

# Economic dispatch of electrical power in South Africa: An application to the Northern Cape province

Thakhani Ravele<sup>1,\*</sup>, Caston Sigauke<sup>1</sup>, Lordwell Jhamba<sup>2</sup>

<sup>1</sup>*Department of Mathematical and Computational Sciences, University of Venda, South Africa*

<sup>2</sup>*Department of Physics, University of Venda, South Africa*

**Abstract** Power utility companies rely on forecasting for the operation of electricity demand. This presents an application of linear quantile regression, non-linear quantile regression, and additive quantile regression models for forecasting extreme electricity demand at peak hours such as 18:00, 19:00, 20:00 and 21:00 using Northern Cape data period 01 January 2000 to 31 March 2014. The variables were selected using the least absolute shrinkage and selection operator. Additive quantile regression models were found to be the best-fitting models for hours 18:00 and 19:00, whereas linear quantile regression models were found to be the best-fitting models for hours 20:00 and 21:00. Out of sample forecasts for seven days (01 to 07 April 2014) were used to solve the unit commitment problem using mixed-integer programming. The unit commitment problem results showed that using all the generating units such as hydroelectric, wind power, concentrated solar power and solar photovoltaic is less costly. This study's main contribution is the development of models for forecasting hourly extreme peak electricity demand. These results could be useful to system operators in the energy sector who have to maintain the minimum cost by scheduling and dispatching electricity during peak hours when the grid is constrained due to peak load demand.

**Keywords** Additive quantile regression, Lasso, Linear quantile regression, Mixed-integer linear programming, Non-linear quantile regression, Unit commitment.

**DOI:** 10.19139/soic-2310-5070-1057

## 1. Introduction

### 1.1. Background

Economic dispatch is essential in power system operation and is defined as the power planning operation with minimum operating costs [1]. The purpose of economic dispatch is to provide optimal power generation at a minimum cost of operation. It also provides the important aspects of power system operation such as meeting load demand at minimum cost by scheduling the committed generating units, reducing the emissions, maintaining the system stability, and security restriction [2]. Electricity load forecasting is important for economic dispatch because it provides future electricity production and consumption [3], which helps electricity utility to maintain the balance of demand and supply [4]. Electricity load forecasting is also important for production planning and trading on the electricity markets. It has different implementations such as energy acquiring and production, load switching, contract rating and infrastructure evolution. Load forecasting helps electric utility management in planning the distribution of electricity [5, 6, 7]. Electricity load forecasting faces rising challenges due to innovative technologies such as smart grids, electric cars, and renewable energy production. The purpose of

---

\*Correspondence to: Thakhani Ravele (Email: ravelethakhani@gmail.com). Department of Mathematical and Computational Sciences, Faculty of Science, Engineering and Agriculture, University of Venda, Private Bag X5050, Thohoyandou, 0950, South Africa.

electricity load forecasting is to predict hourly, daily, weekly and yearly demand or peak demand [8]. However, this paper will model hourly extreme peak electricity demand. The focus is on an estimation of high tail quantiles of the distribution.

Peak electricity is the highest load at a given time. Peak electricity load forecasting is essential because it assures the availability of enough supply. Under forecasting of peak electricity load results in an insufficient capacity for meeting load demand and blackouts. Power blackouts are a problem because they affect the operation of the economy. Extreme peak electricity load forecasting is the solution to the underprediction of the peak electricity load demand. Accurate peak electricity load forecasting is very important as it provides future forecasts which are useful to prevent system failure and power blackouts [9].

### *1.2. An overview of the literature on load forecasting*

Electricity load forecasting has received much attention from industrialists and academics in recent years [8]. Various forecasting techniques for forecasting electricity loads have been developed and they are classified into statistical and artificial intelligence techniques [10]. The type of electricity load forecasting can vary depending on the result desired. It may be spatial if it mainly relates to studying future patterns in a specific region, country, or state. Otherwise, it is temporal if concerned with forecasting hourly, daily, monthly or yearly. Electricity load forecasting is carried out according to their time horizons, such as very short-term load forecasts (VSTLF), short-term load forecasts (STLF), which are from one hour to one week, medium-term load forecasts (MTLF) are generated from a week to a year, and long term load forecasts (LTLF) range up for longer than a year [10, 11, 12]. Reference [9] used a quantile regression model to develop 1.00 quantiles of the distribution to prevent system failure and eliminate power blackouts. In other words, the idea was to provide a model that would avoid under prediction. The estimates from the quantile regression model were compared with the actual demand to investigate the ability of the model to avoid power blackouts (avoid underprediction). The results showed that the prediction of the upper limit for the daily peak demand is accurate when using the 1.00 quantile from the 0.99 to 0.97 quantiles of the distribution [9].

Unit sectors require accurate peak electricity load forecasting, as climate change, technological development, and energy policies contribute to rising peak load. Decision-makers and power generation companies rely on accurate energy demand forecasting regarding policy creation and power generation planning [13]. Accurate forecasting of peak electricity demand is essential in the electricity sector for planning capacity enlargement and medium-term risk assessment [14]. Peak electricity load forecasting and modelling using South African data are discussed in the literature, see [14, 15, 16]. Reference [15] used the additive model that allows non-linear and nonparametric terms to forecast daily winter peak electricity demand in South Africa. The study showed that peak electricity demand is highly sensitive to cold temperatures compared to hot temperatures in South Africa. Modelling extreme peak electricity load is useful for quantifying the amount of electricity that can be shifted from the grid to off-peak periods [16]. Reference [16] modelled extreme daily increases in peak electricity demand in South Africa, focusing on tail quantiles of the distribution of daily peak electricity demand. The forecasting of peak electricity demand using South African data is studied by [14] in which the authors focused on an application of partially linear additive quantile regression models for modelling and forecasting peak electricity demand.

Additive quantile regression (AQR) models were used in South Africa to forecast short-term hourly load. A combination of forecasts from four developed models was done based on pinball loss and quantile regression averaging (QRA). The study found that the AQR model with interactions produced accurate forecasts compared to the QRA model [17]. Factors such as wind speed, solar irradiance, temperature, cloud cover, and seasonal variations significantly affect electricity demand, resulting in uncertain and unpredictable electricity demand patterns. However, it is important to create a robust, intelligent, and adaptive forecast model that accommodates the factors affecting power demand for higher forecast accuracy [18].

Unit commitment (UC), also known as an optimisation problem, helps in managing generating units (when to switch on or off) under various restrictions and environments [19]. UC maximises power systems operational with the minimal costs of production and reserve requirements [20, 21]. A reasonable solution to the optimisation problem is very important because it provides operational planners with the optimal number of required generators. According to [22] there are different methods for determining a reasonable solution to the optimisation problem. Some optimisation techniques used in solving the unit commitment are, Lagrangian relaxation (LR), mixed-integer linear programming (MILP), tabu search (TS), dynamic programming (DP) and stochastic programming (SP) techniques, see [19]. However, reference [23] revealed that many authors used mixed integer programming (MIP) and LR methods. Reference [23] developed an approach that integrates short-term load forecasting with UC, which minimises production costs from one thermal plant in Turkey, the Kutahya region. Applying the LR method [23] provides solutions to a UC problem using forecasts from two developed models for electricity demand. Accurate load forecasts are essential as they are some of the inputs in solving the UC problem.

### 1.3. Motivation and contributions

This study intends to solve the economic dispatch of electricity through the following contributions:

- Forecasting extreme high quantile of the distribution, which improves the accuracy of forecasts,
- Using seven days out of sample forecasts to solve the unit commitment problem,
- Improve the economic dispatch of electricity through the inclusion of renewable energy sources on the unit commitment problem,
- Providing results which help operate electricity utility at a minimal cost.

## 2. The proposed models

### 2.1. Linear quantile regression model

Quantile regression (QR) was introduced by [24], which provides a modelling approach to predict conditional quantiles of the response variable. Linear quantile regression assumes a linear relationship between the response variable and a vector of explanatory variables when estimating the quantiles of the cumulative distribution function of the response variable [25]. Following [25], linear quantile regression is given by:

$$y_{t,h,\tau} = \beta_{\tau} \mathbf{X}_t + \xi_{t,\tau}; \quad \tau \in (0, 1), \quad (1)$$

where  $y_{t,h,\tau}$  is electricity demand on day  $t = 1, \dots, n$  at hour  $h, h = 18 : 00, 19 : 00, 20 : 00, 21 : 00$  at quantile  $\tau, \beta$  denotes a vector of parameters and  $\xi_{t,\tau}$  is a random error term. From [24], quantiles are estimated using asymmetric weights to the mean absolute error. The quantile loss function is represented by:

$$\rho_{\tau}(u) = \begin{cases} \tau(u), & \text{if } u \geq 0, \\ (1 - \tau)u, & \text{if } u < 0, \end{cases} \quad (2)$$

where  $\tau$  is the quantile probability level. The  $\tau^{th}$  quantile is estimated by the quantile regression method and the vector  $\hat{\beta}$  using the minimisation problem given by:

$$\hat{\beta}_{\tau} = \arg \min_{\beta} \sum_{t=1}^n \rho_{\tau}(y_{t,\tau} - \mathbf{X}_t \beta_{\tau}). \quad (3)$$

### 2.2. Non-linear quantile regression model

Non-linear quantile regression assumes a non-linear relationship between the response variable and a vector of explanatory variables when predicting the quantiles of the distribution. In non-linear quantile regression modelling,

a non-linear mapping function  $q_\tau$  converted the vectors of explanatory variables  $X_t$  into a viable higher dimensional feature space [26]. A non-linear quantile regression model is defined by [26, 27, 28]:

$$y_{t,h,\tau} = q_\tau(X_t, \beta_\tau) + \xi_{t,\tau}; \quad \tau \in (0, 1), \quad (4)$$

where  $y_{t,h,\tau}$  is electricity demand on day  $t = 1, \dots, n$  at hour  $h, h = 18 : 00, 19 : 00, 20 : 00, 21 : 00$  at quantile  $\tau$ ,  $\beta_\tau$  denotes a vector of parameters,  $q_\tau$  is the mapping function and  $\xi_{t,\tau}$  denotes random error term. Equation (4) can be estimated by:

$$\hat{\beta}(\tau) = \arg \min_{\beta \in \mathbb{R}} \sum_{t=1}^n \rho_\tau(y_{t,\tau} - q_\tau(X_t, \beta)), \quad (5)$$

### 2.3. Additive quantile regression model

An additive quantile regression (AQR) model is a hybrid model that integrates GAM and QR models. AQR models were initially used by [29] to estimate short term load demand and also extended by [30, 31]. The AQR model is written as [29, 30, 31]:

$$y_{t,h,\tau} = \sum_{i=1}^p s_{i,h,\tau}(x_{ti}) + \varepsilon_{t,h,\tau}; \quad \tau \in (0, 1), \quad (6)$$

where  $y_{t,h,\tau}$  is electricity demand on day  $t = 1, \dots, n$  at hour  $h, h = 18 : 00, 19 : 00, 20 : 00, 21 : 00$  at quantile  $\tau$ ,  $s_{i,h,\tau}$  denote the smooth functions and  $\varepsilon_{t,h,\tau}$  is the error term. The smooth function,  $s$ , is given by:

$$s_i(x) = \sum_{j=1}^q \beta_{ij} b_{ij}(x_{ti}), \quad (7)$$

where  $\beta_{ij}$  represents the  $i^{th}$  unknown coefficient (parameter) and  $b_{ij}(x_{ti})$  are known as spline basis functions. The parameter estimates of Equation (6) are obtained by minimising the function given by:

$$q_{Y|X}(\tau) = \sum_{t=1}^n \rho_\tau \left( y_{t,\tau} - \sum_{i=1}^p s_{i,\tau}(x_{ti}) \right), \quad (8)$$

where

$$\rho_\tau(z) = (\tau - 1)z\mathbf{I}(z < 0) + \tau z\mathbf{I}(z \geq 0)$$

represents the pinball loss function. This study uses the penalised pinball loss function discussed in [31]. Let  $\mu(x_t) = X_t^T \beta$ , with  $x_t$  denoting the  $t^{th}$  row of the  $n \times d$  design matrix  $X$ . The penalised pinball loss function is then defined as [31]:

$$V(\beta, \gamma, \sigma) = \sum_{t=1}^n \frac{1}{\sigma} \rho_\tau [y_t, \tau - \mu(x_t)] + \frac{1}{2} \sum_{i=1}^p \gamma_i \beta^T S_i \beta, \quad (9)$$

where  $\gamma$  denotes a vector of smooth parameters, i.e.  $\gamma = \{\gamma_1, \dots, \gamma_p\}$ ,  $S_i$  represents the positive semi-definite matrices which are meant to penalise the wiggleness of  $\mu(x)$  [31]. The term  $\frac{1}{\sigma}$  represents the learning rate used to determine the weight of the loss and the penalty. For more details see [31].

### 2.4. Error measures for probabilistic forecasting and evaluation of methods

Probabilistic forecasts from proposed models will be evaluated and compared using scoring rules. The scoring rule assigns a penalty score represented by  $S(y, F)$  to the probabilistic forecast, where  $y$  denotes the observation used for forecast assessment and  $F$  represents the forecast distribution [32]. A small score shows a better forecast. In this study, three error measures will be used such as continuous rank probability score (CRPS), Dawid-Sebastiani score (DSS) and the pinball loss function (PLF).

2.4.1. *Continuous rank probability score* The CRPS compares the distance between the predicted and the observed cumulative density functions of scalar variables [32]. The CRPS is defined by:

$$\text{CRPS}(y, F) = \int_0^1 \text{QS}_\tau \left( F^{-1}(\tau), y \right) d\tau, \tag{10}$$

where  $F$  represents the forecast distribution and  $\text{QS}_\tau$  denotes the quantile score defined by:

$$\text{QS}_\tau \left( F^{-1}(\tau), y \right) = 2 \left( \mathbf{I} \left[ y \leq F^{-1}(\tau) \right] - \tau \right) \left( F^{-1}(\tau) - y \right), \tag{11}$$

where  $\mathbf{I}$  denotes an indicator function.

2.4.2. *Dawid-Sebastiani score* The drawback of CRPS is that it is difficult to compute complex forecast distributions. The DDS is the alternative solution because it overcomes the drawback of CRPS by computing the complex forecast distribution easily. It is given by [33]:

$$\text{DSS}_{y,F} = \frac{(y - \mu_F)^2}{(\sigma_F^2)} + 2 \log(\sigma_F), \tag{12}$$

where  $F$  denotes forecast distribution, with the mean and standard deviation given by  $\mu_F$  and  $\sigma_F^2$ , respectively.

2.4.3. *Pinball loss function* The PLF is relatively easy to use and is defined by:

$$L(q_\tau, y_t) = \begin{cases} \tau(y_t - q_\tau), & \text{if } y_t > q_\tau, \\ (1 - \tau)(q_\tau - y_t), & \text{if } y_t \leq q_\tau, \end{cases} \tag{13}$$

where  $q_\tau$  represents quantile forecast and  $y_t$  is the observed value of electricity demand.

### 2.5. Unit commitment

Unit commitment (UC) minimises generating units’ total cost within a specific time, or interval [21]. According to [34] different methods and algorithms used to determine the optimal solution of the UC problems are classified into deterministic, heuristic, and hybrid approaches. This study will use mixed-integer linear programming (MILP), which falls under the deterministic approach. MILP is a special class of linear programming [21, 35], where variables are made of integer and continuous variables [34]. This study envisions demonstrating how to use forecasts to solve the UC problem. Let  $P_{G_i}^{ht}$  be Northern Cape system load at hour  $h$ ,  $h = 18 : 00; 19 : 00; 20 : 00; 21 : 00$  on day  $t$ ,  $t = 1, \dots, n$ ,  $P_{G_i}^{ht}(\min)$  is the lower limit of the power output,  $P_{G_i}^{ht}(\max)$  represents the upper limit of the unit power output,  $x_i^{ht}$  be the 0 - 1 variable (This study assumes that during the peak period all units are up, i.e  $x_i^{ht} = 1$  for all units),  $F_{s_i}$  is the start-up cost of unit  $i$  at hour  $h$  (for this study we assume that the start-up cost is zero),  $P_R^{ht}$  is the power reserve at hour  $h$  on day  $t$  and  $F_i$  is the average production cost of unit  $i$  (cost/MW). This study will use the fuel cost to represent the average production cost per megawatt.

Generating units are grouped into one hydroelectric, eight wind power, eight concentrated solar power (CSP) and 19 solar photovoltaic (PV) power. These will be denoted as:

$$g_{1d}; g_{2w}, g_{3w}, \dots, g_{9w}; g_{10c}, g_{11c}, \dots, g_{17c}; g_{18p}, g_{19p}, \dots, g_{36p}.$$

The objective function to minimise generating unit cost over a specific time is given by [21, 35]:

$$\min \sum_{h=1}^H \sum_{i=1}^m \left[ F_i \left( P_{G_i}^{ht} \right) x_i^{ht} + F_{s_i}(ht) x_i^{ht} \right] = F \left( P_{G_i}^{ht}, x_i^{ht} \right), \tag{14}$$

The constraints are defined as follows:

### Load balance equation

$$\sum_{i=1}^m P_{G_i}^{ht} x_i^{ht} = P_D^{ht}, \quad h = 18, 19, 20, 21, t = 1, \dots, n, \quad (15)$$

### Generator power output limits

$$x_i^{ht} P_{G_i \min} \leq P_{G_i}^{ht} \leq x_i^{ht} P_{G_i \max}, \quad h = 18, 19, 20, 21, t = 1, \dots, n, i = 1, \dots, m \quad (16)$$

In this study  $m = 36$ .

## 3. Empirical results and discussion

This section presents the results of forecasting electricity demand at hours 18:00, 19:00, 20:00, and 21:00. The data used in this study is hourly electricity load demand for the Northern Cape, South Africa, for the period 01 January 2000 to 31 March 2014 giving a sample size of  $n = 5204$ . The load at hours 18:00, 19:00, 20:00 and 21:00, respectively, represents the response variable and the predictor variables are temperature, non-linear trend, and daytype. Hourly temperature data used is from De Aar weather station (Latitude: 30° 38' 59" S and Longitude: 24° 0' 44" E) and Calvinia weather station (Latitude: 31° 28' 14" S and Longitude: 19° 46' 34" E) in Northern Cape, South Africa. The temperature data were aggregated to get each station's maximum, minimum, and average daily temperature. The maximum, minimum, and average daily temperature for both stations were then combined to get the average of the maximum (AveMaxT), minimum (AveMinT), and average daily temperature (AveTem) of the province that they represent. A penalised cubic smoothing spline was fitted to the response variable to determine the non-linear trend (noltrend18, noltrend19, noltrend20, and noltrend21). Daytype variable denotes the days of the week, coded as 1 for Monday, 2 for Tuesday to 7 for Sunday.

### 3.1. Exploratory data analysis

The summary statistics of electricity demand at hours 18:00, 19:00, 20:00, and 21:00 for the sampling period 01 January 2000 to 31 March 2014 is presented in Table 1. The maximum demand is 880, 905, 884, and 855 for each of the four hours. The mean and the median are not equal, which confirms that the distributions of the four hours are not normally distributed. The skewness and kurtosis also confirm that the distributions are non-normal.

Table 1. Summary statistics for electricity demand (MW) at hours 18:00, 19:00, 20:00, and 21:00.

| Hour  | Mean  | Median | Minimum | Maximum | Standard deviation | Skewness   | Kurtosis   |
|-------|-------|--------|---------|---------|--------------------|------------|------------|
| 18:00 | 676.7 | 684    | 397     | 880     | 83.15992           | -0.2860717 | -0.3457701 |
| 19:00 | 707.8 | 714    | 403     | 905     | 86.97296           | -0.3078065 | -0.3036559 |
| 20:00 | 715.9 | 727    | 434     | 884     | 74.94497           | -0.4881771 | -0.2331249 |
| 21:00 | 690.6 | 699    | 439     | 855     | 69.54004           | -0.3849729 | -0.3146267 |

The Northern cape electricity demand time series and density plots at 18:00, 19:00, 20:00, and 21:00 are shown in Figure 1. The left panels of Figure 1 present the time series plots, whereas the right panels show density plots of the demand for the considered hours, respectively. Seasonality patterns of electricity demand in the Northern Cape on the left panels of Figure 1 show that higher electricity demand occurs in winter and lower demand occurs in summer yearly. The densities of the four hours on the right panels of Figure 1 indicate that the distributions do not follow normal distributions, which supports the report of skewness and kurtosis in Table 1. Figure 2 highlighted electricity demand for four hours in box and whisker plots. Figure 3 shows the plots of electricity demand for four hours superimposed with a non-linear trend.

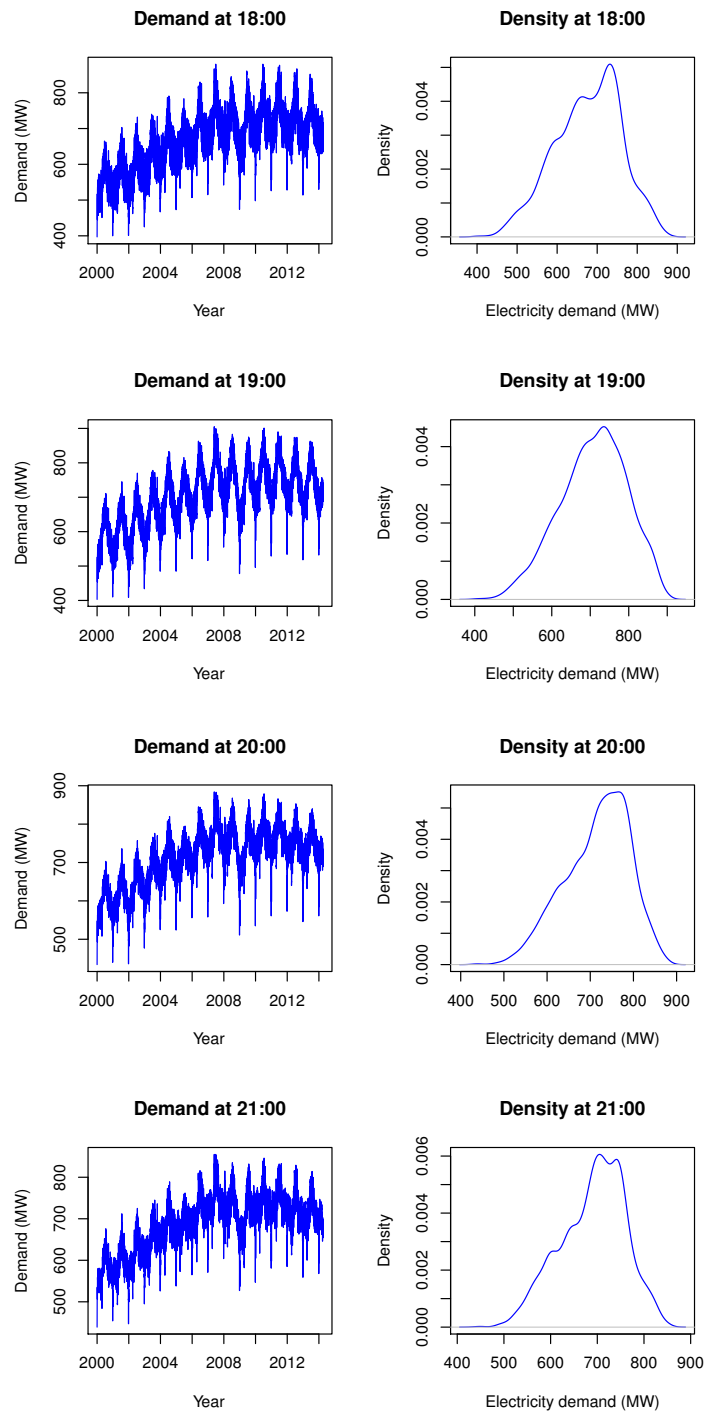


Figure 1. Plot of electricity demand and densities at 18:00, 19:00, 20:00, and 21:00.

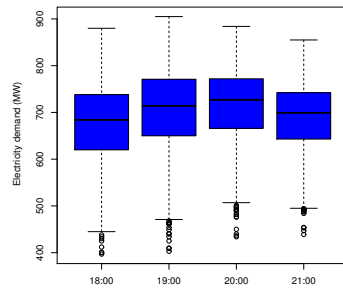


Figure 2. Box plots for hours 18:00, 19:00, 20:00, and 21:00.

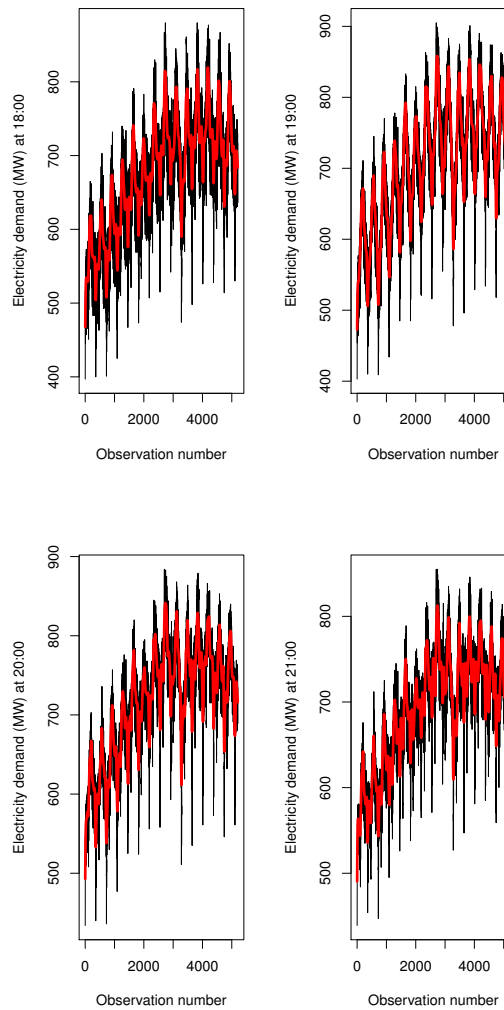


Figure 3. Plot of electricity demand at hours 18:00, 19:00, 20:00, and 21:00 with a non-linear trend (red curve).



### 3.2. Forecasting results

3.2.1. *Forecasting electricity demand when covariates are given* The data is for 01 January 2000 to 31 March 2014, giving a sample size of  $n = 5204$  observations. The data is split into two parts. The data for the period 01 January 2000 to 25 May 2011 is used for training, which is 80% of the data ( $n_1 = 4163$ ), and the period from 26 May 2011 to 31 March 2014 is used for testing, which is 20% of the data ( $n_2 = 1041$ ). The models considered are linear quantile regression (LQR), non-linear quantile regression (NLQR), and additive quantile regression (AQR). Based on the pinball loss presented in Table 2, the best fitting model is AQR for hours 18:00 and 19:00, and LQR for hours 20:00 and 21:00. This is done for  $\tau = 0.9999$  using R-package “qgam” for AQR, and “quantreg” for LQR, and NLQR.

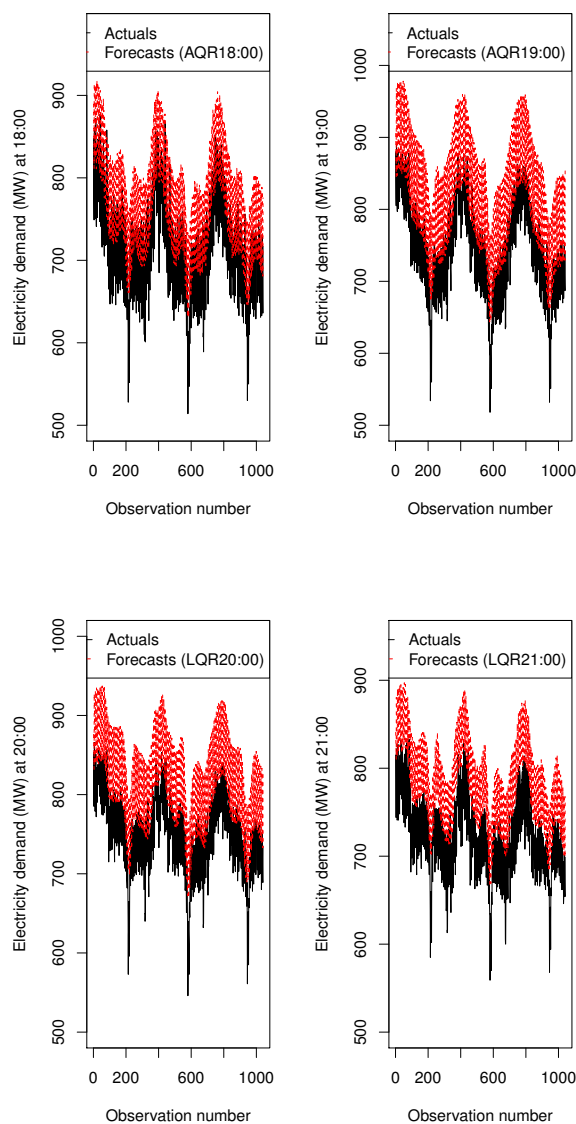


Figure 4. Plot of actual demand and forecasts at 18:00, 19:00, 20:00, and 21:00.

Table 2. Summary table for Pinball loss for  $\tau = 0.9999$

|      | 18:00           | 19:00           | 20:00           | 21:00          |
|------|-----------------|-----------------|-----------------|----------------|
| LQR  | 0.014980        | 0.013776        | <b>0.012544</b> | <b>0.01146</b> |
| NLQR | 0.027425        | 0.020668        | 0.019867        | 0.020587       |
| AQR  | <b>0.013498</b> | <b>0.013441</b> | 0.012827        | 0.011964       |

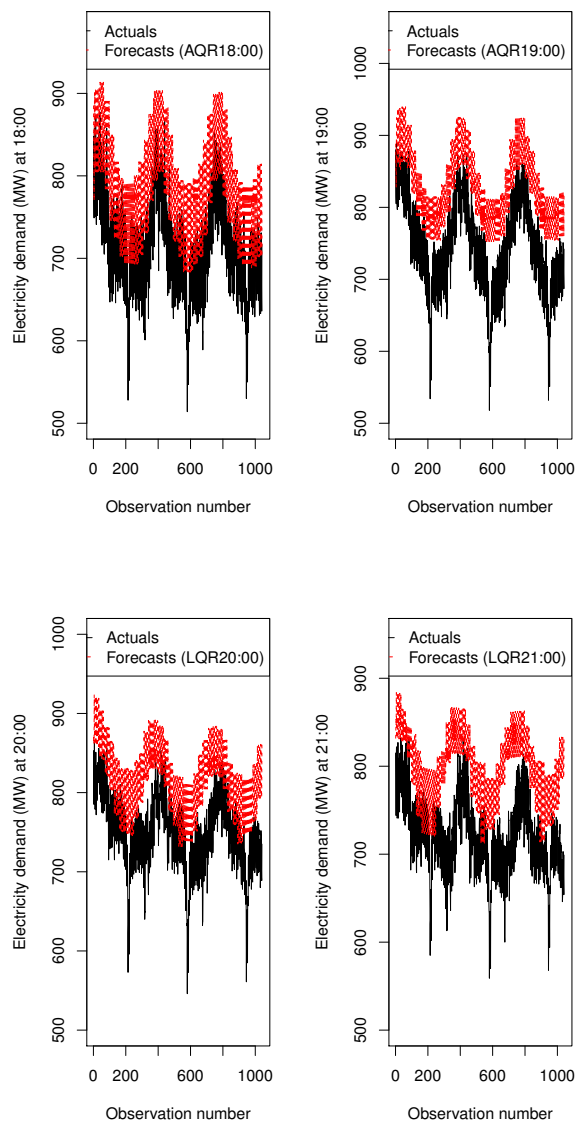


Figure 5. Plot of actual demand and forecasts at hours 18:00, 19:00, 20:00, and 21:00 (Operational forecasts).

The best-fitting models were used to forecast the electricity demand after variable importance and variable selection using the least absolute shrinkage and selection operator (Lasso). The plots of actual demand and forecasts from the developed models are for hours 18:00, 19:00, 20:00, and 21:00 in Figure 4. It is shown in Figure 4 that the forecasts from each model follow the actual demand data remarkably well at a high quantile.

3.2.2. *Forecasting electricity demand when covariates are not given* The predictor variables Daytype, Month, and Trend are used to forecast the unknown variables: AveMaxT, AveMinT, AveTem, noltrend18, noltrend19, noltrend20, and noltrend21. The forecasted predictor variables are then used to forecast electricity demand at hours 18:00, 19:00, 20:00, and 21:00 shown in Figure 5.

Table 3. Base load demand stations.

| Unit | Min Ave prod cost $c_i$ (rands) | Min (MW) | Max (MW) |
|------|---------------------------------|----------|----------|
| g10c | 3753                            | 0        | 50       |
| g11c | 3753                            | 0        | 100      |
| g12c | 3753                            | 0        | 36.652   |
| g13c | 3753                            | 0        | 100      |
| g14c | 3753                            | 0        | 100      |
| g15c | 3753                            | 0        | 100      |
| g16c | 3753                            | 0        | 100      |
| g17c | 3753                            | 0        | 100      |
| g18p | 2518                            | 0        | 60.085   |
| g19p | 2518                            | 0        | 9.65     |
| g20p | 2518                            | 0        | 9.65     |
| g21p | 2518                            | 0        | 10       |
| g22p | 2518                            | 0        | 19.9     |
| g23p | 2518                            | 0        | 19.93    |
| g24p | 2518                            | 0        | 48.3     |
| g25p | 2518                            | 0        | 9.7      |
| g26p | 2518                            | 0        | 40       |
| g27p | 2518                            | 0        | 75       |
| g28p | 2518                            | 0        | 45.4     |
| g29p | 2518                            | 0        | 64       |
| g30p | 2518                            | 0        | 75       |
| g31p | 2518                            | 0        | 74       |
| g32p | 2518                            | 0        | 36.8     |
| g33p | 2518                            | 0        | 75       |
| g34p | 2518                            | 0        | 8.9      |
| g35p | 2518                            | 0        | 75       |
| g36p | 2518                            | 0        | 14.915   |

3.2.3. *Unit commitment results* Table 3 shows the base load demand stations, i.e. CSP and PV. This study used fuel cost to represent the average production cost, given in column 2 of Table 3. Columns 3 and 4 represent the megawatts' minimum and maximum production levels, respectively. Data for peaking stations, i.e. hydroelectric and wind power, are given in Table 4, which is similar to Table 3. The fuel costs data is from [36]. The out-of-sample forecasts for the first seven days of April 2014 obtained using the models for hours 18:00, 19:00, 20:00

and 21:00 are presented in Table 5. The application unit commitment problem is demonstrated using the forecasted demand for Monday, 07 April 2014, shown in bold. UC problem was solved using Lingo version 18. Although the study applied the UC problem for the day 07 April 2014, it is important to note that the procedure can be applied to other hours of the day. The bounded variable MILP model is given by:

$$\text{Min TC} = c_1g_{1d} + \sum_{j=2}^9 c_jg_{jw} + \sum_{j=10}^{17} c_jg_{jc} + \sum_{j=18}^{36} c_jg_{jp}, \tag{17}$$

where TC is the total cost to be minimised. The demand used for hour 18:00, 19:00, 20:00 and 21:00 are  $P_D^{18t} = 821.145$ ,  $P_D^{19t} = 894.737$ ,  $P_D^{20t} = 20:00$  and  $P_D^{21t} = 21:00$  respectively.

Table 4. Peaking stations.

| Unit | Min Ave prod cost $c_i$ (rands) | Min (MW) | Max (MW) |
|------|---------------------------------|----------|----------|
| g1d  | 1470                            | 0        | 77.46    |
| g2w  | 1293                            | 0        | 73.8     |
| g3w  | 1293                            | 0        | 139      |
| g4w  | 1293                            | 0        | 96       |
| g5w  | 1293                            | 0        | 140      |
| g6w  | 1293                            | 0        | 80       |
| g7w  | 1293                            | 0        | 140      |
| g8w  | 1293                            | 0        | 70.555   |
| g9w  | 1293                            | 0        | 70.555   |

Table 5. Out of sample forecasts.

| Date                | 18:00          | 19:00          | 20:00          | 21:00          |
|---------------------|----------------|----------------|----------------|----------------|
| Tue 01-04-14        | 816.721        | 869.6          | 884.442        | 827.603        |
| Wed 02-04-14        | 827.579        | 862.975        | 872.569        | 821.538        |
| Thu 03-04-14        | 840.839        | 856.346        | 859.523        | 815.473        |
| Fri 04-04-14        | 814.269        | 858.669        | 835.682        | 792.56         |
| Sat 05-04-14        | 778.888        | 848.844        | 825.485        | 786.495        |
| Sun 06-04-14        | 735.893        | 835.653        | 818.171        | 780.43         |
| <b>Mon 07-04-14</b> | <b>821.145</b> | <b>894.737</b> | <b>879.355</b> | <b>816.819</b> |

The generator power output limits are given by:

$$0 \leq g_{1d} \leq 77.46 \tag{18}$$

$$\cdot \tag{19}$$

$$\cdot \tag{20}$$

$$\cdot \tag{21}$$

$$0 \leq g_{36p} \leq 14.915 \tag{22}$$

The optimal solution for each of the four hours on Monday 07 April 2014 is presented in Table 6. The amount of electricity in megawatts of the generating units is shown in column 2 of Table 6, which should be produced to

meet the predicted demand while ensuring that the power reserve constraint is met. The minimum cost in rands of producing electricity for each of the four hours is also shown. This study is different from [35] study because it uses extremely high quantile ( $\tau = 0.9999$ ) forecasts to solve the UC problem. This procedure can be used by power utility companies like Eskom with the economic dispatching and UC scheduling of electricity.

Table 6. Optimal Solution using bounded variable MILP.

| Variable | 18:00 (R2,067,643.00) |      | 19:00 (R2,483,463.00) |    | 20:00 (R1,137,006.00) |      | 21:00 (R1,069,858.00) |    |
|----------|-----------------------|------|-----------------------|----|-----------------------|------|-----------------------|----|
|          | Soln                  | RC   | Soln                  | RC | Soln                  | RC   | Soln                  | RC |
| g1d      | -                     | -    | -                     | -  | -                     | -    | 77.46                 | 0  |
| g2w      | -                     | -    | -                     | -  | 73.8                  | 0    | 73.8                  | 0  |
| g3w      | -                     | -    | -                     | -  | 139                   | 0    | 139                   | 0  |
| g4w      | -                     | -    | -                     | -  | 96                    | 0    | 96                    | 0  |
| g5w      | -                     | -    | -                     | -  | 140                   | 0    | 140                   | 0  |
| g6w      | -                     | -    | -                     | -  | 80                    | 0    | 80                    | 0  |
| g7w      | -                     | -    | -                     | -  | 140                   | 0    | 140                   | 0  |
| g8w      | -                     | -    | -                     | -  | 140                   | 0    | 70.555                | 0  |
| g9w      | -                     | -    | -                     | -  | 70.555                | 0    | -                     | -  |
| g10c     | 0                     | 1235 | 50                    | 0  | 0                     | 2460 | -                     | -  |
| g11c     | 0                     | 1235 | 100                   | 0  | 0                     | 2460 | -                     | -  |
| g12c     | 0                     | 1235 | 36.652                | 0  | 0                     | 2460 | -                     | -  |
| g13c     | 0                     | 1235 | -                     | -  | 0                     | 2460 | -                     | -  |
| g14c     | 0                     | 1235 | -                     | -  | 0                     | 2460 | -                     | -  |
| g15c     | 0                     | 1235 | -                     | -  | 0                     | 2460 | -                     | -  |
| g16c     | 0                     | 1235 | -                     | -  | -                     | -    | -                     | -  |
| g17c     | 0                     | 1235 | -                     | -  | -                     | -    | -                     | -  |
| g18p     | 75                    | 0    | 60.085                | 0  | -                     | -    | -                     | -  |
| g19p     | 9.65                  | 0    | 75                    | 0  | -                     | -    | -                     | -  |
| g20p     | 9.65                  | 0    | 28                    | 0  | -                     | -    | -                     | -  |
| g21p     | 10                    | 0    | 75                    | 0  | -                     | -    | -                     | -  |
| g22p     | 19.9                  | 0    | 75                    | 0  | -                     | -    | -                     | -  |
| g23p     | 19.93                 | 0    | 75                    | 0  | -                     | -    | -                     | -  |
| g24p     | 48.3                  | 0    | 75                    | 0  | -                     | -    | -                     | -  |
| g25p     | 9.7                   | 0    | 75                    | 0  | -                     | -    | -                     | -  |
| g26p     | 75                    | 0    | 40                    | 0  | -                     | -    | -                     | -  |
| g27p     | 75                    | 0    | 75                    | 0  | -                     | -    | -                     | -  |
| g28p     | 45.4                  | 0    | 55                    | 0  | -                     | -    | -                     | -  |
| g29p     | 64                    | 0    | -                     | -  | -                     | -    | -                     | -  |
| g30p     | 75                    | 0    | -                     | -  | -                     | -    | -                     | -  |
| g31p     | 74                    | 0    | -                     | -  | -                     | -    | -                     | -  |
| g32p     | 36.8                  | 0    | -                     | -  | -                     | -    | -                     | -  |
| g33p     | 75                    | 0    | -                     | -  | -                     | -    | -                     | -  |
| g34p     | 8.9                   | 0    | -                     | -  | -                     | -    | -                     | -  |
| g35p     | 75                    | 0    | -                     | -  | -                     | -    | -                     | -  |
| g36p     | 14.915                | 0    | -                     | -  | -                     | -    | -                     | -  |

#### 4. Conclusion

Using Northern Cape data, this study applied LQR, NLQR, and AQR models to peak electricity demand forecasting at extreme quantile ( $\tau = 0.9999$ ). The paper's thrust was on peak hourly electricity demand at an extremely high quantile ( $\tau = 0.9999$ ). Lasso was for variable selection. The AQR models were found to be the best fitting models for hours 18:00 and 19:00, whereas LQR models were the best fitting models for hours 20:00 and 21:00. The best-fitting models were used to forecast the demand for each hour. Operational forecasts were also done for each of the four hours. Out of sample forecasts given in Table 5 from the four hours were then used as inputs in solving the UC problem. The Lingo version 18 was used to solve optimisation models. The out of sample forecasts for Monday 07 April 2014 were 821.145 MW, 894.737 MW, 879.355 MW and 816.819 MW for the hours 18:00, 19:00, 20:00 and 21:00 respectively, with the corresponding optimal minimal costs of R2,067,643.00, R2,483,463.00, R1,137,006.00 and R1,069,858.00. The UC problem results showed that using all the hydroelectric generating units, wind power, CSP and PV is less costly. These were all selected as part of the optimal solution. This study's main contribution is the development of models for forecasting hourly extreme peak electricity demand. These results could be useful to system operators in the energy sector who have to maintain the minimum cost by scheduling and dispatching electricity during peak hours when the grid is constrained due to peak load demand.

#### Acknowledgement

The authors acknowledge the DSI – CSIR Inter bursary Support (IBS) Programme for financial support and to both Eskom, South Africa's power utility company and the South African Weather Services for providing the data.

#### REFERENCES

1. E. Arriagada, E. Lopez, M. López, R. Blasco-Gimenez, C. Roa, and M. Poloujadoff, *A probabilistic economic dispatch model and methodology considering renewable energy, demand and generator uncertainties*, Electric Power Systems Research, vol. 121, pp. 325–332, 2015.
2. F.N. Al Farsi, M.H. Albadi, N. Hosseinzadeh, and A.H. Al Badi, *Economic dispatch in power systems*, In 2015 IEEE 8th GCC Conference & Exhibition, pp. 1–6, 2015.
3. S. Fan, and R.J. Hyndman, *Short term load forecasting based on a semiparametric additive model*, IEEE Trans Power Syst 2012, vol. 27, no. 1, pp. 134–41, 2012.
4. A. Pierrot, and Y. Goude, *Short-term electricity load forecasting with generalized additive models*, Proc ISAP Power 2011, pp. 593–600, 2011.
5. F. Cavallaro, *Electric load analysis using an artificial neural network*, International journal of energy research, vol. 29, no. 5, pp. 377–392, 2005.
6. M. Mohandes, *Support vector machines for short-term electrical load forecasting*, International Journal of Energy Research, vol. 26, no. 4, pp. 335–345, 2002.
7. G. Singh, D.S. Chauhan, A. Chandel, D. Parashar, and G. Sharma, *Factor affecting elements and short term load forecasting based on multiple linear*, In International Journal of Engineering Research and Technology, vol. 3, 2014.
8. Y. Goude, R. Nedellec, and N. Kong, *Local short and middle term electricity load forecasting with semi-parametric additive models*, IEEE transactions on smart grid, vol. 5, no. 1, pp. 440–446, 2014.
9. N. Elamin, *Quantile regression model for peak load demand forecasting with approximation by triangular distribution to avoid blackouts*, Annals of Statistics, pp. 65–81, 2018.
10. T. Hong, and S. Fan, *Probabilistic electric load forecasting: A tutorial review*, International Journal of Forecasting, vol. 32, no. 3, pp. 914–938, 2016.
11. E.A. Feinberg, and D. Genethliou, *Load forecasting*, In Applied mathematics for restructured electric power systems, pp. 269–285, 2005.
12. R.J. Hyndman, and S. Fan, *Density forecasting for long-term peak electricity demand*, IEEE Transactions on Power Systems, vol. 25, no. 2, pp. 1142–1153, 2010.
13. Y. Kim, H.G. Son, and S. Kim, *Short term electricity load forecasting for institutional buildings*, Energy Reports, vol. 5, pp. 1270–1280, 2019.
14. N. Maswanganyi, C. Sigauke, and E. Ranganai, *Peak electricity demand forecasting using partially linear additive quantile regression models*, In Annual Proceedings of the South African Statistical Association Conference (Vol. 2017, No. Congress 1, pp. 25–32), 2017. South African Statistical Association (SASA).

15. C. Sigauke, and D. Chikobvu, *Short-term peak electricity demand in South Africa*, African Journal of Business Management, vol. 6, no. 32, pp. 9243–9249, 2012.
16. C. Sigauke, A. Verster, and D. Chikobvu, *Extreme daily increases in peak electricity demand: Tail-quantile estimation*, Energy Policy, vol. 59, pp. 90–96, 2013.
17. C. Sigauke, M.M. Nemukula, and D. Maposa, *Probabilistic hourly load forecasting using additive quantile regression models*, Energies, vol. 11, no. 9, 2208, 2018.
18. MQ. Raza, M. Nadarajah, and C. Ekanayake, *On recent advances in PV output power forecast*, Solar Energy, vol. 136, pp. 125–144, 2016.
19. B. Saravanan, S. Das, S. Sikri, and D.P. Kothari, *A solution to the unit commitment problem—a review*, Frontiers in Energy, vol. 7, no. 2, pp. 223–236, 2013.
20. H. Dai, N. Zhang, and W. Su, *A literature review of stochastic programming and unit commitment*, Journal of Power and Energy Engineering, vol. 3, no. 04, pp. 206, 2015.
21. B. Wright, *A review of unit commitment*, ELENE4511, May, vol. 28, 2013.
22. T. Hong, P. Pinson, S. Fan, H. Zareipour, A. Troccoli, and R.J. Hyndman, *Probabilistic energy forecasting: Global energy forecasting competition 2014 and beyond* 2016.
23. M. Kurban, and U.B. Filik, *Unit commitment scheduling by using the autoregressive and artificial neural network models based on short-term load forecasting*, Proceedings of the 10th International Conference on Probabilistic Methods Applied to Power Systems, pp. 1–5, 2008.
24. R. Koenker, and G. Bassett Jr, *Regression quantiles*, Econometrica: journal of the Econometric Society, pp. 33–50, 1978.
25. P. Lauret, M. David, and H.T. Pedro, *Probabilistic solar forecasting using quantile regression models*, energies, vol. 10, no. 10, 1591, 2017.
26. C. Hwang, and J. Shim, *A simple quantile regression via support vector machine*, International Conference on Natural Computation, Springer, Berlin, Heidelberg, pp. 512–520, 2005.
27. M. Conde Amboage, W. González Manteiga, and C.A. Sánchez Sellero, *Quantile regression: estimation and lack-of-fit tests*, BEIO, Boletín de Estadística e Investigación Operativa, vol. 34, no. 2, pp. 97–116, 2018.
28. R. Koenker, *Quantile regression: 40 years on*, Annual Review of Economics, vol. 9, pp. 155–176, 2017.
29. P. Gaillard, Y. Goude, and R. Nedellec, *Additive models and robust aggregation for GEFCom2014 probabilistic electric load and electricity price forecasting*, International Journal of forecasting, vol. 32, no. 3, pp. 1038–1050, 2016.
30. M. Fasiolo, Y. Goude, R. Nedellec, and S.N. Wood, *Fast calibrated additive quantile regression*, <https://arxiv.org/abs/1707.03307>, 2017.
31. M. Fasiolo, S.N. Wood, M. Zaffran, R. Nedellec, and Y. Goude, *Fast calibrated additive quantile regression*, Journal of the American Statistical Associations, pp. 1–11, 2020.
32. M. Ghofrani, and A. Suherli, *Time series and renewable energy forecasting*, Time series analysis and applications, pp. 77–92, 2017.
33. A.P. Dawid, and P. Sebastiani, *Coherent dispersion criteria for optimal experimental design*, Annals of Statistics, pp. 65–81, 1999.
34. J. Aghaei, A. Ahmadi, H.A. Shayanfar, and A. Rabiee, *Mixed integer programming of generalised hydro-thermal self-scheduling of generating units*, Electrical Engineering, vol. 95, no. 2, pp. 109–125, 2013.
35. M.E. Lebotsa, C. Sigauke, A. Bere, R. Fildes, and J.E. Boylan, *Short term electricity demand forecasting using partially linear additive quantile regression with an application to the unit commitment problem*, Applied Energy, vol. 222, pp. 104–118, 2018.
36. *Monitoring renewable energy performance of power plants*, Tracking progress of 2017 (Issue 11) March 2018 by NERSA, Prepared by: Dr A Gxasheka, Mr H Mokhonoana, Ms R Mokoena.