND OF HOURLY DEMAND: 01:00 TO 24:00

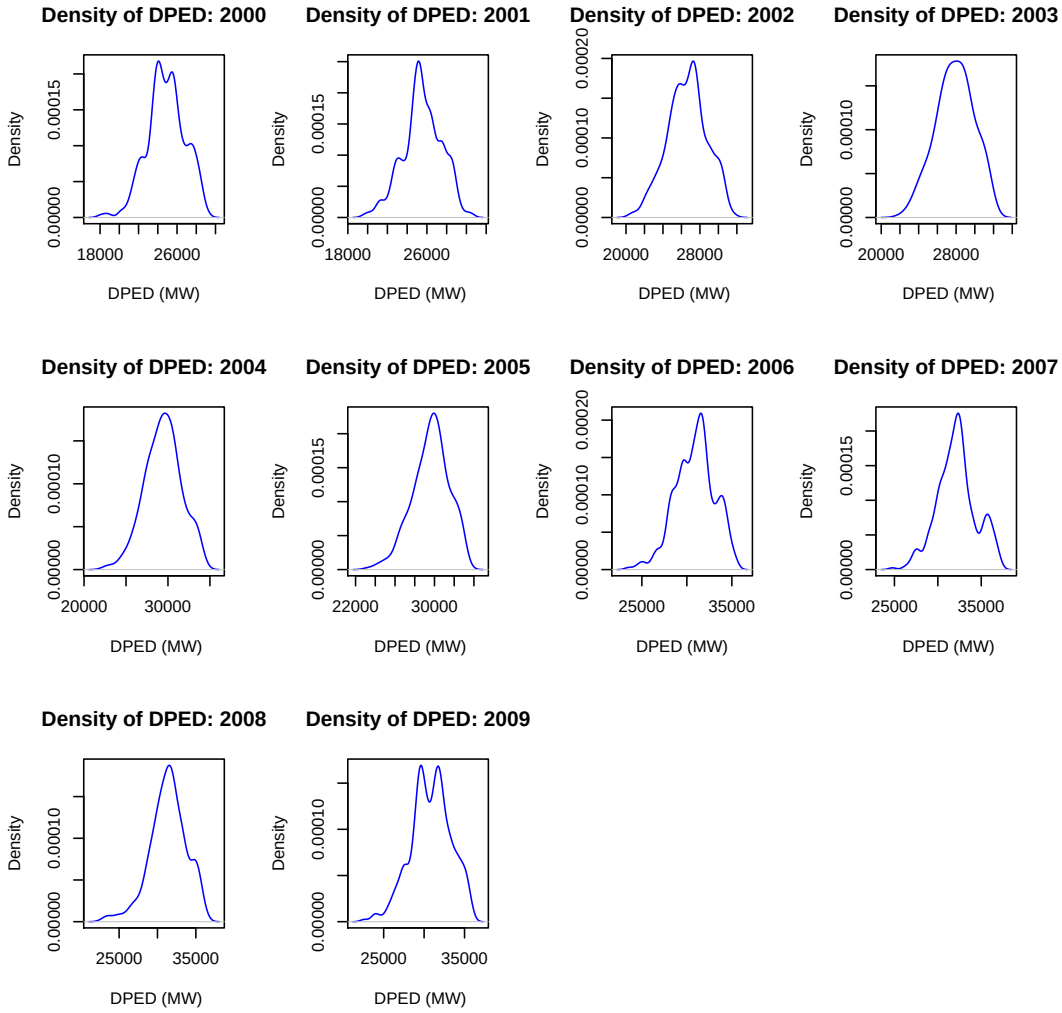

Figure 3.12: Probability density functions of DPED: 2000 to 2009.

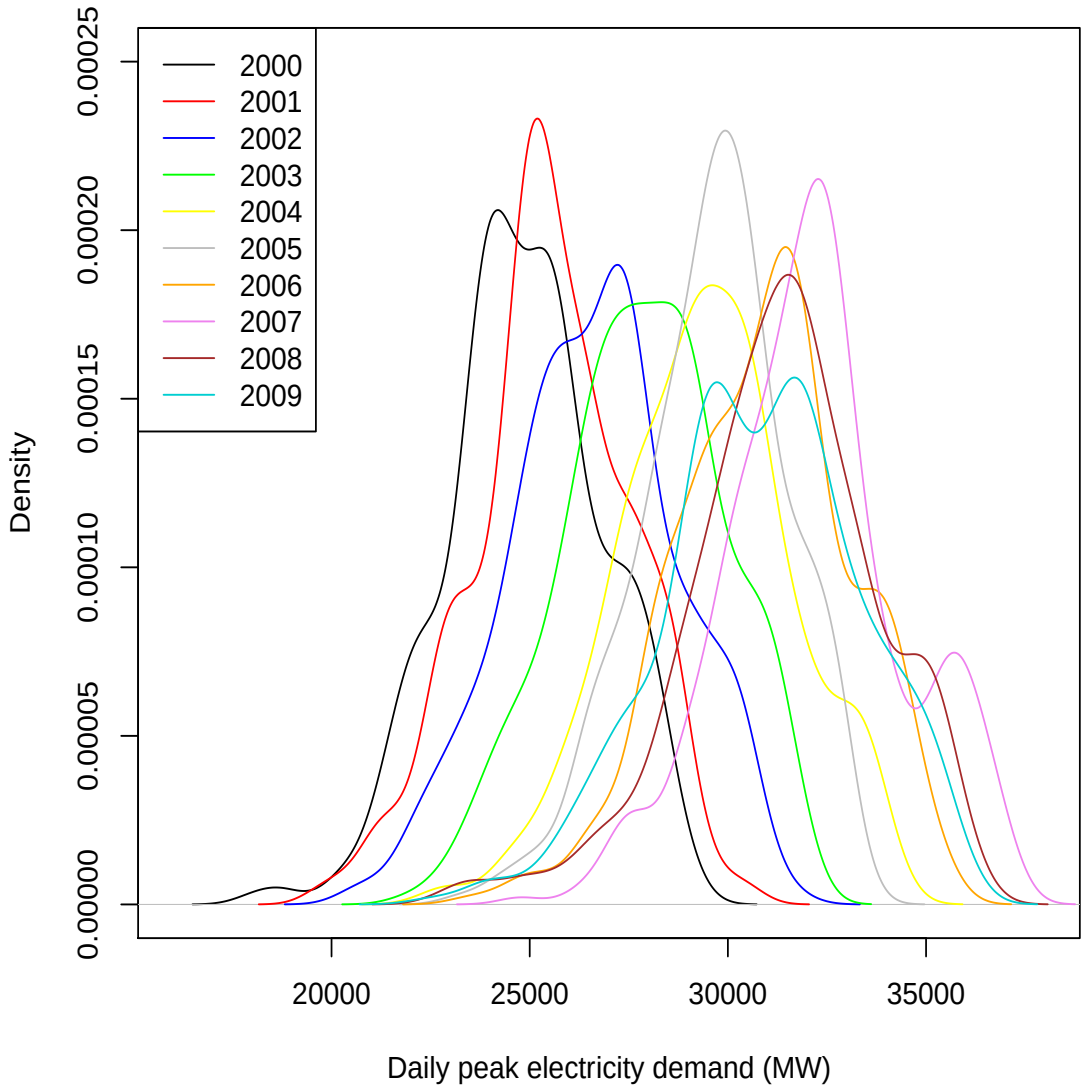

Figure 3.13: Probability density functions of DPED for the years 2000 to 2009.

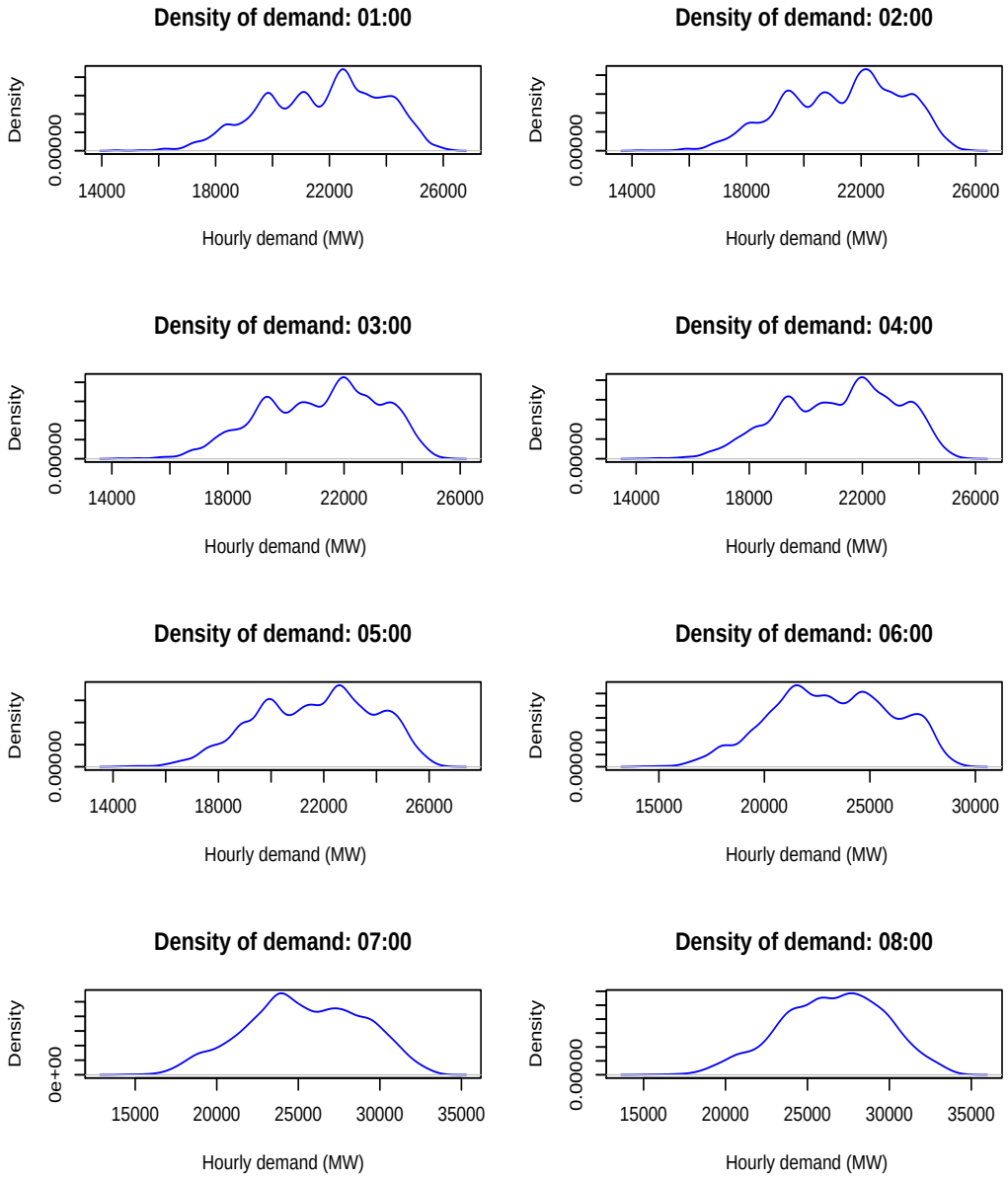

Figure 3.14: Probability density functions of hourly demand: 01:00 to 08:00.

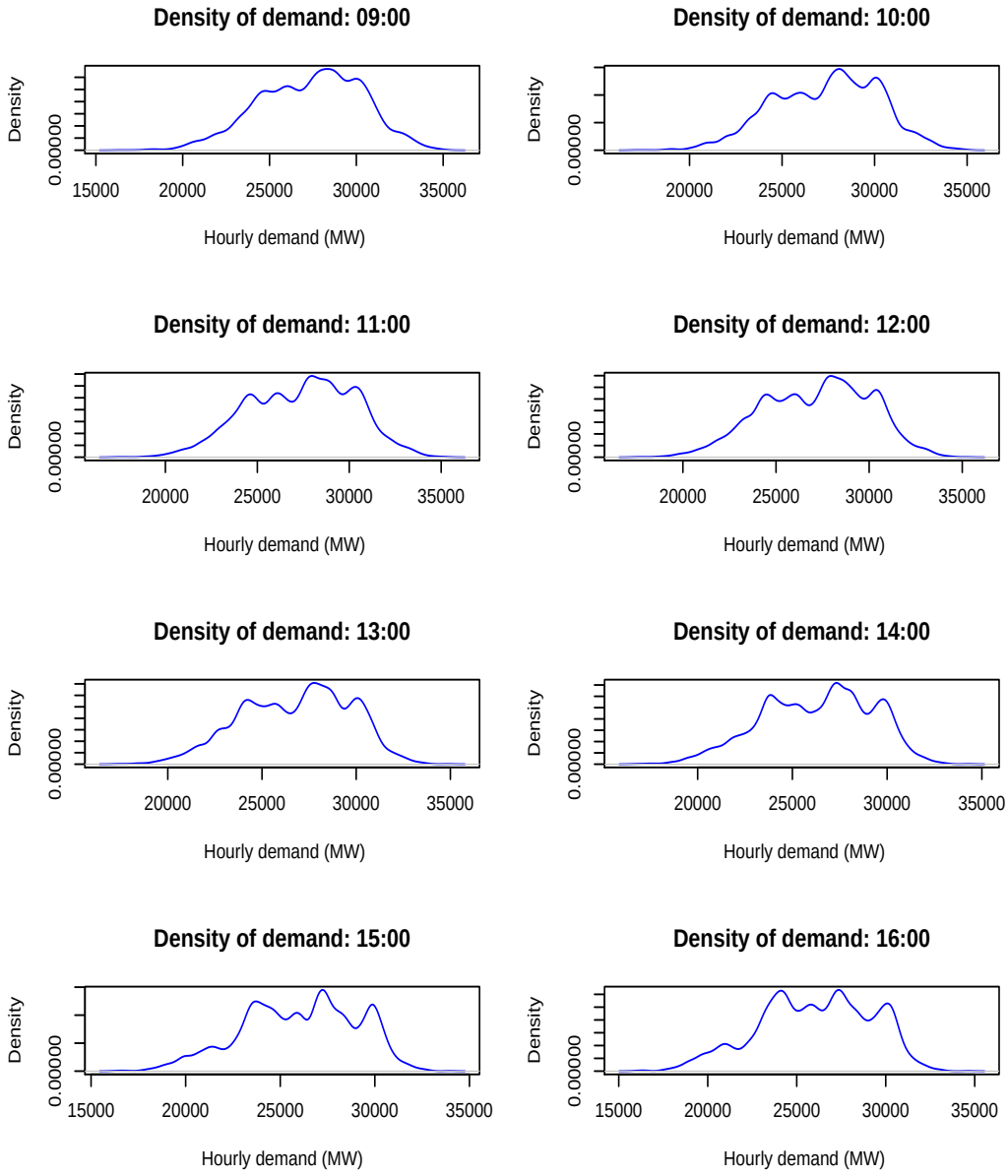

Figure 3.15: Probability density functions of hourly demand: 09:00 to 16:00.

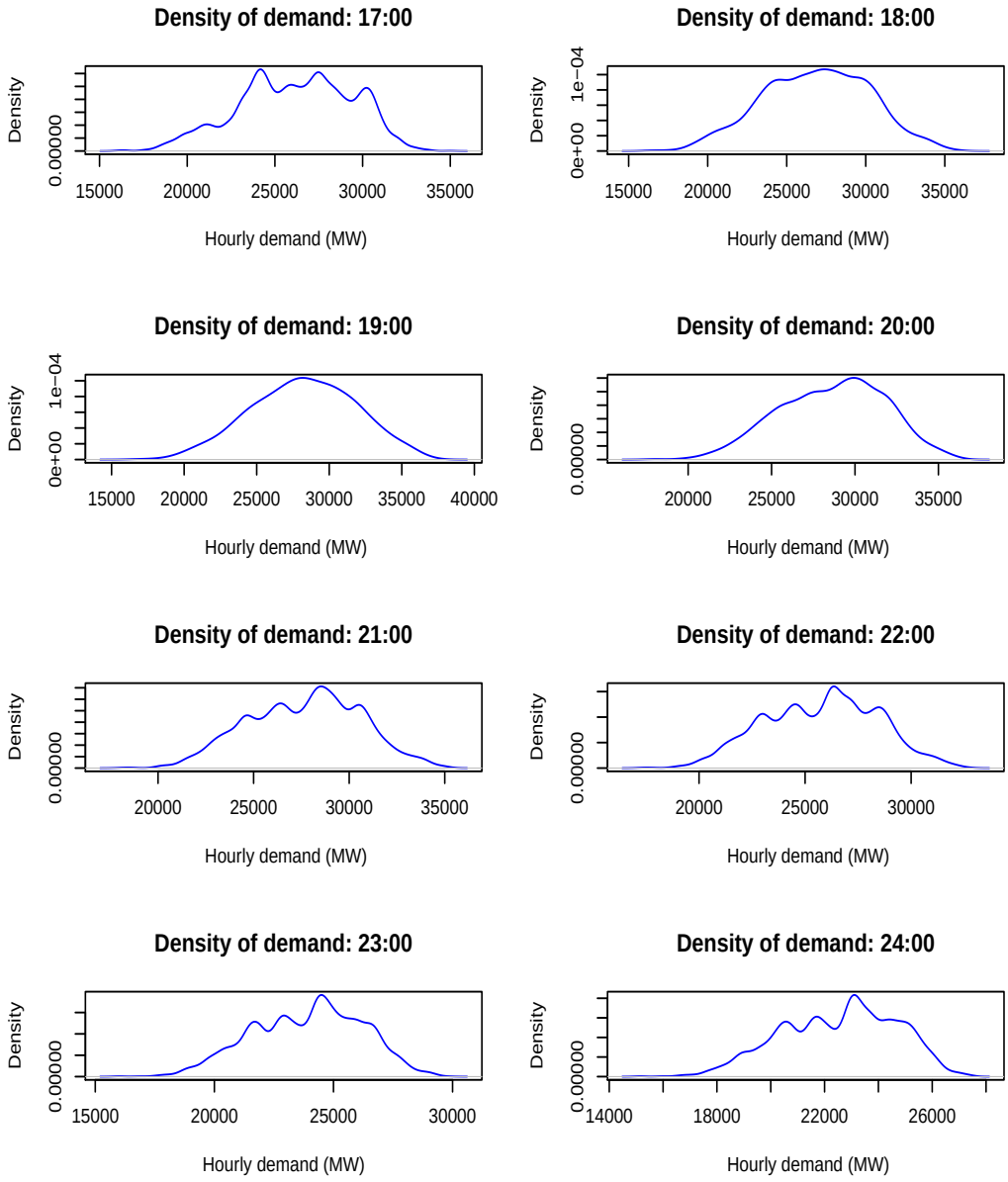

Figure 3.16: Probability density functions of hourly demand: 17:00 to 24:00.

Chapter 4

Modelling the influence of temperature on daily peak electricity demand in South Africa using a piecewise linear regression model

4.1 Introduction

The chapter presents empirical results on the influence of temperature on daily peak electricity demand in South Africa using Piecewise Linear (PL) regression models. In Chapter 3 we determined the threshold temperatures using a MARS model. We extend the models by including lagged error terms and a trend component. The piecewise linear regression model discussed in this

chapter is only used to explore the influence of temperature on electricity demand and not for predictions. In order to assess the influence of temperature on DPED we use peak temperature (T_t^{pt}) instead of the average daily temperature (T_t). Peak temperature is temperature recorded at the hour of peak demand. The values of T_w and T_s determined in Chapter 3 using the MARS model are used in the piecewise linear regression model in Sections 4.2 and 4.3.

4.2 The piecewise linear regression model

The piecewise linear regression model discussed in this chapter includes lagged error terms with a trend component. The piecewise linear regression model is given in equation (4.1).

$$x_t = c_o + c_1 t + c_2(T_t^{pt} - T_w)d_{1t} + c_3(T_t^{pt} - T_s)d_{2t} + R_t \tag{4.1}$$

where T_t^{pt} represents peak temperature (in degrees Celsius) at time t. DPED (in megawatts) observed on day t is denoted by x_t, T_w is the temperature to identify where the winter sensitive portion of demand join the non-weather sensitive demand component, T_s is the temperature to identify where the summer sensitive portion of demand join the non-weather sensitive demand component, c_0 represents the mean daily peak demand observed in the non weather sensitive period ($T_w \leq T_t^{pt} \leq T_s$). It should be noted that daily peak demand during non-weather sensitive days does not depend on temperature. The variable t represents the trend component. R_t is a stochastic disturbance term whose structure is given by $R_t = \sum_{i=1}^{7} \varphi_i R_{t-i} + \varepsilon_t$ for modelling the DPED data and ε_t is a white noise series. The parameters to be estimated are $c_i, i = 0, 1, 2, 3$ and $\varphi_i, i =$

$1, 2, 5, 7$. The dummy variables d_{1t} and d_{2t} are defined as follows

$$d_{1t} = \begin{cases} 1 & \text{if } T_t^{pt} - T_w < 0 \\ 0 & \text{otherwise} \end{cases} \tag{4.2}$$

$$d_{2t} = \begin{cases} 1 & \text{if } T_t^{pt} - T_s > 0 \\ 0 & \text{otherwise} \end{cases} \tag{4.3}$$

The model in equation (4.1) accounts for any residual correlation that may occur as a result of the day-to-day and week-to-week variations in peak electricity demand. Equation (4.1) is based on the following theoretical assumptions:

1. Peak demand on day t is highly correlated with peak demand on day $t+1$.
2. There may be significant correlation between demand with lags of 2 days, 5 days and/or 7 days.

Derivation of the three demand equations for winter-sensitive, summer-sensitive and non-weather sensitive periods

From

$$x_t = c_0 + c_1 t + c_2(T_t^{pt} - T_w)d_{1t} + c_3(T_t^{pt} - T_s)d_{2t} + R_t \tag{4.4}$$

The derivation of the three equations for winter-sensitive, summer-sensitive and non-weather sensitive periods are given as follows:

For winter-sensitive months, i.e. $T_t^{pt} < T_w, d_{1t} = 1, d_{2t} = 0$, we obtain

$$\begin{aligned} E(x_t) &= c_0 + c_1 t + c_2(T_t^{pt} - T_w)(1) + c_3(T_t^{pt} - T_s)(0) \\ &= c_0 + c_1 t + c_2(0, T_t^{pt} - T_w) \\ &= (c_0 - c_2 T_w) + c_1 t + c_2 T_t^{pt} \end{aligned} \tag{4.5}$$

where $(c_0 - c_2 T_w)$ is a constant term. Let $a = (c_0 - c_2 T_w)$ then

$$E(x_t) = a + c_1 t + c_2 T_t^{pt} \tag{4.6}$$

which is a simple linear regression model.

For the summer-sensitive months, i.e. $T_t^{pt} > T_s, d_{1t} = 0, d_{2t} = 1$, we obtain

$$\begin{aligned} E(x_t) &= c_0 + c_1 t + c_2(T_t^{pt} - T_w)(0) + c_3(T_t^{pt} - T_s)(1) \\ &= c_0 + c_1 t + c_3(T_t^{pt} - T_s) \\ &= (c_0 - c_3 T_s) + c_1 t + c_3 T_t^{pt} \end{aligned} \tag{4.7}$$

Similarly $(c_0 - c_3 T_s)$ is also a constant. Let $b = (c_0 - c_3 T_s)$ then

$$E(x_t) = b + c_1 t + c_3 T_t^{pt} \tag{4.8}$$

Finally, for the non-weather sensitive months, $T_w \leq T_t^{pt} \leq T_s, d_{1t} = d_{2t} = 0$, we obtain

$$E(x_t) \quad = \quad c_0 + c_1 t \tag{4.9}$$

where c_0 represents the mean daily demand observed during the non-weather sensitive period.

One of the challenges for South African data is to determine T_w and T_s. If T_w and T_s are known then the parameters of the equation 4.4 can be estimated by the Ordinary Least Squares (OLS) method. The Multivariate Adaptive Regression Splines (MARS) modelling approach was used in Chapter 3 to determine T_w and T_s.

4.3 Empirical results

Using the values of T_w and T_s which were determined in Chapter 3 (18^0C and 22^0C respectively) the piecewise linear regression model is given as

$$x_t = c_0 + c_1 t + c_2 (T_t^{pt} - 18) d_{1t} + c_3 (T_t^{pt} - 22) d_{2t} + R_t \tag{4.10}$$

where c_0, c_1, c_2 and c_3 are parameters, x_t is DPED, R_t is an error term with the following structure, $R_t = \phi_1 R_{t-1} + \phi_2 R_{t-2} + \Phi_1 R_{t-7} + \varepsilon_t + \theta_3 \varepsilon_{t-3} + \Theta_1 \varepsilon_{t-7}$. Let $X_1 = (T_t^{pt} - 18) d_{1t}$ where $d_{1t} = 1$ if $T_t^{pt} < 18$ and zero otherwise. This is equivalent to $X_1 = \min(0, T_t^{pt} - 18)$. Similarly let $X_2 = (T_t^{pt} - 22) d_{2t}$ where $d_{2t} = 1$ if $T_t^{pt} > 22$ and zero otherwise. This is equivalent to $X_2 = \max(0, T_t^{pt} - 22)$. Equation 4.10 can be written as

$$x_t = c_0 + c_1 t + c_2 X_1 + c_3 X_2 + R_t \tag{4.11}$$

After considering several piecewise linear regression models, the best model is given in equation (4.12).

$$\hat{x}_t = 38904 - 1.186t - 185.4X_1 + 11.4X_2 + 0.792\hat{R}_{t-1} - 0.024\hat{R}_{t-2}$$

$$+0.115\hat{R}_{t-3} + 0.043\hat{R}_{t-5} + 0.997\hat{R}_{t-7} - 0.853\varepsilon_{t-7} \qquad (4.12)$$

The parameter c_2 in equation (4.12) is negative, showing that if peak temperature decreases by one degree from 18^0C, electricity demand increases by about 185.4 MW. The parameter c_3 is positive, showing that if the temperature increases by one degree from 22^0C, electricity demand increases by about 11.4MW. This shows that electricity demand is more sensitive to cold temperatures than hot temperatures.

Summary statistics for the piecewise linear regression model given in equation (4.12) are presented in Table 4.1.

Table 4.1: In-sample evaluation.

	Adjusted R-squared	AIC [a]	DW [b]	RMSE [c]	MAPE [d]
PLR model	0.9721	15.4202	1.9977	538.28	1.29

[a] AIC denotes Akaike Information Criterion
[b] DW denotes Durbin Watson statistic
[c] RMSE represents the Root Mean Square Error
[d] MAPE is the Mean Absolute Percentage Error

4.4 Concluding remarks

In this chapter the modelling of the influence of temperature on daily peak electricity demand in South Africa using a piecewise linear regression model is discussed. The developed piecewise linear regression model is not meant for forecasting but to model the effect of temperature on electricity demand. The

study establishes temperature as an important variable in explaining electricity demand. Empirical evidence from this study shows that for temperature values below 18^0C demand for electricity in South Africa increases significantly while for temperature values above 22^0C demand increases slightly. These results are consistent with findings discussed in Chapter 3.

An interesting area for further study would involve the use of Bayesian parameter estimates of the piecewise linear regression model in explaining the effect of temperature on electricity demand. This is discussed in Chapter 5.

Chapter 5

Influence of temperature on daily peak electricity demand in South Africa: A frequentist and Bayesian approach

. . . in terms of forecasting ability, . . . a good Bayesian will beat a non-Bayesian, who will do better than a bad Bayesian.

C.W.J. Granger (1986, p. 16)

5.1 Introduction

The use of the Bayesian parameter estimates to a piecewise linear regression model in explaining the influence of temperature on daily peak electricity demand in South Africa is discussed in this chapter. Probability statements are

made about the uncertainty in the estimated parameters. The chapter concentrates on daily peak electricity demand modelling which is important for providing short term forecasts which will assist in optimal dispatching of electrical energy. Estimation of the parameters of the piecewise linear regression model is done using the least squares method. The next step involves the use of Bayesian statistics in developing posterior densities for the parameter estimates to the piecewise linear regression model. A Bayesian analysis provides a way of taking into account uncertainty in the estimation of the parameters. Uncertainty about the true values of the Bayesian parameter estimates is incorporated into the analysis through the use of a non-informative prior distribution. The Bayesian modelling approach is arguably more informative than the classical statistical modelling approach (Bayarri and Berger, 2004).

5.2 Bayesian modelling framework

5.2.1 Bayesian forecasting

In electricity demand forecasting, models such as multiple linear regression, for example, have important parameters which help in quantifying the marginal contribution of drivers of electricity demand. For example if x denotes electricity demand and z represents gross domestic product, then $x = \beta_0 + \beta_1 z + \varepsilon$ represents a simple model for predicting x given z. The parameters of the model are β_0 and β_1. Let $\boldsymbol{\theta}$ represent a vector of these parameters, i.e. $\boldsymbol{\theta} = \{\beta_0, \beta_1\}$.

Let $\pi(\boldsymbol{\theta})$ be the density of the prior distribution for the parameter $\boldsymbol{\theta} = (\theta_1, ..., \theta_p)$ and $\pi(\boldsymbol{x}|\boldsymbol{\theta})$, where $\boldsymbol{x} = (x_1, ..., x_n)$ be the likelihood function then the posterior distribution, Beirlant *et al.*, (2004), is given by

$$\pi(\boldsymbol{\theta}|\boldsymbol{x}) = \frac{\pi(\boldsymbol{x}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\int_{\boldsymbol{\theta}} \pi(\boldsymbol{x}|\boldsymbol{\theta})\pi(\boldsymbol{\theta}) d\boldsymbol{\theta}} \tag{5.1}$$

$$\pi(\boldsymbol{\theta}|\boldsymbol{x}) \quad \propto \quad \pi(\boldsymbol{x}|\boldsymbol{\theta})\pi(\boldsymbol{\theta}) \tag{5.2}$$

If $\pi(x_{n+1}|\boldsymbol{\theta})$ is the density function of a future observation x_{n+1} the posterior predictive density $\pi(x_{n+1}|\boldsymbol{x})$ is given by

$$\pi(x_{n+1}|\boldsymbol{x}) \quad = \quad \int_{\boldsymbol{\theta}} \pi(x_{n+1}|\boldsymbol{\theta})\pi(\boldsymbol{\theta}|\boldsymbol{x})d\boldsymbol{\theta} \tag{5.3}$$

and the posterior predictive probability of a future observation x_{n+1} greater than a given threshold τ is given by

$$P(X_{n+1} > \boldsymbol{x}|X > \tau) \quad = \quad \int_{\boldsymbol{\theta}} P(X_{n+1} > \boldsymbol{x}|\boldsymbol{\theta})\pi(\boldsymbol{\theta}|\boldsymbol{x})d\boldsymbol{\theta} \tag{5.4}$$

Markov Chain Monte Carlo (MCMC) methods are usually used to find approximate values for equation (5.1). It is given in Beirlant *et al.*, (2004) that if a sample $\theta_1, ..., \theta_k$ is taken from the posterior distribution $\pi(\boldsymbol{\theta}|\boldsymbol{x})$ then the following approximation can be used

$$P(X_{n+1} > \boldsymbol{x}|X > \tau) \quad \propto \quad \frac{1}{k}\sum_{i=1}^{k} P(X_{n+1} > \boldsymbol{x}|\theta_i) \tag{5.5}$$

Gibbs Sampler (the algorithm)

The Gibbs sampler introduced by Geman and Geman (1984) is a special case of the Metropolis-Hastings algorithm. The Gibbs algorithm is given as: Let $\boldsymbol{\theta} = (\theta_1, ..., \theta_p)$. Each $\theta_j^{(i)}$ is sampled from the conditional distribution given all the other components of $\boldsymbol{\theta}$. So if we have the set of values $\left\{\theta_1^{(i)}, ..., \theta_p^{(i)}\right\}$ the

algorithm is as follows:

$$
\begin{aligned}
\text{Draw } \theta_1^{(i+1)} &\sim \pi(\theta_1|\theta_2^{(i)}, ..., \theta_p^{(i)}, x) \\
\text{Draw } \theta_2^{(i+1)} &\sim \pi(\theta_2|\theta_1^{(i+1)}, \theta_3^{(i)}, ..., \theta_p^{(i)}, x) \\
&\vdots \\
\text{Draw } \theta_p^{(i+1)} &\sim \pi(\theta_p|\theta_1^{(i+1)}, ..., \theta_{p-1}^{(i+1)}, x)
\end{aligned}
$$

5.2.2 Prior distributions

Generally prior distributions are classified into noninformative and informative priors. Conjugate priors are a special type of informative priors. Conjugate priors provide computational convenience in that the prior and posterior distributions belong to the same family of distribution (Gelman *et al.*, 2004). Noninformative priors are objective. They are also known as vague, diffuse or flat priors. With a flat prior the likelihood on all possible values of a parameter is equal (Gelman *et al.*, 2004; SAS/STAT9.2 User's Guide Introduction to Bayesian Analysis Procedures (n.d.)). Box and Tiao (1973) discuss in detail development of noninformative priors. Techniques for deriving noninformative priors are discussed in Kass and Wasserman (1996).

Informative priors are used when prior information about the parameter is available (Koch, 2007). Incorporating prior information which is usually based on previous knowledge, previous studies or expert opinions is usually difficult in practice. To overcome this challenge, Ibrahim and Chen (2000) develop a power prior which takes into account previous data.

We discuss briefly two non informative priors used in this thesis, that is the

Jeffreys's prior and the maximal data information prior.

Jeffreys's prior

The Jeffreys's prior is objective and invariant under reparameterization, but violates the likelihood principle (Bernado and Smith, 1994; Beirlant *et al.*, 2004). Jeffreys's prior (Jeffreys, 1961) is defined as

$$\pi(\boldsymbol{\theta}) \quad \propto \quad \sqrt{|I(\boldsymbol{\theta})|} \tag{5.6}$$

where $\boldsymbol{\theta} = (\theta_1, ..., \theta_p)$ and $I(\boldsymbol{\theta})$ is Fisher's information matrix with $(i, j)^{th}$ element.

$$I_{i,j}(\boldsymbol{\theta}) \quad = \quad E\left\{-\frac{\partial^2 \log f(\boldsymbol{X}|\boldsymbol{\theta})}{\partial\theta_i \partial\theta_j}\right\}, i, j = 1, ..., p \tag{5.7}$$

For details see Beirlant *et al.* (2004).

Maximal Data Information Prior

Maximal data information prior is one which provides maximal average data information (Zellner, 1977). Although the prior is not invariant under reparameterization, constraints on the parameter can be built into the prior (Zellner, 1977). The MDI prior is defined as follows (Zellner, 1977; Beirlant *et al.*, 2004):

$$\pi(\boldsymbol{\theta}) \quad \propto \quad \exp\left\{E\left[\log f(\boldsymbol{X}|\boldsymbol{\theta})\right]\right\} \tag{5.8}$$

where $\boldsymbol{\theta}$ is as defined in Section 5.2.1.

5.2.3 Bayesian linear regression model

The normal multiple regression model is given by

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{u} \tag{5.9}$$

where $\boldsymbol{y}$ is an $n \times 1$ vector of observations, $\boldsymbol{X}$ is the $n \times k$ design matrix of observations on $k-1$ independent variables. The vector of regression coefficients is given by $\boldsymbol{\beta}$ and $\boldsymbol{u}$ is an $n \times 1$ vector of error terms. The error terms, $\boldsymbol{u}$ are assumed to be normally and independently distributed each with mean zero and common variance σ^2. The parameter vector is $(\boldsymbol{\beta}, \sigma^2)$. We then estimate $\boldsymbol{\beta}$ and σ^2 together with their respective probability density functions (pdfs). In the frequentist approach the coefficient vector $\boldsymbol{\beta}$ is estimated by the ordinary least squares method using the Moore-Penrose pseudoinverse, $\boldsymbol{\beta} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}$. In contrast to the frequentist approach the Bayesian approach supplements the sample data with additional information which is in the form of a prior distribution.

Joint prior distribution (Jeffreys's prior)

In specifying the prior distribution we assume that the information is vague. We use a non-informative Jeffreys's prior (Zellner, 1971): Let $\boldsymbol{\theta} = (\boldsymbol{\beta}, \sigma^2)$ then

$$\pi(\boldsymbol{\beta}, \sigma^2) \quad \propto \quad \sqrt{|I(\boldsymbol{\beta}, \sigma^2)|}$$

Now

$$
\begin{aligned}
I(\boldsymbol{\beta}, \sigma^2) &= E\left\{-\frac{\partial^2 \log f(\boldsymbol{X}|\boldsymbol{\beta}, \sigma^2)}{\partial \boldsymbol{\beta} \partial \sigma^2}\right\} \\
&= -E\begin{bmatrix} \frac{\partial^2 \log f(\boldsymbol{X}|\boldsymbol{\beta}, \sigma^2)}{\partial \boldsymbol{\beta}^2} & \frac{\partial^2 \log f(\boldsymbol{X}|\boldsymbol{\beta}, \sigma^2)}{\partial \boldsymbol{\beta} \partial \sigma^2} \\ \frac{\partial^2 \log f(\boldsymbol{X}|\boldsymbol{\beta}, \sigma^2)}{\partial \sigma^2 \partial \boldsymbol{\beta}} & \frac{\partial^2 \log f(\boldsymbol{X}|\boldsymbol{\beta}, \sigma^2)}{\partial (\sigma^2)^2} \end{bmatrix} \\
I(\boldsymbol{\beta}, \sigma^2) &= \frac{1}{\sigma^2}\boldsymbol{X}'\boldsymbol{X}
\end{aligned}
$$

Hence

$$
\begin{aligned}
\pi(\boldsymbol{\beta}, \sigma^2) &\propto \sqrt{\left|\frac{1}{\sigma^2}\boldsymbol{X}'\boldsymbol{X}\right|} \\
\pi(\boldsymbol{\beta}, \sigma^2) &\propto \frac{1}{\sigma^2} \qquad (5.10)
\end{aligned}
$$

For a detailed discussion of the Jeffreys's prior of the normal linear regression see Zellner (1971); Ibrahim and Laud (1991); Micali (2007); among others.

Joint likelihood function

Assuming that the random errors are independent and normally distributed, the joint likelihood for y given $\boldsymbol{X}, \boldsymbol{\beta}$ and σ^2 is given as (Zellner, 1971)

$$
\begin{aligned}
\pi(\boldsymbol{y}|\boldsymbol{X}, \boldsymbol{\beta}, \sigma^2) &\propto \frac{1}{\sigma^n}\exp\left[-\frac{1}{2\sigma^2}(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})'(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})\right] \\
&\propto \frac{1}{\sigma^n}\exp\left\{-\frac{1}{2\sigma^2}\left[(\boldsymbol{y} - \hat{\boldsymbol{y}})'(\boldsymbol{y} - \hat{\boldsymbol{y}}) + \left(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}\right)'\boldsymbol{X}'\boldsymbol{X}\left(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}\right)\right]\right\} \\
&\propto \frac{1}{\sigma^n}\exp\left\{-\frac{1}{2\sigma^2}\left[vs^2 + \left(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}\right)'\boldsymbol{X}'\boldsymbol{X}\left(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}\right)\right]\right\} \qquad (5.11)
\end{aligned}
$$

where

$$\begin{aligned} v &= n-k \\ \hat{\boldsymbol{\beta}} &= \left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{y} \\ \hat{\boldsymbol{y}} &= \boldsymbol{X}\hat{\boldsymbol{\beta}} \\ s^2 &= \frac{\left(\boldsymbol{y}-\boldsymbol{X}\hat{\boldsymbol{\beta}}\right)'\left(\boldsymbol{y}-\boldsymbol{X}\hat{\boldsymbol{\beta}}\right)}{v} \end{aligned}$$

and

$$\hat{\boldsymbol{\beta}}|\sigma^2 \sim N\left(\boldsymbol{\beta}, \sigma^2\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\right)$$

Posterior distribution

Combining the prior distribution in expression (5.10) with the likelihood function in expression (5.11) we get the following joint posterior distribution (Zellner, 1971)

$$\begin{aligned} \pi(\boldsymbol{\beta},\sigma^2|\boldsymbol{y},\boldsymbol{X}) &= p(\boldsymbol{y}|\boldsymbol{X},\boldsymbol{\beta},\sigma^2)p(\boldsymbol{\beta},\sigma^2) \\ \pi(\boldsymbol{\beta},\sigma^2|y,\boldsymbol{X}) &\propto \frac{1}{\sigma^n}\exp\left\{-\frac{1}{2\sigma^2}\left[vs^2+\left(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}}\right)'\boldsymbol{X}'\boldsymbol{X}\left(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}}\right)\right]\right\}\times\frac{1}{\sigma^2} \\ \pi(\boldsymbol{\beta},\sigma^2|y,\boldsymbol{X}) &\propto \frac{1}{\sigma^{n+2}}\exp\left\{-\frac{1}{2\sigma^2}\left[vs^2+\left(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}}\right)'\boldsymbol{X}'\boldsymbol{X}\left(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}}\right)\right]\right\} \quad (5.12) \end{aligned}$$

Marginal posterior densities

Integrating expression (5.12) with respect to σ^2 we get the following marginal posterior pdf for the elements of $\boldsymbol{\beta}$ (Zellner, 1971):

$$\begin{aligned} \pi(\boldsymbol{\beta}|\boldsymbol{y},\boldsymbol{X}) &= \int_0^\infty p(\boldsymbol{\beta},\sigma^2|\boldsymbol{y},\boldsymbol{X})\mathrm{d}\sigma^2 \\ \pi(\boldsymbol{\beta}|\boldsymbol{y},\boldsymbol{X}) &\propto \left[vs^2 + \left(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}}\right)'\boldsymbol{X}'\boldsymbol{X}\left(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}}\right)\right]^{-\frac{n}{2}} \end{aligned} \tag{5.13}$$

where $\hat{\boldsymbol{\beta}}$ is the posterior mean vector given the data and prior information. The marginal posterior pdf in expression (5.13) follows a multivariate Student-t distribution. The marginal posterior distribution of β_i follows a Student-t distribution, that is

$$\frac{\beta_i - \hat{\beta}_i}{s(m^{ii})^{\frac{1}{2}}} \sim t_v \tag{5.14}$$

where m^{ii} is the $(i,i)^{th}$ element of $(X'X)^{-1}$. The posterior mean of β_i is $\hat{\beta}_i$ given the observed data and prior information. Now integrating expression (5.12) with respect to the elements of $\boldsymbol{\beta}$ we get the marginal posterior pdf of σ^2, that is

$$\begin{aligned} \pi(\sigma^2|\boldsymbol{y},\boldsymbol{X}) &= \int_{-\infty}^\infty \ldots \int_{-\infty}^\infty p(\boldsymbol{\beta},\sigma^2|\boldsymbol{y},\boldsymbol{X})\mathrm{d}\boldsymbol{\beta} \\ p(\sigma^2|\boldsymbol{y},\boldsymbol{X}) &\propto \frac{1}{\sigma^{v+1}}\exp\left(\frac{vs^2}{2\sigma^2}\right) \end{aligned} \tag{5.15}$$

The marginal posterior pdf in (5.15) follows an inverse gamma pdf (Zellner, 1971).

Posterior predictive density

The predictive density is given by (Zellner, 1971)

$$\pi(\tilde{\boldsymbol{y}}|\boldsymbol{y}, \tilde{\boldsymbol{X}}, \boldsymbol{X}) = \int\int p(\tilde{\boldsymbol{y}}|\boldsymbol{\beta}, \sigma^2, \tilde{\boldsymbol{X}}) p(\boldsymbol{\beta}, \sigma^2|\boldsymbol{y}, \boldsymbol{X}) d\boldsymbol{\beta} d\sigma^2 \tag{5.16}$$

where $\tilde{\boldsymbol{y}}$ denotes the $p-$ step ahead forecasts which is a $p \times 1$ vector, i.e. $\tilde{\boldsymbol{y}} = (y_{t+1}, ..., y_{t+p})$, $\tilde{\boldsymbol{X}}$ are future observations of the independent variables and $\pi(\boldsymbol{\beta}, \sigma^2|\boldsymbol{y}, \boldsymbol{X})$ is the joint posterior distribution of $\boldsymbol{\beta}$ and σ^2 given in expression (5.12). The predictive distribution given in equation (5.16) follows a multivariate Student-t distribution, i.e. $\pi(\tilde{\boldsymbol{y}}|\boldsymbol{y}, \tilde{\boldsymbol{X}}, \boldsymbol{X}) \sim t(n-k, \tilde{\boldsymbol{X}}, \hat{\boldsymbol{\beta}}, s)$ where $s = \hat{\sigma^2}(I_p + \tilde{\boldsymbol{X}}(\boldsymbol{X}'\boldsymbol{X})^{-1}\tilde{\boldsymbol{X}}')$ and $\hat{\boldsymbol{\beta}}$ is the posterior mean of $\boldsymbol{\beta}$ (Zellner, 1971). The predictive distribution for a single component of $\tilde{\boldsymbol{y}}$ follows a univariate Student-t distribution given by $\frac{\tilde{y} - \tilde{X}^k \hat{\beta}_k}{S_{k,k}^{1/2}} \sim t_{n-k}$ (Zellner, 1971).

5.3 Load and temperature data

Daily peak electricity demand (DPED) and temperature data (peak temperature) for the period 2000 to 2010 are used for developing the piecewise linear regression model discussed in this chapter.

5.4 Piecewise linear regression model

Short term electricity load forecasting is very important for system operators who have to ensure that the amount of electricity drawn from the grid and the amount generated balances (Cottet and Smith, 2003; Taylor, 2006). In the absence of blackouts and load-shedding the electricity load is equal to electricity demand (Cottet and Smith, 2003). Short-term load forecasting is important to ensure that there is a balance between demand and supply since electricity cannot be stored (Munoz *et al*., 2010).

Most papers in literature concentrate on point forecasting only. One major drawback of point forecasting only is that it does not fully take into account uncertainty in the estimation of the parameters. One way of overcoming this, is the use of Bayesian analysis and density forecasting.

Initially a piecewise linear regression model is developed. In the second stage Bayesian statistics is used in developing posterior densities for the parameter estimates to the piecewise linear regression model. The reference temperatures (T_w and T_s) which are 18^0C and 22^0C respectively as discussed in chapter 3 are used in the development of the piecewise linear regression model. Within this range of temperature values (18^0C and 22^0C) the assumption is residents would neither use a heater nor a cooling system. The model identifies the winter sensitive, weather neutral and summer sensitive periods. The piecewise linear function used in this chapter is given in equation (5.17).

$$x_t = \beta_0 + \beta_1 t + \beta_2 X_1 + \beta_3 X_2 + \sum_{r=2}^{7} \tau_r D_{rt} + \sum_{l=2}^{12} \lambda_l M_{lt}$$

$$+\mu H_{t-1} + \delta H_t + \gamma H_{t+1} + e_t \qquad (5.17)$$

where x_t is DPED, $X_1 = (T_t^{pt} - 18)d_{1t}$ and $d_{1t} = 1$ if $T_t^{pt} < 18$ and zero otherwise; $X_2 = (T_t^{pt} - 22)d_{2t}$ and $d_{2t} = 1$ if $T_t^{pt} > 22$ and zero elsewhere; t is the trend component which is aimed at taking into account the long-term trend in the daily peak electricity demand; e_t is the error term; T_t^{pt} represents peak temperature (in degrees Celsius). The peak temperature is the temperature recorded at the hour of peak demand on day t; 18^0C is the temperature to identify where the winter sensitive portion of demand join the non-weather sensitive demand component, 22^0C is the temperature to identify where the summer sensitive portion of demand join the non-weather sensitive demand component, β_0 represents the mean daily peak demand observed in the non-weather sensitive

period $(18^0C \leq T_t^{pt} \leq 22^0C)$. The dummy variables H_t, H_{t-1} and H_{t+1} are representing holiday, day before and after a holiday respectively. The day of the week effect is represented by the dummy variable D_{rt}, where D_{rt} equals 1 if in the t^{th} observation, day r (r = Tuesday,...,Sunday) is found and 0 otherwise, with Monday as the base period. The monthly effect is represented by M_{lt}, where M_{lt} equals 1 if in the t^{th} observation the month l (l = February,...,December) is found and 0 otherwise, with January as the base month. To reduce serial correlation we include an autoregressive structure in the error term e_t. The structure of the error term e_t is given by

$$e_t = \varepsilon_t + \phi_1\varepsilon_{t-1} + \phi_3\varepsilon_{t-3} + \phi_6\varepsilon_{t-6} + \phi_7\varepsilon_{t-7} + \phi_{31}\varepsilon_{t-31} + \phi_{175}\varepsilon_{t-175} \tag{5.18}$$

The terms $\phi_1\varepsilon_{t-1}, \phi_3\varepsilon_{t-3}$ and $\phi_6\varepsilon_{t-6}$ will account for day to day variations in electricity peak demand. The week to week, monthly and semi-annual variations in electricity peak demand will be accounted for by $\phi_7\varepsilon_{t-7}, \phi_{31}\varepsilon_{t-31}$ and $\phi_{175}\varepsilon_{t-175}$ respectively.

The developed piecewise linear regression model is given in equation (5.19)

$$x_t = \beta_0 + \beta_1 t + \beta_2 X_1 + \beta_3 X_2 + \tau_5 g_{5t} + \tau_6 g_{6t} + \tau_7 g_{7t} + \mu H_{t-1} + \delta H_t + \gamma H_{t+1}$$

$$+\varepsilon_t + \phi_1\varepsilon_{t-1} + \phi_3\varepsilon_{t-3} + \phi_6\varepsilon_{t-6} + \phi_7\varepsilon_{t-7} + \phi_{31}\varepsilon_{t-31} + \phi_{175}\varepsilon_{t-175} \tag{5.19}$$

where all the variables are as defined in equation (5.17). Substituting the coefficients of the variables we get

$$\hat{x}_t = 25687.61 + 2.13t - 171.47(T_t^{pt} - 18)d_{1t} + 24.85(T_t^{pt} - 22)d_{2t} - 897.36D_{5t}$$

$$-2304.73D_{6t} - 2564.95D_{7t} - 779H_{t-1} - 1779.63H_t - 227.36H_{t+1}$$

$$+0.797\hat{\varepsilon}_{t-1} + 0.095\hat{\varepsilon}_{t-3} + 0.078\hat{\varepsilon}_{t-6} + 0.192\hat{\varepsilon}_{t-7}$$

$$-0.035\hat{\varepsilon}_{t-31} - 0.076\hat{\varepsilon}_{t-175}$$

The coefficient of $X_1 = (T_t^{pt} - 18)d_{1t}$ is negative showing that if peak temperature decreases by one degree from 18^0C, electricity demand will increase by 171.468MW. The coefficient of $X_2 = (T_t^{pt} - 22)d_{2t}$ is positive showing that if temperature increases by one degree from 22^0C, electricity demand will increase by 24.85MW. We conclude that electricity demand in South Africa is more sensitive to the winter period.

The coefficients of the dummy variables holiday, day before holiday and day after holiday are negative indicating that there is a decrease in demand during these periods. Demand for electricity during holidays decreases significantly compared to a normal working day as shown in Figure 5.1.

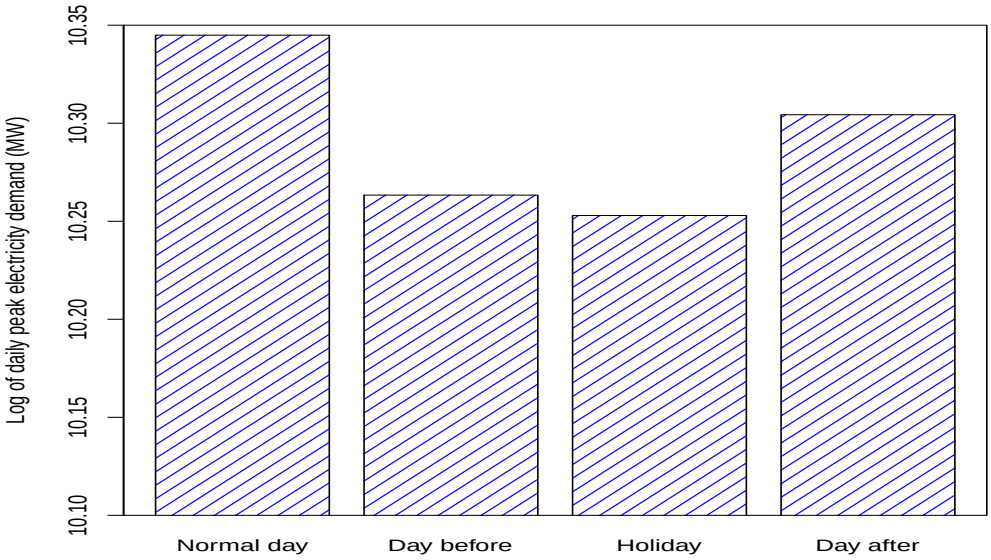

Figure 5.1: Holiday effect. During holidays electricity demand drops significantly compared to a normal day and a day before and after a holiday.

All the coefficients of the dummy variables representing Friday, Saturday and Sunday are negative indicating that there is a decrease in demand during these days. Figure 5.2 shows the day of the week effect. The largest decrease in electricity demand is on Sunday as shown in Figure 5.2.

The root mean square error (RMSE) and mean absolute percentage error (MAPE) are generally used for evaluating the predictive power of short-term load forecasting (STLF) models (Munoz *et al*., 2010). These accuracy measures are used in the in-sample forecasting evaluation of the model. Several models were run and the results of the best model are summarized in Table 5.1. For the period $t = 1, ..., m$, the RMSE, MAPE are calculated as:

$$\text{RMSE} = \sqrt{\frac{1}{m}\sum_{t=1}^{m}(x_t - f_t)^2},$$

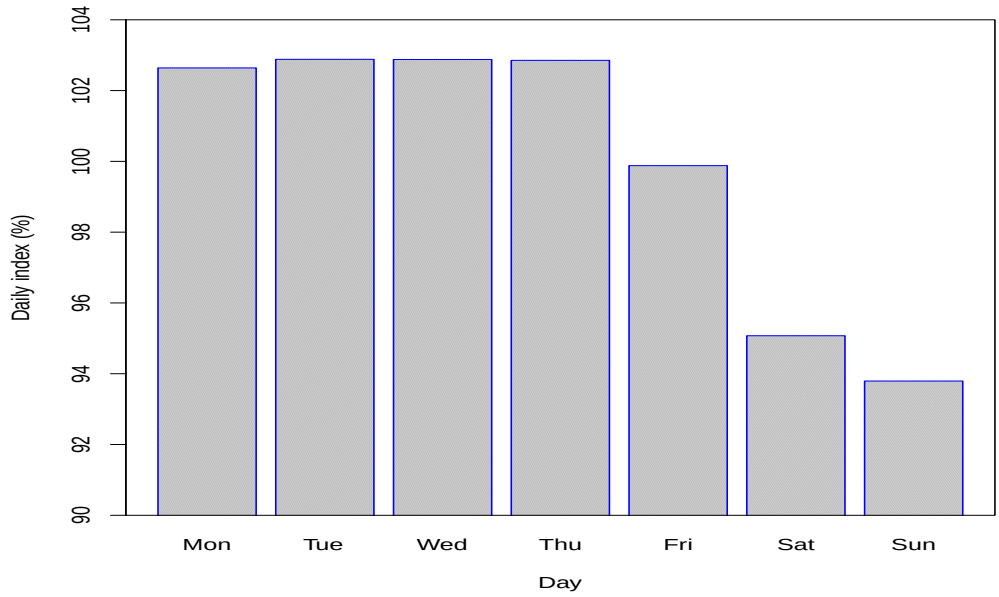

Figure 5.2: Weekly seasonal index plot of DPED. The index for each day is calculated as follows: We calculate the average DPED for each day and then divide by the overall average. Demand for electricity drops significantly during weekends as most companies and industries will be closed.

$$\text{MAPE} = \frac{1}{m}\sum_{t=1}^{m}\frac{|x_t - f_t|}{x_t}100,$$

where x_t and f_t are the actual and predicted daily peak loads at time t.

Table 5.1: In-sample forecasting evaluation.

	Performance criteria (estimation period)	
Forecasting model	MAPE	RMSE
Piecewise linear regression	1.172	422.1

5.5 Bayesian analysis: Posterior distributions of the parameters

The posterior probability density functions (pdfs) for the piecewise linear regression parameters are given in this section. We discuss the results of the

parameters $\beta_0, ..., \beta_3$ and the rest of the posterior distributions of the parameters $\beta_4, ..., \beta_9$ are summarized in Table 5.2 and given in Figure 5.4. The three demand-temperature lines for weekdays without the trend and holiday effects i.e. assuming $(t = D_{5t} = D_{6t} = D_{7t} = H_{t-1} = H_t = H_{t+1} = 0)$ are shown below. For the non-weather sensitive months $(18 \leq T_t^{pt} \leq 22, d_{1t} = d_{2t} = 0)$ we get

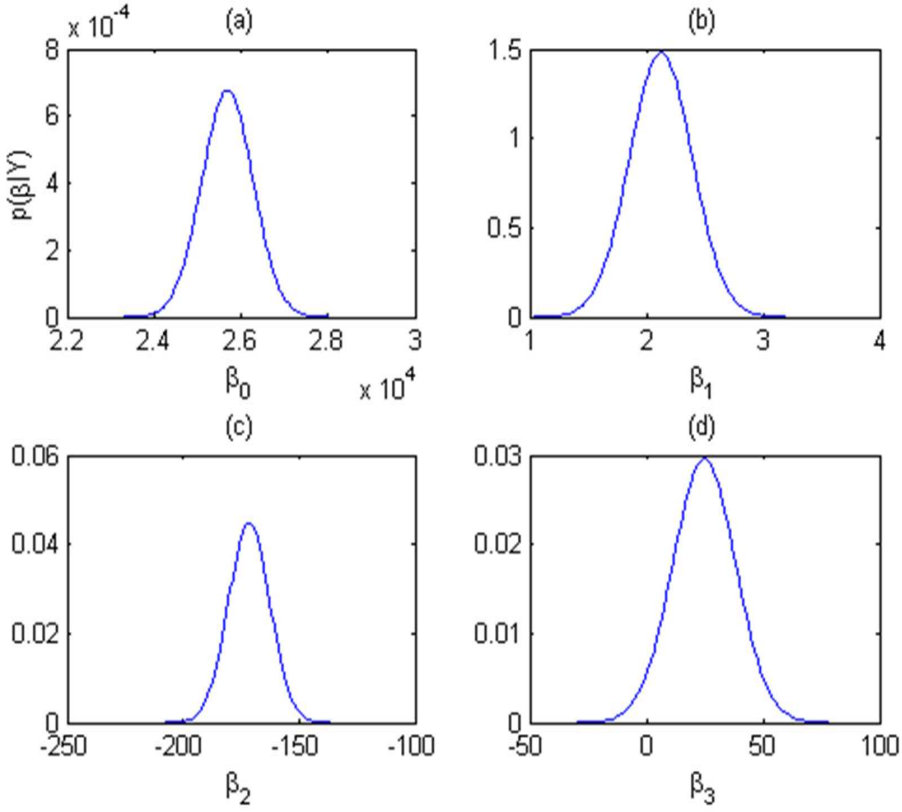

Figure 5.3: Marginal posterior distributions of the model parameters: (a) Top left panel: β_0 (b) Top right panel: β_1, (c) Bottom left panel: β_2 and (d) Bottom right panel: β_3.

$E(\beta_0) = \hat{\beta}_0 = 25687.61$MW is the posterior mean daily peak electricity demand for the non-weather sensitive period. The posterior pdf for β_0 shown in Figure 5.3(a) shows the non-weather sensitive mean daily peak demand could take any value between 23500MW and 28000MW.

There is a persistent increase in daily peak electricity demand with time in South Africa. The slope is around 2MW per day. It is highly unlikely that the

slope exceeds 3 and highly unlikely that it is below 1. The calculated mean slope increase is 2.126MW. Figure 5.3(b) shows the posterior pdf for β_1. For the winter sensitive months ($T_t^{pt} < 18, d_{1t} = 1, d_{2t} = 0$) we get

$$\hat{\text{DPED}} = 25687.61 - 171.468(T_t^{pt} - 18)d_{1t} \tag{5.20}$$

That is if temperature decreases by 1^0C (e.g. from 18^0C to 17^0C) the marginal increase in daily peak demand will be 171.468MW. Because of the uncertainties associated with demand, this increase could be any value between 140MW and 200MW. In other words, it is highly unlikely that this increase will be less than 140MW and highly unlikely that it will exceed 200MW. Figure 5.3(c) shows the posterior pdf for β_2. For the summer sensitive months ($T_t^{pt} > 22, d_{1t} = 0, d_{2t} = 1$) we get

$$\hat{\text{DPED}} = 25687.61 + 24.85(T_t^{pt} - 22)d_{2t} \tag{5.21}$$

That is if temperature increases by 1^0C (e.g. from 22^0C to 23^0C) the marginal increase in demand will be 24.85MW. This marginal increase could be any value between -20MW and 80MW. There is a possibility of no change since zero is included in the interval. The posterior pdf for parameter β_3 is shown in Figure 5.3(d). Table 5.2 summarizes the 99% Bayesian Credible intervals of the posterior distributions of the parameters of the piecewise linear regression model.

5.6 Discussion

A piecewise linear regression model is developed which relates peak demand to temperature, day of the week and the holiday effect which includes the day before and the day after a holiday. The developed model accounted for the week to week, monthly and semi-annual variations in peak electricity demand. This study has shown that a Bayesian analysis provides a way of taking into account

Table 5.2: 99% Bayesian Credible intervals.

Parameter	99% Bayesian Credible intervals	Posterior means of the parameters $\hat{\beta}_i$
β_0	23500 to 28000	$E(\beta_0) = 25687.61$
β_1	1 to 3	$E(\beta_1) = 2.126000$
β_2	-200 to -140	$E(\beta_2) = -171.470$
β_3	-20 to 80	$E(\beta_3) = 24.85000$
β_4	-960 to -820	$E(\beta_4) = -897.360$
β_5	-2400 to -2200	$E(\beta_5) = -2304.73$
β_6	-2640 to -2480	$E(\beta_6) = -2564.95$
β_7	-925 to -650	$E(\beta_7) = -779.490$
β_8	-1950 to -1650	$E(\beta_8) = -1779.63$
β_9	-400 to -125	$E(\beta_9) = -227.360$

uncertainty in the estimation of the parameters. Uncertainty about the true values of the piecewise linear regression parameter estimates is incorporated into the analysis through the choice of a non-informative prior distribution. Empirical results show that if temperature decreases by 1^0C in the winter sensitive period (that is when temperature is less than 18^0C) then the average increase in daily peak electricity load will be 171.47 MW. This increase in daily peak electricity load ranges between 140 and 200 MW. Similarly if temperature increases by 1^0C in the summer sensitive period (that is when temperature is greater than 22^0C) daily peak electricity load will increase by an average of 24.85 MW. This increase will range from -20 to 80MW. These results show that demand of electricity is more sensitive to winter periods than summer periods in South Africa. This quantification of uncertainty about these parameters is important for load forecasters in Eskom, South Africa's power utility company as it helps them in the determination of consistent and reliable supply schedules. The developed model can be used for modelling the influence of temperature on daily peak electricity demand. The mean values of the parameters of the Bayesian approach are the same to the frequentist analogy because of the nature of the non-informative prior employed. It therefore seems that the frequentist properties of the Bayesian inferences of the model based on the prior

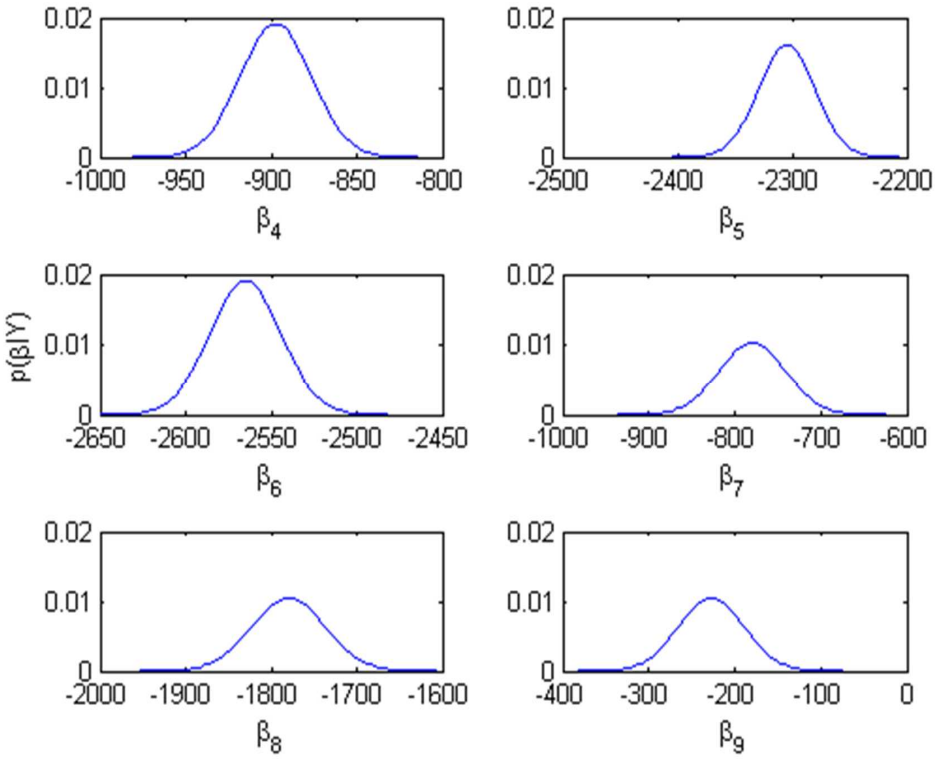

Figure 5.4: Posterior pdf for parameters $\beta_4, ..., \beta_9$.

are adequate. However the resultant posterior may be used in a way that is not always available to a frequentist.

5.7 Concluding remarks

The chapter discussed the use of Bayesian statistics for modelling the influence of temperature on daily peak electricity demand in South Africa. The developed piecewise linear regression model captures demand drivers of electricity demand such as temperature, calendar effects and lagged demand. The model also accounts for residual correlation that may occur as a result of day to day, week to week, monthly and seasonal variations in peak electricity demand. This empirical study provides an extension of point forecasting to den-

sity forecasting so as to take into account uncertainty in the estimation of the parameters. The main contribution in this chapter is the development of a piecewise linear regression model which captures both weather and calendar variables and bringing out the relative contribution of the influence of these important drivers of short-term electricity demand and associated uncertainties in the estimation of parameters.

The model developed in this chapter and the previous two chapters 3 and 4 respectively have been used for modelling the influence of temperature on electricity demand. In the next chapter we are going to develop time series regression models which will be used for short term (up to seven days ahead) out of sample forecasts of daily peak electricity demand.

Chapter 6

Regression-SARIMA modelling of daily peak electricity demand in South Africa

6.1 Introduction

The chapter discusses an application of time series regression models in short term forecasting of up to seven days ahead of Daily Peak Electricity Demand (DPED) using South African data. These models are the seasonal autoregressive integrated moving average (SARIMA), regression-SARIMA (Reg-SARIMA) and regression-SARIMA-GARCH (Reg-SARIMA-GARCH). The Reg-SARIMA modelling framework captures important drivers of electricity demand including multiple seasonality. These drivers of electricity demand are generally divided into weather variables, economic and calendar effects (Hyndman and Fan, 2010; Munoz *et al.*, 2010).

The major challenge in most conventional Reg-SARIMA models is that of selecting a minimum number of predictor variables and ranking them in order of their importance. The proposed models developed in this chapter are designed in such a way that the predictor variables are initially selected based on measures of predictive accuracy which are cross-validation (Arlot and Celisse, 2010), Akaike Information Criterion (AIC), Schwarz Bayesian Information Criterion (SBIC) and Adjusted R^2 (Hyndman and Athanasopoulos, 2012).

6.2 SARIMA and Regression-SARIMA models

6.2.1 SARIMA(*p,d,q*)(*P,D,Q*)[s] Model

The general SARIMA model can be represented analytically as (Box and Jenkins, 1976):

$$\phi_p(B)\Phi_P(B^s)\nabla^d\nabla_s^D x_t \;=\; c+\theta_q(B)\Theta_Q(B^s)\varepsilon_t,\;\; \varepsilon_t \sim N(0,\sigma_\varepsilon^2) \tag{6.1}$$

where x_t represents Daily Peak Electricity Demand (DPED) at time t, where $t = 1, 2, ..., n$, ε_t is the error term at time t, s is the seasonal length, B is a backshift (lag) operator ($Bz_t = z_{t-1}$). The polynomial $\phi_p(B) = (1 - \phi_1 B - ... - \phi_p B^p)$ is the nonseasonal autoregressive (AR) operator, $\Phi_P(B^s) = (1 - \Phi_1 B^s - ... - \Phi_P B^{Ps})$ is the seasonal AR operator, $\theta_q(B) = (1 - \theta_1 B - ... - \theta_q B^q)$ is the nonseasonal moving average (MA) operator, $\Theta_Q(B^s) = (1-\Theta_1 B^s - ... - \Theta_Q B^{Qs})$ is the seasonal MA operator. The operators ∇^d and ∇_s^D are the nonseasonal and seasonal difference operators of order d and D respectively, where $\nabla^d = (1-B)^d$ and $\nabla_s^D = (1 - B^s)^D$.

6.2.2 Estimation of the SARIMA parameters

The Maximum Likelihood (ML) method is used to estimate the parameters of the model given in equation (6.2). Assuming the initial conditions

$$\boldsymbol{x}^* = (x_{1-p}, ..., x_{-1}, x_0)'$$

and

$$\boldsymbol{\varepsilon}^* = (\varepsilon_{1-q}, ..., \varepsilon_{-1}, \varepsilon_0)'$$

the conditional log-likelihood function (Wei, 2006) is given as

$$\log L^* \left(\boldsymbol{\phi}, \mu, \boldsymbol{\theta}, \boldsymbol{\Phi}, \boldsymbol{\Theta}, \sigma_\varepsilon^2\right) = -\frac{n}{2}\log(2\pi\sigma_\varepsilon^2) - \frac{S^*(\boldsymbol{\phi}, \mu, \boldsymbol{\theta}, \boldsymbol{\Phi}, \boldsymbol{\Theta})}{2\sigma_\varepsilon^2} \tag{6.2}$$

where $S^*(\boldsymbol{\phi}, \mu, \boldsymbol{\theta}, \boldsymbol{\Phi}, \boldsymbol{\Theta})$ is the conditional sum of squares function given by

$$S^*(\boldsymbol{\phi}, \mu, \boldsymbol{\theta}, \boldsymbol{\Phi}, \boldsymbol{\Theta}) = \sum_{t=1}^{n} \varepsilon_t^2(\boldsymbol{\phi}, \mu, \boldsymbol{\theta}, \boldsymbol{\Phi}, \boldsymbol{\Theta} | \boldsymbol{x}^*, \boldsymbol{\varepsilon}^*, \boldsymbol{x}) \tag{6.3}$$

where $\boldsymbol{\phi} = (\phi_1, ..., \phi_p)$, $\boldsymbol{\theta} = (\theta_1, ..., \theta_q)$, $\boldsymbol{\Phi} = (\Phi_1, ..., \Phi_P)$, $\boldsymbol{\Theta} = (\Theta_1, ..., \Theta_Q)$ and $\boldsymbol{x} = (x_1, ..., x_n)$. The quantities $\hat{\boldsymbol{\phi}}, \hat{\mu}, \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\Phi}}, \hat{\boldsymbol{\Theta}}$ which maximizes equation (6.2) are the conditional maximum likelihood estimators (Wei, 2006). The model given in equation (6.2) is used for short term forecasting of up to seven days ahead. The modelling is done using South African DPED data.

6.2.3 Regression-SARIMA Model

The regression-SARIMA (Reg-SARIMA) models also known as the ARIMAX models are an extension of the SARIMA model. They are often used in load forecasting to capture important drivers of demand such as weather, economic and calendar variables (Bunn, 1982; Weron, 2006; Suganthi and Samuel, 2012; among others). The developed Reg-SARIMA model will be used to identify im-

portant drivers of daily peak electricity demand in South Africa and also for short term out of sample predictions.

The Reg-SARIMA model is a regression model with SARIMA errors (Hyndman and Athanasopoulos, 2012). The Reg-SARIMA model reduces to a SARIMA model if the independent variables are not used. The Reg-SARIMA model is derived in the following paragraphs. A linear regression model with a SARIMA error structure is given as:

$$x_t = \sum_{f=1}^{k} \beta_f x_{ft} + z_t \tag{6.4}$$

where x_t is the dependent time series, x_{ft} are the explanatory variables, β_f are the regression parameters and z_t is an error term which is assumed to follow a SARIMA model as shown in equation (6.5).

$$\phi_p(B)\Phi_P(B^s)\nabla^d\nabla_s^D z_t = c + \theta_q(B)\Theta_Q(B^s)\varepsilon_t \tag{6.5}$$

Making z_t the subject of equation (6.4) we get

$$z_t = x_t - \sum_{f=1}^{k} \beta_f x_{ft} \tag{6.6}$$

The expression of z_t in equation (6.6) is then substituted in equation (6.5) to get

$$\phi_p(B)\Phi_P(B^s)\nabla^d\nabla_s^D \left[x_t - \sum_{f=1}^{k} \beta_f x_{ft} \right] = c + \theta_q(B)\Theta_Q(B^s)\varepsilon_t$$

$$\phi_p(B)\Phi_P(B^s)\left[\nabla^d\nabla_s^D\left(x_t-\sum_{f=1}^{k}\beta_f x_{ft}\right)\right] = c+\theta_q(B)\Theta_Q(B^s)\varepsilon_t$$

$$\phi_p(B)\Phi_P(B^s)\left[\nabla^d\nabla_s^D x_t-\sum_{f=1}^{k}\beta_f\nabla^d\nabla_s^D x_{ft}\right] = c+\theta_q(B)\Theta_Q(B^s)\varepsilon_t \quad (6.7)$$

Let

$$w_t = \nabla^d\nabla_s^D x_t-\sum_{f=1}^{k}\beta_f\nabla^d\nabla_s^D x_{ft} \quad (6.8)$$

then equation (6.7) can be written as

$$\phi_p(B)\Phi_P(B^s)w_t = c+\theta_q(B)\Theta_Q(B^s)\varepsilon_t \quad (6.9)$$

In order to capture the day of the week effect dummy variables are introduced. Six dummy variables are used with Monday as a base day (Makridakis *et al.*, 1998; Mirasgedis *et al.*, 2006). The dummy variables are defined as follows (as in Chapter 5):

$$D_{rt} = \begin{cases} 1 & r = \text{Tuesday,...,Sunday} \\ 0 & \text{otherwise} \end{cases} \quad (6.10)$$

where t represents the observation number, i.e. $t = 1, 2, ..., n$. The dummy variable D_{rt} takes value 1 if day r is found in observation t and zero otherwise. DPED decreases during holidays. Some companies do close earlier on a day before a holiday. There is a reduction in electricity demand a day before and after a holiday (Ismail *et al.*, 2008). To take into account the effects of holidays

the following dummy variables H_t, H_{t-1} and H_{t+1} are introduced

$$H_t = \begin{cases} 1 & t = \text{a day which is a holiday} \\ 0 & \text{otherwise} \end{cases} \tag{6.11}$$

$$H_{t-1} = \begin{cases} 1 & t = \text{a day before a holiday} \\ 0 & \text{otherwise} \end{cases} \tag{6.12}$$

$$H_{t+1} = \begin{cases} 1 & t = \text{ a day after a holiday} \\ 0 & \text{otherwise} \end{cases} \tag{6.13}$$

South African holidays are: New years day (NY) [1 January], Human rights day (HR) [21 March], Good Friday (GF) (Easter holiday, the date changes from year to year between the months of March and April), Family day (F) (Easter holiday, usually in April), Freedom (FD) day [27 April], Workers (W) day [1 May], Youth (Y) day [16 June], National women's (NW) day [9 August], Heritage (H) day [24 September], Day of reconciliation (DR) [16 December], Christmas (C) day [25 December] and Day of goodwill (DG) [26 December]. South African holidays will be applied to the model given in equation (6.15).

To take into account the monthly seasonality effect a dummy variable M_{lt} is introduced. Eleven dummy variables are used and January is taken as the base month.

$$M_{lt} = \begin{cases} 1 & l = \text{February,...,December} \\ 0 & \text{otherwise} \end{cases} \tag{6.14}$$

where $t = 1, 2, ..., n$. If $t = l$ the dummy variable M_{lt} takes value 1 and zero otherwise.

In order to capture the demand effects we incorporate in the model the following variables; minimum daily electricity demand in the past 24 hours (x_t^-), average daily electricity demand in the past seven days ($\bar{x}_t$). To take into account the effect of temperature on DPED the following variables are included: maximum temperature on day t (T_t^+), maximum temperature on day t squared (T_t^{2+}), minimum temperature on day t (T_t^-), minimum temperature on day t squared (T_t^{2-}), peak temperature on day t (T_t^{pt}) where peak temperature is the temperature recorded at the hour of peak demand on day t, peak temperature on day t squared (T_t^{2pt}), average temperature in the past seven days ($\bar{T}_t$) and average temperature in the past seven days squared ($\bar{T}_t^2$). The Reg-SARIMA model is then written as

$$\phi_p(B)\Phi_P(B^s)(1-B)^d(1-B^s)^D[x_t - \sum_{l=2}^{12}\lambda_l M_{lt} - \sum_{r=2}^{7}\tau_r D_{rt} - \gamma H_{t-1} - \kappa H_t - \rho H_{t+1}$$

$$-\alpha_1 x_t^- - \alpha_2 \bar{x}_t - \alpha_3 T_t^+ - \alpha_4 T_t^{2+} - \alpha_5 T_t^- - \alpha_6 T_t^{2-} - \alpha_7 T_t^{pt}$$

$$-\alpha_8 T_t^{2pt} - \alpha_9 \bar{T}_t - \alpha_{10}\bar{T}_t^2] = c + \theta_q(B)\Theta_Q(B^s)\varepsilon_t \tag{6.15}$$

Equation 6.8 can now be written as

$$w_t = (1-B)^d(1-B^s)^D[x_t - \sum_{l=2}^{12}\lambda_l M_{lt} - \sum_{r=2}^{7}\tau_r D_{rt} - \gamma H_{t-1}$$

$$-\kappa H_t - \rho H_{t+1} - \alpha_1 x_t^- - \alpha_2 \bar{x}_t - \alpha_3 T_t^+ - \alpha_4 T_t^{2+} - \alpha_5 T_t^-$$

$$-\alpha_6 T_t^{2-} - \alpha_7 T_t^{pt} - \alpha_8 T_t^{2pt} - \alpha_9 \bar{T}_t - \alpha_{10}\bar{T}_t^2] \tag{6.16}$$

Combining equations 6.15 and 6.16 we get

$$\phi_p(B)\Phi_P(B^s)w_t = c + \theta_q(B)\Theta_Q(B^s)\varepsilon_t \tag{6.17}$$

The parameters are estimated using the ML method. The conditional log-likelihood function (Wei, 2006) is given as

$$\log L^*\left(\boldsymbol{\phi}, \mu, \boldsymbol{\theta}, \boldsymbol{\Phi}, \boldsymbol{\Theta}, \sigma_\varepsilon^2, \boldsymbol{\tau}, \boldsymbol{\lambda}, \boldsymbol{\alpha}, \gamma, \kappa, \rho\right) = -\frac{n}{2}\log 2\pi\sigma_\varepsilon^2 -$$

$$\frac{S^*(\boldsymbol{\phi}, \mu, \boldsymbol{\theta}, \boldsymbol{\Phi}, \boldsymbol{\Theta}, \boldsymbol{\tau}, \boldsymbol{\lambda}, \boldsymbol{\alpha}, \gamma, \kappa, \rho)}{2\sigma_\varepsilon^2} \tag{6.18}$$

where

$$S^*(\boldsymbol{\phi}, \mu, \boldsymbol{\theta}, \boldsymbol{\Phi}, \boldsymbol{\Theta}, \boldsymbol{\tau}, \boldsymbol{\lambda}, \boldsymbol{\alpha}, \gamma, \kappa, \rho) = \sum_{t=1}^{n}\varepsilon_t^2(\boldsymbol{\phi}, \mu, \boldsymbol{\theta}, \boldsymbol{\Phi}, \boldsymbol{\Theta}, \boldsymbol{\tau}, \boldsymbol{\lambda}, \boldsymbol{\alpha}, \gamma, \kappa, \rho | \boldsymbol{x^*}, \boldsymbol{\varepsilon^*}, \boldsymbol{x}) \tag{6.19}$$

is the conditional sum of squares function. The vector of parameters $\boldsymbol{\alpha}, \boldsymbol{\tau}$ and $\boldsymbol{\lambda}$ are given as follows; $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_8)$, $\boldsymbol{\tau} = (\tau_2, ..., \tau_7)$, $\boldsymbol{\lambda} = (\lambda_2, ..., \lambda_{12})$, and the parameters $\phi, \mu, \boldsymbol{\theta}, \boldsymbol{\Phi}, \boldsymbol{\Theta}, \gamma, \kappa$ and ρ are as defined in Section 6.2.2. The conditional maximum likelihood estimators which maximizes equation (6.18) are $\hat{\boldsymbol{\phi}}, \hat{\mu}, \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\Phi}}, \hat{\boldsymbol{\Theta}}, \hat{\boldsymbol{\tau}}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\alpha}}, \hat{\gamma}, \hat{\kappa}, \hat{\rho}$.

6.3 SARIMA-GARCH and Reg-SARIMA-GARCH models

The Reg-SARIMA modelling approach is extended to a Reg-SARIMA-GARCH model. The developed model is then used for modelling DPED. In all models DPED is taken as the dependent variable.

6.3.1 SARIMA-GARCH Model

Up to now we have assumed the variance to be constant (homoskedastic). In this section we introduce the SARIMA-GARCH model which is meant to accommodate the possibility of serial correlation in volatility since the daily peak demand data exhibits non-constant mean and variance, and multiple seasonality corresponding to weekly and monthly periodicity. The SARIMA-GARCH model is one in which the variance of the error term of the SARIMA model follows a GARCH process. This modelling approach is used to capture the potential conditional heteroskedasticity in electricity data (Bystrom, 2005; Taylor, 2006; among others). The model used for the DPED series can be written as:

$$\begin{aligned}
\phi_p(B)\Phi_P(B^s)x_t &= c + \theta_q(B)\Theta_Q(B^s)\varepsilon_t \\
\varepsilon_t &= \nu_t\sigma_t, \\
\nu_t \sim i.i.d & \qquad \text{with } E(\nu_t) = 0,\ \text{Var}\ (\nu_t) = 1 \\
\sigma_t^2 &= a_0 + \sum_{i=1}^{q} a_i\varepsilon_{t-i}^2 + \sum_{j=1}^{p} b_j\sigma_{t-j}^2 \qquad (6.20)
\end{aligned}$$

where x_t represents the time series as defined in Equation (6.1), p is the order of GARCH process; q is the order of ARCH process; ε_t is the error term; σ_t^2 is the conditional variance of ε_t; ε_{t-i}^2 is the news about the volatility from the i^{th} lag period and σ_{t-j}^2 is the j^{th} lag period forecast error variance, ν_t is a standardized error term; $c; a_0; a_i, i = 1, ..., q; b_j, j = 1, ..., p$ are parameters. The model parameters are estimated using the maximum likelihood method based on the assumption of the conditional distribution of the error term ε_t. The conditional distributions are that the error term either follows a normal (Gaussian) distribution, Student's t-distribution or the Generalized Error Distribution (GED). The estimates are obtained by the Berndt *et al.* (1974) algorithm using numerical derivatives. Using equation (6.20) the log-likelihood function under the assumption that the error term follows a standard Gaussian distribution (Laurent, 2010) is given by

$$\log L_{\text{norm}} = -\frac{1}{2}\sum_{t=1}^{n}\left[\log(2\pi) + \log(\sigma_t^2) + \nu_t\right] \tag{6.21}$$

Under the assumption of a Student-t distribution (Laurent, 2010) the log-likelihood is

$$\begin{aligned}\log L_{\text{Stud}} = n\left\{\log\Gamma\left(\frac{u+1}{2}\right) - \log\Gamma\left(\frac{u}{2}\right) - \frac{1}{2}\log\left[\pi(u-2)\right]\right\} \\ -\frac{1}{2}\sum_{t=1}^{n}\left[\log(\sigma_t^2) + (1+u)\log\left(1 + \frac{\nu_t^2}{u-2}\right)\right]\end{aligned} \tag{6.22}$$

where u is the degrees of freedom. The GED log-likelihood function (Laurent, 2010) is given by

$$\log L_{\text{GED}} = \sum_{t=1}^{n}\left[\log\left(\frac{u}{\lambda_u}\right) - 0.5\left|\frac{\nu_t}{\lambda_u}\right|^u - (1+u^{-1})\log(2) - \log\Gamma\left(\frac{1}{u}\right) - 0.5\log(\sigma_t^2)\right] \tag{6.23}$$

where $u > 0$ and $\lambda_u = \sqrt{\frac{\Gamma\left(\frac{1}{u}\right)2^{-2/u}}{\Gamma\left(\frac{3}{u}\right)}}$. The modelling is done using South African DPED data.

6.3.2 Reg-SARIMA-GARCH Model

There are various factors that influence electricity load demand. Some of these factors discussed include temperature, day of the week, holidays, daily and monthly seasonality. In this section, a Reg-SARIMA-GARCH model is developed which will capture the day of the week, holiday and monthly seasonality effects. The Reg-SARIMA-GARCH model is a regression seasonal ARIMA model with error terms following a GARCH process. The Reg-SARIMA-GARCH model which is an extension of the Reg-SARIMA model given in equation (6.15) can be written as

$$\begin{aligned}
\phi_p(B)\Phi_P(B^s)w_t &= c + \theta_q(B)\Theta_Q(B^s)\varepsilon_t \\
\varepsilon_t &= \nu_t\sigma_t, \\
\nu_t \sim i.i.d & \qquad \text{with } E(\nu_t) = 0,\ \text{Var}\ (\nu_t) = 1 \\
\sigma_t^2 &= \boldsymbol{\gamma q_t'}
\end{aligned}$$

where the vectors represent

$$\begin{aligned}
\boldsymbol{q_t} &= \left(1, \varepsilon_{t-1}^2, ..., \varepsilon_{t-q}^2, \sigma_{t-1}^2, ..., \sigma_{t-p}^2\right), \\
&\quad \text{a vector of parameters}, \\
\boldsymbol{\gamma} &= \left(a_0, a_1, ..., a_q, b_1, ..., b_p\right), \text{a vector of parameters},
\end{aligned}$$

and

$$w_t = (1-B)^d(1-B^s)^D[x_t - \sum_{l=2}^{12}\lambda_l M_{lt} - \sum_{r=2}^{7}\tau_r D_{rt} - \gamma H_{t-1} - \kappa H_t - \rho H_{t+1}$$

$$-\alpha_1 x_t^- - \alpha_2 \bar{x}_t - \alpha_3 T_t^+ - \alpha_4 T_t^{2+} - \alpha_5 T_t^- - \alpha_6 T_t^{2-} - \alpha_7 T_t^{pt}$$

$$-\alpha_8 T_t^{pt2} - \alpha_9 \bar{T}_t - \alpha_{10}\bar{T}_t^2] \quad (6.24)$$

where variables and parameters are as defined in Section 6.2. The model parameters are then estimated using the maximum likelihood method. The estimates are obtained by the Berndt *et al.* (1974) algorithm using numerical derivatives and South African daily peak electricity demand data is used.

6.3.3 Bayesian GARCH(1,1) modelling with Student-t innovations

Bayesian estimation of the GARCH(1,1) model with normal innovations is discussed in Ardia (2006). This work is extended to the Bayesian modelling of the GARCH(1,1) model with Student-t innovations in Ardia and Hoogerheide (2010). We present in this section a summary of the Bayesian GARCH(1,1) modelling with Student-t innovations discussed in Ardia and Hoogerheide (2010).

The GARCH(1,1) model with Student-t innovations for a return series $r_t = x_t - f_t$ where x_t denotes DPED and f_t represents forecasted DPED is given as (see Ardia and Hoogerheide, 2010 for details)

$$\text{Mean equation: } r_t = \varepsilon_t(z\sigma_t\Psi_t)^{\frac{1}{2}} \quad (6.25)$$

$$\text{Variance equation: } \sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2 \quad (6.26)$$

where r_t is the return series, $t = 1, ..., n$, $\omega > 0, \alpha > 0, \beta > 0, z = \frac{\nu-2}{\nu}$, $\Psi_t \sim \mathcal{IG}(\frac{v}{2}, \frac{v}{2})$ represents the inverse gamma distribution and ε_t is the error term which is assumed to follow a Student-t distribution with ν degrees of freedom (Ardia and Hoogerheide, 2010).

Joint prior distribution

The joint prior (Ardia and Hoogerheide, 2010) is given as follows:
Let $\boldsymbol{\theta} = (\boldsymbol{a}, \beta, \nu)$ where $\boldsymbol{a} = (\omega, \alpha)$ then the joint prior on the parameter $\boldsymbol{\theta}$ is given by

$$\pi(\boldsymbol{\theta}) = \pi(\boldsymbol{a})\pi(\beta)\pi(\nu)) \tag{6.27}$$

assuming prior independence (Ardia and Hoogerheide, 2010). Let

$$\boldsymbol{r} \doteq (r_1, ..., r_n)', \boldsymbol{\Psi} \doteq (\Psi_1, ..., \Psi_n)', \boldsymbol{a} \doteq (\omega, \alpha)'$$

and

$$\Sigma \doteq \Sigma(\boldsymbol{\theta}, \boldsymbol{\Psi}) = \text{diag}\left(\left[\Psi_t \frac{\nu-2}{\nu}\sigma_t^2(\boldsymbol{a}, \beta)\right]_{t=1}^{n}\right)$$

where

$$\sigma_t^2(\boldsymbol{a}, \beta) \doteq \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2(\boldsymbol{a}, \beta)$$

The joint prior distribution will therefore be given by $\pi(\boldsymbol{\theta}, \boldsymbol{\Psi})$ (Ardia and Hoogerheide, 2010).

Likelihood function

The likelihood function, $\pi(\boldsymbol{r}|\boldsymbol{\theta}, \boldsymbol{\Psi})$ (Ardia and Hoogerheide, 2010) is given as

$$\pi(\boldsymbol{r}|\boldsymbol{\theta}, \boldsymbol{\Psi}) \propto (\det\sigma)^{-1/2} \exp\left[-\frac{1}{2}\boldsymbol{r}'\sigma^{-1}\boldsymbol{r}\right] \tag{6.28}$$

Posterior distribution

Combining the joint prior and the likelihood function using Bayes theorem we get the following posterior distribution

$$\pi(\boldsymbol{\theta}, \boldsymbol{\Psi}|\boldsymbol{r}) \propto \pi(\boldsymbol{r}|\boldsymbol{\theta}, \boldsymbol{\Psi})\pi(\boldsymbol{\theta}, \boldsymbol{\Psi}) \tag{6.29}$$

The modelling is done for residual return series of the developed Reg-SARIMA models using South African data.

6.4 Data description

The relationship between daily peak load and the predictor variables slowly changes over time and inclusion of a larger training data set is not usually helpful (Fan and Hyndman, 2012). The training period is 1 January 2006 to 31 July 2010. The validation (testing) period is from 1-27 August 2010. The developed models are meant for short-term forecasting up to seven days ahead.

6.5 Regression-SARIMA model results

The Reg-SARIMA model developed in Section 6.2.3 (equation 6.15) is used in forecasting DPED. After carrying out the Box-Cox (1964) class of power transformations we found out that the logarithm best fits the DPED data. The logarithmic transformation helps to stabilize the volatility in the DPED data (Pardo

et al., 2002; Mirasgedis *et al.*, 2006). The equation is

$$\phi_p(B)\Phi_P(B^s)(1-B)^d(1-B^s)^D[\log(x_t) - \sum_{l=2}^{12}\lambda_l M_{lt} - \sum_{r=2}^{7}\tau_r D_{rt} - \gamma H_{t-1}$$

$$-\kappa H_t - \rho H_{t+1} - \alpha_1 x_t^- - \alpha_2 \bar{x}_t - \alpha_3 T_t^+ - \alpha_4 T_t^{2+} - \alpha_5 T_t^- - \alpha_6 T_t^{2-}$$

$$-\alpha_7 T_t^{pt} - \alpha_8 T_t^{pt2} - \alpha_9 \bar{T}_t - \alpha_{10}\bar{T}_t^2] = c + \theta_q(B)\Theta_Q(B^s)\varepsilon_t \quad (6.30)$$

Out of the 30 predictor variables (demand, calendar and temperature variables) 19 are selected for inclusion in the model. The variables selected are: minimum demand in the past 24 hours (x_t^-), average demand in the past seven days ($\bar{x}_t$), maximum temperature squared (T_t^{2+}), maximum temperature (T_t^+), minimum temperature squared (T_t^{2-}), minimum temperature (T_t^-), peak temperature (T_t^{pt}), peak temperature squared (T_t^{pt2}), average temperature in the past seven days ($\bar{T}_t$), average temperature in the past seven days squared ($\bar{T}_t^2$), Tuesday (D_{2t}), Wednesday (D_{3t}), Thursday (D_{4t}), Friday (D_{5t}), Saturday (D_{6t}), Sunday (D_{7t}), Holiday (H_t), before holiday (H_{t-1}) and after holiday (H_{t+1}). The peak temperature (T_t^{pt}) is calculated as a weighted average as given in equation (6.31).

$$T_t^{pt} = \frac{f_1 H9 + f_2 H10 + f_3 H19 + f_4 H20 + f_5 H21}{\sum_{j=1}^{5} f_j} \quad (6.31)$$

where $f_j, j = 1, ..., 5$ are frequencies of DPED occurring in different hours of the day (H9, H10, H19, H20 and H21 respectively, with H9 representing 09:00hrs, H10 is 10:00hrs, H19 is 19:00hrs, etc.) as given in Figure 3.4, in Chapter 3.

6.5.1 Reg-SARIMA model 1

The effect of temperature on electricity demand is known to be non-linear. The Reg-SARIMA model 1 discussed in this section uses the approach discussed in Ramananthan *et al.,* (1997) where we square temperature to capture the non-linear relationship (which is quadratic) between temperature and electricity demand. DPED is initially transformed by taking logarithms. Several Reg-SARIMA models are considered and the best model is selected based on the Adjusted R-squared and AIC values. The developed Reg-SARIMA model 1 is given as

$$\log(x_t) = c + \tau_2 D_{2t} + \tau_3 D_{3t} + \tau_4 D_{4t} + \tau_5 D_{5t} + \tau_6 D_{6t} + \tau_7 D_{7t} + \gamma H_{t-1}$$

$$+\kappa H_t + \rho H_{t+1} + \alpha_2 \bar{x}_t + \alpha_3 T_t^{+} + \alpha_7 T_t^{pt} + \alpha_8 T_t^{pt2} + z_t \qquad (6.32)$$

where the error term z_t follows a SARIMA model given in equation (6.33) with ε_t as a white noise process.

$$(1 - \phi_1 B - \phi_2 B^2)(1 - \Phi_1 B^7 - \Phi_2 B^{14} - \Phi_3 B^{21})(1 - B^7) z_t = \varepsilon_t \qquad (6.33)$$

Combining equations (6.32) and (6.33) we get equation (6.34).

$$(1 - \phi_1 B - \phi_2 B^2)(1 - \Phi_1 B^7 - \Phi_2 B^{14} - \Phi_3 B^{21})(1 - B^7)[\log(x_t) - c$$

$$-\tau_2 D_{2t} - \tau_3 D_{3t} - \tau_4 D_{4t} - \tau_5 D_{5t} - \tau_6 D_{6t} - \tau_7 D_{7t} - \gamma H_{t-1}$$

$$-\kappa H_t - \rho H_{t+1} - \alpha_2 \bar{x}_t - \alpha_3 T_t^{+}$$

$$-\alpha_7 T_t^{pt} - \alpha_8 T_t^{pt2}] = \varepsilon_t \qquad (6.34)$$

In order to reduce the autocorrelation in DPED data and also filter out the strong seasonality the following autoregressive terms are included in the model, AR(1), AR(2), AR(7), AR(14) and AR(21). Table 6.1 shows a summary of the estimates of the variables of Reg-SARIMA model 1 together with the p-values in parentheses. There is a significant reduction in electricity demand during holidays as evidenced by the coefficient of the dummy variable H_t in Table 6.1. After considering several Reg-SARIMA 1 models the best model has a RMSE of 568.41, MAE of 415.14 and a MAPE of 1.42%.

Substituting the coefficients of the parameters in equation (6.34) we get

Table 6.1: Parameter estimates of the Reg-SARIMA Model 1.

Par	c	ϕ_1	ϕ_2	Φ_1	Φ_2	Φ_3	γ
Coef	0.0749	0.8396	-0.0177	-0.5293	-0.2894	-0.1562	-0.0134
	(0.0000)	(0.0000)	(0.4772)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Par	κ	ρ	τ_2	τ_3	τ_4	τ_5	τ_6
Coef	-0.0323	-0.0054	-0.0017	-0.0025	-0.0019	-0.0020	-0.0030
	(0.0000)	(0.0190)	(0.0126)	(0.0039)	(0.0444)	(0.0343)	(0.0008)
Par	τ_7	α_2	α_3	α_7	α_8		
Coef	-0.0019	3.08E-05	-0.0017	-0.0042	0.00014		
	(0.0058)	(0.0000)	(0.0000)	(0.0074)	(0.0007)		

$$\begin{aligned}\log\hat{x}_t = 0.0749 &- 0.0017D_{2t} - 0.0025D_{3t} - 0.0019D_{4t} - 0.0020D_{5t} - 0.0030D_{6t} \\ &- 0.0019D_{7t} - 0.0134H_{t-1} - 0.0323H_t - 0.0054H_{t+1} + 0.00003\bar{x}_t \\ &- 0.0017T_t^{+} - 0.0042T_t^{pt} + 0.00014T_t^{pt2} - 0.8396\hat{x}_{t-1} - 0.0177\hat{x}_{t-2} \\ &- 0.5293\hat{x}_{t-7} - 0.2894\hat{x}_{t-14} - 0.1562\hat{x}_{t-21}\end{aligned} \tag{6.35}$$

The coefficient for the holiday variable indicates that demand during holidays decreases significantly compared to a day before a holiday. The coefficients for the average daily electricity demand over the past seven days and peak temperature squared are both positive and very small showing that their effect

on DPED is minimal.

6.5.2 Reg-SARIMA model 2

In Reg-SARIMA model 2 we model the effect of temperature on DPED using Heating Degree-Days (HDD_t) and Cooling Degree-Days (CDD_t). The functions for HDD_t and CDD_t are given as follows:

$$\text{HDD}_t = \max(T_r - \text{ADT}, 0) \tag{6.36}$$

$$\text{CDD}_t = \max(\text{ADT} - T_r, 0) \tag{6.37}$$

where ADT is the average daily temperature at time t and T_r is the reference temperature. The reference temperature used is 18^0C discussed in Chapter 3. After considering several Reg-SARIMA 2 models the best model is given in equation (6.38).

$$\begin{aligned}\log(x_t) = c + b_1\text{HDD}_t + b_2\text{CDD}_t + b_3\text{HDD}_{t-1} + b_4\text{CDD}_{t-1} + b_5\text{HDD}_{t-2}\\ +b_6\text{CDD}_{t-2} + \tau_2 D_{2t} + \tau_3 D_{3t} + \tau_4 D_{4t} + \tau_5 D_{5t} + \tau_6 D_{6t} + \tau_7 D_{7t} + \gamma H_{t-1}\\ +\kappa H_t + \rho H_{t+1} + \alpha_2 \bar{x}_t + \lambda_7 M_{7t}\\ +\lambda_{10} M_{10t} + z_t\end{aligned} \tag{6.38}$$

where HDD_{t-1}, CDD_{t-1}, HDD_{t-2} and CDD_{t-2} represent the lagged effects of the heating degree-days and cooling degree-days. The error term z_t follows a SARIMA model as given in equation (6.39).

$$(1 - \phi_1 B - \phi_2 B^2)(1 - \Phi_1 B^7 - \Phi_2 B^{14} - \Phi_3 B^{21})(1 - B^7) z_t = \varepsilon_t \tag{6.39}$$

and ε_t is a white noise process. Combining equations (6.38) and (6.39 we get equation (6.40).

$$(1-\phi_1 B-\phi_2 B^2)(1-\Phi_1 B^7-\Phi_2 B^{14}-\Phi_3 B^{21})(1-B^7)[\mathbf{log}(x_t)-c$$

$$-b_1\mathbf{HDD}_t-b_2\mathbf{CDD}_t-b_3\mathbf{HDD}_{t-1}-b_4\mathbf{CDD}_{t-1}-b_5\mathbf{HDD}_{t-2}$$

$$-b_6 CDD_{t-2}+\tau_2 D_{2t}-\tau_3 D_{3t}-\tau_4 D_{4t}-\tau_5 D_{5t}-\tau_6 D_{6t}$$

$$-\tau_7 D_{7t}-\gamma H_{t-1}-\kappa H_t-\rho H_{t+1}-\alpha_2\bar{x}_t-\lambda_7 M_{7t}$$

$$-\lambda_{10} M_{10t}]=\varepsilon_t \qquad (6.40)$$

A summary of the estimates of the variables of Reg-SARIMA model 2 together with the p-values in parentheses are given in Table 6.2. After substituting the

Table 6.2: Parameter estimates of the Reg-SARIMA Model 2.

Par	c	ϕ_1	ϕ_2	Φ_1	Φ_2	Φ_3	γ
Coef	-0.0012	0.8425	-0.0204	-0.5291	-0.2897	-0.1540	-0.0139
	(0.6128)	(0.0000)	(0.5416)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Par	κ	ρ	τ_2	τ_3	τ_4	τ_5	τ_6
Coef	-0.0326	-0.0047	-0.0017	-0.0021	-0.0018	-0.0020	-0.0030
	(0.0000)	(0.1479)	(0.0170)	(0.0107)	(0.0560)	(0.0297)	(0.0006)
Par	τ_7	α_2	b_1	b_2	b_3	b_4	b_5
Coef	-0.0020	3.18E-05	0.0022	0.0012	0.0006	-0.0022	-0.0006
	(0.0072)	(0.0001)	(0.0004)	(0.0877)	(0.5279)	(0.0190)	(0.2218)
Par	b_6	λ_7	λ_{10}				
Coef	0.0011	-0.0121	0.0189				
	(0.0129)	(0.0261)	(0.0011)				

coefficients of the parameters we get

$$\begin{aligned}\log\hat{x}_t = &-0.0012 + 0.0022\mathbf{HDD}_t + 0.0012\mathbf{CDD}_t + 0.0006\mathbf{HDD}_{t-1} \\ &-0.0022\mathbf{CDD}_{t-1} - 0.0006\mathbf{HDD}_{t-2} + 0.0011\mathbf{CDD}_{t-2} - 0.0017D_{2t} \\ &-0.0021D_{3t} - 0.0018D_{4t} - 0.0020D_{5t} - 0.0030D_{6t} - 0.0020D_{7t} \\ &-0.0139H_{t-1} - 0.0326H_t - 0.0047H_{t+1} + 0.0000318\bar{x}_t \\ &-0.0121M_{7t} + 0.0189M_{10t} + 0.8425\hat{x}_{t-1} - 0.0204\hat{x}_{t-2} \\ &-0.5291\hat{x}_{t-7} - 0.2897\hat{x}_{t-14} - 0.1540\hat{x}_{t-21} \qquad (6.41)\end{aligned}$$

After considering several Reg-SARIMA 2 models the best model has a RMSE of 568.27, MAE of 415.31 and a MAPE of 1.42%.

The monthly electricity demand profile given in Figure 6.1 shows that the highest demand is experienced in July which is a winter month in the Southern Hemisphere. In summer the highest demand is experienced in October as shown in Figure 6.1. These two months, July and October are included as dummy variables in Reg-SARIMA model 2 and the p-values of these variables are highly significant. The coefficient for the heating degree days is larger than that of the cooling degree days indicating that demand for electricity is more sensitive to cold temperatures. The value of the coefficients of HDD_{t-1} and CDD_{t-2} are positive showing the significance of the effect of previous day's and two days' temperature on daily peak electricity demand. The coefficient of HDD_t is significantly higher than that of CDD_t, again confirming that the

effect of cold temperature on DPED in South Africa is significantly higher than that of hot temperature. This is consistent with findings discussed in Chapter 3 that in South Africa electricity demand is more sensitive to cold temperatures compared to hot temperatures. The dummy variables representing seasonality including those representing holidays and a day before a holiday are all significant.

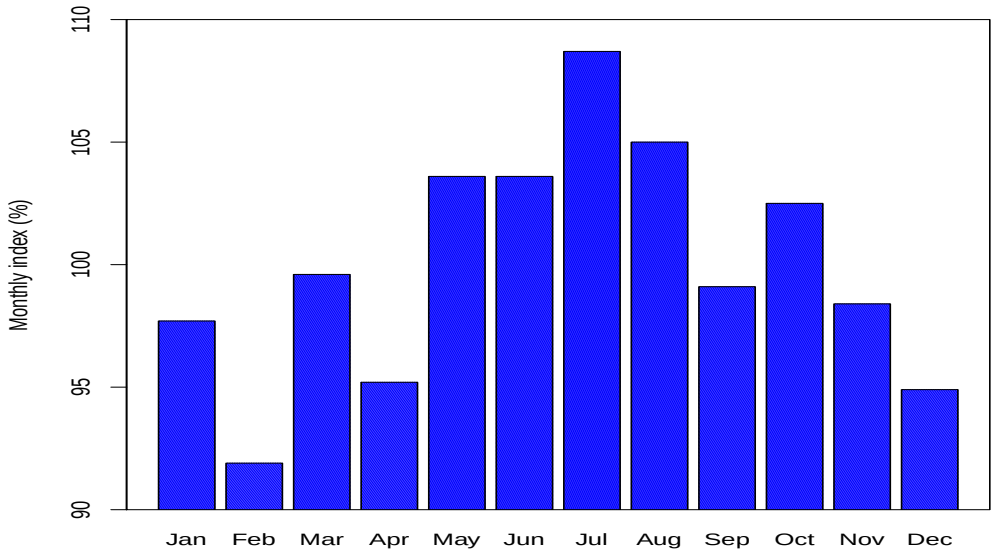

Figure 6.1: Monthly seasonal index plot of DPED. The index for each month is calculated by finding the average DPED for each month and then divide by the overall average.

6.5.3 Reg-SARIMA model 3

In Reg-SARIMA model 3 we model the effect of temperature on DPED using regression splines and the knots are selected using a Multivariate Adaptive Regression Splines (MARS) algorithm developed by Friedman (1991). Reg-

SARIMA model 3 is given in equation (6.42).

$$\begin{aligned}\log(x_t) = c + \tau_2 D_{2t} + \tau_3 D_{3t} + \tau_4 D_{4t} + \tau_5 D_{5t} + \tau_6 D_{6t} + \tau_7 D_{7t} + \gamma H_{t-1} \\ + \kappa H_t + \rho H_{t+1} + \lambda_7 M_{7t} + \lambda_{10} M_{10t} + \alpha_1 x_t^- + \alpha_2 \bar{x}_t \\ + a_1 \max(T_t^+ - k_1, 0) + a_2 \max(k_1 - T_t^+, 0) + a_3 \max(T_t^- - k_2, 0) \\ + a_4 \max(k_2 - T_t^-, 0) + a_5 \max(T_t^{pt} - k_3, 0) + a_6 \max(k_3 - T_t^{pt}, 0) \\ + a_7 \max(\bar{T}_t - k_4, 0) + a_8 \max(k_4 - \bar{T}_t, 0) + z_t \end{aligned} \quad (6.42)$$

where z_t is an error term which is assumed to follow a SARIMA model as shown in equation (6.43).

$$\phi_p(B)\Phi_P(B^s)\nabla^d\nabla_s^D z_t = c + \theta_q(B)\Theta_Q(B^s)\varepsilon_t \quad (6.43)$$

The MARS models' results are given in appendix 6.1 at the end of the chapter. For each temperature-demand relationship two MARS models are given with different knots. After considering different knot combinations based on the predictive power, the following model is selected.

$$\log(x_t) = c + \tau_2 D_{2t} + \tau_3 D_{3t} + \tau_4 D_{4t} + \tau_5 D_{5t} + \tau_6 D_{6t} + \tau_7 D_{7t} + \gamma H_{t-1}$$

$$+\kappa H_t + \rho H_{t+1} + \alpha_1 x_t^- + \alpha_2 \bar{x}_t + a_1 \max(T_t^+ - 10.81, 0)$$

$$+a_2 \max(10.81 - T_t^+, 0) + a_3 \max(T_t^- - (-0.02), 0) + a_4 \max(-0.2 - T_t^-, 0)$$

$$+a_5 \max(T_t^{pt} - 7.12, 0) + a_8 \max(7.16 - \bar{T}_t, 0) + a_6 \max(7.12 - T_t^{pt}, 0)$$

$$+a_7 \max(\bar{T}_t - 7.16, 0) + \lambda_7 M_{7t} + \lambda_{10} M_{10t} + z_t \quad (6.44)$$

After considering several Reg-SARIMA 3 models the best model has a RMSE of 607.18, MAE of 443.71 and a MAPE of 1.42%. The model is given in equation (6.45) and the parameter estimates together with their $p-$values in parentheses are given in Table 6.3.

$$(1 - \phi_1 B - \phi_2 B^2)(1 - \Phi_1 B^7 - \Phi_2 B^{14} - \Phi_3 B^{21} - \Phi_4 B^{28})(1 - B^7)[\log(x_t)$$

$$-c - \tau_2 D_{2t} - \tau_3 D_{3t} - \tau_4 D_{4t} - \tau_5 D_{5t} - \tau_6 D_{6t} - \tau_7 D_{7t} - \gamma H_{t-1} - \kappa H_t$$

$$-\rho H_{t+1} - \alpha_2 \bar{x}_t - a_1 \max(T_t^+ - 10.81, 0) - a_5 \max(T_t^{pt} - 7.12, 0)$$

$$-\lambda_7 M_{7t} - \lambda_{10} M_{10t}] = \varepsilon_t \quad (6.45)$$

The coefficient for holiday shows that demand during holidays decreases significantly compared to a day before and a day after a holiday. The coefficient for the average daily demand over the past seven days is very small indicating

Table 6.3: Parameter estimates of the Reg-SARIMA Model 3.

Par	c	ϕ_1	ϕ_2	Φ_1	Φ_2	Φ_3	Φ_4
Coef	0.0149	0.8550	-0.0191	-0.5627	-0.3492	-0.2163	-0.1129
	(0.0000)	(0.0000)	(0.2664)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Par	γ	κ	ρ	τ_2	τ_3	τ_4	τ_5
Coef	-0.0107	-0.0282	-0.0029	-0.0013	-0.0019	-0.0015	-0.0017
	(0.0000)	(0.0000)	(0.0494)	(0.0012)	(0.0003)	(0.0079)	(0.0032)
Par	τ_6	τ_7	α_2	a_1	a_5	λ_7	λ_{10}
Coef	-0.0026	-0.0018	0.000024	-0.0015	0.0008	-0.0101	0.0216
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0095)	(0.0088)	(0.0000)

that its effect is minimal on DPED.

6.6 Comparative Analysis

MAPE, MAE and RMSE are normally used for comparing models in short-term demand forecasting up to seven days ahead (Munoz *et al.*, 2010; Fan and Hyndman, 2012). Table 6.4 shows a comparative analysis of the Reg-SARIMA models (training period). The Reg-SARIMA 3 model has the least MAPE, showing

Table 6.4: In-sample evaluation of the models.

	Forecasting models		
Performance criteria	Reg-SARIMA 1	Reg-SARIMA 2	Reg-SARIMA 3
MAPE	1.42	1.42	1.42
RMSE	568.95	568.27	607.18
MAE	416.07	415.31	443.71

that it is the best fitting model. Table 6.5 summarizes the error measures for out of sample evaluation of the models. From Table 6.5 Reg-SARIMA model 3

Table 6.5: Out-of-sample evaluation of the models.

	Forecasting models		
Performance criteria	Reg-SARIMA 1	Reg-SARIMA 2	Reg-SARIMA 3
MAPE	1.83	1.81	1.78
RMSE	818.24	815.80	812.41
MAE	610.94	605.69	594.33

has the least MAPE among the three Reg-SARIMA models. The graphical plot of the out of sample forecasts for Reg-SARIMA models 1, 2 and 3 respectively for the first 27 days of August 2010 are given in Figure 6.2. A summary of the out of sample accuracy measures (MAPE and MAE) for Reg-SARIMA models 1, 2 and 3 respectively are given in Table 6.12 in appendix 6.2 at the end of the chapter.

The probability densities of DPED and the forecasted values for the first 27

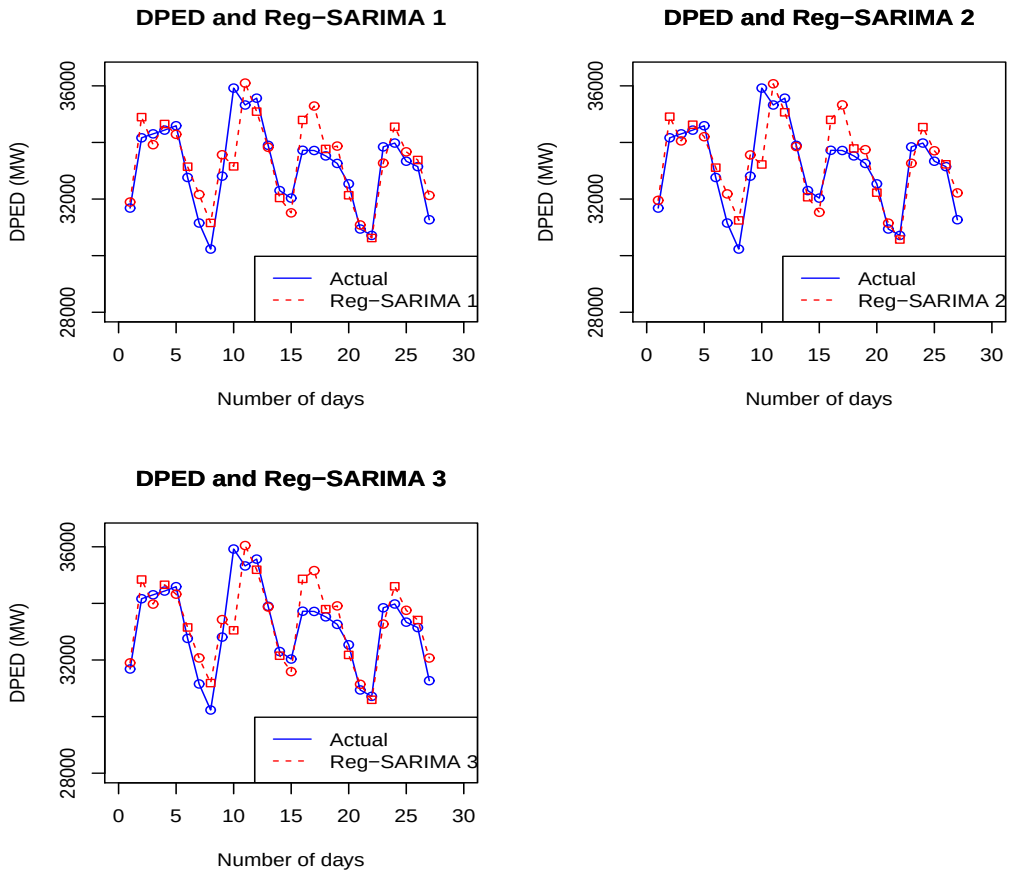

Figure 6.2: Graphical plot of (a) Top left panel: Forecasts (dashed line) using Reg-SARIMA model 1 and actual peak demand (solid line) and (b) Top right panel: Forecasts (dashed line) using Reg-SARIMA model 2 and actual peak demand (solid line) (c) Bottom left panel: Forecasts (dashed line) using Reg-SARIMA model 3 and actual peak demand (solid line), for the first 27 days of August 2010.

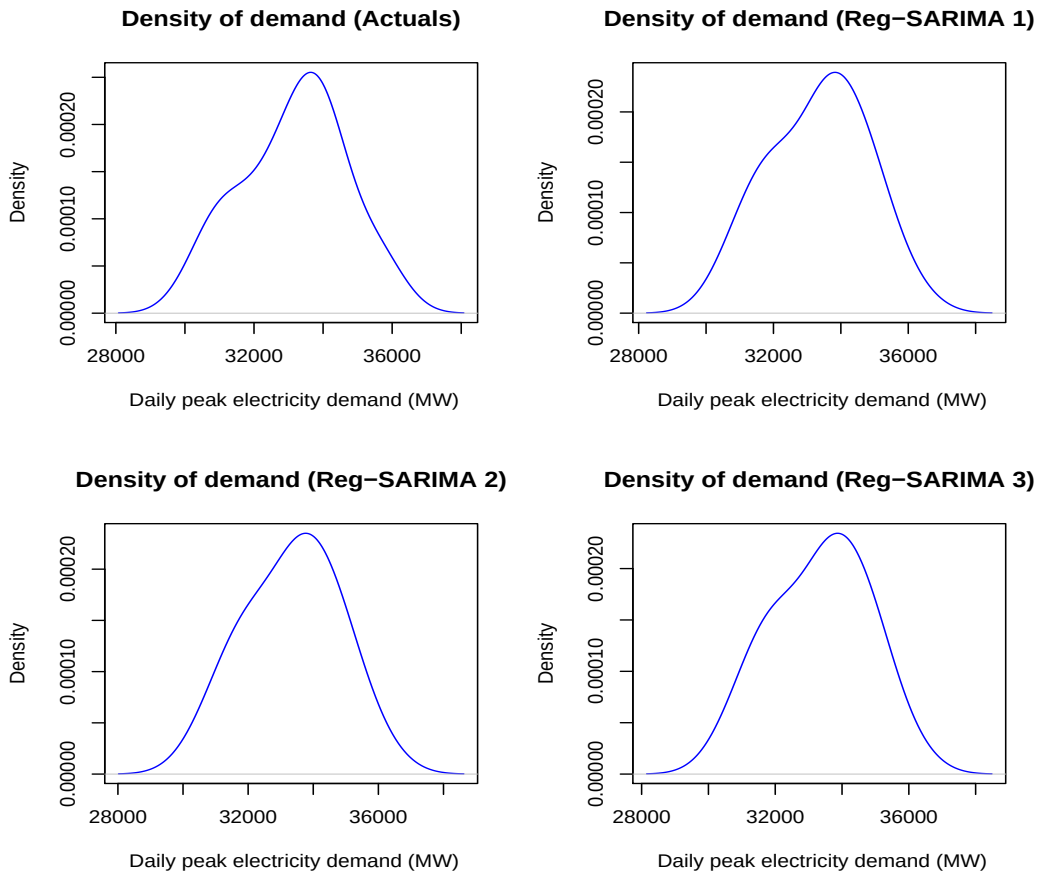

Figure 6.3: Probability densities of actual (Top left panel) and predicted DPED for the first 27 days of August 2010 using: (a) Reg-SARIMA model 1 (Top right panel) (b) Reg-SARIMA model 2 (Bottom left panel) and (c) Reg-SARIMA model 3 (Bottom right panel).

days of August 2010 for Reg-SARIMA models 1, 2 and respectively are shown in Figure 6.3. The actual demand values for both models fall within the region predicted from the forecast distribution.

6.7 Residual distribution of Reg-SARIMA model 3

Table 6.6 shows that the residual series data are non-normal. The Jarque-Bera test is carried out to check whether the skewness and kurtosis are consistent

with a normal distribution. The residual demand series is very volatile as

Table 6.6: Descriptive statistics of the residual return series.

Mean	Median	Max[a]	Min[b]	Std Dev[c]	Skew[d]	Kurt[e]	J-B[f]
9.07E-11	-6.82E-05	0.103	-0.091	0.019	0.054	6.041	625.7(0.00)

[a]Max represents maximum
[b]Min represents minimum
[c]Std Dev represents Standard deviation
[d]Skew denotes skewness
[e]Kurt represents kurtosis
[f]J-B denotes Jarque-Bera

shown in Figure 6.4, top left panel. This calls for risk analysis and modelling of under and over demand predictions for an acceptable level of risk by a power utility company. High quantiles such as the 90^{th}, 95^{th} or even 99^{th} quantiles will be used to calculate peak demand levels which correspond to acceptable levels of risk. We shall in this chapter discuss under demand risk modelling, but it should be noted that the modelling approach can be applied to over demand predictions as well. The residual demand distribution (top right panel of Figure 6.4) can be used to calculate acceptable risk levels of under-demand and over-demand predictions of DPED. We define under demand prediction as:

$$\text{positive residuals} \;=\; \text{actual demand} \;-\; \text{forecasted demand} \;>\; 0$$

and over-demand prediction as

$$\text{negative residuals} \;=\; \text{actual demand} \;-\; \text{forecasted demand} \;<\; 0$$

The bottom right panel of Figure 6.4 shows a plot of the positive residual demand data. We calculate high quantiles of the upper distribution of the residual demand series. Table 6.7 shows some of the calculated quantiles. Using the 97.5^{th} quantile as a confidence level which represents peak demand level which correspond to the acceptable level of risk we carry out a frequency analysis of those positive residuals which exceed 0.04126 (i.e. the 97.5^{th} quantile).

There are 41 observations above 0.04126 as shown in Figure 6.5.

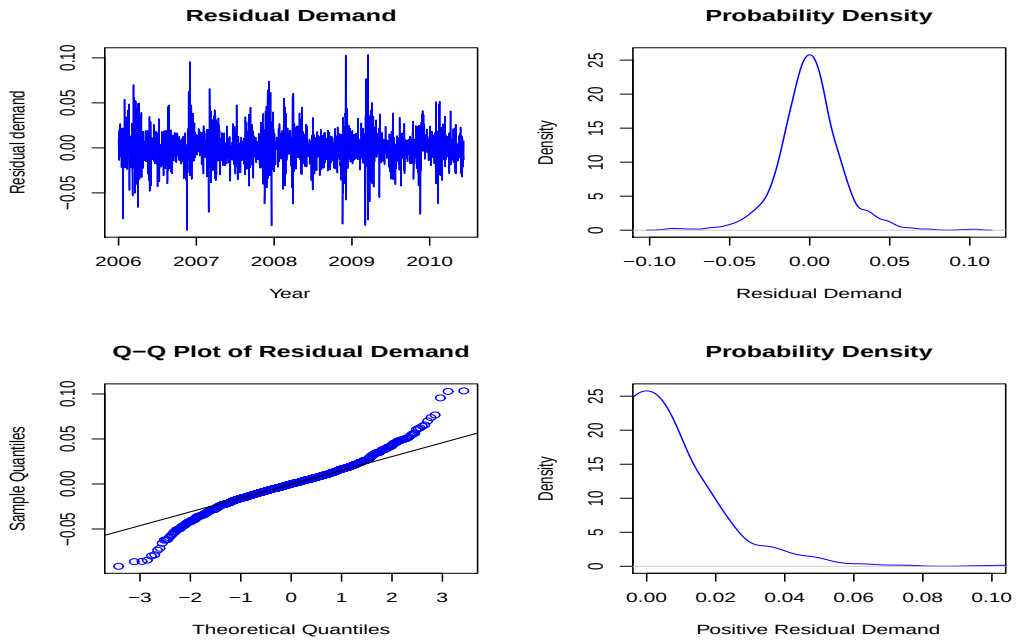

Figure 6.4: (a) Top left panel: Plot of residual demand (b) Top right panel: Probability density of residual demand (c) Bottom left panel: Normal QQ of residual demand and (d) Bottom right panel: Enlarged right tail of density of residual demand.

Table 6.7: Quantiles of residual demand.

	90^{th}	95^{th}	97.5^{th}	99^{th}	99.9^{th}
Quantiles	0.02185	0.03293	0.04126	0.05147	0.09836

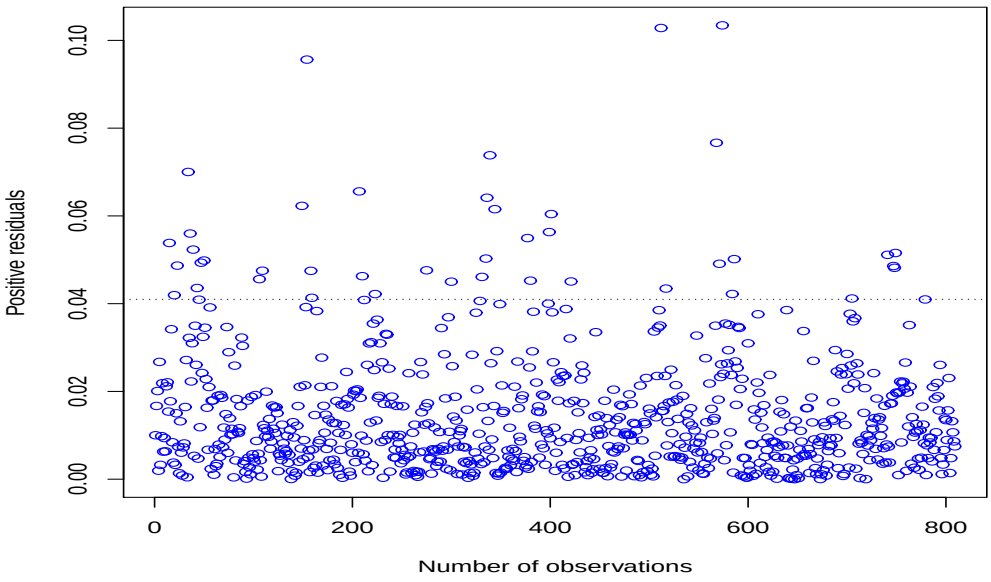

Figure 6.5: Plot of positive residuals. The dotted horizontal line represents the 97.5^{th} quantile which is 0.04126.

6.8 Volatility modelling of DPED

In the previous Sections 6.5 to 6.8, the variance of the disturbance term in each of the developed models is assumed to be constant. Causal inspection of the electricity demand time series plot shown in Figure 3.1 in Chapter 3 suggests that the series does not have a constant variance. The series exhibits phases of high demand (in winter) followed by periods of low demand (in summer). The assumption of homoskedasticity (constant variance) seems inappropriate, since the data exhibits non-constant mean and variance, and multiple seasonality corresponding to weekly and monthly periodicity. The Generalized Autoregressive Conditional Heteroskedasticity (GARCH) modeling methodology is introduced to accommodate the possibility of serial correlation in volatility. Models for volatility forecasting were first developed by Engle (1982). These models known as the Autoregressive Conditional Heteroskedasticity (ARCH) models were developed to capture the non-constant variance. ARCH models were later

extended to Generalized ARCH (GARCH) models by Bollerslev (1986).

6.8.1 Bayesian GARCH modelling with Student-t innovations

In this section we present the empirical results of the Bayesian GARCH(1,1) model with Student-t innovations of the residual return series ($r_t = x_t - f_t$) which was discussed in Sections 6.3.3. and 6.8. The marginal posterior densities of the parameters, ω, α, β and ν of the variance equation (6.46),

$$\text{Variance equation: } \sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2 \tag{6.46}$$

We cannot use the asymptotic symmetric approximation for the parameter estimators since the shapes of the histograms are asymmetric. Figure 6.7 also confirms that the Student-t distribution is a better fit than a normal distribution.

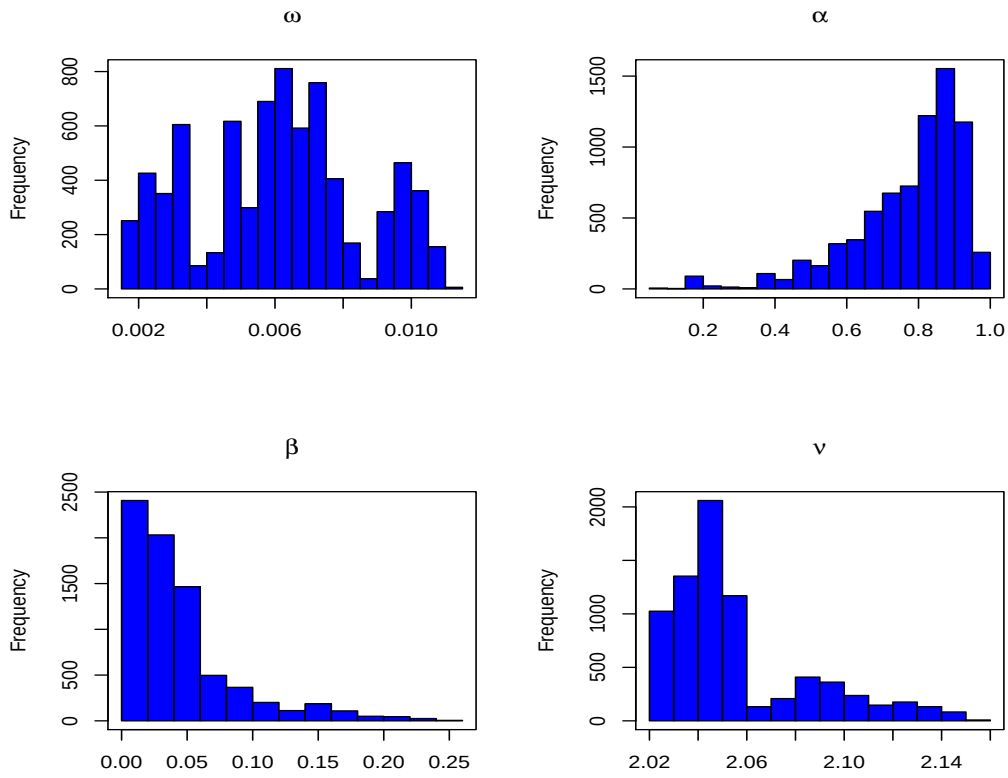

Figure 6.6: (a) Top left panel: Marginal posterior density of the parameter ω (b) Top right panel: Marginal posterior density of the parameter α (c) Bottom left panel: Marginal posterior density of the parameter β and (d) Bottom right panel: Marginal posterior density of the parameter ν.

The modes of the simulated ω's, α's, β's and ν's are calculated as 0.00633, 0.890, 0.00811 and 2.051 respectively and are considered as the estimates of ω, α, β and ν. The sum of α and β is 0.898 which is close to 1. This indicates that past shocks and past variances will have a longer impact on the future conditional variance. The posterior density of the persistence of the squared variance process is given in Figure 6.8.

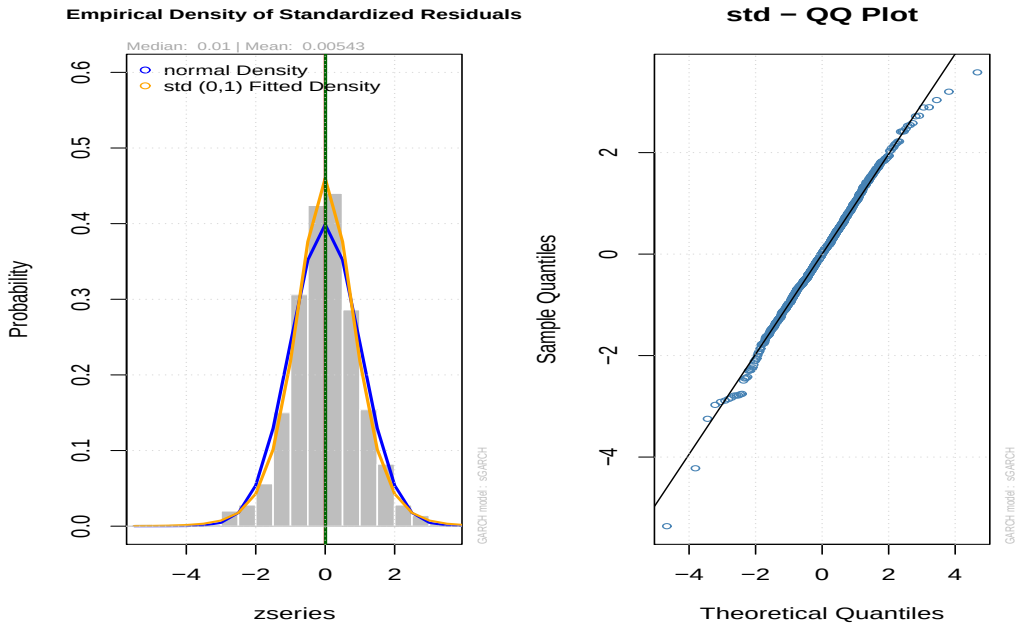

Figure 6.7: (a) Left panel: Return series with Student-t and normal distributions (b) Right panel: Student-t QQ plot.

The Bayesian parameter estimates with Student-t innovations are given in Table 6.8.

Table 6.8: Bayesian parameter estimates.

Parameter	Bayesian parameter estimates with Student-t innovations
ω	0.00633
α	0.89000
β	0.00811
persistence: α+β	0.89800

$$\sigma_t^2 = 0.00633 + 0.890\varepsilon_{t-1}^2 + 0.00811\sigma_{t-1}^2 \tag{6.47}$$

The Bayes parameter estimate of $\alpha = 0.89$ in equation (6.47) shows that the degree of impact of news about volatility of DPED from the previous period is very high. This is true with South African load series data as bad news such as an announcement of a cold front increases significantly the consumption of electricity. This is consistent with previous findings which show that demand

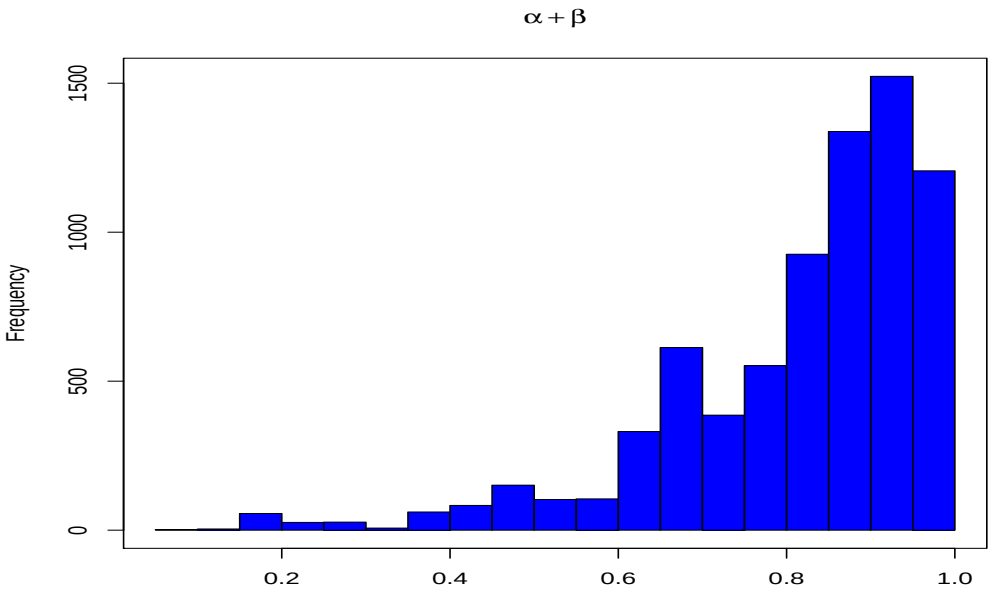

Figure 6.8: Posterior density of the persistence in the squared variance process.

for electricity in South Africa increases significantly for temperature values below the reference temperature of 18^0C and increases slightly above this temperature. The parameter $\beta = 0.00811$ in equation (6.47) shows that the impact of the last period's fluctuation (forecast variance) is very low. Some quantiles for each variable for Bayes GARCH(1,1) with Student-t innovations are given in Table 6.9.

Table 6.9: Quantiles for each variable.

Parameters	2.5%	25%	50%	75%	97.5%
ω	0.00196	0.00467	0.00624	0.00753	0.01047
α	0.36200	0.692	0.822	0.890	0.952
β	0.000823	0.0151	0.0316	0.0540	0.164
ν	2.026	2.037	2.045	2.061	2.132

6.8.2 Empirical results: Reg-SARIMA-TGARCH

In this section we present the empirical results of the Reg-SARIMA-TGARCH model. The DPED return series is given as

$$r_t = \log(x_t) - \log(x_{t-1}) = \log\left(\frac{x_t}{x_{t-1}}\right) \tag{6.48}$$

This represents logarithmic daily changes in DPED. Positive daily changes indicate an increase in demand which poses challenges to system operators who have to ensure that there is an adequate supply of electricity to meet demand.

Several variants of the GARCH type models are fitted to the DPED return series. The best fitting model is found to be the Threshold GARCH (TGARCH) model based on the Akaike Information Criterion (AIC). The Reg-SARIMA-TGARCH model is given as

$$\begin{aligned} \text{Mean equation: } r_t &= c + \tau_5 D_{5t} + \tau_6 D_{6t} + \tau_7 D_{7t} + \gamma H_{t-1} + \kappa H_t - \rho H_{t+1} \\ &\quad + \phi_1 r_{t-1} + \phi_2 r_{t-2} + \phi_3 r_{t-14} + \phi_4 r_{t-28} + \phi_5 r_{t-175} \\ &\quad + \theta_1 \varepsilon_{t-7} \theta_2 \varepsilon_{t-42} + \varepsilon_t \end{aligned} \tag{6.49}$$

where the error term ε_t is given as

$$\begin{aligned} \varepsilon_t &= \nu_t \sigma_t, \\ \nu_t \sim i.i.d & \qquad \text{with } E(\nu_t) = 0, \text{ Var } (\nu_t) = 1 \end{aligned} \tag{6.50}$$

and the variance model is a TGARCH introduced independently by Zakoian (1990) and Glosten *et al.* (1993). The variance equation is a TGARCH(1,1) model given as

$$\text{Variance equation } \sigma_t^2 = a_0 + a_1\varepsilon_{t-1}^2 + \gamma\varepsilon_{t-1}^2 d_{t-1} + b_1\sigma_{t-1}^2 \tag{6.51}$$

where

$$d_t = \begin{cases} 1 & \text{if } \varepsilon_t < 0 \\ 0 & \text{otherwise} \end{cases} \tag{6.52}$$

where the parameters are as defined in Sections 6.2, 6.3 and 6.6, and γ represents the threshold. When $\gamma = 0$ the TGARCH reduces to the GARCH model. In financial time series the effects of good and bad news are measured by ε_t in the variance equation. If $\varepsilon_t > 0$ we have good news meaning that the volatility does not increase significantly and we have bad news if $\varepsilon_t < 0$ resulting in an increase in volatility. The impacts of good and bad news are measured by a and $a + \gamma$ respectively. For $\gamma \neq 0$ the news impact is asymmetric and for $\gamma > 0$ the leverage effect exists meaning that bad news increases volatility.

In load series bad news could be an announcement by the meteorologist of extreme weather conditions (e.g a heat wave or extreme cold temperatures). It should be noted that positive and negative DPED return series of the same size may have different impacts on volatility. We expect a positive return (bad news) to have a larger impact than a negative return (good news).

Table 6.10 shows a summary of the estimates of the parameters of the Reg-SARIMA-TGARCH model developed along with their p-values shown in parentheses. The model parameters are then estimated using the maximum likelihood method. The estimates are obtained by the Berndt *et al.* (1974) algorithm using numerical derivatives. The Engle's LM test is carried out and the

null hypothesis that there are no ARCH effects in the residuals is accepted at the 5% level. From Table 6.11 the coefficients of the dummy variables representing holiday and day before holiday are negative meaning that DPED decreases during these days. This is due to the fact that most companies will be closed during these days. Other Reg-SARIMA-TGARCH models are considered but we are not able to improve further on the heteroskedasticity problem revealed by the squared standardized residuals since we are dealing with high-frequency data. The value of γ which is 0.064 suggests that the leverage effect exists in that bad news will result in an increase in volatility in DPED return series. The impact of good and bad news is asymmetric since $\gamma \neq 0$. For example an announcement of some extreme low temperatures (bad news) due to a cold front by the meteorologist will significantly increase the demand for electricity. The impact of good news will be measured by $a_1 = 0.153$ and bad news by $a_1 + \gamma = 0.153 + 0.064 = 0.217$. The persistence in volatility shocks is $a_1 + b_1 = 0.833$. This suggests that the news impact tends to last for a fairly long period of time.

Table 6.10: Parameter estimates of the Reg-SARIMA-TGARCH Model.

Mean Equation				
Parameter	c	H_{t-1}	H_t	H_{t+1}
Coefficient	0.016(0.000)	-0.029(0.000)	-0.038(0.000)	0.054(0.000)
Parameter	Friday	Saturday	Sunday	ϕ_1
Coefficient	-0.084(0.000)	-0.048(0.000)	0.013(0.107)	-0.134(0.000)
Parameter	ϕ_2	ϕ_3	ϕ_4	ϕ_5
Coefficient	-0.121(0.000)	0.367(0.000)	0.288(0.000)	0.181(0.000)
Parameter	θ_1	θ_2		
Coefficient	0.169(0.000)	0.056(0.015)		
Variance Equation				
Parameter	a_0	a_1	γ	b_1
Coefficient	4.37E-05(0.000)	0.153(0.001)	0.064(0.317)	0.680(0.000)

6.9 Comparative analysis of the models

Table 6.11 shows the number of under demand predictions above some acceptable levels of 500MW, 750MW and 1000MW. Approximate mean absolute percentage errors are calculated for each of the under demand prediction levels of 500MW, 750MW, 1000MW and are found to be 1.60%, 2.38% and 3.18% respectively. For an acceptable under demand estimation of 500MW Reg-SARIMA 3 model has a total of 190 violations (observations above 500MW) compared to the Reg-SARIMA-TGARCH model with 210 violations for the period 2006 to 2009. The two models have the same number of 101 violations for an acceptable level of 750MW, while the Reg-SARIMA-TGARCH model has fewer violations for a 1000MW acceptable level of under demand predictions. Overall Reg-SARIMA 3 has a total of 349 violations compared to 358 violations for the Reg-SARIMA-TGARCH model for the period 2006 to 2009. For the year 2007 Reg-SARIMA 3 model has more violations compared to Reg-SARIMA-TGARCH model. This analysis is important for risk assessment by system operators and decision makers in Eskom, South Africa's power utility company. Both the Reg-SARIMA-TGARCH and Reg-SARIMA models are simple to

Table 6.11: DPED under demand risk estimation for in sample predictions.

Under demand estimation of 500MW	2006	2007	2008	2009	Total
Reg-SARIMA 3 model	15	62	52	61	190
Reg-SARIMA-TGARCH model	23	56	70	61	210
Under demand estimation of 750MW	2006	2007	2008	2009	Total
Reg-SARIMA 3 model	7	35	29	30	101
Reg-SARIMA-TGARCH model	14	25	31	31	101
Under demand estimation of 1000MW	2006	2007	2008	2009	Total
Reg-SARIMA 3 model	5	21	17	15	58
Reg-SARIMA-TGARCH model	2	15	15	15	47

implement, reliable and provide information about the importance of each predictor variable. Using the Reg-SARIMA-TGARCH and Reg-SARIMA models give results which are relatively robust. Plots of standardized residuals from

the Reg-SARIMA-TGARCH and the Reg-SARIMA 3 models for the year 2008 are given in Figure 6.9. The dashed horizontal line denotes the threshold for an acceptable under demand prediction of 1000MW. Using the Reg-SARIMA-

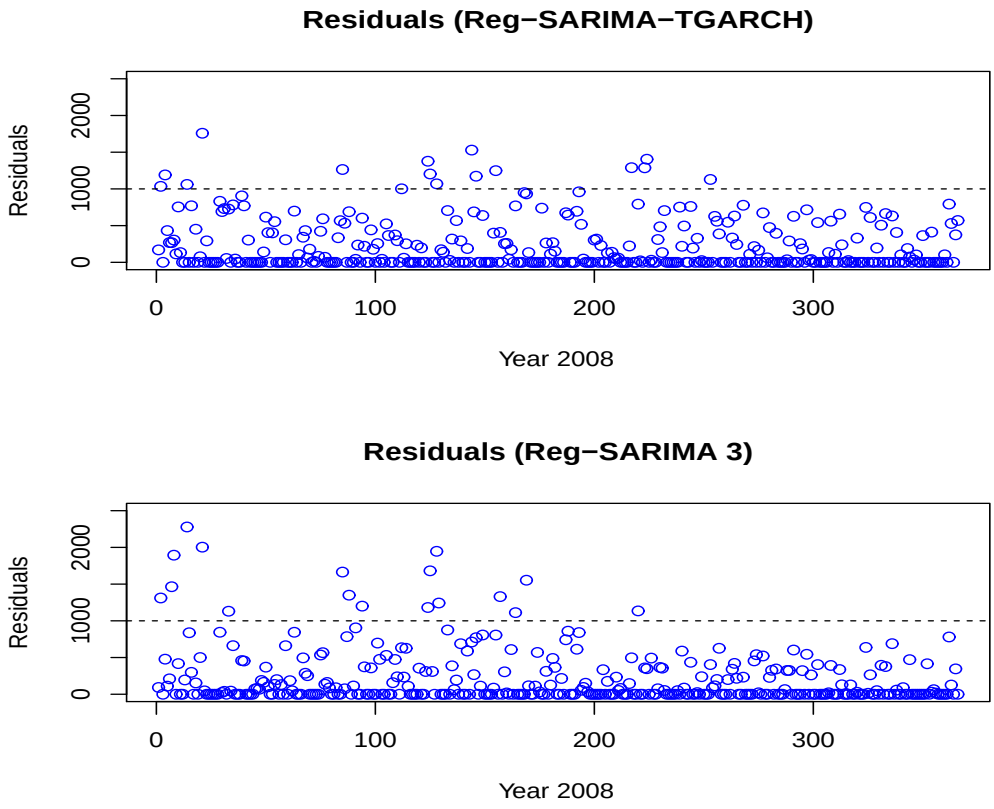

Figure 6.9: Plot of standardized residuals for the year 2008 (a) Upper panel: The Reg-SARIMA-TGARCH model and (b) Lower panel: The Reg-SARIMA 3 model.

TGARCH model there are 15 observations above the threshold of 1000MW and all these exceedances (observations above the threshold) are below 2000MW. There are 17 exceedances above 1000MW threshold using the Reg-SARIMA 3 model with three observations above 2000MW. The Reg-SARIMA-TGARCH seems to capture well the volatility in the residuals.

6.10 Concluding remarks

This chapter has investigated some hybrid models for daily peak electricity demand forecasting. Three regression-seasonal autoregressive integrated moving average (Reg-SARIMA) models were developed. Empirical results show that the Reg-SARIMA model in which the effect of temperature variables are included through regression splines produces the best forecast accuracy. The Reg-SARIMA model is simple to implement, reliable and provides information about the importance of each predictor variable. The results from using a regression-SARIMA model are relatively robust. The model assumes homoskedasticity.

The second part of the chapter discussed volatility forecasting models including volatility modelling of the residual return series data from the Reg-SARIMA model 3. Empirical results of the Bayesian GARCH(1,1) model with Student-t innovations of the residual return series of DPED are presented and compared with results from Bayesian GARCH(1,1) with normal innovations. The Bayesian GARCH(1,1) model with Student-t innovations seem to provide a better description of the volatility of the DPED residual return series data. A Reg-SARIMA-TGARCH model is developed and used in DPED under demand risk estimation for in-sample predictions. A comparative analysis is done with the Reg-SARIMA model which modelled temperature through regression splines. Although both models produced results which were comparable the Reg-SARIMA-TGARCH seem to capture well the volatility in the residuals during the year 2008 when South Africa experienced countrywide blackouts, load shedding and a global recession.

Daily peak electricity demand is not only influenced by the means (averages) but by the tails of probability distributions as well. An analysis of extreme peak load demand forecasting is very important to system operators and de-

cision makers in power utility companies who have to balance supply and demand of electrical energy to ensure grid stability, reliable supply schedules and optimal distribution of electricity. At times there is a need to depart from the average thinking and exploit information provided by the extremes (tails of distributions). Modelling of tail behaviour of DPED through the Extreme Value Theory (EVT) approach is then covered in Chapters 7, 8, 9 and 10.

Appendix 6.1: MARS models for minimum, maximum, average and peak temperatures

$$r(x) = \omega_0 + \sum_{m=1}^{M} \omega_m \prod_{k=1}^{K_m} \left[s_{km}(x_{v(k,m)} - t_{km}) \right] \tag{6.53}$$

$$r(x,k,m) = \begin{cases} x_{v(k,m)} - t_{km}, & \text{if } x_{v(k,m)} > t_{km} \\ t_{km} - x_{v(k,m)}, & \text{if } x_{v(k,m)} < t_{km} \end{cases} \tag{6.54}$$

1. **DPED and Average Daily Temperature**
   (a) DPED = 33287.3 - 380.265 * max(0, $\bar{T}_t$ - 7.16)
   (b) DPED = 27700 + 443 * max(0, 21 - $\bar{T}_t$)

2. **DPED and Minimum Temperature**
   (a) DPED = 33608.8 - 339.079 *max(0, T_t^- - (-0.20))
   (b) DPED = 27600 + 379 * max(0, 16.6 - T_t^-)

3. **DPED and Maximum Temperature**
   (a) DPED = 33646.8 - 345.716 * max(0, T_t^+ - 10.81)
   (b) DPED = 28000 + 464 * max(0, 25.5 - T_t^+)

4. **DPED and Peak Temperature**
   (a) DPED = 33125.5 - 353.961 * max(0, T_t^{pt} - 7.12)
   (b) DPED =27900 + 455 * max(0,20.5 - T_t^{pt})

Appendix 6.2: Out of sample accuracy measures for Reg-SARIMA models 1, 2 and 3 for the period 1-27 August 2010

Table 6.12: Results for out-of-sample test for 1-27 August 2010 (Unit of index: MAPE(%); MAE(MW)).

	Reg-SARIMA model 1		Reg-SARIMA model 2		Reg-SARIMA model 3	
DATE	MAPE	MAE	MAPE	MAE	MAPE	MAE
08/01/10	0.67	212.99	0.86	272.89	0.68	216.86
08/02/10	2.14	730.02	2.18	746.27	1.99	679.24
08/03/10	1.12	382.51	0.73	249.28	0.96	327.72
08/04/10	0.63	216.01	0.55	190.43	0.64	221.8
08/05/10	0.88	303.24	1.12	387.5	0.76	264.02
08/06/10	1.17	381.85	1.06	347.3	1.18	387.65
08/07/10	3.24	1008.55	3.3	1028.78	2.95	919.5
08/08/10	3.06	923.89	3.35	1013.95	3.15	951.64
08/09/10	2.3	755.43	2.29	751.14	1.88	618.1
08/10/10	7.7	2765.82	7.51	2699.41	7.99	2870.93
08/11/10	2.19	774.46	2.12	750.06	2.04	721.23
08/12/10	1.34	475.3	1.4	499.13	1.06	376.17
08/13/10	0.2	68.59	0.13	42.49	0.08	26.19
08/14/10	0.81	260.67	0.73	234.36	0.46	147.38
08/15/10	1.63	520.98	1.56	501.13	1.39	444.03
08/16/10	3.18	1071.32	3.19	1075.09	3.39	1143.23
08/17/10	4.66	1570.91	4.78	1610.74	4.28	1442.17
08/18/10	0.73	246.27	0.78	262.51	0.81	271.29
08/19/10	1.84	613.26	1.46	484.94	1.96	652.71
08/20/10	1.24	402.67	0.95	310.5	1.1	356.95
08/21/10	0.47	144.16	0.67	207.96	0.65	199.69
08/22/10	0.3	91.88	0.47	144.24	0.39	120.96
08/23/10	1.7	574.27	1.72	580.78	1.7	574.44
08/24/10	1.69	575.55	1.66	562.63	1.84	625.21
08/25/10	0.99	331.29	1.11	369.16	1.26	419.08
08/26/10	0.71	234.9	0.25	82.47	0.81	267.52
08/27/10	2.75	858.66	3.03	948.61	2.56	801.29
Average	1.83	610.94	1.81	605.69	1.78	594.33

Chapter 7

Analysis of the effect of temperature on daily electricity demand in South Africa using extreme value theory

7.1 Introduction

The chapter presents an analysis of the effect of extreme low temperature on daily electricity demand in South Africa using the generalized extreme value theory approach. Electricity demand is not only influenced by the means of the distributions but by the tails of distributions as well. The tails of electricity demand distributions provide more information about extreme peak load forecasts. A Generalized Extreme Value Distribution (GEVD) is fitted to the temperature data below the reference temperature (18^0C). The reference tem-

perature which separates the cold and heat regions of the load-temperature relationship was derived in Chapter 3. Extreme Value Theory (EVT) is a powerful and fairly robust framework for modelling the tail behaviour of a distribution (Gencay and Selcuk, 2004).

7.2 Modelling extreme daily electricity demand using extreme value theory

Chapter 3 discussed the modelling of the means of daily electricity demand. Electricity demand is influenced by the tails of probability distributions as well as by means or averages. At times there is a need to depart from the average thinking and exploit information provided by the extremes (tails of distributions). The following sections discuss extreme value theory methods which will be used in the modelling of extreme average daily temperature and its effect on Daily Peak Electricity Demand (DPED).

7.3 Generalized Extreme Value Distribution

The Generalized Extreme Value Distribution (GEVD) is used to fit a distribution to the whole data set so as to model the upper tail of a distribution (maxima). The GEVD appears as a limiting distribution in the Fisher-Tippett Theorem (Beirlant *et al.*, 2004). The Fisher-Tippett Theorem is given in Theorem 7.1.

Theorem 7.1. *: Extreme Value Theorem (Fisher-Tippett, 1928)*
Let $X_1, X_2, ..., X_n$ *be independent and identically distributed random variables with common distribution function* F *and* $M_n = max(X_1, X_2, ..., X_n)$. *If there*

exist sequences of constants $a_n > 0$ *and* b_n *for* $n \geq 1$ *such that*

$$P\left(\frac{M_n - b_n}{a_n} \leq x\right) = F^n(a_n x + b_n) \to G(x)$$

as $n \to \infty$ *for a non-degenerate distribution function* G*, then* G *belongs to the generalized extreme value family of max-stable distributions which are the Gumbel, Frechet and Weibull classes.*

The proof of Theorem 7.1 is given in Gnedenko (1943). The three classes of distributions, Gumbel, Frechet and Weibull are discussed briefly in the following paragraphs. In this thesis we shall use $G_\gamma(x)$ to represent $G(x)$ where γ is the Extreme Value Index (EVI).

Gumbel class of distributions

The distribution function is given as (Fisher and Tippett, 1928):

$$G_\gamma(x) = \exp\left\{-\exp\left[-\frac{x-\mu}{\sigma}\right]\right\}, \gamma = 0 \tag{7.1}$$

For $\gamma = 0$ the distribution function is interpreted as a limit as $\gamma \to 0$. The right tail of these distributions decays exponentially and they have an infinite right end point (McNeil and Frey, 2000). Examples of distributions in this class are Benktander II, Weibull, exponential, gamma, logistic and log-normal (Beirlant *et al.*, 2004).

Fréchet-Pareto class of distributions

The distribution function is given as (Fisher and Tippett, 1928):

$$G_\gamma(x) = \exp\left\{-\left[1+\gamma\left(\frac{x-\mu}{\sigma}\right)\right]^{-\frac{1}{\gamma}}\right\}, \gamma > 0 \tag{7.2}$$

The right tail decays polynomially and is infinite (McNeil and Frey, 2000). Examples of distributions in this class are Pareto, Fréchet, Burr, F, inverse gamma and loggamma (Beirlant *et al.*, 2004).

The (Extremal) Weibull class of distributions

The distribution function is given as (Fisher and Tippett, 1928):

$$G_\gamma(x) = \exp\left\{-\left[1+\gamma\left(\frac{x-\mu}{\sigma}\right)\right]^{-\frac{1}{\gamma}}\right\}, \gamma < 0 \tag{7.3}$$

The distributions in this class have a finite right end point (McNeil and Frey, 2000). Examples of distributions in this class are Beta, uniform, reversed Burr and extreme value Weibull (Beirlant *et al.*, 2004).

The Gumbel, Frechet and Weibull class of distributions can be combined into one parametric family which is the GEVD (von Mises, 1954). The unified GEVD for modelling maxima is given by

$$G_\gamma(x) = \exp\left\{-\left[1+\gamma\left(\frac{x-\mu}{\sigma}\right)\right]^{-\frac{1}{\gamma}}\right\}, \text{ if } 1+\gamma\left(\frac{x-\mu}{\sigma}\right) > 0, \gamma \neq 0 \tag{7.4}$$

where μ and σ are the location and scale parameters respectively. The shape parameter γ also known as the EVI controls the tail behaviour. The EVI, γ is an important parameter which distinguishes the classes of distributions. If $\gamma = 0$, $G_\gamma(x)$ belongs to the light-tailed Gumbel class of distributions (von Mises, 1954). If $\gamma > 0$, $G_\gamma(x)$ belongs to the heavy-tailed Fréchet-Pareto class of distributions (von Mises, 1954). For $\gamma < 0$ we have the short-tailed Weibull class (von Mises, 1954). The probability density function of GEVD is given by

$$g_\gamma(x) = \frac{1}{\sigma}\left[1+\gamma\left(\frac{x-\mu}{\sigma}\right)\right]^{-\frac{1}{\gamma}-1}\exp\left\{-\left[1+\gamma\left(\frac{x-\mu}{\sigma}\right)\right]^{-\frac{1}{\gamma}}\right\} \tag{7.5}$$

Derivation of the quantile function for the GEVD

The GEVD is usually used to estimate extreme tail quantiles. The quantile function for estimating the tail quantiles is derived as follows:
The survival function is given by

$$P(X > x) = \tilde{G}_\gamma(x) = 1 - \exp\left\{-\left[1 + \gamma\left(\frac{x-\mu}{\sigma}\right)\right]^{-\frac{1}{\gamma}}\right\},$$

$$\text{if } 1 + \gamma\left(\frac{x-\mu}{\sigma}\right) > 0, \xi \neq 0 \tag{7.6}$$

Let

$$\begin{aligned}
p = P(X > x) &= 1 - \exp\left\{-\left[1 + \gamma\left(\frac{x-\mu}{\sigma}\right)\right]^{-\frac{1}{\gamma}}\right\} \\
\exp\left\{-\left[1 + \gamma\left(\frac{x-\mu}{\sigma}\right)\right]^{-\frac{1}{\gamma}}\right\} &= 1 - p \\
\left[1 + \gamma\left(\frac{x-\mu}{\sigma}\right)\right]^{-\frac{1}{\gamma}} &= -\log(1-p) \\
1 + \gamma\left(\frac{x-\mu}{\sigma}\right) &= [-\log(1-p)]^{-\gamma} \\
\gamma\left(\frac{x-\mu}{\sigma}\right) &= [-\log(1-p)]^{-\gamma} - 1 \\
x &= \mu + \frac{\sigma}{\gamma}\left\{[-\log(1-p)]^{-\gamma} - 1\right\}
\end{aligned}$$

Therefore the quantile function is given by

$$x_p = \mu + \frac{\sigma}{\gamma}\left\{[-\log(1-p)]^{-\gamma} - 1\right\}, \gamma \neq 0 \tag{7.7}$$

where p is the upper tail probability. For $\gamma = 0$ the GEVD given in equation (7.4) converges to a Gumbel distribution (Fisher and Tippett, 1928) which is

$$G_\gamma(x) = \exp\left\{-e^{-\frac{x-\mu}{\sigma}}\right\} \tag{7.8}$$

and the quantile function is given by

$$x_p = \mu - \sigma \log\left[\log(1-p)\right] \tag{7.9}$$

The quantile functions for the unified GEVD given in equations (7.7) and (7.9) are then used to estimate high quantiles and predicting the probability of exceedance levels using South African electricity demand data.

Modelling Maxima and Minima

The thesis concentrates on modelling of stationary time series processes at extreme levels using block maxima. We define the maximum of n observations in block i as $\{M_{in} = \max(X_{i1}, X_{i2}, ..., X_{in})\}_i, i = 1, 2, ..., b$ and $n \geq 2$, with $X_{i1}, X_{i2}, ..., X_{in}$ as the data in the i-th block. In order to extract upper extreme values an extreme value distribution is then fitted to the block maxima data $\{M_{in}\}_i, i = 1, ..., b$. The probability of M_{in} not exceeding a value x_i is then given as

$$P\{M_{in} \leq x_i\} = P\{\max X_{in} \leq x_i\} = P\{X_{i1} \leq x_i, X_{i2} \leq x_i, ..., X_{in} \leq x_i\}$$

$$= \{F(x_i)\}^{in} \tag{7.10}$$

For modelling minima we use the duality between the distributions for maxima and minima, that is $\tilde{M}_n = \min X_{in} = -\max\{-X_{i1}, -X_{i2}, ..., -X_{in}\}$, $i = 1, ..., b$ and $n \geq 2$ (Coles, 2001). The distribution function of minima is given by

$P\{\min X_{in} \leq x_i\} = 1 - \{1 - F(x_i)\}^{in}$. Extreme maxima theory and methods are then used to model extreme minima.

In this chapter the parameters γ, μ and σ of the GEVD are estimated using the Maximum Likelihood (ML) method. Let $\boldsymbol{\theta} = (\gamma, \mu, \sigma)$, then ML of $\boldsymbol{\theta}$ is given by

$$\begin{aligned} L(\boldsymbol{\theta}, x) &= \prod_{i=1}^{n} g_{\gamma}(x_i) \\ &= \prod_{i=1}^{n} \frac{1}{\sigma}\left[1+\gamma\left(\frac{x_i-\mu}{\sigma}\right)\right]^{-\frac{1}{\gamma}-1} \exp\left\{-\left[1+\gamma\left(\frac{x_i-\mu}{\sigma}\right)\right]^{-\frac{1}{\gamma}}\right\} \end{aligned} \tag{7.11}$$

Let the log-likelihood be denoted by $\ell(\boldsymbol{\theta}, x)$, i.e. $\ell(\boldsymbol{\theta}, x) = \log L(\boldsymbol{\theta}, x)$, then

$$\ell(\boldsymbol{\theta}, x) = -n\log\sigma - \left(1+\frac{1}{\gamma}\right)\sum_{i=1}^{n}\log\left\{1+\frac{\gamma(x_i-\mu)}{\sigma}\right\} - \sum_{i=1}^{n}\left\{1+\frac{\gamma(x_i-\mu)}{\sigma}\right\}^{-\frac{1}{\gamma}} \tag{7.12}$$

The R statistical package, ismev by Heffernan and Stephenson (2013) is used to obtain the ML estimates.

7.4 Piecewise linear regression model

Our modelling approach is in two stages. A piecewise linear regression model is used to explore the influence of temperature on daily electricity demand. In stage two, a generalized extreme value distribution is fitted to the temperature values below the reference temperature. The fitted distribution is then used to estimate the likelihood of extreme low temperatures and consequential extreme increase in electricity demand.

The model for Average Daily Electricity Demand (ADED) developed in chapter 3 is now used to analyse the influence of low temperature on daily electricity

demand using the GEVD. The piecewise linear regression model is given in equation (7.13).

$$\text{ADED} = k_0 + k_1\text{max}(0, T_s - T_t) + k_2\text{max}(0, T_t - T_w) + \varepsilon_t \tag{7.13}$$

where ADED is the average daily electricity demand, T_t is the average daily temperature, T_s and T_w are temperatures which separate summer and winter sensitive periods (hot and cold temperatures) from the weather neutral period respectively. The parameters to be estimated are k_0, k_1 and k_2, and ε_t is the error term with $\varepsilon_t \sim N(0, \sigma_t^2)$. The values for T_w and T_s used in this chapter are 18^0C and 22^0C respectively.

7.5 Data

Average Daily Electricity Demand (ADED) data and average daily temperature[1] (T_t) for the period, years 2000 to 2010.

7.6 Empirical results and discussion

7.6.1 Piecewise linear regression model output

The model identifies the winter sensitive, weather neutral and summer sensitive periods. The model is not used for forecasting electricity demand but rather to explain the influence of temperature on electricity demand. The model is an extension of the one developed in chapter 3 in which we use 18^0C for T_w instead of 16^0C for the ADED-average daily temperature relationship

[1]Average daily temperature for the whole country is usually built into the modelling as weighted average temperatures from different meteorological stations of a country. The weightings should reflect consumption of electricity of each region (province). Population figures are often used for estimating the weights. In this thesis the weightings were not done since only aggregated average daily temperature was available at the time of the study.

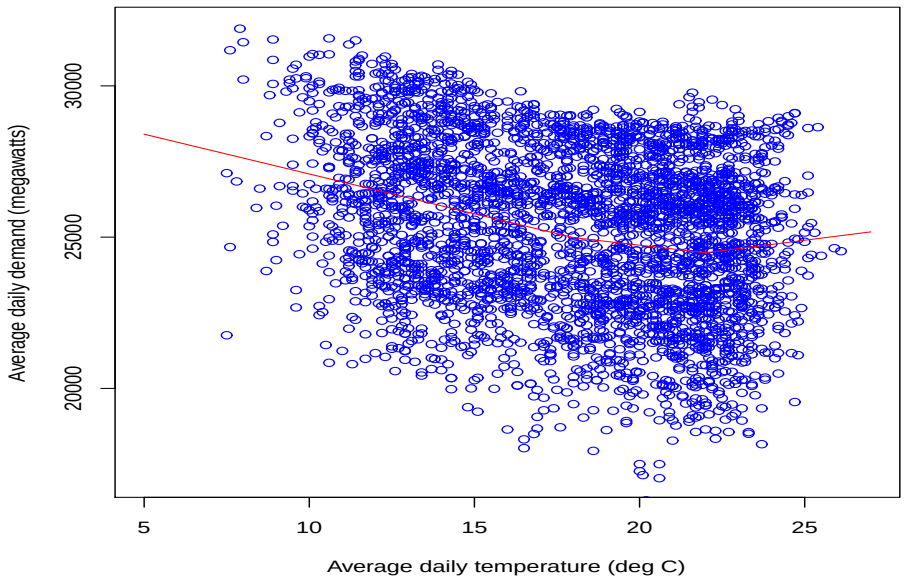

Figure 7.1: Scatter plot of ADED (in megawatts) against T_t (in ^{0}C) with the fitted piecewise regression line of equation (7.14).

model.

$$\hat{\text{ADED}} = 23932 + 263\text{max}(0, 22 - T_t) + 138\text{max}(0, T_t - 18) \tag{7.14}$$

The graphical plot of ADED against T_t is shown in Figure 7.1. The piecewise linear regression plot separates the non-linear response of electricity demand to temperature into three regions: cold for temperatures lower than 18^0C, neutral for temperatures between 18^0C and 22^0C, and hot for temperatures above 22^0C.

The three demand-temperature equations are given in equations 7.15 to 7.17. If average daily temperature is less than or equal to 18^0C we have

$$\hat{\text{ADED}} = 23932 + 263\text{max}(0, 22 - T_t)) \tag{7.15}$$

That is, if the temperature decreases by 1^0C (e.g. from 18^0C to 17^0C) electricity demand will increase marginally by 263MW. If average daily temperature is

greater than or equal to 22^0C we have

$$\hat{\text{ADED}} = 23932 + 138\text{max}(0, T_t - 18) \tag{7.16}$$

If temperature increases by 1^0C (e.g. from 22^0C to 23^0C) electricity demand will increase marginally by 138MW. For the average daily temperature between 18^0C and 22^0C we have

$$\hat{\text{ADED}} = 23932 + 263\text{max}(0, 22 - T_t) + 138\text{max}(0, T_t - 18) \tag{7.17}$$

If temperature decreases by 1^0C (e.g. from 22^0C to 21^0C) electricity demand will increase marginally by 125MW.

7.6.2 Comparative analysis

Figure 7.2 shows a plot of $\hat{\text{ADED}} = 23932 + 263\text{max}(0, 22 - T_t) + 138\text{max}(0, T_t - 18)$). A comparative analysis with the relationship between electricity demand and average temperature for South Australia shows that electricity demand in South Australia is highly sensitive to summer months as shown in Figure 7.3. Although the climate of Australia varies, the largest part is desert and semi-arid and temperature ranges from around 0^0C to 45^0C between winter and summer periods as shown in Figure 7.3. Most people stay in cities which are clustered around coastal areas which are very hot and humid. On the contrary temperatures in South Africa are generally lower than in other countries at similar latitudes such as Australia due to the fact that South Africa is on a greater elevation above sea level (South Africa information on climate (n.d.)).

7.6.3 Modelling extreme minimum temperatures

Modelling of extreme minimum temperatures is important to load forecasters and system operators in South Africa since electricity demand increases signif-

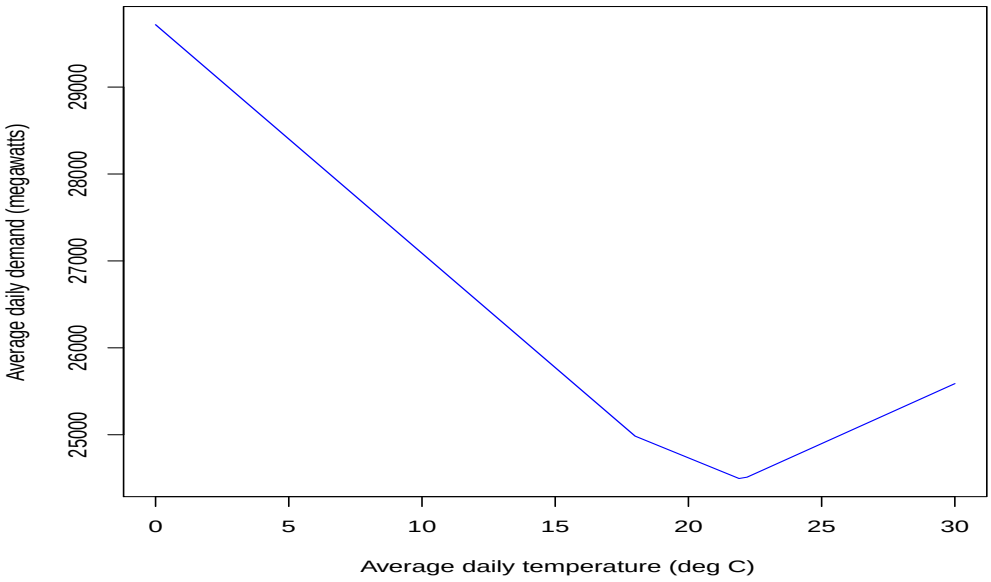

Figure 7.2: Plot of $\widehat{\text{ADED}} = 23932 + 263\textbf{max}(0, 22 - T_t) + 138\textbf{max}(0, T_t - 18))$.

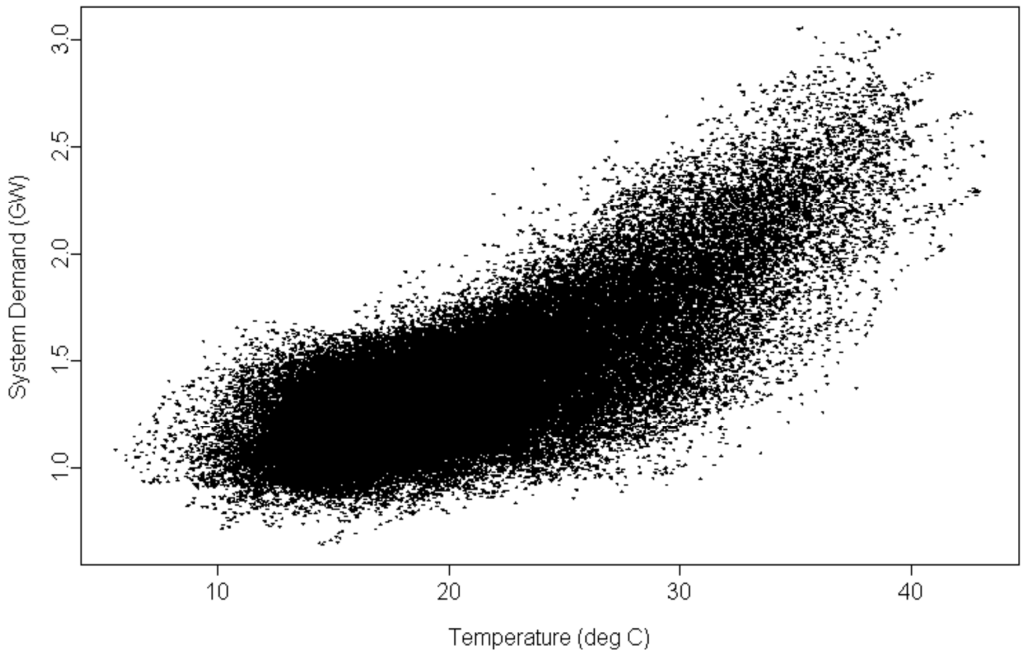

Figure 7.3: Half-hourly South Australian electricity demand in gigawatts (excluding major industrial demand) plotted against temperature (degrees Celsius) (Hyndman and Fan, 2010).

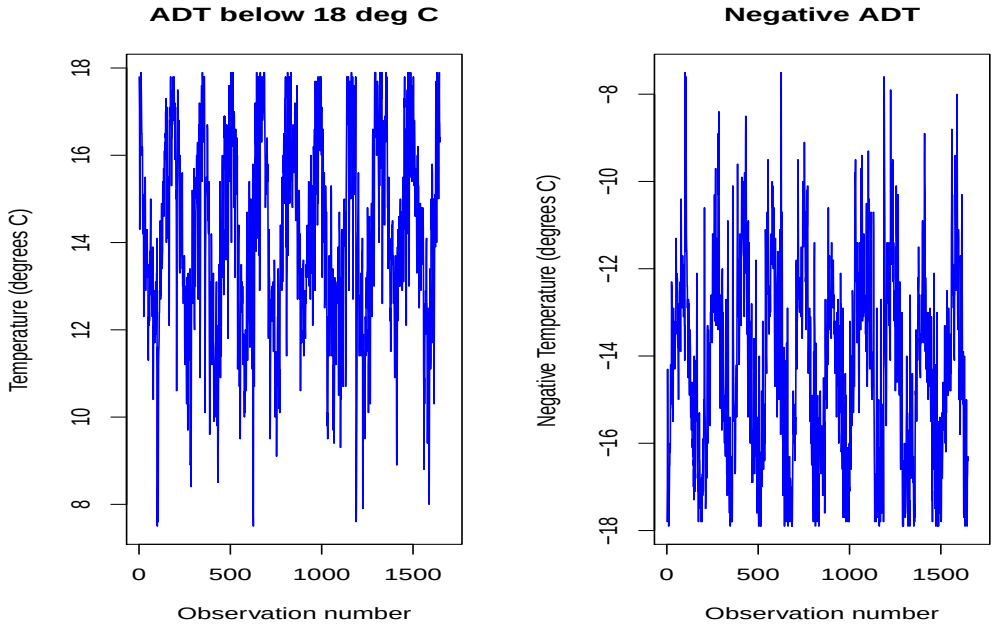

Figure 7.4: (a) Left panel: Average Daily Temperature (ADT or T_t) below 18^0C and (b) Right panel: Inverted graph of ADT below 18^0C. The graph on the left panel is inverted so that we can use the duality between the distributions for maxima and minima.

icantly for average daily temperature values below 18^0C. We use the principle of duality between the distributions of minima and maxima as discussed in Section 7.3. In Figure 7.4 the left panel shows data for average daily temperature (T_t) below 18^0C while the right panel shows the graph of $-T_t$.

7.6.4 Tail quantile estimation

We estimate the extreme tail quantiles of T_t below 18^0C using the GEVD. The data is seasonally adjusted. There are 1649 observations of average daily temperature below 18^0C. The ML estimates of the parameters γ, μ and σ are given

as (the standard errors are given in parentheses):

$$\begin{aligned}\hat{\gamma} &= -0.2004(0.01817),\\ \hat{\mu} &= -15.02544(0.05758),\\ \hat{\sigma} &= 2.0800(0.04176).\end{aligned}$$

The standard errors of the ML estimates of the parameters, are small. This shows that uncertainty about the parameters is small. These results show that the data can be modelled using a Weibull class of distributions, (since $\hat{\gamma} < 0$) and the right endpoint is finite. The right end point is

$$\mu - \frac{\sigma}{\gamma} = -15.202544 - \frac{2.08}{-0.2004} = -4.646$$

This implies that for any degree decrease below 4.6^0C there is unlikely to be any further increase in electricity demand.

The Quantile-Quantile (QQ) and Probability-Probability (PP) plots given in Figure 7.5 show that the Weibull class is a good fit to the data. The return level estimates are inside the 95% confidence interval. This is an indication that the fitted Weibull class is capable of accurately predicting future return levels.

We then use equation (7.7) to estimate high quantiles and predicting the probability of exceedance levels. For example the 95^{th} quantile is obtained as follows

$$-x_{0.05} = -15.025 - \frac{2.08}{0.2004}\left[(-\log(0.95))^{0.2004} - 1\right] = -10.369 \Rightarrow x_{0.05} = 10.369$$

The number of temperature observations that are smaller than the estimated tail quantile $(x_{0.05} = 10.369)$ are then counted and found to be 78. For the observed number of exceedances we get $0.05 \times 1649 = 82.45 \approx 82$ where 1649 is

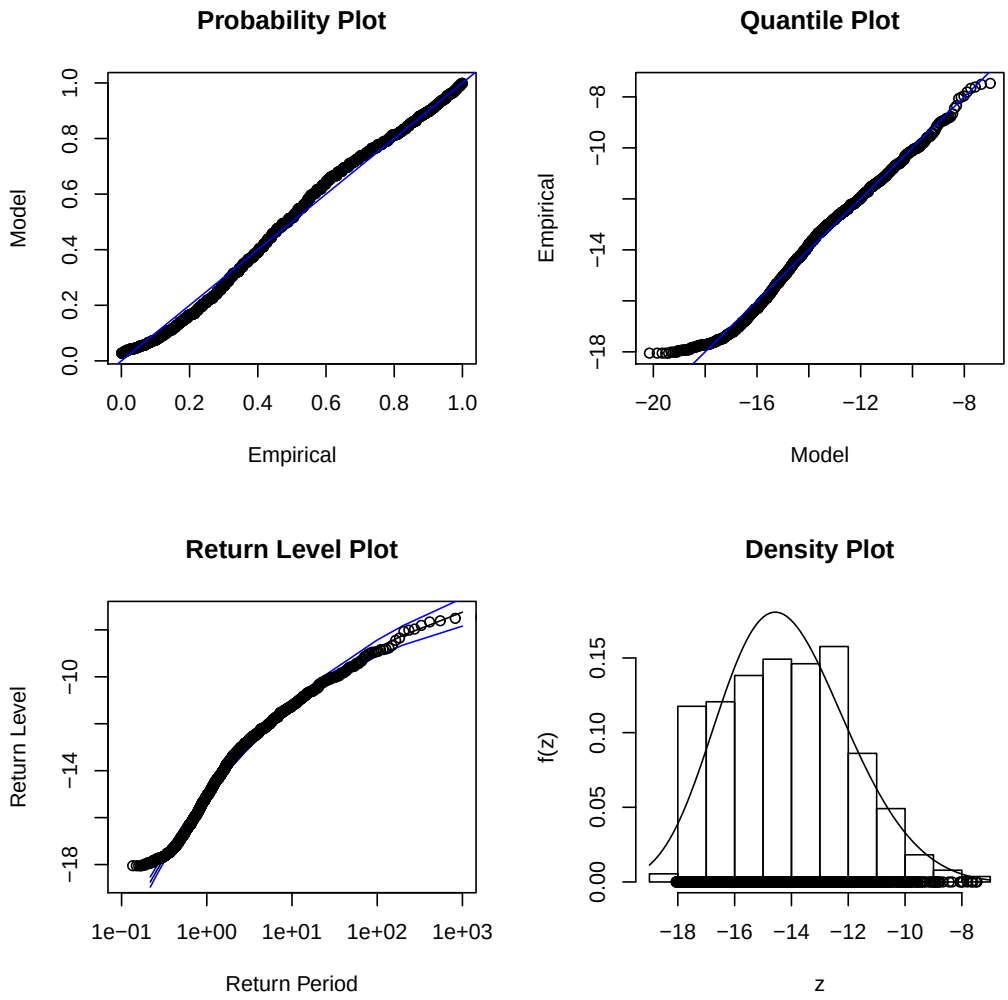

Figure 7.5: Diagnostic plots illustrating the fit of the data (temperature below 18^0C) to the GEVD, (a) Top left panel: Probability plot, (b) Top right panel: Quantile plot, (c) Bottom left panel: Return level plot and (d) Bottom right panel: Density plot.

the number of temperature values below 18^0C. The marginal increase in electricity demand for a drop of temperature from 18^0C to $x_{0.05} = 10.4^0$C is given by $(18 - 10.4) \times 263 = 1998.8$MW where 263 is the marginal increase in demand for a decrease of 1^0C below 18^0C as discussed in Section 7.6.1. It should be noted that this increase in ADED for temperature decreases below 18^0C is bounded. As temperature decreases below 18^0C the increases in ADED reaches a certain value after which any further decrease in temperature will not have any effect on ADED. That is as temperature decreases people will switch on heating systems up to a point where all the heating systems are all switched on and no additional energy is consumed for any further decrease in temperature.

Table 7.1 presents a summary of the estimated tail quantiles at different tail probabilities. The second column shows the calculated quantiles (temperature in degrees C). The number of exceedances using the GEVD are given in column 3 while column 4 shows the observed number of exceedances. In equation (7.16) it is given that for a degree decrease in temperature below 18^0C there will be a marginal increase in demand of 263 MW. For each of the estimated quantiles in column 2, column 5 shows the marginal increase in demand from one quantile to the next, e.g. if temperature drops from 11.3^0C to 10.4^0C there will be an increase in demand of 263(11.3-10.4)=236.7MW. Similarly for a decrease from 10.4^0C to 8.8^0C the increase will be 263(10.4-8.8)=420.8MW. Extreme low average daily temperatures of the order of 8.2^0C are very rare in South Africa. This only occurs about 8 times in a year. Table 7.2 summarizes the temperature values at high quantiles and the corresponding marginal increases while Figure 7.6 shows that the marginal increases converge to 1.58MW when temperature converges to 4.6^0C.

Table 7.1: Estimated tail quantiles at different probabilities.

Quantiles	Temp[a] (x_p)	Expected number of exceedances (assuming GEVD)	Observed number of exceedances	Marginal increase in demand (MW)
90^{th} quantile	11.3^0C	159	164	
95^{th}	10.4^0C	78	82	236.7
99^{th}	8.8^0C	12	16	420.8
99.5^{th}	8.2^0C	8	8	157.8
99.9^{th}	7.2^0C	0	2	263

[a]Temp denotes Temperature

A summary of the monthly frequency of occurrence of average temperature values below 12.4^0C (i.e. above the 95^{th} quantile $(-x_{0.05} = -10.369)$ is given in Table 7.3. Over the sampling period, i.e. years 2000 to 2010 the month of July has the highest number of days with temperature values below 10.4^0C. This is an indication that the month of July is the coldest month in South Africa and the winter period is from May to August of each year.

The bar chart of the monthly frequency of occurrence of exceedances is given in Figure 7.7.

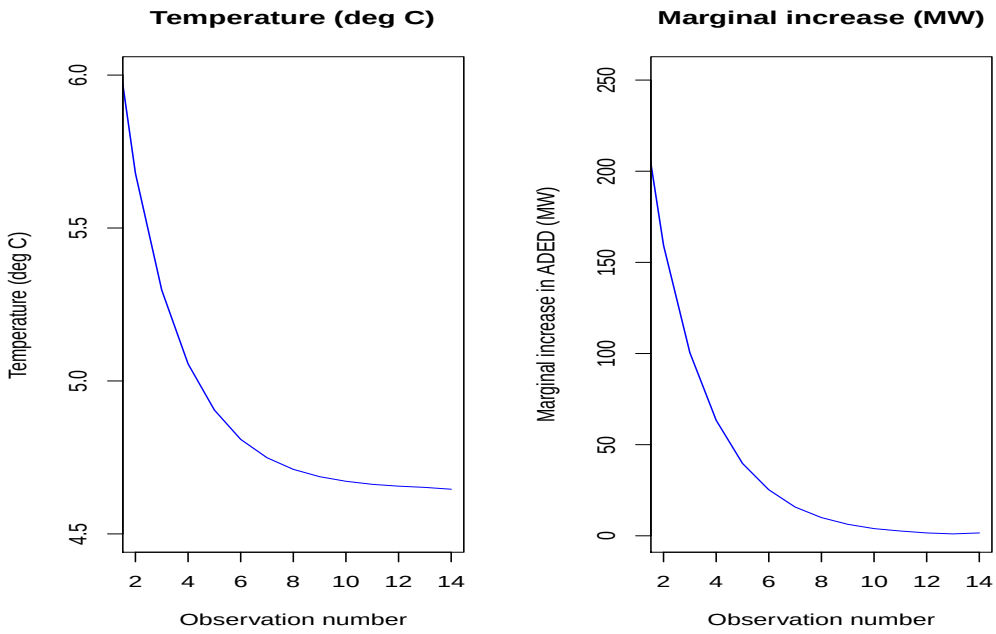

Figure 7.6: (a) Left panel: shows that the gradual decrease in temperature converges to 4.6^0C and (b) Right panel: shows that the marginal increases converge to 1.58MW when temperature converges to 4.6^0C.

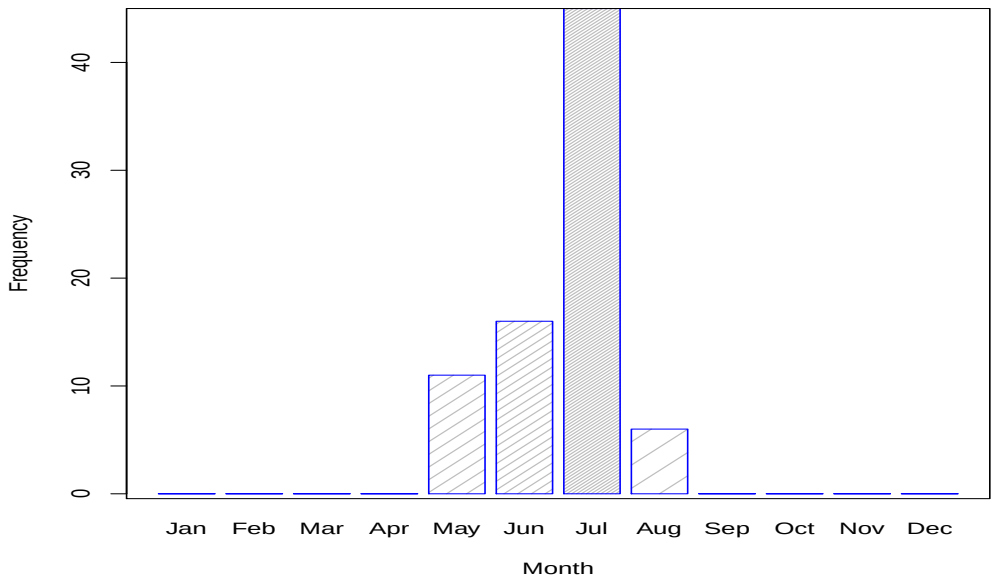

Figure 7.7: Bar chart of the monthly frequency of occurrence of temperature below $x_p = 10.4^0C$ (exceedances). The exceedances are average daily temperature values below 10.4^0C

Table 7.2: Evaluation of estimated tail quantiles at different probabilities.

Number[a]	1	2	3	4	5
Quantile	99.99^{th}	99.999^{th}	99.9999^{th}	$99.999\ 99^{\text{th}}$	$99.999\ 999^{\text{th}}$
Temp (deg C)[b]	6.285	5.679	5.297	5.056	4.905
MI[c] (MW)	252.74	159.38	100.5	63.38	39.71
Number	6	7	8	9	10
Quantile	$99.9999\ 999^{\text{th}}$	$99.9999\ 9999^{\text{th}}$	$99.9999\ 99999^{\text{th}}$	$99.999999\ 9999^{\text{th}}$	$99.999999\ 99999^{\text{th}}$
Temp (deg C)	4.809	4.749	4.711	4.687	4.672
MI (MW)	25.248	15.78	9.994	6.312	3.945
Number	11	12	13	14	
Quantile	$99.999999\ 999999^{\text{th}}$	$99.9999999\ 999999^{\text{th}}$	$99.9999999\ 9999999^{\text{th}}$	$99.99999999\ 9999999^{\text{th}}$	
Temp (deg C)	4.662	4.656	4.652	4.646	
MI (MW)	2.63	1.578	1.052	1.578	

[a]Number denotes the observation number for each of the quantiles
[b]Temp (deg C) represents temperature in degrees Celcius
[c]MI denotes marginal increase in megawatts

Table 7.3: Monthly frequency of exceedances below $x_p = 10.369$ (average temperature below 10.4^0C) by month.

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Freq	0	0	0	0	11	16	45	6	0	0	0	0

7.7 Concluding remarks

An analysis of the intensity and frequency of occurrence of extreme low temperatures is important for load forecasters in the electricity sector. In this chapter the modelling of the influence of temperature on average daily electricity demand in South Africa using a piecewise linear regression model and the extreme value theory modelling framework is discussed. The developed piecewise linear regression model is not meant for forecasting but to model the effect of temperature on electricity demand. Extreme low temperatures can be modelled by the Weibull class of distributions. Extreme low temperatures of the order of 6^0C are very rare in South Africa, but can cause huge increases in

electricity demand. An investigation of expected cooler than normal years is important and helps in guiding planning to decision makers in the electricity sector.

Electricity demand in South Africa is highly sensitive to temperature fluctuations during the winter periods. In contrast, Australia electricity demand is sensitive to summer months (Hyndman and Fan, 2010) but both Australia and South Africa are on the same latitude in the Southern Hemisphere.

The use of extreme value theory provides a lot of information in the modelling of extreme peak electricity demand. Chapter 8 presents a detailed analysis of extreme daily and same day of the week increases in peak electricity demand in South Africa using extreme value theory.

Chapter 8

Analysis of extreme daily and same day of the week increases in peak electricity demand in South Africa

8.1 Introduction

The chapter is divided into two main parts. The first part deals with modelling of extreme daily increases in peak electricity demand. Peak electricity demand is an energy policy concern for all economies throughout the world and if unchecked, excess peak demand may cause blackouts and increase the cost of electricity for consumers (Strengers, 2012). This has resulted in many economies in designing energy efficient and demand side management strategies to either redistribute or reduce energy demand during peak periods. This

requires modelling and accurate assessment of the frequency and level of peak electricity demand (Hor *et al.*, 2008).

The South African Daily Peak Electricity Demand (DPED) data is not stationary as seen in the graphical plot in Chapter 3. In this thesis the data is made stationary by taking first and seventh differences and also by removing the linear trend (detrending). The first difference of DPED will give rise to an analysis of interday (daily) changes in peak electricity demand, while the seventh difference will result in day to day changes. Linear detrending will result in modelling extreme winter peaks in electricity demand. We briefly describe the three ways of making the DPED data stationary in the following paragraphs.

1. Taking the first difference results in daily (interday) changes in peak electricity demand, i.e.

$$\nabla x_t = (1 - B)x_t = x_t - x_{t-1}$$

   The model can be written as ARIMA(0,1,0) which is also known as a random walk or naive model. A plot of the first difference of DPED is given Figure 8.1, top left panel.

2. Taking the seventh difference will give rise to same day of the week changes in peak demand, i.e.

$$\nabla_7^1 x_t = (1 - B^7)x_t = x_t - x_{t-7}$$

   This is also known as a Seasonal Naive (SNAIVE) or SARIMA(0,0,0)(0,1,0)[7] model. Figure 8.1, top right panel shows a plot of the seventh difference of DPED.

3. Detrending
   Initially we fit a trend line to DPED, that is, we regress DPED (x_t) against

time t, for $t = 1, .., n$. The trend line is given by the following equation

$$x_t = \beta_0 + \beta_1 t + \varepsilon_t$$

where β_0, β_1 are parameters and ε_t is the error term. We then subtract the values of the trend line from the original data which results in a time series of residuals, i.e.

$$\varepsilon_t = x_t - \hat{x}_t = x_t - (\hat{\beta}_0 + \hat{\beta}_1 t)$$

The detrended DPED (Figure 8.1 bottom left panel) will be used for modelling extreme winter peak electricity demand in Chapter 9.

Both daily changes and same day of the week changes in peak electricity demand will be discussed in this Chapter.

We define daily increase in peak electricity demand as the positive inter-day change in DPED. Extreme daily increase in peak electricity demand is therefore positive inter-day change above a sufficiently high threshold. Extreme value theory (EVT) is then used to model extreme daily increases in peak electricity demand and also to investigate whether extreme daily increases such as the one experienced in May 2007 (DPED of 37158MW) in South Africa is truly an extreme low-probability event or is one which will appear on a regular basis. The distribution of extreme daily increases in peak electricity is modelled using the Generalized Pareto Distribution (GPD). A comparative analysis is then done using the Generalized Single Pareto Distribution (GSPD). The second part of the chapter discusses an application of the GSPD in predicting the Probability of Exceedance (PoE) levels of the same day of the week increases in peak electricity demand. A comparative analysis is done with a GPD.

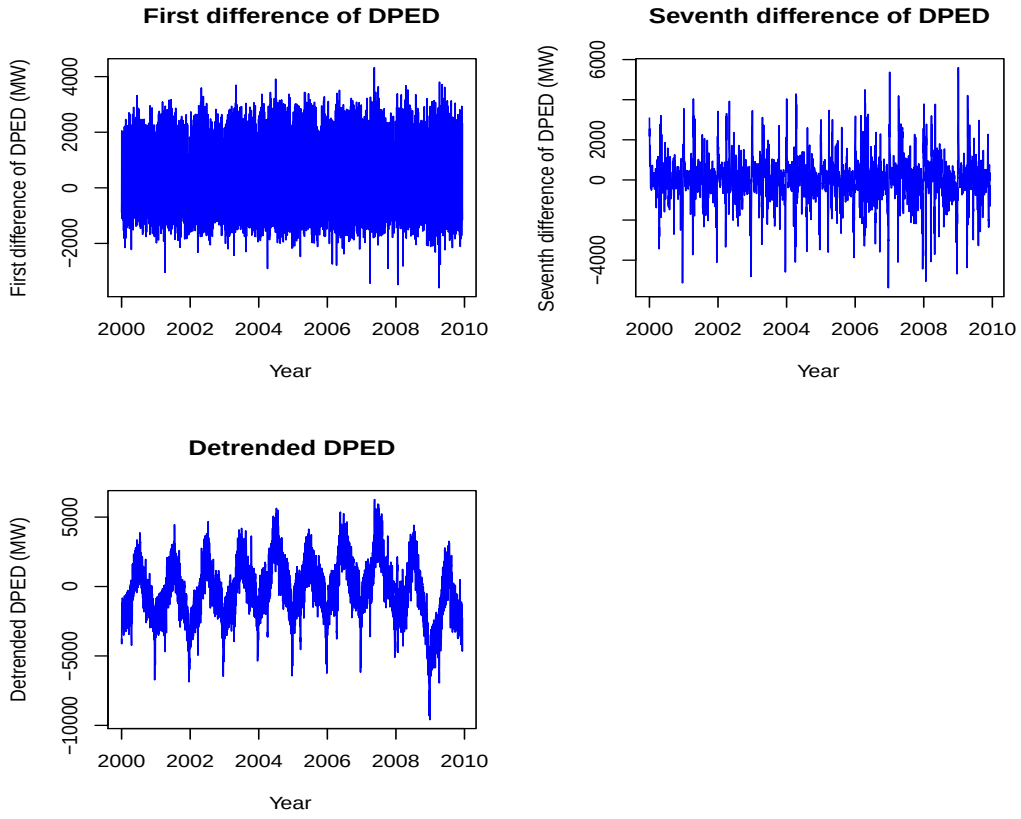

Figure 8.1: (a) Top left panel: First difference of DPED (MW) (b) Top right panel: Seventh difference of DPED (MW) (c) Bottom left panel: Detrended DPED (MW).

Modelling of exceedances above a sufficiently high threshold has been studied for over three decades. Research has shown that the distribution function of the excesses above a high threshold converges to a GPD as the threshold tends to the right endpoint (Balkema and de Haan, 1974; Pickands, 1975).

If we let $X_1, X_2, ..., X_n$ to be a sequence of hourly electricity demand, then

$$M_n = \mathbf{max}(X_1, X_2, ..., X_n)$$

is the maximum hourly electricity demand over an n - observation period where $n \geq 2$. We divide the hourly data into b blocks. Each block consists of 24 hours each, which is a day. The maximum of n observations in block i is then defined as $\{M_{in} = \mathbf{max}(X_{i1}, X_{i2}, ..., X_{in})\}_i, i = 1, 2, ..., b$ and $n \geq 2$, with $X_{i1}, X_{i2}, ..., X_{in}$ the data in the i-th block. The upper extremes from the sequence, $M_{1n}, M_{2n}, ..., M_{bn}$ can be extracted by taking the exceedances over a predetermined high threshold τ. The exceedances y_j over τ are those x_j such that $x_j > \tau$. The values $y_j = x_j - \tau$ are the excesses over τ. Then if $F^{[\tau]}(b_\tau + a_\tau x)$ has a continuous limiting distribution function as τ goes to the right endpoint, $\omega(F)$, then

$$\left|F^{[\tau]}(x) - W_\xi(x)\right| \to 0, as \;\; \tau \to \omega(F) \tag{8.1}$$

for some generalized Pareto distribution function, $W_\xi(x)$, for a suitable choice of constants $b_\tau > 0$ and $a_\tau > 0$ (Balkema and de Haan, 1974; Pickands, 1975). $F^{[\tau]}$ represents the empirical exceedance distribution function at τ.

8.2 Generalized Pareto Distribution (GPD)

The GPD is a Peaks Over Threshold (POT) distribution which is used to model observations above a reasonably high threshold. The unified GPD is given by

$$W_\xi(x) = \begin{cases} 1 - \left(1 + \frac{\xi(x-\tau)}{\sigma}\right)^{-\frac{1}{\xi}} & \text{if } \xi > 0, x > \tau \\ 1 - \exp\left(-\frac{x-\tau}{\sigma}\right), & \text{if } \xi = 0, x > \tau \\ 1 - \left(1 + \frac{\xi(x-\tau)}{\sigma}\right)^{-\frac{1}{\xi}} & \text{if } \xi < 0, \tau < x < \tau - \frac{\sigma}{\xi} \end{cases} \tag{8.2}$$

where $\xi > 0$ is the Extreme Value Index (EVI) (shape parameter), σ is the scale parameter and τ is the threshold. Equation (8.2) shows that when $\xi < 0$ the distribution function of the GPD is bounded above, i.e. the distribution has a finite right end point.

8.2.1 Threshold selection

Several methods are proposed in literature on how to select a threshold when using peaks over threshold distributions such as the generalized Pareto type. In this thesis we use two methods, the mean excess plot and the Pareto quantile plot. One of the benefits of using these methods is that one can inspect the features of the data and assess the model fit (Scarrott and MacDonald, 2012). One of the drawbacks is the lack of a formal assessment of the uncertainty associated with the threshold selected (Scarrott and MacDonald, 2012).

Mean excess plot

One of the methods used in selecting a threshold, τ is the mean excess plot. This was introduced by Davison and Smith (1990). The mean excess function, $e(t)$ is defined as (Beirlant *et al.*, 2004):

$$e(t) = E(X - \tau | X > \tau) \tag{8.3}$$

A mean excess plot is a plot of $x_{n-k,n}$ against the mean excesses $e_{k,n}$. The mean excesses $e_{k,n}$ are defined as (Beirlant *et al.*, 2004):

$$e_{k,n} := \hat{e}_n(x_{n-k,n}) = \frac{1}{k}\sum_{i=1}^{k} x_{n-i+1,n} - x_{n-k,n} \quad k = 2, ..., n \tag{8.4}$$

where $\hat{e}_n$ is the empirical function used to estimate e. The threshold is selected where the plot follows a linear pattern (Beirlant *et al.*, 2004).

Pareto quantile plot

The Pareto quantile plot is defined as the scatter plot of the following points: $(-\log(1-p_i), \log x_i)$ where $i = 1, ..., n$ (Beirlant *et al.*, 1996); Beirlant *et al.*, 2004). Several choices of p_i are used in literature (Beirlant *et al.*, 2004). In this thesis we will use $p_i = \frac{i}{n+1}$. The threshold is taken as the observation on the vertical axis (y-axis) where the plot starts to follow a linear pattern (Beirlant *et al.*, 2004). An analytic relationship between the GEVD and GPD exists and is given in the following theorem (where $\gamma = \xi$) (Reiss and Thomas, 2007).

Theorem 8.1

$$W_\xi(x) = 1 + \log G_\xi(x), \text{ for } \log G_\xi(x) > -1$$

An outline of the proof

Using the right hand side of Theorem 8.1 we get

$$\begin{aligned} & 1 + \log\left[\exp\left\{-\left[1+\xi\left(\frac{x-\tau}{\sigma}\right)\right]^{-\frac{1}{\xi}}\right\}\right] \\ = \; & 1 - \left(1 + \frac{\xi(x-\tau)}{\sigma}\right)^{-\frac{1}{\xi}} \\ = \; & W_\xi(x) \end{aligned}$$

The *pdf* of the GPD is

$$\begin{aligned} w_\xi(x) &= W'_\xi(x) \\ w_\xi(x) &= \frac{1}{\sigma}\left(1+\frac{\xi(x-\tau)}{\sigma}\right)^{-(1+\frac{1}{\xi})} \end{aligned} \tag{8.5}$$

8.3 Generalized Single Pareto Distribution

The distribution function of the GSPD (Verster and De Waal, 2011) is given in equation (8.6).

$$W_\eta(x) = \left\{1+\frac{\eta}{1+\tau\eta}(x-\tau)\right\}^{-\frac{1}{\eta}}, \eta \neq 0, x > \tau \tag{8.6}$$

where η is the EVI and τ is the threshold. The GSPD is a generalized Pareto-type (GP-type) distribution with one parameter to estimate, i.e. the GSPD belongs to the domain of attraction of the GPD written as:

$$W_\eta(x) \in D(W_\xi(x))$$

The *pdf* is

$$\begin{aligned} w_\eta(x) &= W'_\eta(x) \\ w_\eta(x) &= \frac{1}{1+\tau\eta}\left[1+\frac{\eta(x-\tau)}{1+\tau\eta}\right]^{-\frac{1}{\eta}-1} \end{aligned} \tag{8.7}$$

8.4 Estimation of parameters

In this section we discuss the methods for estimating the parameters of GPD and GSPD. We start with a brief discussion of the Hill estimator for estimating

the EVI. This is followed by a discussion of the Maximum Likelihood (ML) and then the Bayesian estimation methods.

8.4.1 Hill estimator

The Hill estimator (Hill, 1975) is an estimator of the EVI for the Generalized Pareto - type (GP-type) distributions. It is based on $k+1$ upper order statistics and is given as

$$H_{k,n} = \frac{1}{k}\sum_{i=1}^{k} \log X_{(n-i+1,n)} - \log X_{(n-k,n)} \tag{8.8}$$

The Hill estimator is ML estimator of the mean excess values of log-transformed data (Beirlant *et al*., 2004). Some of the pitfalls of the Hill estimator are that, for different choices of k the estimator of the EVI changes, i.e. for shifts in the data the estimator is not invariant despite the fact that some invariant estimators are proposed in literature (Pereira, 1994; Drees, 1995, 1998; Fraga Alves, 2001).

8.4.2 Maximum Likelihood Estimation

GPD

Let $\boldsymbol{\alpha} = (\xi, \sigma)$ then ML of $\boldsymbol{\alpha}$ is given by

$$\begin{aligned} L(\boldsymbol{\alpha}, x) &= \prod_{i=1}^{N_\tau} w_\xi(x_i) \\ &= \prod_{i=1}^{N_\tau} \frac{1}{\sigma}\left[1+\xi\left(\frac{x_i-\tau}{\sigma}\right)\right]^{-\left(1+\frac{1}{\xi}\right)} \end{aligned} \tag{8.9}$$

where N_τ is the number of observations above the threshold τ. The log-likelihood is

$$\ell(\boldsymbol{\alpha}, x) = -n\log\sigma - \left(1 + \frac{1}{\xi}\right) \sum_{i=1}^{N_\tau} \log\left[1 + \frac{\xi(x_i - \tau)}{\sigma}\right] \tag{8.10}$$

GSPD

$$\begin{aligned} L(\eta, x) &= \prod_{i=1}^{N_\tau} w_\eta(x_i) \\ &= \prod_{i=1}^{N_\tau} \left\{ \frac{1}{1+\eta\tau} \left[1 + \frac{\eta}{1+\eta\tau}(x_i - \tau)\right]^{-\frac{1}{\eta}-1} \right\} \end{aligned} \tag{8.11}$$

Outline of the ML of the GSPD

Let $L(\eta, x)$ be the maximum likelihood function, then

$$\begin{aligned} L(\eta, x) &= \prod_{i=1}^{N_\tau} w_\eta(x_i) \\ &= \prod_{i=1}^{N_\tau} \frac{1}{1+\tau\eta} \left\{ 1 + \frac{\eta}{1+\tau\eta}\left(x_i - \tau\right) \right\}^{-\frac{1}{\eta}-1} \end{aligned}$$

Now

$$\log L(\eta, x) = \sum_{i=1}^{N_\tau} \log\left[\frac{1}{1+\tau\eta} \left\{ 1 + \frac{\eta}{1+\tau\eta}\left(x_i - \tau\right) \right\}^{-\frac{1}{\eta}-1} \right]$$

Let $\ell(\eta, x) = \log L(\eta, x)$

Then

$$\begin{aligned}
\ell(\eta, x) &= \sum_{i=1}^{N_\tau}\left[-\log(1+\tau\eta) - \frac{1+\eta}{\eta}\log\left\{1+\frac{\eta}{1+\tau\eta}(x_i-\tau)\right\}\right] \\
&= \sum_{i=1}^{N_\tau}\left[-\log(1+\tau\eta) - \frac{1+\eta}{\eta}\log\left\{\frac{1+\eta x_i}{1+\tau\eta}\right\}\right] \\
&= \sum_{i=1}^{N_\tau}\left[-\log(1+\tau\eta) - \frac{1+\eta}{\eta}\left\{\log(1+\eta x_i) - \log(1+\tau\eta)\right\}\right] \\
&= \sum_{i=1}^{N_\tau}\left[-\log(1+\tau\eta) - \frac{1+\eta}{\eta}\log(1+\eta x_i) + \frac{1+\eta}{\eta}\log(1+\tau\eta)\right] \\
&= \sum_{i=1}^{N_\tau}\left[\frac{1}{\eta}\log(1+\tau\eta) - \frac{1+\eta}{\eta}\log(1+\eta x_i)\right] \\
&= \sum_{i=1}^{N_\tau}\left[\frac{1}{\eta}\log(1+\tau\eta)\right] - \sum_{i=1}^{N_\tau}\left[\frac{1+\eta}{\eta}\log(1+\eta x_i)\right]
\end{aligned}$$

Then

$$\frac{\partial \ell(\eta, x)}{\partial \eta} = \sum_{i=1}^{N_\tau}\frac{\partial}{\partial \eta}\left[\frac{1}{\eta}\log(1+\tau\eta)\right] - \sum_{i=1}^{N_\tau}\frac{\partial}{\partial \eta}\left[\frac{1+\eta}{\eta}\log(1+\eta x_i)\right] = 0$$

$$\sum_{i=1}^{N_\tau}\left[-\frac{1}{\eta^2}\log(1+\tau\eta) + \frac{\tau}{\eta(1+\tau\eta)}\right] - \sum_{i=1}^{N_\tau}\left[-\frac{1}{\eta^2}\log(1+\eta x_i) + \frac{\left(1+\frac{1}{\eta}\right)x_i}{1+\eta x_i}\right] = 0$$

$$\frac{\tau N_\tau}{\eta(1+\tau\eta)} - \frac{N_\tau}{\eta^2}\log(1+\tau\eta) + \frac{1}{\eta^2}\sum_{i=1}^{N_\tau}\log(1+\eta x_i) - \frac{1}{\eta}\sum_{i=1}^{N_\tau}\frac{x_i(1+\eta)}{1+\eta x_i} = 0$$

$$\frac{\eta\tau N_\tau}{1+\tau\eta} - N_\tau\log(1+\tau\eta) + \sum_{i=1}^{N_\tau}\log(1+\eta x_i) - \eta(1+\eta)\sum_{i=1}^{N_\tau}\frac{x_i}{1+\eta x_i} = 0 \qquad (8.12)$$

We then obtain $\hat{\eta}$ as the ML estimator of η using the R statistical package. Smith (1985) establishes that the limiting behaviour of the ML depends on the

value of the shape parameter which is the EVI for extremal models. Smith (1985) shows that the classical properties of the ML hold whenever $\xi > -\frac{1}{2}$, but not for $\xi \leq -\frac{1}{2}$ in general for both the GEVD and the GPD. This is important for the stability and asymptotic properties of the ML (Coles, 2001). In general when these regularity conditions fail it is more attractive to use Bayesian methods as they do not depend on these conditions (Beirlant *et al.*, 2004).

8.4.3 Bayesian parameter estimation

In this thesis we use the MDI priors for the GPD and GSPD as they provide maximal data information, easy to implement and that constraints can be built into the prior (Zellner, 1977).

Generalized Pareto Distribution (GPD)

The distribution function of the GPD is given as:

$$W_{\xi,\sigma}(x) = \begin{cases} 1 - \left(1 + \frac{\xi(x-\tau)}{\sigma}\right)^{-\frac{1}{\xi}} & \text{if } \xi > 0, y - \tau > 0 \\ 1 - \exp\left(-\frac{x-\tau}{\sigma}\right) & \text{if } \xi = 0, x - \tau > 0 \\ 1 - \left(1 + \frac{\xi(x-\tau)}{\sigma}\right)^{-\frac{1}{\xi}} & \text{if } \xi < 0, 0 < x - \tau < -\frac{\sigma}{\xi} \end{cases}$$

Maximal Data Information (MDI) prior

The MDI prior (Zellner, 1977), $\pi(\sigma, \xi)$ is defined as follows:

$$\begin{aligned}
\pi(\sigma,\xi) &\propto \exp\left\{E\left[\log f(X|\sigma,\xi)\right]\right\} \\
&= \exp\left\{E\left[\log\left(\frac{1}{\sigma}(1+\xi x)^{-\frac{1}{\xi}-1}\right)\right]\right\} \\
&= \exp\left\{\log\frac{1}{\sigma} - \frac{1+\xi}{\xi}E\left[\log(1+\xi x)\right]\right\} \\
&= \frac{1}{\sigma}\exp\left\{-\frac{1+\xi}{\xi}E\left[\log(1+\xi x)\right]\right\} \\
&= \frac{1}{\sigma}e^{-\xi} \qquad (8.13)
\end{aligned}$$

Likelihood function

The likelihood function is

$$\begin{aligned}
\pi(x|\sigma,\xi) &= \prod_{i=1}^{N_\tau}\frac{1}{\sigma}\left[1+\frac{\xi(x_i-\tau)}{\sigma}\right]^{-\frac{1}{\xi}-1} \\
&= \frac{1}{\sigma^{N_\tau}}\frac{1}{\sigma^{-\frac{N_\tau}{\xi}-N_\tau}}\prod_{i=1}^{N_\tau}\left(\sigma+\xi(x_i-\tau)\right)^{-\frac{1}{\xi}-1} \\
&= \frac{1}{\sigma^{-\frac{N_\tau}{\xi}}}\prod_{i=1}^{N_\tau}\left(\sigma+\xi(x_i-\tau)\right)^{-\frac{1}{\xi}-1} \qquad (8.14)
\end{aligned}$$

Joint posterior distribution

The joint posterior density of ξ and σ is given as follows:

$$\begin{aligned}
\pi(\sigma,\xi|x) &\propto \pi(x|\sigma,\xi)\pi(\sigma,\xi) \\
\pi(\sigma,\xi|x) &\propto \prod_{i=1}^{N_\tau}\frac{1}{\sigma}\left[1+\frac{\xi(x_i-\tau)}{\sigma}\right]^{-\frac{1}{\xi}-1}\pi(\sigma,\xi) \\
&\propto \frac{e^{-\xi}}{\sigma}\frac{1}{\sigma^{-\frac{N_\tau}{\xi}}}\prod_{i=1}^{N_\tau}(\sigma+\xi(x_i-\tau))^{-\frac{1}{\xi}-1} \\
&\propto \frac{e^{-\xi}}{\sigma^{-\left(\frac{N_\tau}{\xi}-1\right)}}\prod_{i=1}^{N_\tau}(\sigma+\xi(x_i-\tau))^{-\frac{1}{\xi}-1} \qquad (8.15)
\end{aligned}$$

Posterior predictive density

The posterior predictive density is given as

$$P(X_0>x_0|x,\tau) \propto \int\int \pi(\sigma,\xi|x)\left\{1+\frac{\xi}{\sigma}(x_0-\tau)\right\}^{-\frac{1}{\xi}}d\sigma d\xi \qquad (8.16)$$

The posterior predictive density given in expression (8.16) cannot be computed analytically. It will be approximated through simulations using the Gibbs sam-

pling procedure.

$$\begin{aligned} P(X_0 > x_0|x,\tau) &= P(X_0 > x_0|x_1,...,x_{N_\tau}) \\ &= E_{(\sigma,\xi)|x}\left(1+\frac{\xi}{\sigma}(x_0-\tau)\right)^{-\frac{1}{\xi}} \\ &= E_{(\sigma,\xi)|x}\left(\frac{\sigma+\xi(x_0-\tau)}{\sigma}\right)^{-\frac{1}{\xi}} \\ &= E_{(\sigma,\xi)|x}\left(\frac{\sigma+\xi(x_0-\tau)}{\sigma}\right)^{-\frac{1}{\xi}} \\ &\approx \frac{1}{N_\tau}\sum_{i=1}^{N_\tau}\left(\frac{\sigma_i+\xi_i(x_0-\tau)}{\sigma_i}\right) \end{aligned} \tag{8.17}$$

Gibbs sampling procedure for GPD

We consider X_i to be GPD distributed, i.e. $X_i \sim$ GPD(σ,ξ). A Gibbs sampler is used to get approximate simulations of the set (ξ,σ) from the joint posterior density given in expression (8.15). The Gibbs sampler considers univariate conditional distributions. This involves simulating σ from its conditional density function given a fixed ξ. The EVI ξ, is then simulated from its conditional density given the selected σ. This process is repeated a large number of times. Let

$$\boldsymbol{\theta} = (\theta_1,\theta_2) = (\sigma,\xi)$$

For each iteration i, each $\theta_j^{(i)}$ is sampled from the conditional distribution given all the other components of $\boldsymbol{\theta}$. Now given the set of values $\left\{\sigma^{(i)},\xi^{(i)}\right\}$, the algo-

rithm proceeds as follows:

$$\begin{aligned} \textbf{Draw } \sigma^{(i+1)} &\sim \pi\left(\sigma|\xi^{(i)}, x\right) \\ \\ \textbf{Draw } \xi^{(i+1)} &\sim \pi\left(\xi|\sigma^{(i+1)}, x\right) \end{aligned} \tag{8.18}$$

Derivation of the quantile function for the GPD

The survival function of the GPD is given as follows:

$$P(X > x|\tau) = \begin{cases} \left(1 + \frac{\xi(x-\tau)}{\sigma}\right)^{-\frac{1}{\xi}} & \textbf{if } \xi > 0, x - \tau > 0 \\ \textbf{exp}\left(-\frac{x-\tau}{\sigma}\right) & \textbf{if } \xi = 0, x - \tau > 0 \\ \left(1 + \frac{\xi(x-\tau)}{\sigma}\right)^{-\frac{1}{\xi}} & \textbf{if } \xi < 0, 0 < x - \tau < -\frac{\sigma}{\xi} \end{cases}$$

$$\begin{aligned} \textbf{Let } p &= P(X > x|X > \tau) \\ \\ p &= 1 - W_{\xi,\sigma} \\ p &= \left(1 + \xi(\frac{x-\tau}{\sigma})\right)^{-\frac{1}{\xi}} \\ \\ 1 + \frac{\xi(x-\tau)}{\sigma} &= p^{-\xi} \\ \\ \frac{\xi(x-\tau)}{\sigma} &= p^{-\xi} - 1 \\ \\ x &= \tau + \frac{\sigma}{\xi}\left(p^{-\xi} - 1\right) \end{aligned}$$

therefore

$$x_p = \tau + \frac{\sigma}{\xi}\left(p^{-\xi} - 1\right), \xi \neq 0, x > \tau \tag{8.19}$$

For details see Beirlant *et al.* (2004).

Generalized Single Pareto Distribution

The MDI prior (Zellner, 1977), $\pi(\eta)$ is defined as follows (Verster and De Waal, 2011):

$$\begin{aligned}
\pi(\eta) &\propto \exp\left\{E\left[\log f(X|\eta)\right]\right\} \\
&= \exp\left\{E\left[\log\left(\frac{1}{1+\tau\eta}\left[1+\frac{\eta(x-\tau)}{1+\tau\eta}\right]^{-\frac{1}{\eta}-1}\right)\right]\right\} \\
&= \exp\left\{E\left[-\log(1+\tau\eta)-\frac{1+\eta}{\eta}\log\left(1+\frac{\eta(x-\tau)}{1+\tau\eta}\right)\right]\right\} \\
&= \frac{e^{-(1+\eta)}}{1+\tau\eta} \\
\pi(\eta) &\propto \frac{e^{-\eta}}{1+\eta\tau} \qquad (8.20)
\end{aligned}$$

The derivation of the MDI prior $\pi(\eta)$ is discussed in detail in Verster and De Waal (2011).

Likelihood function

$$\begin{aligned}
\pi(x|\eta) &= \prod_{i=1}^{N_\tau} w_\eta(x_i) \\
&= \prod_{i=1}^{N_\tau} \frac{1}{1+\eta\tau}\left[1+\frac{\eta(x_i-\tau)}{1+\eta\tau}\right]^{-\frac{1}{\eta}-1} \\
&= \frac{1}{(1+\eta\tau)^{N_\tau}}\prod_{i=1}^{N_\tau}\left(\frac{1+\eta x_i}{1+\eta\tau}\right)^{-\frac{1}{\eta}-1} \qquad (8.21)
\end{aligned}$$

Posterior distribution

$$
\begin{aligned}
\pi(\eta|x) &\propto \pi(x|\eta)\pi(\eta) \\
\pi(\eta|x) &\propto \frac{e^{-\eta}}{(1+\eta\tau)}\frac{1}{(1+\eta\tau)^{N_\tau}}\prod_{i=1}^{N_\tau}\left(\frac{1+\eta x_i}{1+\eta\tau}\right)^{-\frac{1}{\eta}-1} \\
&\propto \frac{e^{-\eta}}{(1+\eta\tau)^{1+N_\tau}}\prod_{i=1}^{N_\tau}\left(\frac{1+\eta x_i}{1+\eta\tau}\right)^{-\frac{1}{\eta}-1} \\
&\propto \frac{e^{-\eta}}{(1+\eta\tau)^{1+N_\tau}}\left[\prod_{i=1}^{N_\tau}\left(\frac{1+\eta x_i}{1+\eta\tau}\right)^{-\frac{1}{\eta}-1}\right] \\
&\propto \left(\frac{e^{-\eta}}{(1+\eta\tau)^{1+N_\tau}}\right)\left(\frac{1}{(1+\eta\tau)^{-\frac{1}{\eta-1}}}\right)^{N_\tau}\left[\prod_{i=1}^{N_\tau}(1+\eta x_i)^{-\frac{1}{\eta}-1}\right] \\
&\propto \frac{e^{-\eta}}{[(1+\eta\tau)^{1+N_\tau}]\left[(1+\eta\tau)^{(-\frac{1}{\eta}-1)N_\tau}\right]}\left[\prod_{i=1}^{N_\tau}(1+\eta x_i)^{-\frac{1}{\eta}-1}\right] \\
&\propto \frac{e^{-\eta}}{(1+\eta\tau)^{-\left(\frac{N_\tau-\eta}{\eta}\right)}}\prod_{i=1}^{N_\tau}(1+\eta x_i)^{-\frac{1}{\eta}-1} \qquad (8.22)
\end{aligned}
$$

Posterior Predictive density

$$
P(X_0 > x_0|x,\tau) = \int \pi(\eta|x)\left[1+\frac{\eta}{1+\eta\tau}(x_0-\tau)\right]^{-\frac{1}{\eta}} d\eta \qquad (8.23)
$$

This posterior predictive density cannot be computed analytically. It will be approximated through simulations.

$$\begin{aligned} P(X_0 > x_0|x_1, ..., x_n) &= P(X_0 > x_0|X_0 > \tau) \\ &= E_{\eta|x}\left(1 + \frac{\eta}{1+\eta\tau}(x_0 - \tau)\right)^{-\frac{1}{\eta}} \\ &\approx \frac{1}{m}\sum_{j=1}^{m}\left(1 + \frac{\eta_j(x_0 - \tau)}{1+\eta_j\tau}\right) \end{aligned} \tag{8.24}$$

Gibbs Sampler (for the GSPD)

$$\text{Draw } \eta^{(i+1)} \sim \pi(\eta|x)$$

Derivation of the quantile function for GSPD

The survival function of the GSPD is given as:

$$\begin{aligned} 1 - W_\eta(x) &= P(X > x|X > \tau) \\ &= \left\{1 + \frac{\eta}{1+\tau\eta}(x - \tau)\right\}^{-\frac{1}{\eta}}, \eta \neq 0, x > \tau \end{aligned}$$

Let $p = P(X > x|X > \tau)$ then

$$\begin{aligned} p &= \left\{1 + \frac{\eta}{1+\tau\eta}(x-\tau)\right\}^{-\frac{1}{\eta}} \\ p^{-\eta} &= 1 + \frac{\eta}{1+\tau\eta}(x-\tau) \\ x &= \tau + \frac{1+\tau\eta}{\eta}(p^{-\eta}-1) \end{aligned}$$

Then

$$x_p = \tau + \frac{1+\tau\eta}{\eta}(p^{-\eta}-1), \eta \neq 0, x > \tau \tag{8.25}$$

8.5 Analysis of extreme daily increases in peak electricity demand

8.5.1 Data Set

Modelling of daily (interday) increases in peak electricity demand is done using data for the period 2000 to 2010. The distribution of the inter-day changes is nonnormal as evidenced by the skewness, kurtosis and Jarque-Bera test given in Table 8.1. A visual inspection of Figure 8.2 shows that the largest interday

Table 8.1: Descriptive statistics of inter-day changes.

Mean	Standard deviation	Skewness	Kurtosis	Jarque-Bera
2.8475	1251.752	0.9299	3.5152	662.66 (0.0000)

change in peak electricity demand over the sampling period was experienced in 2007.

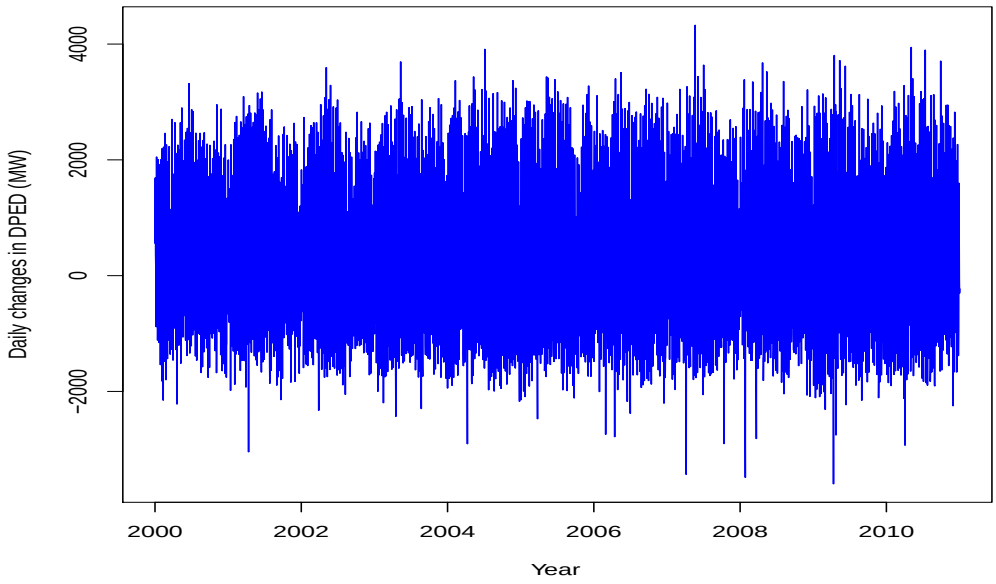

Figure 8.2: Inter-day changes in DPED.

8.5.2 Fitting the Generalized Pareto Distribution

Let $X_1, X_2, ..., X_n$ be a sequence of daily increases in peak electricity demand. The increase in peak demand is relative to the previous day. In order to extract upper extremes from this sequence we take the exceedances over a predetermined high threshold τ. If we let x_t to be DPED on day t and x_{t-1} DPED on day $t-1$, then we define daily increase in peak demand on day t, as $z_t = \max(x_t - x_{t-1}, 0), t = 1, ..., n$. In this definition decreases in demand are ignored. Extremely large daily increases pose challenges to system operators of power utility companies in the dispatching and scheduling of electricity. An initial threshold is set at zero after taking the first differences of DPED and we use observations above zero.

Over the sampling period the highest daily peak electricity demand of 37 158MW was experienced on Thursday 24 May 2007. Plots of hourly electricity demand for Thursday 24 May 2007 and daily peak electricity demand for the month of

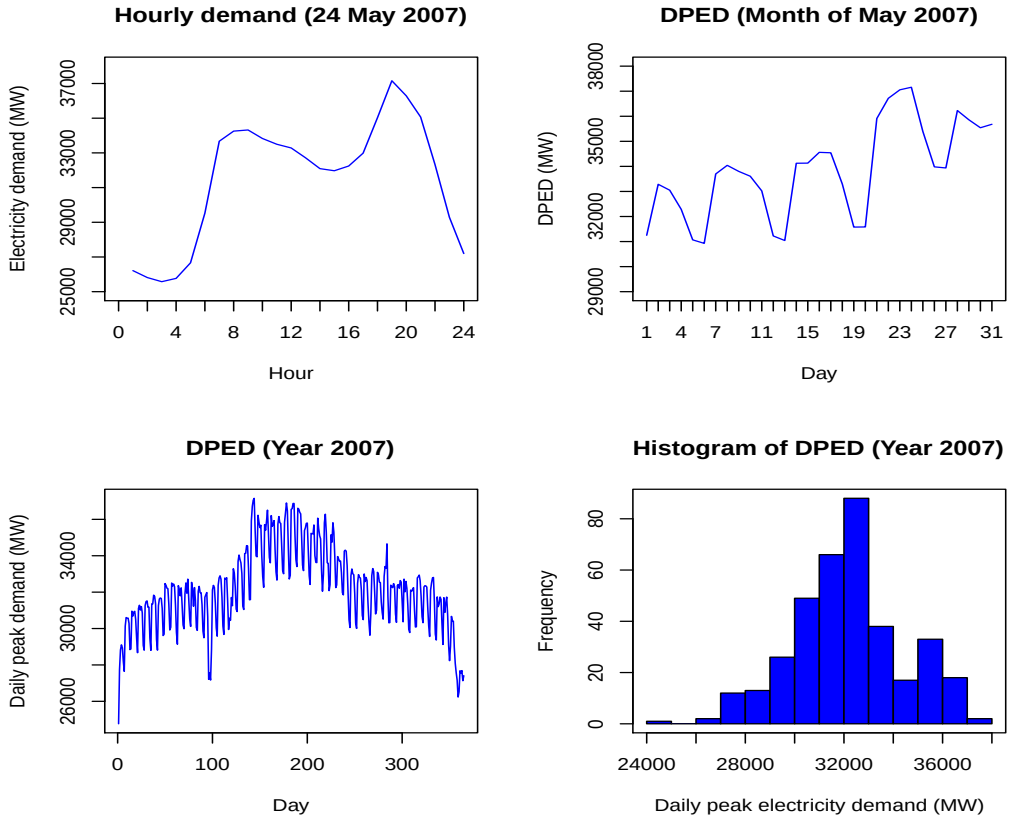

Figure 8.3: The plots are for the year 2007 only. (a) Top left panel: Hourly electricity demand for Thursday 24 May 2007. There are two peaks, one in the morning and the the other is in the evening (b) Top right panel: Daily peak electricity demand for the month of May 2007. The peak demand of 37158MW was on 24 May (c) Bottom left panel: Daily peak electricity demand for the year 2007 and (d) Bottom right panel: Histogram of DPED . The distribution is slightly skewed to the left.

May 2007 are shown in Figure 8.3. Figure 8.4 shows daily increases in peak electricity demand together with the probability density for the whole data set, from 2000 to 2010. The density is estimated using kernel density (Silverman, 1986). The density plot shows that the data is not normally distributed and has a long upper tail suggesting that a heavy tailed distribution could be a good fit to the data. The QQ plot in Figure 8.3 shows evidence of a distribution which has heavier tails compared to those of a normal distribution. This calls for use of extreme value theory distributions to model the upper tail of the distribution

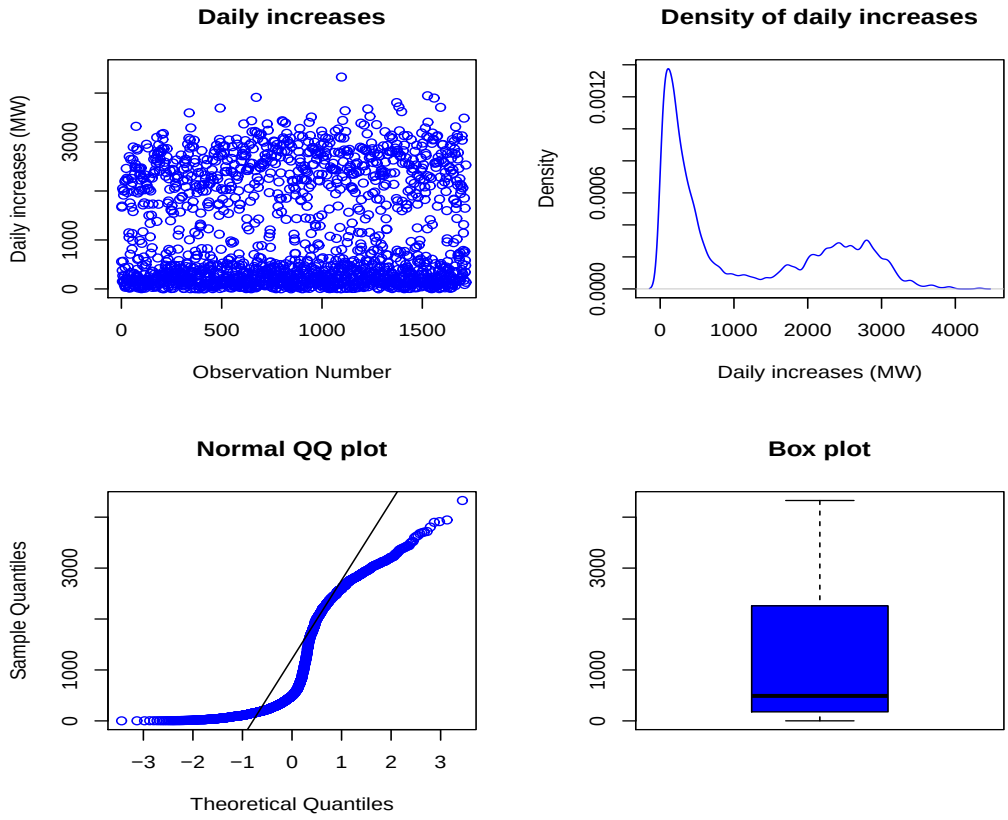

Figure 8.4: The plots are for the whole data set (2000 to 2010). (a) Top left panel: Plot of daily increases, (b) Top right panel: Probability density of daily increases. The distribution is heavily right skewed, (c) Bottom left panel: Normal QQ plot of daily increases and (d) Bottom right panel: Box plot.

of the daily increases in peak electricity demand. A Pareto quantile plot is used to obtain the threshold. The Pareto quantile plot is defined as the scatter plot of the following points: $(-\log(1-p_i), \log z_i)$ where $p_i = \frac{i}{n+1}$ and z_i are the daily increases in peak electricity demand, for all $i = 1, ..., n$ (Beirlant *et al.*, 2004). In a Pareto quantile plot the observation on the vertical axis (y-axis) where the plot starts to follow a linear pattern is taken as the threshold. In this case $\tau = \exp(8) = 2980.958$. The Pareto quantile plot is shown in Figure 8.5 while Figure 8.6 shows the mean excess plot. The plot starts to follow a straight line around $\tau = 3000$MW. We shall use the threshold determined by the Pareto quantile plot (Beirlant *et al.*, 1996, 2004). We now consider the observations

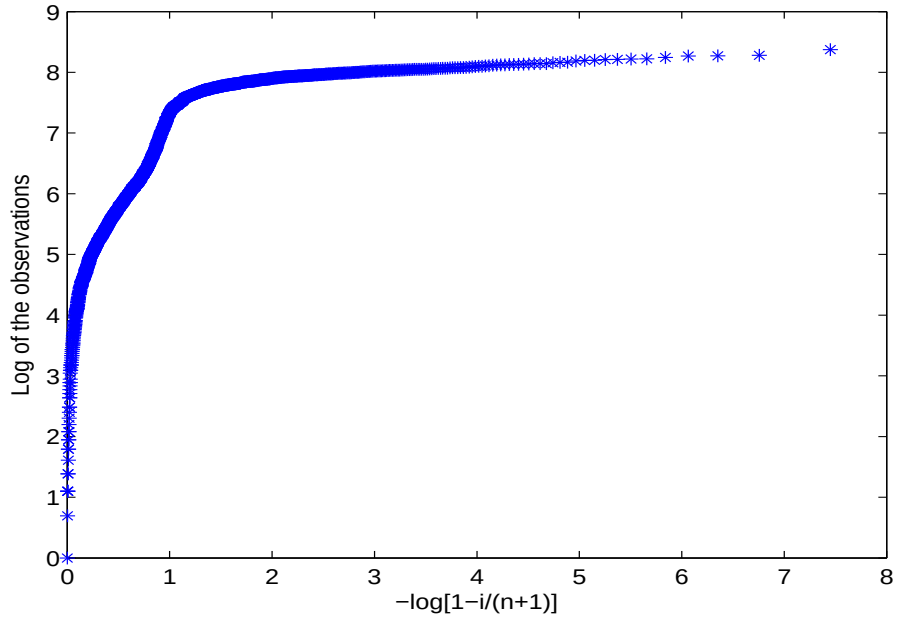

Figure 8.5: Pareto quantile plot of the positive observations. The positive observations are the daily increases in peak electricity demand.

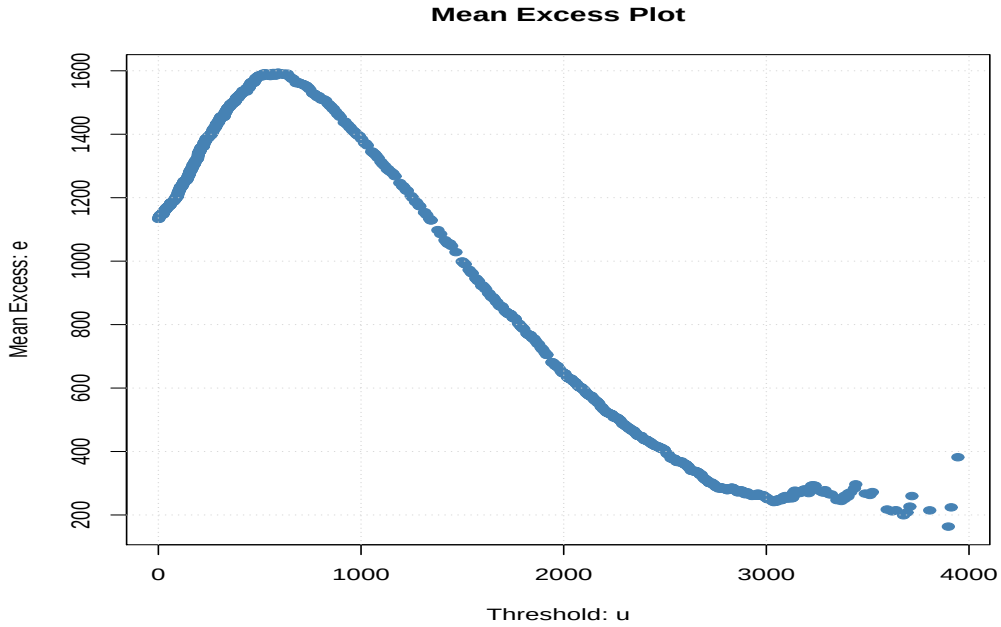

Figure 8.6: Mean excess plot of the daily increases in peak electricity demand where u denotes the threshold τ.

greater than τ to be Generalized Pareto distributed. The parameters ξ and σ are then estimated by simulating from the joint posterior distribution given in equation (8.15).

Figure 8.7 shows a scatter plot of the simulated ξ's against σ's for the data. The means of the simulated ξ's and σ's are calculated as -0.0079 and 269.164 respectively and are considered as the estimates of ξ and σ. The negative sign of ξ, the extreme value index (EVI) indicates that the data has an upper bound which implies that it belongs to the Weibull class. This is probably due to the fact that the interday increases in peak electricity demand between any two consecutive days cannot exceed the maximum in daily peak electricity demand between the two days concerned and also demand cannot exceed supply. The ML estimates of the parameters ξ and σ are -0.048(0.1014) and 274.609(39.073) respectively with the standard errors in parentheses. The ML estimates are comparable to the Bayesian estimates.

The empirical cumulative distribution function (cdf) and theoretical cdf of the GPD with the estimated parameter values is constructed for observations above the threshold and shown in Figure 8.8. From Figure 8.8 it is evident that the theoretical cdf is close to the empirical cdf, indicating the appropriateness of using the GPD to model the observations above the threshold. A QQ plot can be constructed to indicate the goodness of fit. The theoretical quantiles from the GPD with the estimated parameter are plotted against the empirical quantiles. If the plot lies on the 45^0 line it indicates a good fit as is the case in Figure 8.9.

The diagnostic plots given in Figure 8.10 show that the Weibull class of distributions is a good fit to daily increases in DPED.

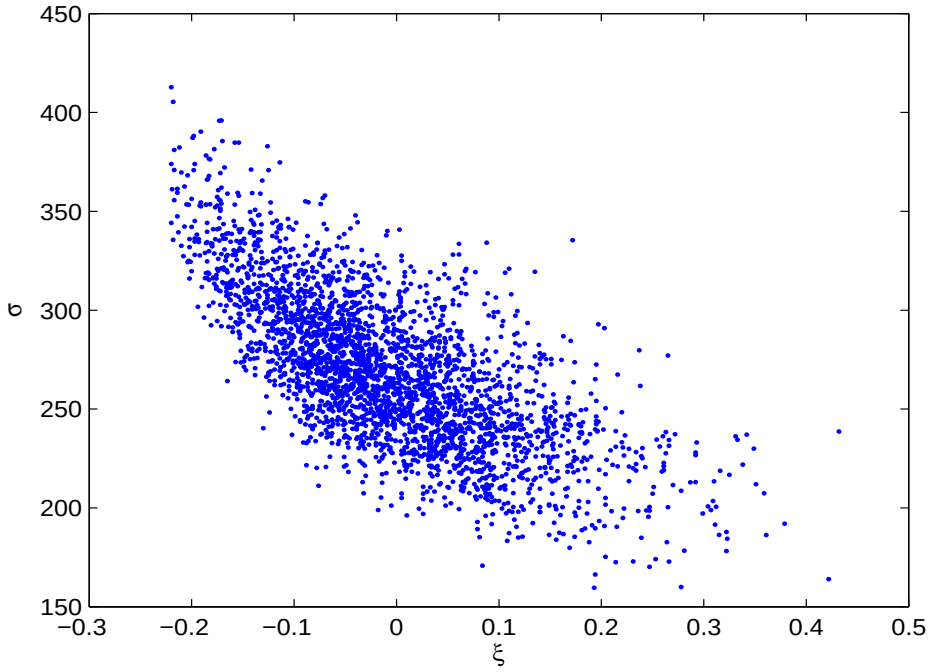

Figure 8.7: Plot of 3000 simulated (ξ, σ) values simulated through Gibbs sampler.

8.5.3 Fitting the Generalized Single Pareto Distribution

The GSPD has one parameter instead of two as in the GPD case. The simultaneous simulations of two parameters can become quite difficult and time consuming. We consider again the threshold obtained previously $\tau = \exp(8) = 2981$MW. For different values of η the posterior of η is constructed in Figure 8.11. The mode of the posterior is an estimate of η which is found to be $\hat{\eta} = 0.08$. Figure 8.12 shows the Hill estimate of the EVI of the daily (interday) increases in peak electricity demand observations above the threshold of 2981MW together with approximate 95% confidence bands.

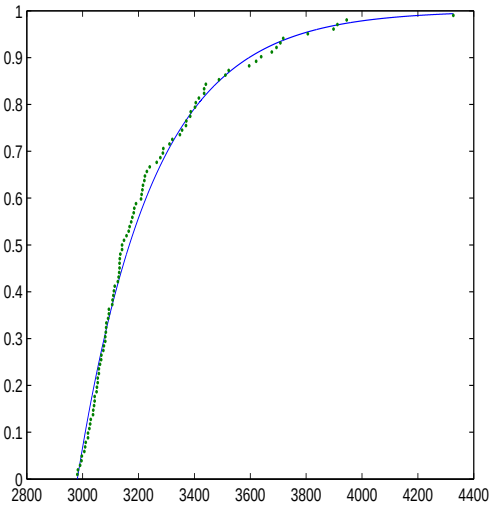

Figure 8.8: Graphical plot of the empirical cdf (dotted curve) and the cdf of the GPD (solid curve) on the exceedances above the threshold $\tau = 2981$. The $x-$axis represents the exceedances and the $y-$axis the cumulative probabilities.

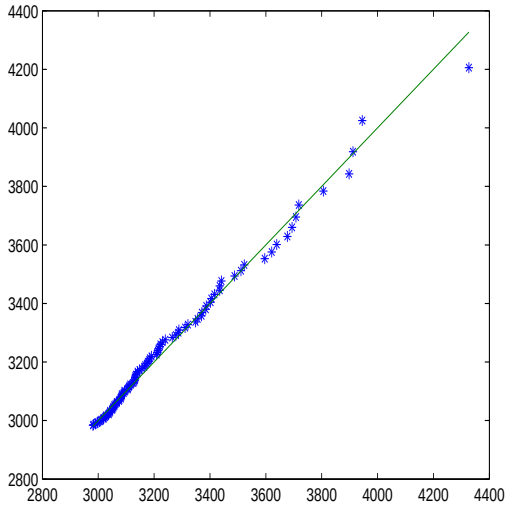

Figure 8.9: A QQ plot of sample data (interday increases above the threshold, $\tau = 2981$MW) versus a GPD. The horizontal axis (x-axis) represents the standard theoretical quantiles while the empirical quantiles are plotted on the vertical axis (y-axis).

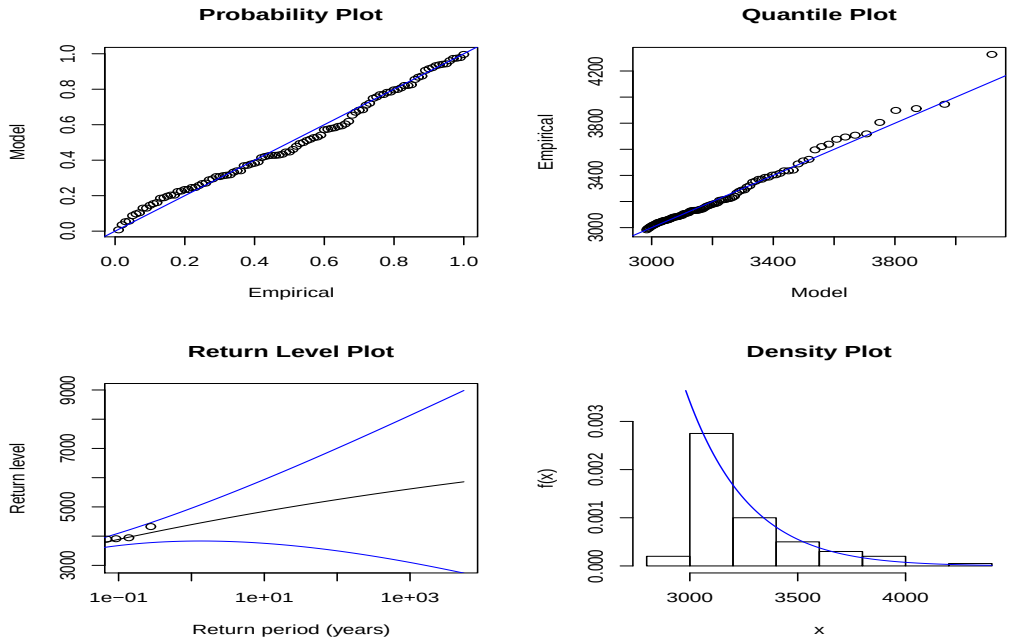

Figure 8.10: Diagnostic plots: (a) Top left panel: Probability plot (b) Top right panel: Quantile plot (c) Bottom left panel: Return level plot and (d) Bottom right panel: Density plot. The distribution is right skewed.

The empirical cdf and theoretical cdf of the GSPD with the estimated parameter values $\hat{\eta} = 0.08$ is constructed for observations above the threshold and shown in Figure 8.13. From Figure 8.13 it is evident that the theoretical cdf is close to the empirical cdf, indicating the appropriateness of using the GSPD to model the observations above the threshold. The QQ plot in Figure 8.14 indicates a good fit of the GSPD to the exceedances above the threshold, $\tau = 2981$MW.

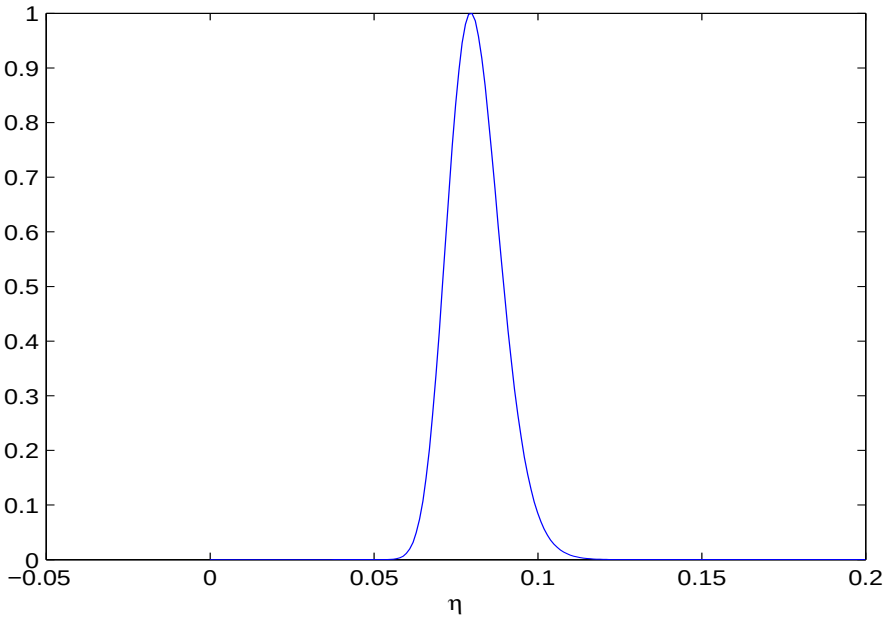

Figure 8.11: The posterior density of η. The x-axis represents the range of values of the parameter η while the y-axis represents the cumulative probabilities.

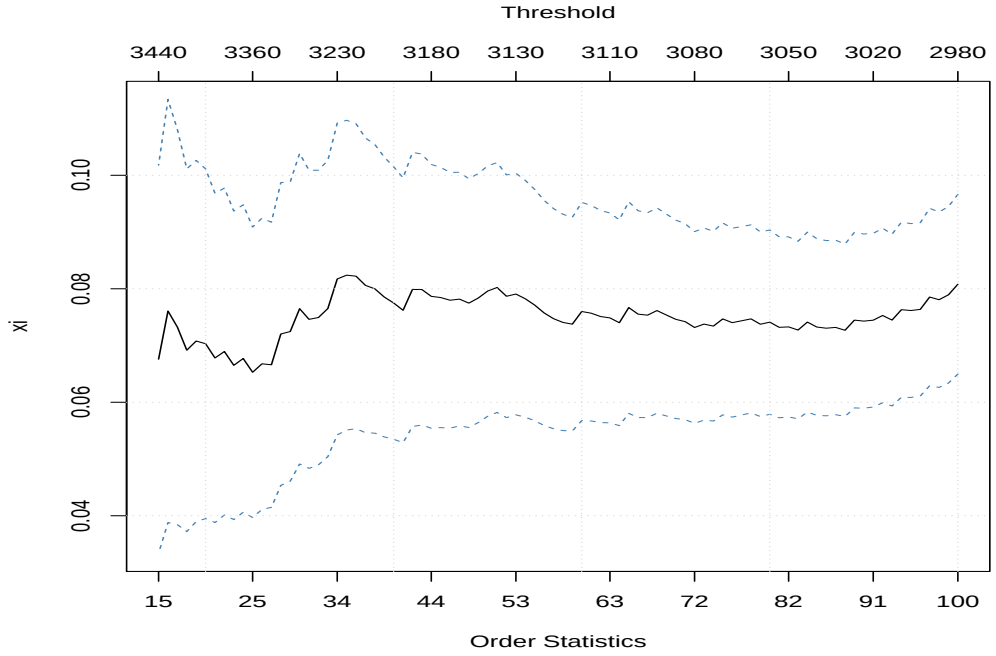

Figure 8.12: Plot of the Hill estimate of the tail index of the daily increases in peak electricity demand observations above the threshold of 2981MW together with approximate 95% confidence bands.

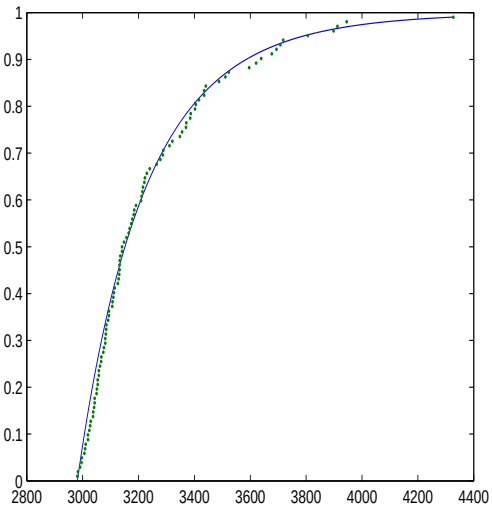

Figure 8.13: Graphical plot of the empirical cdf (dotted curve) and the cdf of the GSPD (solid curve) on the exceedances above the threshold $\tau = 2981$MW. The $x-$axis represents the exceedances and the $y-$axis the cumulative probabilities.

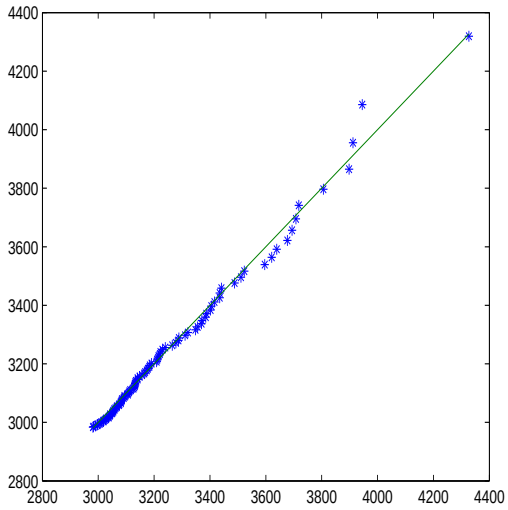

Figure 8.14: A QQ plot of sample data (interday increases above the threshold, $\tau = 2981$MW) versus a GSPD. The horizontal axis (x-axis) represents the standard theoretical quantiles while the empirical quantiles are plotted on the vertical axis (y-axis).

8.5.4 Comparative analysis and discussion

Some of the posterior predicted tail probabilities for various extreme daily increases in peak electricity demand are given in Table 8.2. For the GPD, 1000 ξ's and σ's are simulated and substituted into posterior predictive density given in equation (8.17). Similarly for the GSPD, 1000 η's are simulated and substituted into Equation (8.24). Empirical results show that both the GSPD and the GPD

Table 8.2: Posterior predictive tail probabilities ($\tau = 2981$).

Extreme daily increases (x_0)	$P(X_0 > x_0\|x, \tau)$ GPD	$P(X_0 > x_0\|x, \tau)$ GSPD
3400	0.2004	0.1966
3700	0.0628	0.0688
4000	0.0195	0.0261
4300	0.0060	0.0106

are a good fit to the data. However the QQ plot of the GSPD given in Figure 8.14 incorporates most extreme observations in the tail slightly better than the GPD. One of the main advantages of the GSPD is the estimation of only one parameter instead of two as is the case with GPD. The plot of $P(X_0 > x_0|x, 2981)$ for GPD and GSPD distributions is given in Figure 8.15.

The plot of 100 exceedances (extreme daily increases) above the threshold of 2981MW over the sampling period, 2000 to 2010 is given in Figure 8.16, top panel. Over this sampling period the maximum extreme daily increase is 4327MW and the minimum is 2983MW. The highest extreme daily increase of 4327MW (see Figure 8.16) is the difference between DPED on 21 May 2007 which was 35910MW with that of 20 May 2007 which was 31583MW. This was probably caused by extreme weather conditions. The probability density function of the exceedances, given in Figure 8.16, bottom panel, is estimated using kernel density estimation (Silverman, 1986). The empirical density is skewed to the right and is heavy tailed. All the exceedances of the extreme daily increases occur on Mondays as expected since Sunday has the lowest demand of

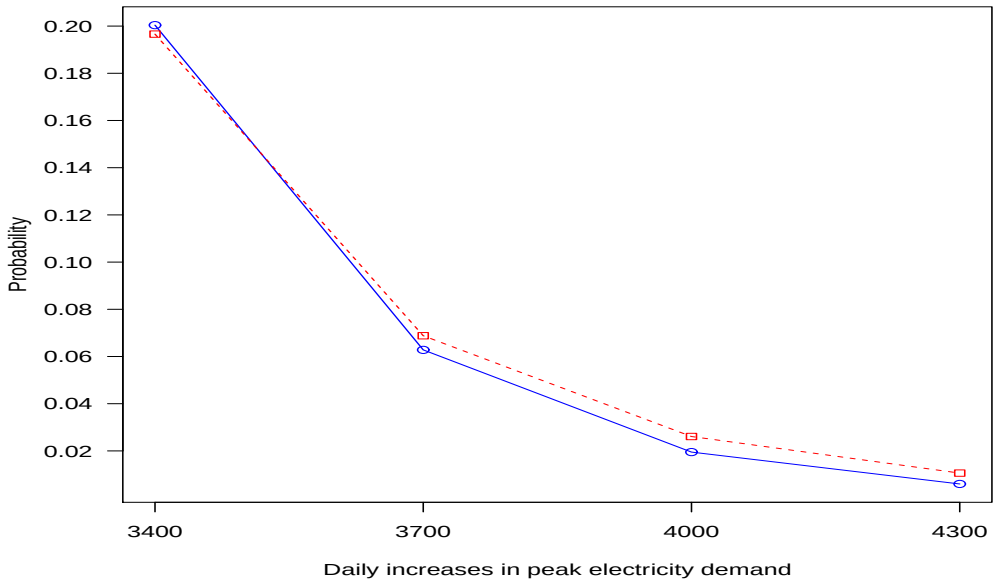

Figure 8.15: Plot of $P(X_0 > x_0|x, 2981)$ for GPD (solid line) and GSPD distribution (dashed line).

electricity. The maximum extreme daily increase is on Monday 21 May 2007. It is interesting to note that on Thursday 24 May 2007 South Africa had the highest daily peak demand over the sampling period 2000 to 2010, which is the same week the country experienced the maximum extreme daily increase. Figure 8.17 shows the density function of the GSPD with threshold of 2981MW and $\hat{\eta} = 0.08$. The density is heavy tailed and skewed to the right, indicating that it can be used to model exceedances above the threshold, $\tau = 2981$MW.

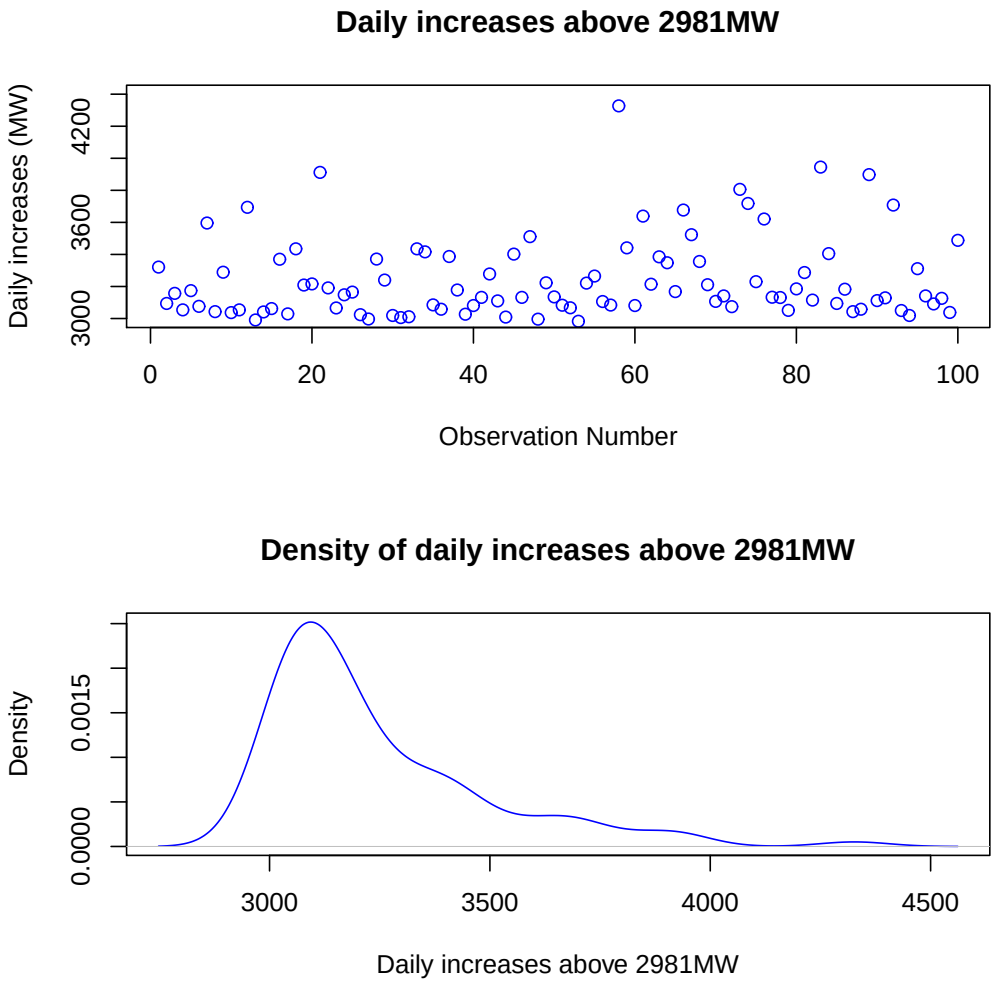

Figure 8.16: (a) Top panel: Plot of extreme daily increases (exceedances) above the threshold $\tau = 2981$MW and (b) Bottom panel: Empirical density function of extreme daily increases (exceedances) above the threshold $\tau = 2981$MW.

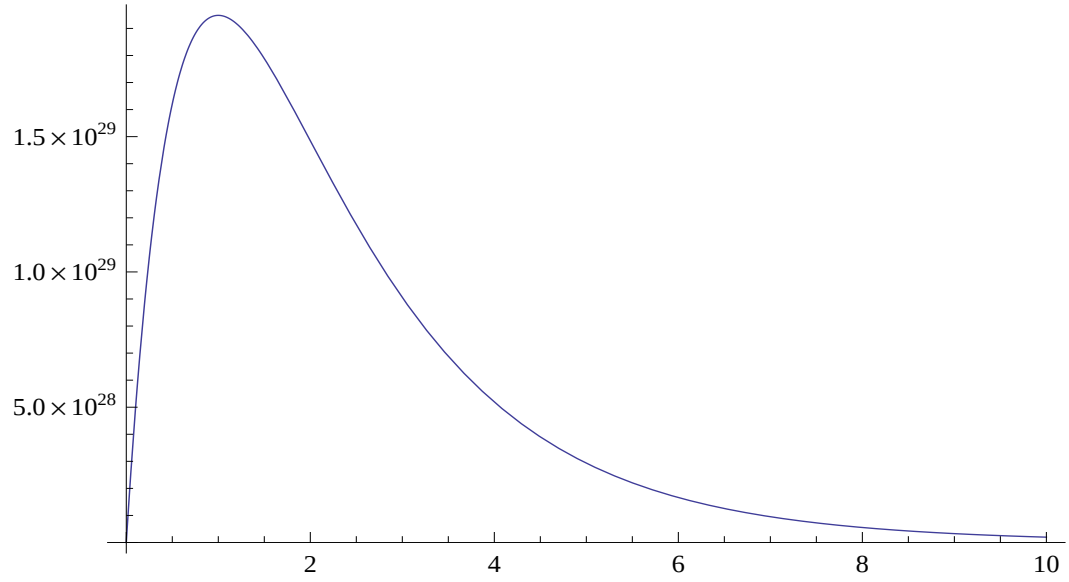

Figure 8.17: Probability density function of the GSPD with threshold of 2981MW and $\hat{\eta} = 0.08$.

The frequency of the occurrence of exceedances per month over the sampling period given in Table 8.3 (Freq represents frequency) shows that May had the highest frequency of 24 with the least being October with a frequency of 2. The months May, June and July have the highest frequencies above the threshold. This is an indication that South Africa experiences a higher peak electricity demand in winter than in summer. This is consistent with results found in chapters 3, 4 and 5. Those days with positive daily changes above the threshold of 2981MW will then be declared as critical peak days and an appropriate critical peak pricing tariff will then be used, e.g. the month of May had a total of 24 critical peak days over the sampling period, years 2000 to 2010. The

Table 8.3: Monthly frequency of extreme daily increases above, $\tau = 2981$MW.

Month	January	February	March	April	May	June
Freq	5	6	7	10	24	12
Month	July	August	September	October	November	December
Freq	13	8	5	2	4	4

bar chart of the frequency of occurrence of extreme daily increases above the threshold, $\tau = 2981$MW is given in Figure 8.18.

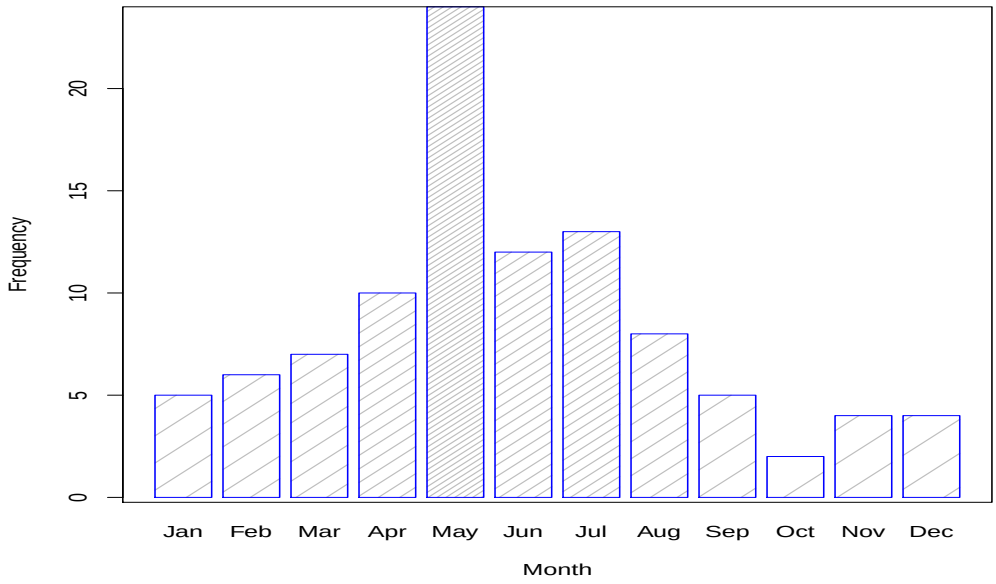

Figure 8.18: Bar chart of the monthly frequency of occurrence of extreme daily increases above the threshold, $\tau = 2981$MW.

A summary of the frequency analysis of the exceedances by year is given in Table 8.4. Over the sampling period the year 2004 has the largest number of exceedances. This analysis is important for prediction of return period of this largest number of exceedances. This will help decision makers in long term strategic planning which will include capacity expansion to ensure that there is enough energy during periods of abnormal peak demand. The bar chart of

Table 8.4: Yearly frequency of extreme daily increases above the threshold, $\tau = 2981$MW.

Year	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Freq	1	4	5	5	14	13	10	10	10	7	10

the frequency of occurrence of exceedances by year is given in Figure 8.19.

Eskom has put in place an integrated demand management programme for all sectors of the economy (Eskom Integrated Demand Management, (n.d.)). This is meant to shift the load curve to lower demand levels (Cousins, 2009). This

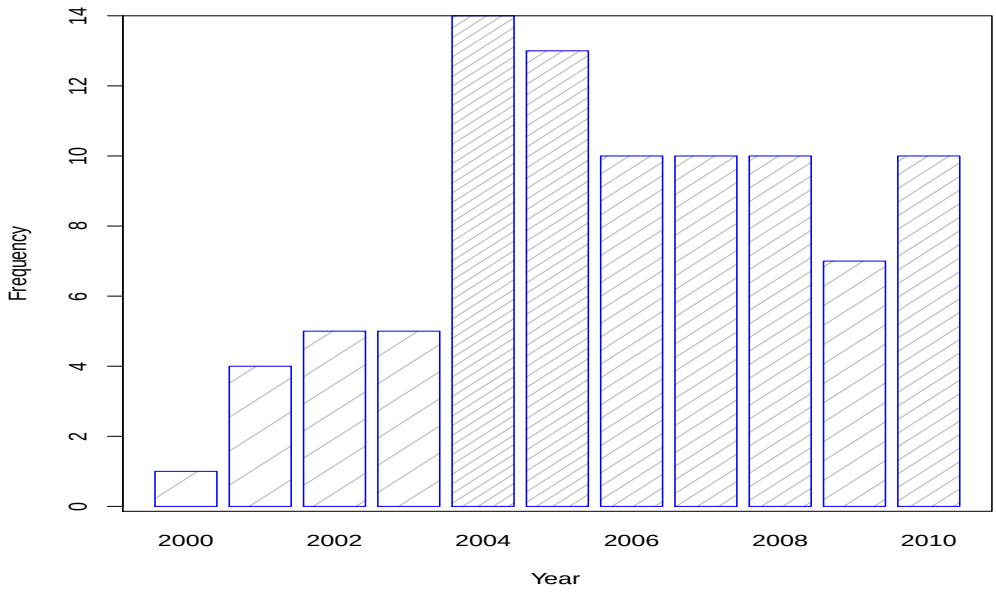

Figure 8.19: Bar chart of the yearly frequency of occurrence of interday increases in peak electricity demand.

is in line with the International Energy Agency (IEA) Demand-side Management (DSM) programme which seeks to reduce the demand peaks, shift the loads between times of day or even seasons and to better utilize existing power resources and reduce overall electricity demand (International Energy Agency Demand Side Management, 2012).

In this Section 8.5, we discussed the modelling and analysis of extreme daily (interday) increases in peak electricity demand. This modelling approach helps to quantify and assess the frequency of occurrence of extreme interday changes in peak electricity demand. This is important to system operators who have to schedule and dispatch electricity to end users, including ensuring grid stability even during periods of extreme peak electricity demand. The modelling approach also assists decision makers in determining critical peak days. Another important and interesting study involves an analysis of the same day of the week increases in extreme peak electricity demand. This analysis will be

discussed in Section 8.6.

8.6 Analysis of same day of the week increases in daily peak electricity demand

8.6.1 Data description

We define same day of the week (day to day) increases in peak electricity demand as follows: Let x_t be DPED on day t and x_{t-7} DPED on day $t-7$, then the same day of the week increase in peak electricity demand on day t is defined as

$$y_t = \max(x_t - x_{t-7}, 0), t = 1, ..., n. \tag{8.26}$$

In this definition we only consider positive increases in peak demand. Electricity demand varies from day to day depending on the day of the week. The demand of electricity is generally higher during the week (Monday to Friday) than during the weekend (Saturday and Sunday). A summary of the demand indices for each day of the week is shown in Table 8.5 while the graphical plot is given in Figure 8.19. The indices are calculated as follows: we calculate the average DPED for each day of the week and divide by the overall average. Day 7 which is Sunday has the lowest weekly seasonal index of 93.793% showing that on average the demand of electricity is 6.027% below average demand. The highest index of 102.883% on day 2 which is Tuesday indicates that there is an above average demand of 2.883%.

Table 8.5: DPED demand indices.

Number	Day	Weekly seasonal index (unit of index (%))
1	Monday	102.64
2	Tuesday	102.883
3	Wednesday	102.878
4	Thursday	102.854
5	Friday	99.879
6	Saturday	95.073
7	Sunday	93.793

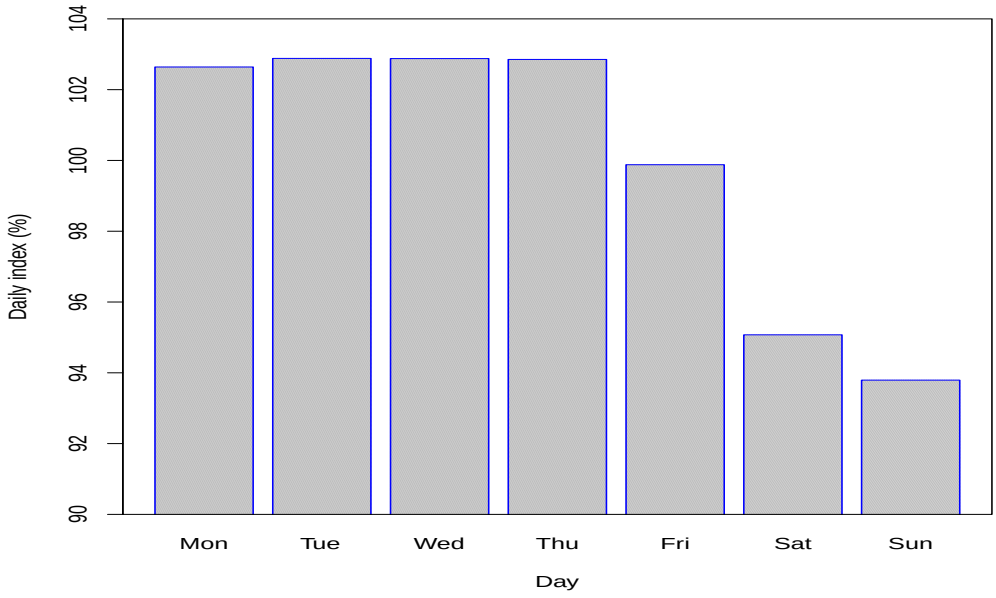

Figure 8.20: Weekly seasonal index plot of DPED.

The plot in Figure 8.20 shows that the same day of the week increases (top left panel) in peak electricity demand is volatile. This volatility may be due to a range of uncertainties, including population growth, changing technology, economic conditions and activity, prevailing weather conditions as well as the general day to day variations in individual usage of electricity (Hyndman and Fan, 2010). Both the histogram (top right panel) and probability density (bottom left panel) of the same day of the week increases in Figure 8.20 shows a heavy tail and right skewed.

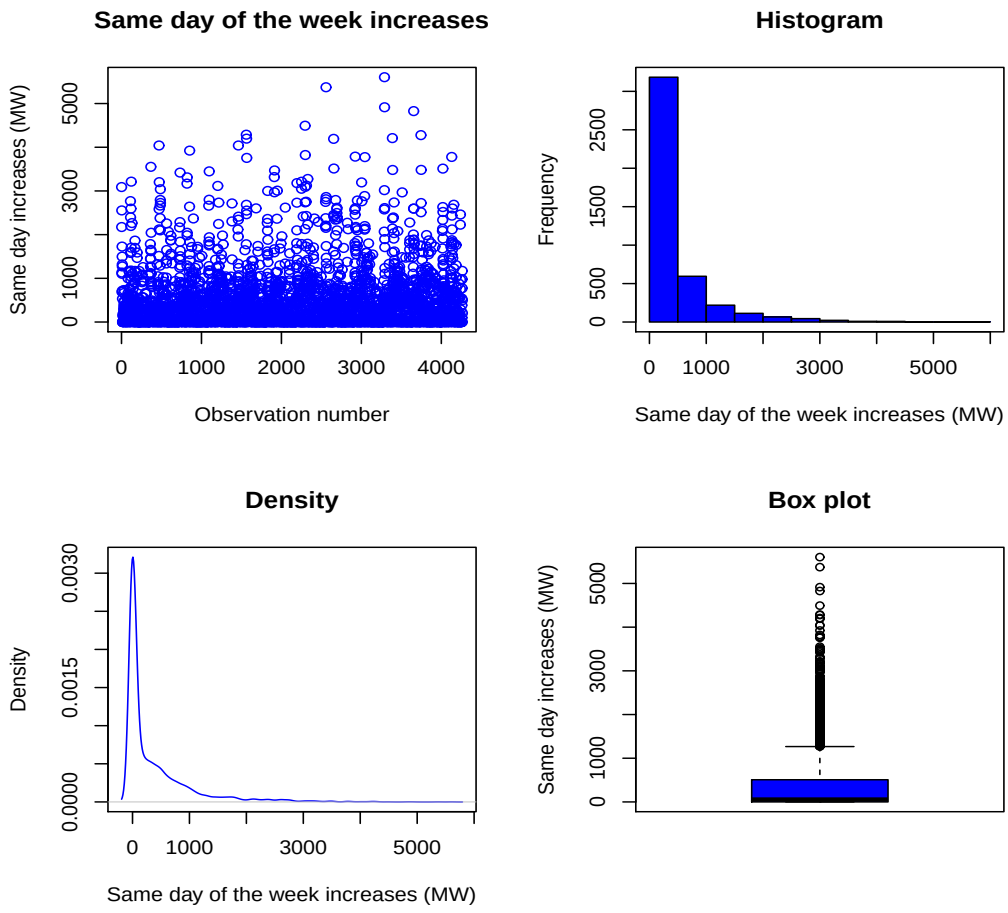

Figure 8.21: Same day of the week increases in peak electricity demand: (a) Top left panel: Time series plot (b) Top right panel: Histogram (c) Bottom left panel: Density plot. The distribution is heavily right skewed and (d) Bottom right panel: Box plot.

8.6.2 Analysis of future observations

Threshold selection

The Pareto quantile plot given in Figure 8.22 is used to obtain the threshold on the same day of the week increases in peak electricity demand. The threshold is $\tau = \exp(8) = 2980.958$. There are 44 exceedances. The exceedances, z_t are those y_t's such that $y_t > \tau$, that is $z_t = y_t - \tau > 0 = \max(x_t - x_{t-7}, 0) - \tau > 0$. The values $z_t = y_t - \tau > 0$ are the excesses over τ. Figure 8.23 shows the mean excess plot of the same day of the week increases in peak electricity demand.

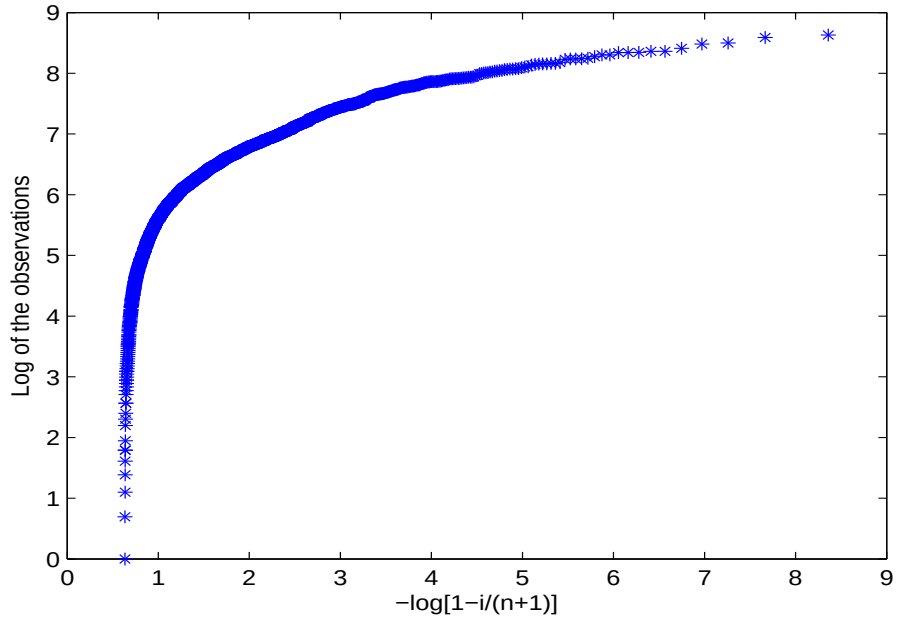

Figure 8.22: Pareto quantile plot on the same day of the week increases in peak electricity demand.

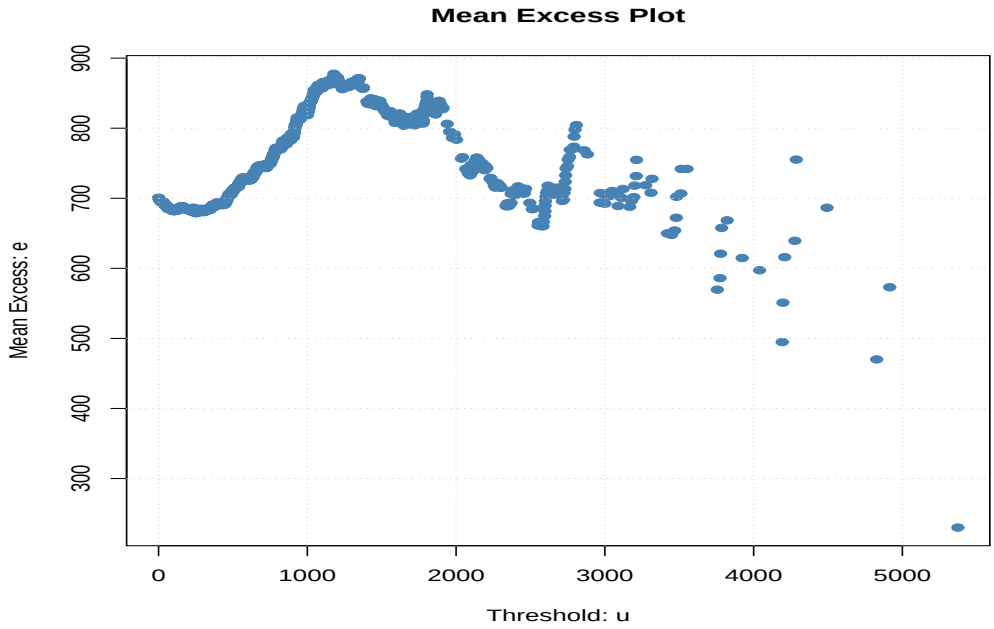

Figure 8.23: Mean excess plot of the same day of the week increases in peak electricity demand.

8.6.3 Analysis of future observations: Fitting the GPD

Figure 8.24 shows a scatter plot of the simulated ξ's against σ's for the data. The means of the simulated ξ's and σ's are calculated as -0.0272 and 740.5557 respectively and are considered as the estimates of ξ and σ. The value of $\hat{\xi}$ is negative indicating the existence of an upper bound. This is probably due to the fact that the same day of the week increase in peak electricity demand cannot exceed supply. Maximum electricity supply is fixed over the short run. The ML estimates of the parameters ξ and σ are -0.1411(0.1658) and 795.551(177.556) respectively with the standard errors in parentheses. The QQ plot of the same day of the week increase in peak electricity demand above the threshold, $\tau = 2981$MW given in Figure 8.25 shows that the empirical distribution is a good fit to the GPD. The diagnostic plots given in Figure 8.26 show that the Weibull class of distributions is a fairly good fit to the same day of the week increases in DPED.

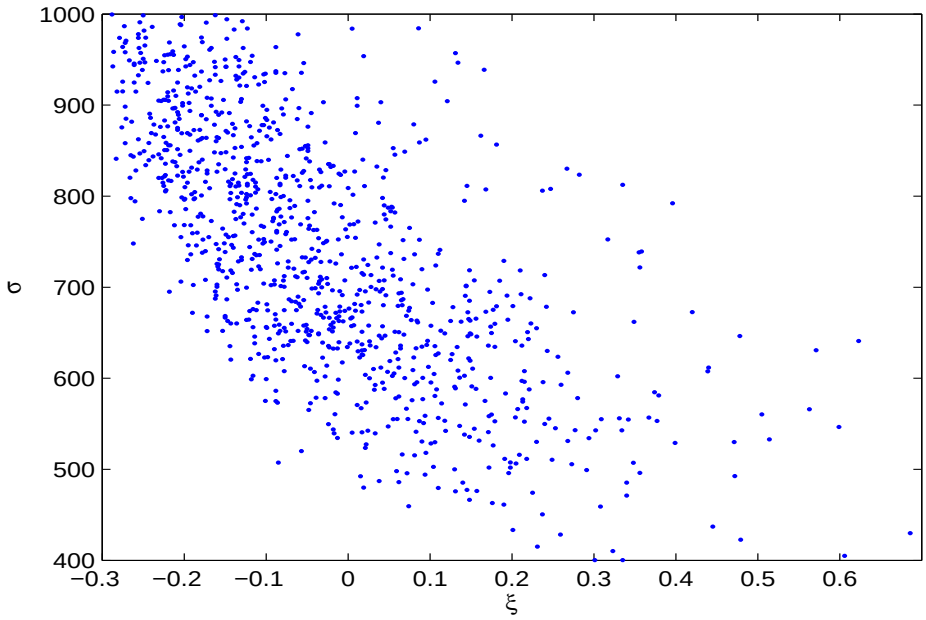

Figure 8.24: Plot of 1000 simulated (σ, ξ) values simulated through Gibbs sampler.

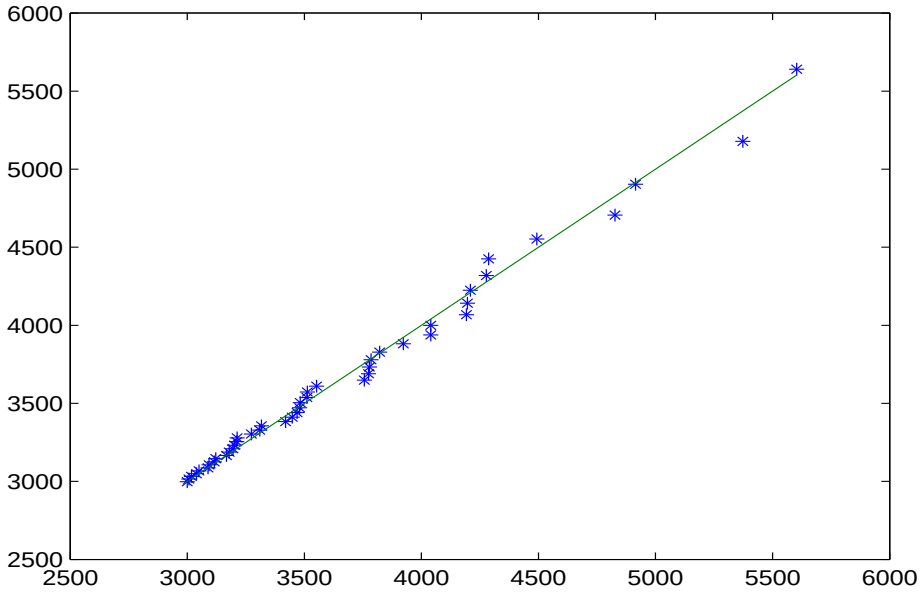

Figure 8.25: A QQ plot of sample data (same day of the week increases in peak demand above the threshold, $\tau = 2981$MW versus a GPD. The horizontal axis (x-axis) represents the standard theoretical quantiles while the empirical quantiles are plotted on the vertical axis (y-axis).

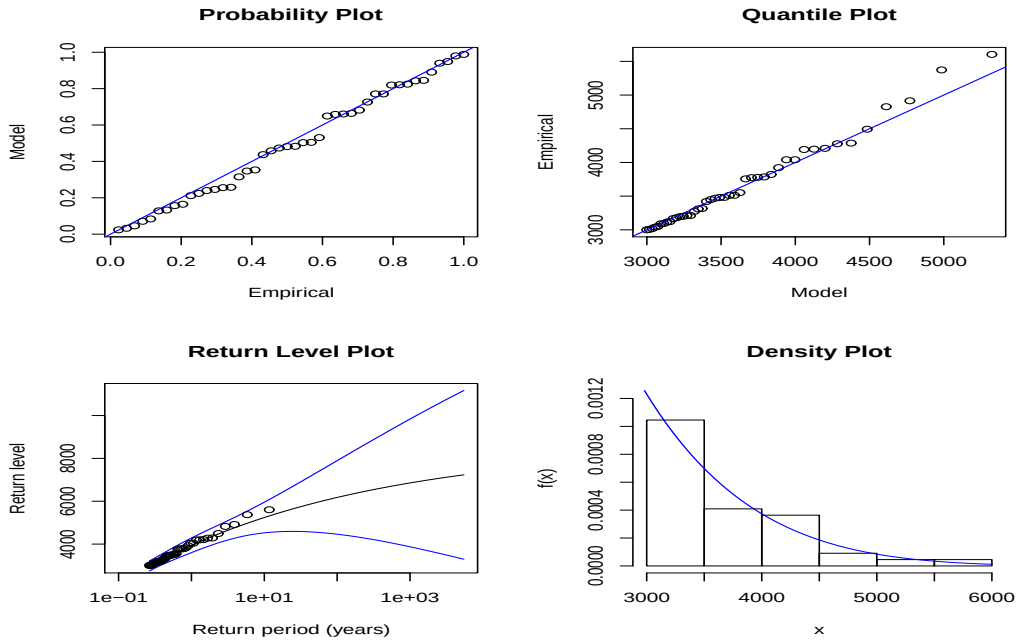

Figure 8.26: Diagnostic plots: (a) Top left panel: Probability plot (b) Top right panel: Quantile plot (c) Bottom left panel: Return level plot and (d) Bottom right panel: Density plot. The distribution is left skewed.

8.6.4 Analysis of future observations: Fitting the GSPD

The threshold is again considered as $\tau = \exp(8) = 2981$ obtained from the Pareto quantile plot in Figure 8.22. The posterior density of η is given in Figure 8.27. The mode of the posterior density is taken as an estimate of η and is $\hat{\eta} = 0.1909$. Figure 8.28 shows a plot of the Hill estimate of the tail index for same day of the week increases in peak electricity demand observations above the threshold of 2981MW together with approximate 95% confidence bands.

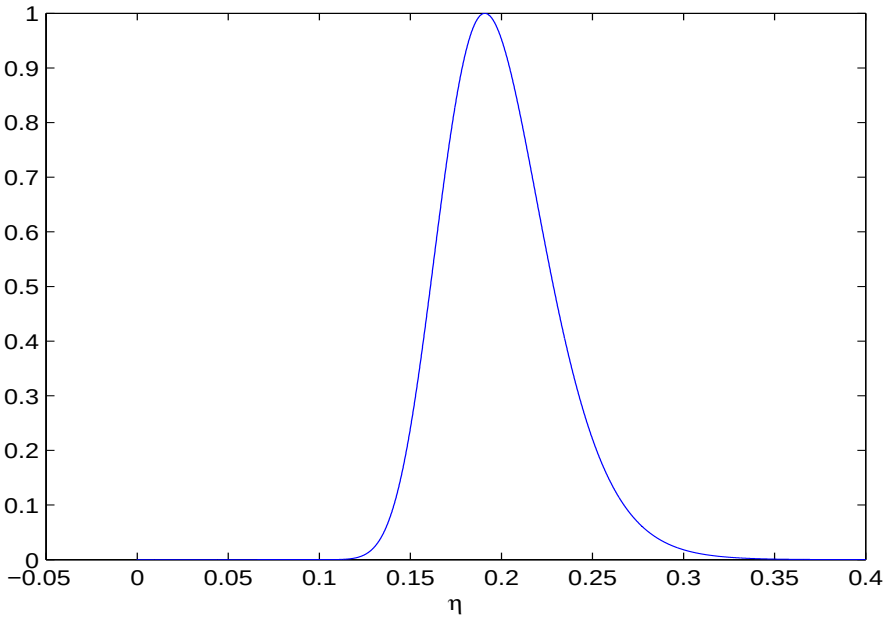

Figure 8.27: The posterior density of η. The x-axis represents the range of values of the parameter η while the y-axis represents the cumulative probabilities.

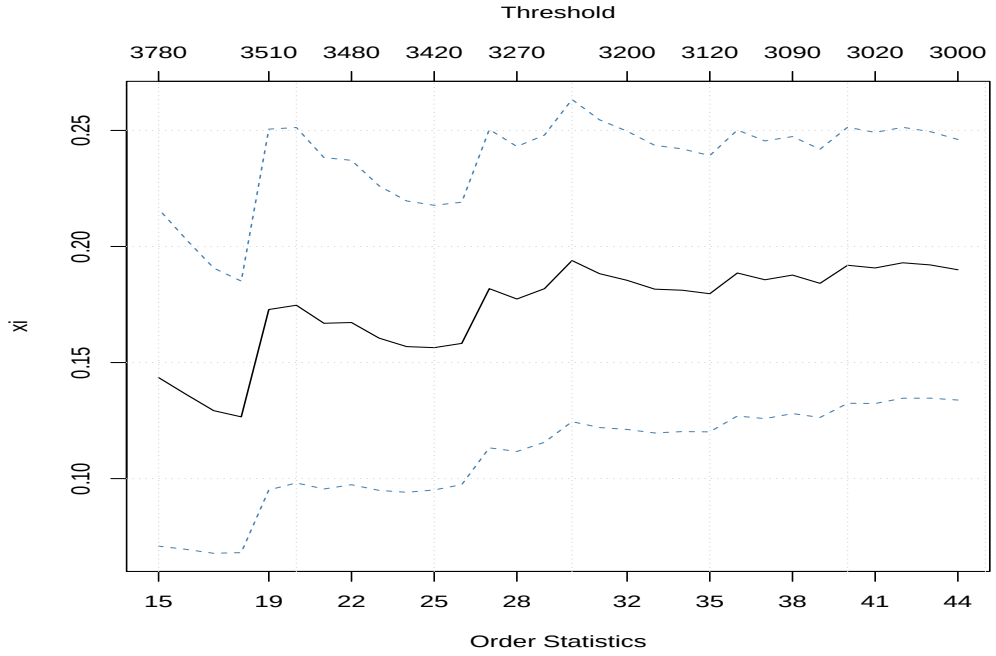

Figure 8.28: Plot of the Hill estimate of the tail index for same day of the week increases in peak electricity demand observations above the threshold of 2981MW together with approximate 95% confidence bands.

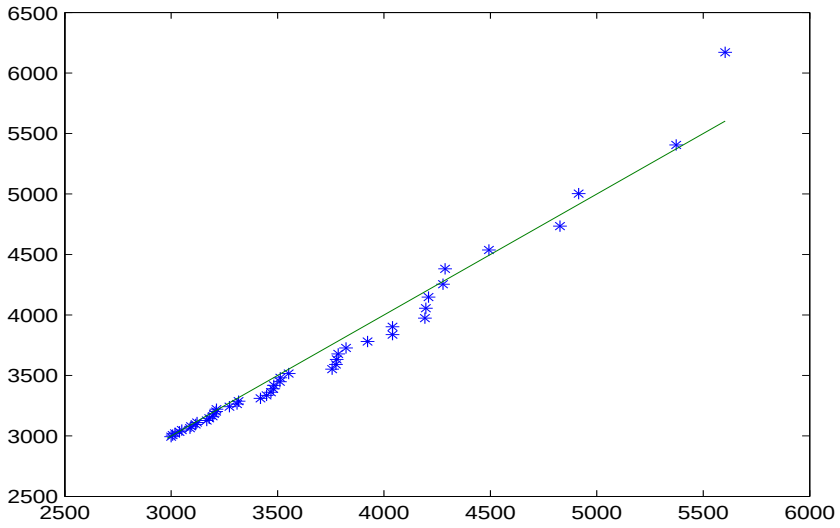

Figure 8.29: A QQ plot of sample data (same day of the week increases in peak demand above the threshold, $\tau = \exp(8) = 2981$MW) versus a GSPD. The horizontal axis (x-axis) represents the standard theoretical quantiles while the empirical quantiles are plotted on the vertical axis (y-axis).

In order to assess whether the GSPD is a good fit to the empirical distribution of the exceedances a QQ plot is constructed. The theoretical quantiles from the GSPD with the estimated parameter are plotted against the empirical quantiles in Figure 8.29. In Figure 8.29 the quantiles seem to lie on the 45^0 line except for the last observation which is a little underestimated. This is a fairly good fit.

8.6.5 Comparative analysis and discussion

Selected posterior predicted tail probabilities for various future extreme same day of the week increases in peak electricity demand are given in Table 8.6. For the GSPD, 1000 η's are simulated and substituted into expression (8.24). Similarly for the GPD, 1000 ξ's and 1000 σ's were simulated and substituted into expression (8.17). After substituting the averages are obtained. Both the GSPD and the GPD are a good fit to the data. As noted earlier, one of the main

Table 8.6: Posterior predictive tail probabilities ($\tau = 2981$).

Observation X_0 (in megawatts)	$P(X_0 > x_0\|x, \tau)$ GSPD	$P(X_0 > x_0\|x, \tau)$ GPD
4500	0.1230	0.1084
5000	0.0707	0.0514
5500	0.0437	0.0221
6000	0.0288	0.0099

advantages of the GSPD is the estimation of only one parameter instead of two as is the case with the GPD. The simultaneous simulations of two parameters can be difficult and time consuming in terms of programming and computing time.

The maximum extreme same day increase above the threshold of 2981MW is 5603MW. This is the difference between the DPED of Thursday 8 January 2009 (28 202MW) and that of Thursday 1 January 2009 (22 599MW). This huge increase is possibly due to weather changes and day to day variations in individual usage of electricity and possibly due to the fact that 1 January is a public holiday. Table 8.7 shows the frequency of same day of the week increases in peak demand. In table 8.7 'Freq' represents frequency. The daily frequency of the occurrence of exceedances given in Figure 8.30 shows that Monday has the highest frequency of 15 followed by Friday with a frequency of 13. This shows that large increases are likely to be experienced mainly on Mondays.

Table 8.7: Same day of the week frequency of exceedances (above $\tau = 2981$MW) over the sampling period 2000-2010.

Day	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Freq	15	3	5	4	13	3	1

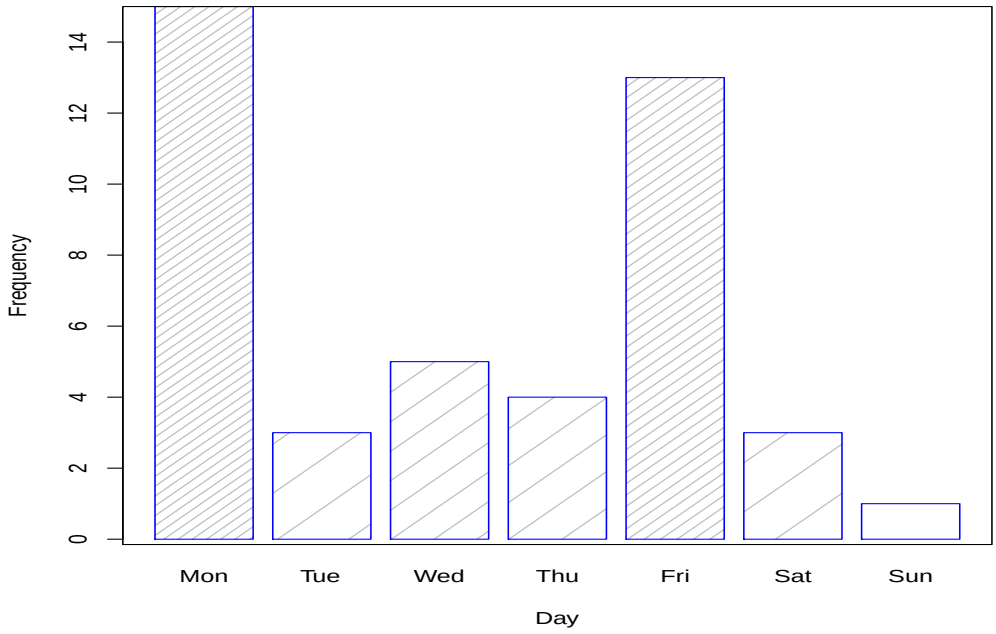

Figure 8.30: Bar chart for frequency of occurrence of same day of the week increases in peak electricity demand above the threshold of $\tau = 2981$MW.

The monthly frequency of same day of the week increases in peak electricity demand above the threshold of 2981MW is given in Table 8.8. An assess-

Table 8.8: Monthly frequency of exceedances (above $\tau = 2981$MW) over the sampling period 2000-2010.

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Freq	14	0	3	20	7	0	0	0	0	0	0	0

ment of the monthly level and frequency of exceedances above the threshold of 2981MW by month is given in Figure 8.31. Figure 8.31 shows that most of the large same day increases above the threshold of 2981MW experienced on Mondays (as shown in Figure 8.30) are found mainly in the month of April. This analysis is important for both load forecasters and system operators of power utilities for planning, load flow analysis and scheduling of short-term electricity demand particularly during periods of peak demand. Figure 8.32 (top panel) shows that same day of the week increases above 2 981 megawatts

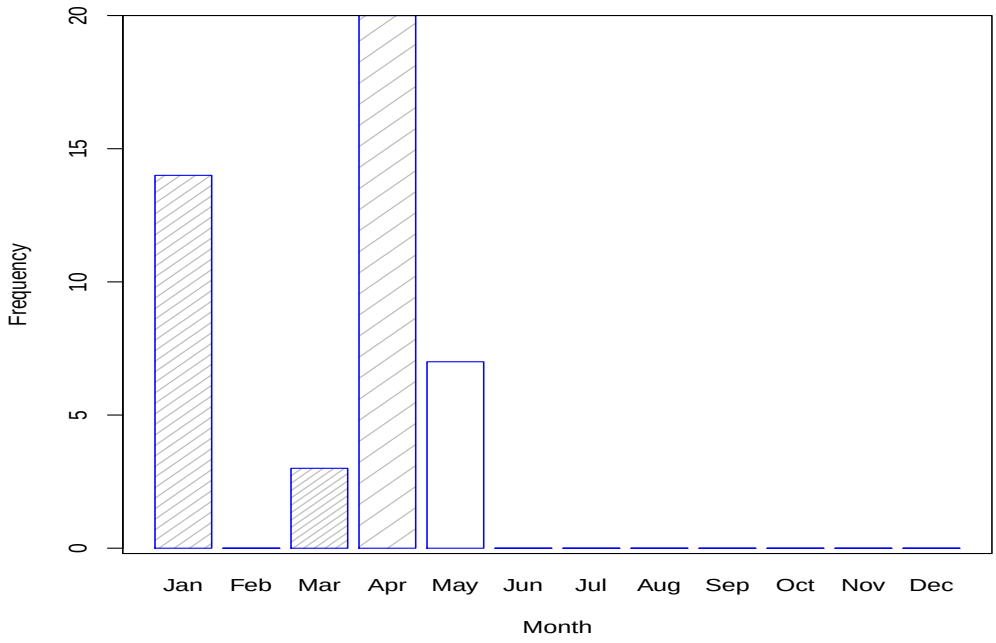

Figure 8.31: Bar chart for monthly frequency of occurrence of same day of the week increases in peak electricity demand above the threshold $\tau = 2981$MW.

are volatile and that there is a weak positive linear trend. This volatility poses challenges to load forecasters and system operators in scheduling peak electricity demand. There is a heavy tail present in the distribution in Figure 8.32 (bottom panel) and is skewed to the right. This calls for accurate predictions of extreme same day of the week peak electricity demand increases as over or underestimation can be very costly to a power utility. Underestimation will result in insufficient generation of electricity which will result in unmet demand (Hyndman and Fan, 2008). Overestimation results in wastage of financial resources.

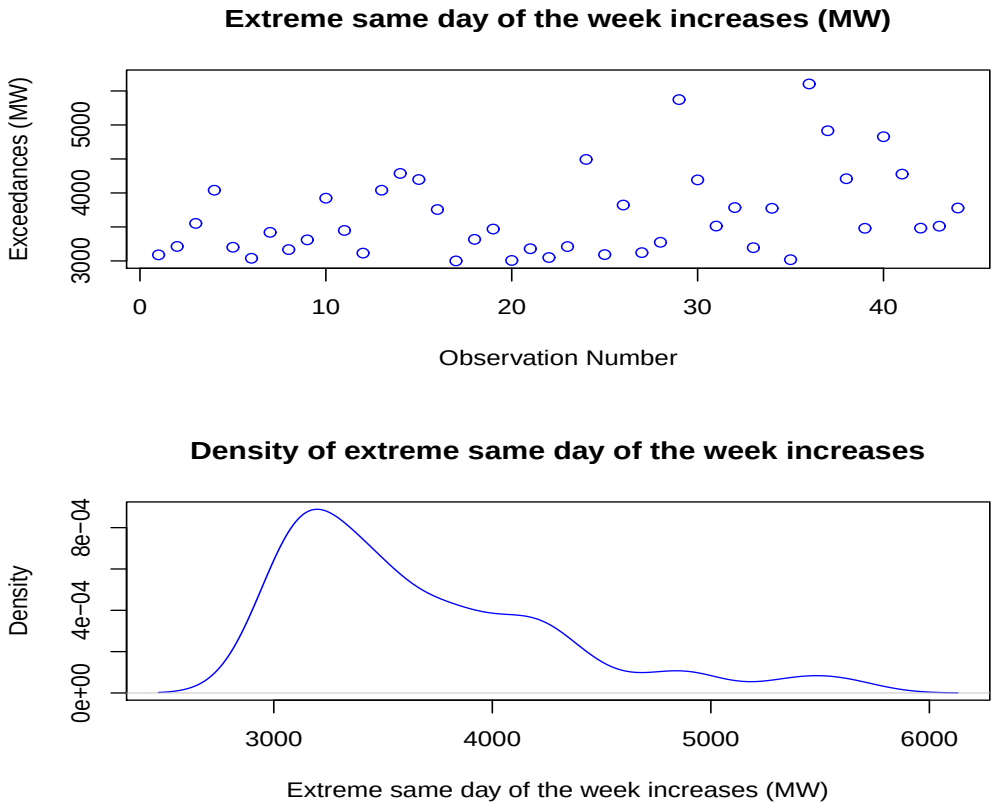

Figure 8.32: (a) Upper panel: Same day of the week increases above 2 981 MW (b) Lower panel: Kernel density estimate of the exceedances. The distribution is heavily right skewed.

8.7 Concluding remarks

The chapter discussed the modelling of extreme interday increases and same day of the week increases in peak electricity demand using South African data. Extreme daily increases in peak electricity demand are modelled using a Generalized Pareto Distribution (GPD) in the first part of the chapter. A comparative analysis of the GPD is done with the Generalized Single Pareto distribution (GSPD). A Pareto quantile plot is used to obtain the optimum threshold. Empirical results show that both the GSPD and the GPD are a good fit to the extreme daily increases in peak electricity demand data.

An application of the GSPD to modelling the same day of the week increases in peak electricity demand using South African data is discussed in the second part of the chapter. The GSPD is used for estimating extreme tail quantiles and probability of exceedance values for various future extreme same day of the week increases in peak electricity demand. A comparative analysis is done with the GPD. The GSPD and the GPD both provide a satisfactory fit to the data. However the GSPD is easier to use since it has only one parameter to estimate instead of two as is the case with the GPD.

Modelling of extreme daily increases in peak electricity demand helps in determining the number of critical peak days and also in improving the reliability of a power network if an accurate assessment of the frequency and level of future extreme load forecasts is carried out (Hor *et al.*, 2008). This also helps to relieve peak capacity loads (Cousins, 2009). Accurate forecasts used together with dynamic electricity pricing programs may help in shifting peak loads to off peak periods (Letzler, 2010; Newsham and Bowker, 2010; among others).

The next chapter focuses on modelling extreme winter peaks in electricity demand.

Chapter 9

Winter peak electricity demand modelling in South Africa

9.1 Introduction

Chapters 3 to 7 established temperature as an important driver of electricity demand in South Africa. Electricity demand increases significantly for temperature values below 18^0C as discussed in Chapters 3 to 7. Modelling of abnormal winter peak electricity demand is important to load forecasters who must provide accurate predictions. In this chapter we model extreme Daily Peak Electricity Demand (DPED) during the winter period in South Africa. The data is initially detrended using the method discussed in Chapter 8. We use a Peaks Over Threshold (POT) distribution to model the winter peaks above a sufficiently high threshold. The POT distribution used is the Generalized Single Pareto Distribution (GSPD). A comparative analysis is then carried out with the Generalized Extreme Value Distribution (GEVD).

9.2 The Generalized Single Pareto Distribution

The survival function of the GSPD distribution is given in equation (9.1).

$$P(X > x|\tau) = \left\{1 + \frac{\eta}{1+\tau\eta}(x-\tau)\right\}^{-\frac{1}{\eta}}, \eta \neq 0, x > \tau \tag{9.1}$$

where x is DPED and τ is the threshold. The Extreme Value Index (EVI) η is estimated using the Bayesian approach. We use the Maximal Data Information (MDI) prior discussed in Zellner (1977), as it provides maximal data information, easy to implement and that constraints can be built into the prior. See Zellner (1977) for more details. The MDI prior for the GSPD derived in Chapter 8 and discussed in detail in Verster and De Waal (2011) is given as:

$$\pi(\eta) \propto \exp\left\{E\left[\log f(X|\eta)\right]\right\} = \frac{e^{-\eta}}{1+\tau\eta} \tag{9.2}$$

The likelihood function is

$$\pi(x|\eta) \propto \prod_{i=1}^{N_\tau} \frac{1}{1+\tau\eta}\left[1 + \frac{\eta(x_i-\tau)}{1+\tau\eta}\right]^{-\frac{1}{\eta}-1} \tag{9.3}$$

i.e. $X_i \sim \text{GSPD}(\eta), i = 1, ..., N_\tau$, N_τ is the number of DPED observations above the threshold τ. The posterior density of η is given as

$$\pi(\eta|x) \propto \prod_{i=1}^{N_\tau} \frac{1}{1+\tau\eta}\left[1 + \frac{\eta(x_i-\tau)}{1+\tau\eta}\right]^{-\frac{1}{\eta}-1} \pi(\eta) \tag{9.4}$$

Future posterior predictive tail probabilities of a future observation, X_0, can be predicted through the following posterior predictive density:

$$P(X_0 > x_0|x,\tau) \propto \int \pi(\eta|x)\left\{1+\frac{\eta}{1+\tau\eta}(x_0-\tau)\right\}^{-\frac{1}{\eta}} d\eta, \eta \neq 0, x > \tau$$

$$P(X_0 > x_0|x,\tau) \approx \frac{1}{m}\sum_{j=1}^{m}\left(1+\frac{\eta_j(x_0-\tau)}{1+\eta_j\tau}\right) \qquad (9.5)$$

The posterior predictive density given in expression (9.5) cannot be computed analytically. We will approximate the posterior predictive density through simulations. Values of η are simulated from the posterior density given in expression (9.4) and substituted into expression (9.5). The average over all the tail probabilities is then used as an estimate of the posterior predictive tail probability.

9.3 Generalized Extreme Value Distribution

The unified GEVD is given by

$$G_\gamma(x) = \exp\left\{-\left[1+\gamma\left(\frac{x-\mu}{\sigma}\right)\right]^{\frac{1}{\gamma}}\right\}, 1+\gamma\left(\frac{x-\mu}{\sigma}\right) > 0, \gamma \neq 0 \qquad (9.6)$$

where γ is the extreme value index. For $\gamma = 0$ equation (9.7) is the Gumbel class of distributions which is given by

$$G_\gamma(x) = \exp\left\{-\exp\left[-\frac{x-\mu}{\sigma}\right]\right\}, \gamma = 0 \qquad (9.7)$$

When $\gamma > 0$ we have the Frechet class of distributions and when $\gamma < 0$ we have the Weibull class of distributions as discussed in Chapter 7.

GEVD: Prior Distribution

For GEVD, Jeffreys' prior is complicated and only exists for $\gamma > -0.5$ (Beirlant *et al.*, 2004). We use the trivariate normal prior (conjugate prior) discussed in Coles and Powell (1996); Stephenson and Ribatet (2006). Let the prior distribution be denoted by $\pi(\boldsymbol{\theta})$ where $\boldsymbol{\theta} = \{\sigma, \mu, \gamma\}$. Let

$$\boldsymbol{\theta}' = \{\log\sigma, \mu, \gamma\}$$

then the prior distribution on $\boldsymbol{\theta}$ is defined as (Stephenson and Ribatet, 2006)

$$\pi(\boldsymbol{\theta}') \propto \frac{1}{\sigma}\exp\left\{-\frac{1}{2}\left(\boldsymbol{\theta}' - \boldsymbol{\nu}\right)^T \Sigma^{-1}\left(\boldsymbol{\theta}' - \boldsymbol{\nu}\right)\right\} \tag{9.8}$$

where $\boldsymbol{\nu}$ is the mean vector and Σ is the covariance matrix.

Likelihood function

The likelihood function is

$$\pi(x|\boldsymbol{\theta}) = \prod_{i=1}^{n}\frac{1}{\sigma}\left[1+\gamma\left(\frac{x_i-\mu}{\sigma}\right)\right]^{-\frac{1}{\gamma}-1}\exp\left\{-\left[1+\gamma\left(\frac{x_i-\mu}{\sigma}\right)\right]^{-\frac{1}{\gamma}}\right\} \tag{9.9}$$

Joint posterior distribution

The joint posterior density is

$$\begin{aligned}
\pi(\boldsymbol{\theta}|x) &\propto \pi(\boldsymbol{\theta})\pi(x|\boldsymbol{\theta}) \\
\pi(\boldsymbol{\theta}|x) &\propto \frac{1}{\sigma}\exp\left\{-\frac{1}{2}\left(\boldsymbol{\theta}'-\boldsymbol{\nu}\right)^T\Sigma^{-1}\left(\boldsymbol{\theta}'-\boldsymbol{\nu}\right)\right\} \\
&\times \prod_{i=1}^{n}\frac{1}{\sigma}\left[1+\gamma\left(\frac{x_i-\mu}{\sigma}\right)\right]^{-\frac{1}{\gamma}-1}\exp\left\{-\left[1+\gamma\left(\frac{x_i-\mu}{\sigma}\right)\right]^{-\frac{1}{\gamma}}\right\} \\
\pi(\boldsymbol{\theta}|x) &\propto \frac{1}{\sigma^{n+1}}\exp\left\{-\frac{1}{2}\left(\boldsymbol{\theta}'-\boldsymbol{\nu}\right)^T\Sigma^{-1}\left(\boldsymbol{\theta}'-\boldsymbol{\nu}\right)\right\} \\
&\times \exp\left\{-\Sigma_{i=1}^{n}\left[1+\gamma\left(\frac{x_i-\mu}{\sigma}\right)\right]^{-\frac{1}{\gamma}}\right\} \\
&\times \prod_{i=1}^{n}\left[1+\gamma\left(\frac{x_i-\mu}{\sigma}\right)\right]^{-\frac{1}{\gamma}-1} \\
\pi(\boldsymbol{\theta}|x) &\propto \frac{1}{\sigma^{n+1}}\exp\left\{-\frac{1}{2}\left(\boldsymbol{\theta}'-\boldsymbol{\nu}\right)^T\Sigma^{-1}\left(\boldsymbol{\theta}'-\boldsymbol{\nu}\right)-\sum_{i=1}^{n}\left[1+\gamma\left(\frac{x_i-\mu}{\sigma}\right)\right]^{-\frac{1}{\gamma}}\right\} \\
&\times \prod_{i=1}^{n}\left[1+\gamma\left(\frac{x_i-\mu}{\sigma}\right)\right]^{-\frac{1}{\gamma}-1} \qquad (9.10)
\end{aligned}$$

Let $\boldsymbol{\theta_0} = \{\sigma_0, \mu_0, \gamma_0\}$ and $s = \{s_\sigma, s_\mu, s_\gamma\}$ denote the initial values of the Markov Chain Monte Carlo (MCMC). We use the Maximum Likelihood (ML) estimates for these initial values. If the chain is at state $\theta_t = \{\sigma_t, \mu_t, \gamma_t\}$, then subsequent states θ_{t+1} are generated as follows (Stephenson and Ribatet, 2006): Let LN denote the log-normal distribution then the algorithm to generate θ_{t+1} is
Let $\mu^* \sim N(\mu_t, s_\mu^2)$

$$\text{Set } \Delta = \frac{\pi(\mu^*, \sigma_t, \gamma_t|\boldsymbol{x})}{\pi(\mu_t, \sigma_t, \gamma_t|\boldsymbol{x})} \qquad (9.11)$$

Set $\mu_{t+1} = \mu^*$ with probability min $\{1, \Delta\}$, else set $\mu_{t+1} = \mu_t$.
Let $\sigma^* \sim LN(\log\sigma_t, s_\sigma^2)$

$$\text{Set } \Delta = \frac{\pi(\mu_{t+1}, \sigma^*, \gamma_t|\boldsymbol{x})}{\pi(\mu_{t+1}, \sigma_t, \gamma_t|\boldsymbol{x})} \frac{\sigma^*}{\gamma_t} \tag{9.12}$$

Set $\sigma_{t+1} = \sigma^*$ with probability min $\{1, \Delta\}$, else set $\sigma_{t+1} = \sigma_t$.
Let $\gamma^* \sim N(\gamma_t, s_\gamma^2)$

$$\text{Set } \Delta = \frac{\pi(\mu_{t+1}, \sigma_{t+1}, \gamma^*|\boldsymbol{x})}{\pi(\mu_{t+1}, \sigma_{t+1}, \gamma_t|\boldsymbol{x})} \tag{9.13}$$

Set $\gamma_{t+1} = \gamma^*$ with probability min $\{1, \Delta\}$, else set $\gamma_{t+1} = \gamma_t$.

Posterior predictive density

Future posterior predictive tail probabilities of a future observation, X_0 can be predicted by the following posterior predictive density:

$$\begin{aligned} P(X_0 > x_0|x_1, ..., x_n) &= \int_{\boldsymbol{\theta}} P(X_0 > x_0|\boldsymbol{\theta})\pi(\boldsymbol{\theta}|x)d\boldsymbol{\theta} \\ P(X_0 > x_0|x_1, ..., x_n) &\approx \frac{1}{n-b+1}\sum_{i=b}^{n} P(X_0 > x_0|\boldsymbol{\theta_i}) \end{aligned} \tag{9.14}$$

where $P(X_0 > x_0)$ is the GEVD evaluated at x_0, b is the burn-in period (Stephenson and Ribatet, 2006).

Let X_M be the DPED over a future period of M years, the posterior predictive distribution is then given as (Stephenson and Ribatet, 2006):

$$P(X_0 > x_0|x_1, ..., x_n) \approx \frac{1}{n-b+1}\sum_{i=b}^{n} P(X_0 > x_0|\boldsymbol{\theta_i})^M \tag{9.15}$$

9.4 Winter demand characteristics and demand data

Winter demand forecasts have been developed using historical data for 11 winters, from 2000 to 2010. The data is from Eskom, South Africa's power utility company. The winter is defined as the 93 day period from 15 May to 15 August of each year. Winter peak demand is the highest level of hourly demand observed during a particular winter. A plot of daily peak electricity demand data from 2000 to 2010 is given in Figure 9.1. The winter period experiences

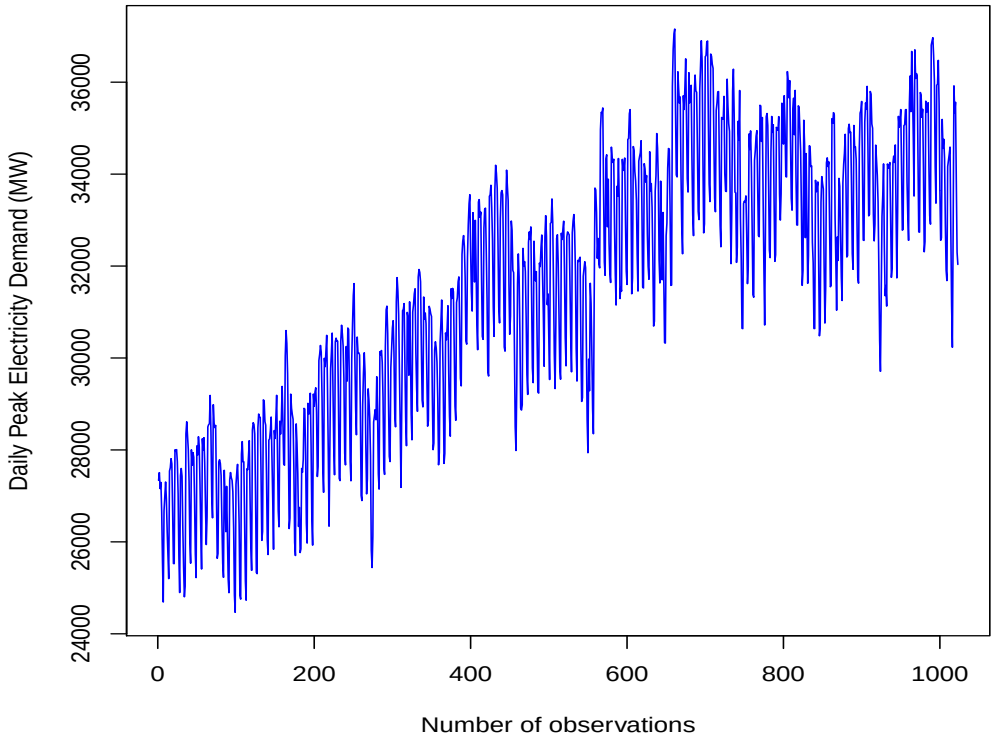

Figure 9.1: Daily peak electricity demand data for the years 2000 to 2010. Only data from 15 May to 15 August of each year are given.

high demand of electricity in South Africa since demand is highly sensitive to temperature changes in winter and less sensitive in summer as discussed in

Chapters 3 to 7. This calls for accurate predictions of peak winter electricity demand as underestimation would result in load interruptions and blackouts. The highest annual peak electricity demand is in July which is a winter month in the Southern Hemisphere, see Figure 9.2.

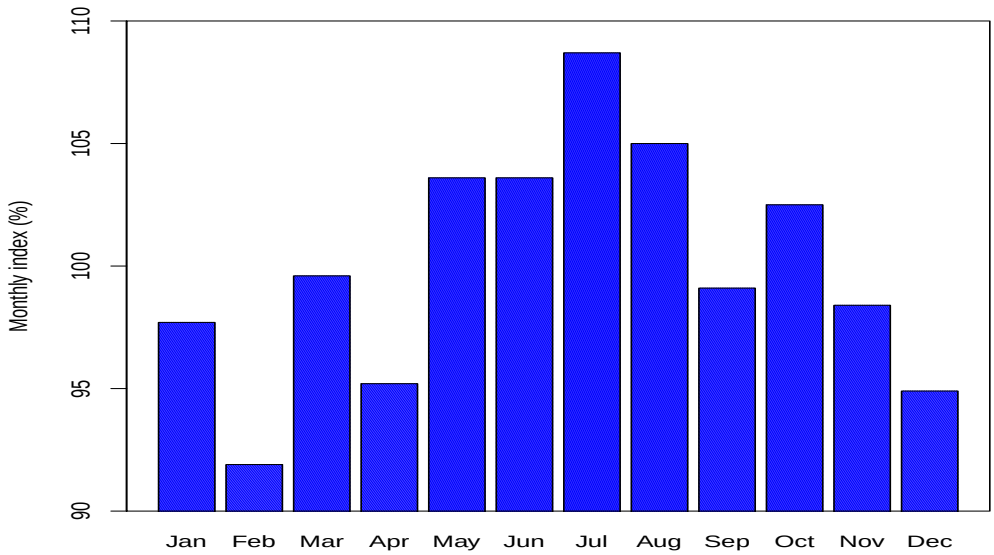

Figure 9.2: Monthly seasonal index plot of DPED. The index for each month is calculated by finding the average DPED for each month and then divide by the overall average.

9.5 Empirical results

Electricity demand is generally subject to an upward positive linear trend. The data is initially made stationary by removing the linear trend (detrending). We fit a trend line to DPED, that is, we regress DPED (x_t) against time t, for $t = 1, .., n$. The trend line is given by the following equation

$$x_t = \beta_0 + \beta_1 t + \varepsilon_t$$

where β_0, β_1 are parameters and ε_t is the error term. We then subtract the values of the trend line from the original data which results in a time series of residuals, i.e.

$$\varepsilon_t = x_t - \hat{x}_t = x_t - (\hat{\beta}_0 + \hat{\beta}_1 t)$$

An initial threshold is set at zero after detrending and all those observations with negative values are excluded. Figure 9.3 shows a plot of the detrended DPED data.

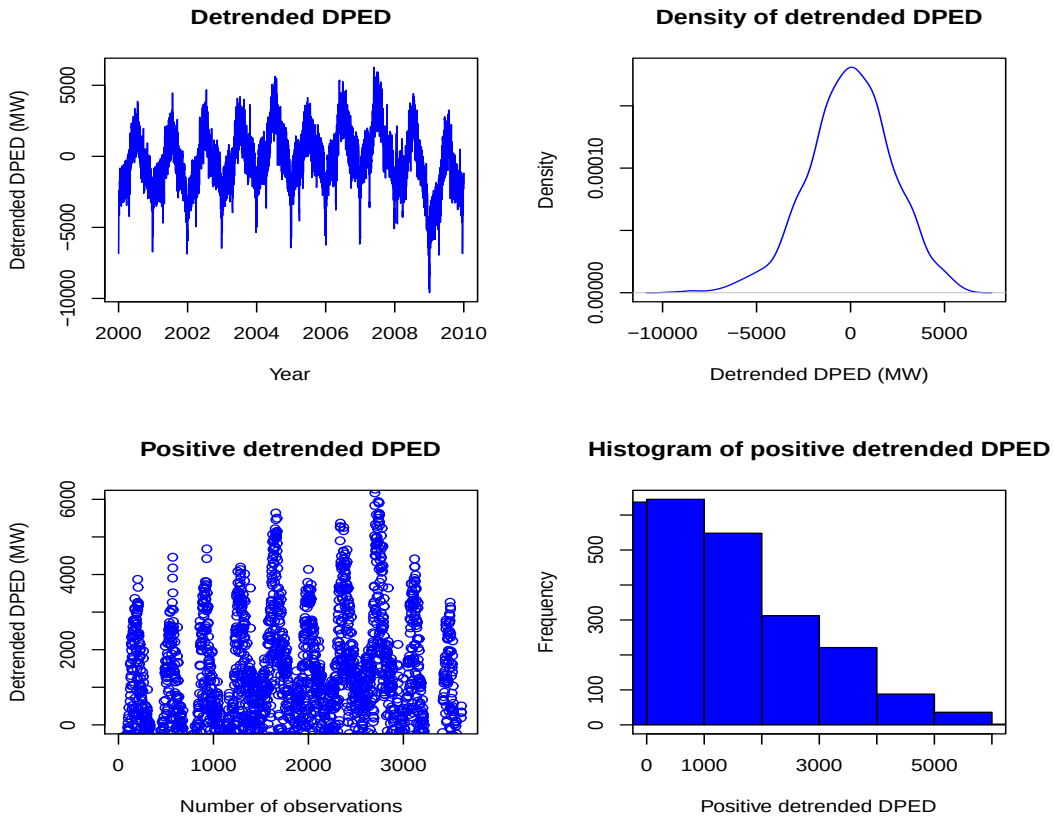

Figure 9.3: (a) Upper left panel: Detrended daily peak electricity demand (b) Upper right panel: Density of detrended DPED. The distribution is left skewed. (c) Lower left panel: Positive detrended daily peak electricity demand (d) Lower right panel: Histogram of positive detrended DPED.

The optimum threshold of these observations is then determined using a Pareto quantile plot shown in Figure 9.4. The threshold is chosen at the observation value where the plot begins to follow a linear pattern, which is $\tau = \exp(8) = 2980.96$.

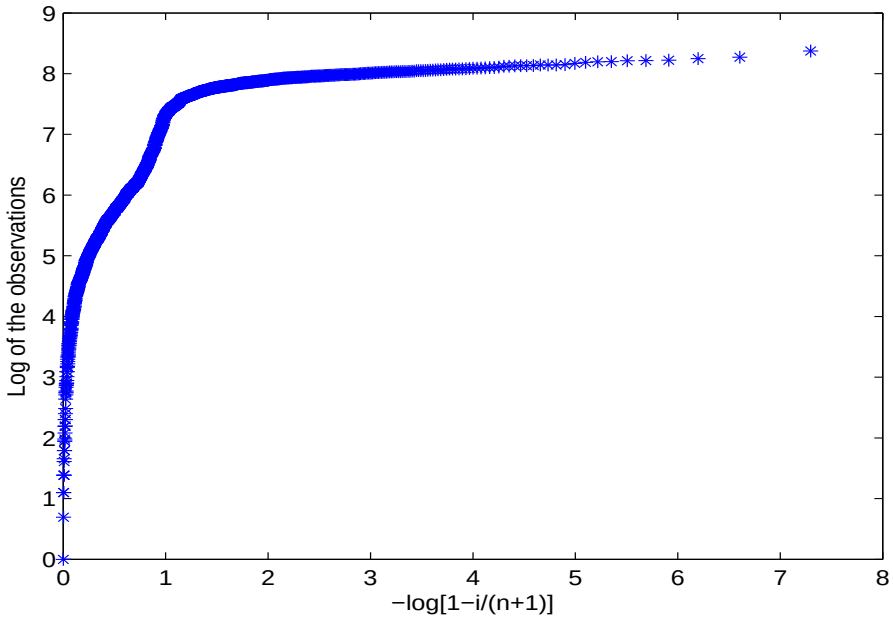

Figure 9.4: Pareto quantile plot on the observations greater than zeros.

9.5.1 Predicting future observations

We now consider the observations larger than the threshold τ to be generalized single Pareto distributed. The parameter η is estimated through a Bayesian approach. The mode of the posterior distribution is taken as the estimate of η which is found to be 0.077. The posterior density of η is given in Figure 9.5.

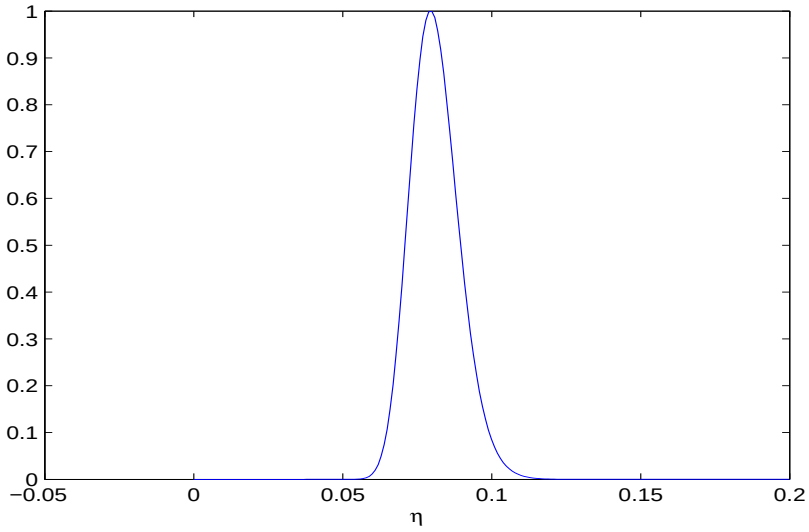

Figure 9.5: The posterior density of η.

The posterior predictive survival distribution of the approximated GSPD given in equation (9.5) is now used to predict posterior tail probabilities. For example we can predict the probability exceeding a large value such as 4500, which is

$$P(X_0 > 4500|x_1, ..., x_n) = \frac{1}{m}\sum_{j=1}^{m}\left(1 + \frac{\eta_j(x_0 - \tau)}{1 + \eta_j\tau}\right) = 0.0052 \tag{9.16}$$

where m number of η values are simulated from the posterior and substituted into equation (9.5). The value on the j-th simulation run is denoted by η_j. Some of the predicted tail probabilities for various future daily peak electricity demands where η was simulated from its posterior are given in Table 9.1.

Table 9.1: Posterior predictive tail probabilities

Observation x_{n+1}	$P(X > x_{n+1}\|X > \tau)$ Posterior tail probabilities (GSPD)
3000	0.9224
3500	0.1310
4000	0.0237
4500	0.0052

The density function of the posterior predictive tail probabilities is given in Figure 9.6.

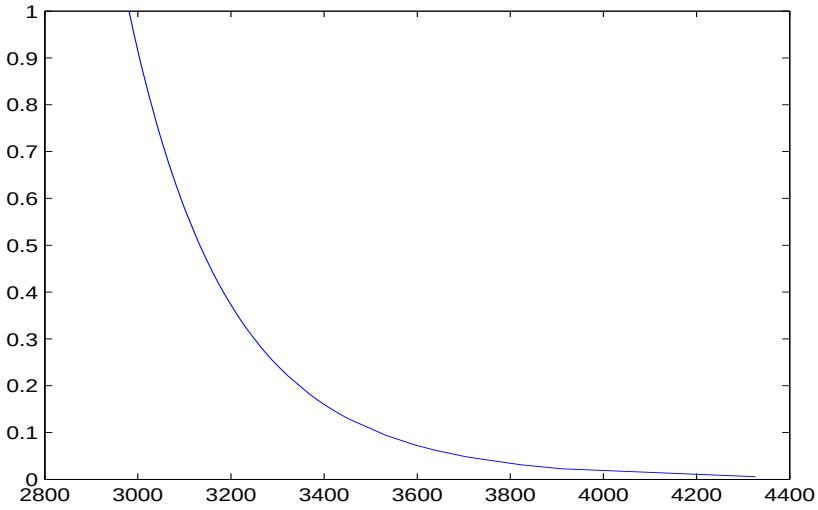

Figure 9.6: Density function for the posterior predictive tail probabilities.

9.5.2 Predicting future observations using the Generalized extreme Value Distribution (GEVD): A comparative analysis

In this section we fit a Generalized Extreme Value Distribution (GEVD) to all the detrended DPED data, $M_{1n}, M_{2n}, ..., M_{bn}$ (not only the observations above a threshold). In order to extract upper extreme values an extreme value distribution is then fitted to the block maxima data $\{M_{in}\}_i, i = 1, ..., b$. The probability of M_{in} not exceeding a value x_i is then given as

$$P\{M_{in} \leq x_i\} = P\{\mathbf{max}X_{in} \leq x_i\} = P\{X_{i1} \leq x_i, X_{i2} \leq x_i, ..., X_{in} \leq x_i\}$$

$$= \{F(x_i)\}^{in} \tag{9.17}$$

and the exceedance probability is given as

$$P\{M_{in} \geq x_i\} = 1 - \{F(x_i)\}^{in} \tag{9.18}$$

The three parameters γ, μ, and σ are estimated for the given dataset through the Bayesian approach. The R statistical package evdbayes by Stephenson and Ribatet (2012) is used for obtaining the parameter estimates. We use the maximum likelihood estimates of μ, σ and γ and their corresponding standard errors as the initial trial solutions. We use simulated annealing (SANN) to maximise $\pi(\boldsymbol{\theta}|\boldsymbol{x}$. Simulated annealing is a meta-heuristic which seeks a global or near global optimal solution. For a detailed discussion of simulated annealing see Deb (1996); Eglese (1990); Kirkpatrick *et al.* (1983); among others. The Bayesian estimates are:

$$\gamma = -0.0401,$$

$$\mu = -540.482,$$

$$\sigma = 1008.22.$$

These results show that the data can be modelled using a Weibull class, ($\gamma < 0$). The quantile function for the GEVD is then used for predicting the probability of exceedance levels. The quantile function is given by

$$x_p = \mu + \frac{\sigma}{\gamma}\left[\{-\log(1-p)\}^{-\gamma} - 1\right], \text{if } \gamma \neq 0 \tag{9.19}$$

The probabilities are calculated using equation (9.20).

$$p = 1 - \exp\left[-\left\{\frac{\gamma(x_p - \mu) + \sigma}{\sigma}\right\}^{-\frac{1}{\gamma}}\right] \tag{9.20}$$

Some of the predicted tail probabilities for various future daily peak electricity demand are given in Table 9.2.

Table 9.2: Posterior predictive tail probabilities

Observation x_{n+1}	Posterior predictive tail probabilities (GEVD)
3000	0.0225
3500	0.0126
4000	0.0069
4500	0.0038

The return level plot of the posterior distribution with 95% Bayesian credible intervals (dashed lines) is given in Figure 9.7.

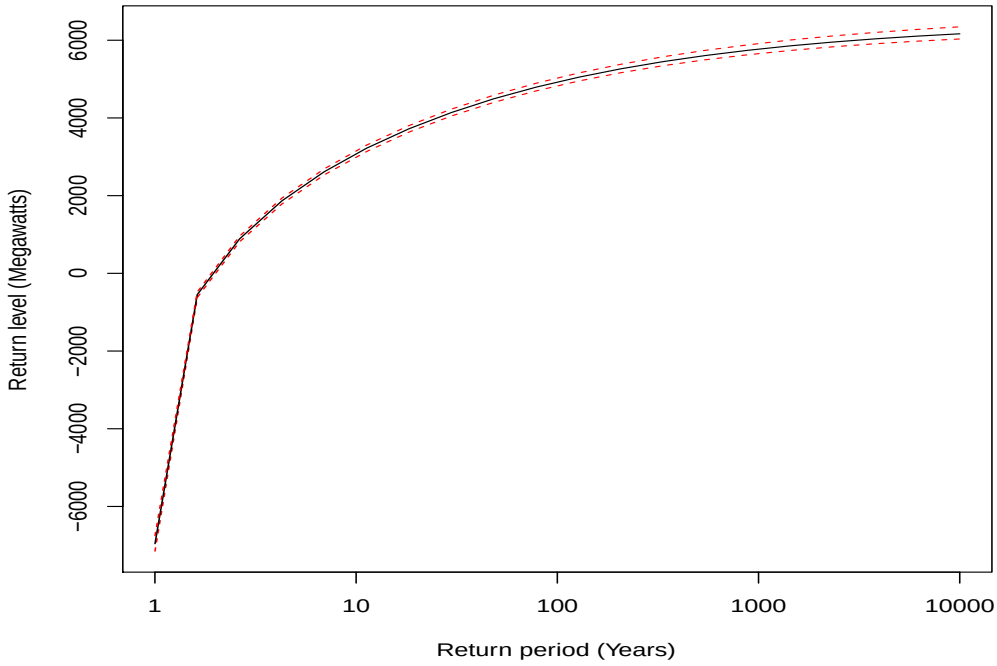

Figure 9.7: Return level plot of posterior distribution with 95% Bayesian credible intervals (dashed lines).

9.6 Comparative analysis

A summary of the posterior predictive tail probabilities from the GSPD and GEVD are given in Table 9.3. The Quantile-Quantile (QQ) plot for the observations above the threshold is constructed in Figure 9.8. The QQ-plot is a method for judging the goodness of fit of the proposed model to the data, see for example Beirlant *et al*. (2004). If the QQ-plot follows a 45^0 line it indicates a good fit. Figure 9.8 therefore indicates a fairly good fit of the GSPD to the data, i.e. $M_{1n}, M_{2n}, ..., M_{bn}$ above $\tau = 2981$.

Table 9.3: Posterior predictive tail probabilities

Observation x_{n+1}	Tail probability (GSPD)	Tail probability (GEVD)
3000	0.9224	0.0225
3500	0.1310	0.0126
4000	0.0237	0.0069
4500	0.0052	0.0038

In Figure 9.9 the QQ-plot (GEVD) for the theoretical quantiles against the empirical quantiles is constructed for all the observations. The results from the GSPD are better than those from fitting a GEVD as shown by the QQ-plots of Figures 9.8 and 9.9 respectively.

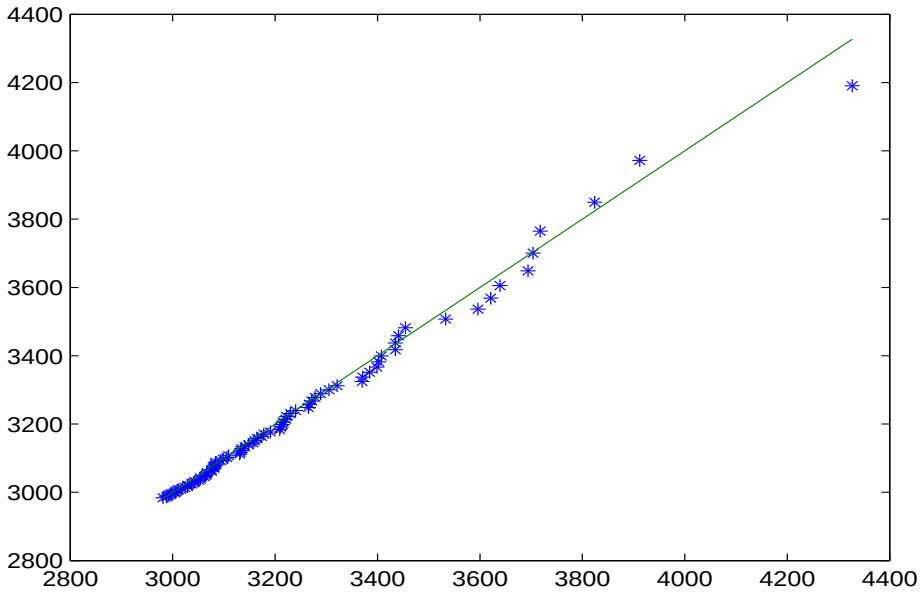

Figure 9.8: QQ plot of DPED above $\tau = 2981$ (GSPD). The horizontal axis represents the standard theoretical quantiles while the empirical quantiles are plotted on the vertical axis.

9.6.1 Frequency analysis of exceedances

Figure 9.10 shows a bar chart of the frequency of occurrence of exceedances above the threshold $\tau = 2981$MW using the Generalized Single Pareto Distribution. Figure 9.10 shows that July has the highest number of exceedances. This indicates that it is the coldest month during the winter period. Modelling of winter peak electricity demand improves the reliability of a power network if an accurate assessment of the level and frequency of future extreme winter peak load forecasts is carried out as peak demand may exceed existing maximum generating capacity of a power utility company.

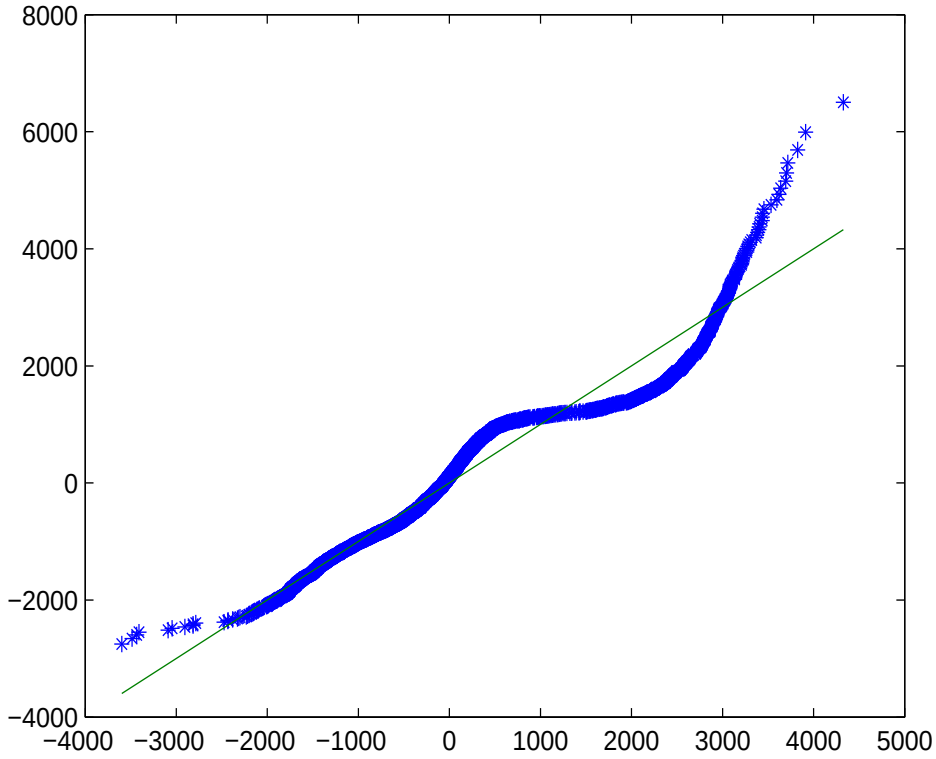

Figure 9.9: QQ plot of DPED above $\tau = 2981$ (GEVD). The horizontal axis represents the standard theoretical quantiles while the empirical quantiles are plotted on the vertical axis.

9.7 Concluding remarks

In this chapter, we presented an application of a Generalized Single Pareto Distribution (GSPD) to the statistical analysis of extreme daily peak electricity demand during the winter period in South Africa. The GSPD seems appropriate in modelling of extreme winter peak electricity demand. The main advantage of this modelling approach is that the GSPD has only one parameter to estimate, given that the threshold is known. A comparative analysis was done by fitting a Generalized Extreme Value Distribution (GEVD). Empirical results show that the Weibull class can be used to model the daily peak electricity demand data. The results from the GSPD are better than those from fitting a

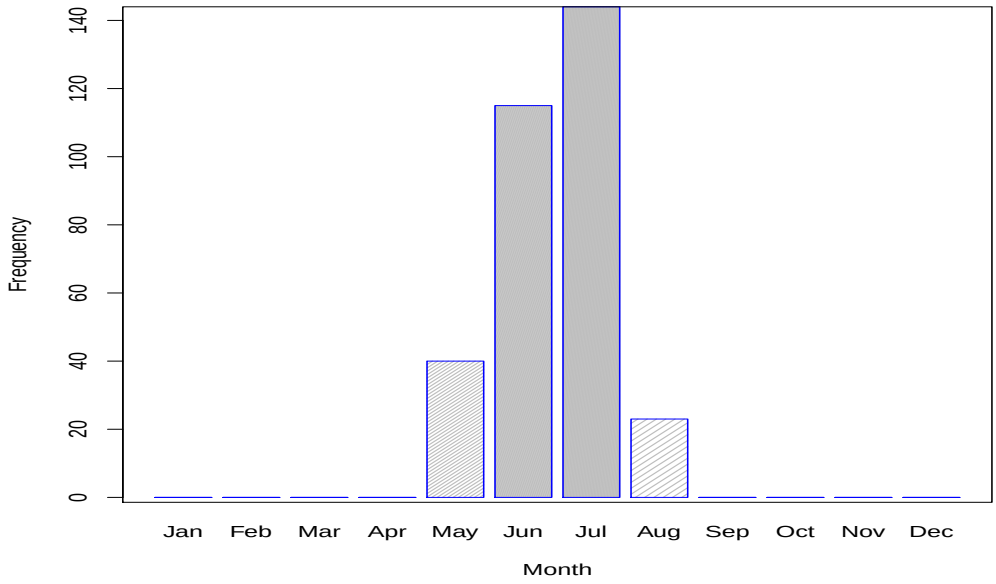

Figure 9.10: Bar chart of the frequency of occurrence of exceedances above the threshold $\tau = 2981$MW using the Generalized Single Pareto Distribution.

GEVD. Extreme value theory is good at analysing tails of distributions.

The focus in the next chapter is an analysis of the tails of the residuals after accounting for changes in the behaviour of the mean demand.

Chapter 10

Assessing peak electricity demand uncertainty in South Africa

10.1 Introduction

Peak electricity demand modelling is a policy concern for countries throughout the world (Strengers, 2012). Many countries are investing heavily in the construction of new generating plants in order to increase electricity supply during peak demand periods. Most countries including those with emerging economies, have embarked on use of new and smart energy saving technologies and have put in place integrated demand side management and energy efficient strategies and policies in an effort to reduce consumption of electricity.

In this chapter we present a probabilistic description of interday changes in

daily peak electricity demand uncertainty and the modelling of extreme under demand predictions using a hybrid model, the Autoregressive Moving Average-Exponential Generalized Autoregressive Conditional Heteroskedasticity-Generalized Single Pareto Distribution (ARMA-EGARCH-GSPD). We define interday changes as daily increase/decrease in Daily Peak Electricity Demand (DPED) where DPED is the maximum hourly demand in a 24-hour period. The Chapter focuses on positive interday changes. Modelling of unexpected extreme positive interday increases is important to load forecasters, systems operators and demand managers in planning, load flow analysis and scheduling of electricity particularly during peak periods. The modelling approach discussed in this chapter is important for assessing and quantifying risk in interday increases in peak electricity demand forecasting.

10.2 The models

We adopt the approach used by McNeil and Frey (2000) and Bystrom (2005). Using a two stage approach, McNeil and Frey (2000) estimate a GARCH model in stage one with a view to filtering the return series to get nearly independent and identical distributed residuals. In stage two, the extreme value theory (EVT) framework is then applied to the standardized residuals.

10.2.1 ARMA-EGARCH model

Electricity returns are highly volatile, exhibit leverage effects and clustering in volatility and display seasonalities in both their mean and as well as volatility, and feature extreme levels of skewness and kurtosis (Chan and Gray, 2006). This requires the use of ARMA-EGARCH extreme value theory modelling framework (McNeil and Frey, 2000; Bystrom, 2005).

Formal unit root tests are conducted using the Augmented-Dickey Fuller test.

Results indicate that the natural logarithm of the first difference of DPED is stationary. Based on the stationarity requirements we calculate the interday percentage changes (r_t) that are called the return series data, as given in equation (10.1).

$$r_t = 100(\log(x_t) - \log(x_{t-1})) \tag{10.1}$$

where x_t, x_{t-1} are the current and one period lagged DPED respectively. The returns (r_t) given in equation (10.1) are explained in detail in appendix 10. Assuming a conditional normal distribution, we adopt an ARMA(*p,q*)-EGARCH(1,1) model with the following mean and variance structures

$$\text{Mean equation} : r_t = \phi_0 + \sum_{i=1}^{p} \phi_i r_{t-i} + \varepsilon_t + \sum_{i=1}^{q} \theta_j \varepsilon_{t-j}, \varepsilon_t \sim N(0, \sigma_t^2) \tag{10.2}$$

$$\text{Variance equation} : \log\sigma_t^2 = \omega + \beta \log\sigma_{t-1}^2 + \alpha \left| \frac{\varepsilon_{t-1}}{\sigma_{t-1}} \right| + \gamma \frac{\varepsilon_{t-1}}{\sigma_{t-1}} \tag{10.3}$$

where r_t is the return series of DPED, as defined in equation (10.1). The EGARCH model was developed to capture the leverage effect in financial time series data (Nelson, 1991). Negative shocks in financial markets (bad news) generally have larger impacts on market volatility than positive shocks (good news). The presence of a leverage effect can be tested by the hypothesis that $\gamma < 0$; if $\gamma \neq 0$, then the impact is asymmetric. The EGARCH (1,1) model is used because the inequality constraints on the parameters, α, β and γ, given in equation (10.3) are not imposed; oscillatory behaviour in the conditional variance is permitted as the coefficient β can either be positive or negative, and the persistence of volatility shocks can be measured easily (Ho and Tsui, 2004). Nelson (1991) discusses in detail the advantages of using the EGARCH approach instead of the standard GARCH model.

10.2.2 The Generalized Single Pareto distribution

The survival function of the Generalized Single Pareto Distribution (GSPD) is given in equation (10.4).

$$P(\mathcal{E}_t > \varepsilon_t|\tau) = \left\{1 + \frac{\eta}{1+\tau\eta}\left(\varepsilon_t - \tau\right)\right\}^{-\frac{1}{\eta}}, \eta \neq 0, \varepsilon_t > \tau \tag{10.4}$$

where ε_t is the error term (i.e $\varepsilon_t = r_t - f_t$), r_t represents return series and predicted return series is denoted by f_t. The Extreme Value Index (EVI) η is estimated using the Bayesian approach. We use the maximal data information (MDI) prior discussed in Zellner (1977). The MDI prior (Verster and De Waal, 2011) is given as:

$$\pi(\eta) \propto \exp\left\{E\left[\ln f(\mathcal{E}_t|\eta)\right]\right\} = \frac{e^{-\eta}}{1+\tau\eta} \tag{10.5}$$

The likelihood function is

$$\begin{aligned}
\pi(\varepsilon_t|\eta) &\propto \prod_{i=1}^{N_\tau} \frac{1}{1+\tau\eta}\left[1 + \frac{\eta(\varepsilon_i - \tau)}{1+\tau\eta}\right]^{-\frac{1}{\eta}-1} \\
\pi(\varepsilon_t|\eta) &\propto \prod_{i=1}^{N_\tau} \frac{1}{1+\tau\eta}\left[\frac{1+\tau\eta+\eta\varepsilon_i-\tau\eta}{1+\tau\eta}\right]^{-\frac{1}{\eta}-1} \\
\pi(\varepsilon_t|\eta) &\propto \prod_{i=1}^{N_\tau} \frac{1}{1+\tau\eta}\left[\frac{1+\eta\varepsilon_i}{1+\tau\eta}\right]^{-\frac{1}{\eta}-1} \\
\pi(\varepsilon_t|\eta) &\propto (1+\tau\eta)^{-n} \prod_{i=1}^{N_\tau}\left[\frac{1+\eta\varepsilon_i}{1+\tau\eta}\right]^{-\frac{1}{\eta}-1} \\
\pi(\varepsilon_t|\eta) &\propto (1+\tau\eta)^{-n}(1+\tau\eta)^{-\frac{1+\eta}{\eta}} \prod_{i=1}^{N_\tau}\left[1+\eta\varepsilon_i\right]^{-\frac{1+\eta}{\eta}} \\
\pi(\varepsilon_t|\eta) &\propto (1+\tau\eta)^{-\frac{\eta^2-\eta-1}{\eta}} \prod_{i=1}^{N_\tau}\left[1+\eta\varepsilon_i\right]^{-\frac{1+\eta}{\eta}}
\end{aligned} \tag{10.6}$$

i.e. $\mathcal{E}_i \sim \text{GSPD}(\eta), i = 1, ..., N_\tau$, N_τ is the number of observations above the threshold. The posterior density of η is given as

$$\begin{aligned}
\pi(\eta|\varepsilon_t) &\propto \pi(\eta)\pi(\varepsilon_t|\eta) \\
\pi(\eta|\varepsilon_t) &\propto \frac{e^{-\eta}}{1+\tau\eta}\prod_{i=1}^{N_\tau}\frac{1}{1+\tau\eta}\left[1+\frac{\eta(\varepsilon_i-\tau)}{1+\tau\eta}\right]^{-\frac{1}{\eta}-1} \\
\pi(\eta|\varepsilon_t) &\propto e^{-\eta}\left(1+\tau\eta\right)^{-1}\left(1+\tau\eta\right)^{-\frac{\eta^2-\eta-1}{\eta}}\prod_{i=1}^{N_\tau}\left[1+\eta\varepsilon_i\right]^{-\frac{1+\eta}{\eta}} \\
\pi(\eta|\varepsilon_t) &\propto e^{-\eta}\left(1+\tau\eta\right)^{-\frac{\eta^2-1}{\eta}}\prod_{i=1}^{N_\tau}\left[1+\eta\varepsilon_i\right]^{-\frac{1+\eta}{\eta}} \qquad (10.7)
\end{aligned}$$

Future posterior predictive tail probabilities of a future observation, $\mathcal{E}_0$, can be predicted through the following formula:

$$\begin{aligned}
P(\mathcal{E}_0 > \varepsilon_0|\varepsilon_t, \tau) &\propto \int \pi(\eta|\varepsilon_t)\left\{1+\frac{\eta}{1+\tau\eta}\left(\varepsilon_0-\tau\right)\right\}^{-\frac{1}{\eta}} d\eta, \eta \neq 0, \varepsilon_t > \tau \\
P(\mathcal{E}_0 > \varepsilon_0|\varepsilon_1, ..., \varepsilon_n) &\approx \frac{1}{m}\sum_{j=1}^{m}\left(1+\frac{\eta_j(\varepsilon_0-\tau)}{1+\tau\eta_j}\right) \\
P(\mathcal{E}_0 > \varepsilon_0|\varepsilon_1, ..., \varepsilon_n) &\approx \frac{1}{m}\sum_{j=1}^{m}\left(\frac{1+\varepsilon_0\eta}{1+\tau\eta_j}\right) \qquad (10.8)
\end{aligned}$$

where m number of η values are simulated from the posterior density given in expression (10.7) and substituted into equation (10.8). The j-th simulation run is denoted by η_j. The average over all the tail probabilities is then used as an estimate of the posterior predictive tail probability.

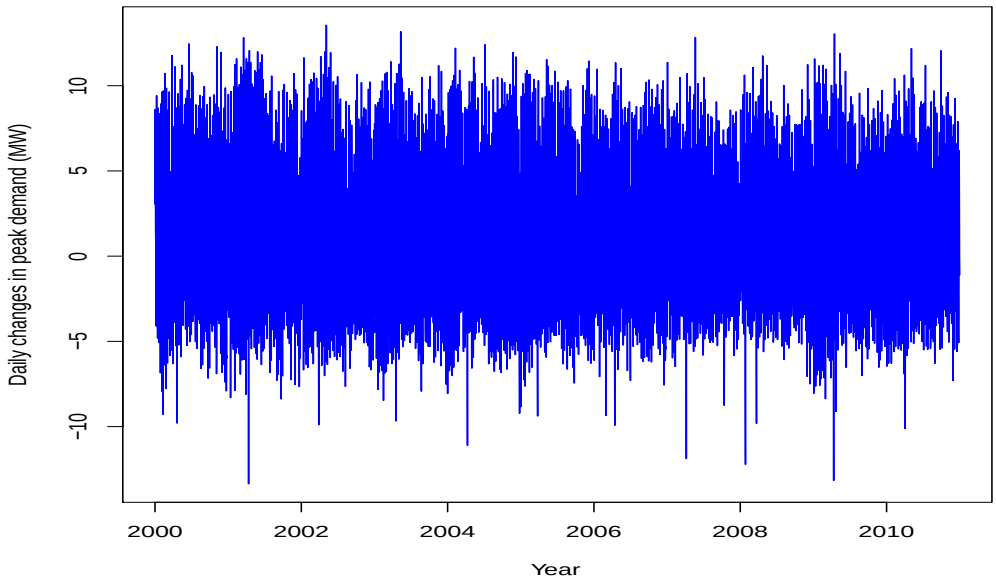

Figure 10.1: Plot of DPED return series (2000-2010).

10.3 Empirical results

Hourly electricity data is collected for years 2000 through to 2010 from Eskom, South Africa's power utility company. The hourly data is then divided into blocks of 24 hours each. All hours in a 24 hour block are from the same date. In each block the maximum hourly demand is recorded, and is referred to as the Daily Peak Electricity Demand (DPED). The DPED data is initially transformed by taking natural logarithms so as to reduce the impact of heteroskedasticity. Formal unit root tests are conducted using the Augmented-Dickey Fuller test. Results indicate that the natural logarithm of the first difference of DPED is stationary.

The DPED return series given in Figure 10.1 shows that volatility occurs in bursts and exhibits the presence of volatility clustering. The kernel density of DPED return series given in Figure 10.2 shows that the empirical distribution of the data is non-normal. The density is estimated using kernel density

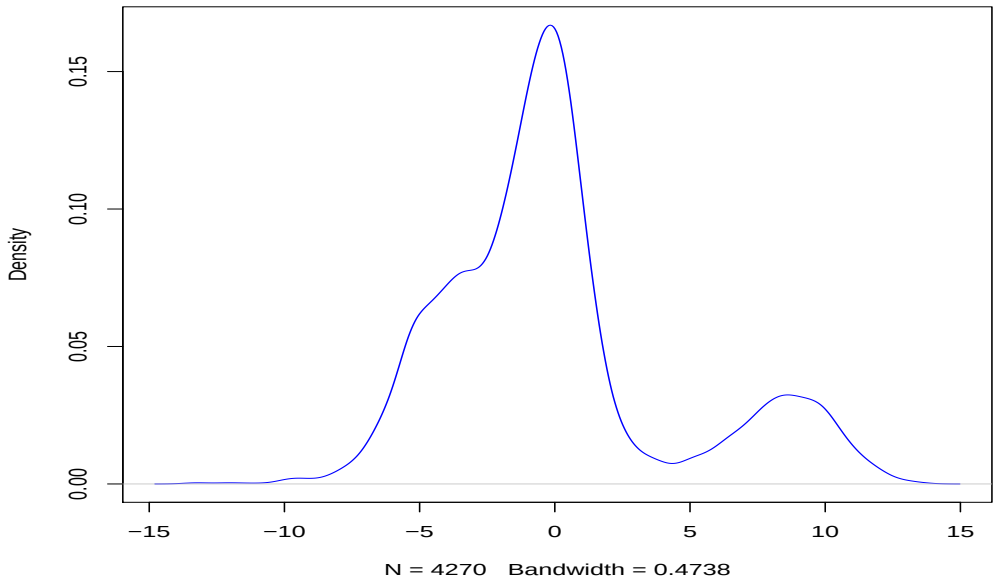

Figure 10.2: Kernel density of DPED return series (2000-2010). Distribution is nonnormal.

estimation (Silverman, 1986).

10.3.1 ARMA(*p,q*)-EGARCH(1,1) model results

In Table 10.1 we present descriptive statistics of the return series data. The skewness and kurtosis presented in Table 10.1 show that the return series data are non-normal. The Jarque-Bera test is carried out to check whether the skewness and kurtosis are consistent with a normal distribution. Our ARMA(*p,q*)-

Table 10.1: Descriptive statistics of the returns.

Mean	Median	Max	Min	Std	Skew	Kurt	Jarque-Bera
0.012	-0.559	13.553	-13.365	4.346	0.908	3.517	635.4 (0.0000)

EGARCH(1,1) model is given as follows:

$$\text{Mean equation} : r_t = \phi_0 + \phi_1 r_{t-1} + \phi_2 r_{t-7} + \phi_3 r_{t-28} + \phi_4 r_{t-365} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-7} \tag{10.9}$$

and

$$\text{Variance equation} : \log\sigma_t^2 = \omega + \beta\log\sigma_{t-1}^2 + \alpha\left|\frac{\varepsilon_{t-1}}{\sigma_{t-1}}\right| + \gamma\frac{\varepsilon_{t-1}}{\sigma_{t-1}} \tag{10.10}$$

As shown in the previous chapters electricity demand in South Africa exhibits strong seasonality. For DPED, seasonality is strong over the week, month and year. The following terms are therefore included AR(7), AR(28), AR(365) and MA(7) in the model given in equation (10.9) in order to filter out this seasonality from the data before fitting the extreme value distributions. Several ARMA(*p,q*)-EGARCH(1,1) models are considered and the model with the smallest Akaike information criterion (AIC) is selected. The model parameters are estimated using the maximum likelihood method under the assumption that the errors are conditionally normally distributed. The estimates are obtained by the Berndt *et al*. (1974) algorithm using numerical derivatives. The parameter estimates of the best model along with their p-values in parentheses are presented in Table 10.2.

The Lung-Box test results given in Table 10.2 indicate that there is some autocorrelation remaining and most of the heteroskedasticity has been removed. It should be noted that it may not be possible to remove all autocorrelation because we are dealing with high-frequency data. Figure 10.3 shows that the residuals are very volatile and exhibit volatility clustering (top left panel), while the top right panel shows the kernel density of residuals, with the bottom left panel showing the normal QQ plot and the bottom right panel shows the kernel density of the positive residuals.

10.3.2 Threshold estimation

We fit a GSPD to the upper tail of the n=1993 positive residuals (under demand predictions, i.e. positive residual = actual demand - predicted demand > 0).

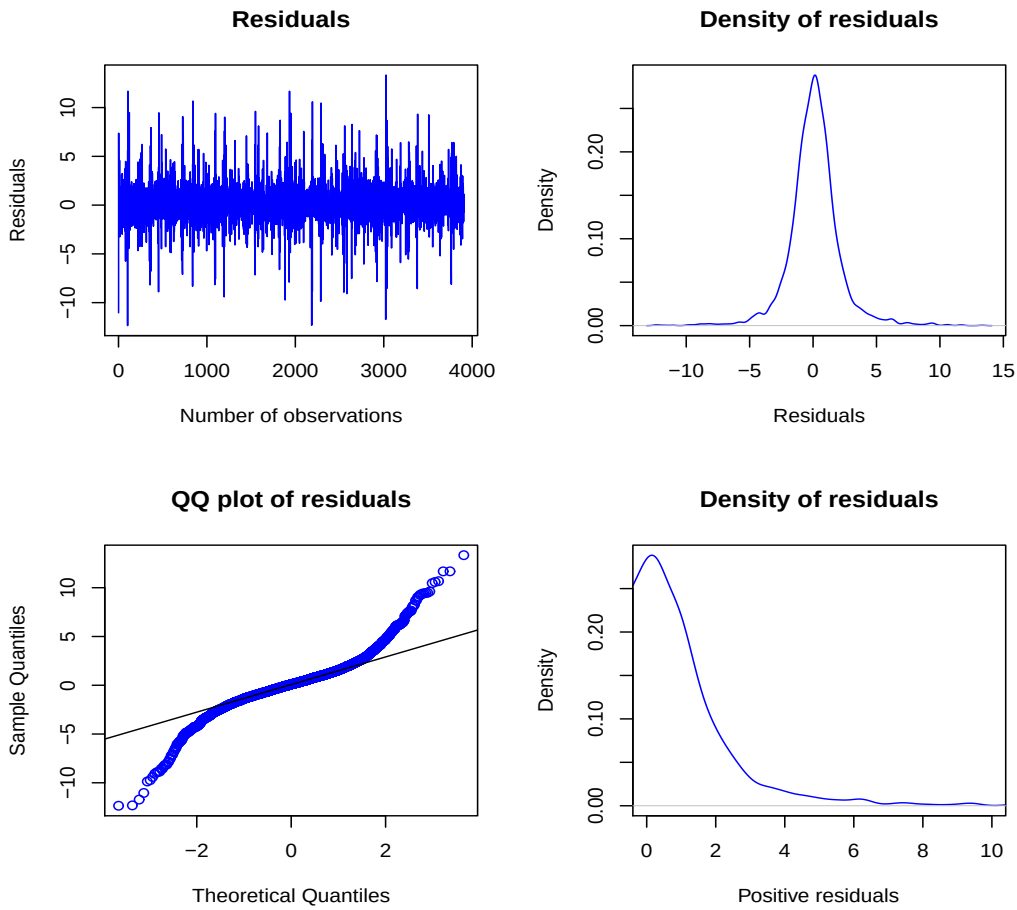

Figure 10.3: (a) Top left panel: Plot of residuals (b) Top right panel: Kernel density of residuals. Distribution is fat tailed (c) Bottom left panel: Normal QQ plot of residuals (d) Bottom right panel: Kernel density of positive residuals.

Table 10.2: ARMA(*p,q*)-EGARCH(1,1) model.

Mean Equation				
Parameter	ϕ_0	ϕ_1	ϕ_2	ϕ_3
Coefficient	-3.851(0.002)	0.006(0.070)	0.979(0.000)	0.021(0.004)
Parameter	ϕ_4	θ_1	θ_2	
Coefficient	0.009(0.009)	-0.135(0.000)	-0.932(0.000)	
Variance Equation				
Parameter	ω	α	β	γ
Coefficient	-0.115(0.000)	0.349(0.000)	0.867(0.000)	-0.249(0.000)
Model Diagnostics				
[a] $Q(7)$	$Q(28)$	$ARCH(7)$	$ARCH(28)$	
248.3(0.000)	376.9(0.000)	0.059(0.032)	0.012(0.384)	

[a] Q(7) is the Ljung-Box tests for serial correlations in the standardized residuals with 7 lags while ARCH(7) is Engle's LM test of ARCH effects up to the 7th order. The p-values are shown in parentheses. In all cases 5% level of significance is used.

This is defined as

$$\varepsilon_t = r_t - f_t > 0$$

where ε_t is the error term, r_t represents return series and predicted return series is denoted by f_t. We have considered the risk of under demand prediction of interday increases in peak electricity demand. However it should be noted that over demand prediction risk analysis (i.e. negative residual = actual demand - predicted demand < 0 or $\varepsilon_t = r_t - f_t < 0$) can be done in the same way. A Pareto quantile plot is used to obtain the threshold. The observation on the y-axis where the plot starts to follow a straight line is taken as the threshold. In this case $\tau = \exp(2) = 7.3891$. There are 26 exceedances. The Pareto quantile plot is shown in Figure 10.4.

10.3.3 GSPD parameter estimates

We now consider the error terms greater than $\tau = 7.3891$ to be generalized single Pareto distributed. The parameter η is estimated, using the Bayesian approach. We use the mode of the the posterior as an estimate of η and is

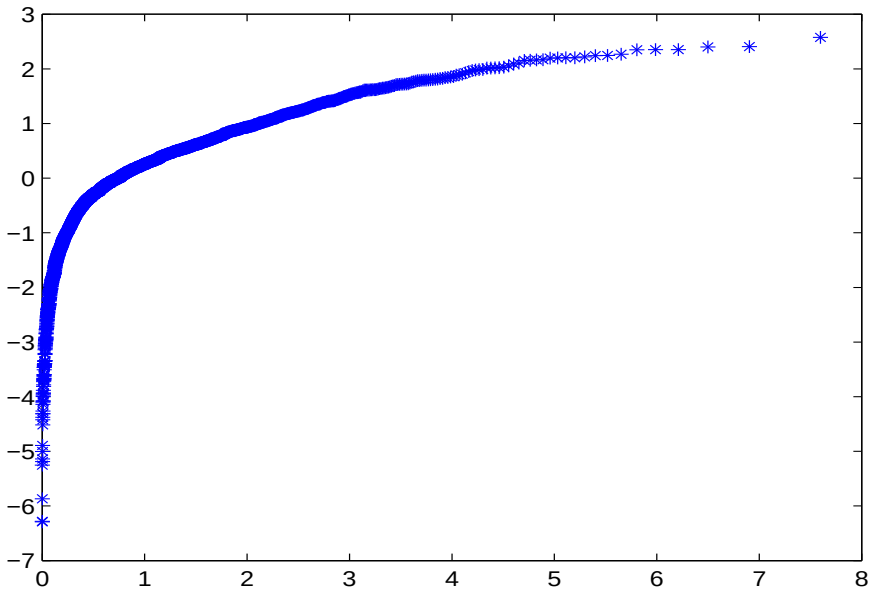

Figure 10.4: Generalized Pareto quantile plot on the positive residual (ε_t) observations.

found to be $\hat{\eta} = 0.072$. The QQ plot of the residual observations in Figure 10.5 suggests that the GSPD is a relatively good fit to the data. The conditional GSPD quantiles of the residual distribution are now estimated using the quantile function $\varepsilon_{t,p}$ given in equation (10.11) after substituting in the estimated parameter values.

$$\varepsilon_{t,p} = \frac{1+\tau\eta}{\eta}\left(p^{-\eta} - 1\right), \eta \neq 0, \varepsilon_{t,p} > \tau \tag{10.11}$$

The upper tail probability p is calculated using equation (10.12).

$$p = \left(\frac{1+\tau\eta+\eta\varepsilon_{t,p}}{1+\tau\eta}\right)^{-\frac{1}{\eta}}, \eta \neq 0, \varepsilon_{t,p} > \tau \tag{10.12}$$

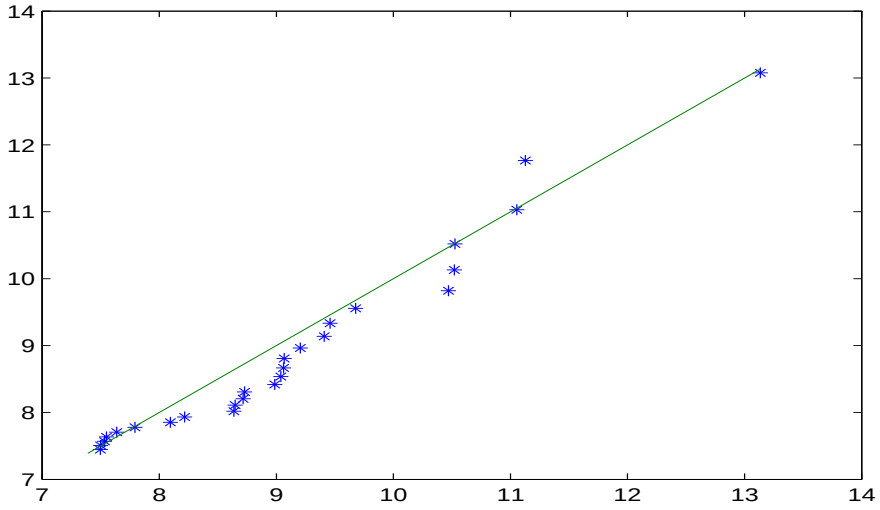

Figure 10.5: Normal QQ plot of $(\varepsilon_{t,p})$ above $\tau = 7.3891$. The horizontal axis represents the standard theoretical quantiles, and the empirical quantiles are plotted on the vertical axis.

The conditional tail quantiles, $r_{t,p}$ of the original return distribution are calculated as

$$r_{t,p} = \phi_0 + \phi_1 r_{t-1} + \phi_2 r_{t-7} + \phi_3 r_{t-28} + \phi_4 r_{t-365} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-7} + \sigma_t \varepsilon_{t,p} \quad (10.13)$$

where $\phi_0 + \phi_1 r_{t-1} + \phi_2 r_{t-7} + \phi_3 r_{t-28} + \phi_4 r_{t-365} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-7}$ and σ_t are the conditional mean and volatility from the ARMA-EGARCH model.

10.3.4 Evaluation of estimated tail quantiles at different probabilities

The estimated tail quantiles at different probabilities using the conditional GSPD are evaluated. The estimated number of exceedances is then compared to the exceedances from fitting ARMA-EGARCH model. A summary of the results is given in Table 10.3. The observed number of exceedances of a 90% quantile which is 3.0353 (n=1993 residuals) is obtained for example as follows:

Table 10.3: Estimated tail quantiles (number of exceedances).

Quantiles	Observed ($\varepsilon_{t,p}$)	ARMA-EGARCH	ARMA-EGARCH-GSPD
90^{th}	3.0353 (199)	2.7118 (255)	3.8369 (138)
95^{th}	4.5264 (100)	3.4505 (172)	5.1123 (75)
98.715^{th}	7.3891 (26)	3.8961(131)	7.8196 (20)
99^{th}	7.7914 (20)	4.8360 (84)	8.3474 (18)
99.5^{th}	9.0606 (10)	5.3433 (69)	9.8598 (6)
99.9^{th}	10.5256 (2)	6.3892 (39)	13.6757 (0)

($0.1 \times 1993 = 199$). The density of the residuals from the ARMA-EGARCH model in Figure 10.3 top right panel is used for estimating the quantiles given in column 3 of Table 10.3.

Using the quantile function given in equation (10.11) the 90^{th} quantile using the GSPD is given as:

$$\varepsilon_{t,0.1} = \frac{1 + 7.3891(0.072)}{0.072}\left(0.1^{-0.072} - 1\right) = 3.8369$$

where 7.3891 is the threshold and 0.072 is the Bayesian estimate of η. The number of observations that are larger than the estimated tail quantile ($\varepsilon_{t,0.1} = 3.8299$) are then counted and found to be 138. Overall the ARMA-EGARCH-GSPD model produces more accurate estimates of extreme tails than a pure ARMA-EGARCH model above the threshold, $\tau = 7.3891$ as shown in Table 10.3 and Figure 10.6.

10.3.5 Frequency analysis of exceedances (by month)

There are 26 exceedances above the threshold ($\tau = 7.3891$). A summary of the monthly frequency analysis of the exceedances over the sampling period (2000-2010) is presented in Table 10.4 and the bar chart is given in Figure 10.7.

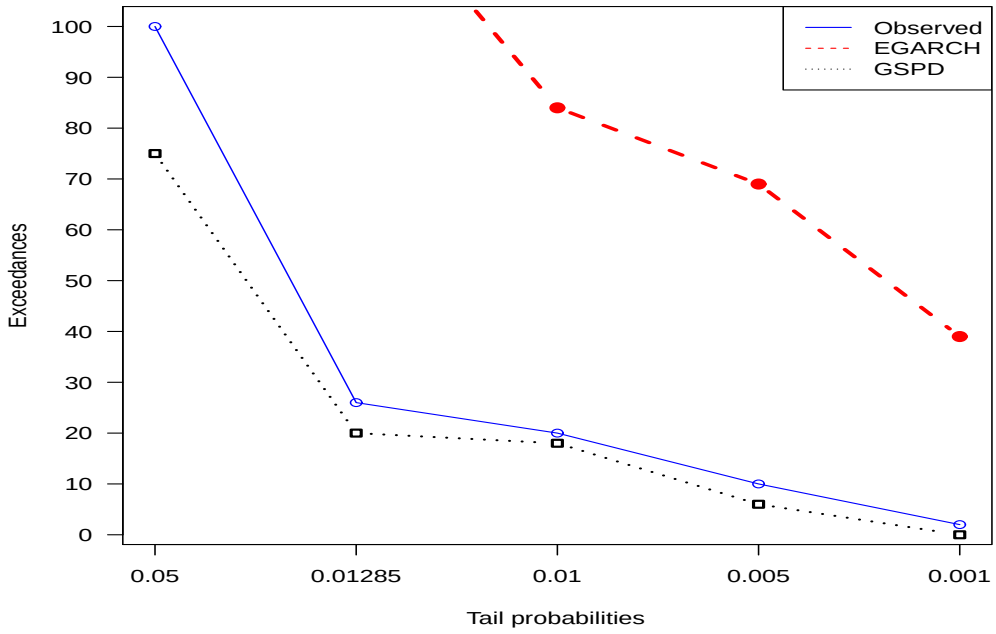

Figure 10.6: Plot of $P(X_0 > x_0|\boldsymbol{x}, \tau)$ for Observed (solid line), GSPD (dotted line below solid line) and ARMA-EGARCH (dashed line above solid line).

Table 10.4: Monthly frequency of exceedances above the threshold ($\tau = 7.3891$).

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Freq	7	0	2	9	3	1	0	2	0	0	0	2

Over the sampling period large under demand predictions from the developed model (ARMA-EGARCH-GSPD) are most frequently experienced in April followed by January. This frequency analysis of extreme under demand prediction is important to system operators and decision makers in the electricity sector as it helps them in planning and scheduling electricity.

The upper panel of Figure 10.8 shows a plot of exceedances $(\varepsilon_{t,p})$ above the threshold of $\tau = 7.3891$ and the lower panel shows the kernel density of the exceedances.

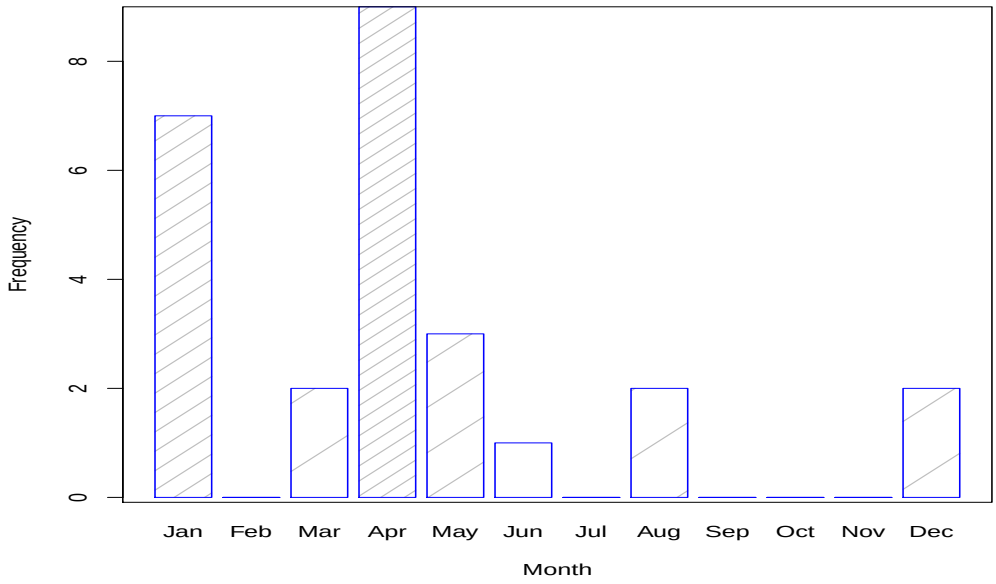

Figure 10.7: Bar chart of the monthly frequency of occurrence of exceedances above the threshold ($\tau = 7.3891$).

10.4 Concluding remarks and policy implications

In this chapter we have considered the risk of under demand prediction of interday increases in peak electricity demand (i.e. positive residual = actual demand - predicted demand > 0) using an ARMA-EGARCH-GSPD model. Under demand predictions can severely affect the operations of all the sectors of an economy and may lead to grid instability, blackouts and load shedding. Although we analysed under demand predictions it should be noted that over demand prediction risk analysis (i.e. negative residual = actual demand - predicted demand < 0) can be done in the same way. Underestimation demand analysis (positive residual analysis) provides power utility decision makers with an important and useful tool which can assist in assessing and quantifying the risk of under demand predictions of peak electricity demand particularly during periods of extreme weather conditions. Empirical results show that the ARMA-EGARCH-GSPD model produces more accurate estimates of

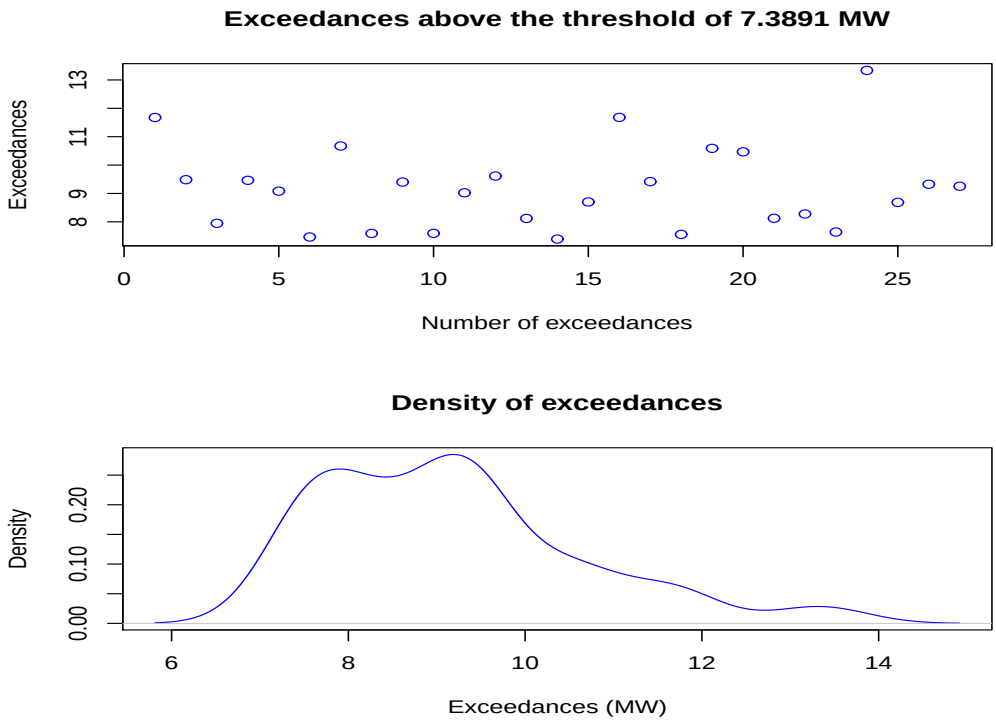

Figure 10.8: (a) Upper panel: plot of exceedances above the threshold of $\tau = 7.3891$ and (b) Lower panel: Kernel density of the exceedances. The distribution is right skewed.

extreme tails than a pure ARMA-EGARCH model.

Policy implications derived from this study are that demand side managers, decision and policy makers in the electricity sector should play a leading role in achieving behavioural change in electricity usage particularly during peak periods. Policy makers should design demand side management strategies where users are exposed to electricity price-based incentives such as tax incentives initiated by the South African National Energy Development Institute (SANEDI, (n.d.)). One of these incentives could be in the form of an energy efficiency savings allowance to users who save electricity in peak periods. The allowance could be in the form of a tax rebate (South African National Energy ACT, 2008). All energy users should be encouraged to use energy efficient technologies. There is need for energy efficient target monitoring systems. Use of smart technologies (e.g. smart meters) in all sectors of the economy should be encouraged. Properly designed time of use tariffs can be used to discourage high electricity usage during peak periods.

The chapter and the other previous chapters concentrated on short term peak demand modelling which is important for load flow analysis, scheduling and dispatching of electrical energy. For strategic decision making such as planning for construction of new generating plants, long term forecasts are required. We present in chapter 10 a brief discussion on long term peak demand forecasting.

Appendix 10: Log return

The log return is defined as

$$r_t = (1 - B)\log x_t = \log x_t - \log x_{t-1} = \log \frac{x_t}{x_{t-1}} = \log(1 + R_t)$$

where x_t, x_{t-1} are the current and one period lagged DPED respectively, B is the backward shift operator and R_t is the simple return defined as

$$R_t = \frac{x_t - x_{t-1}}{x_{t-1}}$$

For the average return, $r_t(k)$ over k periods we have

$$\begin{aligned} r_t(k) &= \log\{1 + R_t(k)\} = \frac{1}{k}\log\prod_{j=0}^{k-1}(1 + R_{t-j}) \\ &= \frac{1}{k}\sum_{j=0}^{k-1}\log(1 + R_{t-j}) \\ &= \frac{1}{k}\sum_{j=0}^{k-1} r_{t-j} \end{aligned}$$

Chapter 11

Annual peak electricity demand forecasting in South Africa

11.1 Introduction

Load forecasting is carried out for various time frames, from short term to long term forecasting. Short term forecasting is important for scheduling of electricity and for daily system operations (Fan and Hyndman, 2012). On the other hand long term electricity demand is important for medium term risk assessment and strategic planning for capacity expansion, power system planning, power security and supply reliability (Ismail *et al*., 2009; Hyndman and Fan, 2010). Long term peak electricity demand forecasting ranges from several months to several years ahead. This thesis would be incomplete without mentioning the long term electricity demand forecasting.

The use of the Generalized Gamma Distribution (GGD) in predicting annual

hourly peak electricity loads is discussed in this chapter. We model the annual peak load as a function of Gross Domestic Product (GDP) using the Bayesian framework. An accurate prediction of peak annual hourly electricity demand is important to decision makers in the electricity sector for strategic planning particularly in capacity expansion and risk assessment.

11.2 The Generalized Gamma Distribution (GGD)

The Generalized Gamma Distribution (GGD) was introduced by Stacy (1962). It is a flexible family of distributions which includes the gamma, exponential, Weibull, including the lognormal as a limiting distribution. If $Z \sim \text{GGD}(k, \alpha, \beta)$ then the probability density function is given by (Beirlant *et al.*, 2002)

$$f(z) = \frac{\alpha^k}{\beta\Gamma(k)} e^{-\alpha z^{\frac{1}{\beta}}} z^{\frac{k}{\beta}-1}, z > 0 \tag{11.1}$$

where k is the nuisance (shape) parameter, β is the tail index (shape parameter) and the scale parameter is α. When $\beta = k = 1$ we get the exponential distribution given in equation (11.2)

$$f(z) = \alpha e^{-\alpha z}, z > 0 \tag{11.2}$$

If $\beta = 1$ the GGD reduces to a gamma distribution

$$f(z) = \frac{\alpha^k}{\Gamma(k)} e^{-\alpha z} z^{k-1}, z > 0 \tag{11.3}$$

and the Weibull distribution when $k = 1$

$$f(z) = \frac{\alpha}{\beta} e^{-\alpha z^{\frac{1}{\beta}}} z^{\frac{1}{\beta}-1}, z > 0 \tag{11.4}$$

As $k \longrightarrow \infty,\ f(z)$ tends to the lognormal distribution (Lawless, 1980).
The GGD can be used in predicting peak electricity demand given predictor variables such as economic, calendar and weather. In this thesis we consider annual hourly peak electricity demand with GDP as a regressor variable. Let Z (annual hourly peak electricity demand) be the dependent variable where $Z \sim$ GGD(α, β, k) and Y denote the independent variable (GDP). To obtain parameter estimates on Z we introduce Z into the model in the form of a linear regression, by using the fact that $\mu = -E(\log(Z))$ and $\sigma = \sqrt{Var(-\log(Z))}$. Let $-E(\log(Z)) = \beta_0 + \beta_1 \log(\text{GDP})$. This equates to the following regression:

$$-E(\log(Z)) \quad = \quad \beta_0 + \beta_1 \log(\text{GDP}) = \mu \tag{11.5}$$

The parameters β_0 and β_1 are then expressed in terms of the shape parameter, k. A distribution function for k is then constructed by applying Bayesian inference. A detailed discussion of this modelling approach is found in De Waal *et al*. (2011). We use the Maximal Data Information (MDI) prior discussed in

Zellner (1977). The MDI prior for k is given by

$$\begin{aligned}
\pi(k) &\propto \exp\left\{E\left[\log(z|k)\right]\right\} \\
&\propto \exp\left\{E\left[\log\left(\frac{\alpha^k}{\beta\Gamma(k)}e^{-\alpha z^{\frac{1}{\beta}}}z^{\frac{k}{\beta}-1}\right)\right]\right\} \\
&\propto \exp\left\{E\left[k\log(\alpha)-\log(\beta\Gamma(k))-\alpha z^{\frac{1}{\beta}}+\left(\frac{k}{\beta}-1\right)\log(z)\right]\right\} \\
&\propto \exp\left[k\log(\alpha)-\log(\beta\Gamma(k))-\alpha E\left(z^{\frac{1}{\beta}}\right)+\left(\frac{k}{\beta}-1\right)E\left(\log(z)\right)\right] \\
&\propto \frac{1}{\Gamma(k)}\times\frac{\alpha^k}{\beta}\exp\left[-\alpha E\left(z^{\frac{1}{\beta}}\right)+\left(\frac{k}{\beta}-1\right)E\left(\log(z)\right)\right] \\
&\propto \frac{1}{\Gamma(k)}e^{-k+(k-1)\Psi(k)} \qquad (11.6)
\end{aligned}$$

where $\Psi(k)=\frac{\partial\log\Gamma(k)}{\partial k}$ represents the digamma function. The likelihood function is

$$\begin{aligned}
\pi(z|k) &= \prod_{i=1}^{n} f(z_i,k) \\
&= \prod_{i=1}^{n}\frac{\alpha^k}{\beta\Gamma(k)}e^{-\alpha z_i^{\frac{1}{\beta}}}z_i^{\frac{k}{\beta}-1} \\
&= \left(\frac{\alpha^k}{\beta\Gamma(k)}\right)^n\prod_{i=1}^{n}e^{-\alpha z_i^{\frac{1}{\beta}}}z_i^{\frac{k}{\beta}-1} \\
&= \left(\frac{\alpha^k}{\beta\Gamma(k)}\right)^n e^{-\sum_{i=1}^{n}\alpha z_i^{\frac{1}{\beta}}}\prod_{i=1}^{n}z_i^{\frac{k}{\beta}-1} \\
&= \frac{\alpha^{nk}}{\beta^n\Gamma(k)^n}e^{-\alpha\sum_{i=1}^{n}z_i^{\frac{1}{\beta}}}\prod_{i=1}^{n}z_i^{\frac{k}{\beta}-1} \qquad (11.7)
\end{aligned}$$

Combining the prior in expression (11.6) with the likelihood in equation (11.7) we get the following posterior distribution of k:

$$\begin{aligned}
\pi(k|z) &\propto \pi(z|k)\pi(k) \\
&\propto \frac{1}{\Gamma(k)} e^{-k+(k-1)\Psi(k)} \times \frac{\alpha^{nk}}{\beta^n \Gamma(k)^n} e^{-\alpha \sum_{i=1}^n z_i^{\frac{1}{\beta}}} \prod_{i=1}^n z_i^{\frac{k}{\beta}-1} \\
&\propto \left(\frac{\alpha^k}{\beta\Gamma(k)}\right)^n \times \frac{e^{-k+(k-1)\psi(k)}}{\Gamma(k)} \times e^{-\alpha \sum_{i=1}^n z_i^{\frac{1}{\beta}}} \prod_{i=1}^n z_i^{\frac{k}{\beta}-1} \\
&\propto \frac{1}{\beta^n (\Gamma(k))^{n+1}} \times \alpha^{nk} e^{-\left[k+(k-1)\psi(k)+\alpha \sum_{i=1}^n z_i^{\frac{1}{\beta}}\right]} \times \left(\prod_{i=1}^n z_i\right)^{\frac{k}{\beta}-1} \qquad (11.8)
\end{aligned}$$

A large number of k's are then simulated from the posterior distribution of k. The predictive density of Z_{n+1} given Y is given as

$$\begin{aligned}
\text{Pred}(Z_{n+1}|Y_{n+1}) &= E\left[f(z|k,\alpha,\beta)\right] \\
&\approx \frac{1}{m} \sum_{j=1}^m f(k_j,\alpha,\beta) \qquad (11.9)
\end{aligned}$$

11.3 Data

Aggregated hourly Annual Peak Electricity Demand (APED) data is used for the industrial, commercial and residential sectors of South Africa. Modelling of APED is done using data for fifteen years from 1996 to 2010. APED is the maximum hourly demand in a year, i.e. in 8760 hours. In South Africa the annual peak electricity is usually in July (which is a winter month in the Southern Hemisphere) as shown by the monthly seasonal demand index plot in Figure 11.1.

Whilst there are several drivers of electricity demand, in this chapter we

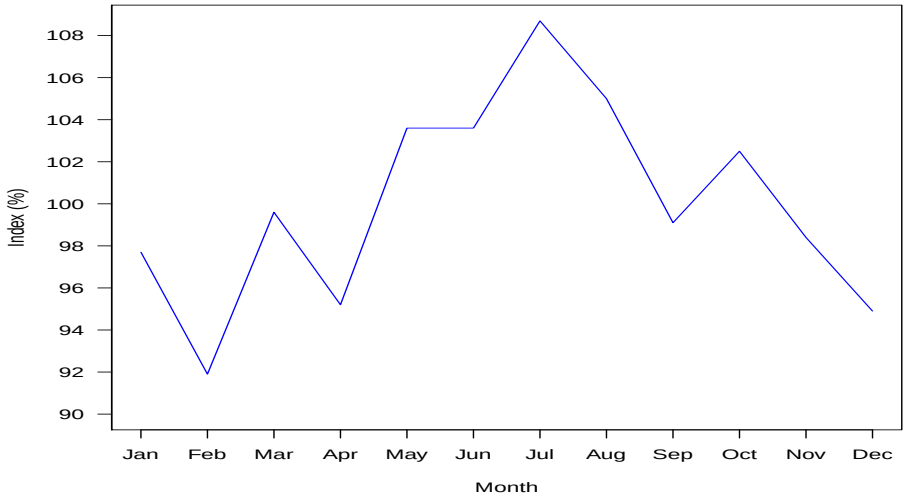

Figure 11.1: Monthly seasonal index plot of Daily Peak Electricity Demand (DPED). The index for each month is calculated by finding the average DPED for each month and then divide by the overall average. Annual peak electricity demand is in July of each year.

have GDP as the major driver. There is a strong positive relationship between GDP and APED as shown in Figure 11.2. The Pearson's correlation coefficient between GDP and APED over the sampling period is 0.975. Figure 11.2 shows that APED and GDP both have an upward positive trend and are non-stationary. Generally there are four cases to consider when working with non-stationary variables. We shall consider a simple linear regression model[1], $z_t = \beta_0 + \beta_1 y_t + u_t$ where z_t is the dependent variable, y_t is the explanatory variable and u_t is the error term. The four cases to consider are:

1. when both variables, z_t and y_t are stationary then the estimated coefficients will be appropriate,

2. If both z_t and y_t are integrated of different orders then regressing z_t on y_t is meaningless.

[1]"An important consideration in estimating a regression with ARMA errors is that all variables in the model must first be stationary" (Hyndman and Athanasopoulos, 2012)

3. when both variables are integrated of the same order and the residuals are not white noise we have a case of spurious regression and the results are meaningless and

4. when both z_t and y_t are are integrated of the same order and the residuals follow a white noise process, then the variables are said to be cointegrated and the regression model is appropriate.

The non stationary APED and GPD are integrated of order 1 (after first differences they both become stationary).

Over the years 1996 to 2010 APED has been increasing steadily, an indication of increased demand for electricity by all sectors of the economy after South Africa had gained democracy in 1994. This sharp increase in electricity demand is a policy concern to decision makers in Eskom, South Africa's power utility company as they have to balance supply with demand. This calls for accurate predictions of APED to enable decision makers to make strategic decisions with regards to capacity expansion and risk assessment. Fitting the GGD to APED before differencing we get the following parameter estimates:

$$k = 1.0079$$

$$\alpha = 85.121$$

$$\beta = 396.27$$

Since $k \approx 1$ the density of APED follows the Weibull class of distributions. The quantile-quantile (QQ) and probability-probability (PP) plots given in Figure 11.3 show that the Weibull class of distributions is a good fit to APED.

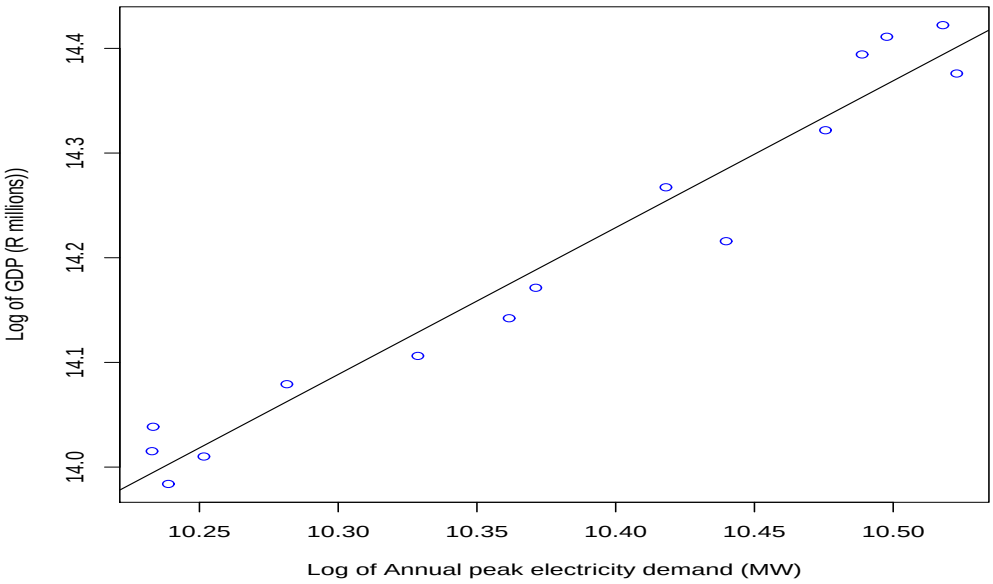

Figure 11.2: Plot of logarithm of GDP against logarithm of APED.

11.4 Empirical results

Estimating $\boldsymbol{\beta} = \{\beta_0, \beta_1\}$ by least squares method from equation (11.5) we get $\beta_0 = 0.0253$ and $\beta_1 = 0.3029$. The parameters β_0 and β_1 are then expressed in terms of the shape parameter, k of the GGD. A distribution function for k is then constructed by applying Bayesian inference as discussed in Section 11.2.

The Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) are used for comparing the models. Table 11.1 shows a comparative analysis of the Bayes (GGD) model with a Seasonal Naive (SNAIVE) model. The SNAIVE model is given as

$$(1 - B^s)Y_t = \varepsilon_t$$

$$Y_t = Y_{t-s} + \varepsilon_t \tag{11.10}$$

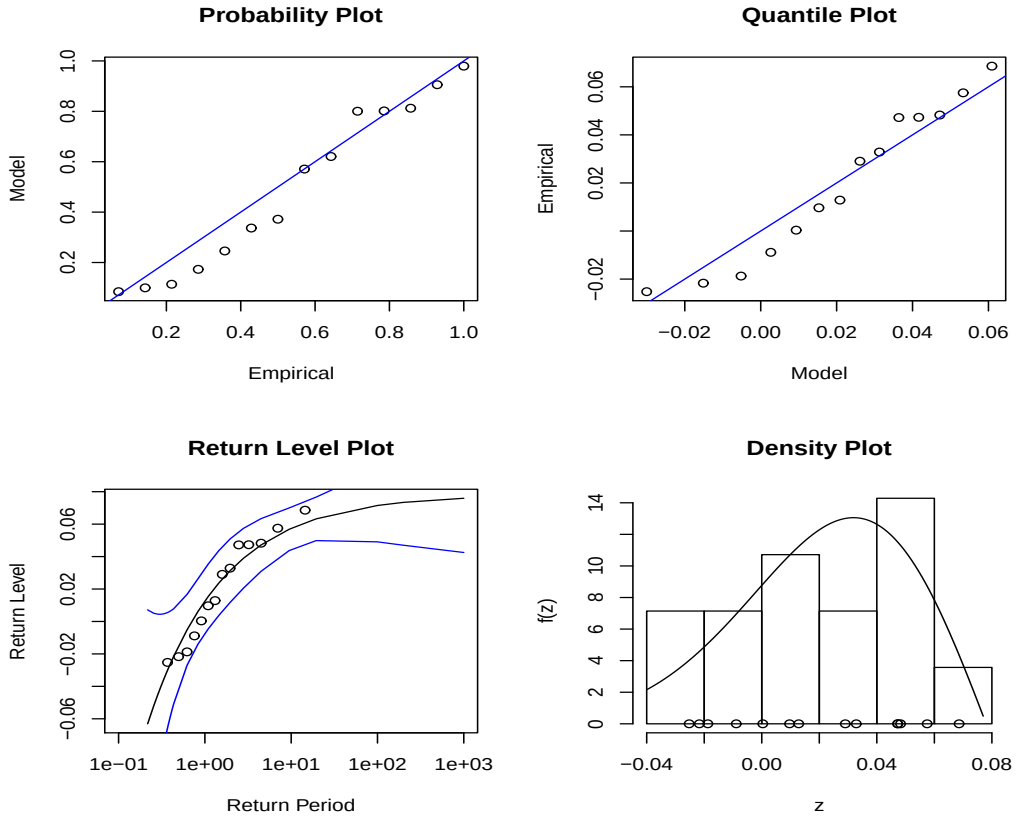

Figure 11.3: Diagnostic plots: (a) Top left panel: Probability plot (b) Top right panel: Quantile plot (c) Bottom left panel: Return level plot and (d) Bottom right panel: Density plot. The distribution is left skewed.

where B is lag operator, s is the seasonal length, Y_t denotes the annual peak electricity demand and ε_t is the error term. The Bayes (GGD) model has the

Table 11.1: In-sample evaluation.

Method	RMSE	MAE	MAPE (%)
Bayes method (GGD)	799.1	686.9	2.13
Seasonal naive method	1156.3	936.2	2.81

least MAPE, RMSE and MAE, showing that it can be used for forecasting annual hourly peak electricity demand.

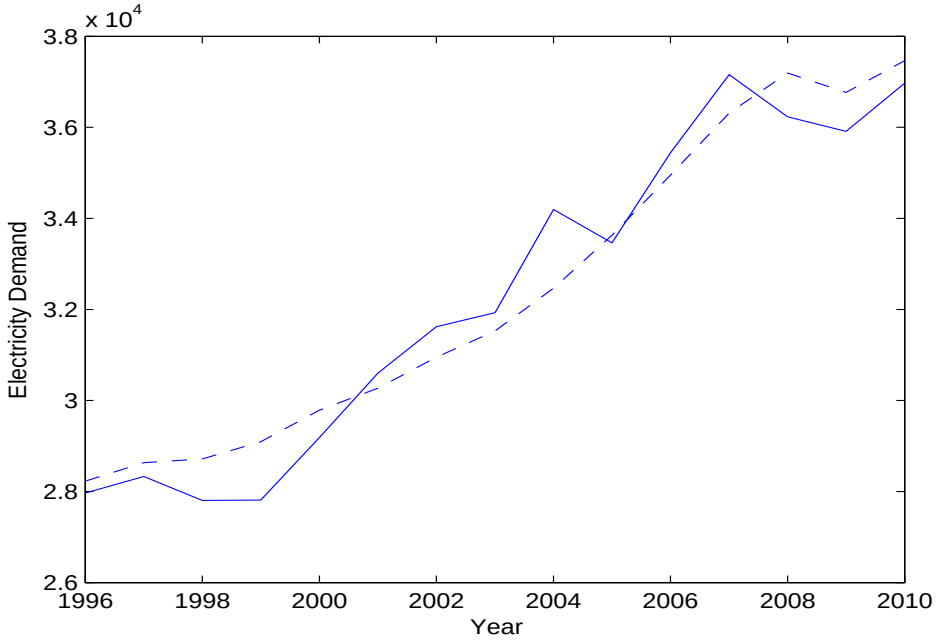

Figure 11.4: Plot of APED (solid line) with the fitted APED using GGD (broken line).

11.5 Concluding remarks

In this chapter we explored the use of the Generalized Gamma Distribution (GGD) in modelling hourly annual electricity peak demand in South Africa. Empirical results show that the GGD can be used in predicting annual peak electricity demand into the long term.

Areas for further study would include the addition of other explanatory variables. It will also be interesting to explore other methods for long term forecasting such as semi-parametric additive models (Debba *et al.*, 2010; Hyndman and Fan, 2010; Hong *et al.*, 2013; among others).

Chapter 12

Conclusion

12.1 Introduction

In this thesis we presented and analysed a number of statistical models for modelling daily peak electricity demand in South Africa. The thesis concentrated on short term daily peak electricity demand forecasting and modelling of extreme daily peak electricity demand. Modelling of short term peak electricity demand is very important to system operators and decision makers in the electricity sector in load flow analysis, scheduling of start-up times for peak stations, in the determination of consistent and reliable supply schedules particularly during peak periods and optimal distribution of electrical energy. Modelling of extreme peak electricity demand helps in determining critical peak days, risk management including load shifting between transmission substations which is important for the stability of a power network.

12.2 Summary and concluding remarks

Modelling the influence of temperature on daily peak electricity demand using Multivariate Adaptive Regression Splines (MARS) and piecewise linear regression models was discussed in Chapter 3. Using the MARS algorithm the threshold temperatures which separate the weather neutral period from the winter and summer sensitive periods were determined. The threshold temperatures determined by the MARS algorithm were found to be 18^0C and 22^0C. Above 22^0C electricity demand tends to rise slightly and below 18^0C electricity demand increases significantly. Within the weather neutral period residents would neither use a heater nor a cooling system. Chapter 3 established temperature as an important variable in explaining electricity demand in South Africa.

Chapter 4 extended the analysis discussed in Chapter 3 by modelling the effect of temperature on average daily electricity demand in South Africa using a piecewise linear regression model. Empirical results were found to be consistent with findings discussed in Chapter 3 that electricity demand in South Africa is more sensitive to cold temperatures than to hot temperatures.

The use of the Bayesian parameter estimates to a piecewise linear regression model in explaining the effect of temperature on daily peak electricity demand in South Africa was discussed in Chapter 5. Probability statements were made about the uncertainty in the estimated parameters. The modelling approach exploits patterns in electricity demand. This empirical study provides an extension of point forecasting to density forecasting so as to take into account uncertainty in the estimation of the parameters.

Chapter 6 investigated some hybrid models for daily peak electricity demand forecasting. Three regression-seasonal autoregressive integrated moving aver-

age (Reg-SARIMA) models were developed. The Reg-SARIMA model in which the effect of temperature variables were included through regression splines produced the best forecast accuracy amongst the three Reg-SARIMA models discussed. The Reg-SARIMA model is simple to implement, reliable and provides information about the importance of each predictor variable. The results from using a regression-SARIMA model are relatively robust. The model assumes homoskedasticity. The second part of the chapter discussed volatility forecasting models. Empirical results of the Bayesian GARCH(1,1) model with Student-t innovations of the residual return series of DPED are presented and compared with results from Bayesian GARCH(1,1) with normal innovations The Bayesian GARCH(1,1) model with Student-t innovations seem to provide a better description of the volatility of the DPED residual return series data. A Reg-SARIMA-TGARCH model was developed and used in DPED under demand risk estimation for in-sample predictions. A comparative analysis was done with the Reg-SARIMA model which modelled temperature through regression splines. Although both models produced results which were comparable the Reg-SARIMA-TGARCH seem to capture well the volatility in the residuals during the year 2008 when South Africa experienced countrywide blackouts and load shedding.

Chapters 3 to 6 established that electricity demand in South Africa is highly sensitive to cold temperatures (below 18^0C) than to hot temperatures (above 22^0C). Electricity demand in South Africa is not only influenced by average demand but by extreme demand as well. In order to have a better understanding of peak electricity demand for temperature values below 18^0C Chapter 7 presented an analysis of the effect of extreme low temperature on daily electricity demand in South Africa using the generalized extreme value theory approach. Empirical results showed that extreme low temperatures can be modelled by the Weibull class of distributions. Extreme low average daily temperatures

of the order of 8.2^0C are very rare in South Africa. They only occur about 8 times in a year and result in huge increases in electricity demand. Electricity demand in South Africa is highly sensitive to temperature fluctuations during the winter periods. A comparative analysis was done with Australia on the effect of temperature on electricity demand. It was noted that in Australia electricity demand is highly sensitive to summer months (Hyndman and Fan, 2010) despite the fact that both Australia and South Africa are on the same latitude in the Southern Hemisphere. It was noted that this difference could be due to the fact that the largest part of Australia is desert and semi-arid while South Africa is on a greater elevation above sea level.

Chapter 7 established that the use of extreme value theory provides a lot of information in the modelling of peak electricity demand. Chapter 8 then presented a detailed analysis of extreme daily and same day of the week increases in peak electricity demand in South Africa using extreme value theory. Extreme daily increases in peak electricity demand were modelled using a Generalized Pareto Distribution (GPD) in the first part of Chapter 8. A comparative analysis of the GPD was done with the GSPD. A Pareto quantile plot was used to obtain the optimum threshold. Empirical results showed that both the GSPD and the GPD are a good fit to the extreme daily increases in peak electricity demand data. However the use of the GSPD is recommended. The main advantage of this modelling approach is that the GSPD has only one parameter to estimate, given that the threshold is known. The GSPD was used for estimating extreme tail quantiles and probability of exceedance values for various future extreme same day of the week increases in peak electricity demand in the second part of Chapter 8. A comparative analysis was done with the GPD. The GSPD and the GPD both provide a satisfactory fit to the data. Empirical results show that the modelling of extreme daily increases in peak electricity demand helps in determining the number of critical peak days and

also in improving the reliability of a power network if an accurate assessment of the frequency and level of future extreme load forecasts is carried out.

In an attempt to have a better understanding of peak electricity demand during the winter period Chapter 9 discussed an application of a Generalized Single Pareto Distribution (GSPD) to the statistical analysis of extreme daily winter peak electricity demand. The GSPD was found to be appropriate in modelling of South African extreme daily peak electricity demand data. A comparative analysis was done by fitting a Generalized Extreme Value Distribution (GEVD). Empirical results show that the Weibull class can be used to model the daily peak electricity demand data. The results from the GSPD were found to be better than those from fitting a GEVD.

Chapter 10 provided an extension of Chapters 8 and 9 to the modelling of the risk of under demand prediction of interday increases in peak electricity demand. Under demand predictions can severely affect the operations of all the sectors of an economy and may lead to grid instability, blackouts and load shedding. Although we analysed under demand predictions in this thesis it should be noted that over demand prediction risk analysis can be done using the same modelling approach. Underestimation demand analysis (positive residual analysis) provides power utility decision makers with an important and useful tool which can assist in assessing and quantifying the risk of under demand predictions of peak electricity demand particularly during periods of extreme weather conditions. Empirical results show that the ARMA-EGARCH-GSPD model produces more accurate estimates of extreme tails than a pure ARMA-EGARCH model. Some policy implications were discussed.

Chapters 3 to 10 focused on modelling short term daily peak electricity demand which is important for operational decision making such as optimal dis-

patching of electrical energy to consumers. For strategic decision making which involves, e.g. construction of new generating plants, long term peak demand forecasts are needed. Chapter 11 explored the use of the Generalized Gamma (GG) distribution in modelling hourly annual electricity peak demand in South Africa. Empirical results show that the GG distribution can be used in predicting annual peak electricity demand.

Use of frequentist and Bayesian methods produce results which are fairly comparable. However the Bayesian approach is more attractive since it provides a way of taking into account uncertainty in the estimation of the parameters. The Bayesian approach is more informative than the classical approach.

12.3 Summary of key findings and contributions

The key findings are summarized as follows:

1. Demand for electricity in South Africa is more sensitive to cold temperatures (less than 18^oC) than to hot temperatures (more than 22^oC). Electricity demand increases significantly for temperature values below 18^oC and rises slightly for temperature values above 22^oC. Extreme low temperatures of the order of 5^oC are very rare in South Africa, but can cause huge increases in electricity demand.

2. Modelling under and over demand electricity forecasts provides a basis for risk assessment and quantification of the risk associated with forecasting uncertainty including demand variability in South Africa.

3. An accurate assessment of the frequency and level of future extreme peak load forecasts help system operators and decision makers in determining the number of critical peak days as well as scheduling, load flow analysis and dispatching of electricity in South Africa.

4. Accurate forecasts of the magnitude and timing of peak electricity demand can help to determine the time to shift peak loads to off peak periods and also helps in planning for maintenance.

12.4 Limitations of the Thesis

One of the limitations of this thesis is that we only used aggregated hourly electricity demand for all the sectors of the South African economy which are: residential, industrial, commercial and agricultural sectors. It would have been interesting to analyse electricity demand by each sector of the South African economy and also by each of the nine provinces of the country. Average temperature data for the country was made available. The ideal average temperature would be the weighted temperature by, e.g. population size of each town or city would be a reasonable weight.

12.5 Future Research Directions

The results of this thesis provide possible areas for future research. The following possible future research directions are proposed:

- Future research may include use of a full Bayesian MARS model in which we place a prior distribution on the whole MARS model space, treating the number of splines, knot points and all other parameters as unknown.
- In long term forecasting it is important to capture uncertainty involving forecasts. This can be done through use of density forecasting of long term peak electricity demand as this provides full probability distributions of possible future load forecasts, (Hyndman and Fan, 2010). It would be interesting to carry out this study using South African data.
- In the thesis Jeffreys and maximal data information priors are used. The

use of robust Bayesian analysis in which a class of prior distributions is used instead of a single prior distribution (Chaturvedi, 1996) would be an interesting approach to use.

- In this thesis aggregated national hourly electricity data is used. Future research should look at modelling electricity demand for each of the nine provinces of South Africa separately including analysis by sector (i.e. residential, commercial, agricultural and industrial sectors). Some provinces are in the coastal areas and electricity demand patterns are expected to be different from provinces which are inland. Some of the provinces such as Gauteng are densely populated and heavily industrialized as compared to others.

- Another area which requires future research is a probabilistic description and modelling of extreme peak loads using Poisson point process. This modelling approach helps in estimating the frequency of occurrence of peak loads (Chidodo and Lauria, 2012).

- A sensitivity analysis with respect to daily and seasonal peak electricity demand performed for each of the provinces of South Africa and the development of a two-stage stochastic integer recourse models with the objective of optimizing electricity distribution is an interesting future research direction with South African data.

References

[1] AGUIRRE, L.A., RODRIQGUES, D.D., LIMA, S.T. and MARTINEZ, C.B. (2008). Dynamic prediction and pattern mapping in short-term load forecasting. *Electrical Power and Energy Systems*, 30, 73–82.

[2] ALFONS, A., HOLZER, J. and TEMPL, M. (2013). Laeken: R package version 0.4.4. `http://cran.r-project.org/web/packages/laeken/index.html`

[3] ALLEN, P.G. and MORZURCH, B.J. (1995). Comparing probability forecasts derived from theoretical distributions. *International Journal of forecasting*, 11, 147–157.

[4] AMARAL, L.F., SOUZA, R.C. and STEVENSON, M. (2008). A smooth transition periodic autoregressive (STPAR) model for short-term load forecasting. *International Journal of Forecasting*, 24, 603–615.

[5] AMJADY, N. (2001). Short-Term Hourly Load Forecasting Using Time Series Modeling with Peak Load Estimation Capability. *IEEE Transactions on Power Systems*, 16(3), 498–505.

[6] AMIN-NASERI, M.R. and SOROUSH, A.R. (2008). Combined use of unsupervised and supervised learning for daily peak load forecasting. *Energy Conversion and Management* 49, 1302–1308.

[7] AMUSA, H., AMUSA, K. and MABUGU, R. (2009). Aggregate demand for electricity in South Africa: An analysis using the bounds testing approach to cointegration. *Energy Policy*, 37, 4167–4175.

[8] APS http://www.aps.com/aps_services/residential/rateplans/ResRatePlans_28.html (Accessed 3 August 2012).

[9] ARDIA, D. (2011). BayesGARCH: Bayesian Estimation of the GARCH(1,1) Model with Student-t Innovations in R. R package version 1-00.10. http://CRAN.R-project.org/package=bayesGARCH

[10] ARDIA, D. and HOOGERHEIDE, L.F. (2010). Bayesian Estimation of the GARCH(1,1) Model with Student-t Innovations. The R Journal 2(2), pp.41-47. http://journal.r-project.org/

[11] ARDIA, D. (2006). Bayesian Estimation of the GARCH(1,1) Model with Normal Innovations. Student 5(3-4), pp.283-298. http://ssrn.com/abstract=1543409

[12] ARLOT, S. and CELISSE, A. (2010). A survey of cross-validation procedures for model selection. *Statistics Surveys*, 4, 40–79.

[13] AZADEH, A., GHADERI, S.F. and SOHRABKHANI, S. (2007). Forecasting electrical consumption by integration on neural network, time series and ANOVA. *Applied Mathematics and Computation*, 186, 1753–1761.

[14] BAKHAT, M. and ROSSELLO, J. (2011). Estimation of tourism-induced electricity consumption: The case study of Balearics Islands, Spain. *Energy Economics*, 33, 437–444.

[15] BALKEMA, A. and DE HAAN, L. (1974). Residual life time at great age. *Annals of Probability*, 2, 792–804.

[16] BARTHELMIE, R.J., MURRAY, F. and PRYOR, S.C. (2008). The economic benefit of short-term forecasting for wind energy in the UK electricity market. *Energy Policy*, 36, 1687–1696.

[17] BAYARRI, M.J. and BERGER, J.O. (2004). The Interplay of Bayesian and Frequentist Analysis. *Statistical Science*, 19(1), 58–80.

[18] BEHRENS, C.N., LOPES, H.F. and GAMERMAN, D. (2004). Bayesian analysis of extreme events with threshold estimation. *Statist. Mod.*, 4, 227–244.

[19] BERGER, J.O. (1985). *Statistical Decision Theory and Bayesian Analysis*, Second Edition, New York: Springer-Verlag.

[20] BERGER, J.O. and WOLPERT, R. (1988). *The Likelihood Principle*, 9, Second Edition, Hayward, California: Institute of Mathematical Statistics, monograph series.

[21] BEIRLANT, J., DE WAAL, D.J. and TEUGELS, J.L. (2002). The Generalised Burr-Gamma family of distributions with applications in extreme value analysis. *Limit theorems in Probability and Statistics I* (I. Berkes, E. Csaki, M. Csorgo, eds.). Budapest, 113–132.

[22] BEIRLANT, J., VYNCKIER, P. and TEUGELS, J.L. (1996). Tail index estimation, Pareto quantile plots and regression diagnostics. *J. Am. Statist. Assoc.*, 91, 1659–1667.

[23] BEIRLANT, J., GOEDGEBEUR, Y., SEGERS, J. and TEUGELS, J. (2004). *Statistics of Extremes Theory and Applications*: Wiley, England.

[24] BEIRLANT, J. and GOEGEBEUR, Y. (2003). Regression with response distributions of Pareto-type. *Computational Statistics and Data Analysis*, 42, 595–619.

[25] BEIRLANT, J. DE WET, T. and GOEGEBEUR, Y. (2006). A goodness-of-fit statistic for Pareto-type behaviour. *Journal of Computational and Applied Mathematics*, 186, 99–116.

P. de Zea Bermudez,

[26] BERMUDEZ, P. DE ZEA and KOTZ, S. (2010). Parameter estimation of the generalized Pareto distribution-Part II. *Journal of Statistical Planning and Inference*, 140, 1374–1388.

[27] BERNARDO, J.M. and SMITH, A.F.M. (1994). *Bayesian Theory*, New York: John Wiley & Sons.

[28] BERNDT, E.K., HALL, B.H., HALL, R.E. and HAUSMAN, J.A. (1974). Estimation and inference in nonlinear structural models. Annals of Economic and Social Measurement 4, 653–665.

[29] BOLLERSLEV, T. (1986). Generalized autoregressive conditional heteroscedasticity. Journal of Econometrics 31, 307–327.

[30] BORAINE, H. and YADAVALLI, V.S.S. (2003). Electricity load forecasting with artificial neural networks. *SA Journal of Industrial Engineering*, 14(2), 23–35.

[31] BOX, G.E.P. and JENKINS, G.M. (1976). *Time series analysis: forecasting and control*. Revised Edition, Holden-Day, California.

[32] BOX, G.E.P., JENKINS, G.M. and REINSEL, G.C. (1994). *Time Series Analysis: Forecasting and Control*, third ed. Prentice Hall, New Jersey.

[33] BOX, G.E.P. and TIAO, G.C. (1973). *Bayesian Inference in Statistical Analysis*, Wiley Classics Library Edition, published 1992, New York: John Wiley & Sons.

[34] BUNN, D.W. (1982). Short-term forecasting: a review of procedures in the electricity supply industry. *J Oper Res Soc*, 33, 533–545.

[35] BYSTROM, H.N.E. (2005). Extreme value theory and extremely large electricity price changes. *International Review of Economics and Finance*, 14(1), 41–55.

[36] CABRAS, S. and CASTELLANOS, M.E. (2010). An objective Bayesian approach for threshold estimation in the peaks over the threshold model, Technical Report TR2010.10, An*á*lisis de riesgo.

[37] CARLIN, B.P. and LOUIS, T.A. (2000). *Bayes and Empirical Bayes Methods for Data Analysis*, Second Edition, London: Chapman & Hall.

[38] CARREAU, J. and BENGIO, Y. (2009). A hybrid Pareto model for asymmetric fat-tailed data: the univariate case. *Extremes*, 12, 53–76.

[39] CASTILLO, E. and HADI, A.S. (1997). Fitting the Generalized Pareto Distribution to Data. *Journal of the American Statistical Association*, 92(440), 1609–1620.

[40] CHAN, K.F. and GRAY, P. (2006). Using extreme value theory to measure value-at-risk for daily electricity spot prices. *International Journal of Forecasting*, 22, 283–300.

[41] CHAN, K.F. GRAY, P. and VAN CAMPEN, B. (2008). A new approach to characterising and forecasting electricity price volatility. *International Journal of Forecasting*, 24, 728–743.

[42] CHEN, L-A. (1997). Multivariate regression splines. *Computational Statistics and Data Analysis*, 26, 71–82.

[43] CHEN, H., LI, F., WAN, Q. and WANG, Y. (2011). Short term load forecasting using regime-switching GARCH models. *Power and Energy Society General Meeting, EEE*, 1–6. DOI: 10.1109/PES.2011.6039457

[44] CHIODO, E. and LAURIA, D. (2012). Probabilistic description and prediction of electric peak power demand. *Electrical Systems for Aircraft, Rail-*

way and Ship Propulsion (ESARS), 1–7. DOI: 10.1109/ESARS.2012.6387418

[45] CHO, H., GOUDE, Y., BROSSAT, X. and YAO, Q. (2013). Modeling and Forecasting Daily Electricity Load Curves: A Hybrid Approach. *Journal of the American Statistical Association*, 108(501), 7–21.

[46] CHRISTOPHER, M. and LEE, H. (2004). Mitigating supply chain risk through improved confidence. *Int J Phys Distrib Logist Mngt*, 4, 388–396.

[47] Climate in Australia (Accessed on 10 June 2012) http://en.wikipedia.org/wiki/Climate_of_Australia

[48] ClimateTemp.info. (Accessed on 10 June 2012) http://www.climatetemp.info/south-africa/

[49] COLES, S.G. and POWELL, E.A. (1996). Bayesian methods in extreme value modelling: A review and new developments. *Int. Statist. Rev.*, 64, 119–136.

[50] COLES, S. (2001). *An introduction to statistical modeling of extreme values. Springer Series in Statistics*. London: Springer-Verlag London Ltd.

[51] CONSTABLE, S., HAMILTON, J. and PFAFF, T.J. (2012). A case study on regional impacts of climate change: peak loads on the power grid in Rochester, New York. *Journal of Environmental Studies and Sciences*. DOI: 10.1007/s13412-012-0097-5

[52] COOLEY, D.S. (2005). *Statistical Analysis of Extremes Motivated by Weather and Climate Studies: Applied and Theoretical Advances*. PhD theses, University of Colorado.

[53] COTTET, R. and SMITH, M. (2003). Bayesian modeling and forecasting of intraday electricity load. *Journal of the American Statistical Association*, 98, 839–849.

[54] COUSINS, T. (2009). Using time of use (TOU) tariffs in industrial, commercial and residential applications effectively. *TLC Engineering Solutions*. http://www.tlc.co.za/white_papers/pdf/using_time_of_use_tariffs_in_industrial_commercial_and_residential_applications_effectively.pdf. (Accessed on 21 August 2012)

[55] CRAVEN, P. and WAHBA, G. (1979). Smoothing noisy data with spline functions: Estimating the correct degree of smoothing by the method of generalized cross-validation. *Numberische Mathematik*, 31, 317–403.

[56] DANIELSSON, J., DE HAAN, L., PENG, L. and DE VRIES, C.G. (2001). Using a bootstrap method to choose the sample fraction in tail index estimation. *J. Mult. Anal.*, 76, 226–248.

[57] DAVISON, A.C. and SMITH, R.L. (1990). Models for exceedances over high thresholds (with discussion). *J. R. Statist. Soc. B*, 52, 237–254.

[58] DEB, K. (1996). *Optimisation for engineering design: algorithms and examples*, New Delhi: Prentice-Hall of India.

[59] DEBBA, P., KOEN R., HOLLOWAY, J.P., MAGADLA, T., RASUBA, M., KHULUSE, S. and ELPHINSTONE, C.D. (2010). Forecasts for electricity demand in South Africa (2010–2035) using the CSIR sectoral regression model. *CSIR* (http://www.energy.gov.za/IRP/irp%20files/CSIR_model_IRP%20forecasts%202010_final_v2.pdf)

[60] DE HAAN, L. and FERREIRA, A. (2006). *Extreme Value Theory: An Introduction*. New York: Springer Science and Business Media, LCC.

[61] DE WAAL, D.J., BEKKER, S.J. and PRETORIUS, J.H. (2011). Modelling inflows into the Gariep Dam : case study. *South African Statistical Journal*, 45(1), 135–147.

[62] DE ZEA BERMUDEZ, P. and KOTZ, S. (2010). Parameter estimation of the generalized Pareto distribution-Part II. *J. Statist. Plan Inf.*, 140, 1374–1388.

[63] DORDONNAT, V., KOOPMAN, S.J., OOMS, M., DESSERTAINE, A. and COLLET, J. (2008). An hourly periodic state space model for modelling French national electricity load. *International Journal of Forecasting*, 24, 566–587.

[64] DLAMINI, N.G. and CROMIERES, F. (2012). Implementing peak load reduction algorithms for household electrical appliances. *Energy Policy*, 44, 280–290.

[65] DREES, H. (1995). Refined Pickands estimators of the extremal index. *Ann. Statist.*, 23, 2059–2080.

[66] DREES, H. (1998). Optimal rates of convergence for estimates of the extreme value index. *Ann. Statist.*, 26, 434–448.

[67] DREES, H., DE HAAN, L. and RESNICK, S.I. (2000). How to make a Hill plot. *Ann. Statist.*, 28, 254–274.

[68] EGLESE, R.W. (1990). Simulated Annealing, A tool for Operational Research. *European Journal of Operational Research*, 46, 271 – 281.

[69] Energy Efficiency and Environment `http://www.energy.gov.za/EEE/Projects.html` (Accessed on 11 January 2013).

[70] ENGLE, R.F. (1982). Autoregressive Conditional Heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50(4), 987–1006.

[71] Eskom DSM, (2006). The SA Demand Side Management Programme–an operational guide for energy services companies.

[72] Eskom Integrated Demand Management. http://www.eskom.co.za/c/23/idm/ (Accessed on 14 January 2013).

[73] Eskom Critical Peak Day Pricing-Pilot Project http://www.eskom.co.za/c/article/975/critical-peak-day-pricing-pilot-project/ (Accesses on 24 October 2012).

[74] FAN, S. and HYNDMAN, R.J. (2012). Short-term load forecasting based on a semi-parametric additive model. *IEEE Transactions on Power Systems*, 27(1), 134–141.

[75] FERREIRA, A., DE HAAN, L. and PENG, L. (2003). On optimising the estimation of high quantiles of a probability distribution. *Statistics*, 37, 401–434.

[76] FISHER, R. and TIPPETT, L. (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.

[77] FRAGA ALVES, M.I. (2001). A location invariant Hill-type estimator. *Extremes*, 4(3), 199–217.

[78] FRANCO, G. and SANSTAD A.H. (2008). Climate change and electricity demand in California. *Climatic Change*, 87(Suppl 1): S139–S151.

[79] FRIEDMAN, J.H. (1991). Multivariate adaptive regression splines. *The Annals of Statistics*, 19(1), 1–141.

[80] GELMAN, A., CARLIN, J., STERN, H. and RUBIN, D. (2004). *Bayesian Data Analysis*. Second edition. Chapman and Hall, Boca Raton, FL.

[81] GEMAN, S. and GEMAN, D. (1984). Stochastic relaxation, Gibbs distribitions and Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(2), 721–741.

[82] GENCAY, R. and SELCUK, F. (2004). Extreme value theory and value-at-risk: relative performance in emerging markets. *International Journal of Forecasting*, 20, 287–303.

[83] GLOSTEN, L.R., JAGANNATHAN, R. and RUNKLE, D.E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. Econometrica 48(5), 1779-1801.

[84] GNEDENKO, R. (1943). Sur la distribution limite du terme maximum dne srie alatoire. *Ann. of Math.*, 44, 423–453.

[85] GOIA, A., MAY, C. and FUSAI, G. (2010). Functional clustering and linear regression for peak load forecasting. *International Journal of Forecasting*, 26, 700–711.

[86] GONZLEZ-ROMERA, E., JARAMILLO-MORN, M.A. and CARMONA-FERNNDEZ, D. (2008). Monthly electric energy demand forecasting with neural networks and Fourier series. *Energy Conversion and Management*, 49, 3135–3142.

[87] GOEGEBEUR, Y., BEIRLANT, J. and DE WET, T. (2008). Linking Pareto-tail kernel goodness-of-fit statistics with tail index at optimal threshold and second order estimation. *Revstat*, 6, 51–69.

[88] GRANGER, C.W.J. and JEON, Y. (2007). Long-term forecasting and evaluation. *International Journal of Forecasting*, 23, 539–551.

[89] GRANGER, C.W.J. (1986). Comment (on McNees, 1986). *Journal of Business and Economic Statistics*, 4, 16–17.

[90] HAHN, H., MEYER-NIEBERG, S. and PICKL, S. (2009). Electric load forecasting methods: Tools for decision making. *European Journal of Operational Research*, 199, 902–907.

[91] HAKSOZ, C. and SESHADRI, S. (2007). Supply chain operations in the presence of a spot market: a review with discussion. *Journal of Operational Research Society*, 58, 1412–1429.

[92] HALL, P.G. (1990). Using the bootstrap to estimate mean squared error and select smoothing parameter in nonparametric problems. *J. Mult. Anal.*, 32, 177–203.

[93] HARVEY, A. and KOOPMAN, S.J. (1993). Forecasting Hourly Electricity Demand using Time-Varying Splines. *JASA*, 88, 1228–1236.

[94] HEFFERNAN, J.E. and STEPHENSON, A.G. (2013). Ismev: R package version 1.39. `http://www.ral.ucar.edu/~ericg/softextreme.php`

[95] HEKKENBERG, M., BENDERS, R.M.J., MOLL, H.C. and SCHOOT, A.J.M. (2009). Indications for a changing electricity demand pattern: The temperature dependence of electricity demand in the Netherlands. *Energy Policy*, 37, 1542–1551.

[96] HENNINGSEN, A. and TOOMET, O. (2011). maxLik: A package for maximum likelihood estimation in R. *Computational Statistics*, 26(3), 443–458. DOI: `10.1007/s00180-010-0217-1`.

[97] HILL, B.M. (1975). A simple general approach to inference about the tail of a distribution. *Ann. Statist.*, 3, 1163–1174.

[98] HIPPERT, H.S., PEDREIRA, C.E. and SOUZA, R.C. (2001). Neural networks for short-term load forecasting: a review and evaluation. *IEEE Trans Power Syst*, 16(1), 44–55.

[99] HONG, T. (2010). *Short Term Load Forecasting*. PhD Thesis, North Carolina University, United States of America.

[100] HONG, T., WILSON, J. and XIE, J. (2013). Long Term Probabilistic Load Forecasting and Normalization With Hourly Information. *IEEE Transactions on Smart Grid*, Article in Press.

[101] HONGXIAO, W., CHUANWEN, J., NENGLING, T. and ZHIJIAN, H. (2004). The application of grey neural network model to city electricity demand forecasting. *The Journal of Grey System*, 4, 357–365.

[102] HO, K.Y. and TSUI, A.K.C. (2004). Analysis of real GDP growth rates of greater China: An asymmetric conditional volatility approach. *China Economic Review*, 5(4), 424–442.

[103] HOR, C-L, WATSON, S.J. and MAJITHIA, S. (2006). Daily Load Forecasting and Maximum Demand Estimation using ARIMA and GARCH. *9th International Conference on Probabilistic Methods Applied to Power Systems KTH, Stockholm, Sweden June 11-15*. http://labplan.ufsc.br/congressos/PMAPS/files/pdf/2.1/2.1_hor.pdf

[104] HOR, C., WATSON, S., INFIELD, D. and MAJITHIA, S. (2008). Assessing load forecast uncertainty using extreme value theory. *16th PSCC*, Glasgow, Scotland, July 14–18. http://www.pscc-central.org/uploads/tx_ethpublications/pscc2008_571.pdf (Accessed on 27 January 2012).

[105] HOSKING, J.R.M. and WALLIS, J.R. (1987). Parameter and quantile estimation for the Generalized Pareto Distribution. *Technometrics*, 29(3), 339–349.

[106] HUGO, J. (2012). *Bayesian tolerance intervals for variance component models*. PhD Thesis, University of the Free State, South Africa.

[107] HUANG, J., LI, Y. and LIU, Y. (2012). Summer daily peak load forecasting considering accumulation effect and abrupt change of temperature.

Power and Energy Society general meeting, IEEE, DOI: 10.1109/PESGM.2012.6345263

[108] HSIEH, D.A. (1989). The statistical properties of daily foreign exchange rates: 1974-1983. *Journal of International Economics*, 24, 129–145.

[109] HYNDMAN, R.J. and FAN, S. (2008). Density forecasting for long-term peak electricity demand. Working Paper 06/08, http://www.buseco.monash.edu.au/depts/ebs/pubs/wpapers/

[110] HYNDMAN, R.J. and FAN, S. (2009). Forecasting long-term peak half-hourly electricity demand for South Australia. Report for Electricity Supply Industry Planning Council (South Australia), http://www.buseco.monash.edu.au/units/forecasting

[111] HYNDMAN, R.J. and FAN, S. (2010). Density forecasting for long-term peak electricity demand. *IEEE Transactions on Power Systems*, 25(2), 1142–1153.

[112] HYNDMAN, R.J. and ATHANASOPOULOS, G. (2012). Forecasting: principles and practice, online textbook, accessed on 26 November 2012. http://otexts.com/fpp/

[113] HYNDMAN, R.J., ATHANASOPOULOS, G., RAZBASH, S., SCHMIDT, D., ZHOU, Z., KHAN, Y. and BERGMEIR, C. (2013). Forecast: R package version 4.05. http://robjhyndman.com/software/forecast/

[114] IBRAHIM, J. and CHEN, M. (2000). Power prior distributions for regression models. *Statistical Science*, 15, 46–60.

[115] IBRAHIM, J.G. and LAUD, P.W. (1991). On Bayesian Analysis of Generalized Linear Models Using Jeffreys's Prior. *Journal of the American Statistical Association*, 86(416), 981–986.

[116] International Energy Agency Demand Side Management. (Accessed on 7 January 2013). http://www.ieadsm.org/Files/Exco%20File%20Library/Participation/Final%20strategy%202008-2012.pdf

[117] INGLESI-LOTZ, R. (2011). The evolution of price elasticity of electricity demand in South Africa: A Kalman filter application. *Energy Policy*, 39, 3690-3696.

[118] ISLAM, S.M., AL-ALAWI, S.M. and ELLITHY, K.A. (1995). Forecasting monthly electric load and energy for a fast growing utility using an artificial neural network. *Electric Power Systems Research*, 34, 1–9.

[119] ISMAIL, Z., YAHYA, A. and MAHPOL, K.A. (2009). Forecasting peak load electricity demand using statistics and rule based approach. *American Journal of Applied Sciences*, 6(8), 1618–1625.

[120] JARQUE, C.M. and BERA, A.K. (1987). A test for normality of observations and regression residuals. *International Statistical review*, 55(2), 163–172.

[121] KASS, R.E. and WASSERMAN, L. (1996). Formal Rules of Selecting Prior Distributions: A Review and Annotated Bibliography. *Journal of the American Statistical Association*, 91, 343–370.

[122] KATZ, R., PARLANGE, M. and NAVEAU, P. (2002). Statistics of extremes in hydrology. *Advances in Water Resources*, 25, 1287–1304.

[123] KESAVABHOTLA, C.B., HARAGOPAL, V.V. and BABU, A.V. (2012). Application of Analytics using R Programming for Forecasting Day-ahead Electricity Demand. *IJCSI International Journal of Computer Science Issues*, 9(6), 341–349.

[124] KHAN, G.M., KHAN, S. and ULLAH, F. (2011). Short-term daily peak load forecasting using fast learning neural network. *Intelligent Systems*

Design and Applications (ISDA), 11th International Conference. DOI: 10.1109/ISDA.2011.6121762

[125] KIARTZIS, S., KEHAGIAS, A., BAKIRTZIS, A. and PETRIDIS, V. (1997). Short term load forecasting using a Bayesian combination method. *Electrical Power & Energy Systems*, 19(3), 171–177.

[126] KIRKPATRICK, S., GELATT, C.D. JR. and VECCHI, M.P. (1983). Optimisation by simulated annealing, *Science*, 220, 671-680.

[127] KOCH, K-R. (2007). Introduction to Bayesian Statistics, Second edition: Springer, Berlin.

[128] LAURENT, S. (2010). G@RCH-site-Book63. http://www.core.ucl.ac.be/~laurent/G@RCH/site/Book63.html

[129] LAWLESS, J.F. (1980). Inference in the generalized gamma and log gamma distributions. Technometrics 22(3):409-419.

[130] LEADBETTER, M.R. (1974). On extreme values in stationary sequences. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 28, 289–303.

[131] LEADBETTER, M.R. (1983). Extremes and local dependences in stationary sequences. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 65, 291–306.

[132] LEADBETTER, M.R., LINDGREN, G. and ROOTźEN, H. (1983). *Extremes and related properties of random sequences and processes*. Springer Verlag, New York.

[133] LETZLER, R. (2010). Using incentive preserving rebates to increase acceptance of critical peak electricity pricing. *CSEM University of California Energy Institute*. [Online], [Cited: 21 August 2012], Available: http://www.ucei.berkeley.edu/PDF/csemwp162r.pdf.

[134] LJUNG, G.M. and BOX, G.E.P. (1978). On measure of lack of fit in time series models. *Biometrika*, 65, 297–303.

[135] MACDONALD, A., SCARROTT, C.J., LEE, D., DARLOW, B., REALE, M. and RUSSELL, G. (2011). A flexible extreme value mixture model. *Computational Statistics and Data Analysis*, 55, 2137–2157.

[136] MACDONALD, A. (2012). Extreme value mixture modelling with medical and industrial applications, PhD thesis, University of Canterbury, New Zealand.

[137] MACKAY, E.B.L., CHALLENOR, P.G. and BAHAJ, A.S. (2011). A comparison of estimators for the generalised Pareto distribution. *Ocean Engineering*, 38, 1338–1346.

[138] MAKRIDAKIS, S., WHEELWRIGHT, S. and HYNDMAN, R. (1998). *Forecasting methods and applications*. 3rd ed. New York: Wiley.

[139] MCNEIL, A.J. and FREY, R. (2000). Estimation of Tail-Related Risk Measures for Heteroscedastic Financial Time Series: An Extreme Value Approach. *Journal of Empirical Finance*, 7, 271–300.

[140] MCSHARRY, P.E., BOUWMAN, S. and BLOEMHOF, G. (2005). Probabilistic forecasts of the magnitude and timing of peak electricity demand. *IEEE Transactions on Power Systems*, 20(2), 1166–1172.

[141] MARIN, J.M, DIEZ, R.M. AND INSUA, D.R. (2003). BAYESIAN METHODS IN PLANT CONSERVATION BIOLOGY. *Biological Conservation*, 113, 379–387.

[142] MESTEKEMPER, T., KAUERMANN, G. and SMITH, M.S. (2013). A comparison of periodic autoregressive and dynamic factor models in intraday energy demand forecasting. *International Journal of Forecasting*, 29, 1–12.

[143] MICALI, V. (2007). *On the use of extreme value theory in energy markets*. PhD thesis. http://etd.uovs.ac.za/ETD-db/theses/available/etd-11162007-075128/unrestricted/MicaliV.pdf

[144] MILBORROW, S. (2013). Earth: R package version 3.2-6. http://www.milbo.users.sonic.net/earth

[145] MIRASGEDIS, S., SARAFIDIS, Y., GEORGOPOULOU, E., LALAS, D.P., MSCHOVITIS, M., KARAGIANNIS, F. and PAPAKONSTANTINOU, D. (2006). Models for mid-term electricity demand forecasting incorporating weather influences. *Energy*, 31, 208–227.

[146] MORAL-CARCEDOA, J. and VICéNS-OTERO, J. (2005). Modelling the non-linear response of Spanish electricity demand to temperature variations. *Energy Economics*, 27, 477–494.

[147] MUNOZ, A., SANCHEZ-UBEDA, E.F., CRUZ, A. and MARIN, J. (2010). Short-term forecasting in power systems: a guided tour. *Energy Systems*, 2, 129–160.

[148] MUNOZ, J. and FELICISIMO, A.M. (2004). Comparison of statistical methods commonly used in predictive modeling. *Journal of Vegetation Science*, 15, 285–292.

[149] National Energy ACT. 2008. *Regulations on the allowance for energy efficiency savings*. http://www.info.gov.za/view/DownloadFileAction?id=150701 (Accessed on 29 May 2012).

[150] National Response to South Africa's Electricity Shortage http://www.info.gov.za/otherdocs/2008/nationalresponse_sa_electricity1.pdf (Accessed on 25 February 2010)

[151] NELSON, D.B. (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica*, 59(2), 347–370.

[152] NEWSHAM, G.R. and BOWKER, B.G. (2010). The effect of utility time-varying pricing and load control strategies on residential summer peak electricity use: A review. *Energy Policy*, 38, 3289–3296.

[153] NEWSHAM, G.R., BIRT, B.J. and ROWLANDS, I.H. (2011). A comparison of four methods to evaluate the effect of a utility residential air conditioner load control program on peak electricity use. *Energy Policy*, 39 (10), 6376–6389.

[154] NIUA, D., SHI, H. and WU, D.D. (2012). Short-term load forecasting using bayesian neural networks learned by Hybrid Monte Carlo algorithm. *Applied Soft Computing*, 12, 1822–1827.

[155] OHTSUKA, Y., OGA, T. and KAKAMU, K. (2010). Forecasting electricity demand in Japan: A Bayesian spatial autoregressive ARMA approach. *Computational Statistics and Data Analysis*, 54, 2721–2735.

[156] Pardo, A., Meneu, V. and Valor, E. (2002). Temperature and seasonality influences on Spanish electricity load. *Energy Econ*, 24, 55–70.

[157] PEREIRA, T.T. (1994). Second order behaviour of domains of attraction and the bias of generalized Pickands' estimator. In "Extreme Value Theory and Applications III", *Proceedings of Gaithersburg Conference*, 1993 (J. Galambos, L. Lechner and E. Simiu, Eds.), 866, 165–177.

[158] PICKANDS, J. (1975). Statistical Inference using extreme order statistics. *The Annals of Statistics*, 3, 119–131.

[159] PILLI-SIHVOLA, K., AATOLA, P., OLLIKAINEN, M. and TUOMENVIRTA, H. (2010). Climate change and electricity consumption-Witnessing increasing or decreasing use and costs, Energy Policy, 38(5), 2409–2419.

[160] PSILOGLOU, B.E., GIANNAKOPOULOS, C., MAJITHIA, S. and PETRAKIS, M. (2009). Factors affecting electricity demand in Athens, Greece and London, UK: A comparative assessment. *Energy*, 34, 1855–1863.

[161] RAMANATHAN, R., ENGLE, R., GRANGER, C.W.J, VAHID-ARAGHI, F. and BRACE, C. (1997). Short-run forecasts of electricity loads and peaks. *International Journal of Forecasting*, 13, 161–174.

[162] REISS, R.D. and THOMAS, M. (2007). *Statistical Analysis of Extreme Values*, Third Edition, Berlin: Birkhauser.

[163] ROBERT, C.P. (2001). *The Bayesian Choice*, Second Edition, New York: Springer-Verlag.

[164] SAILOR, D.J. and MUNOZ, J.R. (1997). Sensitivity of electricity and natural gas consumption to climate in the U.S.A: methodology and results for eight states. *Energy*, 22 (10), 987–998.

[165] Salford Predictive Modeller version 7.0 `http://www.salford-systems.com/`

[166] SANEDI `http://www.sanedi.org.za/` (Accessed on 29 May 2012).

[167] SAS/STAT9.2 User's Guide Introduction to Bayesian Analysis Procedures (Book Excerpt). `http://support.sas.com/documentation/cdl/en/statugbayesian/61755/PDF/default/statugbayesian.pdf` (Accessed on 29 February 2011).

[168] SHAW, R., AHREE, M., JACKSON, T. and KAY, M. (2009). The value of reducing distribution by domestic load-shifting: a network perspective. *Energy Policy*, 37, 3159–3167.

[169] SOARES, L.J. and MEDEIROS, M.C. (2008). Modelling and Forecasting short-term electricity load: A comparison of methods with an application to Brazilian data. *International Journal of Forecasting*, 24, 630–644.

[170] SUGANTHI, L. and SAMUEL, A.A. (2012). Energy models for demand forecasting: A review. *Renewable and Sustainable Energy Reviews*, 16, 1223–1240.

[171] SMITH, R. (1987). Estimating tails of probability distributions. *The Annals of Statistics*, 15(3), 1174–1207.

[172] SouthAfrica.info. (Accessed on 10 June 2012) http://www.southafrica.info/travel/advice/climate.htm

[173] STACY, E.W. (1962). A generalization of the gamma distribution. Annals of Mathematical Statistics 33, 1187–1192.

[174] STEPHENSON, A.G. and RIBATET, M.A. (2006). A User's Guide to the evdbayes Package (Version 1.1). http://cran.r-project.org/

[175] STEPHENSON, A.G. and RIBATET, M.A. (2012). Bayesian Analysis in Extreme Value Theory. R package version 1.1-0. http://cran.r-project.org/web/packages/evdbayes/

[176] STRAPP, J., KING, C. and TALBOTT, S. (2007). Ontario Energy Board Smart Price Pilot Final Report. Prepared by IBM Global Business Services and eMeter Strategic Consulting for the Ontario Energy Board. http://www.oeb.gov.on.ca/AF9CF86B-333A-4635-B899-C02E2D5B6508/FinalDownload/DownloadId-B9DDC81F75CF02F29E22108EA3370D59/AF9CF86B-333A-4635-B899-C02E2D5B6508/documents/cases/EB-2004-0205/smartpricepilot/OSPP%20Final%20Report%20-%20Final070726.pdf.

[177] Strategic Plan for the IEA Demand-Side Management Programme 2008–2012. http://www.ieadsm.org/Files/Exco%20File%20Library/Participation/Final%20strategy%202008-2012.pdf

[178] STRENGERS, Y. (2012). Peak electricity demand and social practice theories: Reframing the role of change agents in the energy sector. *Energy Policy*, 44, 226–234.

[179] TASSEL, A. (2003). South Africa-is power supply crisis developing?: Electricity supply. *African Energy Journal*, 5(5).

[180] TAYLOR, J.W. and BUIZZA, R. (2003). Using Weather Ensemble Predictions in Electricity Demand Forecasting. *International Journal of Forecasting*, 19, 57–70.

[181] TAYLOR, J.W. (2006). Density forecasting for the efficient balancing of the generation and consumption of electricity. *International Journal of Forecasting*, 22, 707–724.

[182] TAYLOR, J.W. (2008). An evaluation of methods for very short-term load forecasting using minute-by-minute British data. *International Journal of Forecasting*, 24, 645–658.

[183] TAYLOR, J.W. and MCSHARRY, P.E. (2008). Short-term load forecasting methods: an evaluation based on European data. *IEEE Trans Power Syst*, 22, 2213–2216.

[184] TAYLOR, J.W. (2010). Triple seasonal methods for short-term electricity demand forecasting. *European Journal of Operational Research*, 204, 139–152.

[185] THOMPSON, P., CAI, Y., REEVE, D. and STANDER, J. (2009). Automated threshold selection methods for extreme wave analysis. *Coastal Engineering*, 56, 1013–1021.

[186] TRUONG, N-V, WANG, L. and WONG, P.K.C. (2008). Modelling and short-term forecasting of daily peak power demand in Victoria using two-

dimensional wavelet based SDP models. *Electrical Power and Energy Systems*, 30, 511–518.

[187] TSEKOURAS, G.J., DIALYNAS, E.N., HATZIARGYRIOU, N.D. and KAVATZA, S. (2007). A non-linear multivariable regression model for midterm energy forecasting of power systems. *Electric Power Systems Research*, 77, 1560–1568.

[188] VERSTER, A. and DE WAAL, D.J. (2011). A method for choosing an optimum threshold if the underlying distribution is generalized Burr-Gamma. *South African Statistical Journal*, 45, 273–292.

[189] VON MISES, R. (1936). La distribution de la plus grande de n valeurs, *reprinted in Selected Papers II, Amer. Math. Soc., Providence, Rhode Island, 1954*, 271–294.

[190] WASSERMAN, L. (2004). *All of Statistics: A Concise Course in Statistical Inference*, New York: Springer-Verlag.

[191] WERON, R. (2006). *Modeling and forecasting electricity loads and prices: a statistical approach*. Wiley, Chichester.

[192] WEI, W.W.S. (2006). *Time Series Analysis: Univariate and Multivariate Methods*, Second Edition. Pearson, Addison Wesley, New York.

[193] WINTERS, P.R. (1960). Forecasting sales by exponentially weighted moving averages. *Manag Sci*, 6(3), 324–342.

[194] XU, K., ZHOU, J., ZHANG, Y. and GU, R. (2010). The Long and Short Term Volatility Modeling for Load Series based on ARCH Type Models. *2nd International Conference on Information and Multimedia Technology (ICIMT 2010) IPCSIT* 42, IACSIT Press, Singapore DOI: 10.7763/IPCSIT.2012.V42.29

[195] ZAKOIAN J.M. 1994. Threshold heteroscedastic models. Journal of Economic Dynamics and Control 18, 931–955.

[196] ZELLNER, A. (1977). *Bayesian analysis in econometrics and statistics*: Edward Elgar, Lyme US.

[197] ZELLNER, A. (1977). Maximal data information prior distributions, *In: A. Aykac and C. Brumat, eds., New developments in the applications of Bayesian methods (North-Holland, Amsterdam)*, 21, 211–232.

[198] ZIRAMBA, E. (2008). The demand for residential electricity in South Africa. *Energy Policy*, 36, 3460–3466.

Appendix

SOME SELECTED R CODES

R code for Bayes GARCH (1,1) model with Student-t innovations

The code is run on bayesGARCH package in R

import data

win.graph()

RUN THE SAMPLER (1 chain)

addPriorConditions $<-$ function(psi)psi[2] + psi[3] $<$ 1

set.seed(1234)

MCMC $<-$ bayesGARCH(Residual, control = list(n.chain = 1, l.chain = 8000,

addPriorConditions = addPriorConditions))

MCMC ANALYSIS (using coda)

plot(MCMC)

summary(MCMC, quantiles = c(0.025, 0.25, 0.5, 0.75, 0.975))

autocorr.diag(MCMC)

1-rejectionRate(MCMC)

FORM THE POSTERIOR SAMPLE

smpl $<-$ formSmpl(MCMC, l.bi = 500)

POSTERIOR STATISTICS

summary(smpl,quantiles = c(0.025, 0.25, 0.5, 0.95, 0.99,0.995),)

```
smpl < − as.matrix(smpl)
summary(smpl)
pairs(smpl,col="blue")
```

R code for GEVD using the Bayesian approach

The code is run on evdbayes package in R

```
attach(dpdRESDPD)
head(dpdRESDPD)
win.graph()
y=ts(RESDPD1, start=2000,freq=365)
plot(y, xlab = "Year", ylab = "Daily Peak Electricity Demand (MW)", col="blue")
mat < − diag(c(10000, 10000, 100))
pn < − prior.norm(mean = c(0,0,0), cov = mat)
n < − 2000 ; t0 < − c(-1200,2100,-0.2) ; s < − c(40,20,0.005)
ptpmc < − posterior(n, t0, prior = pn, lh = "gev", data = RESDPD1, psd = s)
attributes(ptpmc)$ar
# MCMC works with coda
ptp.mcmc < − mcmc(ptpmc, start = 0, end = 2000)
plot(ptp.mcmc, den = FALSE, sm = FALSE, main="  ")
# best initial values of to
maxpst < − mposterior(t0, prior = pn, lh = "gev", method = "SANN",data =
RESDPD1)
round(maxpst$par, 2)
# best initial values of s
t0 < − c(-1196.17, 2101.25,-0.25); psd < − rep(0.01, 3)
psd < − ar.choice(init = t0, prior = pn, lh = "gev", data = RESDPD1, psd =
psd, tol = rep(0.02, 3))$psd
round(psd, 2)
```

```
# using best t0 and s
mat < − diag(c(10000, 10000, 100))
pn < − prior.norm(mean = c(0,0,0), cov = mat)
n < − 2000 ; t0 < − c(-1196.17, 2101.25,-0.25) ; s < − c(85.39,0.02, 0.01)
ptpmc < − posterior(n, t0, prior = pn, lh = “gev”, data = RESDPD1, psd = s)
attributes(ptpmc)$ar
# MCMC works with coda
ptp.mcmc < − mcmc(ptpmc, start = 0, end = 2000)
plot(ptp.mcmc, den = FALSE, sm = FALSE, main=“ ”)
# TRACE OF THE PARAMETERS
t0 < − c(-1196.17, 2101.25,-0.25) ; s < − c(85.39,0.02, 0.01)
ptpmc < − posterior(n, t0, prior = pn, lh = “gev”, data = RESDPD1, psd = s)
ptp.mcmc < − mcmc(ptpmc, start = 0, end = 2000)
plot(ptp.mcmc, den = FALSE, sm = FALSE,col=“blue”, main=“ ”)
# diagnostic checks Geweke
ptp.mcmc < − window(ptp.mcmc, start = 100)
geweke.diag(ptp.mcmc)
geweke.diag(ptp.mcmc, 0.2, 0.4)
# Marginal posterior densities and summary statistics
bwf < − function(x) sd(x)/2
plot(ptp.mcmc, trace = T, bwf = bwf,col=“blue”, main=“ ”)
summary(ptp.mcmc)
#Return level plot of posterior distribution with 95% Bayesian credible
# intervals
rl.pst(ptp.mcmc, npy, lh = “gev”, ci = 0.95, lty = c(2,1), col = c(2,1),
xlab = “Return period (Years)”, ylab = “Return level (Megawatts)”)
mc.quant(ptp.mcmc, p=c(0.9796,0.96011,0.9332,0.895), lh = “gev”)
```

R codes for fitting extreme value distributions (GEVD and GPD)

Both codes are run on the R package ismev

R code for fitting GEVD

ppfit $<-$ gev.fit(resid.)

gev.diag(ppfit)

R code for fitting GPD

rnfit $<-$ gpd.fit(RESDPD, 2981)

gpd.diag(rnfit)

R codes for fitting a Pareto quantile plot

The code is run on the R package laeken

attach(paretoqp)

win.graph()

library(laeken)

paretoQPlot(lnYt, w = NULL, xlab = "-log[$1 - i/(n + 1)$]",

ylab = "Log of the observations", main=" ",

interactive = TRUE, x0 = NULL, theta = NULL,

pch = par("pch"), cex = par("cex"), col ="blue",

bg = "transparent", ylim=c(4,9))

The R code for MLE of GSPD

The code is run on maxLik: A package for maximum likelihood estimation in R

Import data

win.graph()

library(“maxLik”)

fn $<-$ function(η){

sum((1/η)*log(1+(τ*η)))-sum((1+η)/(η)*log(1+η*data))

}

mle $<-$ maxLik(logLik = fn, start= c(η=0.08), method=“SANN”)

summary(mle)

Summary

Peak electricity demand is an energy policy concern for all countries throughout the world, causing blackouts and increasing electricity tariffs for consumers. This calls for load curtailment strategies to either redistribute or reduce electricity demand during peak periods. This thesis attempts to address this problem by providing relevant information through a frequentist and Bayesian modelling framework for daily peak electricity demand using South African data. The thesis is divided into two parts. The first part deals with modelling of short term daily peak electricity demand. This is done through the investigation of important drivers of electricity demand using *(i)* piecewise linear regression models, *(ii)* a multivariate adaptive regression splines (MARS) modelling approach, *(iii)* a regression with seasonal autoregressive integrated moving average (Reg-SARIMA) model *(iv)* a Reg-SARIMA model with generalized autoregressive conditional heteroskedastic errors (Reg-SARIMA-GARCH). The second part of the thesis explores the use of extreme value theory in modelling winter peaks, extreme daily positive changes in hourly peak electricity demand and same day of the week increases in peak electricity demand. This is done through fitting the generalized Pareto, generalized single Pareto and the generalized extreme value distributions.

One of the major contributions of this thesis is quantification of the amount of electricity which should be shifted to off peak hours. This is achieved through accurate assessment of the level and frequency of future extreme load forecasts. This modelling approach provides a policy framework for load curtailment and

determination of the number of critical peak days for power utility companies. This has not been done for electricity demand in the context of South Africa to the best of our knowledge. The thesis further extends the autoregressive moving average-exponential generalized autoregressive conditional heteroskedasticity model to an autoregressive moving average exponential generalized autoregressive conditional heteroskedasticity-generalized single Pareto distribution. The benefit of this hybrid model is in risk modelling of under and over demand predictions of peak electricity demand.

Some of the key findings of this thesis are *(i)* peak electricity demand is influenced by the tails of probability distributions as well as by means or averages, *(ii)* electricity demand in South Africa rises significantly for average temperature values below 18^0C and rises slightly for average temperature values above 22^0C and *(iii)* modelling under and over demand electricity forecasts provides a basis for risk assessment and quantification of such risk associated with forecasting uncertainty including demand variability.

Key words and phrases: Bayesian inference, extreme value theory, frequentist, load management, maximal data information prior, peak electricity demand, risk management, temperature.

Opsomming

Spitselektrisiteitsaanvraag is vir alle lande regdeur die wêreld 'n energiebeleidskwessie en veroorsaak verdonkerings en toenemende elektrisiteitstariewe vir allle verbruikers. Dit verg ladinginkortingstrategieë om óf elektrisiteitsaanvraag tydens spitstye te hervers prei óf om dit te verminder. Diéverhandeling poog om hierdie probleem aan te spreek deur toepaslike inligting deur 'n frekwentistiese en Bayes-modelraamwerk vir daaglikse spitselektrisiteitsaanvraag te verskaf met behulp van Suid-Afrikaanse data. Die verhandeling word in twee dele verdeel. Die eerste gedeelte handel oor modellering van korttermyn daaglikse spitselektrisiteitsaanvraag. Dit word bereik deur die belangrike dryfvere van elektrisiteitsaanvraag te ondersoek met behulp van (i) stuksgewyse lineêre regressiemodelle, (ii) 'n meerveranderlike aanpassende regressielatmodelbenadering (MARS in Engels), (iii) 'n regressie met seisoenale outoregressiewe geïntegreerde bewegingsgemiddeld-model (Reg-SARIMA in Engels) en (iv) 'n Reg-SARIMA-model met veralgemeende outoregressiewe voorwaardelike heteroskedastiese foute (Reg-SARIMA-GARCH in Engels). Die tweede gedeelte van die verhandeling verken die gebruik van ekstreemwaardeteorie in die modellering van winterspitstye, ekstreem-daaglikse positiewe veranderinge in uurlikse spitselektrisiteitsaanvraag en selfde dag van die week verhogings in spitselektrisiteitsaanvraag. Dit word bereik deur die veralgemeende Pareto, veralgemeende enkel-Pareto en die veralgemeende ekstreemwaardeverdeling, in te pas.

Een van die hoofbydraes van die verhandeling is die kwantifisering van die

aantal elektrisiteit wat na niespitstydperke toe geskuif moet word. Dit word bereik deur akkurate assessering van die vlak en herhaling van toekomsitge ladingvoorspellings. Di$\acute{e}$ modelbenadering verskaf 'n beleidsraamwerk vir ladinginkorting en die bepaling van die aantal kritieke spitsdae vir krag-utiliteitsmaatskappye. Dit is, na ons beste wete, nie vir elektrisiteitsaanvraag in die Suid-Afrikaanse konteks gedoen nie. Die verhandeling brei verder die outoregressiewe bewegende gemiddelde eksponensiaal veralgemeende outoregressiewe voorwaardelike heteroskedastisiteitsmodel uit na 'n outoregressiewe bewegende gemiddelde eksponensiaal veralgemeende outoregressiewe voorwaardelike heteroskedastisiteit-veralgemeende enkel-Paretoverdeling. Die voordeel van hierdie hibriede model is in risikomodellering van onder-en ooraanvraagvoorspellings van spitselektrisiteitsaanvraag.

Sommige van die hoofbevindinge van die verhandeling is dat (i) spitselektrisiteitsaanvraag beïnvloed word deur die waarskynlikheidsverdelings asook deur die gemiddelde, (ii) elektrisiteitsaanvraag in Suid-Afrika vir gemiddelide temperatuurwaardes onder 18^0C aansienlik vermeerder en vir gemiddelde temperatuurwaardes bo 22^0C effens vermeerder en (iii) modellering van onder-en ooraavraag-elektrisiteitsvoorspellings 'n basis vir risiko-assessering en kwantifisering van sulke risikos wat verband hou met voorspellingsonsekerheid, insluitend aanvraagveranderlikheid, verskaf.

Sleutelwoorde en-frases: Bayes-inferensie, ekstreemwaardeteorie, frekwentisties, beladingsbestuur, maksimale data-inligting inligting a priori-verdeling, spitselektrisiteitsaanvraag, risikobestuur, temperatuur.