Renewable and Sustainable Energy Reviews 73 (2017) 1104–1122

Contents lists available at ScienceDirect

Renewable and Sustainable Energy Reviews

journal homepage: www.elsevier.com/locate/rser

A review and analysis of regression and machine learning models on commercial building electricity load forecasting

B. Yildiz, J.I. Bilbao*, A.B. Sproul

School of Photovoltaics and Renewable Energy Engineering, University of New South Wales, Sydney, NSW 2052, Australia

ARTICLE INFO

Keywords:
Short term load forecasting for commercial buildings
Review of regression models
Machine learning
Neural Networks
Support Vector Regression
Regression Trees

ABSTRACT

Electricity load forecasting is an important tool which can be utilized to enable effective control of commercial building electricity loads. Accurate forecasts of commercial building electricity loads can bring significant environmental and economic benefits by reducing electricity use and peak demand and the corresponding GHG emissions. This paper presents a review of different electricity load forecasting models with a particular focus on regression models, discussing different applications, most commonly used regression variables and methods to improve the performance and accuracy of the models. A comparison between the models is then presented for forecasting day ahead hourly electricity loads using real building and Campus data obtained from the Kensington Campus and Tyree Energy Technologies Building (TETB) at the University of New South Wales (UNSW). The results reveal that Artificial Neural Networks with Bayesian Regulation Backpropagation have the best overall root mean squared and mean absolute percentage error performance and almost all the models performed better predicting the overall Campus load than the single building load. The models were also tested on forecasting daily peak electricity demand. For each model, the obtained error for daily peak demand forecasts was higher than the average day ahead hourly forecasts. The regression models which were the main focus of the study performed fairly well in comparison to other more advanced machine learning models.

1. Introduction

According to reports from the International Energy Agency (IEA) [1], the commercial building sector accounts for 32% of the final electricity consumption in OECD countries. In particular, this number was reported as 29% for European countries and in the USA, more recent reports showed that commercial buildings accounted for over 35% of end-use electricity consumption [2]. In Australia, commercial buildings accounted for around 30% of the electricity end-use consumption [1] and 10% of the total greenhouse gas emissions in 2013 [3]. Most of these buildings have inefficiencies in energy use due to their physical nature. The Rocky Mountain Institute has stated that there is the potential to reduce commercial building energy use by 20% in the USA and other reports indicate that there is a reduction potential of about 29% [4]. These numbers suggest the importance of focusing efforts on understanding and reducing the energy use and demand of commercial buildings. Furthermore, it's well known that reducing peak electricity demand is a clear pathway to achieve economic and environmental benefits. For example, peak demand is identified as the main driver for the growing investments in network infrastructure which exerts upward pressure on electricity prices [5]. The Energy Supply Association of Australia estimates that 80% of the investment in grid upgrades was required to meet the growing peak demand in Sydney [6]. Hence, the implementation of accurate and robust electricity load forecast methods both at distribution network and end user levels can assist demand management and energy efficiency activities which can be considered as alternative solutions to electricity network augmentation [7].

Commercial buildings equipped with modern monitoring and metering systems along with building management systems are well

Abbreviations: AR, Auto Regressive; MA, Moving Average; ARMA, Auto Regressive Moving Average; ARIMA, Auto Regressive Integrated Moving Average; ANN, Artificial Neural Network; NARX, Nonlinear Autoregressive Network with Exogenous Inputs; SVM, Support Vector Machine; SVR, Support Vector Regression; SLR, Single Linear Regression; MLR, Multivariate Linear Regression; PRISM, The Princeton Scorekeeping Method; R^2, Coefficient of Determination; R_{adj}^2, Adjusted Coefficient of Determination; CV, Coefficient of Variance; RMSE, Root Mean Squared Error; CV-RMSE, Percentage RMSE by the mean; MAPE, Mean Absolute Percentage Error; MPE, Mean Percentage Error; TMY, Typical Meteorological Year; DDCAV, Dual Duct Under Constant Air Volume; DDVAV, Dual Duct Under Variable Air Volume; TRCAV, Terminal Reheat Under Constant Air Volume; TRVAV, Terminal Reheat Under Variable Air Volume; DBT, Dry Bulb Temperature; T_{dp}, Dew Point Temperature; RH, Relative Humidity; q_{sol}, Solar Heat Gains; q_i, Sensible Heat Gains; I, Indicator Variable; WWR, Window to Wall Ratio; UNSW, University of New South Wales

* Corresponding author.
E-mail address: j.bilbao@unsw.edu.au (J.I. Bilbao).

http://dx.doi.org/10.1016/j.rser.2017.02.023
Received 29 March 2016; Received in revised form 2 December 2016; Accepted 4 February 2017
Available online 11 February 2017
1364-0321/

suited to implement electricity load reduction activities. Moreover, potential economic benefits brought about by reducing the demand can become more significant for prosumers – customers who produce as well as consume energy. However, energy systems in commercial buildings can be complex systems, particularly in buildings with large heating, ventilation, and air-conditioning (HVAC) systems. This complexity makes the evaluation and forecasting of electricity demand quite challenging. The main cause of the difficulty can be attributed to the variation in the energy consumption profiles within buildings [8]. The problem increases in buildings that have a mix of areas which have different HVAC and lighting requirements such as offices, laboratories, lecture theatres, operating theatres, event rooms, and data centre and manufacturing facilities. In addition, building electricity loads vary with internal factors such as occupancy and scheduling. Last but not least, building loads are also susceptible to the changes in external weather parameters such as temperature, solar radiation and humidity.

Because of the reasons outlined above, numerous attempts have been made to accurately forecast commercial building electricity loads. Different techniques such as thermal models, statistical regression models, time series models and machine learning models have been used in forecasting commercial building electricity loads for various climates and time horizons (short-term, mid-term and long-term). Shor-term forecasts (minutes to a week ahead) can have an immediate impact on a building's operation and scheduling and is a crucial component for building energy management systems. Mid-term (a week to a year ahead) and long-term (more than a year ahead) forecasts have greater importance for longer term planning. Existing review papers have provided a good foundation for classifying the work done in terms of model types, forecast horizon and scale (single building to regional or national level) [9–13], while some articles provide a particular focus on certain methods [11,14,15].

Among the different forecasting methods, regression models are simple to develop, use, and interpret, in comparison to other more complex techniques hence, they have been commonly used for building load forecasting [9,16]. Regression models correlate the energy consumption with external weather and internal building parameters. These models can be developed by using real historical load data [16–20] or simulated load data [21–25]. To the authors' knowledge, there hasn't been a detailed review particularly focused on regression models in commercial building load forecasting, although there have been comprehensive studies where regression models were used in commercial building electricity load forecasting and its performance was compared with other methods [26–28]. Our study therefore presents a thorough review on regression models and aims to inform the reader about the range of different applications where these models can be successfully used. For clarity, regression models are classified under different categories such as their area of application, commonly used input parameters, methods to improve model performance and comparison with other models.

In order to extend our study beyond theory, regression models are implemented in a day ahead hourly electricity load forecast analysis by applying the methodologies discussed in the review section. Studies where regression models were used for commercial building electricity load forecasting and their performances were compared with other methods, were mainly limited to single building level and the results were only analysed for the overall data-set. This study extends the analysis by implementing forecast models for different scales: single building and university Campus level which allows us to observe the impact of load scale on forecast performance. Furthermore, our study also allows the comparison of model performance for different seasons. The analysed models are trained and tested not only for the overall data set but also with seasonal sub-sets. In addition to the day ahead hourly load forecasts, the analysis is broadened to forecast daily peak loads by modifying the models which gave another opportunity to compare model performance on different target loads.

The remainder of the paper is organized as follows. In Section 2, a brief review of models apart from regression models is introduced. In Section 3, a detailed review on the use of regression models is presented. Following the review sections, in Section 4 the implementation of regression models for forecasting day ahead hourly and daily peak electricity loads, for a single university building (Tyree Energy Technologies Building – TETB) and a university Campus with around 50 buildings on a 38 ha site (Kensington Campus – UNSW) is discussed. Following the regression analysis, four other machine learning models are used for the forecast analysis: Artificial Neural Networks (ANN) with Levenberg Marquitd (LM) and Bayesian Regulation (BR) Backpropagation, Nonlinear Autoregressive Network with Exogenous Inputs (NARX) with LM and BR Backpropagation, Regression Trees (RT) and Support Vector Regression (SVR). ANN and SVR are the most commonly used machine learning models in the area [9,14]. Implementation of regression trees is also not uncommon [26,27,29] whereas NARX is a relatively new method and to our knowledge its implementation has been limited, thus we wanted to compare this method with other popular machine learning models. The performance, ease of use, and interpretability of the models are compared in Section 5.

2. Models used in commercial building electricity load forecasting

2.1. Thermal models

Thermal models calculate heat transfers and energy behaviour on a sub or whole building level [9]. Heat transfer calculations are based on the interaction of the building envelope with internal and external environments. Comprehensive thermal models may require a high number of inputs in comparison to simpler models. Historical data is typically not required for these models [30].

Analytical thermal modelling software such as DOE-2, Energy Plus, BLAST, and ESP-r has been developed for evaluating energy consumption and efficiency in buildings. This type of software has been widely used for developing building energy standards and analysing energy consumption and conservation measures in buildings. Although these models are quite powerful, they require detailed data regarding the building envelope, external weather, occupant behaviour and interior equipment performance, which may not be accessible to some users [30]. Section 3.1 gives examples of regression models developed by using thermal model simulation data.

2.2. Auto regressive models

Auto regressive models analyse sets of data points in a time series and correlate the future value of a certain variable with its past values [30]. It is possible to correlate different input variables to the output such that in commercial building load forecasting, load can be correlated with other important weather and building parameters. For example, Espinoza et al. [31] used a periodic auto regression model for 245 substations where each substation had four years of hourly data points. The method yielded satisfactory results for short term forecasting (225 substations out of 245 showed an R^2 higher than 90%).

One of the most commonly used forecasting techniques amongst the time series models is based on the Box-Jenkins methodology, which combines the Auto Regressive (AR) order p and Moving Average (MA) order q of the time series. The model is called ARIMA when an additional differencing order d is integrated into the model in order to remove the possible non-stationarities within the data [30]. A study done by Amjady [32] uses a novel Box Jenkins method for short term and peak load forecasts. This modified ARIMA method uses an initial forecast input and combines it with temperature and load data for the regression analysis. The method can accurately forecast hourly and peak loads and gives better results than the standard ARIMA model (for three operators located in different climatic zones of Iran, the hourly forecast MAPE was between 2.2% and 4.3% for the ARIMA model and 1.5% and 2.0% for the modified ARIMA).

2.3. Machine learning models

2.3.1. Artificial Neural Networks (ANN)

Artificial Neural Networks are often referred to as the most widely used machine learning models in building load forecasting [11]. ANN models are effective when dealing with complex, nonlinear problems. ANN's are used for estimating heating/cooling loads, total electricity consumption and sub-level components operation and optimization.

A study by Neto and Fiorelli [8] compared a simple ANN model with a thermal model developed in Energy Plus, for forecasting hourly loads in a university administration building. The thermal model daily MAPE was less than 13% whereas the ANN model performed with a daily MAPE of 10%. The study further compared the simple ANN model (temperature was taken as the only influence parameter) with a more complex ANN model (temperature, relative humidity and solar radiation were taken as influence parameters). The complex ANN had an average MAPE of 9.5% which was slightly better than the simpler ANN. It was also concluded that the effects of humidity and radiation on energy consumption were less significant than those of external temperature for buildings with air conditioning. A more recent study by Huang et al. [33] used an ANN model to model HVAC energy of an airport terminal building with four thermal zones. In contrast to single zone approaches, the study was capable of taking dynamic heat transfer interactions between multi-zones into account by the use of NARX model. The proposed method and single zone approach showed an average RMSE of 0.32 °C and 0.38 °C respectively for an hour up to a week ahead, zone temperature forecasts. The higher accuracy of the temperature forecasts led to daily energy savings over the entire building of 28%.

2.3.2. Support Vector Machines (SVM)

SVMs are highly effective models in machine learning and have the capability of solving non-linear problems, even with small quantities of training data [9]. They can be used both for classification and regression problems where the latter case is called Support Vector Regression (SVR).

SVMs have an advantage over other machine learning methods such as ANNs, due to their ability to locate global minima rather than local minima in the solution space [34]. In a recent study, Borges et al. [35] compared different machine learning methods for three different commercial buildings in the Eastern Slovakian region. Although the authors concluded that there is no one best model to fit every scenario, the SVR model gave more accurate day ahead hourly forecast results than the ANN model for each of the three buildings after bias correction. Furthermore, Chen et al. entered the EUNITE (European Network on Intelligence Technologies for Smart Adaptive Systems) competition for mid-term electricity load forecasting and won the competition by using a SVR model [36].

2.3.3. Regression Trees

Regression Trees is another machine learning model which can be preferred over linear regression models when the data has many features which interact in complicated and nonlinear ways. Linear regression models use a global predictive formula holding the entire data-space whereas Regression Trees sub-divide the space into smaller regions and further partition the sub-divisions and assigns to its nodes (leaves) where interactions are more manageable. Regression Trees regress decisions in a tree form, starting from the root node down to a leaf node where the leaf node contains the response [37].

Fan et al. used various models for predicting day ahead hourly and peak electricity loads of a commercial building in Hong-Kong. Regression trees were the best performing model over other regression, machine learning, and time series models [27].

3. Literature review on regression models

Regression models are statistical methods for estimating the relationship between the output and the variables which have influence on the output, also referred to as influence parameters. An example of a regression equation is given below:

$$\hat{y} = a_1 \times x_1 + a_2 \times x_2 + \ldots + a_n \times x_n \tag{1}$$

$$y = \hat{y} + e \tag{2}$$

where, y is the real output, $\hat{y}$ represents the regression model output, x_1 to x_n represent the influence parameters, a_1 to a_n represent the coefficients for the corresponding influence parameters, and e is the associated error term.

Data for y and x can be obtained from historical values or can be simulated by thermal modelling software. Usually, the objective of the regression model is to minimize the sum of squared errors by varying the coefficients a_1 to a_n. For electricity load forecasting in buildings, regression models correlate a relationship between the historical values of the load with the influence parameters to predict the future value of the load. The models can vary with the number and selection of influence parameters (single or multivariate regression models), inclusion of change point parameters (single, multivariate or change point regression models), forecasting horizon (yearly, monthly, daily, and hourly forecasts) and selection of the data (historical data or simulated data).

Regression models are widely used not only for load forecasting [16,20,23,38,39] but also for monitoring building energy consumption, measurement and verification of energy efficiency methods [40–42], identifying operation and maintenance (O & M) problems, and analysing HVAC system [16,18].

3.1. Regression models based on historical data

Regression models are built based on real data whenever historical data is available for the electricity load and other influence parameters such as weather variables (temperature, humidity, solar radiation and wind). A frequent hypothesis states that the load forecasting models should use real data when it is available, otherwise the evaluation of energy consumption might be highly under or overestimated [8].

A multivariate regression model (MLR) based on historical data was developed by Ramanathan et al. [20] for short term electricity load and peak forecasting. With the given historical load and weather data, a number of MLR models were developed for each hour of the day. The forecasts were made from 16 to 40 h into the future. The MLR model has the form of:

$$\begin{aligned} Load_{hour1} = {} & (a \times Deterministic) + (b \times Temperature) + (c \times Load) \\ & + (d \times PastErrors) + e \end{aligned} \tag{3}$$

The '*a, b, c*, and *d*' terms represent the constants of the influence parameters, '*e*' represents the residual error and the '*Deterministic*' parameter refers to variables such as year, month, week of the month and day of the week. The MAPE of the 24 individual hourly regression models varied between 4.04% and 5.66% for the given five month period.

Another study was carried to improve the regression model used for the short term system load and peak forecasting by the Pacific Gas and Electric Company, San Francisco, California [38]. The initial model consisted of an ARIMA peak model and a MLR peak model which used historical data for the last 15 days. The improved model removed the ARIMA peak forecast model and only relied on a MLR model which produces a daily peak forecast and uses it as an input for the hourly forecast. The model daily mean percentage error (MPE) ranged between 0.25% and 0.32% for the weekday peak loads whereas the new model improved the MPE range to −0.04% and 0.09%.

A recent study by Fan et al. [27] used a MLR model along with eight other models to predict the electricity load of the tallest commercial building in Hong-Kong. One year of historical hourly electricity load and 12 other climate variables were used to develop the models. The MLR model showed 4.23% and 6.08% MAPE for the total daily and next-day peak load forecasts respectively and performed better than the ARIMA and ANN models, but worse than five other machine learning and time series models. The results also showed that the MLR model required the least computation time by far compared to the other methods.

Another recent study done by Braun et al. [17] developed regression models for electricity and gas consumption of a supermarket located in the UK, in order to predict its long term energy consumption. The regression models were developed by using weather and energy use data from the base year of 2012 and tested with the long term average weather and consumption data from 1961 to 1990. The models were then used to predict consumption in 2040 and estimated a 5.5% rise for electricity and 28% fall for gas.

3.2. Regression models based on simulated data

For some commercial buildings, historical load data may not be readily available. In this case, analytical thermal modelling software, as mentioned in Section 2.1, can become an effective tool by simulating data for regression models.

Turiel et al. [23] utilized a database of previous DOE 2.1-A simulations to develop a simplified method for commercial building load analysis. An office building in Denver, Colorado was used for the simulations and 11 influence parameters were identified using sensitivity analysis of building and system control parameters (see Section 3.7.2). The annual heating, cooling and total energy loads were then compared to the DOE simulations results. For most of the test runs, the model prediction and actual DOE 2.1-A simulations differed by less than 15%. The study proved to be very useful for generic building types which require similar analysis but separate models would be required for buildings which have different aspects, HVAC and climates than the studied types.

Lam et al. [21] generated a load database through a series of building energy simulation runs in DOE-2 for five different buildings in five different climates in China. A sensitivity analysis identified 12 key building design variables. Typical Meteorological Year (TMY) data which was calculated for an earlier study [22], was used in the energy simulation with the identified building influence parameters. The regression models were developed to predict annual building energy use and were compared with DOE-2 simulation program results. The model R^2 varied between 89% and 97% for different climates. Warmer climates proved to have a stronger correlation between annual building electricity use and influence parameters. The difference between the regression forecasted and DOE simulated annual building electricity was mainly within 10%. The study proved to be useful for estimating energy savings during the initial design stages when different building schemes and design concepts are being considered, however, it might not be adequate for applications where more accurate and higher time resolution forecasts are desired.

Hygh et al. [24] used Energy Plus to develop MLR models. A medium size office building from DOE standards was chosen for the study and 27 building parameters were identified. TMY climate data was used in the model to find the annual cooling, heating and total energy use for the same office building in four different climatic locations. The model was developed using this data set and its predictions were compared with the simulation results. The accuracy of the regression model was further improved by a forward stepwise regression which resulted in an adjusted R^2, (see Section 4.2.1) values exceeding 96%. Table 1 shows the influence parameters used in each regression model based on simulated data.

Window to wall ratio (WWR), window heat loss coefficient, and wall heat loss coefficient, are the three influence parameters used in all three studies. Hygh et al. [24] used the highest number of influence parameters which gave the highest R_{adj}^2 values (see Appendix A), however it is not as straightforward to conclude that a higher number of influence parameters will result in higher R_{adj}^2 as each study has differences in the influence parameters, building types and climates. Having a higher number of influence parameters might not improve accuracy; in fact, in some cases it can complicate the models and reduce the forecast performance (see Section 4).

Table 1
Influence parameters used in each study (An asterisk indicates that the influence parameter is analysed for each orientation N, S, E, W).

Influence Parameters	Turiel et al. [23]	Yang et al. [22]	Hygh et al. [24]
Roof heat loss coeff.	X		X
Wall heat loss coeff.	X	X	X*
Window heat loss coeff.	X	X	X*
Window SHGC	X		X*
Window to wall ratio	X	X	X*
Window shading coeff.		X	X*
Lighting Load	X	X	
Equipment Load	X	X	
HVAC air intake	X	X	
Fan efficiency		X	
Chiller COP		X	
Boiler efficiency		X	
Heating set point	X	X	
Cooling set point	X	X	
Night thermostat	X		
Total building area			X
Number of stores			X
Depth			X
Building orientation			X
Roof colour and emissivity			X

3.3. Regression models for predicting HVAC loads in commercial buildings

The study by Katipamula et al. [18] aimed to derive steady state functional relationships between weather parameters with hourly, whole building HVAC thermal energy use (cooling and heating) in medium and large commercial buildings. Analytic equations for each of the four most widespread HVAC systems were developed: Dual duct under constant and variable air volume (DDCAV and DDVAV) and terminal reheat under constant variable air volume (TRCAV and TRVAV). The equations included building envelope characteristics (heat loss coefficient, internal loads), system parameters (airflow rate, hot & cold duct temperatures) and weather parameters. The study revealed which parameters are mostly affected by HVAC system types and gave results on the realistic range of variation of system parameters along with a set of typical values (see Table 1).

Another study done by Katipamula et al. [16] analysed cooling energy loads for five different commercial buildings in Texas, USA for two different HVAC systems, DDCAV and DDVAV, using historical load and weather data. Regression models were developed for cooling energy consumption for the five buildings at monthly, daily, and hourly time scales. The study found that MLR models produced a 33% decrease in the coefficient of variance (CV) when compared to single regression models. The study further investigated the partial R^2 of each influence parameter of the MLR models. Table 2 below shows the results of the MLR models implemented on five building cooling loads for the discussed HVAC systems.

3.4. Regression models for measurement and verification of retrofit savings

The Princeton Scorekeeping Method (PRISM) was introduced in

Table 2
Partial R^2 of Influence Parameters on both HVAC systems for five different buildings (See nomenclature for the influence parameters [16]).

	DDVAV MLR PARTIAL R^2 (%)			DDCAV MLR PARTIAL R^2 (%)		
Building No	**1**	**2**	**3**	**1**	**4**	**5**
Total Model	96.4	96.2	91.3	91.9	96.9	95.0
DBT	83.2	79.1	84.1	87.1	6.6	4.5
T_{dp}	7.2	8.1	5.5	4.0	90.3	88.6
q_i	4.1	2.3	0.6	0.8	0.0	0.9
I	1.1	2.3	0.2	0.0	0.0	0.1
I*DBT	0.8	0.8	0.9	0.0	0.0	0.9
q_{sol}	0.0	0.0	0.0	0.0	0.0	0.0

1986 to determine Normalized Annual Consumption (NAC) of residential units in New Jersey, US. PRISM uses utility meter readings from before and after a retrofit with average outdoor temperatures in order to derive the total retrofit savings (also known as weather adjusted retrofit savings) [43]. Some extensions to the PRISM model have been introduced by adding solar gains, occupancy data, wind speed, and internal gains on top of the main influence parameter, the dry bulb temperature DBT [25]. The models can incorporate a change point temperature term, which are then called change point or signature models. Eq. (4) is an example for a three parameter change point regression model with β_3 as the change point and $(\ )^+$ indicates that the value within the brackets is set to zero when its value is positive.

$$E = \beta_1 + \beta_2(T_0 - \beta_3)^+ \tag{4}$$

Kissock et al. [41] used linear and change point regression models to estimate weather adjusted retrofit savings in commercial buildings. Sever et al. [40] used a modified change point regression model (inverse energy signature model) to estimate retrofit savings in industrial buildings. The method uses both actual and TMY data. The change point regression model results for expected savings were compared to the savings predicted by detailed hourly simulations and the MAPE was within 14%. More recently, Walter et al. [42] used multivariate regression models to build a global model on predicting retrofit savings. The authors used data from a large building stock which included information of building energy use, building use type, location, operational characteristics, and energy systems. The study quantified the impact of certain building envelope components and equipment on energy usage by finding their regression coefficients. This made it possible to predict the potential savings from retrofitting a certain component. In particular, savings from retrofitting walls and windows were quantified for the studied building stock.

3.5. Influence parameters

3.5.1. Ambient Dry Bulb Temperature (DBT)

According to Kissock et al. [41], DBT is recognized as the primary contributing weather variable in forecasting building electricity loads. There are some advantages in using the ambient temperature as the only influence parameter for regression analysis as it eliminates statistical problems due to multicollinearity (see Section 3.8) and reduces data collection requirements to a single, accurately measured and widely available parameter. Because of these advantages, temperature based regression models are widely used in electricity load forecasting [17,24,27,38,44] and estimating weather normalized retrofit savings [40,43,45]. However, regression models using ambient temperature as the only input have produced forecasted daily heating and cooling loads of commercial building with a CV-RMSE of 15% [41]. This is a reasonably good forecast error, however hourly CV-RMSEs under 10% are not uncommon in the literature [30].

DBT can also be used in other forms, such as cooling and heating degree days (CDD & HDD) or cooling and heating degree hours (CDH & HDH). Hygh et al. [24] used annual HDD and CDD functions for Energy Plus simulations and achieved R^2 values of 96%. Lam et al. [44] compared two simple regression models, one with DBT as the sole input and the other with the CDH as the sole input (a base temperature of 19 °C was used). The study revealed that monthly electricity use showed a strong correlation with both mean monthly ambient temperature and CDH. The simple regression model with DBT had an average R^2 of 0.87 while the model with CDH had an average R^2 of 0.88 for 20 different commercial buildings.

3.5.2. Humidity

Humidity is a commonly used climate parameter in regression models [16–19,21,22,39]. In the study done by Nassif [19], specific humidity showed the lowest partial R^2 for buildings located in sunny and inland climates. Whereas in hot and humid climates, where the cooling load was dominant year round, humidity had a significant impact on electricity consumption for air conditioned buildings.

Katipamula et al. [16] used specific humidity as one of the influence parameters in the MLR modelling for cooling energy for different buildings (which was previously discussed in Section 3.3). Reddy & Claridge [39] used humidity as well as DBT and solar radiation as the climatic influence parameters. The study derived functional relationships that exist between hourly whole building air side HVAC thermal energy use, building envelope and climate parameters.

Kissock et al. [41] , used humidity as one of the regression parameters as well as DBT and internal load, for generating a synthetic data sequence, representative of the daily energy use in large institutional buildings in four different climates. In the first part of the study, the effects of humidity on constant and variable air volume HVAC systems were investigated, where HVAC loads showed an increase with the increased amount of latent load. Braun et al. [17] used humidity in the form of the humidity ratio, the ratio of DBT to the relative humidity. It was noted that due to multicollinearity between the humidity ratio and DBT, DBT only regression analysis was preferred. The R^2 was 0.95 for regression with both DBT and relative humidity and it was only reduced to 0.92 when the humidity ratio was dropped from the regression equation.

3.5.3. Solar gain

Solar gain is another weather parameter used in regression models. The effect of solar gain can be significant especially for buildings in sunny climates and with high window to wall ratio (WWR). Reddy & Claridge [39] found that solar gain had a larger influence than humidity for three of the four investigated buildings, each located in a different climate. For the building located in a dry, high solar radiation location (Albuquerque), solar gain, partial R^2 was three times that of humidity for daily building load, whereas in a high humidity location (Miami), humidity showed a larger partial R^2 than solar gain. In another study, solar gain through windows was identified as the largest component of the cooling loads in commercial buildings during the summer month July for five different cities in China [22].

3.5.4. Scheduling (weekdays/weekends/holidays)

The effect of scheduling is important for commercial buildings. The load profiles can vary significantly between weekdays, weekends and holidays. Therefore weekdays generally need to be distinguished from weekends and holidays [25]. In the literature, scheduling is mostly taken into account by either using separate models for weekdays, weekends, and holidays or by using an indicator variable (binary variable) within the same model. For example, Ramanathan et al. [20] developed 24 separate forecasting equations for each hour of the day. These equations were further divided for weekends and weekdays so that there were actually 48 separate models.

The multivariate linear regression models developed by Papalexopoulos & Hesterberg [38] took weekdays, weekends, and holidays into account by using binary variables. It was concluded that taking these effects into account can significantly improve the accuracy of the forecasting algorithm. Kissock et al. [41] also used an indicator variable in the regression model which was assigned to 1 for weekdays and 0 for weekends.

3.5.5. Time resolution

The time resolution of the regression models reviewed in this paper ranged from hourly to annual. The weather influence parameters such as DBT, T_{dp}, and solar gains can significantly vary from hour to hour. This variation becomes much lower over a daily or monthly time period as the positive and negative changes from the mean values average each other out. Therefore, building electricity loads which have a strong correlation with the weather parameters show much less scatter at monthly and daily intervals compared to hourly. The same logic can be applied when comparing monthly and yearly variation. As the monthly variations average out over the year, annual results show less deviation than monthly results [16]. Hence, obtaining high accuracy becomes a more difficult task with increased time resolution.

3.6. Single vs multivariate regression models

MLR models provide higher accuracy than single variable regression models for modelling HVAC energy use in commercial buildings, since the consumption in large commercial buildings is a complex function of climatic conditions, building characteristics, HVAC system operation and occupant behaviour [16].

Nassif [19] compared SLR models with different MLR models (three parameters, and four parameters) with and without change point parameters for the data collected from 225 schools in Central Florida. The SLR models used DBT as the only parameter whereas the MLR models had additional parameters such as T_{dp} and a binary variable (weekdays/weekend). Table 3 shows the average CV of MLR and SLR for the 225 schools for data analysed over a three year period.

3.7. Methods for improving regression models

MLR influence parameters may be highly correlated with each other which causes a multicollinearity problem. When multicollinearity exists in a model, influence parameters do not indicate their true relative importance. Furthermore, the uncertainties of the regression coefficients (usually the parameter's standard error) could be so large that the model's usefulness might be compromised [41].

3.7.1. Principal Component Analysis

Principal Component Analysis (PCA) is one of the methods developed to deal with the multicollinearity problem and it analyses covariance and correlation structuring. PCA can not only summarize data (reducing the number of influence parameters) but can effectively remove the multicollinearity effects from influence parameters. Reddy & Claridge [39] made a broad PCA evaluation on a three parameter MLR model (DBT, humidity, solar radiation) and compared this model with a single parameter MLR model. It was shown that PCA is more demanding in time and effort than the single parameter MLR approach; however, it can produce more reliable and physically plausible models. It was concluded that PCA can be recommended when one of the influence parameters has a correlation coefficient strength of 0.5 or higher, and when the MLR R^2 value is low (around 0.5).

Table 3
Average coefficient of variance as a percentage for SLR and MLR [19].

CV (%)	2006	2007	2008
SLR	11.6	15.4	14.9
MLR	8.3	12.1	13.5

3.7.2. Sensitivity analysis

When regression analysis is implemented using thermal model, usually a high number of variables are required. Additionally this can be a time intensive and challenging approach to run sufficient simulations for high numbers of influence parameters. In this case, a sensitivity analysis can be carried out in order to measure the relative importance of the influence parameters, making it possible to eliminate the parameters that have negligible effect on the output variable. Many studies have used sensitivity analysis in order to identify the most important variables from the initially nominated variables [20–24].

Another approach is the Seemingly Unrelated Regressions (SUR) model that can also be used in regression models. However, Ramanathan et al. [20] investigated this approach for the common influence parameters used in the MLR but found the improvements in accuracy achieved by this approach to be negligible.

3.7.3. Stepwise regression

Stepwise regression is another way to improve the accuracy of a regression model. Backward stepwise regression eliminates parameters from the model, whereas forward stepwise regression considers adding new influence parameters into the model one at a time. The overall ${R_{adj}}^2$ (see section 4.3.2) is observed each time a new variable is introduced or removed from the model; the variable which does not increase the ${R_{adj}}^2$ is left out from the model. The stepwise regression quantifies in decreasing order of importance, the parameters that explain a portion of the remaining error [46].

Hygh et al. [24] performed forward stepwise regression to select the combination of variables which generated the best model among the original variables and their cross product terms. For the four analysed buildings, the initial regression model yielded an ${R_{adj}}^2$ between 0.498 and 0.816 for the heating load, and 0.917–0.977 for the cooling load, whereas the model obtained by the forward stepwise regression yielded an ${R_{adj}}^2$ greater than 0.96 for both loads.

4. Electricity load forecast analysis with regression and machine learning models

UNSW Kensington Campus is located on a 38 ha site and accommodates around 50,000 students. The Campus includes various lecture theatres, tutorial rooms, academic and administration offices, laboratories, restaurants, shops, and a fitness & aquatic centre. The Tyree Energy Technologies Building (TETB) consists of five levels and has a floor area of approximately 18,000 m^2. The building includes teaching, research and office spaces, energy intensive laboratories, and a café. A 150 kW_p Photovoltaic array is located on the roof and the building design achieved a 6 Star, Green Star Design Rating [47].

4.1. Preliminary analysis of campus and TETB hourly electricity loads

Hourly electricity load data and minute interval weather data including DBT and relative humidity (RH) was obtained from the Campus and TETB electricity meters and a local weather station respectively. Complimentary weather data was obtained from the Sydney Observatory Hill Weather Station which is located at 7 km from UNSW. The data analysis was carried out using Matlab_R2015b and the associated Statistical and Neural Network toolboxes and LibSVM [48].

Figs. 1 and 2 show the boxplots for the Campus and TETB hourly electricity loads for each day of the week and hour of the day and

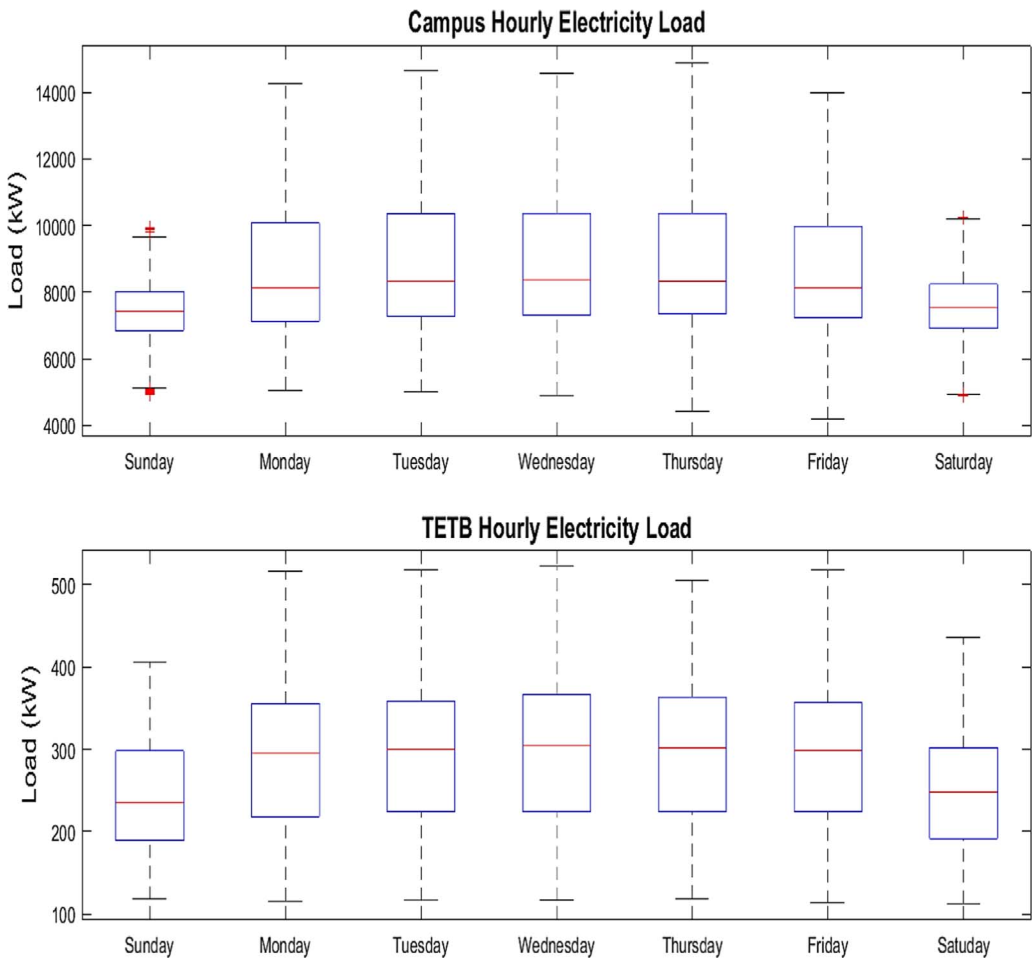

Fig. 1. Boxplots for the Campus and TETB Hourly Electricity Loads vs Day of Week.

Table 4 shows the results of the preliminary statistical analysis. The analysed hourly data set is from 01-Jan-2013 to 09-Sep-2014.

As expected, working days show higher electricity consumption than weekends for both the Campus and the TETB. The distribution of the hourly electricity load shows that the Campus experiences peak loads between 1 pm and 4 pm and the TETB between 12 pm and 4 pm. The analysed period shows that the TETB and the Campus hourly electricity loads followed a similar profile. Fig. 3 shows the hourly load profiles for an example period from March 03 to April 20, 2014. In order to compare the profiles on the same plot, both the Campus and TETB loads are scaled according to their ranges. The TETB was observed to experience daily peaks with a slight delay from the peak demand of the Campus.

4.2. Regression analysis

Following the preliminary analysis, SLR and MLR models were trained and tested on the hourly electricity loads of the Campus and the TETB. Seasonal analysis consisted of training the models with data from summer, autumn, winter and spring of 2013 and testing with the corresponding seasons of 2014. In addition to seasonal sub-sets, a fifth model was created using the complete data, which was randomly partitioned and then 10 fold cross validated. The performance metrics used for the forecast analysis are: root mean squared error as a percentage of the mean test-set load (RMSE%), mean bias error as a percentage of the mean test-set load (MBE%), mean absolute percentage error (MAPE), coefficient of determination (R^2), and adjusted coefficient of determination (R_{adj}^2). The formulas and justification for using these metrics are presented in Appendix A.

4.2.1. Single regression with climate parameters

Single regression models were previously discussed in Section 3 and their convenience and ease of use were emphasized. Therefore, the analysis starts with training single regression models using climate parameters such as DBT, RH, human discomfort parameter (humidex) and enthalpy. The latter two parameters can be calculated using DBT and RH. Humidex is used as a measure of discomfort caused by the temperature and relative humidity and it is one of the most commonly used discomfort indexes [49]. Humidex (*HI*) is calculated as follows:

$$HI = DBT + \frac{5}{9}(e - 10) \tag{5}$$

where *e* is the water vapour pressure of the air in hPa [49]. For the calculation of enthalpy values the equations reported by Padfield [50] were used. The results for each model (trained with 2013 data and tested using 2014 data) can be found in Table 4, where R^2 represents the wellness of the model fit on the training set, and the other metrics represent the model performance on the test set.

The DBT based single regression models showed the highest R^2 for almost every training set. However, this higher R^2 performance didn't result in significantly better results than other climate parameter based models on test sets. In fact the error terms were quite similar for the autumn and winter seasons for all the models. Moreover, RH based models achieved better results than DBT based models for the summer and spring seasons. It was observed that errors were significantly higher for the TETB models, which is mainly caused by the higher MBE. Investigating the mean seasonal load values of 2013 and 2014, it was found out that the TETB and the Campus's electricity loads increased around 40% and 8% respectively. The main cause of the

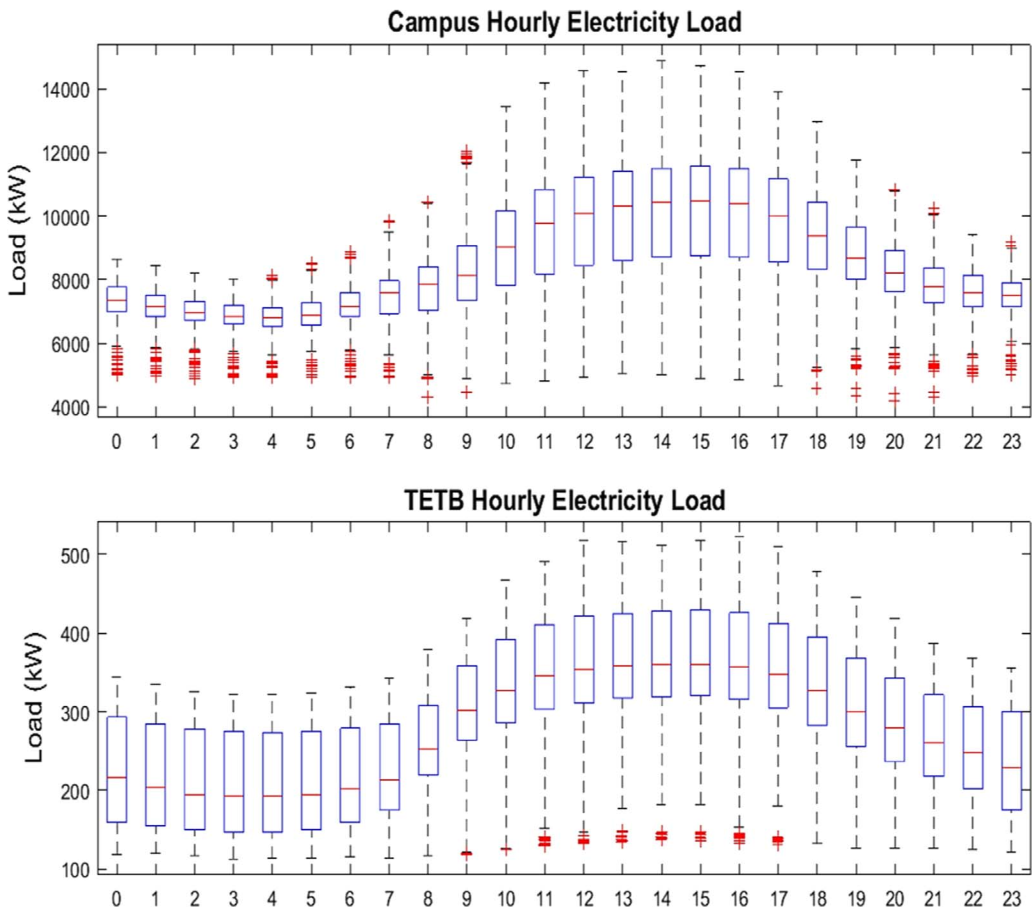

Fig. 2. Boxplots for the Campus and TETB Hourly Electricity Loads vs Hour of Day.

increase in TETB electricity loads can be attributed to the significant increase in the number of students and lab activities. Furthermore, the TETB had an active tri-generation plant which produced electricity for the whole Campus. This unit was shut down by the end of 2013 which is another major factor for this annual increase between the two years. The climate based single regression models were not capable of detecting these changes and hence resulted in poor performances.

In order to investigate the performance of the models without these step changes, a bias correction was implemented where each model output was simply increased by an amount equal to the difference between average seasonal electricity loads. Bias corrected single regression model results are shown in Table 5.

It can be seen that the bias corrected SLR models show significantly better performance than the previous models due to the smaller MBE.

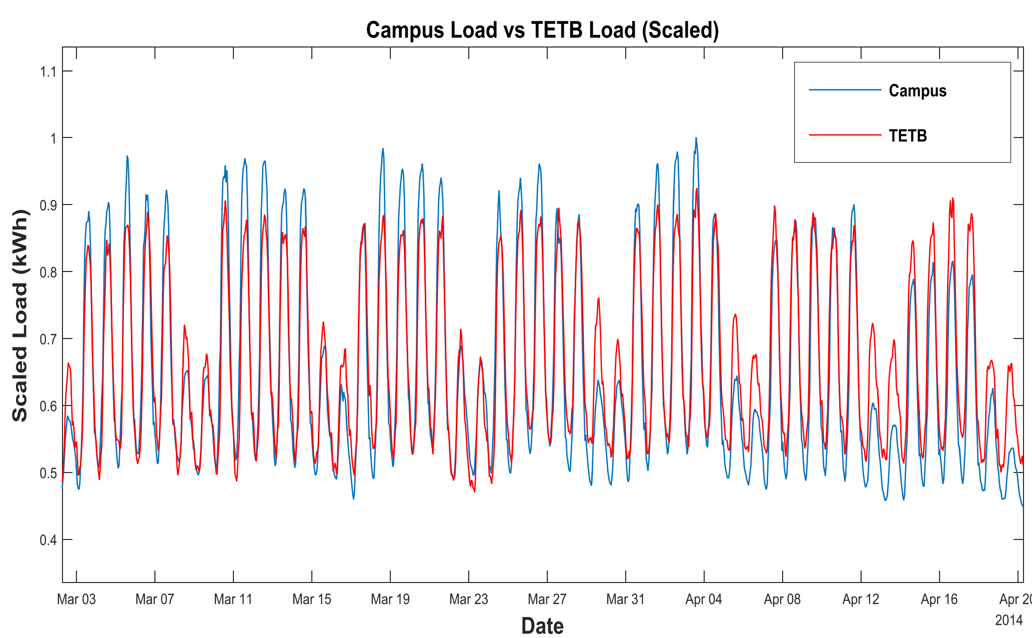

Fig. 3. Campus and TETB electricity load profiles March 03-April 20, 2014.

Table 4
Results of Seasonal Single Regression Models.

	Summer				Autumn				Winter				Spring			
	R^2	RMSE%	MBE%	MAPE	R^2	RMSE%	MBE%	MAPE	R^2	RMSE%	MBE%	MAPE	R^2	RMSE%	MBE%	MAPE
Campus Load																
DBT	0.32	21.55	−10.81	14.57	0.33	19.67	−6.79	12.59	0.29	20.89	−6.62	13.55	0.47	26.28	−15.67	18.07
RH	0.15	22.08	−8.11	14.51	0.17	22.76	−7.83	14.27	0.16	20.20	−5.30	13.81	0.19	21.87	−6.94	14.76
Humidex	0.23	23.22	−11.78	15.38	0.24	20.80	−6.61	13.26	0.21	20.87	−6.90	13.53	0.34	29.38	−17.04	19.59
Enthalpy	0.12	24.41	−11.66	15.51	0.16	22.00	−6.61	13.95	0.09	20.52	−6.81	13.38	0.12	21.87	−8.73	14.61
TETB Load																
DBT	0.21	38.51	−31.30	31.43	0.21	36.67	−31.55	31.55	0.34	41.87	−34.60	34.61	0.35	42.83	−37.25	37.25
RH	0.26	35.82	−28.08	28.69	0.19	37.95	−33.08	33.09	0.18	39.97	−32.84	32.84	0.14	34.53	−27.80	27.81
Humidex	0.09	40.09	−31.67	31.68	0.13	37.45	−31.45	31.45	0.25	42.04	−35.02	35.05	0.28	46.60	−39.35	39.35
Enthalpy	0.02	40.36	−31.02	31.02	0.06	38.15	−31.60	31.60	0.11	41.53	−34.95	34.96	0.12	36.53	−30.52	30.52

Table 5
Seasonal single regression model results after bias correction.

	Summer				Autumn				Winter				Spring			
	R^2	RMSE%	MBE%	MAPE	R^2	RMSE%	MBE%	MAPE	R^2	RMSE%	MBE%	MAPE	R^2	RMSE%	MBE%	MAPE
Campus Load																
DBT	0.32	17.11	2.81	13.31	0.33	17.34	4.73	14.99	0.29	18.75	2.53	15.50	0.47	22.01	−8.34	15.08
RH	0.15	19.27	5.52	15.62	0.17	19.91	3.70	15.86	0.16	18.55	3.84	15.72	0.19	19.68	0.39	15.79
Humidex	0.23	18.31	1.85	14.10	0.24	18.55	4.92	16.08	0.21	18.59	2.25	15.32	0.34	25.22	−9.72	17.04
Enthalpy	0.12	19.63	1.96	15.14	0.16	19.72	4.91	16.90	0.09	18.21	2.34	15.14	0.12	19.08	−1.40	14.77
TETB Load																
DBT	0.21	19.30	3.45	15.43	0.21	16.10	4.07	13.77	0.34	20.46	2.59	16.98	0.35	21.99	−9.85	15.43
RH	0.26	19.81	6.66	16.47	0.19	15.94	2.53	13.26	0.18	19.52	4.36	16.85	0.14	17.96	−0.40	14.42
Humidex	0.09	20.95	3.08	16.82	0.13	17.33	4.16	14.92	0.25	20.08	2.17	16.63	0.28	26.44	−12.00	17.78
Enthalpy	0.02	21.90	3.73	17.89	0.06	18.11	4.01	15.65	0.11	19.06	2.25	16.10	0.12	18.04	−3.11	13.91

Importantly for the TETB models, the errors were reduced. Furthermore, the performance difference between the TETB and Campus models was greatly reduced. On the other hand, the RMSE and MAPE were still on average 19% and 16% respectively, which is higher than errors reported for SLR models in the literature [16,19]. Our SLR models were not capable of capturing the changes and variations in electricity loads of both the TETB and the Campus. Hence, more advanced models are required for more accurate results. Although the bias correction method was helpful in reducing the error for the SLR models, it requires an additional step and may not be a feasible implementation for real time forecast applications. Therefore the bias correction will be omitted for the following methods and we will investigate if the models are able to adequately describe the step change.

4.2.2. Multiple Linear Regression (MLR) model with climate and temporal parameters

For MLR models, the forward stepwise regression method was implemented. In a forward stepwise regression, the initial model does not include any of the regression predictors. Besides the R_{adj}^2 criteria which were previously discussed in Section 3.7.3, each predictor has to have a p-value of less than 0.05 in order to be included in the model. In addition to the climate parameters which were used in the single regression models, new temporal predictors are introduced and the initial candidate parameters consist of:

- Previous day same hour load (X1)
- Previous week same hour load (X2)
- Previous 24 h average load (X3)
- Working day/holiday binary indicator (X4)
- DBT (X5)
- RH (X6)
- Humidex (X7)
- Enthalpy (X8)
- Hour of the day (X9)
- Day of the week (X10)

The temporal predictors are introduced in order to capture the relationship between the load's present value with its past. The working day/holiday parameter is a binary indicator such that for working days it is one and for the weekends and holidays it is zero.

Five MLR models were built using 2013 data for training, four using sub-sets for each season, and one using the whole data set using 10 fold cross validation. Fig. 4 shows the stepwise regression process implemented for the summer MLR model of the Campus load. The figure also shows the coefficients, t-statistics, p-values and confidence intervals of each parameter. The parameters shown in blue are those which are included in the final model.

The model history shows the progress of RMSE with the addition of each parameter to the summer model. The parameter X1 was first included as it resulted in a R_{adj}^2 value of 0.73. Next parameters X2, X5, X4, X10, and X3 were added as all improved the overall model R_{adj}^2. The parameters X6, X7, X8 and X9 were not included in the model as each had p-values greater than 0.05. In particular parameters X7 and X8 (Humidex and Enthalpy) had p-values close to 1, which indicates that both have very little significance as predictors for the Campus electricity load when they are used with other chosen parameters. The parameter X1 was the most significant predictor for all models (R_{adj}^2~0.8), followed by X2. The parameter DBT was included in all other seasonal models whereas RH was only included in the spring model.

In the case of seasonal models for the TETB, RH was included in the summer, winter and spring models, whereas DBT was only included in autumn and winter models. It is important to note that in all of the models, the first four model parameters to be chosen by the stepwise regression consisted of temporal parameters and working

Fig. 4. Stepwise regression summary for the summer MLR model for the campus load.

day/holiday binary indicator (X1–X4) and whenever a climate parameter was included in the model it only had the 5th or lower priority. The fact that RH and enthalpy were not included in three of the Campus and two of the TETB seasonal models is unusual given the fact that around 60% of the TETB building load is due to HVAC. Furthermore, the Campus has many buildings like the TETB and all of them have high HVAC loads. This could be an indication that including the temporal load parameters such as X1, X2, and X3 for day ahead hourly forecasts, brings information that is originally contained within the climate parameters. This can hinder the inclusion of some climate parameters in the MLR models when using the forward stepwise regression method, even though they proved to be a significant parameter by themselves in single regression analysis.

Once the stepwise MLR training was completed and final model parameters were determined, the models were tested on the corresponding 2014 seasonal test sets. In addition to the SLR analysis, MLR models were also tested on random training and test sub sets with 10 fold cross-validation method. The results are presented in Table 6.

It is clear that MLR models outperformed all SLR models. The additional temporal predictors greatly improved the wellness of the fit resulting in smaller errors. Furthermore, temporal parameters were able to capture the step change to a certain degree without needing a bias correction, as was required for the SLR analysis. It can be seen that the models showed a slightly better performance for the Campus loads than the TETB. Also the best performance was shown on the winter data set.

4.2.3. Principle Component Analysis

In addition to the stepwise analysis, a correlation matrix was plotted in order to measure the linear correlation between each of the MLR parameters, as shown in Table 7. This was done to diagnose highly correlated parameters and hence prevent any possible collinearity problems.

The climate parameters DBT, humidex and enthalpy proved to have high linear correlation with each other (highest between humidex and enthalpy at 0.97) which was expected as the latter two are functions of DBT. This high correlation was not observed for RH, since it was only indirectly included in the calculation of humidex and enthalpy [50]. For the temporal parameters, previous day same hour load and previous week same hour load showed high correlation at 0.84.

If a parameter is a linear combination of other parameters then the model has multicollinearity (as introduced in Section 3), and the variance inflation factor (VIF) method can be used to assess the level of multicollinearity. For each parameter, the VIF is calculated by dividing the variance of its regression coefficient within the full model to the variance of its regression coefficient for its single regression fit. As a rule of thumb, VIF values larger than 10 is an indication of

Table 6
Seasonal and cross validation MLR Model Results for day ahead hourly electricity load forecasts.

	Campus				TETB			
Model	**R_{adj}^2**	**RMSE (%)**	**MBE (%)**	**MAPE**	**R_{adj}^2**	**RMSE (%)**	**MBE (%)**	**MAPE**
Summer	0.87	7.95	−1.68	5.50	0.88	10.08	5.17	7.71
Autumn	0.83	9.43	−6.13	6.99	0.91	6.82	−4.54	5.16
Winter	0.93	4.32	1.97	3.49	0.96	5.03	2.69	4.06
Spring	0.89	6.98	−2.50	4.55	0.90	5.44	3.64	4.29
CrossVal	0.89	6.98	0.43	4.75	0.94	7.75	0.61	5.50

Table 7
Correlation matrix of MLR parameters.

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10
X1	1.00	0.84	0.49	0.11	0.15	−0.36	0.04	−0.05	0.40	0.24
X2	0.84	1.00	0.42	0.30	0.13	−0.34	0.03	−0.06	0.40	0.03
X3	0.49	0.42	1.00	0.40	−0.22	−0.09	−0.26	−0.29	0.00	0.27
X4	0.11	0.30	0.40	1.00	0.00	−0.02	−0.01	−0.02	0.00	0.00
X5	0.15	0.13	−0.22	0.00	1.00	−0.29	0.95	0.85	0.24	0.03
X6	−0.36	−0.34	−0.09	−0.02	−0.29	1.00	−0.01	0.22	−0.25	−0.03
X7	0.04	0.03	−0.26	−0.01	0.95	−0.01	1.00	0.97	0.18	0.02
X8	−0.05	−0.06	−0.29	−0.02	0.85	0.22	0.97	1.00	0.12	0.02
X9	0.40	0.40	0.00	0.00	0.24	−0.25	0.18	0.12	1.00	0.01
X10	0.24	0.03	0.27	0.00	0.03	−0.03	0.02	0.02	0.01	1.00

multicollinearity [51]. In this case, all the climate parameters (X5–X8) showed VIF values of the order of 100.

The correlation matrix and the VIF method showed that there is multicollinearity within the chosen predictor parameters. As discussed in Section 3.7.1, one way to tackle the multicollinearity problem is by using the Principle Component Analysis (PCA). PCA creates new predictors by using linear combinations of the existing parameters where multicollinearity does not exist within the new parameters. Since the predictor parameters have different units and have different variance, they are scaled by dividing them by their respective variances [52]. Using the correlation matrix previously shown, ten principle components (eigenvectors) were found. These principle components were sorted according to their eigenvalues, where the relative importance of the principle components is shown in Fig. 5. The first seven principle components were sufficient to explain almost 100% variance within the predictor parameters. These seven principle components were used to construct the new predictor matrix, consisting of seven new parameters instead of ten.

Using the new seven parameters, training and testing were implemented for each of the seasonal models. However the new principle components did not result in any significant improvement in models R_{adj}^2 and accuracy. In fact, the obtained errors were very similar to the stepwise regression results shown previously. This indicates that the stepwise method was an adequate parameter selection method for the dataset, and didn't require additional steps to tackle the multicollinearity problem.

4.3. Regression analysis on daily peak loads using MLR models

As mentioned previously, forecasting peak loads is of paramount importance for maximum demand management activities, and so its analysis requires special attention. Table 8 shows the summary statistics for the daily peak electricity loads of the Campus and TETB for the years 2013 and 2014. The data is analysed for the working/business days and separated between four seasons (Holidays/Weekends are presented separately at the end of the table). The TETB experienced the highest peak load during winter; however spring has the highest average value. The Campus on the other hand experienced the highest peak during autumn which has the highest average as well. This makes sense since in Australia, the Autumn season starts on the 1st of March, which coincides with the beginning of the university semester.

There is a significant increase in average daily peak loads between 2013 and 2014 (11% and 27% of the 2013 averages for the Campus and the TETB respectively). As previously mentioned a tri-generation unit was operating during weekdays at a rated output of 746 kW which used natural gas as the fuel. This unit was shut down at the beginning of 2014 hence around 7% of the daily Campus peak increase of grid electricity, can be attributed to the closure of this unit. In addition, the increase in the lab and classroom activities within the TETB in 2014 explain some portion of the increase in peak loads experienced by the building.

The parameters used in the previous section were modified in order to incorporate peak and minimum demand. The new peak MLR model consisted of:

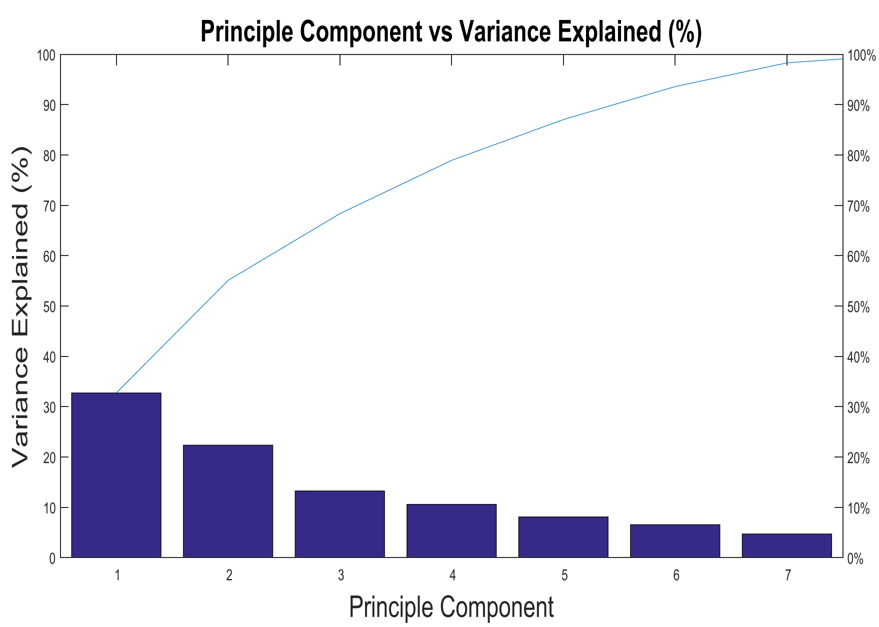

Fig. 5. New principle components vs variance explained within parameters.

Table 8
Summary Statistics for Daily Peak Loads.

	Max	Min	Mean	Std
Campus (kW)				
Summer	14,256	5435	10,710	1736
Autumn	14,874	7743	11,931	1453
Winter	12,523	6889	10,831	829
Spring	13,762	9249	11,148	906
Year 2013	14,030	5435	10,636	1266
Year 2014	14,874	7113	11,818	1302
Weekends/Holidays	10,250	5316	8389	834
TETB (kW)				
Summer	518	117	362	73
Autumn	493	218	394	61
Winter	522	162	399	71
Spring	517	313	414	41
Year 2013	462	117	348	50
Year 2014	522	267	443	45
Weekends/Holidays	434	123	285	68

- Previous Day Peak Load (X1)
- Previous Day Minimum Load (X2)
- Previous Week Peak Load (X3)
- Previous Week Minimum Load (X4)
- Holiday/ Business-day Binary Indicator (X5)
- DBT (X6)
- RH (X7)
- Humidex (X8)
- Enthalpy (X9)

- Hour of the day (X10)
- Day of the week (X11)

Using these parameters, four seasonal MLR models were built to predict daily peak loads. The stepwise analysis showed that the first model parameter to be included was X5 which indicates the importance of the Holiday/Business day binary indicator in explaining the variation of daily peak loads (the difference between the average daily peak for business days & holidays was around 1500 kW and 60 kW for the Campus and the TETB respectively). The parameters X5; X3, X1, X4 and X10 were the next most significant predictors of daily peak loads. The only climate variable included in the MLR models was DBT (X6) since no other climate variables had p-values less than 0.05.

The seasonal MLR models showed a fairly good performance for the Campus, however the MBE was observed to be significant for the TETB peak loads (above 20%). This high bias error was an indicator for an under-fit model, which failed to capture the increase of the TETB daily peak loads between the years 2013 and 2014. In order the tackle the problem two additional methods were tried [53]:

- Introducing new parameters to the model
- Using a larger training set

Hence, additional second and third order polynomial parameters were introduced to the model for X1, X2, X3 and X4. The minimum errors were obtained by using X1 as the polynomial parameter; however the bias problem was not solved by this method. In order to test the effect of using a larger training set, the models were trained with a whole year of data (2013) and tested on 2014 seasons. The larger data set was implemented on the initial models without using any polynomial terms. Using a larger training set resulted in reduced errors. Table 9 summarizes the obtained results for the seasonal Campus model (trained with corresponding 2013 seasons) and the TETB model (trained with the whole 2013 data set).

It can be observed that four out of five models performed better at predicting Campus peak loads. Moreover, models showed a better performance in predicting day ahead hourly loads than daily peak loads for both TETB and the Campus (see Table 6).

4.3.1. Daily peak loads vs DBT & RH

It was noted previously that electricity consumption driven by HVAC cooling was identified as the main contributor to the peak loads [54] in NSW, Australia. Hence this leads to an expectation that the highest peak loads would occur during summer months (December, January and February). However, this wasn't the case for the Campus and TETB building for the analysed periods. Fig. 6a and c show the relationship between daily peak loads with the corresponding daily maximum DBT for working days. It can be observed that there is no significant relationship between daily peak loads and daily maximum DBT (the R^2 of linear fits were observed to be around 0.09 and 0.03 for the Campus and TETB respectively). Both for the Campus and the TETB the highest peaks occurred at daily maximum DBT (27 °C and 19 °C respectively) rather than at a higher summer DBT. This weak relationship indicates that other factors influence the peak loads other than DBT for these particular cases.

Fig. 6b and d further investigate the relationship of daily peak loads to the corresponding daily maximum relative humidity values. Once again, the relationship between daily maximum RH and daily peak load for both the Campus and the TETB was found to be weak. Normally, an increase in RH would result in increased HVAC latent load which might lead to the expectation of higher peak electricity demand.

Since there are many laboratories, classrooms and other facilities within a university environment, occupancy becomes a very important factor which can significantly affect the peaks. This factor is most likely the reason why the Campus and TETB don't experience the highest daily peaks during summer (typically when highest DBT and RH occur) as most students and staff are on holidays.

4.4. Machine learning model analysis

Machine learning models can have different objectives in terms of minimizing errors (also known as the cost function). All the implemented machine learning models utilized in this study used the least-squares method for the cost function.

Similar to the MLR analysis, all machine learning models were initially trained with seasonal sub-sets of the 2013 data and tested on the corresponding seasonal sub-sets of 2014. In addition, another model was trained using the entire data set, such that the complete data was randomly partitioned and 10 fold cross validated. All the training data sets were further partitioned and 70% of the data was allocated to training and 30% to cross validation. The results obtained from seasonal data sub-sets were represented by the seasons summer, autumn, winter and spring; while the results obtained by using the complete data set was represented by '2013–2014'.

4.4.1. Artificial Neural Networks (ANN)

Following the MLR analysis, ANN models were implemented for day ahead hourly and daily peak load forecasts. The ANN models incorporate different input, hidden and output layers to explain complex and non-linear relationships between inputs and outputs. For this study, the input layer consisted of the same predictor variables used in the MLR analysis which are activated by a hyperbolic tangent (sigmoid) function, and are then fed into the single hidden layer. This study followed the common practice of using a single hidden layer network instead of multiple layers for forecasting applications [55]. On the other hand, in order to choose the optimal number of neurons (learned input parameters) within the hidden layer, training was done with different number of neurons (from 5 to 30). The optimum number of neurons was found to be 10, based on cross-validation errors. The output layer consisted of one neuron which is activated by a linear regression function.

The ANN models used in our analysis were the Levenberg-Marquardt Backpropagation (LM) and the Bayesian Regularization

Table 9
Results for daily peak loads by MLR models.

Model	Campus (trained with seasonal data)				TETB (trained with whole year data)			
	R_{adj}^2	RMSE (%)	MBE (%)	MAPE	R_{adj}^2	RMSE (%)	MBE (%)	MAPE
Summer	0.85	7.34	2.12	6.18	0.86	11.94	5.45	9.35
Autumn	0.84	8.12	−1.90	6.45	0.87	9.02	6.10	7.23
Winter	0.81	6.29	−2.55	5.27	0.85	6.45	2.29	5.63
Spring	0.90	12.03	−11.28	11.28	0.86	9.13	5.69	7.56
CrossVal	0.85	6.74	0.47	5.15	0.90	7.28	0.85	6.50

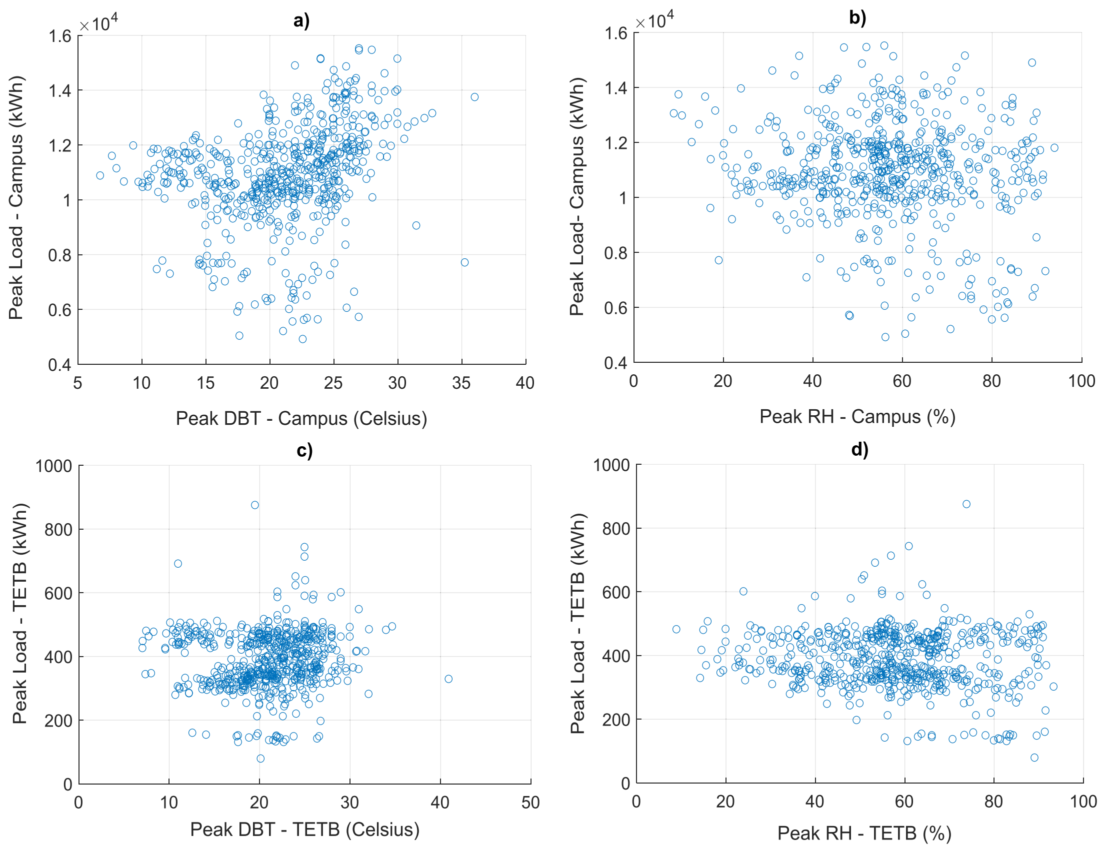

Fig. 6. Daily peak loads vs daily maximum DBT (°C) and RH (%) for the Campus and TETB, a) Peak DBT for Campus load, b) Peak RH for Campus load, c) Peak DBT for TETB load, and d) Peak RH for Campus load.

Backpropagation (BR) models. Fig. 7 demonstrates the best performing ANN architecture where w and b corresponds to the weight and bias unit respectively (bias here represents the unity vector for representing the constant terms in the hidden layer). LM uses a standard back propagation technique to calculate the Jacobian matrix without computing the Hessian matrix, which is a much more complex task. In the BR method, the network weights are treated as random variables (Gaussian distribution) and the optimization function parameters are defined by Bayes' theorem. Please refer to [56] for more detailed information and comparison of these methods.

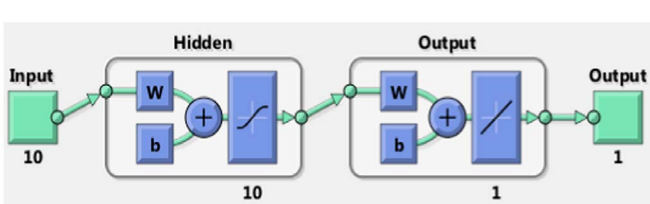

Fig. 7. Best performing ANN architecture with a single hidden layer and 10 neurons [57].

An important feature of ANN is the possibility of using a regularization parameter especially when coping with overfitting problems. A range of different regularization parameter values for the network were tried. However, this did not achieve significant improvements compared to the default network with no regularization parameter. The results are shown in Tables 10–13.

4.4.2. Dynamic Artificial Neural Networks with exogenous inputs

The nonlinear autoregressive network with exogenous inputs (NARX) is a mixture of Neural Network and Time Series methods and is based on the linear Auto Regressive Model with Exogenous inputs (ARX). With the inclusion of the time series, the ANN becomes a dynamic network where the output is regressed on its previous values and previous values of other exogenous variables (e.g. climate variables in our example). The architecture of the NARX method is shown in Fig. 8. The additional parameter for such a network is the time delay steps for the training set. The training of the network is done on an open loop and the test is done on the closed loop where the load is recurrently fed into itself as an input. For such a configuration, the NARX can perform as many predictions as the input series has time

Table 10
RMSE (%) results of all the models on day ahead hourly and daily peak electricity load.

Model	MLR	NN LM	NN BR	NARX LM	NARX BR	RT	SVR	MLR	NN LM	NN BR	NARX LM	NARX BR	RT	SVR
	Campus Hourly Forecast RMSE (%)							**TETB Hourly Forecast RMSE (%)**						
Summer	7.95	4.72	4.36	2.34	1.82	7.09	7.43	10.08	7.09	6.66	2.83	2.74	8.94	8.63
Autumn	9.43	3.97	3.29	1.85	1.88	7.02	5.97	6.82	3.70	3.28	2.11	1.98	9.15	5.17
Winter	4.32	2.34	2.05	1.56	1.30	5.96	4.58	5.03	2.94	2.81	1.99	1.89	7.95	4.41
Spring	6.98	4.09	3.10	2.15	1.76	7.49	5.02	5.44	4.69	3.56	2.41	2.33	9.54	6.50
2013–2014	6.98	4.25	3.89	1.92	1.87	6.18	5.99	7.75	4.81	4.53	2.41	2.22	7.99	6.01
	Campus Daily Peak RMSE (%)							**TETB Daily Peak RMSE (%)**						
Summer	7.66	7.83	5.53	9.73	7.58	12.39	8.24	11.94	10.80	9.21	11.74	9.29	17.15	13.49
Autumn	9.28	4.50	4.00	8.87	6.38	8.25	7.03	9.02	6.82	5.63	7.09	6.38	8.67	7.38
Winter	6.44	2.62	2.60	9.34	2.71	5.02	6.95	6.45	6.03	5.64	10.42	5.60	8.11	6.89
Spring	12.64	5.00	6.93	8.21	6.20	7.53	8.01	9.13	4.77	2.61	10.80	7.08	9.63	8.78
2013–2014	6.83	6.34	4.35	8.85	4.92	7.83	6.41	8.01	7.08	5.96	8.98	6.35	9.47	10.94

steps [58]. Similar to the previous analysis, both BR and LM algorithms were used for day ahead hourly and peak load forecasts. It was observed that the BR model outperformed the LM model in this case. However, this time the implementation of the BR model required significantly more computation time than the latter. Results are shown in Tables 10–13.

4.4.3. Regression Trees

Regression tree analysis is carried out by using the same influence parameters for both day-ahead hourly and daily peak load forecasts. Two important parameters affecting the performance of the regression trees were optimized during the training. The first one was the minimum leaf size of the tree where each leaf has at least the number of minimum leaf size observations per tree leaf [59]. A deep tree with many leaves usually over-fits the data, and while it shows high accuracy on the training set it may fail to show a similar performance on the test set. A shallow tree might not achieve a training accuracy as high as a deep tree; however, its test accuracy is more similar to its training accuracy [59]. Therefore, an optimum leaf size has to be found. Fig. 9 shows the cross validation error of the regression tree for the Campus day ahead hourly load with varying leaf size. The optimum leaf size for the Campus and the TETB day ahead hourly electricity loads was 5 and 43 respectively.

The other important parameter is the pruning level of the tree, which adjusts the depth (leafiness) of the regression tree by merging the leaves on the same branch [59]. The optimum pruning level can also be found by observing the cross validation error, similar to the leaf size. After the pruning process the optimum prune levels were within 40–60, whereas the original prune levels were of the order of 100. Once these two parameters were optimized for each seasonal model, regression trees are built. The results of the regression tree models are given in Tables 10–13.

4.5. Support Vector Regression

Support Vector Regression (SVR) is a method of Support Vector Machines used for solving numerical regression problems. Similar to the ANN models, input vectors are transformed into higher dimensional spaces. The transformation can be done using similarity functions (Kernels). Amongst various different Kernel methods, the Gaussian (RBF) Kernel was used for the study due to its wide use within the literature [60]. In comparison to linear regression, SVR defines new margins denoted by ε which regress along the line with the margins (SVR tube). Points which fall within the boundaries of the tube are not considered errors but the ones which fall outside are (denoted by ξ) [34] (Fig. 10).

The SVR analysis was carried using LibSVM [48] within the MATLB environment. LibSVM required scaling (also known as feature scaling) of the influence parameters; hence, before starting the analysis, all the predictor variables were scaled using the methodology described in [61]. Following the feature scaling, input matrices were trained and 10 fold cross validated for each seasonal model. Similar to the ANN analysis, important SVR parameters were adjusted to obtain the minimum RMSE on the cross validation set. In addition to the margins ε, two other important parameters were also optimized on cross validation sets; C, the cost function regularization parameter, and γ, the Kernel parameter. This was done by introducing a range of different values for each parameter within in for loops whilst observing the RMSE error. The optimum parameters were found to be: C=500 and 1000, ε=5 and 1, γ=16 and 0.5 for day ahead hourly and daily peak loads respectively. The results for the SVR analysis are presented in Tables 10–13.

5. Results & discussions

The following tables show the RMSE (%), MAPE, R^2 and MBE (%)

Table 11
MAPE results of all the models on day ahead hourly and daily peak electricity load.

Model	MLR	NN LM	NN BR	NARX LM	NARX BR	RT	SVR	MLR	NN LM	NN BR	NARX LM	NARX BR	RT	SVR
	Campus Hourly Forecast MAPE							**TETB Hourly Forecast MAPE**						
Summer	5.50	4.64	4.60	1.59	1.36	3.90	5.42	7.71	4.64	4.60	2.54	2.19	5.23	5.42
Autumn	6.99	2.61	2.64	1.52	1.40	3.25	3.82	5.16	2.61	2.64	1.49	1.51	3.25	3.82
Winter	3.49	3.42	2.31	1.25	1.04	2.31	3.69	4.06	3.42	2.31	1.43	1.45	5.63	3.69
Spring	4.55	3.21	3.11	1.57	1.33	5.41	5.55	4.29	3.21	3.11	2.14	1.79	8.60	5.55
2013–2014	4.75	3.49	3.29	1.36	1.33	4.61	4.32	5.50	3.49	3.29	1.75	1.73	7.55	4.32
	Campus Daily Peak MAPE							**TETB Daily Peak MAPE**						
Summer	6.18	5.24	3.51	9.30	4.97	6.02	7.38	9.35	7.93	8.19	9.74	4.14	9.30	9.53
Autumn	6.45	3.66	1.98	6.06	2.03	6.53	6.11	7.23	4.06	3.77	7.90	5.23	6.01	6.36
Winter	5.27	2.12	1.18	3.07	1.01	3.23	4.19	5.63	4.62	1.76	7.56	1.23	4.59	4.39
Spring	11.28	3.25	1.18	7.11	6.01	5.23	6.02	7.56	3.55	0.95	8.29	1.74	7.28	7.23
2013–2014	5.15	4.36	3.31	4.99	3.77	4.71	5.82	6.50	4.49	3.62	6.29	4.44	5.94	7.53

Table 12
Training R^2 results of all the models.

Model	MLR	NN LM	NN BR	NARX LM	NARX BR	RT	SVR	MLR	NN LM	NN BR	NARX LM	NARX BR	RT	SVR
	Campus Training R^2							**TETB Training R^2**						
Summer	0.87	0.95	0.97	0.99	0.99	0.88	0.96	0.88	0.97	0.96	0.99	0.99	0.86	0.96
Autumn	0.83	0.97	0.98	0.99	0.99	0.87	0.96	0.91	0.99	0.99	0.99	0.99	0.85	0.95
Winter	0.93	0.98	0.99	0.99	0.99	0.91	0.96	0.96	0.97	0.99	0.99	0.99	0.86	0.94
Spring	0.89	0.92	0.97	0.98	0.99	0.86	0.94	0.90	0.98	0.98	0.98	0.99	0.84	0.94
2013–2014	0.89	0.97	0.97	0.99	0.99	0.90	0.93	0.94	0.98	0.97	0.99	0.99	0.87	0.95
	Campus Training R^2							**TETB Training R^2**						
Summer	0.85	0.92	0.86	0.74	0.83	0.86	0.78	0.86	0.86	0.86	0.74	0.85	0.82	0.79
Autumn	0.84	0.90	0.94	0.81	0.91	0.84	0.80	0.87	0.93	0.93	0.81	0.89	0.84	0.82
Winter	0.81	0.95	0.97	0.85	0.92	0.92	0.83	0.85	0.98	0.98	0.85	0.94	0.83	0.85
Spring	0.90	0.90	0.94	0.68	0.69	0.87	0.76	0.86	0.98	0.98	0.68	0.88	0.85	0.82
2013–2014	0.85	0.89	0.93	0.85	0.94	0.88	0.79	0.90	0.95	0.95	0.85	0.93	0.87	0.93

results for all the models implemented on both hourly and daily peak forecasts. It can be seen that NARX with Bayesian Regulation backpropagation showed the best overall performance on day ahead hourly load forecasts and ANN with Bayesian Regulation backpropagation showed the best overall performance on daily peak load forecasts. The ANN models showed very high training R^2 performance and their MBE (%) were much smaller compared to other machine learning models.

Fig. 11a, b, c and d show scatter plots of actual loads vs model predictions for SVR, MLR, RT and NARX BR models respectively on day ahead hourly predictions for the Campus load data for the test period from June 16 to June 23, 2014. Day ahead hourly forecast plots for all the models on the Campus load for the same period is shown in Fig. 12. Furthermore, Figs. 13 and 14 show the best and worst performing models, NARX BR and RT, for the TETB day ahead hourly forecasts, again using the same example week period.

The ANN models showed very small bias errors over all the seasonal models (< 1%). The MLR models showed initially high bias on daily peak forecasts (20% avg.). However, the error was reduced to an average value of 5% after training the models with a whole year's worth of data rather than a single seasonal data set. Both the SVR and RT models initially suffered from bias errors on both hourly and daily peak loads for the Campus loads (> 15% for 2014 Autumn and 10% on average). As previously discussed, there was a step change in the Campus electricity demand from the grid due to the closure of the tri-generation unit in 2014; hence, the high bias error could be attributed to the models not being able to capture this step change. When the training data set was increased to involve some portion of the 2014 data (test sets were still left independent of the training set), the bias errors reduced significantly to levels below < 5% on average. In addition to using larger data sets, SVR models were tested with smaller numbers of the influence parameters. The models showed a better performance when the climate variables were removed and the models

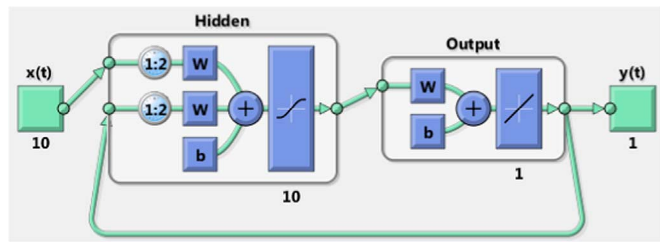

Fig. 8. NARX Architecture for closed loops [57].

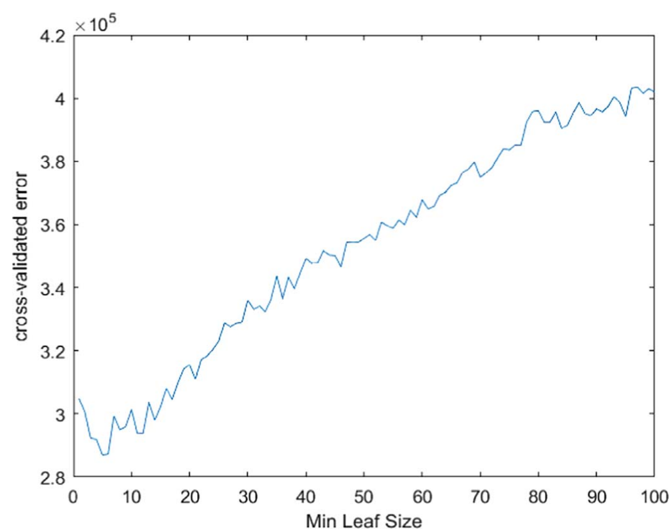

Fig. 9. Leaf Size vs Cross-Validation Error on Campus Hourly Load.

Table 13
MBE (%) results of all the models on day ahead hourly and daily peak electricity load.

Model	MLR	NN LM	NN BR	NARX LM	NARX BR	RT	SVR	MLR	NN LM	NN BR	NARX LM	NARX BR	RT	SVR
	Campus Hourly Forecast MBE (%)							**TETB Hourly Forecast MBE (%)**						
Summer	−1.68	0.05	−0.04	0.00	0.00	1.12	1.39	3.94	0.01	0.01	0.03	0.00	0.51	1.39
Autumn	−6.80	0.01	−0.19	−0.08	−0.01	−0.83	2.33	−4.57	0.04	0.01	0.00	−0.02	−0.92	2.33
Winter	1.97	0.03	0.00	−0.01	0.00	−0.53	2.21	2.49	−0.08	0.00	0.04	−0.01	−1.43	2.21
Spring	−3.88	−0.10	−0.01	0.11	−0.01	−1.17	1.21	1.41	0.04	0.01	−0.05	−0.01	−6.16	5.24
2013–2014	1.21	−0.11	−0.01	0.03	−0.01	−1.84	5.24	0.95	−0.03	0.01	0.00	0.01	−5.10	2.15
	Campus Daily Peak MBE (%)							**TETB Daily Peak MBE (%)**						
Summer	3.09	1.28	−0.30	0.16	0.34	3.76	2.65	5.45	−1.65	−1.14	1.73	−0.14	4.63	2.64
Autumn	−6.35	−0.04	−0.06	−2.93	0.01	−3.05	3.32	6.10	0.30	−0.05	1.24	0.08	−2.85	1.63
Winter	−3.84	0.22	0.06	0.19	−0.16	−0.78	2.47	2.29	−0.40	−0.05	−0.19	−0.28	−2.31	1.09
Spring	−12.02	0.17	−0.75	−1.78	0.06	−2.54	0.45	5.69	0.14	0.21	2.65	1.13	−3.09	0.57
2013–2014	0.47	1.18	−0.06	0.60	−0.23	0.08	1.58	0.85	1.89	−0.01	0.56	0.13	2.22	2.64

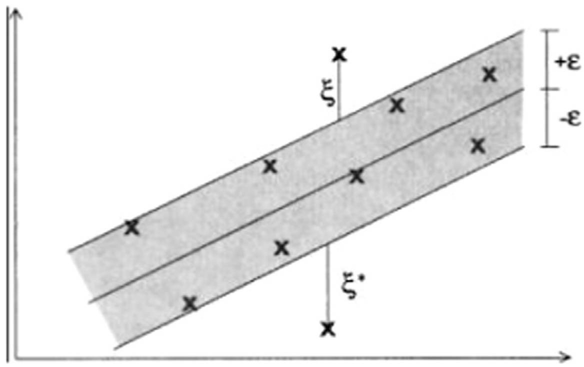

Fig. 10. SVR Tube with new margins and errors [34].

were trained with only the temporal parameters and holiday & business day indicator.

It was also noted that when models were trained with seasonal sets for forecasting daily peak loads, due to the relatively smaller size of the data set (90 points per season) increasing the number of influence parameters helped to improve the performance of the models (especially SVR and RT).

A limitation of this study is not being able to incorporate ARIMA models in to the analysis. Various combinations of *p, d* and *q* (see Section 2.2) and different transformations such as log and square root were investigated in order to find an improved model for this study. This process was done after observing autocorrelation plots ACF and PACF and Dickey-Fuller tests [62]. However, the models did not produce meaningful results on the test sets. Some authors have pointed out the difficulties involved in forecasting data sets which have complex seasonalities [63,64]. This study's unsuccessful attempts at using ARIMA models for forecasting can be attributed to not being able to remove the seasonal non-stationarities within the data set.

6. Conclusion

From a review of the literature it was found that regression models are commonly used not only for forecasting total building electricity loads but also for HVAC loads and retrofit savings. The DBT was the most pronounced climate variable used in previous regression studies and our analysis also confirmed this finding. Through the single regression analysis, DBT was found to be the most significant predictor for both the Campus and the TETB electricity loads amongst the other weather parameters. When the temporal parameters were included in the MLR stepwise selection, most of the weather parameters failed to give p-values smaller than 0.05 and were not included in the final models. However DBT was the only exception as it was included in almost all of the models. Another result from our study that supported the literature findings was the significance of holiday and business day binary indicator as a predictor for building electricity loads. The significance of this parameter was even greater for the daily peak load MLR models.

In contrast to the literature findings where DBT based single regression models showed fair forecasting capabilities [16,44], our single regression models which consisted of different climate parameters failed to show adequate forecast performance. This raised the importance of other parameters which have impact on the load. Especially for university Campuses and buildings, occupancy can have a significant effect on electricity loads therefore it should be taken into

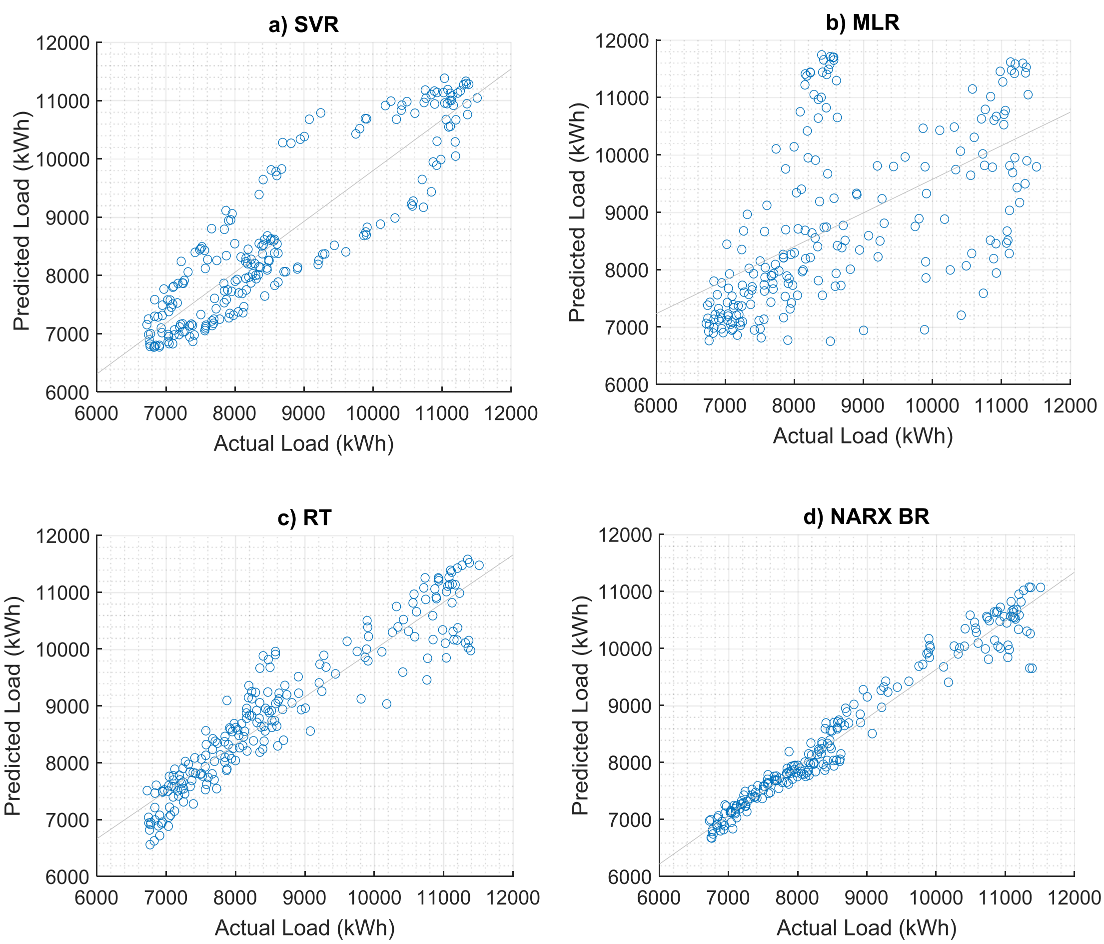

Fig. 11. Predicted vs actual Campus load scatter plots for four of the models between June 16 to June 23, 2014.

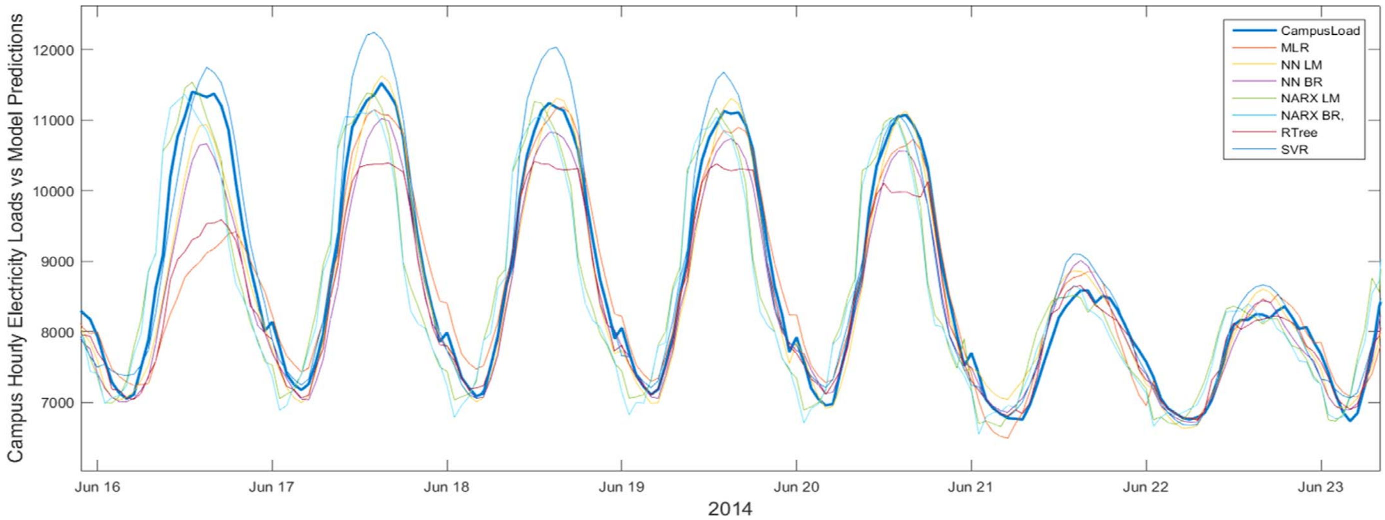

Fig. 12. Campus hourly electricity load forecasts vs model predictions from June 16 to June 23, 2014.

account whenever possible. In stepwise regression analysis, temporal parameters were observed to be the most significant predictors and the previous day same hour was the first included variable during the selection. It is possible that some proportion of the variation brought about by variations in occupancy can be captured by using the past values of the load in MLR models. However this needs to be further tested with the actual occupancy data.

Most of the machine learning models used in this study showed a better forecast performance than the MLR models however the MLR analysis enabled greater user engagement and control over the forecast analysis which is a comparative advantage of regression models over machine learning models. Machine learning models used for the study were pre-made, highly complex optimization packages, which are often referred as 'black box models'. Although the mathematical background of these models does not need to be known in detail, the user is expected to be able to know the operational principles and model diagnosis. Our analysis proved that model diagnosis can be a cumbersome and time consuming process, yet the effort spent can bring significant improvements in model accuracy and performance. In particular, utilizing a larger training set, both increasing and reducing the number of influence parameters and optimizing the relevant model parameters were found to improve the model performance.

Another advantage of regression analysis over machine learning models was found to be their straightforward implementation and relative ease of use. However for applications where forecast accuracy is highly crucial, machine learning models can have a significant advantage over MLR models. Since forecasting peak electricity loads has significant importance for many commercial buildings, machine learning models should be considered.

This study showed that forecasting daily peak electricity loads is a more difficult task than forecasting the day ahead hourly electricity loads since all the models showed higher RMSE and bias for the former. In addition, almost all the models showed a better forecast performance for the Campus demand in comparison to the single TETB demand for both day ahead hourly and peak loads. The forecast analysis for smaller scale applications such as residential load forecasting has been acknowledged as a more difficult task than commercial load building forecasts [65] . The same analogy can be seen in our example where, as the scale gets bigger, smaller and instantaneous changes in consumption have a smaller impact on the overall load of the Campus, therefore models can analyse and forecast this steadier load data more accurately.

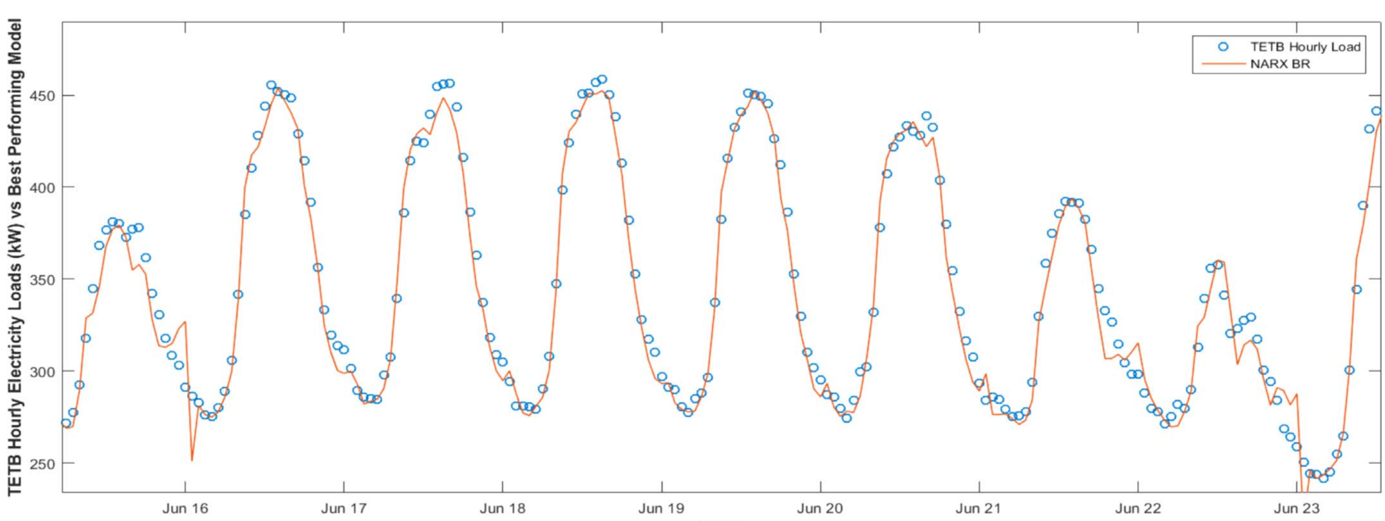

Fig. 13. TETB hourly electricity load vs best performing model prediction with NARX BR from June 16 to June 23, 2014.

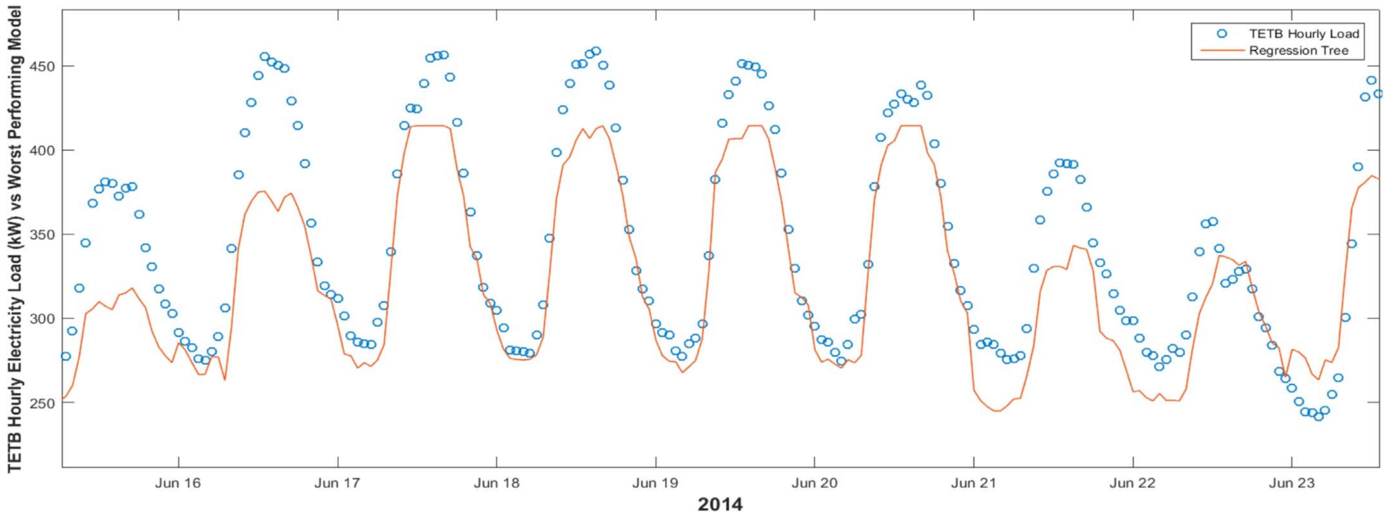

Fig. 14. TETB hourly electricity load vs worst performing model prediction with Regression Trees from June 16 to June 23, 2014.

Acknowledgements

Baran Yildiz would like to thank UNSW Facilities Management for the provided data-set and access to the equipment used to conduct this research.

Appendix A

Error terms

To measure and compare the performance of the models the following error terms are used:

$$R^2 = 1 - \frac{SSE}{SST} \quad \text{(Coefficient of Determination)} \quad R^2_{adj} = 1 - \frac{SSE}{SST} \times \frac{n-1}{n-k-1} \quad \text{(Adjusted Coefficient of Determination)}$$

$$RMSE\ (\%) = \frac{\sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n}}}{\bar{y}} \quad \text{(Root Mean Squared Error)} \quad MBE\ (\%) = 1/n \sum_{i=1}^{n} \frac{(y_i - \hat{y}_i)}{y_i} \times 100 \quad \text{(Mean Bias Error)}$$

$$MAPE = 1/n \sum_{i=1}^{n} \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100 \quad \text{(Mean Absolute Percentage Error)}$$

In the above equations n represents the number of observations, k is the number of influence parameters (predictors), SSE and SST are the unexplained and total variability of the measured output (load) respectively [66].

R^2 and $R_{adj}{}^2$ are used to measure the wellness of the fit by the trained models. $R_{adj}{}^2$ is a more useful parameter for MLR than R^2 since it is a better indicator of whether introducing a new parameter adds any value to the model (regardless of the relationship with the parameter and output, adding a new parameter to the model results in an increase in R^2 however $R_{adj}{}^2$ only increases if the parameter is actually a good predictor for the output). For large datasets, both have similar values since the penalty term ($\frac{n-1}{n-k-1}$) approaches unity when n is large [66].

MAE is a commonly used average error metric which is the mean value of the sum of absolute errors. RMSE is one of the most commonly used metric for describing uncertainty and it is a function of both MAE and the distribution of error magnitudes. RMSE penalizes the larger error terms and tends to become increasingly larger than MAE as the distribution of the errors magnitude becomes more variable. Previous papers have discussed the advantages of using MAE over RMSE [67] and others have favoured the use of RMSE over MAE [68]. We therefore decided to present both metrics as well as the mean bias error (MBE) which indicates by how much the model predictions are on average above or below the average measured output value.

For our forecasting study, p-value is used to measure the statistical significance of an influence parameter on the load during stepwise regression. Note that p-values smaller than 0.05 represent the rejection of the null hypothesis where the null hypothesis can be defined as: There is no statistically significant relationship between a particular influence parameter and load. Therefore influence parameters which had p-values smaller than 0.05 rejected the null hypothesis and were included in the model.

References

[1] International Energy Agency. Electricity Information. Paris; 2012.
[2] EIA . Electric power monthly: with data for January 2015. US Energy Inf Adm; 2015.
[3] Pitt Sherry. Final report : quantitative assessment of energy savings from building energy efficiency; 2013.
[4] Granade HC, Creyts J, Derkach A, Farese P, Nyguist S, Ostrowski K. Unlocking energy efficiency in the U.S. economy. McKinsey Glob Energy Mater 2009:1–165.
[5] Lister M Scaling the peaks : demand management in network infrastructure planning key questions for the stage 3 review; 2011.
[6] Smith M, Hargroves K, Stasinopoulos P, Stephens R, Desha CH. Energy transformed: sustainable energy solutions for climate change mitigation; 2007.

[7] Steinfeld J, Bruce A, Watt M. Peak load characteristics of Sydney office buildings and policy recommendations for peak load reduction. Energy Build 2011;43:2179–87. http://dx.doi.org/10.1016/j.enbuild.2011.04.022.

[8] Neto AH, Fiorelli FAS. Comparison between detailed model simulation and artificial neural network for forecasting building energy consumption. Energy Build 2008;40:2169–76. http://dx.doi.org/10.1016/j.enbuild.2008.06.013.

[9] Zhao H, Magoulès F. A review on the prediction of building energy consumption. Renew Sustain Energy Rev 2012;16:3586–92. http://dx.doi.org/10.1016/j.rser.2012.02.049.

[10] Alfares HK, Nazeeruddin M. Electric load forecasting: literature survey and classification of methods. Int J Syst Sci 2002;33:23–34. http://dx.doi.org/10.1080/00207720110067421.

[11] Foucquier A, Robert S, Suard F, Stephan L, Jay A. State of the art in building modelling and energy performances prediction: a review. Renew Sustain Energy Rev 2013;23:272–88. http://dx.doi.org/10.1016/j.rser.2013.03.004.

[12] Suganthi L, Samuel AA. Energy models for demand forecasting – a review. Renew Sustain Energy Rev 2012;16:1223–40. http://dx.doi.org/10.1016/j.rser.2011.08.014.

[13] Harish VSKV, Kumar A. A review on modeling and simulation of building energy systems. Renew Sustain Energy Rev 2016;56:1272–92. http://dx.doi.org/10.1016/j.rser.2015.12.040.

[14] Ahmad AS, Hassan MY, Abdullah MP, Rahman HA, Hussin F, Abdullah H, et al. A review on applications of ANN and SVM for building electrical energy consumption forecasting. Renew Sustain Energy Rev 2014;33:102–9. http://dx.doi.org/10.1016/j.rser.2014.01.069.

[15] Wang Z, Srinivasan RS A review of artificial intelligence based building energy use prediction : contrasting the capabilities of single and ensemble prediction models. Renew Sustain Energy Rev; n.d. 0–1. http://dx.doi.org/10.1016/j.rser.2016.10.079.

[16] Katipamula S, Reddy T, Claridge DE. Multivariate regression modeling. J Sol Energy Eng 1998;120:177. http://dx.doi.org/10.1115/1.2888067.

[17] Braun MR, Altan H, Beck SBM. Using regression analysis to predict the future energy consumption of a supermarket in the UK. Appl Energy 2014;130:305–13. http://dx.doi.org/10.1016/j.apenergy.2014.05.062.

[18] Katipamula S, Kissock JK, Claridge DE. The functional basis of steady- state tliermal energy use in air- side HVAC Equipment; 1995. p. 117.

[19] Nassif N. Regression models for estimating monthly energy consumptions in schools in hot and humid climates. ASHRAE Trans 2012;118:225–32. http://dx.doi.org/10.1080/01998595.2013.10750242.

[20] Ramanathan R, Engle R, Granger CWJ, Vahid-Araghi F, Brace C. Short-run forecasts of electricity loads and peaks. Int J Forecast 1997;13:161–74. http://dx.doi.org/10.1016/S0169-2070(97)00015-0.

[21] Lam JC, Wan KKW, Liu D, Tsang CL. Multiple regression models for energy use in air-conditioned office buildings in different climates. Energy Convers Manag 2010;51:2692–7. http://dx.doi.org/10.1016/j.enconman.2010.06.004.

[22] Yang L, Lam JC, Liu J, Tsang CL. Building energy simulation using multi-years and typical meteorological years in different climates. Energy Convers Manag 2008;49:113–24. http://dx.doi.org/10.1016/j.enconman.2007.05.004.

[23] Turiel I, Boschen R, Seedall M, Levine M. Simplified energy analysis methodology for commercial buildings. Energy Build 1984;6:67–83. http://dx.doi.org/10.1016/0378-7788(84)90008-2.

[24] Hygh JS, DeCarolis JF, Hill DB, Ranji Ranjithan S. Multivariate regression as an energy assessment tool in early building design. Build Environ 2012;57:165–75. http://dx.doi.org/10.1016/j.buildenv.2012.04.021.

[25] Bauer M, Scartezzini J-L. A simplified correlation method accounting for heating and cooling loads in energy-efficient buildings. Energy Build 1998;27:147–54. http://dx.doi.org/10.1016/S0378-7788(97)00035-2.

[26] Tso GKF, Yau KKW. Predicting electricity energy consumption: a comparison of regression analysis, decision tree and neural networks. Energy 2007;32:1761–8. http://dx.doi.org/10.1016/j.energy.2006.11.010.

[27] Fan C, Xiao F, Wang S. Development of prediction models for next-day building energy consumption and peak power demand using data mining techniques. Appl Energy 2014;127:1–10. http://dx.doi.org/10.1016/j.apenergy.2014.04.016.

[28] Gómez JR, Gregio K, Santo D, Antonio M. Linear and non-linear methods for prediction of peak load at University of São Paulo, vol. 78; 2016. p. 187–201. http://dx.doi.org/10.1016/j.measurement.2015.09.053.

[29] Basu K, Hawarah L, Arghira N, Joumaa H, Ploix S. A prediction system for home appliance usage. Energy Build 2013;67:668–79. http://dx.doi.org/10.1016/j.enbuild.2013.02.008.

[30] Lazos D, Sproul AB, Kay M. Optimisation of energy management in commercial buildings with weather forecasting inputs: a review. Renew Sustain Energy Rev 2014;39:587–603. http://dx.doi.org/10.1016/j.rser.2014.07.053.

[31] Espinoza M, Joye C, Belmans R, De Moor B. Short-term load forecasting, profile identification, and customer segmentation: a methodology based on periodic time series. IEEE Trans Power Syst 2005;20:1622–30. http://dx.doi.org/10.1109/TPWRS.2005.852123.

[32] Amjady N. Short-term hourly load forecasting using time-series modeling with peak load estimation capability. IEEE Trans Power Syst 2001;16:798–805. http://dx.doi.org/10.1109/59.962429.

[33] Huang H, Chen L, Hu E. A neural network-based multi-zone modelling approach for predictive control system design in commercial buildings. Energy Build 2015;97:86–97. http://dx.doi.org/10.1016/j.enbuild.2015.03.045.

[34] Clarke SM, Griebsch JH, Simpson TW. Analysis of support vector regression for approximation of complex engineering analyses. J Mech Des 2005;127:1077. http://dx.doi.org/10.1115/1.1897403.

[35] Borges CE, Penya YK, Fernández I. Optimal combined short-term building load forecasting. 2011 IEEE PES Innov Smart Grid Technol ISGT Asia 2011 Conference Smarter Grid Sustain Afford Energy Futur; 2011 ⟨doi:10.1109/ISGT-Asia.2011.6167091⟩.

[36] Chen B-J, Chang M-W, Lin C-J. Load forecasting using support vector machines: a study on EUNITE competition 2001. IEEE Trans Power Syst 2004(19):1821–30. http://dx.doi.org/10.1109/TPWRS.2004.835679.

[37] Tibshirani R, Wasserman L. Statistical machine learning. Carnegie Mellon Univ; 2006. p. 1–7.

[38] Papalexopoulos AD, Hesterberg TC. A regression-based approach to short-term system load forecasting – power systems. IEEE Trans Power Syst 1990;5:1535–50.

[39] Reddy T, Claridge D. Using synthetic data to evaluate multiple regression and principal component analyses for statistical modeling of daily building energy consumption. Energy Build 1994;21:35–44. http://dx.doi.org/10.1016/0378-7788(94)90014-0.

[40] Sever F, Kissock K, Brown D, Mulqueen S. Estimating industrial building energy savings using inverse simulation. ASHRAE Trans 2011;117:348–55.

[41] Kissock JK, Reddy TA, Claridge DE. Ambient-temperature regression analysis for estimating retrofit savings in commercial buildings. J Sol Energy Eng 1998;120:168. http://dx.doi.org/10.1115/1.2888066.

[42] Walter T, Sohn MD. A regression-based approach to estimating retrofit savings using the building performance database. Appl Energy 2016;179:996–1005. http://dx.doi.org/10.1016/j.apenergy.2016.07.087.

[43] Fels M. PRISM: an introduction. Cent Energy Environ Stud Princt Univ, vol. 9; 1986. p. 5–18.

[44] Lam JC, Chan RYC, Li DHW. Simple regression models for fully air-conditioned public sector office buildings in subtropical climates. Arch Sci Rev 2002;45:361–9. http://dx.doi.org/10.1080/00038628.2002.9696951.

[45] Kissock K, Joseph HB, McBride JR. Effects of varying indoor air temperature and heat gain on the measurement of retrofit savings. ASHRAE Trans 1998;104:895–900.

[46] Morgan MG. Uncertainty: a guide to dealing with uncertainty in quantitative risk and policy analysis. New York; 1990.

[47] Green Building Council of Australia. Introducing Green Star; 2013.

[48] Chih-Wei Hsu, Chih-Chung Chang, C-JL . Library for support vector machines. ACM Trans Intell Syst Technol 2011;2:27. http://dx.doi.org/10.1177/02632760022050997.

[49] Beccali M, Cellura M, Lo Brano V, Marvuglia A. Short-term prediction of household electricity consumption: assessing weather sensitivity in a Mediterranean area. Renew Sustain Energy Rev 2008;12:2040–65. http://dx.doi.org/10.1016/j.rser.2007.04.010.

[50] Padfield T. Equations describing the physical properties of moist air. Creat Commons 2015:1–4.

[51] James G, Witten D, Hastie T, Tibshirani R. An introduction to statistical learning. vol. 102; 2006. http://dx.doi.org/10.1016/j.peva.2007.06.006.

[52] Smith LI. A tutorial on principal components analysis introduction. Department of Computer Science Univ Otago, vol. 51; 2002. p. 52. http://dx.doi.org/10.1080/03610928808829796.

[53] NG A. Advice for applying machine learning. Stanford Univ; 2015 ⟨https://www.coursera.org/learn/machine-learning/lecture/zJTzp/deciding-what-to-do-next-revisited⟩.

[54] Sinclair . Identification and investigation of peak demand reduction opportunities – Sydney CBD area. Sinclair Kn Merz Consult; 2006.

[55] NG A. Neural networks learning in machine learning. Stanford Univ; 2015 ⟨https://www.coursera.org/learn/machine-learning/lecture/Wh6s3/putting-it-together⟩.

[56] Jesus DE, Hagan MT. Backpropagation through time for a general class of recurrent network. Proc Int Jt Conf Neural Networks, Washingtonm DC; 2001. p. 2638–42.

[57] MATLAB. Nerual network toolbox; 2015.

[58] Beale MH, Hagan MT, Demuth HB. Neural network toolbox^{TM} user's guide 2015. http://dx.doi.org/10.1016/j.neunet.2005.10.002.

[59] MathWorks . Decision trees. MATLAB Doc; 2015 ⟨http://au.mathworks.com/help/stats/classification-trees-and-regression-trees.html#bsw6a96⟩.

[60] Ng A. Support vector machines, 1. Stanford Univ; 2000. p. 1–25. http://dx.doi.org/10.1016/j.aca.2011.07.027.

[61] Sauptik Dhar, Han Tai S. Support Vector Machine Regression. Univ Minnesota, CSci, Mach Learn; 2011. ⟨http://www.ece.umn.edu/users/cherkass/ee4389/sw/scaleSVM.m⟩.

[62] Shumway RH, Stoffer DS. Time series analysis and its applications with R examples; 2011. http://dx.doi.org/10.1007/978-1-4419-7865-3.

[63] Taylor JW. Short-term electricity demand forecasting using double seasonal exponential smoothing. J Oper Res Soc 2003;54:799–805. http://dx.doi.org/10.1057/palgrave.jors.2601589.

[64] Livera AMDe, Hyndman RJ, Snyder RD. Forecasting time series with complex seasonal patterns using exponential smoothing Forecasting time series with complex seasonal patterns using exponential smoothing.October 2010:to appear.

[65] Sevlian R, Rajagopal R. Short term electricity load forecasting on varying levels of aggregation; 2014.

[66] Diez DM, Barr CD, Çetinkaya-Rundel M. Open intro statistics, Third ed. Creat Commons; 2015.

[67] Willmott CJ, Matsuura K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. Clim Res 2005;30:79–82. http://dx.doi.org/10.3354/cr030079.

[68] Chai T, Draxler RR. Root mean square error (RMSE) or mean absolute error (MAE) -Arguments against avoiding RMSE in the literature. Geosci Model Dev 2014;7:1247–50. http://dx.doi.org/10.5194/gmd-7-1247-2014.