

# Modelling Summer Daily Peak Loads in South Africa Using Discrete Time Markov Chain Analysis

Molete Mokhele<sup>1,\*</sup>, Caston Sigauke<sup>1,2</sup>

<sup>1</sup>School of Statistics and Actuarial Science, University of the Witwatersrand, Johannesburg, Private Bag 3, Wits 2050, South Africa

<sup>2</sup>Department of Statistics, University of Venda, Private Bag X5050, Thohoyandou 0950, South Africa

\*Corresponding Author: mokhelemolete01@gmail.com

Copyright ©2015 by authors, all rights reserved. Authors agree that this article remains permanently open access under the terms of the Creative Commons Attribution License 4.0 International License

**Abstract** Electricity demand exhibits a large degree of randomness in South Africa, particularly in summer. Its description requires a detailed analysis using statistical methodologies, in particular stochastic processes. The paper presents a Markov chain analysis of peak electricity demand. The data used is from South Africa's power utility company Eskom, for the period 2000 to 2011. This modelling approach is important to decision makers in the electricity sector particularly in scheduling maintenance and refurbishments of power-plants. The randomness effect is accountable to meteorological factors and major electricity appliance usage. Aggregated data on daily electricity peak demand is used to develop the transition probability matrices, steady-state probabilities, mean return- and the first passage times. Such analysis are important to Eskom and other energy companies in planning load-shifting, load analysis and scheduling of electricity particularly during peak period in summer.

**Keywords** Daily Peak Electricity Demand, Discrete Time Markov Chain, Steady State Probability, Passage Time, Mean Return Time, Transition Matrix

---

## 1 Introduction

Electricity demand exhibits a large degree of randomness, particularly in South Africa. South Africa is faced with the challenge to appropriately schedule power-station plant maintenance and insufficient potential to meet the electricity demand for consumers; which is mainly due to the meteorological influences, use of major electricity appliances and the demand also varies with consumers' behavior.

South Africa is heavily reliant on Eskom for the supply of Electricity. Eskom generates approximately 95% of the electricity consumed in South Africa and approximately 45% of the electricity consumed in Africa, [6].

Electricity is the most important commodity, a social advancement and a core driver of the economy. Hence the development of electricity infrastructure and refurbishment of the power-station plant is vital.

While huge investments have already been made on new power-station plants for generation of electricity, accelerated maintenance expansion is needed to unlock the growth potential of sectors such as Tourism and Agriculture. However, there is great uncertainty of consumers demand usage.

Despite there being sufficient generation capacity of electricity accommodating a large fraction of the South African population for now, the demand for electricity by consumers keeps on increasing due to population growth, rural electrification and industrial development. Electricity demand in South Africa vary from sector to sector, and the major driver is mainly temperature as discussed by [4]. Under these circumstances, South Africa's power utility company is faced with the challenge to keep up with the rising electricity demand. As noted by [2] the electricity demand varies with season. Demand is significantly higher in winter than in summer, but the individual behaviour has confirmed demand in summer can be significantly higher. Due to the heavy demand and stochastic fluctuation of the electricity demand by consumers in winter, summer is the appropriate time for the grid companies to perform restorations and maintenance on power-station plants.

The use of electricity major appliances (refrigerators, ventilators, air-conditioners, e.t.c) account for a large proportion of overall daily demand in summer. This becomes problematic for Eskom to perform maintenance and refurbishments on the power-station plants. The paper focuses on the use of Markov chain analysis to model summer daily peak demand to help schedule maintenance of power plants.

Scheduling maintenance means knowing the exact occurrences of extreme demand changes (electricity demand above sufficiently high threshold). This paper contributes towards understanding the impact of daily extreme electricity demand in South Africa during the

summer season and the various interactions that exist between the various factors and variables that drive demand.

We model the summer extreme positive daily changes in peak electricity demand, using discrete time Markov chains (DTMCs) analysis. We calculate the transition probabilities, mean return times, first passage times and the steady state probabilities. Extreme Value Theory (EVT) is then used to calculate the summer extreme increases in peak electricity demand. This tool provides decision makers in the electricity sector to help in scheduling and planning and risk management.

The rest of the paper is organized as follows. Section 2 provides the models. A practical application and discussion of the tail quantile estimation and calculation of probability of exceedance levels using the Generalized Pareto Distribution (GPD). The last section concludes and provides short overview on the findings.

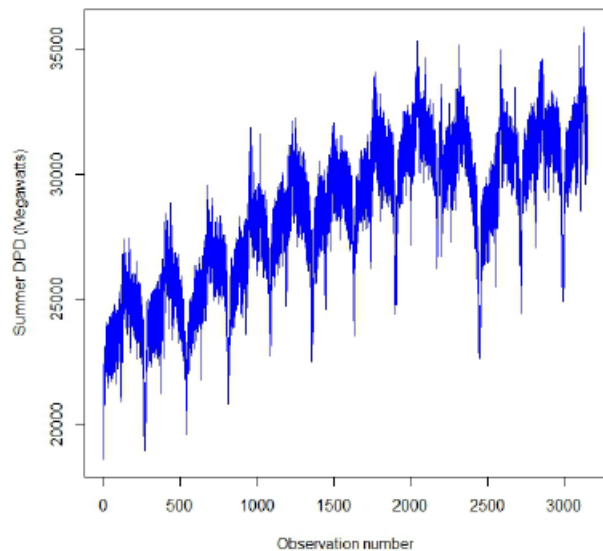
### 1.1 Discrete time Markov chains

A Markov chain, in general, is a way to describe a movement of state from one position to another. Additionally, discrete time Markov Chain (DTMC) analysis is a type of Markov chain in which a finite state space exists and one of the advantages of using the Markov chain models captures all dominant factors of uncertainty [9]. In detail, DTMC involves modelling transition probabilities of the daily peak electricity demand changes in different hours. Firstly, we define the two states; positive daily changes (increase) and negative daily changes (decrease), and the probability distribution of the transitions from one state to another is represented by a transition matrix  $P$ , where each element of position  $\{i, j\}$  represents the probability  $p_{ij}$ .

Notably, this technique is characterized by the Markov property which is fully explained in section 1.C. Furthermore, we calculate the steady state probabilities, mean return- and first passage times of the daily peak electricity demand in different hours. Henceforth, the two-state problem is then extended to the three-state problem, where the daily electricity demand changes in different hours are split into the following states: small increases, extreme increases of the hourly peak demands and last, negative hourly changes are classified as decreases. Prior, this is done by fitting a Pareto quantile plot to the positive daily changes to determine a sufficiently high threshold ( $\tau$ ) with all the exceedances regarded as extreme positive hourly demand changes and those below  $\tau$  as small increases. Further, a transition probability matrix is developed from which the steady state probabilities, mean return- and first passage times are calculated. We use Generalized Pareto Distribution (GPD) to calculate extreme tail quantiles and Probability of Exceedance (PoE) of the extreme peak electricity demand in different hours.

### 1.2 Description of the data

The transition matrices, steady-state probabilities, mean return and first passage times are developed from the summer daily electricity demand (SDPED) historical data of electricity demand from Eskom (South African



**Figure 1.** A Time-series plot of the summer daily peak demand. The horizontal axis ( $x$ -axis) represents the observation number while the summer daily peak electricity demand is plotted on the vertical axis ( $y$ -axis) for the period 16 August to 14 May of each year (2000-2011).

utility company). Eskom is responsible for facilitating and supplying electricity to the whole South African population (about 50 million people) and other parts of Southern Africa. The historical data comprises of the aggregated hourly data (in MegaWatts) for the period 2000-2010. Aggregated data comes from various sectors (manufacturing, mining, tourism, communication, tourism, transport, finance and business services, e.t.c.) that contribute to the well being of South Africa's economic growth. The winter period in South Africa is defined as the date from 15 May- 14 August of every year and 15 August- 14 May is considered the summer period.

Figure 1 shows a time series plot of the summer daily peak electricity demand, and exhibits both a positive upward trend and strong seasonality. As clearly depicted in Figure 1, the data is not stationary. In view of making the data stationary there are several ways available and in this study we use the differencing. The data (summer daily peak electricity demand) is stationary after the first difference as shown in Figure 2(a). The first difference of the series is given as:

$$z_t = \nabla x_t = (1 - \beta)x_t = x_t - x_{t-1} \quad (1)$$

where,

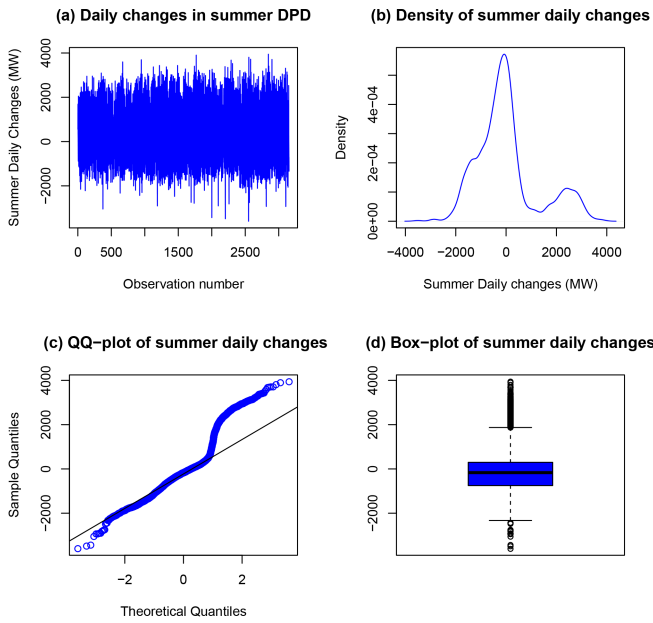
$z_t$  denotes summer daily changes in peak electricity demand, and

$x_t$  denotes summer daily peak electricity demand on day  $t$ .

In order for us to determine the states, we initially select the threshold to be zero and this results in a two-state problem with the states defined as:

state 1: when  $x_t - x_{t-1} > 0$  (increase state)

state 2: when  $x_t - x_{t-1} \leq 0$  (decrease state)



**Figure 2.** (a) Plot of summer daily peak demand (top left panel), (b) Density plot of daily changes in SDPED (top right panel), (c) Normal QQ-plot of daily changes in SDPED (bottom left panel), (d) Box plot of daily changes in SDPED (bottom right panel)

### 1.3 Markov Property

The Markov property states that the distribution of the forthcoming state  $X_{n+1}$  depends only on the current state  $X_n$  and does not depend on the previous ones  $X_{n1}, X_{n2}, \dots, X_1$ . This is a sequence of random variables characterized by a memorylessness property. Correspondingly, the transition probabilities of an increase or decrease of summer daily peak electricity demand (SDPED) depend only on the current state (either a decrease (state 2) or an increase (state 1)) and not on past evolution of DPED. This can be represented mathematically as:

$$\begin{aligned} Pr(X_{n+1} = x_{n+1} | X_1 = x_1, \dots, X_n = x_n) \\ = Pr(X_{n+1} = x_{n+1} | X_n = x_n) \end{aligned} \tag{2}$$

### 1.4 Estimation of transition probabilities (daily changes)

The chain successively moves from one state to another (this change is named either 'transition' or 'step'), [7].

Let,  $p_{ij}$  = the probability of moving from state  $i$  to state  $j$ , where  $i, j = 1, 2$ . Mathematically this can be represented as:

$$p_{ij} = Pr(X_1 = s_j | X_0 = s_i) \tag{3}$$

where,

state 1: increase (1)

state 2: decrease (2)

The two state transition matrix is given as:

$$P = \frac{1}{2} \begin{pmatrix} 1 & 2 \\ p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix} \tag{4}$$

It can be observed that if we let  $p_{12} = \gamma$  then  $p_{11} = 1 - \gamma$ . Equally, if we let  $p_{21} = \psi$  then  $p_{22} = 1 - \psi$ . The

transition matrix above can be written as:

$$P = \frac{1}{2} \begin{pmatrix} 1 & 2 \\ 1 - \gamma & \gamma \\ \psi & 1 - \psi \end{pmatrix} \tag{5}$$

The state probabilities of the transition matrix,  $p_{ij}$  are estimated using the formulae,

$$\hat{p}_{ij} = \frac{n_{ij}}{n_i} \tag{6}$$

where  $n_{ij}$  is the number of observed transitions from state  $i$  to state  $j$  and  $n_i = \sum_{j=1}^n n_{ij}$ .

The set of all possible states  $S = (S_1, S_2)$  of the summer daily peak electricity demand is the state space of the chain.

### 1.5 n-step transition probabilities

The entries,  $p_{ij}$  in the transition matrix  $P$  represent the transition probabilities. Then the  $n$ -step transition probability of a Markov Chain is defined as the probability that the system goes from state  $i$  to state  $j$  in  $n$  steps. Consequently, the  $n$ -step transition probability is given by the equation

$$P_{ij}^{(n)} = Pr(X_{m+n} = j | X_m = i) = P(X_n = j | X_0 = i) \tag{7}$$

and,  $p^{(n)} = \{p_{ij}^{(n)}\}$ . Then the associated 2-step transition probability of the two-state problem is given by;

$$P = \frac{1}{2} \begin{pmatrix} 1 & 2 \\ p_{11}^{(n)} & p_{12}^{(n)} \\ p_{21}^{(n)} & p_{22}^{(n)} \end{pmatrix} \tag{8}$$

Notably, as  $n$  approaches infinity,  $P^{(n)}_{ij}$  converges and begins to steadily settle to some limiting value ( $\eta$ ), independent of some initial value.  $\eta_i$  is the limiting steady probability of state  $i$ , and the steady state probability of the Markov chain can be represented mathematically as:

$$r_{di}(n) = P(X_n = i | X_0 = d) = \sum_{k=1}^n r_{dk}(n-1)P_{k1} \tag{9}$$

$$\lim_{x \rightarrow \infty} r_{12}(n) = \sum_{k=1}^n \lim_{x \rightarrow \infty} r_{dk}(n-1)P_{k1} \tag{10}$$

$$\Rightarrow \eta_i = \sum_{k=1}^n \eta_2 P_{k1} \tag{11}$$

Now for the two-state problem, the steady state probabilities of the states is given as:

$$\boldsymbol{\eta} = (\eta_1, \eta_2) \tag{12}$$

### 1.6 Mean Return times

For an ergodic (irreducible) Markov chain system, the mean return time of state  $l$  is the first visit time  $T_l$  after a number of transitions. This is represented mathematically as:

$$MR_l = E(T_l) = \frac{1}{\eta_l} \tag{13}$$

where  $\boldsymbol{\eta} = [\eta_1, \dots, \eta_r]$  is the stationary probability vector of the transition probability matrix  $\mathbf{P}$ , and  $r$  is the number of states.

### 1.7 First Passage times

The "first passage time" (FPT),  $(T_{ij})$  from state  $i$  to state  $j$  is the number of transitions (steps), taken by the Markov chain until it arrives for the first time at state  $j$  given that the initial state was  $i$ ,  $(X_0 = i)$ . The probability distribution of the first passage time is given by equations (10) and (11).

$$h_{ij}^{(n)} = P(T_{ij} = n) \tag{14}$$

$$= P(X_n = j, X_{n-1} \neq j, \dots, X_1 \neq j | X_0 = i) \tag{15}$$

Now for  $n = 1$  in equation (10) we have,

$$h_{ij}^{(1)} = p_{ij} \tag{16}$$

and for  $n \geq 2$ , the first passage time is given by:

$$h_{ij}^{(n)} = \sum_{k \in S-j} p_{ik} h_{kj}(n-1) \tag{17}$$

## 2 Empirical results and discussion

Using the R-statistical software called "Markovian" package developed by [7], the transition matrix for summer daily peak electricity demand (SDPED) for states 1 and 2 is:

$$P = \begin{matrix} & \begin{matrix} 1 & 2 \end{matrix} \\ \begin{matrix} 1 \\ 2 \end{matrix} & \begin{pmatrix} 0.4295830 & 0.5704170 \\ 0.3846971 & 0.6153029 \end{pmatrix} \end{matrix}$$

As clearly depicted by the transition matrix  $\mathbf{P}$ , the probability that the current peak electricity demand changes from state 1 (increase) and back to state 1 again the next day is 0.4295830. In detail, there is an approximate 43.0 % chance that the system experiences an increase from state 1 and back to state 1 again, and while the probability of moving from state 1 to state 2 is 0.5704170. The transition matrix  $\mathbf{P}$  indicates, the probability that the system moves from state 2 (decrease) and back to state 2 again is 0.6153029. This kind of information is crucial in proper planning, analysis and facilitation of power system operations in order to ensure that outages, blackouts and load shedding do not occur. The steady state probabilities of the two-state problem are:

$$\boldsymbol{\eta} = (\eta_1, \eta_2) = (0.4027761, 0.597224)$$

The mean return times for the two-state problem are :

$$MR_1 = E(T_1) = \frac{1}{0.4027761} = 2.5 \text{ days for state 1 (increase) and}$$

$$MR_2 = E(T_2) = \frac{1}{0.597224} = 1.7 \text{ days for state 2 (decrease).}$$

The results indicate that, given the current state of the system is an increase, the system can expect to experience another increase in about 2.5 days (approximately 60 hours). Similarly, given the current state of the system is a decrease (state 2) then we expect to have another decrease in about 1.7 days, (approximately 41 hours).

**Table 1.** First passage time probability given current state 1 (increase) for the next seven days

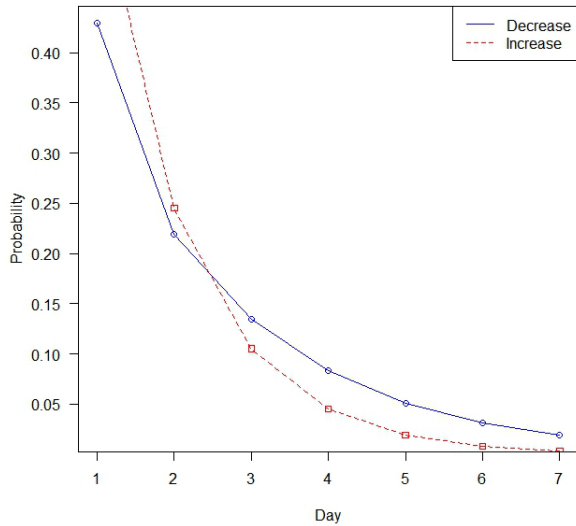
Day	State 1 (increase)	State 2 (decrease)
1	0.42958300	0.570417000
2	0.21943777	0.245041446
3	0.13502069	0.105265640
4	0.08307862	0.045220329
5	0.05111852	0.019425885
6	0.03145337	0.008345030
7	0.01935335	0.003584883

A short-term summary of the first passage time probabilities (for the next seven days) is depicted in Table 1. The results indicate, given that the current state of the system is state 1, for the next upcoming seven days the probabilities of the system to continue to experience state 1 (increase) is increasingly decreasing exponentially. Meanwhile, given the current state of the system is state 1, the probability is 42.9% that the next day the system will be in state 1 again and meanwhile the probability that the system will be in state 2 is 57.0%. Notably, this information is very important to the forecasters to determine which periods are most appropriate to undertake maintenance and monitor the security of the national grid. It should be noted energy security is very important for the economic growth of any country [9].

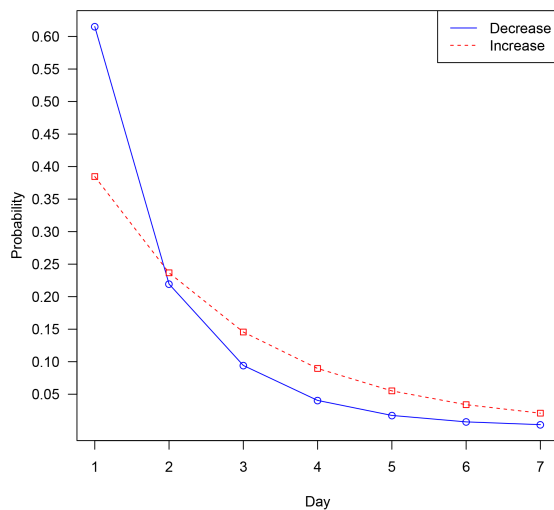
**Table 2.** First passage time probability given current state 2 (increase) for the next seven days

Day	State 1 (increase)	State 2 (decrease)
1	0.38469710	0.615302900
2	0.23670524	0.219437766
3	0.14564542	0.094266734
4	0.08961605	0.040495386
5	0.05514102	0.017396130
6	0.03392843	0.007473082
7	0.02087626	0.003210309

In Table 2, a summary of the short-term first passage time probabilities (for the next seven days) given the current state of the system is state 2 is depicted. The probabilities for the first passage times exhibited by the system for the next upcoming seven days show sign of decreasing. The probability that the system will be in state 2 (decrease) during the first day is approximately 62% and the second day is approximately 21.9%.



**Figure 3.** First passage time given current state is an increase (state 1)



**Figure 4.** First passage time given current state is a decrease (state 2)

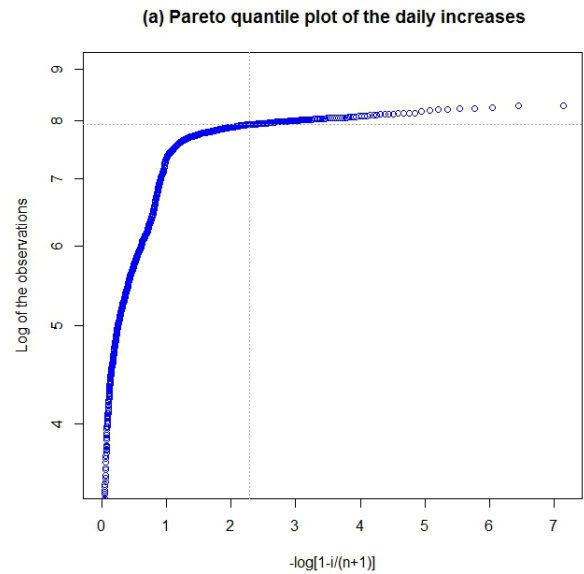
Figure 3 shows that the shape of the two states (increase and decrease) distributions of positive daily changes is similar. The distribution of the probabilities exhibits some exponential distribution with time (days). Similarly figure 4 shows that the shape of the two states (increase and decrease) distributions of positive daily changes is nearly similar. The distribution of the probabilities exhibits some exponential distribution with time (days).

### 2.1 Modelling extreme peaks for the two-state problem

We fit a Pareto quantile plot to the positive summer daily changes to determine a sufficiently high threshold,  $\tau$ . All the observations above the determined threshold (exceedances) are then defined as extreme daily positive changes which will be classified as (state 1) and for all the observations less than or equal to the threshold ( $\tau$ ) as (state 2), i.e  $z_t \leq \tau$ .

A 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 changes in summer peak electricity demand, for all  $i = 1, \dots, n$ .

Figure 5 shows that the threshold,  $\tau$  is given by  $\tau = \exp(7.8846) = 2656.1$  MW with 164 number of observations in the tail.



**Figure 5.** (a) Pareto quantile plot of the daily increases in summer peak electricity demand.

The transition matrix for the two states which are: state 1: extreme increases (observations above  $\tau$ ), state 2: no extreme increase (observations below  $\tau$ ), is given as:

$$P = \begin{matrix} & \begin{matrix} 1 & 2 \end{matrix} \\ \begin{matrix} 1 \\ 2 \end{matrix} & \begin{pmatrix} 0.00000000 & 1.00000000 \\ 0.05522088 & 0.9447791 \end{pmatrix} \end{matrix}$$

and the steady state probabilities are:

$$\eta = (\eta_1, \eta_2) = (0.05233111, 0.9476689) \quad (18)$$

The mean return times for the two states are:

$$MR_1 = \frac{1}{\eta_1} = \frac{1}{0.05233111} = 19.1 \text{ days for state 1 (extreme increase)}$$

This gives approximately 19 days in the whole year and approximately 14 days in the whole of summer period (272 days) and

$$MR_2 = \frac{1}{\eta_2} = \frac{1}{0.9476689} = 1.06 \text{ days for state 2 (no extreme increase)}$$

**Table 3.** First passage time probability given current state 1 for the next seven days ( $z_t > 2656.1$  MW)

Day	State 1 (extreme increase)	State 2 (no extreme increase)
1	0.00000000	1
2	0.05522088	0
3	0.05217153	0
4	0.04929057	0
5	0.04656870	0
6	0.04399714	0
7	0.04156758	0

Given the current state is an extreme increase, Table 3 shows the probabilities of change in the two states(

state 1 and state 2). It is less intuitive, but true that the probability of having another extreme increase the following day is 0 while there is absolute certainty that the following day will result in no extreme increase.

**Table 4.** First passage time probability given current state 2 for the next seven days ( $z_t \leq 2656.1$  MW)

Day	State 1 (extreme increase)	State 2 (no extreme increase)
1	0.05522088	0.94477910
2	0.05217153	0.05522088
3	0.04929057	0.00000000
4	0.04656870	0.00000000
5	0.04399714	0.00000000
6	0.04156758	0.00000000
7	0.03927218	0.00000000

Given the current state is no extreme increase, Table 4 shows that the probability that the next day we will observe an extreme increase is 0.05135045 while the probability that it will be in the same state (no extreme increase) is 0.94864950.

### 2.2 Three-state Problem

We extend the two-state problem to the three-state problem. The three-state problem involves categorization of positive and negative summer daily changes. The positive daily changes are categorized into two states; small changes and extreme daily changes and the negative daily changes are categorized as decreases. The three states are defined as:

- State 1: observations above  $\tau$  (extreme increase), i.e.  $z_t > 2656.1$  MW .
  - State 2: observations between 0 and  $\tau$  (small increase), i.e.  $(0 < z_t < 2656.1$  MW),
  - State 3: observations below zero (decreases in peak electricity demand, i.e.  $z_t \leq 0$ ).
- The transition matrix for the three state problem is given as:

$$P = \begin{matrix} & \begin{matrix} 1 & 2 & 3 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \end{matrix} & \begin{pmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{pmatrix} \end{matrix} \quad (19)$$

After computing the probabilities we get the transition probabilities.

$$P = \begin{matrix} & \begin{matrix} 1 & 2 & 3 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \end{matrix} & \begin{pmatrix} 0.00000000 & 0.5151515 & 0.4848485 \\ 0.03342367 & 0.3830172 & 0.5835592 \\ 0.06804891 & 0.3173844 & 0.6145667 \end{pmatrix} \end{matrix}$$

and the steady state probabilities of the three-state problem is given as:

$$\eta = (\eta_1, \eta_2, \eta_3) = (0.05234204, 0.3507571, 0.5969009)$$

The mean return times of the three-state problem is given as:

$$MR_1 = \frac{1}{\eta_1} = \frac{1}{0.05234204} = 19.1 \text{ days for the extreme increase state } (z_t > 2656.1).$$

The return period is  $\simeq 19.1$  days giving approximately  $\frac{272}{19.1} = 14$  days in a year of extreme increases during the summer period  $MR_2 = \frac{1}{\eta_2} = \frac{1}{0.3507571} = 2.9$  days for the small increase state ( $0 < z_t \leq 2656.1$ ).

This gives approximately 94 days in the whole of summer period.

$$MR_3 = \frac{1}{\eta_3} = \frac{1}{0.5969009} = 1.7 \text{ days for the decrease state } (z_t \leq 0).$$

This gives approximately 160 days in the whole of the summer session. Table 5 shows that given the current state is an extreme increase, the probability that the next day we will observe a small increase is 0.51515150 and the probability that we will observe an extreme increase is 0.000000.

**Table 5.** First passage time given current state is a decrease ( $z_t \leq 0$ ).

Day	State 1 (extreme increase)	State 2 (small increase)	State 3 (decrease)
1	0.06804891	0.31738440	0.614566700
2	0.05242875	0.23010938	0.218205998
3	0.04888758	0.15188916	0.096539911
4	0.04613874	0.10093811	0.040733573
5	0.04357423	0.06704454	0.017263908
6	0.04115385	0.04453364	0.007313735
7	0.03886799	0.02958092	0.003098540

**Table 6.** First passage time given current state is an extreme increase ( $z_t > 2656.1$  MW).

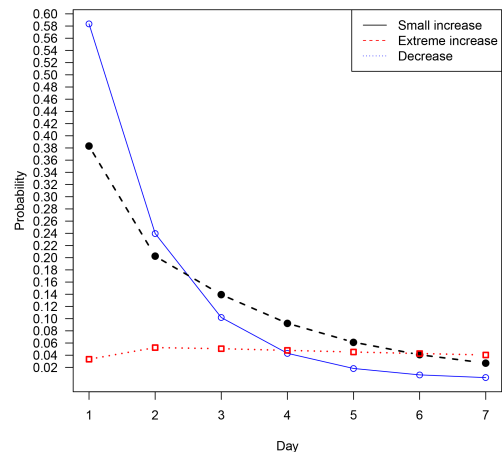
Day	State 1 (extreme increase)	State 2 (small increase)	State 3 (decrease)
1	0.00000000	0.51515150	0.484848500
2	0.05021167	0.15388335	0.300621397
3	0.05247184	0.11156819	0.123491410
4	0.04982559	0.07364323	0.052475510
5	0.04707233	0.04893969	0.022225329
6	0.04445850	0.03250645	0.009416220
7	0.04198913	0.02159207	0.003989256

Given the current state is an extreme increase in the system, the probability that the next day we will observe an extreme increase in summer daily changes is 0 as shown in Table 6 and on day 2 is 0.05021167.

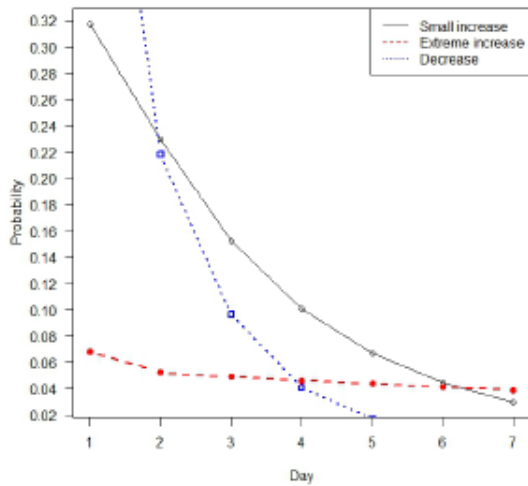
**Table 7.** First passage time given current state is a small increase ( $0 < z_t < 2656.1$  MW).

Day	State 1 (extreme increase)	State 2 (small increase)	State 3 (decrease)
1	0.03342367	0.38301720	0.583559200
2	0.05251241	0.20243084	0.239718627
3	0.05070843	0.13942579	0.101864228
4	0.04795100	0.09236533	0.043143287
5	0.04529075	0.06136479	0.018278545
6	0.04277528	0.04076020	0.007743849
7	0.04039938	0.02707450	0.003280752

Table 7 shows that the probability that the next day we will observe an extreme increase is 0.03342367.



**Figure 6.** First passage time given current state is a small increase (state 2)



**Figure 7.** First passage time given current state decrease (state 3)

Figure 7 present the first passage time, given the current state is a decrease (state 3), the probability distribution of state 1 and state 2 decrease exponentially for the next 7 days. Figure 6 present a first passage time given current state of the system is state 2 (small increase), the probability distribution shape of state 1 (small-increase) and state 3 (decrease) seem to be decreasing exponentially and state 2 is slowly increasing decreasingly.

### 3 Probability of exceedance levels using Generalized Pareto distribution (peaks-over-threshold distribution)

We consider the peaks (summer daily peak electricity demand) above the threshold  $\tau = 2656.1$  MW. There are several ways to model the observations above the threshold (exceedances). We use the Generalized Pareto Distribution (GPD) to model the observations above the threshold discussed in [3]. Using the Maximum Likelihood method discussed in section 3, we will estimate the parameters  $(\sigma, \xi)$ .

Using R statistical package "ismev" by [5]. The estimates and their respective standard errors are given in Table 8.

**Table 8.** MLE Estimates of excesses above the threshold

Parameters	Estimates (MLE)	Standard Errors (SE)
$\hat{\sigma}$	380.8571547	38.45899130
$\hat{\xi}$	-0.1990201	0.06589998

From Table 8, we can see that the standard errors for the parameter estimates are small. This confirms that the level of uncertainty in estimating the parameters is small. The 95 percent confidence intervals for the parameters  $\sigma$  and  $\xi$  are  $(\hat{\sigma} \pm Z_{\alpha/2} S.E) = (305.478; 456.237)$  and  $(\hat{\xi} \pm Z_{\alpha/2} S.E) = (-0.3282; -0.06986)$  respectively.

Since the shape parameter  $\xi < 0$  then the Generalized Pareto Distribution of the excesses belongs to the Weibull class of distributions (Beta Distribution). The distribution has a short tail with finite upper end point

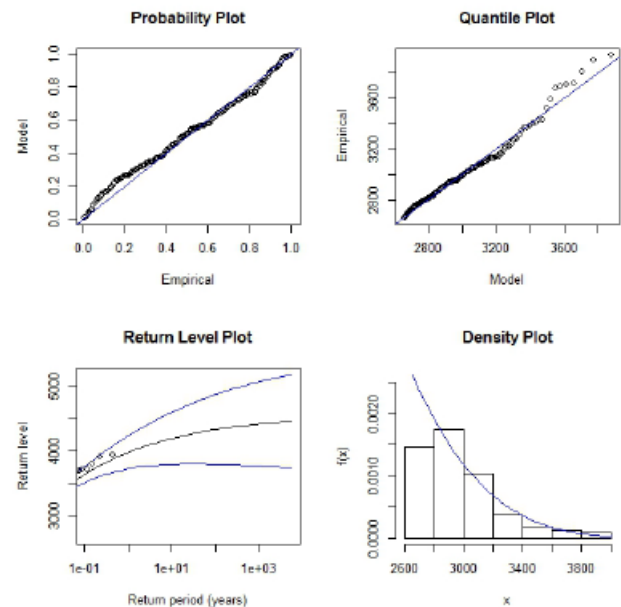
$\tau - \frac{\sigma\tau}{\xi}$  if  $\xi < 0$ . The upper end point is  $\hat{\xi} - \frac{\hat{\sigma}\tau}{\hat{\xi}} = 2656.1 - \frac{380.8}{0.199} = 4569.67$ MW.

The quantiles of the distribution can be estimated from the inverse function of the the Generalized Pareto Distribution. We let:

$$p = \left[ 1 + \xi \left( \frac{z_p - \tau}{\sigma} \right) \right]^{-\frac{1}{\xi}} \tag{20}$$

where  $p$  is the upper tail probability and  $z_p$  is the return level with period of  $1/p$ . Making  $z_p$  subject of the formula we get,

$$z_p = \tau + \frac{\sigma}{\xi} (p^{-\xi} - 1) \tag{21}$$



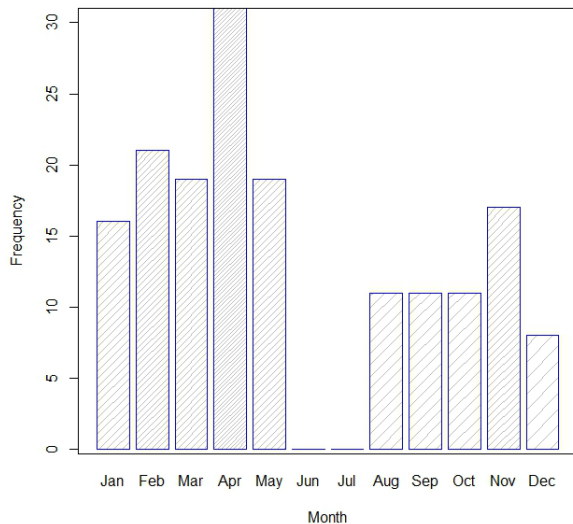
**Figure 8.** Diagnostic plots illustrating the fit of exceedances to the Generalized Pareto Distribution, (a) Probability plot (top left panel), (b) Quantile plot (top right panel), (c) Return level plot (bottom left panel) and (d) Density plot (bottom right panel)

The diagnostic plots of the exceedances above the threshold  $\tau = 2656.1$  MW are shown in Figure 8. Figure 8 shows that the GPD is a good fit to the exceedances.

**Table 9.** Estimated tail quantiles at different probabilities (Number of exceedances)

Tail Probability	Quantiles	estimates ( $z_p$ )	Observed ( $z_p$ )
0.1	90 <sup>th</sup>	3358.03 (16)	3371 (16)
0.05	95 <sup>th</sup>	3513.62 (9)	3677 (8)
0.025	97.5 <sup>th</sup>	3649.17 (7)	3806 (4)
0.01	99 <sup>th</sup>	3801.93 (3)	3903 (2)

Table 9 shows the number of exceedances relative to the corresponding tail probabilities. At 99<sup>th</sup> percentile, the value is 3801.93. The number of observations that are greater than the estimated tail quantile ( $Z_{0.05} = 3801.93$ ) are then counted and found to be 3 for the number of exceedances, we get  $0.05 * 164 = 8.2 \simeq 8$ , where 164 is the number of observations above the threshold.



**Figure 9.** Bar chart of the monthly frequency of occurrence of summer extreme daily increases in peak electricity demand above the threshold,  $\tau = 2656.1$  MW

Figure 9 present the frequency analysis of the summer daily peak electricity demand changes over the sampling period of 12 years. Clearly in South Africa, extreme summer daily changes occur mostly in April and the least daily changes are observed in December. More importantly, an accurate assessment of the frequency and analysis of summer daily peak loads help system operators and decision makers in determining the critical peak days, scheduling maintenance, load flow analysis and dispatching of electricity demand.

## 4 Conclusion

The paper discussed a Markov chain analysis of summer daily changes in peak electricity demand. Empirical results show that the steady state probability of an extreme increase ( $z_t > 2656.1$  MW) in daily peak electricity demand is 0.05233111. This resulted in a mean return time of about 19 days. We extended the two-state problem to the three-state problem. The steady state probability of an extreme increase was calculated as 0.05234204 with a return period of 19.1 days, giving approximately 14 days of extreme increases in the summer period. Extreme daily electricity demand changes occurred mostly in April while December had the least occurrence of extreme daily changes for all summer months. December appears to be an appropriate time to carry out maintenance on power-station plants. Such an analysis is important; they help system operators and decision makers in determining the critical peak days in order to schedule maintenance avoiding unnecessary

power outages.

Discrete time Markov chain analysis is very important in modelling electricity demand, however further research is necessary. Areas for further research include incorporating various economic variables (Gross Domestic Product), hourly electricity demand, natural gas price and involve long-term predictions of the electricity demand. Comparative analysis with the Bayesian inferential estimates will be done as well in the future.

## Acknowledgements

The authors are grateful and thankful to the National Research Foundation (NRF) in South Africa for funding this research and Eskom for providing the data and numerous people for their comments and suggestions.

## REFERENCES

- [1] Ardakanian, Omid and Keshav, Srinivasan and Rosenberg, Catherine. (2011). Markovian models for home electricity consumption, *Proceedings of the 2nd ACM SIGCOMM workshop on Green networking.*, pp. 31-36.
- [2] Balachandra, P. and Chandru, Vijay (1999). Modelling electricity demand with representative load curves, *Energy*, Vol. 24, pp. 219-230, *Elsevier*.
- [3] Beirlant, J., Goegebeur, Y., Segers, J. and Teugels, J. (2006). *Statistics of extremes: theory and applications*, John Wiley & Sons, England.
- [4] Chikobvu, D. and Sigauke, C. (2013). Modelling influence of temperature on daily peak electricity demand in South Africa, *Journal of Energy in Southern Africa*, Vol. 24, No. 4, pp. 63-70, Energy Research Centre.
- [5] Hefferman, J.E. , Stephenson, A.G. and Gilleland, E., (2012). *Ismev: an introduction to statistical modeling of extreme values*, *R package version*, Vol. 1.
- [6] Kielty, V. and Daynes, W. (2014). *Business Excellence*, [online], available: [//www.bus-ex.com/article/eskom-kusile-power-station-project](http://www.bus-ex.com/article/eskom-kusile-power-station-project), [09 May 2015].
- [7] Spedicato, G. A., Kang, T. S. , Thoralf, M. and Spedicato, M. G. A. and Rcpp, L. (2015). *R Package markovchain*.
- [8] Yin, G George and Zhang, Qing. (2006), *Discrete-time Markov chains: two-time-scale methods and applications*, vol. 55, Springer Science & Business Media.
- [9] C Sigauke.(2012), *Modelling electricity demand in South Africa*, PhD Thesis, University of the Free-State.