
Research article

Long-term peak electricity demand forecasting in South Africa: A quantile regression averaging approach

Norman Maswanganyi^{1,*}, Edmore Ranganai² and Caston Sigauke³

¹ Department of Statistics and Operations Research, University of Limpopo, Private Bag X1106, Sovenga, 0727, South Africa

² Department of Statistics, University of South Africa, Private Bag X6, Florida 1710, South Africa

³ Department of Statistics, University of Venda, Private Bag X5050, Thohoyandou, 0950, Limpopo, South Africa

* **Correspondence:** Email: nmaswanganyi72@gmail.com; Tel: +27152683680.

Abstract: Forecasting electricity demand in South Africa remains an increasingly national challenge as the government does not sufficiently take into account the impact of the electricity prices in their electricity demand forecast. Effective measures to rapidly reduce the demand of electricity are urgently needed to deal with future electricity prices and government policies uncertainties within the energy industry. Moreover, long-term peak electricity demand forecasting methods are needed to quantify the uncertainty of future electricity demand for better electricity security management. The prediction of long-term electricity demand assists decision makers in the electricity sector in planning for capacity generation. This paper presents an application of quantile regression averaging (QRA) approach using South African monthly and quarterly data ranging from January 2007 to December 2014. Variable selection is done in a comparative manner using ridge, least absolute shrinkage and selection operator (Lasso), cross validation (CV) and elastic net. We compare the forecasting accuracy of monthly peak electricity demand (MPED) and quarterly peak electricity demand (QPED) forecasting models using generalised additive models (GAMs) and QRA. The coefficient estimates for ridge, Lasso and elastic net are estimated and compared using MPED and QPED data.

Keywords: additive quantile regression; cubic smoothing splines; long term peak demand forecasting; penalised regression methods; quantile regression averaging model; temperature

1. Introduction

1.1. Context

In an effort to deal with long-term uncertainties within the energy sector, more accurate long-term peak electricity demand forecasting methods are made use of in decision making and planning purposes. Inaccurate forecasting may seriously result in poor decision making and bad policies establishment. For instance, the South African government, electricity consumers, producers and industries need accurate forecasts for planning future events to assist in long-term development plans for the country in order to minimise risks. The long-term electricity demand forecasting methods are dependent on the choice of variable selections and very importantly, assumptions made for the long run. Researchers have dealt with long-term electricity demand forecasting using different models. For example; Optimal Forecasting Quantile Regression (OFQR), Multiple Simple Linear Regression (MSLR), Artificial Neural Networks (ANNs), Quantile Regression (QR) and time series analysis models have been used for long-term electricity demand forecasting ([1–5] and [6] among others).

Short-term electricity demand forecasting has attracted substantial attention than medium and long-term forecasting [3]. However, QRA is a new approach and produces more accurate long-term point forecasts for electricity peak demand. The model also performs well in practice for both short and medium-term peak electricity demand. It also gives the smallest values based on prediction intervals (PIs). This has been confirmed in the following literature; for example, [5, 7] and [8] among others.

The popularity of long-term peak electricity demand forecasting is mainly motivated by the need to understand the various uncertainties associated with the decision making processes in the South African industries. According to [9], the development of renewable energy policy activities by the South African government is greatly influenced by factors such as the periodical occurrence of different policy crises as a result of economic instability. The key interest for this study is identifying the linear relationship between the response and predictor variables using GAMs, OLS and QR models. OLS estimates quantify the relationship around the mean of the distribution while QR quantifies the relationship across all the various quantiles. Moreover, QR is robust in handling extreme value points of the response variable. Secondly, the electricity demand in South Africa is higher in winter than in summer and influenced especially by irregular events like extreme weather conditions. Also, generally electricity demand is non-stationary and exhibits an upward trend due to mainly economic and population growth. The long-term electricity demand forecasting is essential to avoid the shortage of electricity that may hamper the South African economic growth. South Africa needs reliable forecasts of electricity demand to mitigate the risk of shortages in the future so as to guarantee sustainability.

1.2. Literature review

Although the short-term electricity demand is most widely applied, the long-term peak electricity demand is increasing in popularity with papers such as those by [4] who estimate long-term peak demand by comparing different techniques using a typical fast growing methodology with dynamic electricity demand characteristics and typical developing methodology. Empirical results show that the model is superior, although it uses single electricity demand observation and is capable of giving several electricity demand peak forecasts corresponding at different levels of maximum temperature

and some level of social activities. However, it is not easy to rely on a single methodology for determining the electricity demand of such a fast growing and changing system.

The work done by [3] describes in a comparative manner that electricity demand forecasting is divided into three groups, namely, short, medium and long-term electricity demand forecasting. Long-term electricity demand forecasting normally deals with long time horizons. These include one year to ten years and sometimes up to three decades ahead forecast. Furthermore, it is often characterized by large uncertainties created by these distant horizons.

On modelling and forecasting long-term daily peak electricity demand, applied partially linear additive quantile regression (PLAQR) model is applied by [10] using South African data ranging from January 2007 to December 2013. Some data obtained from the 28 South African weather stations divided into coastal and inland regions are used. Variable selection is carried out using Lasso. The empirical results are presented with a comparative analysis of two developed models namely; with pairwise interactions and without interactions. Based on the pinball loss function, the average loss suffered by PLAQR with pairwise interaction is less than that of PLAQR without interaction. Moreover, based on RMSE, MAE and MAPE, PLAQR with pairwise interaction is better than PLAQR without interaction. The empirical results have also shown the usefulness of PLAQR models.

In recent paper, the proposed approaches such as GAMs, additive quantile regression (AQR) with and without interactions and QRA models are used for short-term probabilistic hourly load forecasting in [8]. The authors consider using average pinball loss function for model comparisons and also constructed PI indices such as prediction interval with nominal confidence (PINC), prediction interval widths (PIWs), prediction interval normalised average widths (PINAWs) and prediction interval normalised average deviations (PINADs) for the comparative evaluation of three different models. In their discussion, they indicated that QRA model gives the superior results as it produces the most accurate forecasts compared to other models. It also produces the smallest PINAW and PINAD values.

The study performed by [11] utilizes QR model for peak load demand forecasting to create the estimates of daily peak load demand. The proposed method is used based on its ability to avoid underprediction of the peak load demand and risk of power blackouts. The empirical results show that constructing the quantile from 0.97 to 0.99 is sufficient to estimate the upper limit for the daily peak demand. The authors proposed that the QR model performances should also be compared in other empirical data (for example, economic and financial data). However, in contrast to our study, the authors did not utilize QRA and GAMs in their analysis. This paper considers QRA approach as a comprehensive forecasting solution for long-term peak electricity demand using monthly and quarterly data.

1.3. Contributions

Considering an overview on electricity demand forecasting in section 1.2, the major contributions of this study are summarised in this manner: (1) the study is concentrated on QRA methodology to forecast long-term electricity demand; (2) the developed method should be useful in producing improved forecasting results; (3) the study could answer many challenges and increase the potential to develop a strategy in generating capacity planning and system reliable assessment of dealing with long-term peak electricity demand in South Africa; (4) the long-term electricity demand forecasting solution is discussed and the best forecasting model for future peak electricity demand suggested and (5) the

peak electricity demand literature seems scarce in South Africa on the issue of long-term electricity demand and this closes a gap by introducing new methodology to forecast long-term peak electricity demand.

The study considers monthly and quarterly data for peak electricity demand forecasting using quantile regression methods based on the following arguments. Firstly, the paper evaluates accuracy measures of both OLS and QR models. A comparative analysis is done using generalised additive models. The final structure of the paper is organised as follows: In Section 2, we first present the generalised additive models, followed by quantile regression models and then later explores more on additive quantile regression and quantile regression averaging. In Section 3, we present variable selection methods using cross validation (CV), Least absolute shrinkage and selection operator (Lasso) and elastic net. In Section 4, we briefly summarise the empirical results as well as discussing further the general remarks of the results. The results in this section are found using the statistical package ‘r’. Lastly, in Section 5, we wrap up with the conclusion.

2. Models

Focusing on the main interest of this paper, we lay out the methodology for fitting the monthly and quarterly peak electricity demand with OLS, GAMs, QR and QRA. Then combine the forecasts using some forecast combination methods.

2.1. Generalized additive model

Generalized additive model (GAM) is developed by [12]. GAM is known to be a simple powerful technique that is easy to interpret, more flexible because the relationships between dependent and independent variables are not assumed to be linear and the regularization of predictor functions assist to avoid overfitting effect [13]. In generalized additive models the response variable is expressed as sum of smooth functions. GAMs capture the common non linear patterns which the linear regression technique fails to do. Let y_t denote hourly peak demand and $\mu_t = E(y_t)$, then the model can be written as:

$$g(\mu_t) = \mathbf{x}_t^* \Theta + f_1(x_{1t}) + f_2(x_{2t}) + f_3(x_{3t}, x_{4t}) + \dots, \quad (2.1)$$

where \mathbf{x}_t^* is a row of the model matrix for any strictly parametric model components, Θ is the corresponding parameter vector, f_i are smooth functions of the covariates, x_{it} and $f_1, f_2, f_3 \dots$ are estimated by maximizing a penalized log-likelihood function analogous to the penalized least squares criterion. The GAMs are built around penalised regression smoothers and one of them is called cubic spline smoother [14]. Cubic splines are smoothest interpolators that minimize the penalized least squares. However, they have many free parameters to be smoothed and this seems to be a wasteful exercise. We usually estimate $g(x_t)$ by minimizing

$$\sum_{t=1}^n \{y_t - f(x_t)\}^2 + \lambda \int f''(x)^2 dx, \quad (2.2)$$

where λ is a non-negative smoothing parameter. If λ tends to infinity, the penalty term becomes more important and forces $f''(x)$ to zero. The bigger the value of λ , the more the curves become smoother. In this paper, cubic smoothing splines as a solution to optimization problem is used.

2.2. Quantile regression model

QR was introduced by [15] and is described in detail in [16] as an extension of classical least squares estimation of conditional mean models to conditional quantile function. An updated QR theory captures this method as a particular center of the distribution, minimizing the weighted absolute sum of deviations [17]. The conditional quantile function is given by

$$Q_\tau(y_t|\mathbf{x}_t) = \mathbf{x}_t^T \boldsymbol{\beta}_\tau + \varepsilon_t, \quad (2.3)$$

where, $0 < \tau < 1$ and $Q_\tau(y_t|\mathbf{x}_t)$ denotes the conditional function for the τ^{th} quantile of the electricity demand. Given a set of training data $G = (y_t, \mathbf{x}_t^T)^T$ for $t = 1, \dots, G$, the parameter $\boldsymbol{\beta}_\tau$ can be estimated as

$$\widehat{\boldsymbol{\beta}}_\tau = \underset{\boldsymbol{\beta}_\tau}{\operatorname{argmin}} \frac{1}{G} \left\{ \tau \sum_{t: y_t \geq \mathbf{x}_t^T \boldsymbol{\beta}_\tau} |y_t - \mathbf{x}_t^T \boldsymbol{\beta}_\tau| + (1 - \tau) \sum_{t: y_t < \mathbf{x}_t^T \boldsymbol{\beta}_\tau} |y_t - \mathbf{x}_t^T \boldsymbol{\beta}_\tau| \right\}. \quad (2.4)$$

The parameter $\widehat{\boldsymbol{\beta}}_\tau$ in Eq 2.4 is usually estimated by using linear programming methods called least absolute deviation (LAD) regression. If $\tau = 0.5$, Eq 2.4 is now given by

$$\widehat{\boldsymbol{\beta}}_\tau = \frac{1}{G} \sum_{t=1}^G |y_t - \mathbf{x}_t^T \boldsymbol{\beta}_\tau|, \quad (2.5)$$

giving the L_1 (median regression) estimator. Eq 2.4 can also be written as

$$\widehat{\boldsymbol{\beta}}_\tau = \frac{1}{G} \sum_{t=1}^G \rho_\tau(y_t - \mathbf{x}_t^T \boldsymbol{\beta}_\tau) = \frac{1}{G} \sum_{t=1}^G (\tau - \mathbf{1}_{y_t - \mathbf{x}_t^T \boldsymbol{\beta}_\tau < 0}) (y_t - \mathbf{x}_t^T \boldsymbol{\beta}_\tau), \quad (2.6)$$

where ρ_τ is a check function such that $\rho_\tau(b) = \tau(b)$ if $b \geq 0$ and $\rho_\tau(b) = (\tau - 1)b$ if $b < 0$, $\mathbf{1}_B$ is the indicator function of the event B .

2.3. Additive quantile regression

The study done by [18] proposes the boosting AQR procedure for forecasting uncertainty using electricity smart meter data. The authors consider boosting procedure to estimate an AQR model for a set of quantiles of the future distribution. According to [18], the k^{th} -step ahead forecast for the condition at τ^{th} -quantile of the electricity demand in equation (2.3) can be estimated using the pinball loss function defined by

$$L(y_t, Q_\tau) = \begin{cases} \tau(y_t - \widehat{Q}_\tau) & \text{if } y_t \geq \widehat{Q}_\tau; \\ (1 - \tau)(\widehat{Q}_\tau - y_t) & \text{if } y_t < \widehat{Q}_\tau, \end{cases} \quad (2.7)$$

where \widehat{Q}_τ denotes the quantile forecast and y_t represents the actual value of peak electricity demand.

2.4. Quantile regression averaging

QRA is another recent forecast combination technique used to compute the prediction interval. Despite its popularity and simplicity, it is noted that combining forecasts has not been performed

widely in the area of long-term peak electricity demand forecasting using quarterly or monthly data. However, it has been performing well in the practice of electricity price forecasting (see, [7, 19], among others). According to [18] the QRA model is given by

$$y_t = g_k(\mathbf{x}_t) + \varepsilon_t, \quad (2.8)$$

where $\mathbf{x}_t = (\mathbf{y}_t, \mathbf{z}_t)$; \mathbf{z}_t is a vector of exogenous variables known at time t ; \mathbf{y}_t is a vector of past demand occurring prior to time t ; ε_t denotes the model error term with $E[\varepsilon_t] = 0$ and $E[\mathbf{x}_t \varepsilon_t] = 0$.

3. Monthly and quarterly demand models

3.1. Variable selection methods

The following section briefly discusses three variable selection methods for penalised regression which includes Lasso, CV and elastic net. Lasso guarantees both variable selection and shrinkage simultaneously. Unlike other penalised regression techniques, in Lasso some of the coefficients may be shrunken all the way to zero. The Lasso estimates for the OLS regression model is given by

$$\hat{\boldsymbol{\beta}} = \operatorname{argmin} \sum_{j=1}^p (\mathbf{y} - \mathbf{x}\boldsymbol{\beta}_j)^2 + \lambda_1 \sum_{j=1}^p |\boldsymbol{\beta}_j|, \quad (3.1)$$

where, $\mathbf{y} = (\mathbf{y}_1, \dots, \mathbf{y}_n^T$ is the response vector and $\mathbf{x}_j = (\mathbf{x}_{1j}, \dots, \mathbf{x}_{nj})^T$, $j = 1, \dots, p$ are the linear independent predictor, λ is the Lasso regularization parameter, $\sum_{j=1}^p |\boldsymbol{\beta}_j| \leq t$ is called Lasso penalty and t is the Lasso turning parameter [20]. If $p > n$, the Lasso selects at most n variables. There is however an adaptive Lasso method and this method modifies the Lasso penalty by applying weights to each parameter that forms the Lasso constraint. Elastic net is a regularization and variable selection method suggested by [21]. The method removes the limitation on the number of selected variables and also encourages grouping effect that Lasso fails to do. The elastic net estimator is given by

$$\hat{\boldsymbol{\beta}} = \operatorname{argmin} \sum_{j=1}^p (\mathbf{y} - \mathbf{x}\boldsymbol{\beta}_j)^2 + \lambda_1 \sum_{j=1}^p |\boldsymbol{\beta}_j| + \lambda_2 \sum_{j=1}^p \boldsymbol{\beta}_j^2, \quad (3.2)$$

where, $\sum_{j=1}^p |\boldsymbol{\beta}_j| \leq t_1$ and $\sum_{j=1}^p \boldsymbol{\beta}_j^2 < t_2$ are elastic net penalties, t_1 and t_2 are elastic net turning parameters. Cross validation is another most commonly used selection criterion method that uses part of the training data set to fit the model and the remaining part estimate the prediction error. The data consist of 28 South African weather stations divided into interior and coastal regions with monthly and quarterly peak electricity demand (MPED and QPED) data, temperature, lagged demand and calendar effects. The MPED and QPED are used as the dependent variables while temperature, lagged demand and calendar effects are used as predictor variables. For MPED, the temperature effects for coastal are AMTC, maxTC and minTC defined as average monthly coastal, average maximum and minimum coastal temperatures respectively. The effects for interior include AMTI, maxTI and minTI defined as average monthly interior, average maximum and minimum interior temperatures respectively.

Likewise, for QPED the temperature effects for coastal are AQTC, maxTC and minTC defined as average quarterly coastal, average maximum and minimum coastal temperatures respectively. The

effects for quarterly interior include AQT_I, max_{TI} and min_{TI} defined as average quarterly interior, average maximum and minimum interior temperatures respectively. The temperature effects for both monthly and quarterly peak electricity demand coastal and interior are given by average minimum of coastal and interior temperatures (AminTCI), average of average coastal and interior temperatures (AATCI), average maximum of coastal and interior temperatures (AmaxTCI), difference between average maximum of coastal and interior temperatures (DmaxTCI), difference between average minimum of coastal and interior temperatures (DminTCI) and the difference between average of average monthly coastal and interior temperatures (DAAMTCI) and the difference between average of average quarterly coastal and interior temperatures (DAAQT_{CI}). Furthermore, day type (day of the week), day before holiday (DBH), day after holiday (DAH) and day holiday (DH) are defined as calendar effects.

3.1.1. Monthly model

In forecasting electricity demand at inland and coastal regions, we use long-term peak electricity demand models. The QR models are used for both monthly and quarterly data. This study will explore the inclusion of monthly temperature covariates that will also be included in the peak electricity demand parameters through heating and cooling degree days which can be calculated as:

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

and

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

The total monthly heating degree-days (MHDD_{*t*}) and the total monthly cooling degree-days (MCDD_{*t*}) are calculated as:

$$\text{MHDD}_t = \max \left[\sum_{j=1}^m \text{HDD}_{t,j} - \sum_{j=1}^m \text{CDD}_{t,j}, 0 \right]$$

and

$$\text{MCDD}_t = \max \left[\sum_{j=1}^m \text{CDD}_{t,j} - \sum_{j=1}^m \text{HDD}_{t,j}, 0 \right]$$

which reduces to

$$\text{MHDD}_t = \max \left[\sum_{j=1}^m (\max(T_{\text{ref}} - T_{t,j}, 0)) - \sum_{j=1}^m (\max(T_{t,j} - T_{\text{ref}}, 0)), 0 \right]$$

and

$$\text{MCDD}_t = \max \left[\sum_{j=1}^m (\max(T_{t,j} - T_{\text{ref}}, 0)) - \sum_{j=1}^m (\max(T_{\text{ref}} - T_{t,j}, 0)), 0 \right],$$

respectively, where m is the number of days in month t , T_t is average daily temperature on day t and T_{ref} is the reference temperature which separates cold temperatures from hot temperatures. Based on the description of variable selection given in section 3.1, the following MPED models for electricity demand are proposed:

Model 1:

Linear quantile regression (LQR) without interactions model is given as

$$\begin{aligned} Q_\tau(\text{MPED}|\mathbf{x}_t) = & \alpha_0 + \alpha_1(DH) + \alpha_2(DAH) + \alpha_3(DBH) + \alpha_4(\text{Daytype}) \\ & + \alpha_5(\text{CDDAminTCI}) + \alpha_6(\text{CDDAmaxTCI}) + \alpha_7(\text{HDDAAMTCI}) \\ & + \alpha_8(\text{HDDAminTCI}) + \alpha_9(\text{noltrend}) + \varepsilon_{t,\tau}, \end{aligned} \quad (3.3)$$

where $\alpha_0, \alpha_1, \dots, \alpha_9$ are constants, CDDAminTCI and CDDAmaxTCI are cooling degree days average minimum and maximum coastal and interior temperatures, HDDAAMTCI is monthly heating degree days average of average coastal and interior temperature, HDDAminTCI is heating degree days average minimum coastal and interior temperature, noltrend is a non-linear trend, $\varepsilon_{t,\tau}$ is residual error.

Model 2: LQR with interactions model is given by

$$\begin{aligned} Q_\tau(\text{MPED}|\mathbf{x}_t) = & \alpha_0 + \alpha_1(DAH) + \alpha_2(DBH) + \alpha_3(\text{Daytype}) \\ & + \alpha_4(\text{CDDAminTCI} * \text{Daytype}) + \alpha_5(\text{CDDAmaxTCI}) \\ & + \alpha_6(\text{HDDAAMTCI}) + \alpha_7(\text{HDDAminTCI}) + \alpha_8(\text{noltrend}) \\ & + \varepsilon_{t,\tau}. \end{aligned} \quad (3.4)$$

Model 3: QRA is given in Eq 3.5

$$\text{MPED} = \alpha_0 + \alpha_1(\text{fOLS}) + \alpha_2(\text{fGAM}) + \alpha_3(\text{fQR}) + \varepsilon_{t,\tau}, \quad (3.5)$$

where fOLS, fGAM and fQR are ordinary least squares, generalised additive model and quantile regression forecasts respectively.

3.1.2. Quarterly model

Similarly we can derive the formulae for calculating the total quarterly heating and cooling degree-days $QHDD_t$ and $QCDD_t$ respectively as in the monthly formulae. The quarterly temperature covariate will be included in the long-term peak electricity demand parameters through heating and cooling degree-days. Based on the description of variable selection given in section 3.1, the following QPED models for electricity demand are proposed:

Model 4: LQR without interactions model is given as

$$\begin{aligned} Q_\tau(\text{QPED}|\mathbf{x}_t) = & \alpha_0 + \alpha_1(DH) + \alpha_2(DAH) + \alpha_3(\text{Daytype}) + \alpha_4(\text{CDDAminTCI}) \\ & + \alpha_5(\text{CDDAmaxTCI}) + \alpha_6(\text{CDDAAQTTCI}) + \alpha_7(\text{noltrend}) \\ & + \varepsilon_{t,\tau}, \end{aligned} \quad (3.6)$$

where HDDAAQTCI is quarterly heating degree days average of average coastal and interior temperature.

Model 5: LQR with interactions model is given as

$$\begin{aligned} Q_{\tau}(QPED|\mathbf{x}_t) = & \alpha_0 + \alpha_1(DH) + \alpha_2(DAH) + \alpha_3(Daytype) + \alpha_4(CDDAminTCI) \\ & + \alpha_5(CDDAmaxTCI) + \alpha_6(HDDAAQTCI * Dytype) \\ & + \alpha_7(noltrend) + \varepsilon_{t,\tau}. \end{aligned} \quad (3.7)$$

Model 6: QRA is given in equation (3.8)

$$QPED = \alpha_0 + \alpha_1(fOLS) + \alpha_2(fGAM) + \alpha_3(fQR) + \varepsilon_{t,\tau}. \quad (3.8)$$

3.2. Parameter estimation

The QR model defined in Eq 2.3 will be helpful in this analysis. The parameter for OLS over $\boldsymbol{\beta}$ is estimated by minimizing:

$$\sum_{t=1}^n r(y_t - \mathbf{x}_t^T \boldsymbol{\beta}) = \sum_{t=1}^n (y_t - \mathbf{x}_t^T \boldsymbol{\beta})^2, \quad (3.9)$$

where r is the quadratic loss function and the parameter for MLE over $\boldsymbol{\beta}$ based on the sample $\{\mathbf{x}_t, y_t\}_{t=1}^n$ is calculated by maximizing

$$L(\boldsymbol{\beta}) \propto \exp\left\{-\frac{1}{2\sigma^2} \sum_{t=1}^n (y_t - \mathbf{x}_t^T \boldsymbol{\beta})^2\right\}. \quad (3.10)$$

3.3. Benchmark models

3.3.1. Stochastic gradient boosting

Stochastic gradient boosting (SGB) is a modification of Gradient boosting (GB) which is a machine learning technique. It fits an additive model in a stage-wise way. The additive model can take the form given in Eq 3.11 ([22]).

$$f(x) = \sum_{m=1}^M \beta_m b(x; \gamma_m), \quad (3.11)$$

where $b(x; \gamma_m) \in \mathbb{R}$ are functions of x which are characterised by the expansion parameters γ_m, β_m . The parameters β_m and γ_m are fitted in a stage-wise way, a process which slows down over-fitting ([22]). With SGB, a random sample of the training data set is taken without replacement. A more detailed discussion of this is found in [23].

3.3.2. Support vector regression

Support vector regression (SVR) which is based on support vector machines (SVM) uses different kernel functions which map low dimensional data to high dimensional space. SVR was introduced by [24] and estimates a regression function of the form given in Eq (3.12).

$$f(x) = \mathbf{x}^T \mathbf{w} + b, \quad (3.12)$$

where b is a scalar and ω is a vector of weights. Eq (3.12) can be reformulated as a quadratic programming problem (QPP) by introducing an ε -insensitive loss function as given in Eq (3.13) ([24]).

$$\begin{aligned} \min_{(\omega, b, \xi_1, \xi_2)} \quad & \frac{1}{2} \omega^T \omega + C(e^T \xi_1 + e^T \xi_2) \\ \text{s.t.} \quad & y_i - x_i^T \omega - b \leq \varepsilon + \xi_1 \\ & x_i^T \omega + b - y_i \leq \varepsilon + \xi_2 \\ \text{and} \quad & \xi_{1i}, \xi_{2i} \geq 0, i = 1, \dots, m, \end{aligned} \quad (3.13)$$

where $C > 0, \varepsilon > 0$ are input parameters, $\xi_1 = (\xi_{11}, \dots, \xi_{1m})^T$ and $\xi_2 = (\xi_{21}, \dots, \xi_{2m})$ are slack variables.

4. Empirical results and discussion

4.1. Exploratory data analysis

In this study, the South African monthly and quarterly electricity demand data from Eskom ranging from January 2000 to March 2014 are used. The aggregate energy data with 5204 daily observations are used to create 171 monthly and 57 quarterly data for analysis. Both training data (Jan 2000–Dec 2008) and testing (Jan 2009–Mar 2014) are used for MPED and QPED modelling. The temperature data for coastal and interior regions from 28 South African weather stations are considered. The study uses the most common variables that have been included in the electricity demand models which include, average monthly and quarterly electricity demand, non linear trend and temperature.

It further analyses three different benchmark models which are GAMs, SVR and SGB. The use of at least two benchmark models is consistent with current trends in forecasting. Current trends in forecasting use a variety of models, statistical models, including comparative analysis with machine learning models. In this study we use one statistical learning benchmark model which is the generalized additive model and two machine learning models, stochastic gradient boosting and support vector regression. Using at least two benchmark models helps us to see how good our model is against well-known methods which produce accurate forecasts. See for instance, [25].

Figures 1 and 2 show monthly and quarterly peak electricity demand (MPED and QPED) plots for the period 2000 to 2014 with density, q-q and box plots respectively. Both MPED and QPED plots show upward trends with strong seasonality. The presence of increasing trends are indicated by the overall upward drifts in the time series. The seasonal variations produce the regular fluctuations from year to year but appear similar across months and quarters.

Table 1 shows the summary statistics for both monthly and quarterly electricity demand data ranging from 1 January 2000 to 31 March 2014. The minimum and maximum values are compared and also used to assess the spread behaviour of both quarterly and monthly electricity demand. The minimum values of MPED and QPED from January–March 2000 and on Monday, January 2000 are 24083 and 24760 MW respectively. Moreover, the maximum value of both MPED and QPED from April–June 2007 and also on Thursday, 24 May 2007 is 37158 MW. The kurtosis values of MPED and QPED are -0.31 and -0.21 respectively. The values indicate that both MPED and QPED data do not perfectly follow the normal distribution. Furthermore, negative kurtosis shows that the distribution has lighter tails and flatter peaks compared to the normal distribution. Table 1 also reveals that both monthly and quarterly electricity demand data are left skewed.

Tables 2 and 3 present the summary results of OLS forecast and QR, QRA, GAMs forecasts for both MPED and QPED models based on RMSE, MAE and MAPE respectively. In Table 2, M1 and M2 represent OLS forecast without and with interactions, respectively, for MPED model, while M4 and M5 represent OLS forecast without and with interactions, respectively, for QPED model.

Likewise, in Table 3, M1 and M2 represent linear QR forecast without and with interactions, respectively, for MPED model while M4 and M5 represent linear QR forecast without and with interactions, respectively, for QPED model, M3 and M6 are QRA forecast for MPED and QPED model respectively, M7 and M8 are GAMs forecast for both MPED and QPED models. The values of RMSE = 79.95 MW, MAE = 72.45 MW and MAPE = 0.22 MW for M4 in Table 2 are less compared to that of M1, M2 and M5 models. Likewise, the values of RMSE = 76.55 MW, MAE = 51.61 MW and MAPE = 0.16 MW for M6 in Table 3 are less compared to that of M1, M2, M3, M4, M5, M7 and M8 models. Moreover, the values of RMSE, MAE and MAPE for M6 (QRA for QPED) are less compared to that of M4 (OLS for MPED). The results of RMSE, MAE and MAPE for M6 (QRA for QPED) also confirm that the model is doing well compared to that of M3 (QRA for MPED). In addition, the values of RMSE, MAE and MAPE for benchmark (SVR and SGB) models in Table 4 are larger than that of M6 model.

Figures 8 and 10 show the results of cross validation estimates of the mean squared prediction error for ridge, Lasso and elastic net models in dependence on the regularization parameter (log of lamda). The numbers on top of each figures shows the number of non-zero regression coefficients. The mean squared prediction errors suggest the smaller values of the parameter that shrinks the coefficients to optimal. Figures 9 and 11 show the coefficient paths for ridge, Lasso and elastic net model independence on log of lamda for $\alpha = 0, 1$ and 0.5 respectively. The analysis was also done repeated using different values of α .

The comparison of MPED cross validation estimates of the mean squared prediction error for ridge, lasso and elastic net in Figure 8 shows that the variable selection method called Lasso is superior compared to other methods. There are also noticeable differences in cross validation estimates for all the methods for which Lasso gives the best estimates. Using Lasso and elastic net regression model for model selection, only one non-zero coefficient shows that the function has chosen the second vertical line on the cross validation. This shows that the model might be doing a good job. In Figure 9, all coefficients of ridge regression model are essentially zero when the log of lamda is 15. However, as we relax lamda, the coefficients moves away from zero in a nice smooth way. In fitting Lasso, a lots of the r squared what are explained as quite heavily shrunk coefficients but towards the end coefficients increases (grow very large). This may suggest that at the end of the path, the model might be overfitted. Furthermore, elastic net regression model looks like Lasso with slight differences and this confirms that it is an extension of Lasso proposed by [20]. However, elastic net does not perform satisfactory because there are two shrinkage procedures such as ridge and Lasso in it. Double shrinkage might introduce unnecessary bias. Likewise, Figure 10 shows the QPED model selection of ridge, Lasso and elastic net where the dotted red lines show the cross validation curve. The two dashed lines in Figure 10 also show log of lambda values that gives the minimum mean square errors for cross validation (left dashed line) and that is within one standard error(right dashed line). Figure 11 shows all coefficients of ridge regression model to be zeros when the value of lambda is 14 or more. It can be seen from figure 11 that Lasso regression shrinks the regression coefficients to zero as lambda increases. For instance, when log of lambda = 2 there are 6 non zero coefficients and when log of lambda = 0 all coefficients are

zero.

Table 1. Summary statistics for monthly and quarterly electricity demand data (1 January 200–31 March 2014).

	Mean	Median	Min	Max	Kurtosis	Skewness	St.Dev
Monthly data	31560	31822	24083	37158	-0.31	-0.42	30.44
Quarterly data	32452	32675	24760	37158	-0.21	-0.54	30.71

Table 2. OLS forecast model comparisons.

	M1	M2	M4	M5
RMSE	327.91	338.12	79.95	172.17
MAE	252.06	262.83	72.45	128.41
MAPE	0.77	0.80	0.22	0.37

Table 3. QR ,GAM and QRA forecast models comparison.

	M1	M2	M3	M4	M5	M6	M7	M8
RMSE	324.89	345.38	315.46	87.92	87.97	76.55	315.96	165.94
MAE	249.31	273.41	231.70	73.55	73.64	51.61	235.01	123.57
MAPE	0.76	0.83	0.71	0.22	0.22	0.16	0.72	0.36

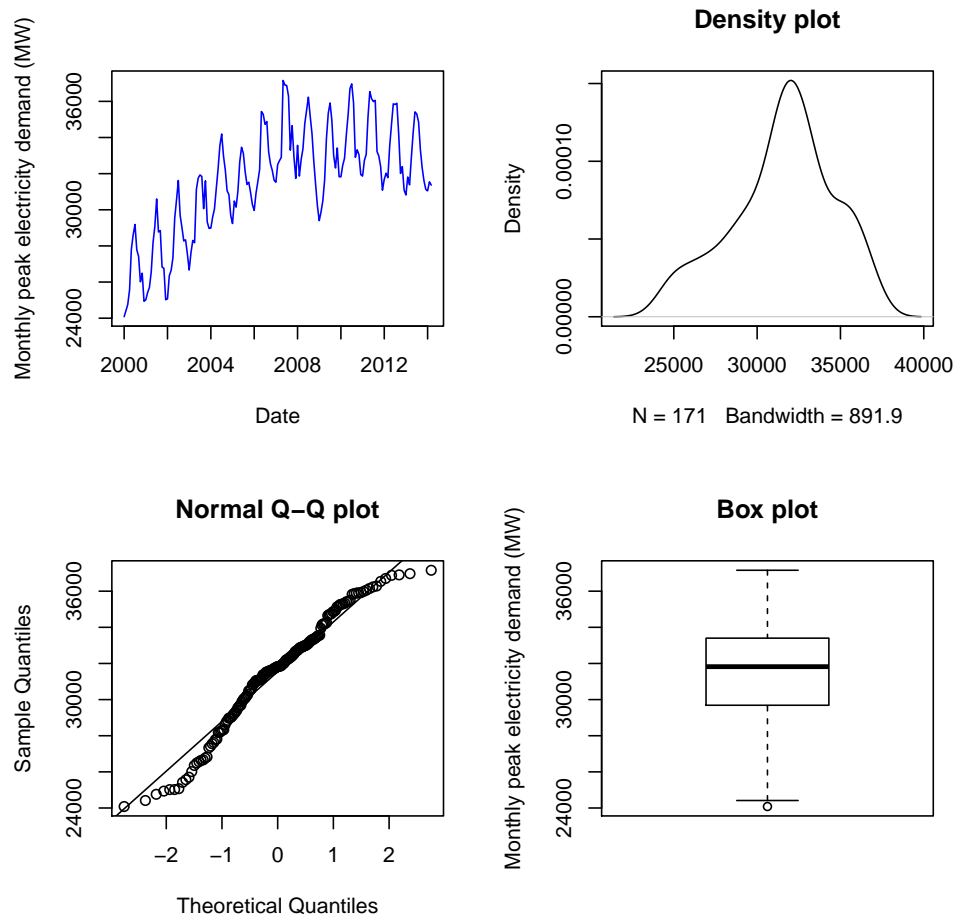


Figure 1. Monthly peak electricity demand(top left), Density plot (top right),Q-Q plot (bottom left) and Box plot (bottom right).

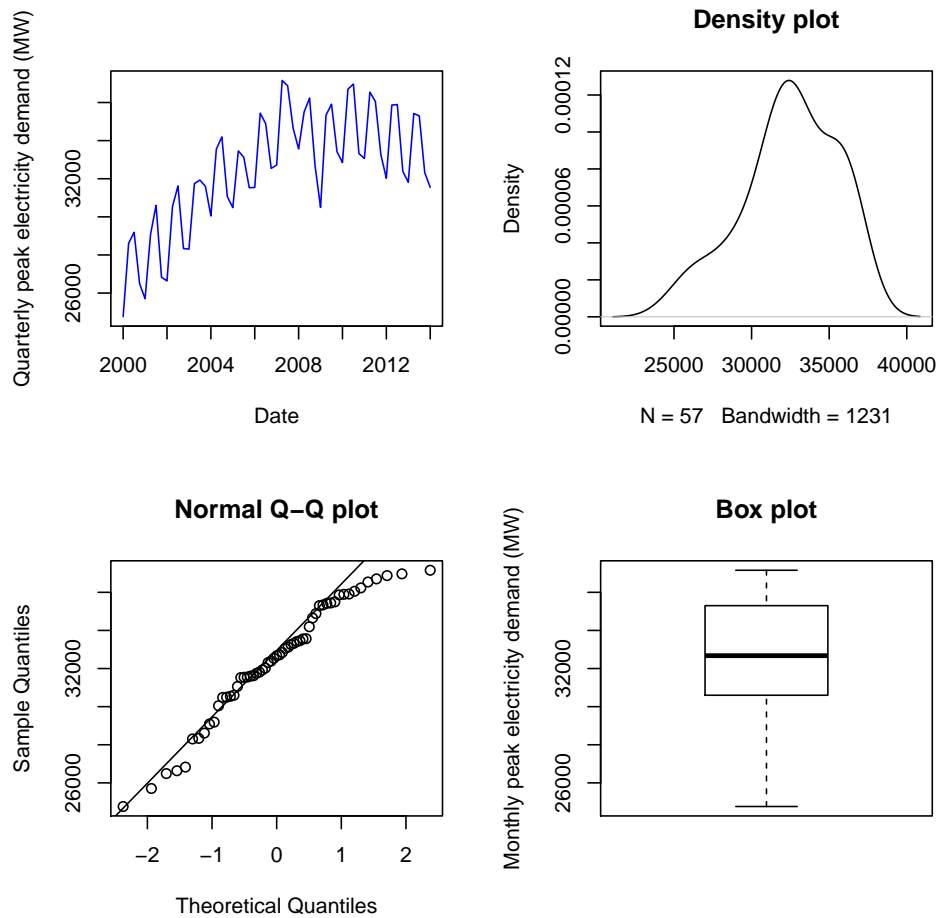


Figure 2. Quarterly peak electricity demand(top left), Density plot (top right), Q-Q plot (bottom left) and Box plot (bottom right).

4.2. Benchmark models

GAMs are considered as simple and powerful technique because they provide more flexible predictor functions that uncover hidden patterns in the data and regularize the predictor functions. Moreover, this is one of the techniques that is mostly indispensable in the analysis of long-term peak electricity demand forecasting. This machine learning technique is considered as one of benchmark models in MPED and QPED analysis.

Model 7: Monthly data model GAM is given as

$$\begin{aligned} MPED = & \alpha_0 + s(\text{Daytype}) + \alpha_1(DH) + \alpha_2(CDDAminTCI) + \alpha_3(CDDAmaxTCI) \\ & + \alpha_4(HDDAAMTCI) + \alpha_5(HDDAminTCI) + s(\text{noltrend}) + \varepsilon_{t,\tau}, \end{aligned} \quad (4.1)$$

Model 8: Quarterly data model GAM is given as

$$\begin{aligned} QPED = & \alpha_0 + \alpha_1(DH) + \alpha_2(DAH) + s(\text{Daytype}) \\ & + \alpha_3(CDDAminTCI) + \alpha_4(CDDAmaxTCI) + \alpha_5(HDDAAQTCI * \text{Daytype}) \\ & + s(\text{noltrend}) + \varepsilon_{t,\tau}, \end{aligned} \quad (4.2)$$

where s represents smoothing.

Table 4. Model comparisons: Monthly peak electricity demand.

	SVR (benchmark)	SGB(benchmark)
RMSE	398.35	728.19
MAE	325.73	558.98
MAPE	0.98	1.66

4.3. Discussion of results from models

Figures 1 and 2 display monthly and quarterly peak electricity demand plots for the period 2000 to 2014 respectively. Hence, Figure 3 shows monthly and quarterly average temperature plots. It also shows strong seasonal effects. The smallest values of RMSE, MAE and MAPE in OLS forecast model (M4) for both MPED and QPED are 79.95 MW, 72.45 MW and 0.22 MW respectively. Likewise, the smallest RMSE, MAE and MAPE in QRA model (M6) are 76.5 MW, 51.61 MW and 0.16 MW respectively.

Table 4 shows the error levels (RMSE = 398.35 MW, MAE = 325.73 MW, MAPE = 0.98 MW) of SVR to be lower than the error level (RMSE = 728.19 MW, MAE = 558.98 MW, MAPE = 166 MW) of SGB model. Figures 6 and 7 show the SVR and SGB benchmark model forecasts for MPED. Clearly, we can see that SVR forecasts in Figure 6 fit very well compared to SGB forecasts in Figure 7. Figures 4 and 5 show the actuals of both MPED and QPED with quantile regression forecast (fQR), generalised additive model forecast (fGAM) and quantile regression averaging forecast (fQRA) respectively. Figure 5 shows fQRA fits very well in QPED data.

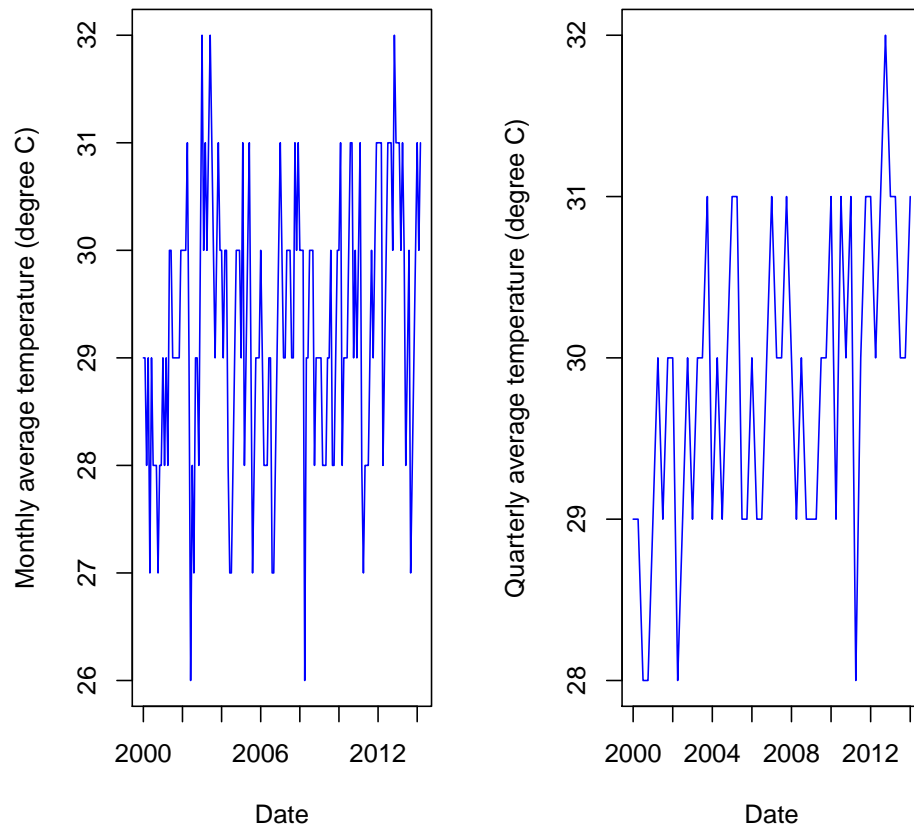


Figure 3. Monthly average temperature plot(left side) and Quarterly average temperature plot(on the right).

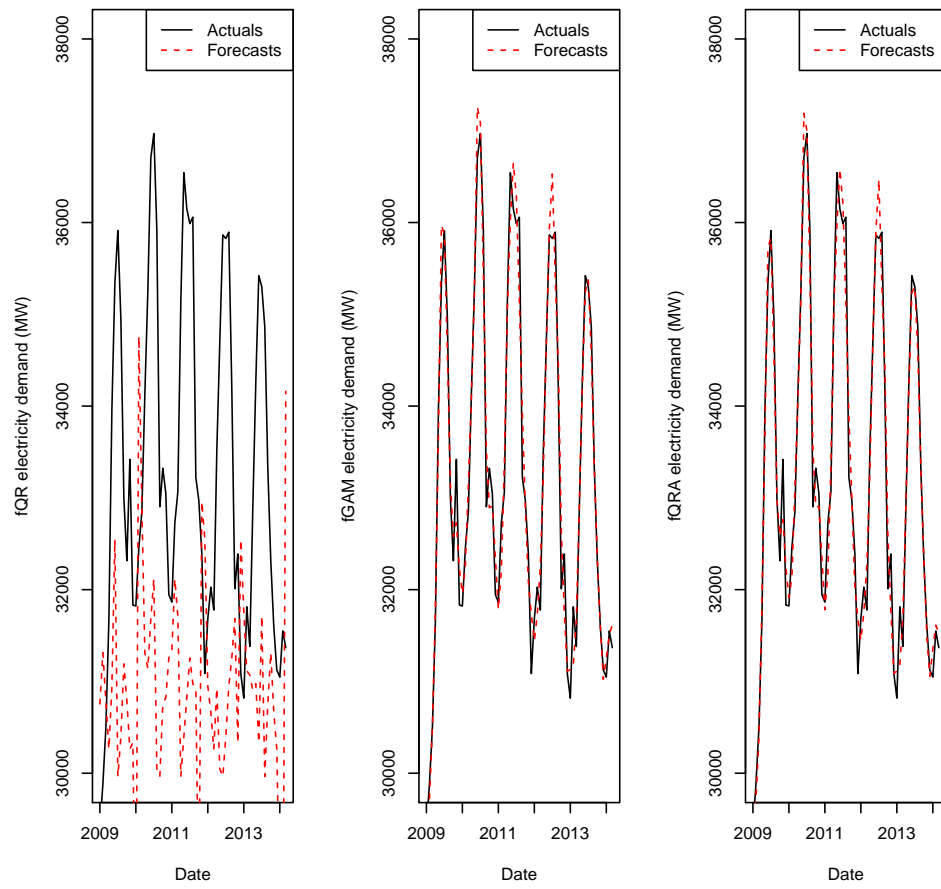


Figure 4. Actuals and forecasts for MPED.

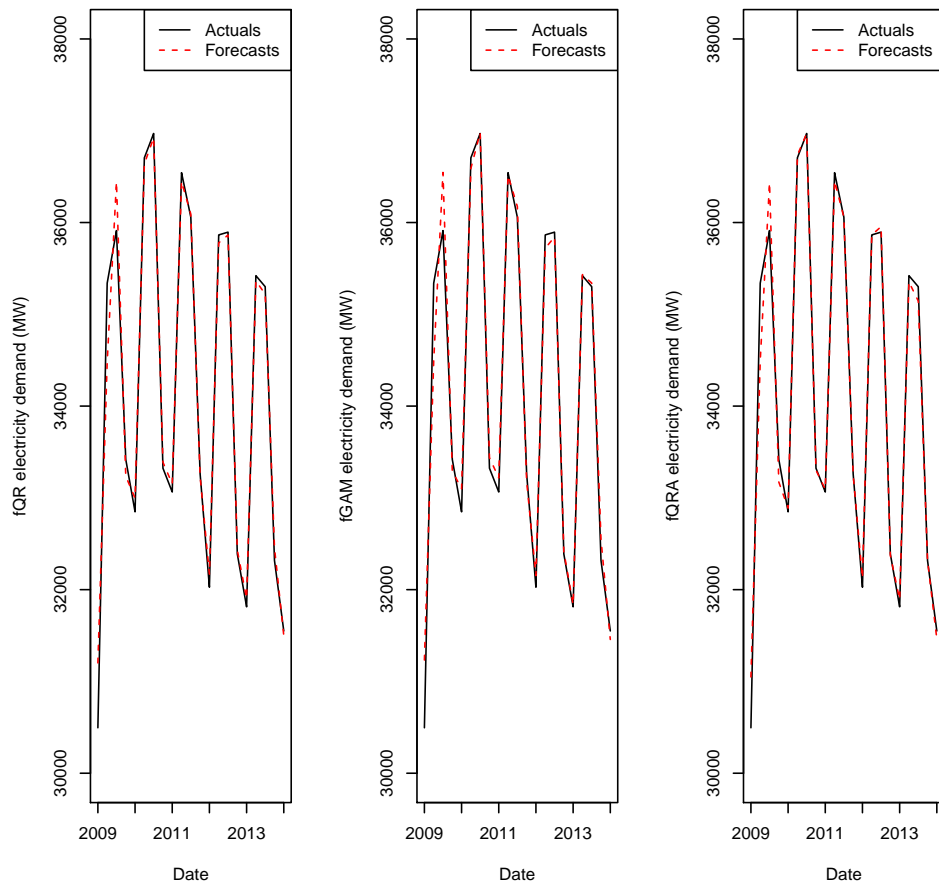


Figure 5. Actuals and forecasts for QPED.

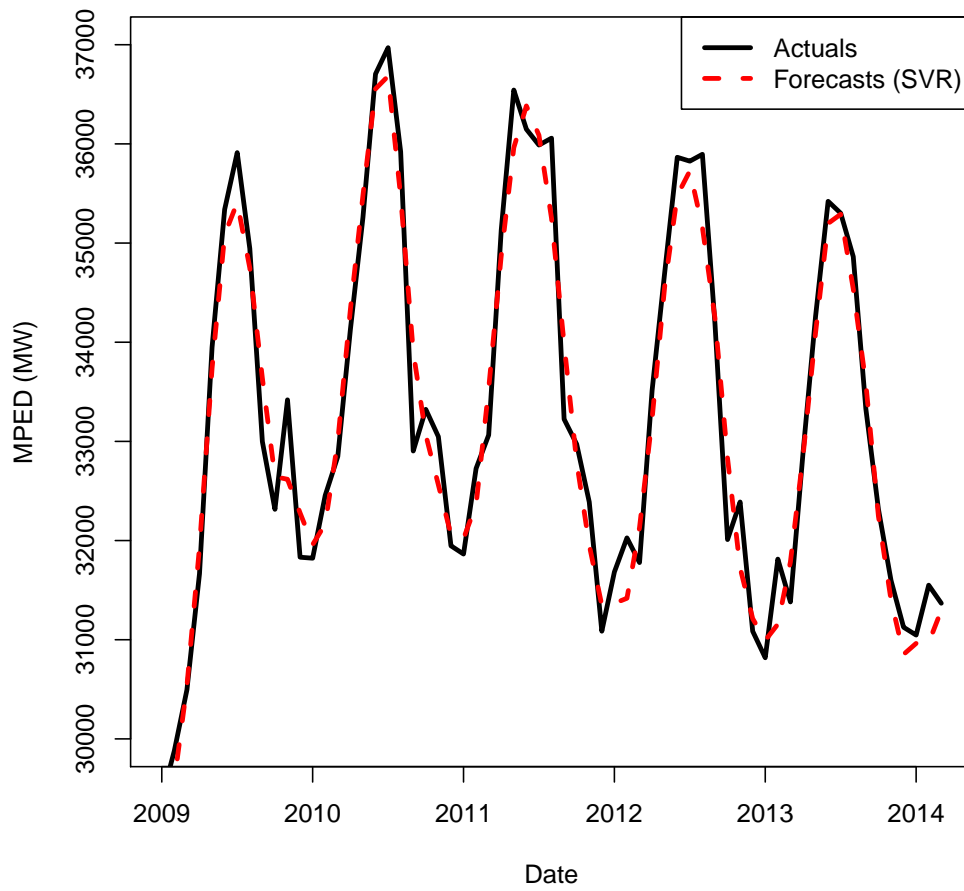


Figure 6. SVR forecasts for MPED.

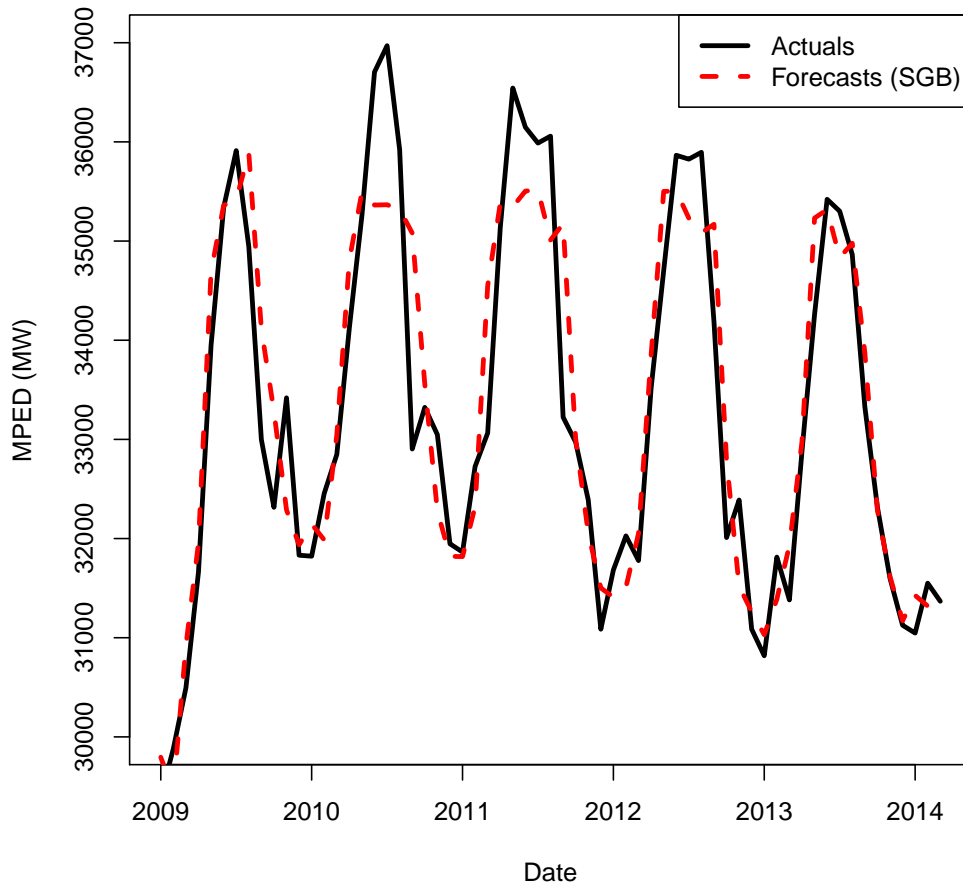


Figure 7. SGB forecasts for MPED.

5. Conclusions

In this study the long-term monthly and quarterly peaks electricity demand forecasting using South African data have been presented and therefore the policies to control the supply of electricity are required to make sure that the electricity demand is met to support the South African community. By using monthly and quarterly data and then later applying QR models, the more precise long-term forecasting models are developed. The QR models are becoming more attractive to long-term electricity demand modelling this days as it enables us to construct new quantile functions easily. We believe this study will make important contribution in all research application areas more especially in short, medium or long-term peak electricity demand methodology. QRA model in QPED has the smallest RMSE, MAE and MAPE compared to other models. The Lasso model gives best estimates compared to other variable selection methods. The empirical results of proposed models are suitable not just for long-term electricity demand forecasting, but also for testing policies geared towards the expansion of the electricity infrastructure that should be intensified in South Africa in order to cope

with the increasing demand exercised by the country’s economic growth and rapid industrialisation programme. QRA is a new technique and performed well in QPED data for long-term electricity demand and it did not perform bad in MPED data also. However, it is not advisable to rely only on a single approach for determining the electricity demand in such a fast growing and changing system of South African economy.

Acknowledgements

The authors are grateful to the National Research Foundation of South Africa for funding this research, to Eskom, South Africa’s power utility company for providing the data, University of Limpopo for their resources and to the numerous people for helpful comments on this paper.

Appendix

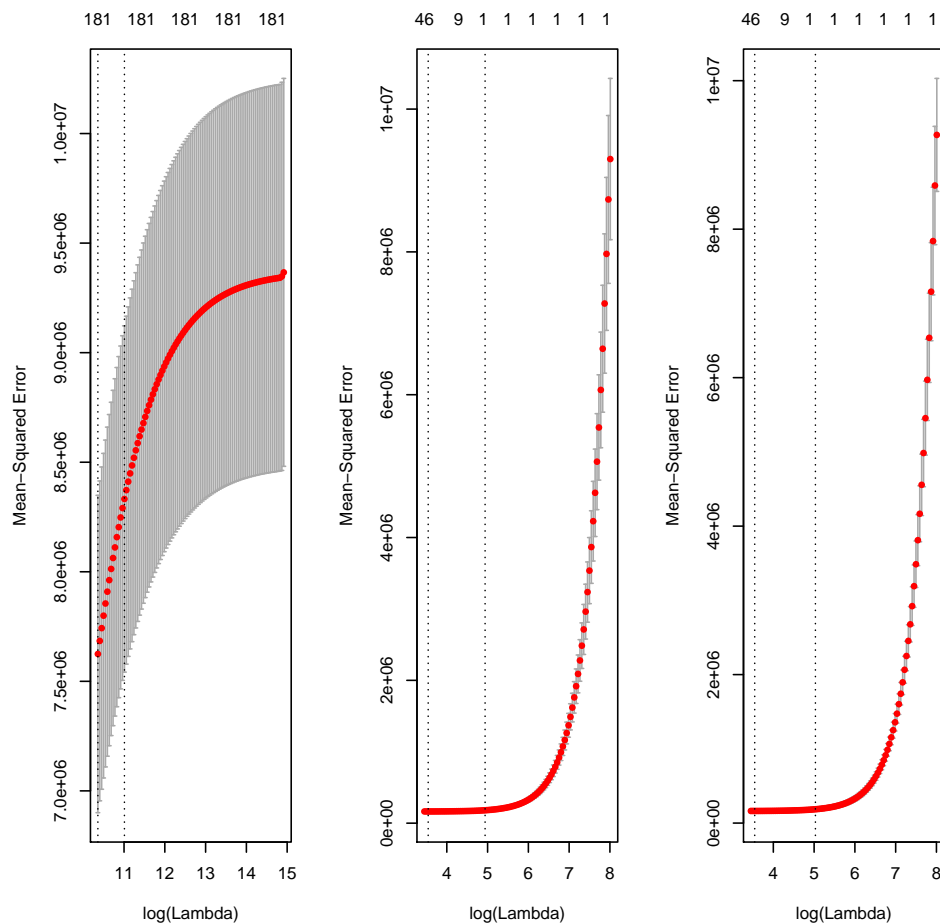


Figure 8. MPED cross validation estimates of the mean squared prediction error for ridge(left side), Lasso (center) and elastic net (on the right) as the function of $\log \lambda$.

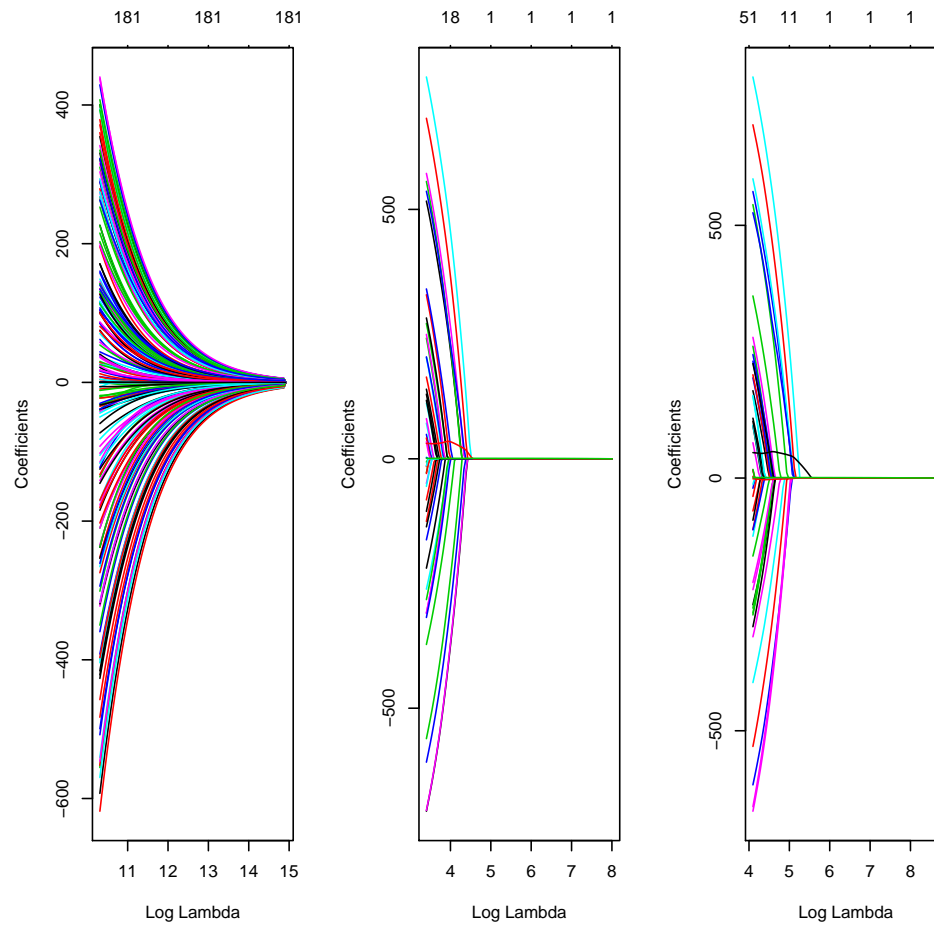


Figure 9. Coefficient estimates for ridge(left side), Lasso(centre) and elastic net (right side) for MPED data plotted versus $\log \lambda$.

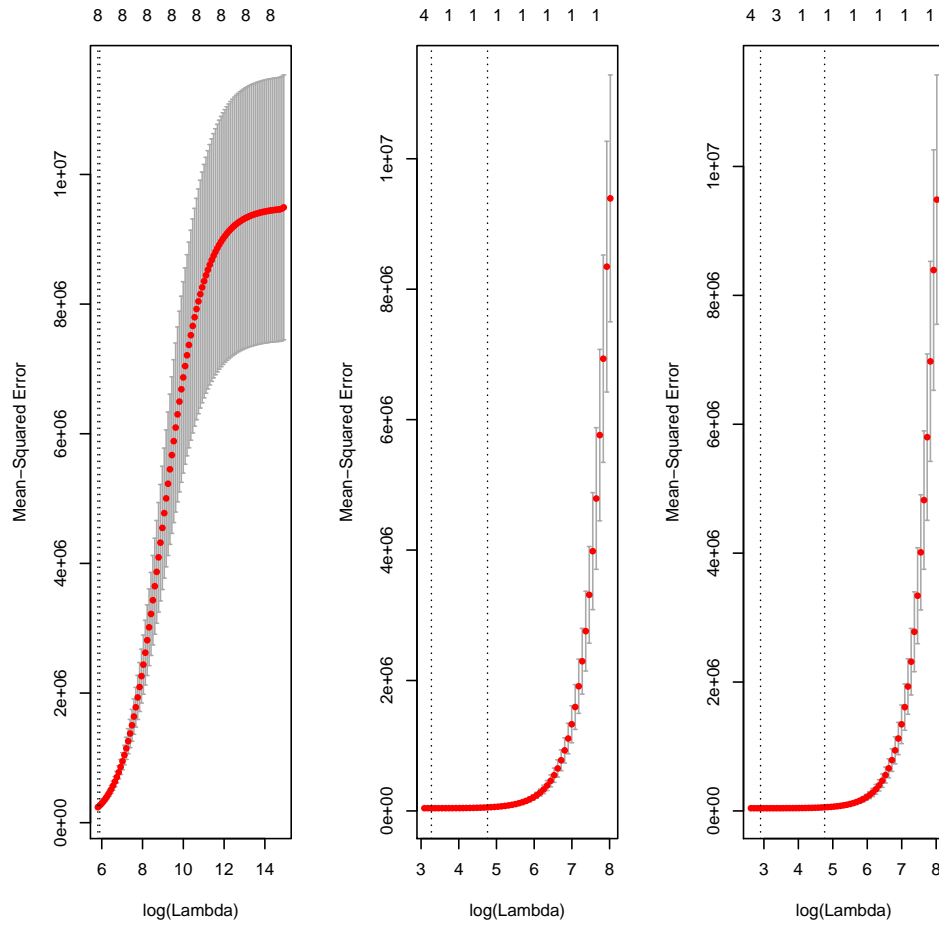


Figure 10. QPED cross validation estimates of the mean squared prediction error for ridge(left side), Lasso (center) and elastic net (on the right) as the function of λ .

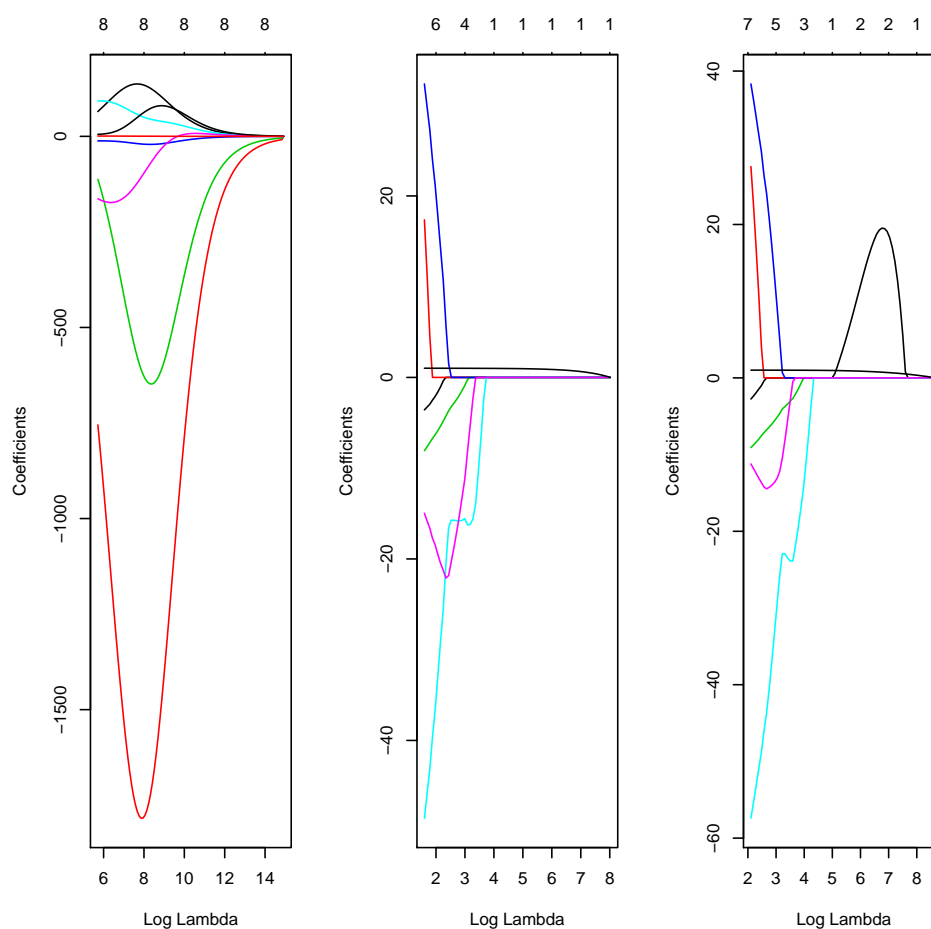


Figure 11. Coefficient estimates for ridge(left side), Lasso(centre) and elastic net (right side) for QPED data plotted versus $\log \lambda$.

Conflict of interest

The authors declare no conflict of interest.

References

1. Al-Hamadi HM, Soliman SA (2005) Long-term/mid-term electric load forecasting based on short-term correlation and annual growth. *Electr Power Syst Res* 74: 353–361.
2. Al-Saba T, El-Amin I (1999) Artificial neural networks as applied to long-term demand forecasting. *Artif Intell Eng* 13: 189–197.
3. Hyndman RJ, Fan S (2010) Density forecasting for long-term peak electricity demand. *IEEE Trans Power Syst* 25: 1142–1153.
4. Kandil MS, El-Debeiky SM, Hasanien NE (2002) Long-term forecasting for fast Developing Utility Using a knowledge-Based Expert System. *IEEE Trans Power Syst* 17: 491–496.

5. Mokilane P, Galpin J, Sarma Yadavalli VS, et al. (2018) Density forecasting for long-term electricity demand in South Africa using quantile regression. *South African J Econ Manage Sci* 21: 1–14.
6. Ringwood JV, Bofelli D, Murray FT (2001) Forecasting electricity demand on short, medium and long time scales using neural networks. *J Intell Rob Syst* 31: 129–147.
7. Nowotarski J, Weron R (2015) Computing electricity spot price prediction intervals using quantile regression and forecast averaging. *Comput Stat* 30: 791–803.
8. Sigauke C, Nemukula M, Maposa D (2018) Probabilistic hourly load forecasting using additive quantile regression models. *Energies* 11: 2208.
9. Marquard A (2006) The origins and development of South African energy policy, PhD thesis, University of Cape Town.
10. Maswanganyi N, Sigauke C, Ranganai E (2017) Peak electricity demand forecasting using partially linear additive quantile regression models. *Annual Proceedings of the South African Statistical Association Conference 2017*: 25–32.
11. Elamin N, Fukushige M (2018) Quantile regression model for peak load demand forecasting with approximation by triangular distribution to avoid blackouts. *Int J Energy Econ Policy* 8: 119–124.
12. Hastie T, Tibshirani R, et al. (2000) Bayesian backfitting (with comments and a rejoinder by the authors). *Stat Sci* 15: 196–223.
13. Larsen K (2016) Gam: The predictive modeling silver bullet. *Multithreaded. Stitch Fix* 30.
14. Wood SN, Augustin NH (2002) GAMs with integrated model selection using penalized regression splines and applications to environmental modelling. *Ecol Modell* 157: 157–177.
15. Koenker R, Bassett G (1978) Regression quantiles. *Econometrica: J Econ Soc* 33–50.
16. Koenker R (2005) *Quantile regression*. Cambridge university press, 38.
17. Hao L, Naiman DQ (2007) *Quantile regression*, Sage 149.
18. Taieb SB, Huser R, Hyndman RJ, et al. (2016) Forecasting uncertainty in electricity smart meter data by boosting additive quantile regression. *IEEE Trans Smart Grid* 7: 2448–2455.
19. Maciejowska K, Nowotarski J, Weron R (2016) Probabilistic forecasting of electricity spot prices using Factor Quantile Regression Averaging. *Inter J Forecasting* 32: 957–965.
20. Zou H, Hastie T, Tibshirani R, et al. (2007) On the ‘degrees of freedom’ of the lasso. *Ann Stat* 35: 2173–2192.
21. Zou H, Hastie T (2005) Regularization and variable selection via the elastic net. *J R stat soc: Ser B (Stat Methodol)* 67: 301–320.
22. Franklin, J (2005) The elements of statistical learning: data mining, inference and prediction. *Math Intell* 27: 83–85.
23. Friedman JH (2002) Stochastic gradient boosting. *Comput Stat Data Analysis* 38: 367–378.
24. Vapnik V (2013) *The nature of statistical learning theory*, Springer science and business media.
25. Makridakis S, Spiliotis E, Assimakopoulos V (2019) Predicting/hypothesizing the findings of the M4 Competition. *Int J Forecasting*.



AIMS Press

©2019 Norman Maswanganyi et al., licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)