Applied Mathematics & Information Sciences
An International Journal

http://dx.doi.org/10.18576/amis/130206

Bayesian Structural Time-Series Approach to a Long-Term Electricity Demand Forecasting

Paul Mokilane[1,2,*], *Pravesh Debba*[1,2], *Venkata S. S. Yadavalli* [3] *and Caston Sigauke*[4]

[1] Council for Scientific and Industrial Research, Pretoria, South Africa
[2] School of Statistics and Actuarial Science, University of the Witwatersrand, Johannesburg, South Africa
[3] Department of Industrial and Systems Engineering,School of Engineering, University of Pretoria, Pretoria, South Africa
[4] Department of Statistics, University of Venda, Thohoyandou, South Africa

Received: 5 Oct. 2018, Revised: 12 Nov. 2018, Accepted: 23 Nov. 2018
Published online: 1 Mar. 2019

Abstract: The paper presents an application of Bayesian structural time-series model to forecast long-term electricity demand. Accurate trend specification in long-term forecasting is important; otherwise erroneous forecasts could be obtained especially in South Africa where it is difficult to determine if the trend would continue a downward trajectory or would revert to an upward trajectory. Long-term probabilistic electricity demand forecasts in South Africa from 2013 to 2023 are presented in this paper. The findings are; (a) electricity demand in South Africa is less likely to exceed the highest historical hourly demand of 36 826 kW until 2023 (b) South African demand from Eskom is more likely to maintain the downward trend until 2023 (c) electricity demand lies between 15 849 kW and 39 810 kW with a 90% probability between 2013 and 2023. The contributions of the paper are; (a) application of BSTS to long-term electricity demand forecasting (b) use of autocorrelation plot to determine the number of time lags in long-term electricity demand forecasting (c) long-term forecasting of electricity demand using South African data with their uncertainties quantified.

Keywords: Bayesian, probabilistic forecasts, time series, uncertainties

1 Introduction

Electricity load is defined as an amount of electricity that balances the amount generated with that drawn from the electricity grid. In the absence of black-outs, load-shedding and availability of electricity generated from the renewable sources, electricity load is equivalent to electricity demand in South Africa. Therefore, in this study, the hourly electricity demand is defined as an amount o electricity in kW sent out every hour by Eskom to meet consumersdemand. Eskom is the main supplier of electricity to the South Africans and it generates approximately 95% of the total electricity consumed in South Africa while municipal-owned power plants and independent power producers (IPPs) generate the remaining 5% [27].

The economic growth of any country is dependent on its energy security. For planning purposes, it is important to ensure that long-term electricity demand is accurately forecasted, which would avoid wasted investment in the construction of excess generation facilities while ensuring that future electricity demand is sufficiently met. As indicated in [23], there is a need for probabilistic forecasting of hourly electricity demand in South Africa due to the inherent uncertainties related to long-term forecasting. Discussions of probabilistic forecasting can be found in [6], [19] and [21] while more information on probabilistic forecasting of peak electricity demand can be found in [7], [9] and [13]. The point forecasts of electricity demand represent the mean of the demand distribution and hence they inevitably under predict some peak electricity demand.

It must be accepted by forecasters and people using forecasts that all forecasts will inevitably differ with the actuals. Hong and Fan argue that researchers have long been pursuing the most accurate forecasts which resulted in researchers hoping to find the best forecasting technique which is a myth [8]. In search of the best forecasting technique, some researchers combined different modeling techniques to form hybrid models

* Corresponding author e-mail: PMokilane@csir.co.za

[14], [18] and [20]. According to Hong and Fan most of the hybrid techniques are of minimal value for load forecasting practice [8]. Hong and Fan pointed out that researchers and practitioners must understand that a universally best technique does not exist [8].

Forecasting, by nature, is a stochastic problem, but most of the utilities are still developing and using point forecasts, which are very difficult to defend, instead of using probabilistic forecasts [7]. One of the advantages of probabilistic forecasts is that they provide the full probability distribution of possible future electricity demand, and most importantly, the uncertainties in the forecasts are quantifiable.

Hourly electricity demand could be modelled through a Bayesian or a frequentist approach. In the frequentist approach, the attention is restricted to phenomena that are repeatable under identical conditions and the probability of an event is defined to be the limiting relative frequency with which the outcome would occur in n repetitions as n goes to infinity. The model parameters in this framework are considered to be fixed while the data are random. Under this framework, Quantile Regression (QR) is an attractive approach for probabilistic forecasting, as shown in [23] in which the demand was forecasted at the specified quantiles of the demand distribution.

A Bayesian approach on the other hand is a natural approach to probabilistic forecasting, where the data are considered to be fixed and the model parameters random. Under this framework, the model parameters are estimated with their full distributions. There are two approaches to Bayesian modelling which are full and empirical Bayes. In the full Bayes modelling, the prior parameters or the distribution of the prior parameters are assumed known while in the empirical Bayes modelling, the observed data are used to estimate the prior parameters and then proceeding as though the priors were known. In both Bayes approaches, the prior distribution represents what is known about the parameters before observing the data which can be based on past experience (expert information) of previous related studies. If the expert information is not available, then non-informative priors are used which give the data more weight in the posterior distribution. The next step is to collect data and form a likelihood function, and then the prior is combined with the likelihood function to form the posterior distribution.

Cottet and Smith applied Bayesian model to forecast intraday electricity load using a multi-equation approach [5]. They used a first-order Vector Autoregression (VAR) to model the errors and a Bayesian model selection methodology to explore the intraday correlation structure. Sigauke and Chikobvu used a Bayesian approach to forecast the short-term extreme electricity demand in South Africa [18]. According to Carstens, Xia and Yadavalli one of the disadvantages of point forecasts is that the confidence interval either contain the value, or it does not and the probability is either zero or one [28]. The confidence interval is therefore not associated with probability.

Short-term forecasts according to McSharry, Bouwman and Bloemhof are future values ranging from five minutes to one week ahead, and are used to ensure system stability, medium-term forecasts are mainly used for maintenance scheduling and they range from one week to six months ahead, while the long-term forecasts are used for capital planning and they range from six months to many years ahead [13]. There is a vast body of literature on forecasting of annual electricity demand as well as peak electricity demand in South Africa [2], [10], [11], [15], [18] and [22]. Maharaj and Yadavalli modelled the stochastic behaviour of the price dynamics of the electricity spot prices with a three-regime Markov switching model and they concluded that seasonality and uncertainties play important roles in predicting hourly spot prices [29].

The literature on Bayesian structural time series (BSTS) is very limited; however, a BSTS modelling approach would be ideal for forecasting long-term electricity demand in South Africa. This is because the electricity demand trend has shown changes in the recent past and therefore it may be difficult to make assumptions regarding the shape of the future electricity demand trend. The BSTS modelling approach makes it possible to determine the contribution of each component of the model and therefore it makes it possible to investigate the origin of the problem if results appear to be counter-intuitive.

The objective of this paper is to use the BSTS modelling approach to forecast the long-term (ten years) electricity demand in South Africa in the face of uncertainties which could among others emanate from increased technologies making use of electricity, population growth, general randomness in individual usage of electricity, prevailing economic patterns, change in weather conditions, escalating costs of electricity, use of power saving electrical appliances, the growing sources of renewable electricity and the possible market penetration of electric vehicles. To the best of our knowledge the BSTS has not been applied for forecasting the long-term electricity demand.

2 Methodology

The model parameters in the BSTS model are estimated through the Bayesian modelling approach. In general, the formulation followed in this study is that of Marin, Diez and Insua [12]. The posterior distribution of electricity demand in this study can be defined as in (1);

$$\pi(\theta|x) = \pi(x|\theta)\pi(\theta) / \int_\theta \pi(x|\theta)\pi(\theta)d\theta \qquad (1)$$

where; θ denotes the parameter space, that is $\theta = \theta_1, \theta_2, \ldots, \theta_n$ representing the set of demand for electricity parameters to be estimated; $x = x_1, x_2, \ldots, x_n$ are the drivers of demand for electricity; $\pi(x|\theta)$ is the function generating the demand for electricity data given the model parameters (the likelihood function), $\pi(\theta)$ is the prior distribution representing information on the model parameters before the electricity demand data are observed and $\int_\theta \pi(x|\theta)\pi(\theta)d\theta$ is a normalising constant which makes it difficult to compute posterior in a closed form, it gives evidence that the demand for electricity data is generated by the model, therefore the posterior distribution in (1) can be written as;

$$\pi(\theta|x) \propto \pi(x|\theta)\pi(\theta).$$

The posterior distribution summarises all available information about the parameters of the drivers of electricity demand after observing the demand for electricity data. The posterior mean of the parameters of the demand for electricity are estimated by;

$$E(\theta|x) = \int_\theta \theta\pi(\theta|x)d\theta \qquad (2)$$

which serves as the point estimate of the parameters θ given the data. The future values of (x) represented by (x_{n+1}) can also be predicted by using the posterior predictive distribution;

$$\pi(x_{n+1}|x) = \int_\theta \pi(x|\theta)\pi(\theta|x)d\theta \qquad (3)$$

The detailed formulation of the Bayesian model is given by Marin, Diez and Insua [12]. The complicated integrals in (1) make it difficult to compute the posterior distribution hence, the numerical method, the Markov Chain Monte Carlo (MCMC) is used to estimate the model parameters through sampling from the posterior distribution. The idea is to attain a chain whose equilibrium distribution matches that of the hourly electricity demand posterior distributin. The sampling stops when the chain reaches its equilibrium. In time series, De Jong suggested the simulation smoother procedure for simulating from the posterior density of the states given a Gaussian state space time series [25].

According to Brodersen, et al. and Scott and Varian the BSTS is formulated as a state space model [4] and [16]. The state space models are attractive because they are modular as the independent state components can be combined by concatenating the observation vectors Z_t [16]. There are two pieces to this model, that is, the observation equation that links the observed hourly demand for electricity (y_t) with the unobserved latent state (α_t). The observation equation is represented in (4).

$$y_t = Z_t^T \alpha_t + \varepsilon_t\ , \quad \varepsilon_t \sim N(0, H_t), \quad t = 1, 2, \ldots, 24 \quad (4)$$

where; y_t is the demand for electricity at hour t; α_t is a vector of latent variables (a set of unknown parameters); H_t and Z_t are structural parameters where H_t is a constant diagonal matrix with diagonal elements $\sigma_\varepsilon{}^2$ and Z_t is the observation vector. The transition equation is given in (5);

$$\alpha_{t+1} = T_t\alpha_t + R_t\eta_t\ , \quad \eta_t \sim N(0, Q_t), \quad t = 1, 2, \ldots, 24 \qquad (5)$$

where; T_t, R_t and Q_t are structural parameters and η_t is a lower dimension than α_t. Q_t is a constant diagonal matrix with diagonal elements $\sigma_u{}^2$, $\sigma_\omega{}^2$ and $\sigma_\varphi{}^2$. $\eta_t = (v_t, \varphi_t, \omega_t)$ are the independent components of the Gaussian random noise. The state space model representation of the BSTS model used in this study is presented in (6) and its components are represented by (7), (8) and (9);

$$y_t = \mu_t + \gamma_t + \beta^T X_t + \varepsilon_t \qquad (6)$$

$$\mu_t = \mu_{t-1} + \delta_{t-1} + \varphi_t \qquad (7)$$

$$\delta_t = \delta_{t-1} + v_t \qquad (8)$$

$$\gamma_t = -\sum_{s=1}^{s-1} \gamma_{t-s} + \omega_t \qquad (9)$$

where; y_t is the demand for electricity at hour t.

The μ_t is the local-level trend of the demand for electricity series. The trend is assumed to follow a random walk with φ_t representing some errors where; $\varphi_t \sim N(0, \sigma_\varphi{}^2)$. The prior is on $\sigma_\varphi{}^2$.

The γ_t is the seasonal component of the demand for electricity series. The seasonal model can be thought of as a regression on n seasons dummy variables with coefficients constrained to sum to one. If there are S seasons then the state vector γ_t is of dimension $S-1$. ω_t are seasonality errors, $\omega_t \sim N(0, \sigma_\omega{}^2)$ and the prior is on $\sigma_\omega{}^2$.

The ε_t is the irregular component, $\beta^T X_t$ is the regression component, the local-level term which defines how the latent state evolves over time and it is often referred to as an unobserved trend; the level is μ_t plus slope δ_t, and seasonal $S-1$ dummy variables with time varying coefficients that are expected to sum up to zero and y_t is the response series representing demand for electricity at hour t.

The ε_t is the irregular component (noise) and are assumed to follow an autoregressive process of order one (AR(1)) and it is distributed as $\varepsilon_t \sim N(0,\sigma^2)$. The prior coefficients for the AR(1) component are set at $\beta = 0$, $\sigma^2 = 1$ and the stationarity is ensured by setting $-1 < \phi_1 < 1$.

In this model, a time series component captures the general trend (μ_t), seasonality (γ_t) and irregular patterns (u_t) in the data. The priors for the time series component of the model are specified during the creation of the state specification of the model.

The $\beta^T X_t$ represents the regression component of the BSTS model where β represents the vector of regression coefficients of the demand for electricity model. The concept of sparse modelling asserts that if there are many possible predictors for the demand for electricity variable, some of them are expected to have zero coefficients. The natural way to represent sparsity in Bayesian modelling is through a spike and slab prior on the regression coefficients [16]. In this way, possible drivers of demand for electricity are selected. The prior is on the set of β coefficients.

Since the BSTS model in this paper has a regression component, the spike and slab priors are used to select the model variables. The spike part governs the probability of a given variable to be selected into the model (having a non-zero coefficient). The slab part shrinks the non-zero coefficients towards the prior expectation (often zero).

Let τ denotes a vector of ones and zeros where; one indicates that the variable is selected into the model (non-zero coefficient) and zero indicates that the variable is not selected into the model. If β_τ is the subset of elements of β where, $\tau_k = 1$ if $\beta_k \neq 0$ and $\tau_k = 0$ if $\beta_k = 0$, then the spike and slab prior for the electricity demand model can be written as in (10);

$$p(\tau,\beta,\sigma_\varepsilon^2) = p(\tau)p(\sigma_\varepsilon^2|\tau)p(\beta_\tau|\tau,\sigma_\varepsilon^2) \tag{10}$$

where; σ_ε^2 is the error variance, the marginal distribution $\rho(\tau)$ is the spike representing probability of choosing a given variable for inclusion into the demand for electricity model and it is assumed to follow a Bernoulli distribution given in (11);

$$p(\tau) = \prod_{j=1}^{J} {\pi_j}^{\tau_j}(1-\pi_j)^{1-\tau_j} \quad j = 1,2,\ldots,J \tag{11}$$

where; π_j is a prior probability of variable j being included into the demand model. For practical consideration, all π_j are assumed to have the same value π. If there are h non-zero predictors, then $\pi = h/k$, where k is the dimension of x_t. For the specific values of j, some variables can be forced to be included or excluded by setting $\pi_j = 1$ for inclusion and $\pi_j = 0$ for exclusion. For the *slab* part, a normal prior is used for β which leads to inverse gamma prior for σ_ε^2 and it is given by;

$$1/(\sigma_\varepsilon^2) \sim G(v_\varepsilon/2, s_\varepsilon/2), \tag{12}$$

where; v_ε is the prior sample size (shrinkage parameter) which can be thought of as a prior sample size for learning the variance parameter of each coefficient, and s_ε is the prior sum of squares.

$$s_\varepsilon/v_\varepsilon = (1-R^2)s_y^2 \tag{13}$$

where; s_y^2 is the sample variance and R^2 is the amount of variation explained by the regression model.

In this paper, the spike and slab priors for the regression component are set in such a way that all potential independent variables are given equal chances (50%) of being included in the model ($\pi_j = 0.5$). The prior for the overall variation explained by the model is set at 0.5 ($R^2 = 0.5$) and a shrinkage parameter is set at 1% ($v_\varepsilon = 0.01$). The full Bayes approach is used with a minor violation as s_y^2 is data-determined. The detailed formulation of BSTS is given by Brodersen, et al. and Scott and Varian [4] and [16]. Since this is a Bayesian model, the parameters are estimated through MCMC.

The time-lags in time-series analysis are important because values within time-series tend to correlate with previous values in the series, this is called autocorrelation. Using a regression-based technique like BSTS; it becomes a challenge to include time-lags in long-term forecasting and to determine the number of lags to be included. In this model, the number of time-lags is determined by using the autocorrelation plot in Figure 1. The highest correlations are at every 24th data point (Figure 1). The correlation cycle stops at the 168th point and the new cycle starts, hence the number of time-lags to include is the number of data points at every 24th point in the cycle (i.e. 24, 48, 72, 96, 120, 144 and 168). Each cycle represents a week. With one time-lag, only one future value can be forecasted, while two time-lags would only allow two future values to be forecasted. Therefore, n time-lags would enable n future values to be forecasted. If the first lag is n, then the next 6 lags are (n+24, n+48, n+72, n+96, n+120 and n+144). The time-lags are done in such a way that the shape of the autocorrelation pattern is maintained and the leap years coincide.

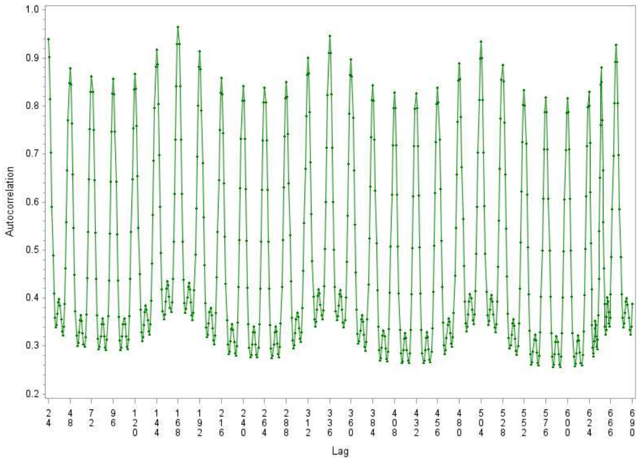

Fig. 1: Autocorrelation of hourly electricity demand

3 Data, Results and Discussion

3.1 Data

The South African total hourly electricity demand data from 1997 to 2015 are provided by Eskom. During this period, the highest hourly electricity demand is 36 826 kW and reached in 2011, while the minimum is 13 533 kW in 1998. The whole dataset provided is used in the modelling (data from 1997 to 2006 are used in the formation of time-lags).

The model is built from the logarithmically-transformed time-series. Logarithmic transformations are convenient means of transforming a highly-skewed variable into one that is approximately normal [1]. After carrying out a number of transformations of hourly electricity demand from the Box and Cox family, the logarithm is found to best fit the data [3]. Therefore, the logarithmic hourly electricity demand data is modelled. The electricity demand data between 2007 and 2012, inclusively, are used to train the model, while the data from 2013 to 2015 are withheld and used in model validation. The hourly electricity demand is forecasted from 2013 to 2023. The forecasts from 2013 to 2023 are out of sample forecasts.

A number of time-related variables, the Fourier series or harmonic transformation terms (to capture cycles inherent in the time-series) and time-lag variables are included as covariates to harness the correlation between hourly electricity demand data. All variables considered are given in Figures 7 and 8 in the appendix.

3.2 Model selection

The predictor variables in the regression component are selected using the spike and slab priors. The first step in variable selection is to estimate the coefficients of all potential regression models given the covariates. The averages of the coefficients of each variable from different models are then calculated. The regression component is dominated by eight variables (Friday, Monday, Saturday, Sunday, lag70200, lag70224 and lag70248) because their average coefficients are by far greater than the rest (Figure 5) in the appendix. Other variables have coefficients either close to zero or zero which would likely be insignificant.

The variables that are selected into all possible models are determined. In this paper, only variables that have got selected into all possible models (inclusion probability of one in Figure 8 in the appendix) are considered for modelling. The good model with trend component contributing less to the overall model is preferred. Fruhwirth-Schnatter discussed how to discriminate between different linear Gaussian state space models for a given time-series using the Bayesian approach which chooses the model that minimizes the expected loss [24]. In this way, a possible candidate Gaussian state space model for modelling a time-series could be selected.

3.3 Assessment of the forecasts

The forecasted daily profiles at the 50 th percentile of the electricity demand distribution in Figures 2 could be used to assess the accuracy of the forecasts. If the forecasted daily profiles capture the actual daily profiles over the years and the forecasts at the 50 th percentile are close to the actual electricity demand, then it would indicate that the model is forecasting well. The forecasted daily profiles are consistently compared well with the actual daily profiles as shown in Figure 2. The MAPE between the actual and the forecasts at the 50 th percentile of the electricity demand distribution are mostly below 4% in all hours, as shown in Table 1. Lewis indicates that a MAPE of less than 10% can be classified as highly-accurate forecast [26]. The MAPE in Table 1 are calculated from the untransformed hourly electricity demand data from 2013 to 2015.

3.4 Results and discussion

For illustration purposes, four days in June are selected in such a way that the day with the highest hourly electricity demand of the year is included with some days closer to it. The highest hourly electricity demand in 2013 is 35 393 kW on the 18 June while in 2014, it is 36 039 kW on the 12 June and it is 34 481 kW in 2015 on the 11 June

Table 1: Mean absolute percentage error (Forecasts at 50 th percentile)

Hours	2013	2014	2015	Average
0	3.714	2.190	4.093	3.332
1	3.531	2.221	4.349	3.367
2	3.456	2.320	4.841	3.539
3	3.603	2.324	4.702	3.543
4	4.282	2.569	3.809	3.553
5	6.331	4.691	4.162	5.061
6	5.970	4.699	4.722	5.130
7	4.135	2.673	4.266	3.692
8	4.299	2.337	3.361	3.332
9	4.303	2.315	3.036	3.218
10	4.480	2.367	3.437	3.428
11	4.226	2.337	3.798	3.454
12	3.927	2.362	4.340	3.543
13	3.840	2.463	4.753	3.685
14	3.829	2.717	5.065	3.870
15	4.127	2.881	4.921	3.976
16	4.501	3.012	4.443	3.985
17	4.523	3.187	4.020	3.910
18	4.099	3.333	3.399	3.611
19	3.951	2.997	3.125	3.358
20	4.593	2.771	3.442	3.602
21	4.653	2.479	3.845	3.659
22	4.272	2.256	3.815	3.448
23	3.941	2.189	3.912	3.347
Average	4.274	2.737	4.069	3.693

and they all happen at 18h00. The forecasted electricity demand profiles have well captured the actual electricity demand profiles. The interval between the black graphs in Figures 2, 3 and 4 represents the 90% credible intervals for the forecasted hourly electricity demand.

Electricity demand in Figure 2 has been forecasted to lie between 26 316 kW (logdemand = 4.4202) and 39 496 kW (logdemand = 4.5966) with a 90% probability on the 18 th June 2013 at 18h00 while the actual electricity demand is 35 393kW (logdemand = 4.5489) and the forecasted electricity demand at the 50 th percentile is 32 126 kW (logdemand = 4.5065). The blue graph represents the forecasted hourly electricity demand at the 50 th percentile of the demand distribution while the red circles represents the actual hourly electricity demand. The black graphs represent the lower and upper 90% credible intervals for hourly electricity demand.

By comparing electricity demand density functions over the years, insight into expected shifts in electricity demand patterns can be obtained, that is, whether the distribution of electricity demand in Figure 5 is expected to shift towards higher or lower electricity demand over the years until 2023. The forecasted electricity demand distributions in Figure 3 for the period investigated, suggest that electricity demand from Eskom is likely to shift towards lower demand over the years until 2023. There are some signs of electricity demand shifting towards lower demand in the historical data in the last four years until 2015 which could be attributed to the following: the growing renewable sources of electricity in

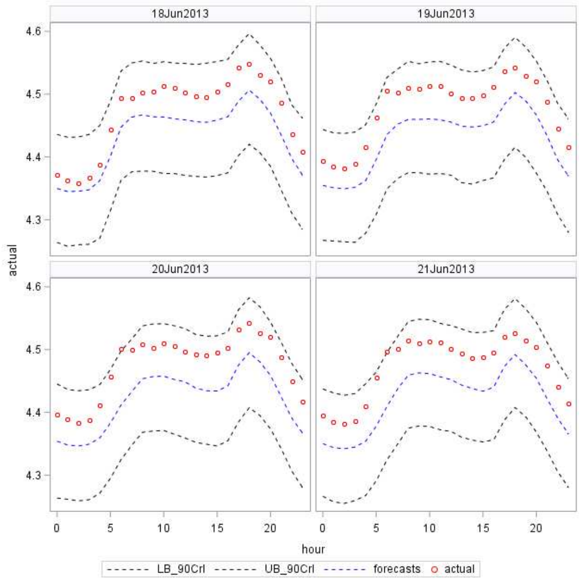

Fig. 2: Comparison of actual and forecasted hourly demand at 50 th percentile with 90% credible interval: 2013 actual hourly demand against forecasted

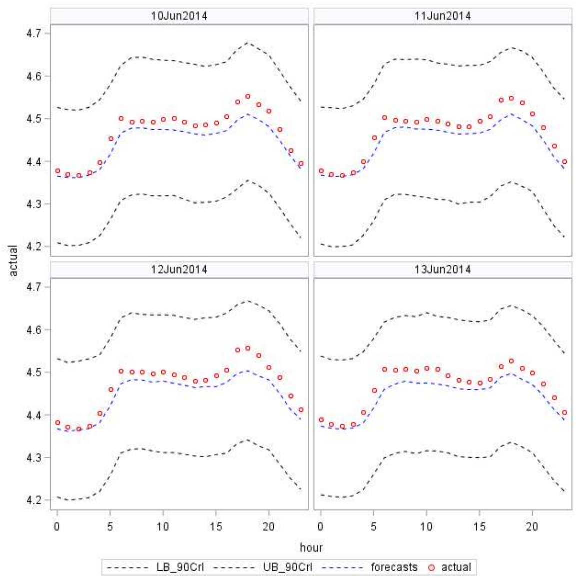

Fig. 3: Comparison of actual and forecasted hourly demand at 50 th percentile with 90% credible interval: 2014 actual hourly demand against forecasted

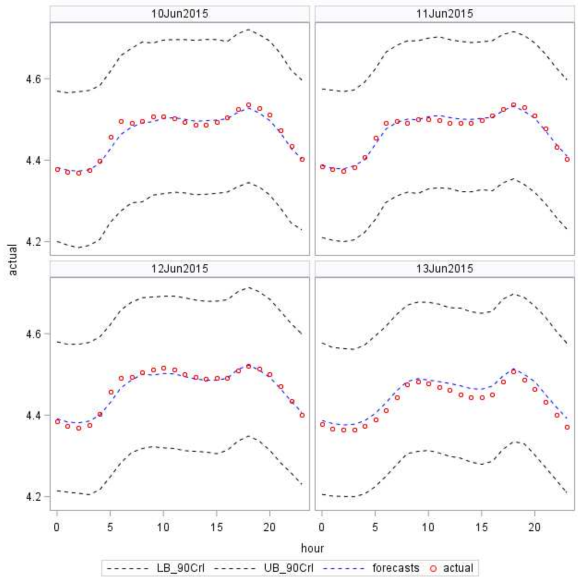

Fig. 4: Comparison of actual and forecasted hourly demand at 50 th percentile with 90% credible interval: 2015 actual hourly demand against forecasted

South Africa, the sluggish economic growth, the steep increases in electricity tariffs in the recent passed, and the market penetration of energy efficient appliances.

The density functions in Figure 5 could also be used to calculate the probabilities of exceeding certain electricity demand values by integrating the density functions. For example, to calculate the probability of exceeding the highest historical hourly electricity demand in future and to determine if demand can be met with the existing generation infrastructure.

Figure 6 represents the full forecasted distribution of hourly electricity demand in South Africa with its 90% credible interval. The actual hourly electricity demand from 2007 to 2012 (training data) are represented by the black graph; while the blue graph represents the forecasted posterior mean of the hourly electricity demand distribution between 2013 and 2023. The interval represented by the green graphs is the 90% credible interval of the forecasted hourly electricity demand.

Figure 6 shows that electricity demand during the period between 2013 and 2023 would lie between 15 849 kW (logdemand = 4.2) and 39 810 kW (logdemand = 4.6) with 90% probability. The blue graph in Figure 6 indicates that the hourly electricity demand on average is forecasted to lie between 18 000 kW and 35 000 kW between 2013 and 2023 in South Africa.

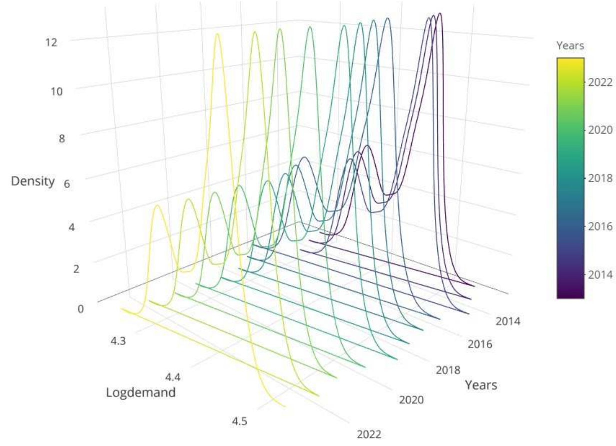

Fig. 5: Comparisons of density functions of electricity demand over years

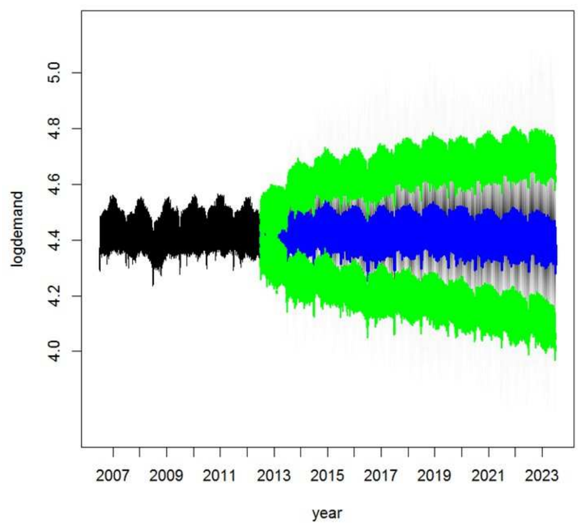

Fig. 6: Predicted distributions of demand with 90% credible interval

3.5 Managerial implications of electricity demand in South Africa

The long-term electricity demand is forecasted under uncertainties emanating from the increased technologies making use of electricity, population growth, general randomness in individual usage of electricity, seasonal effects, prevailing economic patterns, change in weather conditions, escalating costs of electricity, use of power saving electrical appliances, the growing sources of renewable electricity (whose data are not adequately collected in South Africa) and the market penetration of electric vehicles among others. Probabilistic modelling accounts for uncertainties in electricity demand forecasting; hence, enabling decision makers to make better decisions in the face of such uncertainties. The decisions could have far-reaching consequences because the decision for infrastructure expansion could result in the construction of unnecessary power-generating facilities while the decision without expanding on the current infrastructure could result in failure to meet future electricity demand. For planning purposes, probabilistic forecasts of long-term electricity demand can help managers to investigate a range of possible future electricity demand values, and so that the more informed decisions regarding the need for more (or fewer) power stations could be taken.

4 Conclusions

The hourly electricity demand forecasts have shown that it is unlikely for future electricity demand from Eskom to exceed the highest historical electricity demand of 36 826 kW. The electricity demand is likely to shift towards smaller demand over the years until 2023 and it is estimated to lie between 15 849 kW (logdemand = 4.2) and 39 810 kW (logdemand = 4.6) during the period between 2013 and 2023 with a 90% probability.

One of the advantages of the BSTS model is that it provides the contributions of individual components (e.g. trend, seasonal and regression components) of the overall model as standard output. For the South African electricity demand data, it would be ideal for the trend component to contribute less to the model as there are uncertainties with regard to whether the trend would continue the apparent decline trajectory as shown in the last four years until 2015 or if it would stabilise or would revert to the old upward trajectory going forward. If the trend component contributes more to the model and it were wrongly specified then the long-term electricity demand forecasts could vastly deviate from the actual electricity demand and this could have serious planning implications.

The BSTS model uses the spike and slab priors to minimise the correlation between variables in the model if it exists and this is a powerful way of solving the problem of multicollinearity and the resultant model is always parsimonious.

One of the main advantages of the probabilistic forecasts is that the planner could assess the chances of exceeding the current electricity generation capability of the system in future. If there are such chances, then the probability of exceeding the maximum that the system can generate currently can be calculated.

The BSTS model has a huge potential of accurately forecasting the long-term electricity demand especially in the case of a complicated electricity demand trend such as that of the South African electricity demand data. The uncertainties are quantifiable in the BSTS modelling approach. Therefore, BSTS modelling approach is recommended for forecasting the long-term electricity demand in South Africa.

Acknowledgement

The research was conducted during a contract research project funded by Eskom and some funding from Thuthuka fund of the National Research Fund (NRF). I would therefore, like to thank Eskom and the NRF for funding the research. I would also like to thank the following Eskom staff members for supplying the data; Moonlight Mbata and Ntokozo Sigwebela. I would like to thank the following CSIR staff members for their technical support; Nontembeko Dudeni-Tlhone, Nosizo Sebake, Renee Koen, Jenny Holloway and Sibusisiwe Makhanya.

The authors are grateful to the anonymous referees for a careful checking of the details and for helpful comments that improved this paper.

References

[1] K. Benoit, Linear regression models with logarithmic transformations, London School of Economics, London, Vol. 22, pp. 23-36 (2011).

[2] H. Boraine and V.S.S. Yadavalli, Electricity load forecasting with artificial neural networks, South African Journal of Industrial Engineering, Vol. 14, No. 2, pp. 23-35 (2003).

[3] G.E. Box and D.R. Cox, An analysis of transformations, Journal of the Royal Statistical Society, Series B (Methodological), Vol. 26, pp. 211-253 (1964).

[4] K. H. Brodersen, F. Gallusser, J. Koehler, N. Remy, S.L. Scott and Others, Inferring causal impact using Bayesian structural time-series models, The Annals of Applied Statistics, Vol. 9, No. 1, pp. 247-274 (2015).

[5] R. Cottet and M. Smith, Bayesian modeling and forecasting of intraday electricity load, Journal of the American Statistical Association, Vol. 98, No. 464, pp. 839-849 (2003).

[6] T. Gneiting and M. Katzfuss, Probabilistic forecasting, Annual Review of Statistics and Its Application, Vol. 1, pp. 125-151 (2014).

[7] T. Hong, J. Wilson and J. Xie, Long term probabilistic load forecasting and normalization with hourly information, IEEE Transactions on Smart Grid, Vol. 5, No. 1, pp. 456-462 (2014).

[8] T. Hong and S. Fan, Probabilistic electric load forecasting: A tutorial review, International Journal of Forecasting, Vol. 32, No. 3, pp. 914-938 (2016).

[9] 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).

[10] R. Inglesi-Lotz, The evolution of price elasticity of electricity demand in South Africa: A Kalman filter application, Energy Policy, Vol. 39, No. 6, pp. 3690-3696 (2011).
[11] R. Koen , T. Magadla and P. Mokilane, Developing long-term scenario forecasts to support electricity generation investment decisions, In 43rd Annual Conference of the Operations Research Society of South Africa, pp. 9 (2014).
[12] J.M. Marin, R.M. Diez and D.R. Insua, Bayesian methods in plant conservation biology, Biological Conservation, Vol. 113, No. 3, pp. 379-387 (2003).
[13] P.E. McSharry, S. Bouwman and G. Bloemhof, Probabilistic forecasts of the magnitude and timing of peak electricity demand, IEEE Transactions on Power Systems, Vol. 20, No. 2, pp. 1166-1172 (2005).
[14] H. Mori and A. Takahashi, Hybrid intelligent method of relevant vector machine and regression tree for probabilistic load forecasting, In Innovative Smart Grid Technologies (ISGT Europe), 2nd IEEE PES International Conference and Exhibition, IEEE, pp. 1-8 (2011).
[15] P. Debba, R. Koen, J.P. Halloway, T. Magadla, M. Rasuba, S. Khuluse and C.D. Elphinstone, Forecasts for electricity demand in South Africa (2010-2035) using the CSIR sectoral regression model, Pretoria (2010).
[16] S.L. Scott and H.R. Varian, Predicting the present with bayesian structural time series, Available at SSRN 2304426, (2013).
[17] C. Sigauke, Modelling electricity demand in South Africa, (PhD thesis, University of the Free State), (2014).
[18] C. Sigauke and D. Chikobvu, Prediction of daily peak electricity demand in South Africa using volatility forecasting models, Energy Economics, Vol. 33, No. 5, pp. 882-888 (2011).
[19] R. Weron, Electricity price forecasting: A review of the state-of-the-art with a look into the future, International Journal of Forecasting, Vol. 30, No. 4, pp. 1030-1081 (2014).
[20] T. Xiong, Y. Bao and Z. Hu, Interval forecasting of electricity demand: a novel bivariate EMD-based support vector regression modeling framework, International Journal of Electrical Power and Energy Systems, Vol. 63, pp. 353-362 (2014).
[21] Y. Zhang, J. Wang and X. Wang, Review on probabilistic forecasting of wind power generation, Renewable and Sustainable Energy Reviews, Vol. 32, pp. 255-270 (2014).
[22] E. Ziramba, The demand for residential electricity in South Africa, Energy Policys, Vol. 36, No. 9, pp. 3460-3466 (2008).
[23] P. Mokilane, J. Galpin, V.S.S. Yadavalli, P. Debba, R. Koen and S. Sibiya, Density forecasting for lon-term electricity demand in South Africa using quantile regression, South African Journal of Economic Management and Sciences, Vol. 21, No. 1, pp. 1-14 (2018).
[24] S. Fruhwirth-Schnatter, Bayesian model discrimination and Bayes factors for linear Gaussian state space models, J.R. Statist. Soc. B, pp. 237-246 (1995).
[25] P. De Jong, The simulation smoother for time series models, Biometrika , Vol. 82, No. 2, pp. 339-350 (1995).
[26] C.D. Lewis, Industrial and business forecasting methods: A radical guide to exponential smoothing and curve fitting, Colin David, (1982).
[27] Deloitte, An overview of electricity consumption and pricing in South Africa, An analysis of the historical trends and policies, key issues and outlook in 2017 , http://www.eskom.co.za/Documents/EcoOverviewElectricitySA-2017.pdf, (2017).
[28] H. Carstens, X. Xia and V.S.S. Yadavalli, Bayesian Energy Measurement and Verification Analysis, Energies, Vol. 11, No. 2, (2018).
[29] Y. Maharaj and V.S.S. Yadavalli, Modelling electricity spot prices with a three-regime Markov model, Industrial Engineering and Engineering Management (IEEM), IEEE International Conference, pp. 2153-2158 (2017).

4.1 Appendix

The regression component of the model was dominated by eight variables (Friday, Monday, Saturday, Sunday, weekday, lag70200, lag70224 and lag70248) because their average coefficients were by far greater than the rest (Figure 7). Other variables had coefficients either close to zero or zero which would likely be insignificant.

Figure 8 represented the model variables with their proportions of all possible models in which they would be selected given the variables under consideration. There were 20 variables that were selected in 100% of all possible models which indicated that, given the variables under consideration, 20 of them would be included in all possible models. In this study, only variables that appeared in all possible models were considered for modelling resulting into a parsimonious model.

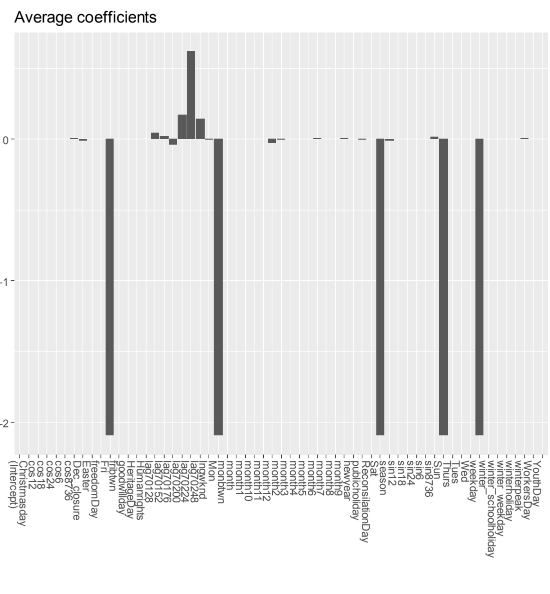

Fig. 7: Average coefficients of potential variables

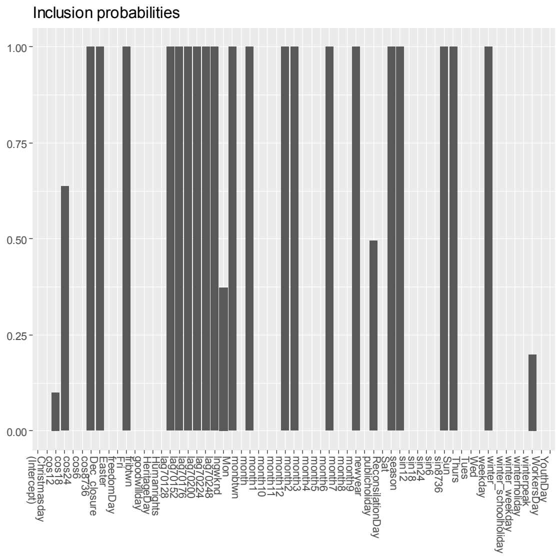

Fig. 8: Spike-and-slab variable inclusion probabilities

Paul Mokilane is a Senior Statistician in Spatial Planning and Systems within the Built Environment at the Council for Scientific and Industrial Research (CSIR). He is a PhD candidate at the University of the Witwatersrand in the School of Statistics and Actuarial Sciences. He serves in the board of the Institute of Certificated and Chartered Statisticians of South Africa (ICCSSA) and he is a member of the South African Statistical Association (SASA).

Pravesh Debba is currently the Manager of Spatial Planning and Systems within the Built Environment at the Council for Scientific and Industrial Research (CSIR). Prior to joining CSIR in 2008, he was an Associate Professor in Statistics at the University of South Africa (UNISA). He holds a PhD in Statistics from the International Institute for Geo-Information Science and Earth Observation (ITC) and Wageningen University in the Netherlands and an MSc degree in Biostatistics from the University of Hasselt in Belgium. He has been visiting professor in the School of Statistics and Actuarial Sciences at the University of the Witwatersrand (Wits) since 2015 and has successfully co-supervised several masters and currently co-supervising several doctoral students in Statistics. He has served as president of the South African Statistical Association (SASA), and as chairperson of the Institute of Certificated and Chartered Statisticians of South Africa (ICCSSA), elected member of the International Statistical Association (ISI) and editorial board member of the Spatial Statistics Journal published by Elsevier. He has published around 40 research papers in journals, conference proceedings or book chapters.

Venkata S.S. Yadavalli is a Professor and Head of Department of Industrial & Systems Engineering, at University of Pretoria, South Africa. He obtained his Ph D from Indian Institute of Technology, Chennai, in 1983. His main areas of interest include, Stochastic modeling towards, Reliability, Queuing, Inventory, Manpower planning systems and Finance. He has published over 150 research papers in these areas. He published articles in various national and international journals including, IEEE Transactions on Reliability, Applied Mathematics and Information Sciences, Computers & Industrial Engineering, Stochastic Analysis and Applications, International Journal of Systems Science, International Journal of Production Economics, International Journal of Production Research, Applied Mathematics & Computation, etc. He was past Vice-President and President of Operations Research Society of South Africa. He is in the editorial board of various national and international journals. He is a Fellow of the South African Statistical Association. He received various awards from various professional bodies like IEOM (Industrial Engineering & Operations Management, USA), South African Institute of Industrial Engineers, South African Institute of Mining and Metallurgy etc. for his contributions towards research, and ICQRIT.

Caston Sigauke holds a PhD in Statistics from the University of the Free State and an MSc in Operations from the National University of Science and Technology in Zimbabwe. He is a member of the: International Institute of Forecasters (IIF), Member ID: 7178; Operations Research Society of South Africa (ORSSA), Member ID 511 and South African Statistical Association (SASA). His fields of professional expertise are lecturing, research and consulting. His fields of expertise are forecasting and time series, statistics of extremes, statistical learning and modelling, exploratory data analysis with applications in probabilistic forecasting in energy and environmental systems. Current research is on probabilistic electricity demand forecasting and solar energy modelling.