

Modelling influence of temperature on daily peak electricity demand in South Africa

Delson Chikobvu

Department of Mathematical Statistics and Actuarial Science, University of the Free State, South Africa

Caston Sigauke

School of Statistics and Actuarial Science, University of the Witwatersrand, South Africa

Abstract

The paper discusses the modelling of the influence of temperature on average daily electricity demand in South Africa using a piecewise linear regression model and the generalized extreme value theory approach for the period - 2000 to 2010. Empirical results show that electricity demand in South Africa is highly sensitive to cold temperatures. Extreme low average daily temperatures of the order of 8.20C are very rare in South Africa. They only occur about 8 times in a year and result in huge increases in electricity demand.

Keywords: Extreme value theory, piecewise linear regression, regression splines, temperature

1. Introduction

Drivers of electricity demand are generally divided into economic factors, calendar effects, weather variables and lagged demand variables. Inclusion of these factors in electricity demand models improves the predictive power of the models and also enables system operators and load forecasters to have a better understanding of the factors that have a greater impact on electricity demand. Weather variables such as temperature, solar radiation, humidity, wind speed and rainfall are often used as explanatory variables in regression based load forecasting models. Most authors, however, use temperature as the main driver (Munoz *et al.*, 2010).

The influence of temperature on daily electricity load forecasting has been studied extensively in the energy sector using classical time series, regression based methods including artificial neural networks (Miragedis *et al.*, 2006; Hekkenberg *et al.*, 2009; Psiloglou *et al.*, 2009; Munoz *et al.*, 2010; Pilli-

Sihvola *et al.*, 2010; among others). The paper discusses the modelling of the effect of average daily temperature on daily electricity demand in South Africa using a piecewise linear regression modelling framework and the generalized extreme value theory approach. A generalized extreme value distribution (GEVD) is fitted to the temperature data below the reference temperature. Extreme value theory (EVT) is a powerful and fairly robust framework for modelling the tail behaviour of a distribution (Gencay and Selcuk, 2004). Extreme value theory has been applied in various fields such as flood frequency analysis, environmental sciences, modelling extreme temperatures, finance and insurance including material and life sciences. The family of extreme value distributions is called the generalized extreme value distribution. GEVD consists of the Gumbel, Frechet and Weibull class distributions which are also known as the type I, II and III extreme value distributions respectively.

The rest of the paper is organized as follows. The models are discussed in Section 2. In Section 3 we briefly describe the data used. Empirical results are presented in Section 4 while Section 5 concludes.

2. The models

Our modelling approach is in two stages. A piecewise linear regression model is used to explore the effect of temperature on daily electricity demand. In stage two, we fit a generalized extreme value distribution to the temperature values below the reference temperature. The fitted distribution is then used to estimate extreme low temperatures and calculating the corresponding marginal increases of electricity demand.

2.1 Piecewise linear regression model

The piecewise linear regression model used for modelling the influence of temperature on electricity demand is given in equation (1).

$$ADED = \alpha_0 + \alpha_1 \max(0, t_h - ADT) + \alpha_2 \max(0, ADT - t_c) + \varepsilon_t \quad (1)$$

where ADED (Average Daily Electricity Demand), ADT (Average Daily Temperature), t_h and t_c 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 α_0, α_1 and α_2 and ε_t is the error term with $\varepsilon_t \sim N(0, \sigma_t^2)$. A Multivariate Adaptive Regression Splines (MARS) algorithm developed by Friedman (1991) is used to estimate the two reference temperatures t_c and t_h which are 18°C and 20°C respectively and discussed in Chikobvu and Sigauke (2012).

2.2 Generalized extreme value distribution

The GEVD consists of the Gumbel, Frechet and Weibull class of distributions. The unified GEVD for modelling maxima is given by

$$G_\xi(x) = \exp \left\{ - \left[1 + \xi \left(\frac{x-\mu}{\sigma} \right) \right]^{-\frac{1}{\xi}} \right\}, \text{ if } 1 + \xi \left(\frac{x-\mu}{\sigma} \right) > 0 \text{ and } \xi \neq 0 \quad (2)$$

where μ and σ are the location and scale parameters respectively. The shape parameter ξ also known as the extreme value index (EVI) determines the rate of tail decay. If $\xi > 0$, $G_\xi(x)$ belongs to the heavy-tailed Frechet class of distributions (Beirlant *et al.*, 2004). For $\xi < 0$ we have the short-tailed Weibull class and is bounded above by $\mu - \frac{\sigma}{\xi}$ (Beirlant *et al.*, 2004). If $\xi = 0$, $G_\xi(x)$ belongs to the light-tailed Gumbel class of distributions (Beirlant *et al.*, 2004).

In order to model minima we use the duality between the distributions for maxima and minima. If $M_n = \min\{x_1, x_2, \dots, x_n\}$ where $x_i, i = 1, 2, \dots, n$ represents temperature, then $\tilde{M}_n = \max\{-x_1, -x_2, \dots, -x_n\}$. Extreme maxima theory and methods are then used to model extreme minima. The Maximum Likelihood (ML) method is used to estimate the parameters ξ, μ and σ . The survival distribution of the GEVD given in equation (2) is given by:

$$P(X > x) = 1 - G_\xi(x) = 1 - \exp \left\{ - \left[1 + \xi \left(\frac{x-\mu}{\sigma} \right) \right]^{-\frac{1}{\xi}} \right\}, \text{ if } 1 + \xi \left(\frac{x-\mu}{\sigma} \right) > 0 \text{ and } \xi \neq 0 \quad (3)$$

If we let $p = P(X > x)$ and rearranging equation (3) to make x the subject we get the quantile function:

$$x_p = \mu + \frac{\sigma}{\xi} \left[\{-\ln(1-p)\}^{-\xi} - 1 \right], \xi \neq 0 \quad (4)$$

Now as $p \rightarrow 0$ and $\xi < 0$ we get $x_p = \mu - \frac{\sigma}{\xi}$. The quantile function for the unified GEVD given in equation (4) is then used to estimate high quantiles and predicting the probability of exceedance levels.

3. Data

National daily electricity data for the industrial, commercial and domestic sectors of South Africa is used in this study. The data is from Eskom, South Africa's power utility company. Figure 1 shows that ADED data exhibits strong seasonality and has a positive upward trend.

Aggregated ADT from 32 meteorological stations of South Africa representing all provinces (regions) of the whole country is used in the analysis.¹

Figure 2 shows that ADT has strong seasonality and is stationary. The minimum ADT and maximum ADT over the sampling period (2000-2010) are 7.5°C and 26.1°C respectively.

4. Empirical results and discussion

4.1 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.

$$E(ADED) = 23932 + 263 \max(0, 22 - ADT) + 138 \max(0, ADT - 18) \quad (5)$$

The graphical plot of ADED against ADT is shown in Figure 3. The three demand-temperature equations are given in equations 6-8. If average daily temperature is less than or equal to 18°C equation (5) reduces to

$$E(ADED) = 23932 + 263 \max(0, 22 - ADT) \quad (6)$$

That is, if the temperature decreases by 1°C (e.g. from 18°C to 17°C) electricity demand will increase marginally by 263 MW. A fall in ADT of 1°C (say, from 16°C to 15°C) would result in an increase of about 1.03% in electricity consumption.

If average daily temperature is greater than or equal to 22°C equation (5) reduces to

$$E(ADED) = 23932 + 138 \max(0, ADT - 18) \quad (7)$$

If temperature increases by 1°C (e.g. from 22°C to 23°C) electricity demand will increase marginally by 138 MW. For a rise in average daily temperature of 1°C (say, from 25°C to 26°C) would result in an increase of about 0.55% in electricity consumption.

For the average daily temperature between 18°C and 22°C we use the full model given in equation (5), i.e.

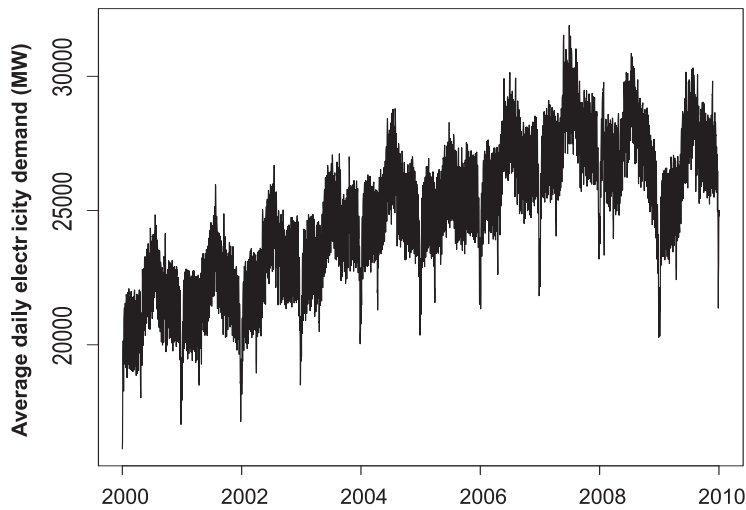


Figure 1: Average daily electricity demand (megawatts) from 2000-2010

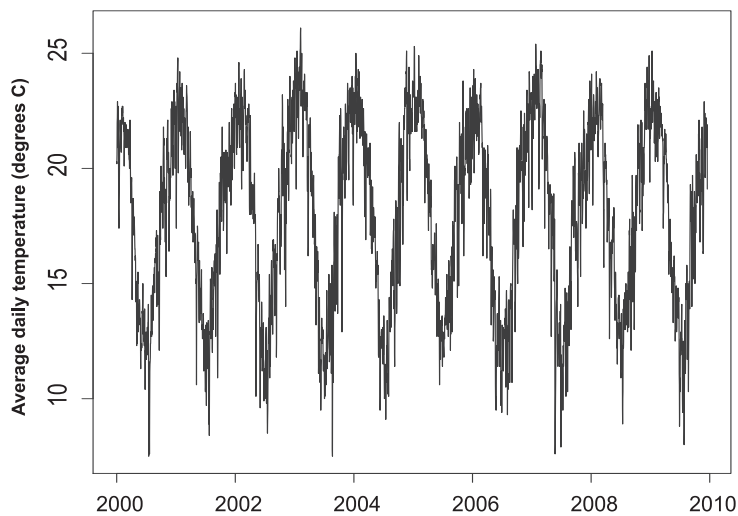


Figure 2: Time series plot of average temperature (degrees C)

$$E(ADED = 23932 + 263 \max(0, 22 - ADT) + 138 \max(0, ADT - 18) \quad (8)$$

$$HDD_t = \max(T_r - ADT_t) \quad (9)$$

$$CDD_t = \max(ADT - T_r, 0) \quad (10)$$

If temperature decreases by 1°C (e.g. from 22°C to 21°C) electricity demand will increase marginally by 125 MW. A decrease of ADT from 20°C to 19°C would result in an increase of about 0.51% in electricity consumption.

This analysis shows that electricity demand in South Africa is highly sensitive to cold temperature (see Figure 3). There is a non-linear relationship between temperature and electricity demand as shown in Figure 3. This non-linear relationship is modelled in literature using heating degree days (HDD) and cooling degree days (CDD). Modelling of this relationship between temperature and electricity demand is discussed in literature (Mirasgedis, 2006; Franco and Sanstad, 2008; Psilogu *et al.*, 2009; Munoz *et al.*, 2010; Pilli-Sihvola *et al.*, 2010; among others). HDD and CDD are calculated using the following functions:

where ADT is the average daily temperature at time and T_r is the reference temperature.

In this paper a piecewise linear regression model is used with two reference temperature values which are estimated using the MARS algorithm (Friedman, 1991).

Figure 3 shows the plot of the model in equation (5). The piecewise linear regression plot separates the non-linear response of electricity demand to temperature into three regions: cold for temperatures lower than 18°C, neutral for temperatures between 18°C and 22°C, and hot for temperatures above 22°C.

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).

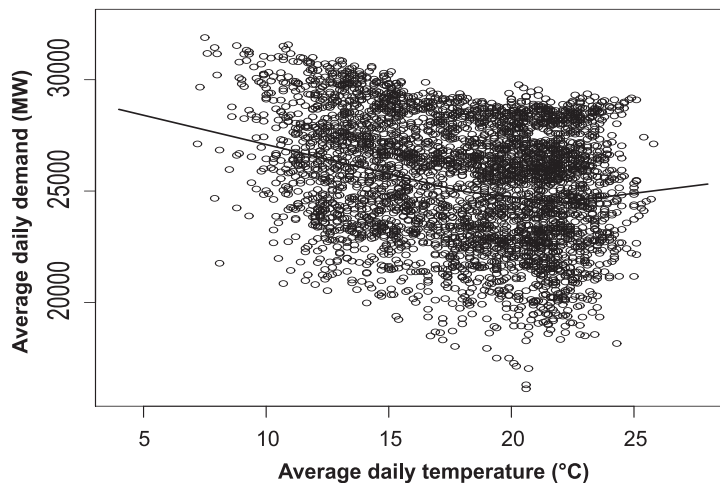


Figure 3: Scatter plot of ADED (seasonally adjusted) against ADT with the fitted piecewise regression line of equation (5)

4.2 Modelling the minimum daily temperature (tail quantile estimation)

In section 4.1 it was noted that demand of electricity is more sensitive to cold temperatures (less than 18°C) than to hot temperatures (more than 22°C). Modelling of extreme minimum temperatures is therefore important to load forecasters and system operators for planning, load flow analysis and scheduling of electricity. In this section, we estimate the extreme tail quantiles of ADT below 18°C using the GEVD. The data is seasonally adjusted. There are 1649 observations below 18°C. Figure 4 shows data for temperature below 18°C.

4.3 Tail quantile estimation

We use the principle of duality between the distributions of minima and maxima as discussed in Section 2.2. Figure 5 shows the graph of $-x_i$, $i = 1, \dots, n$ where x_i represents temperature below 18°C. The R statistical package Ismev (Heffernan and Stephenson, 2013) is used to obtain the ML estimates. The estimates are given as (the standard errors are given in parentheses):

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

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 $\hat{\xi} < 0$) and the right endpoint is finite and is given by:

$$\mu - \frac{\sigma}{\xi} = -15.02544 - \frac{2.08}{-0.2004} = -4.646$$

This implies that for any degree decrease below 4.6°C there won't be any further increase in electricity demand.

The quantile-quantile (QQ) and probability-probability (PP) plots given in Figure 6 show that a Weibull distribution is a good fit to the data. The return level estimates are inside the 95% confi-

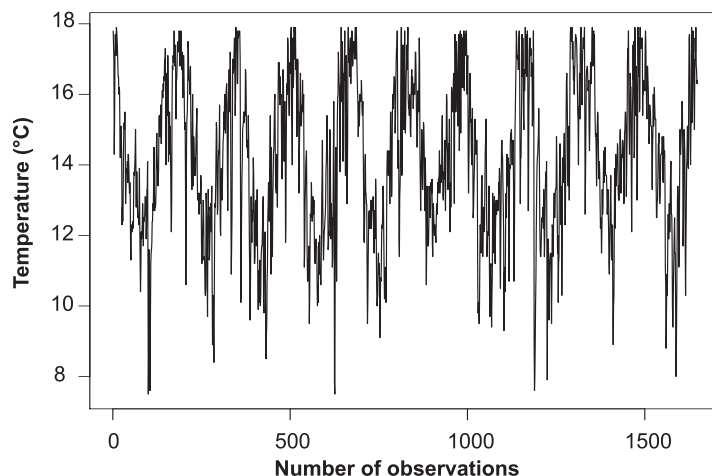


Figure 4: ADT below 18°C (2000 to 2010)

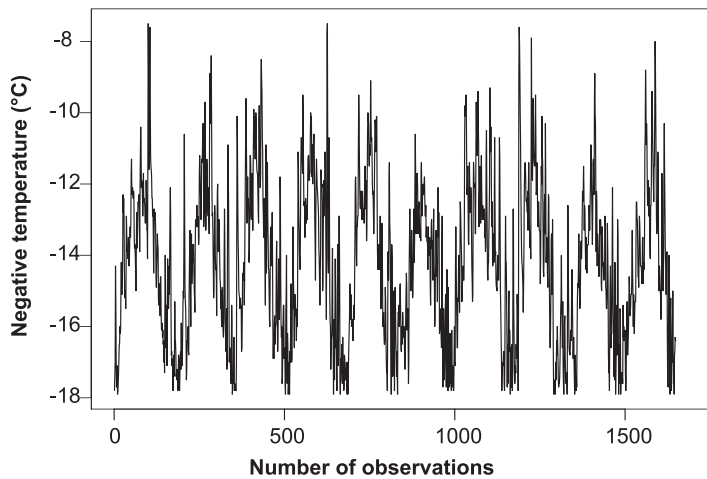


Figure 5: Inverted graph of Figure 4. Figure 4 is inverted so that we can use the duality between the distributions for maxima and minima

dence interval. This is an indication that the fitted distribution is capable of accurately predicting future return levels. We then use equation (4) to estimate the future return levels for different return periods. The return level is the quantile of the GEVD (Weibull distribution). For example, the 95th quantile is obtained as follows:

$$\begin{aligned}
 -x_{0.05} &= -15.025 - \frac{2.08}{0.2004} \left[\{-\ln(0.95)\}^{0.2004} - 1 \right] \\
 &= -10.369 \Rightarrow x_{0.05} = 10.369
 \end{aligned}$$

The number of 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 the number of temperature values below 18°C. The increase in electricity demand for a drop of temperature from 18°C to $x_{0.05} = 10.4$ °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°C below 18°C as discussed in Section 4.1. It should be noted that this

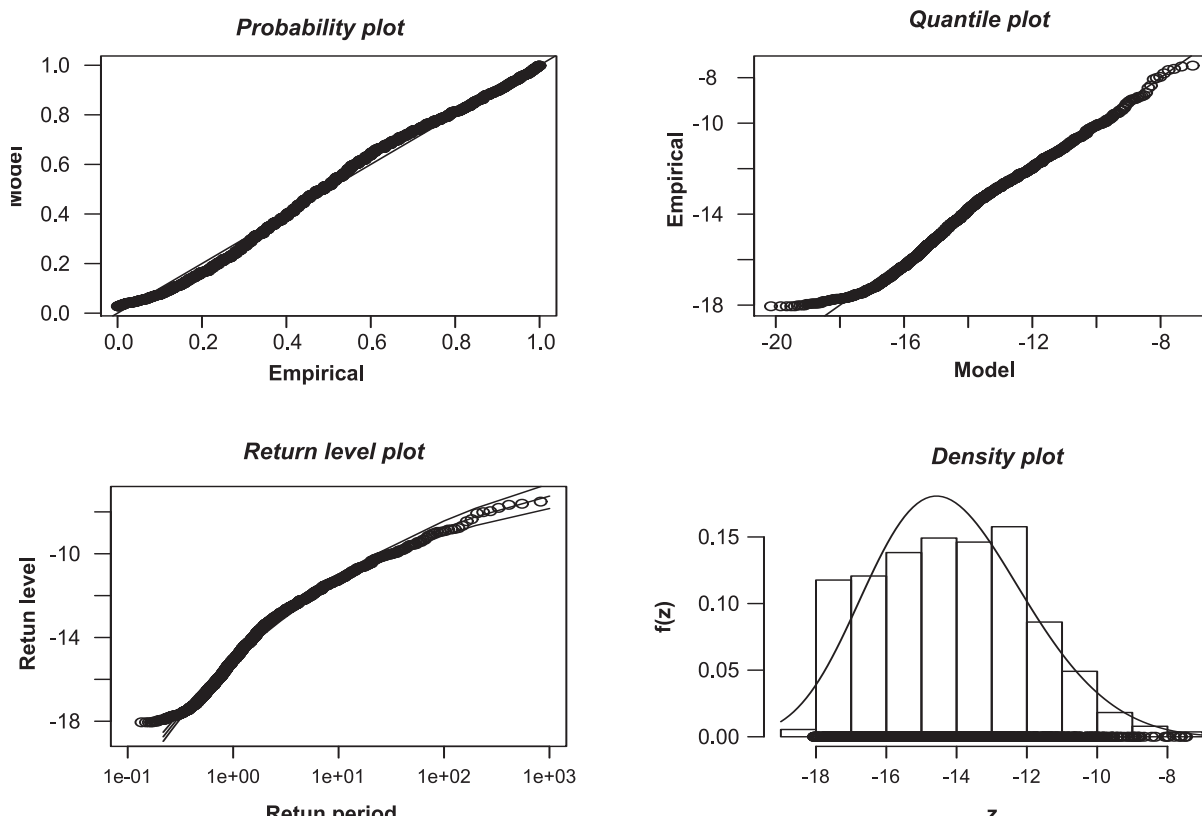


Figure 6: Diagnostic plots illustrating the fit of the data (temperature below 18°C) to the GEVD, (a) Probability plot (top left panel), (b) Quantile plot (top right panel), (c) Return level plot (bottom left panel) and (d) Density plot (bottom right panel)

Table 1: In-sample evaluation of estimated tail quantiles at different probabilities (number of exceedances)

Quantiles	Temp (xp)	Observed no. of exceedances	GEVD (no. of exceedances)	Marginal increase in demand (MW)
90 th	11.3 ^o C	164		159
95 th	10.4 ^o C	82	78	236.7
99 th	8.8 ^o C	16	12	420.8
99.5 th	8.2 ^o C	8	8	157.8

Table 2: Monthly frequency of exceedances below $x_{0.05} = 10.4^{\circ}\text{C}$

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

increase in ADED for temperature decreases below 18^oC is bounded above. As temperature decreases below 18^oC, the increases in ADED reaches a certain maximum 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 when all the heating systems are all switched on and no additional energy is consumed.

Table 1 presents a summary of the estimated tail quantiles at different tail probabilities. The tail quantiles (temperature) are given in column 2. The observed number of observations (temperature) that are smaller than the estimated tail quantiles are shown in column 3 while column 4 shows the corresponding number estimated using GEVD. In equation (5) it is given that for a degree decrease in temperature below 18^oC 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 increases from one quantile to the next, e.g. if temperature drops from 11.3^oC to 10.4^oC there

will be an increase in demand of 263(11.3 – 10.4) = 236.7MW. Similarly, for a decrease from 10.4^oC to 8.8^oC the marginal increase will be 263(10.4 – 8.8) = 420.8MW. Extreme low average daily temperatures of the order of 8.2^oC are very rare in South Africa. This only occurs about 8 times in a year.

Table 2 given in the Appendix summarizes the temperature values at high quantiles and the corresponding marginal increases while Figure 7 shows that the marginal increases converge to 1.58 MW when temperature converges to 4.6^oC.

A summary of the monthly frequency of occurrence of temperature values below 12.4^oC (i.e. above the 95th quantile $-x_{0.05} = 10.369$) is given in Table 2. 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^oC. 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 8.

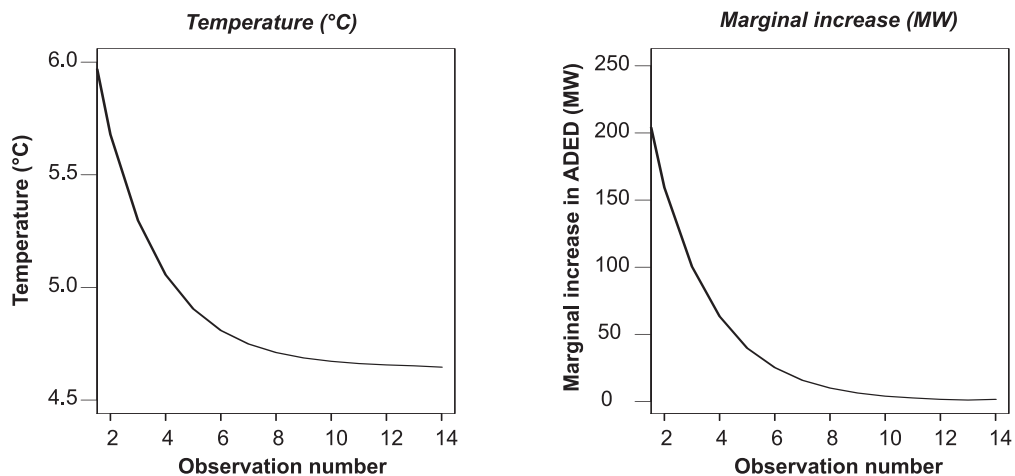


Figure 7: (a) Left panel shows that the gradual decrease in temperature converges to 4.6^oC and (b) Right panel shows that the marginal increases converge to 1.58MW when temperature converges to 4.6^oC

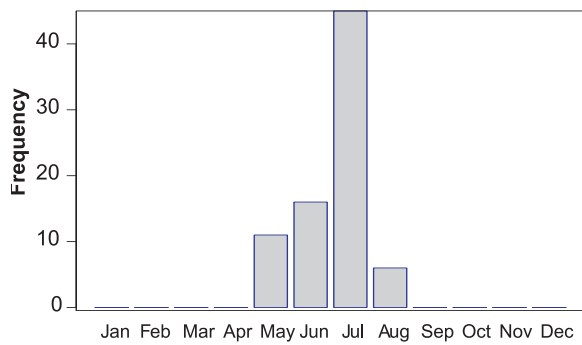


Figure 8: Bar chart of the monthly frequency of occurrence of exceedances ($x_{0.05} = 10.4$). The exceedances are average daily temperature values below 10.4°C

5. Conclusion

An analysis of the intensity and frequency of occurrence of extreme low temperatures is important for load forecasters in the electricity sector. In this paper, 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. The study establishes temperature as an important variable in explaining electricity demand. Empirical evidence from this study shows that for temperature values below 18°C demand for electricity in South Africa increases significantly while for temperature values above 22°C demand increases slightly. This analysis is important for decision makers in Eskom, South Africa's power utility company. Extreme low temperatures can be modelled by the Weibull class of distributions. Extreme low temperatures of the order of 6°C are very rare in South Africa, but can cause huge increases in electricity demand. An investigation of expected

cooler or warmer than typical years is important and helps in guiding planning to decision makers in the electricity sector.

Areas for future research would include a comparative analysis of the generalized extreme value distribution with a generalized Pareto distribution and a generalized single Pareto distribution in modelling extreme low temperatures in South Africa. These areas will be studied elsewhere.

Note

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 research, the weightings were not done since only aggregated average daily temperature was available.

Acknowledgments

The authors are grateful to Eskom for providing the data and to the numerous people who assisted in making comments on this paper.

References

- Chikobvu, D. and Sigauke, C., (2012). A frequentist and Bayesian regression analysis to daily peak electricity load forecasting in South Africa. *African Journal of Business Management*, 6(40): 10524-10533.
- ClimateTemp.info. (Accessed on 10 June 2012) www.climateemp.info/south-africa/.
- Franco, G. and Sanstad, A.H. (2008). Climate change and electricity demand in California. *Climatic Change*, 87 (Suppl 1): S139-S151.
- Heffernan, J.E. and Stephenson, A.G. (2013). Ismev: R package version 1.39.
- Hekkenberg, M., Benders, R.M.J., Moll, H.C. and

Appendix

Table 2: Evaluation of estimated tail quantiles using the GEVD quantile function

Observation number	1	2	3	4	5
Quantile	99.99 th	99.999 th	99.9999 th	99.99999 th	99.999999 th
Temp (deg C)	6.285	5.679	5.297	5.056	4.905
Marginal increase (MW)	252.74	159.38	100.5	63.38	39.713
Observation number	6	7	8	9	10
Quantile	99.9999999 th	99.99999999 th	99.999999999 th	99.9999999999 th	99.99999999999 th
Temp (deg C)	4.809	4.749	4.711	4.687	4.672
Marginal increase (MW)	25.248	15.78	9.994	6.312	3.945
Observation number	11	12	13	14	
Quantile	99.999999999999 th	99.9999999999999 th	99.99999999999999 th	99.99999999999999 th	
Temp (deg C)	4.662	4.656	4.652	4.646	
Marginal increase (MW)	2.63	1.578	1.052	1.578	

- 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.
- 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.
- Hyndman, R.J., Fan, S. (2010). Density forecasting for long-term peak electricity demand. *IEEE Transactions on Power Systems*, 25(2): 1142-1153.
- 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.
- Moral-Carcedo, J. and Vicéns-Otero, J. (2005). Modelling the non-linear response of Spanish electricity demand to temperature variations. *Energy Economics*, 27, 477-494.
- 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.
- 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), pp. 2409-2419.
- 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.
- SouthAfrica.info. (Accessed on 10 June 2012) southafrica.info/travel/advice/climate.htm.
- www.ral.ucar.edu/~ericg/softextreme.php.

Received 1 October 2012; revised 21 November 2013