Advanced Engineering Informatics 35 (2018) 1–16

Contents lists available at ScienceDirect

Advanced Engineering Informatics

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

Full length article

Short-term electricity demand forecasting with MARS, SVR and ARIMA models using aggregated demand data in Queensland, Australia

Mohanad S. Al-Musaylh[a,b,*], Ravinesh C. Deo[a,d,*], Jan F. Adamowski[c], Yan Li[a]

[a] School of Agricultural, Computational and Environmental Sciences, Institute of Agriculture and Environment (IAg&E), University of Southern Queensland, QLD 4350, Australia
[b] Management Technical College, Southern Technical University, Basrah, Iraq
[c] Department of Bioresource Engineering, Faculty of Agricultural and Environmental Science, McGill University, Québec H9X 3V9, Canada
[d] Cold and Arid Regions Environmental and Engineering Research Institute, Chinese Academy of Sciences, Lanzhou, China

ARTICLE INFO

Keywords:
Electricity demand forecasting
Machine learning
SVR
MARS
ARIMA

ABSTRACT

Accurate and reliable forecasting models for electricity demand (G) are critical in engineering applications. They assist renewable and conventional energy engineers, electricity providers, end-users, and government entities in addressing energy sustainability challenges for the National Electricity Market (NEM) in Australia, including the expansion of distribution networks, energy pricing, and policy development. In this study, data-driven techniques for forecasting short-term (24-h) G-data are adopted using 0.5 h, 1.0 h, and 24 h forecasting horizons. These techniques are based on the Multivariate Adaptive Regression Spline (MARS), Support Vector Regression (SVR), and Autoregressive Integrated Moving Average (ARIMA) models. This study is focused in Queensland, Australia's second largest state, where end-user demand for energy continues to increase. To determine the MARS and SVR model inputs, the partial autocorrelation function is applied to historical (area aggregated) G data in the training period to discriminate the significant (lagged) inputs. On the other hand, single input G data is used to develop the univariate ARIMA model. The predictors are based on statistically significant lagged inputs and partitioned into training (80%) and testing (20%) subsets to construct the forecasting models. The accuracy of the G forecasts, with respect to the measured G data, is assessed using statistical metrics such as the Pearson Product-Moment Correlation coefficient (r), Root Mean Square Error ($RMSE$), and Mean Absolute Error (MAE). Normalized model assessment metrics based on $RMSE$ and MAE relative to observed means ($RMSE_{\overline{G}}$ and $MAE_{\overline{G}}$), Willmott's Index (WI), Legates and McCabe Index (E_{LM}), and Nash–Sutcliffe coefficients (E_{NS}) are also utilised to assess the models' preciseness. For the 0.5 h and 1.0 h short-term forecasting horizons, the MARS model outperforms the SVR and ARIMA models displaying the largest WI (0.993 and 0.990) and lowest MAE (45.363 and 86.502 MW), respectively. In contrast, the SVR model is superior to the MARS and ARIMA models for the daily (24 h) forecasting horizon demonstrating a greater WI (0.890) and MAE (162.363 MW). Therefore, the MARS and SVR models can be considered more suitable for short-term G forecasting in Queensland, Australia, when compared to the ARIMA model. Accordingly, they are useful scientific tools for further exploration of real-time electricity demand data forecasting.

Abbreviations: MW, Megawatt; G, Electricity load (demand, Mega Watts); MARS, Multivariate Adaptive Regression Splines; SVR, Support Vector Regression; ARIMA, Autoregressive Integrated Moving Average; r, Correlation Coefficient; $RMSE$, Root Mean Square Error (MW); MAE, Mean Absolute Error (MW); $RMSE_{\overline{G}}$, Relative Root Mean Square Error, %; $MAE_{\overline{G}}$, Mean Absolute Percentage Error, %; WI, Willmott's Index of Agreement; E_{NS}, Nash–Sutcliffe Coefficient; E_{LM}, Legates and McCabe Index; ANN, Artificial Neural Network; RBF, Radial Basis Function for SVR; σ, Kernel Width for SVR Model; C, Regulation for SVR Model; $BF_m(X)$, Spline Basis Function for MARS; GCV, Generalized Cross-Validation; p, Autoregressive Term in ARIMA; D, Degree of Differencing in ARIMA; Q, Moving Average Term in ARIMA; AEMO, Australian Energy Market Operator; NEM, National Electricity Market; ACF, Auto-Correlation Function; PACF, Partial Auto-Correlation Function; MSE, Mean Square Error (MW); R^2, Coefficient of Determination; AIC, Akaike Information Criterion; L, Log Likelihood; σ^2, Variance; G_i^{for}, i^{th} Forecasted Value of G, MW); G_i^{obs}, i^{th} Observed Value of G, MW); Q_{25}, Lower Quartile (25th Percentile); Q_{50}, Median Quartile (50th Percentile); Q_{75}, Upper Quartile (75th Percentile); d, Degree of Differencing in ARIMA

* Corresponding authors at: School of Agricultural, Computational and Environmental Sciences, Institute of Agriculture and Environment (IAg&E), University of Southern Queensland, QLD 4350, Australia.
E-mail addresses: MohanadShakirKhalid.AL-Musaylh@usq.edu.au, mohanadk21@gmail.com (M.S. Al-Musaylh), ravinesh.deo@usq.edu.au (R.C. Deo).

https://doi.org/10.1016/j.aei.2017.11.002
Received 8 April 2017; Received in revised form 18 November 2017; Accepted 20 November 2017
Available online 01 December 2017
1474-0346/

1. Introduction

Electricity load forecasting (also referred to as demand and abbreviated as *G* in this paper, MW) plays an important role in the design of power distribution systems [1,2]. Forecast models are essential for the operation of energy utilities as they influence load switching and power grid management decisions in response to changes in consumers' needs [3]. *G* forecasts are also valuable for institutions related to the fields of energy generation, transmission, and marketing. The precision of *G* estimates is critical since a 1% rise in load forecasting error can lead to a loss of millions of dollars [4–6]. Over- or under-projections of *G* can endanger the development of coherent energy policies and hinder the sustainable operation of a healthy energy market [7]. Furthermore, demographic, climatic, social, recreational, and seasonal factors can impact the accuracy of *G* estimates [1,8,9]. Therefore, robust forecasting models that can address engineering challenges, such as minimizing predictive inaccuracy in *G* data forecasting, are needed to, for example, support the sustainable operation of the National Electricity Market (NEM).

Qualitative and quantitative decision-support tools have been useful in *G* forecasting. Qualitative techniques, including the Delphi curve fitting method and other technological comparisons [6,10,11], accumulate experience in terms of real energy usage to achieve a consensus from different disciplines regarding future demand. On the other hand, quantitative energy forecasting is often applied through physics-based and data-driven (or black box) models that draw upon the inputs related to the antecedent changes in *G* data. The models' significant computational power has led to a rise in their adoption [12]. Data-driven models, in particular, have the ability to accurately forecast *G*, which is considered a challenging task [6]. Having achieved a significant level of accuracy, data-driven models have been widely adopted in energy demand forecasting (*e.g.*, [13,14]). Autoregressive Integrated Moving Average (ARIMA) [15], Artificial Neural Network (ANN) [16], Support Vector Regression (SVR) [17], genetic algorithms, fuzzy logic, knowledge-based expert systems [18], and Multivariate Adaptive Regression Splines (MARS) [19] are among the popular forecasting tools used by energy researchers.

The SVR model, utilised as a primary model in this study, is governed by regularization networks for feature extraction. The SVR model does not require iterative tuning of model parameters [20,21]. Its algorithm is based on the structural risk minimization (SRM) principle and aims to reduce overfitting data by minimizing the expected error of a learning machine [21]. In the last decades, this technique has been recognized and applied throughout engineering, including in forecasting (or regression analysis), decision-making (or classification works) processes and real-life engineering problems [22]. Additionally, the SVR models have been shown to be powerful tools when a time-series (*e.g.*, *G*) needs to be forecasted using a matrix of multiple predictors. As a result, their applications have continued to grow in the energy forecasting field. For example, in Turkey (Istanbul), several investigators have used the SVR model with a radial Basis Kernel Function (RBF) to forecast *G* data [23]. In eastern Saudi Arabia, the SVR model generated more accurate hourly *G* forecasts than a baseline autoregressive (AR) model [24]. In addition, different SVR models were applied by Sivapragasam and Liong [25] in Taiwan to forecast daily loads in high, medium, and low regions. In their study, the SVR model provided better predictive performance than an ANN approach for forecasting regional electric loads [29]. Except for one study that confirmed SVR models' ability to forecast global solar radiation [17], to the best of the authors' knowledge, a robust SVR forecasting model has been limitedly applied for energy demand. Thus, additional studies are needed to explore SVR modelling in comparison to other models applied in *G* forecasting.

Contrary to the SVR model, the MARS model has not been widely tested for *G* forecasting. It is designed to adopt piecewise (linear or cubic) basis functions [26,27]. In general, the model is a fast and flexible statistical tool that operates through an integrated linear and non-linear modelling approach [28]. More importantly, it has the capability of employing a set of basic functions using several predictor variables to assess their relationship with the objective variable through non-linear and multi-collinear analysis. This is important for demand forecasting based on interactions between different variables and the demand data. Although the literature on MARS models applied in the field of *G* forecasting is very scarce, this model has proven to be highly accurate in several estimation engineering challenges. Examples may be drawn from studies that discuss doweled pavement performance modelling, determination of ultimate capacity of driven piles in cohesionless soil, and analysis of geotechnical engineering systems [29–31]. In Ontario (Canada), the MARS model was applied, through a semiparametric approach, for forecasting short-term oil prices [32] and investigating the behaviour of short-term (hourly) energy price (HOEP) data through lagged input combinations [8]. Sigauke and Chikobvu [19] tested the MARS model for *G* forecasting in South Africa; this demonstrated its capability of yielding a significantly lower Root Mean Square Error (*RMSE*) when compared to piecewise regression-based models. However, despite its growing global applicability (*e.g.*, [26,27,33–35]), the MARS model remains to be explored for *G* forecasting in the present study region.

In the literature, the ARIMA model has generated satisfactory results for engineering challenges including the forecasting of electricity load data [15], oil [32], and gas demand [36]. A study in Turkey applied a co-integration method with an ARIMA model for *G*-estimation and compared results with official projections. It concluded that approximately 34% of the load was overestimated when compared to measured data from the ARIMA model [8]. Several studies have indicated that the ARIMA model tends to generate large errors for long-range forecasting horizons. For example, a comparison of the ARIMA model, the hybrid Grey Model (GM-ARIMA), and the Grey Model (GM(1, 1)) for forecasting *G* in China showed that GM (1, 1) outperformed the ARIMA model [37]. Similarly, a univariate ARAR model (i.e., a modified version of the ARIMA model) outperformed a classical ARIMA model in Malaysia [38]. However, to the best of the authors' knowledge, a comparison of the MARS, SVR, and ARIMA methods, each having their own merits and weaknesses, has not been undertaken in the field of G forecasting.

To explore opportunities in *G* forecasting, this paper discusses the versatility of data-driven techniques (multivariate MARS and SVR models and the univariate ARIMA model) for short-term half-hourly (0.5 h), hourly (1.0 h) and daily (24 h) horizon data. The study is beneficial to the field of power systems engineering and management since energy usage in Queensland continues to face significant challenges, particularly as it represents a large fraction (i.e., 23%) of the national 2012–2013 averaged energy demand [39]. The objectives of the study are as follows: (1) To develop and optimise the MARS, SVR, and ARIMA models for *G* forecasting using lagged combinations of the state-aggregated *G* data as the predictor variable; (2) To validate the optimal MARS, SVR, and ARIMA models for their ability to generate *G* forecasts at multiple forecasting horizons (i.e., 0.5, 1.0 and 24 h); and (3) To evaluate the models' preciseness over a recent period, [01-01-2012 to 31-12-2015 (dd-mm-yyyy)], by employing robust statistical metrics comparing forecasted and observed *G* data obtained from the Australian Energy Market Operator (AEMO) [40]. To evaluate and reach these objectives, this paper is divided into the following sections: Section 2 describes the theory of SVR, MARS, and ARIMA models; Section 3 presents the materials and methods including the *G* data and model development and evaluation; Section 4 presents the results and discussion; and Section 5 further discusses the results, research opportunities, and limitations. The final section summarizes the research findings and key considerations for future work.

2. Theoretical background

2.1. Support Vector regression

An SVR model can provide solutions to regression problems with multiple predictors $X = \{x_i\}_{i=1}^{i=n}$, where n is the number of predictor

variables and each x_i has N variables. These are linked to an objective variable $y = \{y_i\}_{i=1}^{i=N}$. The matrix X is converted to a higher-dimensional feature space, in accordance with the original, but constitutes a lower-dimensional input space [41,42]. With an SVR model, a non-linear regression problem is defined as [43]:

$$y = f(X) = \omega \cdot \varnothing(X) + b \tag{1}$$

where b is a constant, ω is the weighted vector, and $\varnothing(X)$ denotes the mapping function employed in the feature space. The coefficients ω and b are estimated by the minimisation process below [43]:

$$\text{Minimize} \quad \frac{1}{2}\|w\|^2 + C\frac{1}{N}\sum_{i=1}^{N}(\xi_i + \xi_i^*) \tag{2}$$

$$\text{Subject to} \begin{cases} |y_i - (w, x_i + b) \geqslant \varepsilon + \xi_i \\ \langle w, x_i \rangle + b - y_i \leqslant \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geqslant 0 \end{cases} \tag{3}$$

where C and ε are the model's prescribed parameters. The term of $\frac{1}{2}\|w\|^2$ measures the smoothness of the function and C evaluates the trade-off between the empirical risk and smoothness. ξ and ξ^* are positive slack variables representing the distance between actual and corresponding boundary values in the ε-tube model of function approximation.

After applying Lagrangian multipliers and optimising conditions, a non-linear regression function is obtained [43]:

$$f(X) = \sum_{i=1}^{i=N}(\alpha_i - \alpha_i^*)K(x_i, x_j) + b \tag{4}$$

where α_i and α_i^* are Lagrangian multipliers and the term $K(x_i, x_j)$ is the kernel function describing the inner product in D-dimensional feature space, x_i and $x_j \in X$ [43]. Under Kuhn-Tucker conditions, a limited number of α_i and α_i^* coefficients will be non-zero [17]. The associated data points, termed the "support vectors", lie the closest to the decision surface (or hyperplane) [17]. The radial basis function (RBF) employed in developing the SVR model in this study, can be expressed as [44]:

$$K(x_i, x_j) = \exp\left(\frac{-\|x_i - x_j\|^2}{2\sigma^2}\right) \tag{5}$$

where x_i and x_j are the inputs in the i^{th} and j^{th} respective dimensions and σ is the kernel width. Over the training period, the support vectors' area of influence with respect to input data space is determined by kernel width (σ) and regulation (C). Deducing these can represent a critical task for achieving superior model accuracy [17]. This is performed through a grid-search procedure (Section 3.2).

2.2. Multivariate adaptive regression splines

The MARS model, first introduced by Friedman [28], implements the piecewise regression process for feature identification of the input dataset. In addition, it has the capability to flexibly and efficiently analyse the relationships between a given predictand (*i.e.*, the G in context of the present study) and a set of predictor variables (*i.e.*, the lagged combinations of G). In general, the MARS model can analyse non-linearities in predictor-predictand relationships when forecasting a given predictand [45].

Assuming two variable matrices, X and y, where X is a matrix of descriptive variables (predictors) over a domain $D \subset \mathbb{R}^n$, $X = \{x_i\}_{i=1}^{i=n}$, and y is a target variable (predictand), there are then N realizations of the process $\{y_i, x_{1i}, x_{2i}, \ldots, x_{ni}\}_1^N$ [8]. Consequently, the MARS model relationship between X and y is demonstrated below [28]:

$$y = \widehat{f}(X) = a_0 + \sum_{m=1}^{M} a_m BF_m \tag{6}$$

where a_0 is a constant, $\{a_m\}_1^M$ are the model coefficients estimated to produce data-relevant results, M is the number of subregions $R_m \subset D$ or the equivalent basis functions in MARS, and $BF_m(X)$ is a spline function defined as $C(X|s, t_1, t, t_2)$. In the latter, $t_1 < t < t_2$, and s have a value of $+1$ or -1 for a spline basis function or its mirror image, respectively.

The Generalized Cross-Validation criterion (GCV) used by the MARS model assesses the lack-of-fit of the basis functions through the Mean Square Error (MSE) [28] and is expressed as:

$$GCV = MSE \Big/ \left[1 - \frac{\widetilde{G}(M)}{N}\right]^2 \tag{7}$$

where $MSE = \frac{1}{N}\sum_{I=1}^{N}[y_i - \widehat{f}(X_i)]^2$ and $\left[1 - \frac{\widetilde{G}(M)}{N}\right]^2$ is a penalty that accounts for an increasing variance from a complex model. Furthermore, $\widetilde{G}(M)$ is defined as [28]:

$$\widetilde{G}(M) = C(M) + v \cdot M \tag{8}$$

where v is a penalty factor with a characteristic value of $v = 3$ and $C(M)$ is the number of parameters being fitted.

The MARS model with the lowest value of the GCV for the training dataset is considered the optimal model.

2.3. Autoregressive integrated moving average

Relying on the antecedent data to forecast G, the ARIMA model constitutes a simplistic, yet popular approach applied for time-series forecasting. ARIMA was popularized by the work of Box and Jenkins [46]. To develop the ARIMA model, two types of linear-regressions are integrated: the Autoregressive (AR) and the Moving Average (MA) [46].

The AR model is written as [46]:

$$y_t = c + a_1 y_{t-1} + \ldots + a_p y_{t-p} + u_t \tag{9}$$

where $a_1, \ldots, a_p$ are the AR parameters, c is a constant, p is the order of the AR, and u_t is the white noise.

Likewise, the MA model can be written as [46]:

$$y_t = \mu + u_t + m_1 u_{t-1} + \ldots + m_q u_{t-q} \tag{10}$$

where $m, \ldots, m_q$ are the MA parameters, q is the order of MA, $u_t, u_{t-1}, \ldots, u_{t-q}$ are the white noise (error) terms, and μ is the expectation of y_t.

By integrating these models with the same training data, the ARIMA model [ARIMA (p,q)] becomes [46]:

$$y_t = c + a_1 y_{t-1} + \ldots + a_p y_{t-p} + u_t + m_1 u_{t-1} + \ldots + m_q u_{t-q} \tag{11}$$

where p and q are the autoregressive and moving average terms, respectively.

The basic premise of this model is that time-series data incorporates statistical stationarity which implies that measured statistical properties, such as the mean, variance, and autocorrelation remain constant over time [47]. However, if the training data displays non-stationarity, as is the case with real-life predictor signals (*e.g.*, G data), the ARIMA model requires differenced data to transform it to stationarity. This is denoted as ARIMA (p,d,q) where d is the degree of differencing [37].

3. Materials and methods

3.1. Electricity demand data

In this study, a suite of data-driven models was developed for short-term G forecasting in Queensland, Australia. The predictor data, comprised of half-hourly (48 times per day) G records for a period between 01-01-2012 to 31-12-2015 (dd-mm-yyyy), was acquired from the Australian Energy Market Operator (AEMO) [40]. The AEMO database aims to provide G data, in terms of relevant energy consumption, for the Queensland region of the NEM. Hence, these data have been previously used in various forecasting applications (*e.g.*, [48,49]). However, they have not been employed in machine learning models as attempted in the present study.

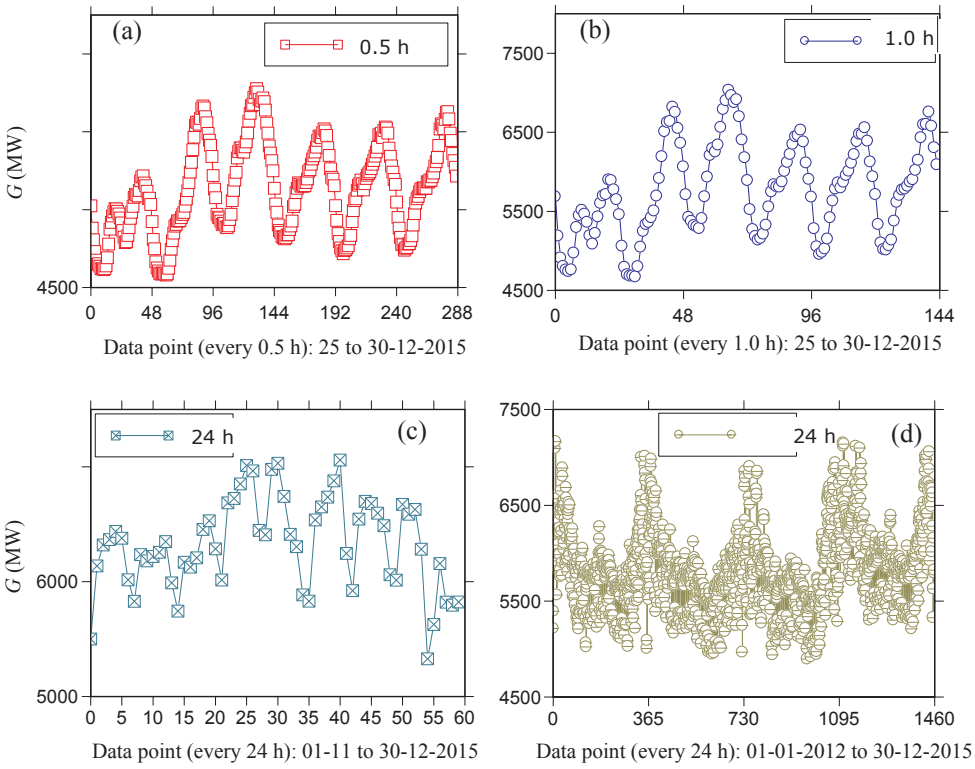

Fig. 1. Time-series of electricity demand (G) data and various forecasting periods.

In the present study, the 0.5 h time-step corresponds to the NEM settlement periods 1 (0:00 h–0:30 h) through 48 (23:30 h–24:00 h). The 0.5 h interval readings, reported in other research works (*e.g.*, [48,49]), were thus used for short-term forecasting of the G data. To expand the forecasting horizon to 1.0 h and 24 h periods to obtain G values, an arithmetic averaging of the half-hourly data was performed. The MARS, SVR, and ARIMA models considered in this paper, developed and evaluated 0.5 h, 1.0 h and 24 h forecasts utilising data from periods 01-12-2015 to 31-12-2015, 01-11-2015 to 31-12-2015, and 01-01-2012 to 31-12-2015, respectively. In principle, the number of predictive features remained similar throughout (*i.e.*, approximately 1460 data points for each horizon).

Fig. 1(a–d) depicts plots of the aggregated G data for the Queensland region, whereas Table 1 provides its associated descriptive statistics. The stochastic components, present in G data at the 0.5 h and 1.0 h time-scales, exhibit fluctuations due to the change in consumer electricity demands. This is confirmed by the large standard deviation and high degree of skewness observed for the 0.5 h and 1.0 h scale when compared to those associated with the 24 h scale in Table 1.

3.2. Forecast model development

Data-driven models incorporate historical G data to forecast future G values. The initial selection of (lagged) input variables to determine the predictors is critical for developing a robust multivariate (SVR or MARS) model [17,26]. The literature outlines two input selection methods for determining the sequential time series of lagged G values that provide an optimal performance. These are (i) trial and error and (ii) an auto-correlation function (ACF) or partial auto-correlation function (PACF) approach. For this study, patterns were analysed in historical G data from the training period using the ACF and PACF to extract correlation statistics [50–52]. This approach employed time-lagged information to analyse the period between current and antecedent G values at specific points in the past (*i.e.*, applying a time lag) and assessed any temporal dependencies existing in the time-series. Subsequently, inputs for each time lag (0.5 h, 1.0 h, 24 h) were identified by statistical verification of lagged G combinations and their respective correlation coefficient (r).

The PACF for G data, depicted in Fig. 2, aided in identifying potential inputs for data-driven models. The method computed a time-series regression against its n-lagged-in-time values that removed the dependency on intermediate elements and identified the extent to which G was correlated to the antecedent timescale value. Consequently, the statistically correlated signal $G(t)$ and the respective n-lagged signals were selected. This procedure developed forecast models that considered the role of memory (*i.e.*, antecedent G) in forecasting the current G. The 15 modelling scenarios, presented in Table 2, were developed based on the MARS and SVR algorithms.

For the 0.5 h and 24 h forecasting horizons, the models employed half-hourly and daily data from the 1-12-2015 to 31-12-2015 ($\approx$1488 data points) and 1-1-2012 to 31-12-2015 ($\approx$1461 data points) time periods, respectively. The MARS and SVR models were built with 1–3 statistically significant lagged input combinations (3 representing the maximum number of lags of significant G data) and denoted as T_1,T_2 and T_3 for 0.5 h, and D_1,D_2 and D_3 for 24 h, respectively. Similarly, the 1.0 h forecasting horizon for the MARS and SVR models were constructed from data over the period 1-11-2015 to 31-12-2015 ($\approx$1464 data points), built with 1–6 statistically significant lagged input

Table 1
Descriptive statistics of the electricity demand (G) (MW) data aggregated for the Queensland (QLD) study region.

Forecast horizon (h)	Data Period (dd-mm-yyyy)	Minimum (MW)	Maximum (MW)	Mean (MW)	Standard deviation (MW)	Skewness	Flatness
0.5	01-12 to 31-12-2015	4660.55	8402.56	6318.42	802.67	0.17	−0.85
1.0	01-11 to 31-12-2015	4668.66	8393.81	6323.48	806.06	0.11	−0.83
24	01-01-2012 to 31-12-2015	4896.05	7165.54	5827.85	414.81	0.54	0.36

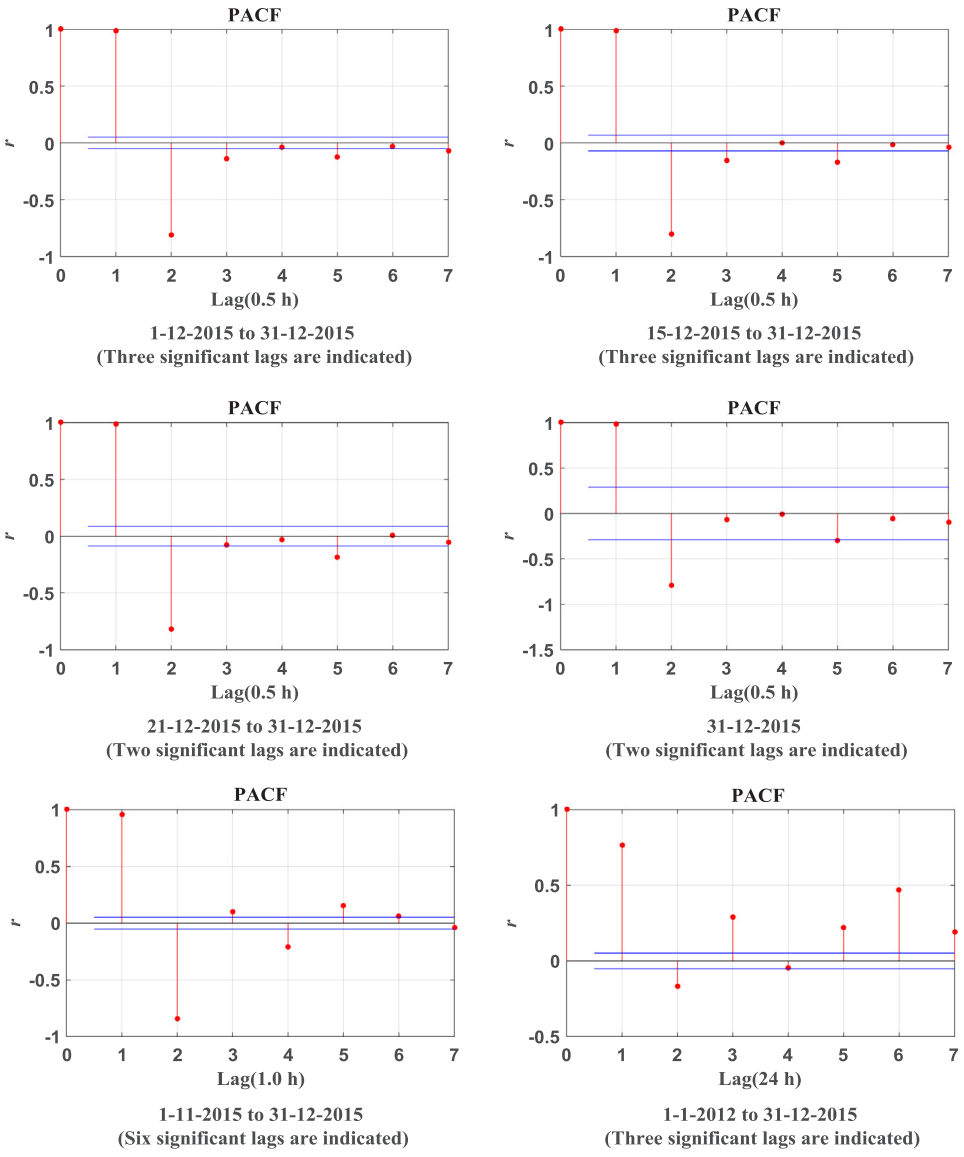

Fig. 2. Correlation coefficient (r) based on the partial autocorrelation function (PACF) of predictors (*i.e.*, electricity demand, G) used for developing the support vector regression (SVR) and multivariate regression splines (MARS) models. Statistically significant lags at the 95% confidence interval are marked (blue). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

combinations (6 representing the maximum number of significant lagged G values), and denoted as H_1,...,H_6, respectively.

To determine the effect of data length, the short-term (0.5 h) forecasting horizon scenario was studied using data from the 15-12-2015 to 31-12-2015 period for the SVR and MARS models. A total of 817 data points with 1–3 statistically significant lags were applied and denoted as the T^a model. Furthermore, the T^b and T^c models used data from period 21-12-2015 to 31-12-2015 and single-day data for 31-12-2015 which consisted of 529 data points and 48 data points with 1 or 2 statistically significant lags, respectively.

On the other hand, the univariate ARIMA model's mechanism differs as it creates its own lagged data through the p and q parameters developed in its identification phase seen in Table 3. Therefore, all historical G data were used as a single input (with no lags) to identify the ARIMA model for all forecasting horizons.

Table 2 and Fig. 2 contain further details regarding the forecast models and their nominal designation. It should be noted that for the baseline models, the input variables had a total of 1461–1488 data points.

There is no single method for dividing training and evaluation data [17]. To deduce optimal models for G forecasting, data were split into subsets as follows: 80% for training and 20% for evaluation (testing). Given the chaotic nature of the input where changes in G seem to occur at a higher frequency, the trained data required appropriate scaling to avoid predictor values (and associated patterns/attributes) with large numeric ranges from dominating attributes with narrower ones [53,54]. Data were therefore normalized and bounded by zero and one through the following expression [17]:

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{12}$$

where x is any given data value (input or target), x_{min} is the minimum value of x, x_{max} is the maximum value of x, and x_{norm} is the normalized value of the data.

The SVR models were developed by the MATLAB-based Libsvm toolbox (version 3.1.2) [55]. The RBF (Eq. (5)) was used to map nonlinear input samples onto a high dimensional feature space because it examines the non-linearities between target and input data [53,54] and outperforms linear-kernel-based models in terms of accuracy [42,56]. The RBF is also faster in the training phase [57,58] as demonstrated in [41]. An alternative linear kernel is a special case of the RBF [56], whereas the sigmoid kernel behaves as the RBF kernel for some model parameters [54].

Furthermore, the selection of C and σ values is crucial to obtain an accurate model [59]. For this reason, a grid search procedure, over a wide range of values seeking the smallest *MSE,* was used to establish the optimal parameters [53]. Fig. 3(a) illustrates a surface plot of the *MSE*

Table 2
Model designation for the MARS, SVR and ARIMA for 0.5 h, 1.0 h and 24 h forecast horizons.

Model		Period of G data studied (dd-mm-yyyy) No. data points						No. significant lags (* = all lags)					
		1-1-2012 to 31-12-2015 1461	1-11-2015 to 31-12-2015 1464	31-12-2015 48	21-12-2015 to 31-12-2015 529	15-12-2015 to 31-12-2015 817	1-12-2015 to 31-12-2015 1488	1	2	3	4	5	6
Half-hourly (0.5 h) forecast horizon													
MARS and SVR	T_1						×	×					
	T_2						×		×				
	T_3						×			×*			
	T^a					×				×*			
	T^b				×				×*				
	T^c			×					×*				
	ARIMA						×						
	ARIMA^a					×							
	ARIMA^b				×								
	ARIMA^c			×									
Hourly (1.0 h) forecast horizon													
MARS and SVR	H_1		×					×					
	H_2		×						×				
	H_3		×							×			
	H_4		×								×		
	H_5		×									×	
	H_6		×										×*
	ARIMA		×										
Daily (24 h) forecast horizon													
MARS and SVR	D_1	×						×					
	D_2	×							×				
	D_3	×								×*			
	ARIMA	×											

with respect to different regularisation constants C and σ (kernel width) values for the SVR model used in 1.0 h forecasting. In this case, the optimal model H_4 attained an MSE of $\cong 0.0001$ MW^2 for $C = 1.00$ and $\sigma = 48.50$. Table 3 lists the optimal values of C and σ that are unique to each SVR model.

Alternatively, the MARS model adopted the MATLAB-based ARESLab toolbox (version 1.13.0) [60]. Two types of MARS models are possible and employ cubic or linear piecewise formula as their basis functions. In this study, a piecewise cubic model was adopted since it provided a smoother response in comparison to a linear function [61]. Moreover, generalized recursive partitioning regression was adopted for function approximation given its capacity to handle multiple predictors [8]. Optimisation operated in two phases: forward selection and backward deletion. In the forward phase, the algorithm ran with an initial 'naïve' model consisting of only the intercept term. It iteratively added the reflected pair(s) of basis functions to yield the largest reduction in training the MSE. The forward phase was executed until one of the following conditions was satisfied [62]:

Table 3
Parameters for the SVR and ARIMA model presented in the training period for 0.5 h, 1.0 h and 24 h forecast horizons.

SVR*	C	σ	MSE (MW^2)	ARIMA**	p	d	q	R^2	σ^2	L	AIC	$RMSE$ (MW)	$MAPE$ (%)
0.5 h Forecast horizon													
T_1	0.19	256.0	0.0012										
T_2	1.74	256.0	0.0004										
T_3	1.00	147.0	0.0004	ARIMA	2	1	6	0.993	4966	−6738.5	13494.9	70.44	0.829
T^a	1.00	84.5	0.0005	ARIMA^a	5	1	6	0.993	4042	−3623.2	7270.4	63.53	0.785
T^b	0.57	147.0	0.0004	ARIMA^b	6	1	3	0.994	3553	−2319.6	4659.2	59.54	0.768
T^c	1.00	9.2	0.0011	ARIMA^c	6	0	1	0.991	2170	−203.7	425.3	46.59	0.660
1.0 h Forecast horizon													
H_1	0.19	147.0	0.0041										
H_2	0.57	256.0	0.0010										
H_3	0.33	147.0	0.0010										
H_4	1.00	48.5	0.0001										
H_5	0.57	48.5	0.0008										
H_6	0.33	27.9	0.0007	ARIMA	5	1	5	0.981	12613	−7159.7	14341.2	112.26	1.366
24 h Forecast horizon													
D_1	0.06	3.0	0.0134										
D_2	0.19	27.9	0.0122										
D_3	0.33	27.9	0.0093	ARIMA	8	1	3	0.805	34015	−7736.7	15497.3	184.35	2.298

* C = cost function, σ = kernel width.
** d = degree of differencing, p = autoregressive term, q = moving average term, R^2 = coefficient of determination, σ^2 = variance, L = log likelihood, AIC = Akaike information criterion, MAPE = mean absolute percentage error, RMSE = root mean square error.

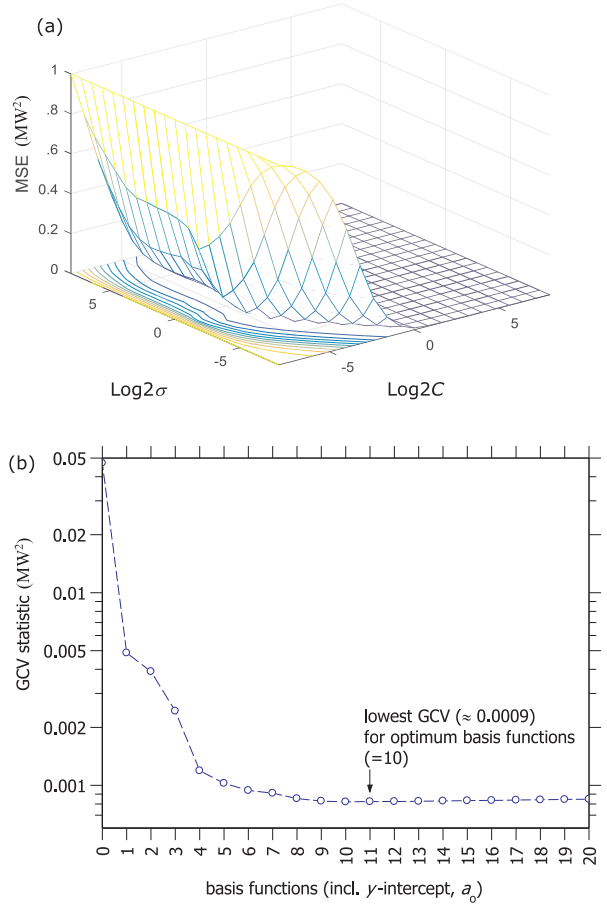

Fig. 3. Illustration of SVR and MARS model parameters for 1.0 h forecast horizon, (H_4) model.

(i) the maximum number of basis functions reached threshold rule min [200, max(20, 2n) + 1], where n = the number of inputs;
(ii) adding a new basis function changed the coefficient of determination (R^2) by less than 1×10^{-4};
(iii) R^2 reached ≈1;
(iv) the number of basis functions including the intercept term reached the number of data observations; or
(v) the effective number of parameters reached the number of observed data points.

In the deletion phase, the large model, which typically over-fits the data, was pruned back one-at-a-time to reduce *RMSE* until only the model's intercept term remained. Subsequently, the model with lowest Generalized Cross-Validation (*GCV*) was selected. The MARS model (H_4) used for the 1.0 h forecasting horizon had 20 basis functions and the lowest *GCV* at the pruning stage was indicated with 10 functions (Fig. 3(b)). Table 4 shows the forecasting equations (in training periods) with optimum basis functions (BF_m) and the *GCV* for all forecast horizons. A MARS model's *GCV* statistic after the pruning stage should be relatively small.

To offer a comparative framework for the SVR and MARS models, the ARIMA model was developed using the *R* package [46,63]. Table 3 displays the ARIMA model's architecture. Since many model identification methods exist, a selection technique was implemented that considered the coefficient of determination(R^2), Akaike information criterion (*AIC*) [64], log likelihood (*L*) [64] and the lowest variance (σ^2).

Since *G* data was non-stationary as observed in Fig. 2, a differencing process was applied to convert the *G* data to stationarity and satisfy the ARIMA model's input requirements as previously mentioned [46,63]. The requirement was confirmed by ensuring the results of *autoarima* (AR) function [65] obtained the lowest standard deviation and *AIC* with the highest *L*.

Additionally, the autoregressive (p), differencing (d), and moving average terms (q) were determined iteratively [46]. The estimates of p and q were obtained by testing reasonable values and evaluating how the criteria, *L* AIC, σ, and R^2, were satisfied. The fitted ARIMA model was then optimised with 'trial' values of p,d, and q. The training

Table 4
The MARS model forecast equation, $y = a_0 + \sum_{m=1}^{M} a_m BF_m(X)$ with optimum basis functions (BF_m), and generalized cross validation statistic (*GCV*) in MW² for all horizons, in the training period.

MARS model	Model Equation	Opt. Basis Functions	***GCV*** **(MW²)**
0.5 h Forecast horizon			
T_1	$y = 0.461 + 0.992BF_1 - 0.984BF_2$	2	0.00109
T_2	$y = 0.456 + 1.67BF_1 - 1.911BF_2 - 0.681BF_3 + 0.944BF_4$	4	0.00037
T_3	$y = 0.480 + 1.587BF_1 - 1.834BF_2 - 0.484BF_3 + 0.790BF_4 - 0.104BF_5$	5	0.00036
T^a	$y = 1.251 + 1.475BF_1 - 1.641BF_2 - 0.481BF_3 + 0.525BF_4 - 0.176BF_5 + 0.110BF_6$	6	0.00043
T^b	$y = 1.035 + 1.711BF_1 - 1.806BF_2 - 0.791BF_3 + 0.834BF_4$	4	0.00038
T^c	$y = 0.656 - 1.689BF_1 + 0.857BF_2 + 0.785BF_3$	3	0.00168
1.0 h Forecast horizon			
H_1	$y = 0.236 + 0.47BF_1 + 1.784BF_2 - 1.453BF_3 - 0.849BF_4$	4	0.0039
H_2	$y = 0.139 + 0.277BF_1 - 0.837BF_2 + 1.319BF_3 + 1.538BF_4 - 2.171BF_5 - 0.324BF_6$	6	0.0010
H_3	$y = 0.926 + 3.131BF_1 - 3.417BF_2 + 1.011BF_3 - 0.339BF_4 - 1.616BF_5 + 2.092BF_6 + 0.201BF_7 - 0.465BF_8 + 2.773BF_9 - 1.417BF_{10} - 1.829BF_{11}$	11	0.0009
H_4	$y = -0.144 + 0.537BF_1 + 1.702BF_2 - 2.298BF_3 - 0.194BF_4 - 0.434BF_5 + 0.437BF_6 + 0.063BF_7 - 0.749BF_8 + 0.896BF_9 - 0.261BF_{10}$	10	0.0009
H_5	$y = -0.021 + 0.010BF_1 - 0.932BF_2 + 1.371BF_3 + 0.618BF_4 - 0.771BF_5 + 1.707BF_6 - 2.332BF_7 + 0.522BF_8 - 0.213BF_9 - 0.544BF_{10} + 0.314BF_{11} + 0.016BF_{12} + 0.113BF_{13} - 0.555BF_{14}$	14	0.0009
H_6	$y = 0.686 + 2.418BF_1 - 2.417BF_2 - 0.792BF_3 + 1.655BF_4 - 0.288BF_5 - 0.721BF_6 + 0.432BF_7 - 0.826BF_8 - 0.391BF_9 + 0.581BF_{10} - 0.076BF_{11} - 0.058BF_{12}$	12	0.0008
24 h Forecast horizon			
D_1	$y = 0.176 + 0.617BF_1 + 0.749BF_2 - 0.448BF_3$	3	0.01339
D_2	$y = 0.214 + 0.98BF_1 + 0.486BF_2 - 1.183BF_3 - 0.538BF_4 - 0.764BF_5 - 0.158BF_6 + 1.820BF_7$	7	0.01269
D_3	$y = 0.092 + 1.106BF_1 - 0.487BF_2 - 0.387BF_3 + 0.592BF_4 + 1.872BF_5 - 0.864BF_6 + 0.400BF_7 + 0.750BF_8 - 0.819BF_9 - 1.197BF_{10} - 1.528BF_{11}$	11	0.01187

performance was unique for each forecasting horizon and in accordance with the goodness-of-fit parameters shown in Table 3.

3.3. Model performance evaluation

Error criteria were adopted to establish the accuracy of the data-driven models. [66–71]. These include the Mean Absolute Error (*MAE*), *RMSE*, relative error (%) based on *MAE* and *RMSE* values ($MAE_{\overline{G}}$ and $RMSE_{\overline{G}}$), correlation coefficient (r), Willmott's Index (*WI*), the Nash–-Sutcliffe coefficient (E_{NS}), and Legates and McCabe Index (E_{LM}) [41,67–69,72–74] represented below:

$$r = \frac{\sum_{i=1}^{i=n}[(G_i^{obs} - \overline{G^{obs}})(G_i^{for} - \overline{G^{for}})]}{\sqrt{\sum_{i=1}^{i=n}(G_i^{obs} - \overline{G^{obs}})^2} \cdot \sqrt{\sum_{i=1}^{i=n}(G_i^{for} - \overline{G^{for}})^2}} \quad (13)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{i=n}(G_i^{for} - G_i^{obs})^2} \quad (14)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{i=n}|G_i^{for} - G_i^{obs}| \quad (15)$$

$$RMSE_{\overline{G}} = 100 \times \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{i=n}(G_i^{for} - G_i^{obs})^2}}{\overline{G^{obs}}} \quad (16)$$

$$MAE_{\overline{G}} = 100 \times \frac{1}{n}\sum_{i=1}^{i=n}\left|\frac{G_i^{for} - G_i^{obs}}{G_i^{obs}}\right| \quad (17)$$

$$WI = 1 - \left[\frac{\sum_{i=1}^{i=n}(G_i^{for} - G_i^{obs})^2}{\sum_{i=n}^{i=n}(|G_i^{for} - \overline{G^{obs}}| + |G_i^{obs} - \overline{G^{obs}}|)^2}\right], \text{ and } 0 \leqslant WI \leqslant 1 \quad (18)$$

$$E_{NS} = 1 - \left[\frac{\sum_{i=1}^{i=n}(G_i^{for} - G_i^{obs})^2}{\sum_{i=1}^{i=n}(G_i^{obs} - \overline{G^{obs}})^2}\right], \text{ and } \infty \leqslant E_{NS} \leqslant 1 \quad (19)$$

$$E_{LM} = 1 - \left[\frac{\sum_{i=1}^{i=n}|G_i^{obs} - G_i^{for}|}{\sum_{i=1}^{i=n}|G_i^{obs} - \overline{G^{obs}}|}\right], \text{ and } (\infty \leqslant E_{LM} \leqslant 1) \quad (20)$$

where n is the total number of observed (and forecasted) values of G, G_i^{for} is the i^{th} forecasted value of G, $\overline{G^{for}}$ is the mean of forecasted values, G_i^{obs} is the i^{th} observed value of G, $\overline{G^{obs}}$ is the mean of observed values.

The model statistics, obtained through equations (13)–(20), aimed to assess the accuracy of the G forecasts with respect to observed G values. For instance, the covariance-based metric r served to analyse the statistical association between G_i^{for} and G_i^{obs} where $r = 1$ represents an absolute positive (ideal) correlation; $r = -1$, an absolute negative correlation; and $r = 0$, a lack of any linear relationship between G_i^{for} and G_i^{obs} data. According to the work of Chai and Draxler [70], the *RMSE* is more representative than the *MAE* when the error distribution is Gaussian. However, when it is not the case, the use of *MAE*, *RMSE,* and their relative expressions, $MAE_{\overline{G}}$ and $RMSE_{\overline{G}}$, can yield complementary evaluations. Since other metrics can also assess model performance [70], the E_{NS} and *WI* were also calculated. A value of E_{NS} and *WI* near 1.0 represents a perfect match between G_i^{for} and G_i^{obs}, while a complete mismatch between the G_i^{for} and G_i^{obs} results in values of ∞ and 0, respectively. For example, when E_{NS}, which is the ratio of the mean square error to the variance in the observed data, equals 0.0, it indicates that $\overline{G^{obs}}$ is as good a predictor as G_i^{for}, however, if E_{NS} is less than 0.0, the square of the differences between G_i^{for} and G_i^{obs} is as large as the variability in G_i^{obs} and indicates that $\overline{G^{obs}}$ is a better predictor than G_i^{obs} [74,75]. As a result, using a modified version of *WI,* which is the Legates and McCabe Index ($\infty \leqslant E_{LM} \leqslant 1$) [74], can be more advantageous than the traditional *WI*, when relatively high values are expected as a result of squaring of differences [68,73]. On the other hand, the $MAE_{\overline{G}}$ and $RMSE_{\overline{G}}$ were applied to compare forecasts at different timescales that yield errors of different magnitudes (*e.g.*, Fig. 2). According to [41,42,76,77], a model can be considered excellent when $RMSE_{\overline{G}} < 10\%$, good if the model satisfies $10\% < RMSE_{\overline{G}} < 20\%$, fair if it satisfies $20\% < RMSE_{\overline{G}} < 30\%$, and poor if $RMSE_{\overline{G}} > 30\%$.

4. Results and discussion

Evaluation of the data-driven models' ability to forecast the electricity demand (G) data for the 0.5 h, 1.0 h, and 24 h horizons is presented in this section using the statistical metrics from Eqs. (13)–(20). Only optimum models with lowest *MAE* and largest r and *WI* are shown in Table 5. Between the SVR and ARIMA models, the MARS model yielded better G forecasting results for the 0.5 h and 1.0 h horizons. This was evident when comparing the MARS (T^b) model's accuracy statistics ($r = 0.993$, $WI = 0.997$, and $MAE = 45.363$ MW) with the equivalent SVR (T^b) and ARIMA^b models' results ($r = 0.990$, $WI = 0.995$ and $MAE = 55.915$ MW) and ($r = 0.423$, $WI = 0.498$ and $MAE = 362.860$ MW), respectively.

While both the MARS and SVR models yielded accurate G forecasts when predictor variables were trained for the data period from 21-12-2015 to 31-12-2015, the ARIMA model attained the highest accuracy for data trained in period 31-12-2015 (*i.e.*, model ARIMA^c; $r = 0.976$, $WI = 0.702$ and $MAE = 237.746$ MW). Despite being significantly inferior to the MARS and SVR models for longer periods, the ARIMA models' performance improved when a shorter data set (*i.e.*, 31-12-2015) was utilised. When the four ARIMA models for 0.5 h forecasting horizons (developed in Table 3) were evaluated, an increase in the correlation coefficient (0.128–0.976) was identified. In addition, a respective decrease was observed in *MAE and RMSE* values (475.087–237.746 MW) and (569.282–256.565 MW) respectively, with parallel changes in *WI* and E_{NS} values.

The analysis based on Fig. 1(a) confirmed that the ARIMA model was most responsive in forecasting G data when input conditions had lower variance, as detected in single day's data (31-12-2015) in comparison to longer periods (1–12 to 31-12-2015). Therefore, the SVR and MARS models had a distinct advantage over the ARIMA model when a lengthy database was used for G forecasting. Furthermore, when models were evaluated for the 1.0 h forecasting horizon (Table 5), the MARS and SVR models (H_4), with four sets of lagged input combinations, were the most accurate and outperformed the best ARIMA model.

The MARS model was significantly superior to the SVR and ARIMA

Table 5
Evaluation of the optimal models attained for 0.5 h, 1.0 h and 24 h forecast horizons in the test period.

Model	Model Accuracy Statistics*				
	r	*WI*	E_{NS}	*RMSE* (MW)	*MAE* (MW)
0.5 h Forecast horizon					
MARS(T^b)	0.993	0.997	0.986	57.969	45.363
SVR(T^b)	0.990	0.995	0.980	70.909	55.915
ARIMA^b	0.423	0.498	0.080	476.835	362.860
ARIMA^c	0.976	0.702	−0.233	256.565	237.746
1.0 h Forecast horizon					
MARS(H_4)	0.990	0.994	0.978	106.503	86.502
SVR(H_4)	0.972	0.981	0.930	189.703	124.453
ARIMA	0.401	0.381	0.144	665.757	555.637
24 h Forecast horizon					
MARS(D_3)	0.753	0.859	0.543	256.000	200.426
SVR(D_3)	0.806	0.890	0.647	225.125	162.363
ARIMA	0.289	0.459	−1.018	538.124	474.390

* r = correlation coefficient, E_{NS} = Nash–Sutcliffe coefficient, MAE = mean absolute error, RMSE = root mean square error, WI = Willmott's index.

models for the 1.0 h forecasting horizon. Based on the *r*, *WI*, and *MAE* metrics, the MARS model ($r = 0.990$, $WI = 0.994$ and $MAE = 86.502$ MW) outperformed the SVR model ($r = 0.972$, $WI = 0.981$ and $MAE = 124.453$ MW). The MARS model's *WI*, a more robust statistic than the linear dependence measured by *r* [66], was 1.33% greater than the SVR model's. This was supported by the MARS model's lower *RMSE* and *MAE* values, 78.12% and 43.87%, respectively. In contrast, the ARIMA model displayed an inferior performance ($r = 0.401$, $WI = 0.381$ and $MAE = 555.637$ MW) as seen in Table 5.

For a 24 h forecasting horizon, the SVR ($r = 0.806$, $WI = 0.890$ and $MAE = 162.363$ MW) outperformed the MARS model (D_3) by a small margin ($r = 0.753$, $WI = 0.859$ and $MAE = 200.426$ MW) (Table 5). Similarly to the hourly scenario, the ARIMA model performed poorly ($r = 0.289$, $WI = 0.459$ and $MAE = 474.390$ MW). It is important to consider that the ARIMA models for hourly and daily forecasting horizons were developed using the long time-series: 1-11-2015 to 31-12-2015 and 1-1-2012 to 31-12-2015, respectively. The predictor (historical *G*) data exhibited significant fluctuations over these long-term periods compared to the single day *G* data of 31-12-2015 (ARIMA^c).

In conjunction with statistical metrics and visual plots of forecasted *vs.* observed *G* data, the $MAE_{\overline{G}}$,$RMSE_{\overline{G}}$, and E_{LM} (*e.g.,* [17,41,42,78]) are used to show the alternative 'goodness-of-fit' of the model-generated *G* in relation to observed *G* data. The MARS model yielded relatively high precision (lowest $MAE_{\overline{G}}$ and $RMSE_{\overline{G}}$ and the highest E_{LM}) followed by the SVR and ARIMA models (Table 6). For the MARS model, $MAE_{\overline{G}}/RMSE_{\overline{G}}$ for the 0.5 h and 1.0 h forecasting horizons were 0.77/0.99% (T^b) and 1.45/1.76% (H_4), respectively. On the other hand, the SVR model resulted in 0.95/1.21% (T^b) and 2.19/3.13% (H_4). Likewise, $E_{LM}$was utilised in combination with other performance metrics for a robust assessment of models [74]. The respective value for both 0.5 h and 1.0 h forecasting horizons was determined to be greater for the MARS model (0.887/0.857) than for the SVR model (0.861/0.794). Although the MARS models outperformed the SVR models for the 0.5 h and 1.0 h horizons, the SVR model surpassed the MARS model for the 24 h horizon (13.73%/23.63% lower $RMSE_{\overline{G}}/MAE_{\overline{G}}$ and 45.42% higher E_{LM}). It is evident that both the MARS and SVR models, adapted for *G* forecasting in the state of Queensland, exceeded the performance of the ARIMA model and thus, should be further explored for use in electricity demand estimation.

Nevertheless, despite the ARIMA model faring slightly worse for most of the *G* forecasting scenarios in this paper, specifically for the case of 1.0 h and 24 h horizons ($RMSE_{\overline{G}} = 11.0\%$ and 9.04%, respectively), its performance for the 0.5 h horizon using a single day's data (ARIMA^c) exhibited good results. This is supported by an $RMSE_{\overline{G}}$ value of approximately 4.18% (Table 6). Therefore, it is possible that a large degree of fluctuation in the longer training dataset could have led the ARIMA model's autoregressive mechanism to be more prone to cumulative errors than to a situation with a shorter data span.

Table 6
The relative root mean square error $RMSE_{\overline{G}}$(%), mean absolute percentage error $MAE_{\overline{G}}$(%) and Legates & McCabes Index (E_{LM}) for the optimal models in the test datasets.

Model	E_{LM}	$MAE_{\overline{G}}$ (%)	$RMSE_{\overline{G}}$ (%)
0.5 h Forecast horizon			
MARS(T^b)	0.887	0.765	0.990
SVR(T^b)	0.861	0.945	1.211
ARIMA^b	0.098	6.487	8.140
ARIMA^c	−0.238	3.939	4.184
1.0 h Forecast horizon			
MARS(H_4)	0.857	1.446	1.760
SVR(H_4)	0.794	2.192	3.134
ARIMA	0.080	9.350	11.000
24 h Forecast horizon			
MARS(D_3)	0.295	3.359	4.300
SVR(D_3)	0.429	2.717	3.781
ARIMA	−0.668	8.193	9.039

In contrast to previous studies on the MARS, SVR, or ARIMA models, the forecasting models developed in this study achieved a relatively high precision for short-term *G* forecasting. For example, a study that forecasted daily *G* data for South Africa using the MARS model attained an *RMSE* of 446.01 MW [19], whereas the present study's MARS model resulted in an *RMSE of* 256.00 MW (see MARS(D_3) in Table 5). Likewise, 24 h lead time forecasts of *G* in Istanbul (Turkey) using an RBF-based SVR model [23] yielded an $MAE_{\overline{G}}$ of 3.67%, whereas the $MAE_{\overline{G}}$ value obtained in the present study was 2.72% (see SVR(D_3) in Table 6). For the same forecast horizon, the ARIMA model

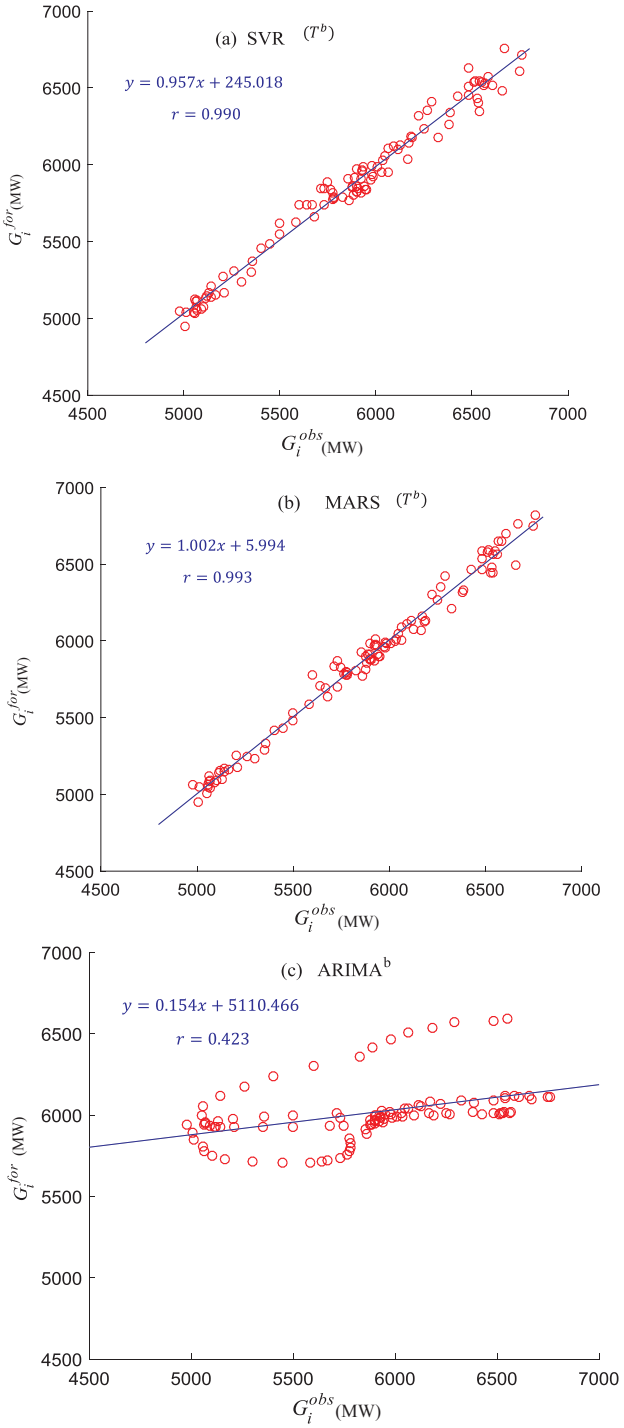

Fig. 4. Scatterplot of the forecasted, G_i^{for}*vs.* theobserved,G_i^{obs} electricity demand data in the testing period for the 0.5 h forecast horizon, (a) SVR(T^b) (b) MARS(T^b) and (c) ARIMA^b. A linear regression line, $y = G_i^{for} = a'G_i^{obs} + b'$ with the correlation coefficient, *r* is included.

reported in [38], denoted as $(p,d,q) = (4,1,4)$,yielded an *RMSE* value of 584.72 MW compared to a lower *RMSE* of 538.12 MW achieved with the present ARIMA model denoted as $(p,d,q) = (8,1,3)$. Furthermore, a study that forecasted *G* data in New South Wales, Queensland and Singapore [79], used singular spectrum analysis, gravitational search, and adaptive particle swarm optimization following a gravitational search algorithm (APSOGSA) to forecast *G*. The APSOGSA model yielded an *MAE/RMSE* of 115.59/133.99 MW and an $MAE_{\overline{G}}$ of 2.32%. Equivalent models in this study seem to exceed the others' performance as evidenced in Tables 5 and 6. The analysis for MARS(T^b) and SVR(T^b)resulted in an *MAE/RMSE* of 45.36/57.97 and 55.92/70.91 MW and $MAE_{\overline{G}}$ of 0.77 and 0.95%, respectively.

Separately, Figs. 4–6 depict scatterplots of G_i^{for} vs. G_i^{obs} for the 0.5 h, 1.0 h and 24 h forecasting horizons using optimal MARS, SVR, and ARIMA models (see Table 5). A least square regression line, $y = G_i^{for} = a'^{G_i^{obs}} + b'$, and *r* value are used to illustrate the relationship between G_i^{for} and G_i^{obs} data, where a' is the slope and b' is the *y*–intercept. Both are used to describe the model's accuracy [17].

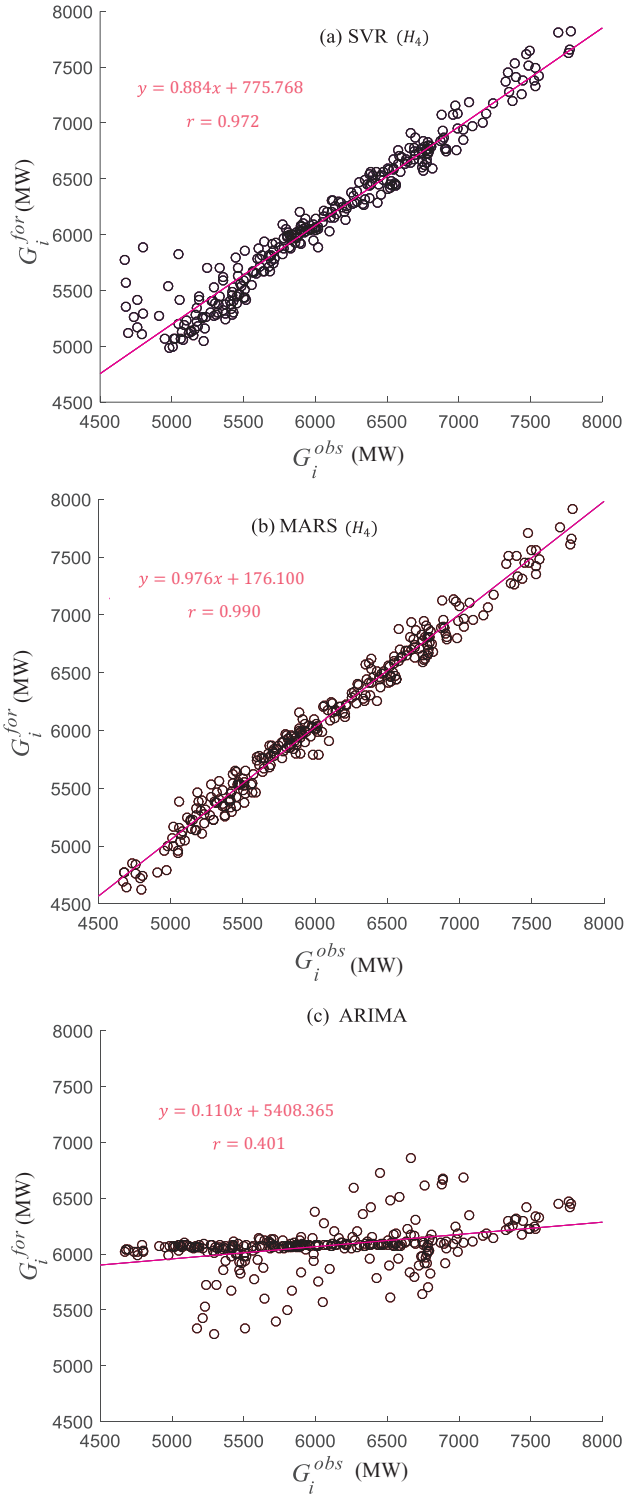

Fig. 5. The caption description is the same as that in Fig. 4 except for the 1.0 h forecast horizon, (a) SVR(H_4), (b) MARS(H_4) and (c) ARIMA.

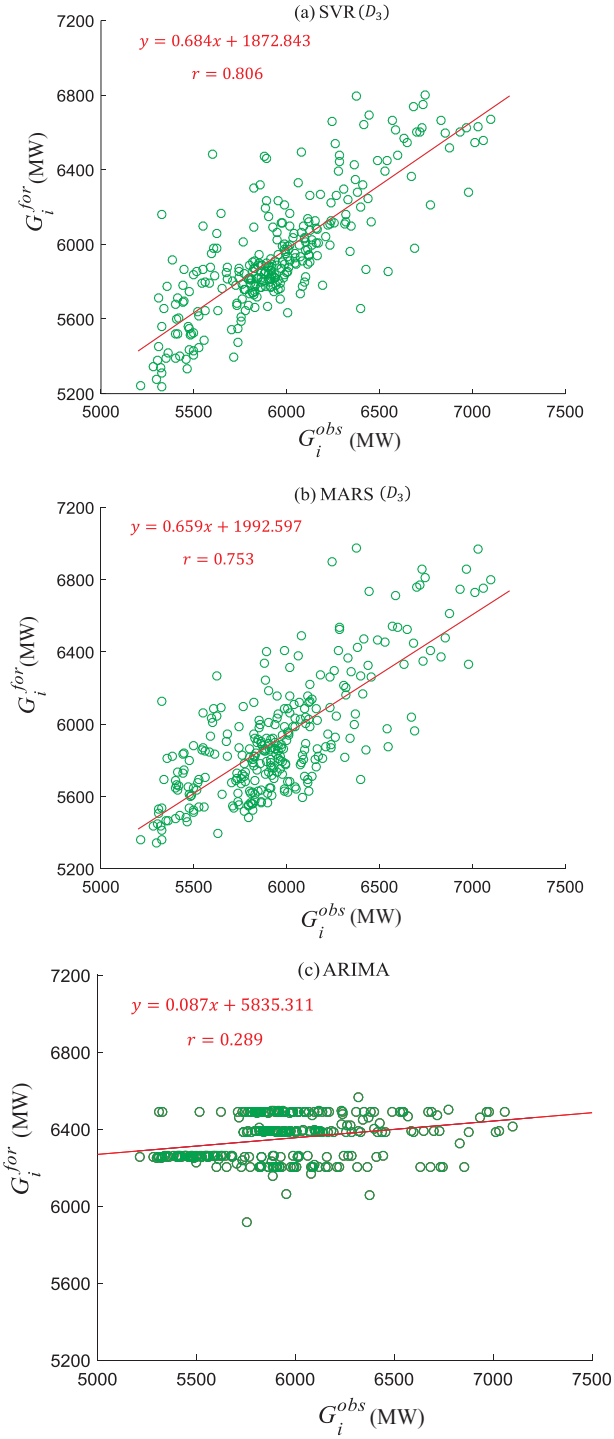

Fig. 6. The caption description is the same as that in Fig. 4 except for the 24 h forecast horizon, (a) SVR(D_3) (b) MARS(D_3) and (c) ARIMA.

For the 0.5 h horizon, the optimal SVR and MARS models yielded near unity a' values of 0.957 and 1.002, respectively. On the contrary, the a' for the ARIMA model (ARIMA^b) was 0.154 deviating significantly from an ideal value of 1 (Fig. 4a–c). The deviation of forecasted G data from observations (*i.e.*, 1:1 line or reference a'-value of 1) was largest in the case of the ARIMA model, approximately 0.846. In the case of the SVR and MARS models, these deviations were 0.043 and 0.002, respectively.

Consistent with the level of scattering, the r value for the MARS model exceeded the SVR and ARIMA models' values. In concordance with the r value trends, similar results for a' values were attained for 1.0 h forecasting where the optimum MARS, SVR, and ARIMA models (Fig. 5a, b and c) yielded 0.976, 0.884 and 0.110, respectively. Additionally, for the 24 h forecasting horizon (Fig. 6a, b and c), the SVR model (r = 0.806, a' = 0.684, b' = 1872.843) outperformed the MARS model (r = 0.753, a' = 0.659, b' = 1992.597). Both models provided

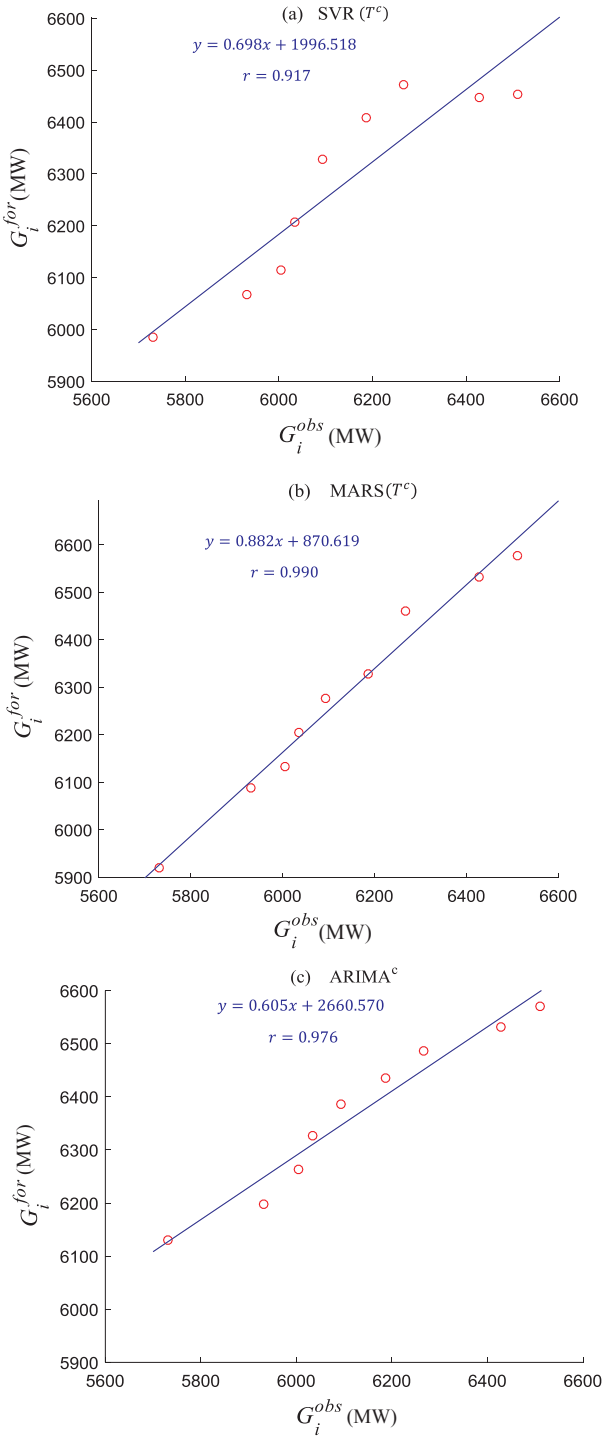

Fig. 7. The caption description is the same as that in Fig. 4 except for the 0.5 h forecast horizon, (a) SVR(T^c), (b) MARS(T^c) and (c) ARIMA^c.

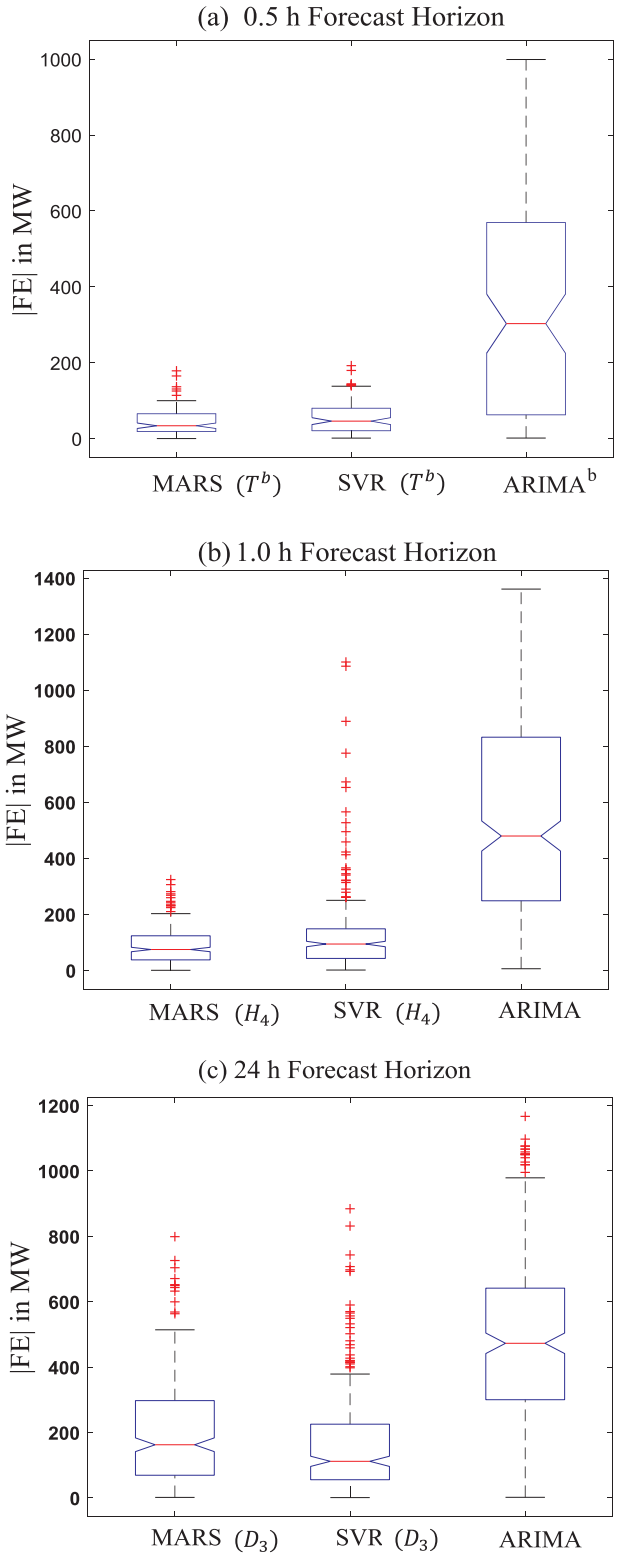

Fig. 8. Boxplots of the absolute forecasted error, $|FE| = |G_{FOR,i} - G_{OBS,i}|$ for: (a) 0.5 h, (b) 1.0 h and (c) 24 h forecast horizons using the MARS, SVR and ARIMA models.

better results than the ARIMA model (r = 0.289, a' = 0.087 and b' = 5835.311).

On the other hand, Fig. 7 compares the performance for the shortest horizon (0.5 h) using G data gathered over a single day (*i.e.*, 31-12-2015) partitioned into training and testing phases. The MARS model (r = 0.99) outperformed the SVR (r = 0.917) and ARIMA (r = 0.976) models. However, it is important to note that the performance of the ARIMA, for the shorter dataset, was better than its performance for longer datasets (Table 5). This suggests that the ARIMA model's performance deteriorated as the forecasting period increased. This concurs with its auto-regressive and integrated averaging nature since the sum of preceding errors is used for forecasting the next G value [46]. Although the cause is not yet clear, the ARIMA[c] model's better performance could be attributed to greater fluctuations in longer-term predictor data drawn upon in the hourly and daily models (Table 2 and Fig. 2).

Boxplots showing the error distribution for absolute values of forecasted error statistics, $|FE| = |G_i^{for} - G_i^{obs}|$, reveal a greater amount of detail about the models' precision, where the whiskers (Fig. 8) represent the extremes of the forecasted and the observed G values. The lower end of each boxplot represents the lower quartile, Q_{25} (25th percentile); the upper end shows the upper quartile, Q_{75} (75th percentile); and the central line shows the second quartile, Q_{50} (*i.e.*, 50th percentile) or the median value. Two horizontal whiskers are also extended from Q_{25} to the smallest non-outlier and from Q_{75} to the largest non-outlier, respectively. Based on the box plots, Table 7 summarizes statistical properties of the forecasted and observed G data.

For all forecasting horizons considered, the MARS and SVR models performed better than the ARIMA model and therefore, demonstrated significant differences. In terms of the maximum absolute error, the MARS model was most precise for the 0.5 h horizon. For example, theMARS(T^b) resulted in a maximum $|FE|$ of 178.54 MW (Fig. 8a and Table 7) and the smallest median value ($Q_{50} \approx$ 33.77 MW)) relative to any other model. Similarly, for the 1.0 forecasting scenario, statistics indicated the superiority of the MARS model over the SVR and ARIMA models (Table 7; Fig. 8b).

When the errors for the 24 h forecasting horizon were analysed, the MARS and SVR resulted in similar maximum values but distinctly lower than for the ARIMA model. When the median errors were compared, the SVR model (111.76 MW) generated more accurate forecasts than the MARS model (162.41 MW). These median errors differed significantly from those of the ARIMA model (479.66 MW; Table 7; Fig. 8c).

Fig. 9(a–c) illustrates the percentage of the absolute value of forecasted error statistics ($|FE|$) encountered through the empirical cumulative distribution function (ECDF) for optimal models at different forecasting horizons. With respect to the percentage of errors located in the smallest error bracket (*i.e.*, 0 to ± 50 MW), the ECDF demonstrated that the MARS and SVR models outperformed the ARIMA model for all forecasting horizons. Based on this error bracket, the MARS performed slightly better than the SVR model (*i.e.*, about 60% *vs.* 57% and 34% *vs.* 28% for 0.5 h and 1.0 h forecasts, respectively). Within the error bracket of 0 to ± 100 MW for the 0.5 h horizon, the MARS model recorded about 94% of all forecasted errors, whereas the SVR model only 83%. Additionally, for the 1.0 h horizon, the MARS model performed better than the SVR model (*i.e.*, about 63% *vs.* 53% of errors within the 0 to ± 100 MW bracket). However, data for the 24 h horizon recorded comparable values between the two in the smaller error bracket. Nonetheless, better percentage was yielded for the SVR (about 49%) against MARS (about 38%) in the 0 to ± 100 MW bracket.

Since the MARS and SVR models illustrated similar performance in several cases, a statistical t-test was utilised to demonstrate whether the differences in the mean of $|FE|$ were significant. For the 0.5 h, 1.0 h, and 24 h forecasting horizons, we could reject the null-hypothesis that the means are the same (p-value $<$ 0.05). Consequently, the differences in the means are statistically significant for the absolute values of the forecasted errors generated by the MARS and the SVR model.

Based on Table 5, the ARIMA model proved highly inaccurate for the short-term 0.5 h G forecasting horizon as nearly 60% of the errors in the testing period fell in the error range magnitude of greater than 100 MW (Fig. 9a). Similar observations were evident for about 90% of the hourly and daily ARIMA forecasts (Fig. 9b and c). The ARIMA models' forecasting accuracy for the 0.5 h horizon exceeded those for 1.0 h or 24 h horizons as the percentage of errors received from ECDF in the smallest category (0.5 h) was nearly double. This concurred with earlier results (Table 5) where overall evaluation metrics demonstrated the greatest correlation between the observed and ARIMA-forecasted G, including higher WI and E_{NS} and lower $RMSE/MAE$ values.

Ultimately, the versatility of data-driven models was also examined with respect to the forecasting errors for peaks in G by plotting the ten greatest relative errors (Fig. 10). Except for one data point, it was apparent that the MARS models consistently yielded the lowest percent errors for the 0.5 h and 1.0 h forecasting horizons compared to the SVR or ARIMA models (Fig. 10a and b). In contrast, for the 24 h forecasting horizon, the ten highest relative error values were very similar between the MARS and SVR models, but dramatically lower for the ARIMA model (Fig. 10c). The accuracy of the present data-driven models appeared to deteriorate as the forecasting period was extended. This was demonstrated by the relative performance errors (Table 6), the top error values (Fig. 10), and the statistical distribution of the errors (Fig. 8 and Table 7).

Table 7
Evaluation of the differences in the absolute value of forecast error statistics based on observed and forecasted G in the test period for the optimal models.

Error Statistic[a] (MW)	Forecast horizon (h)								
	0.5 h			1.0 h			24 h		
	MARS(T^b)	SVR(T^b)	ARIMA[b]	MARS(H_4)	SVR(H_4)	ARIMA	MARS(D_3)	SVR(D_3)	ARIMA
Maximum	178.54	192.20	999.10	324.09	1100.50	1360.80	798.97	884.25	1177.40
Minimum	0.02	1.21	1.34	0.27	1.20	5.85	1.41	0.78	2.72
Q_{25}	18.54	20.61	62.54	37.47	42.65	248.39	69.15	55.47	270.17
Q_{50}	33.77	45.97	302.97	74.47	94.03	479.54	162.41	111.76	479.66
Q_{75}	65.37	79.99	569.29	123.46	148.31	832.24	297.44	225.36	667.07
Range	178.52	190.99	997.76	323.82	1099.30	1355.00	797.56	883.47	1174.70
Skewness	1.23	0.91	0.46	0.99	3.68	0.45	1.00	1.77	0.09
Flatness	4.59	3.14	1.87	3.98	21.13	2.15	3.82	6.61	2.34

[a] Lower quartile (Q_{25}), median (Q_{50}), upper quartile (Q_{75}).

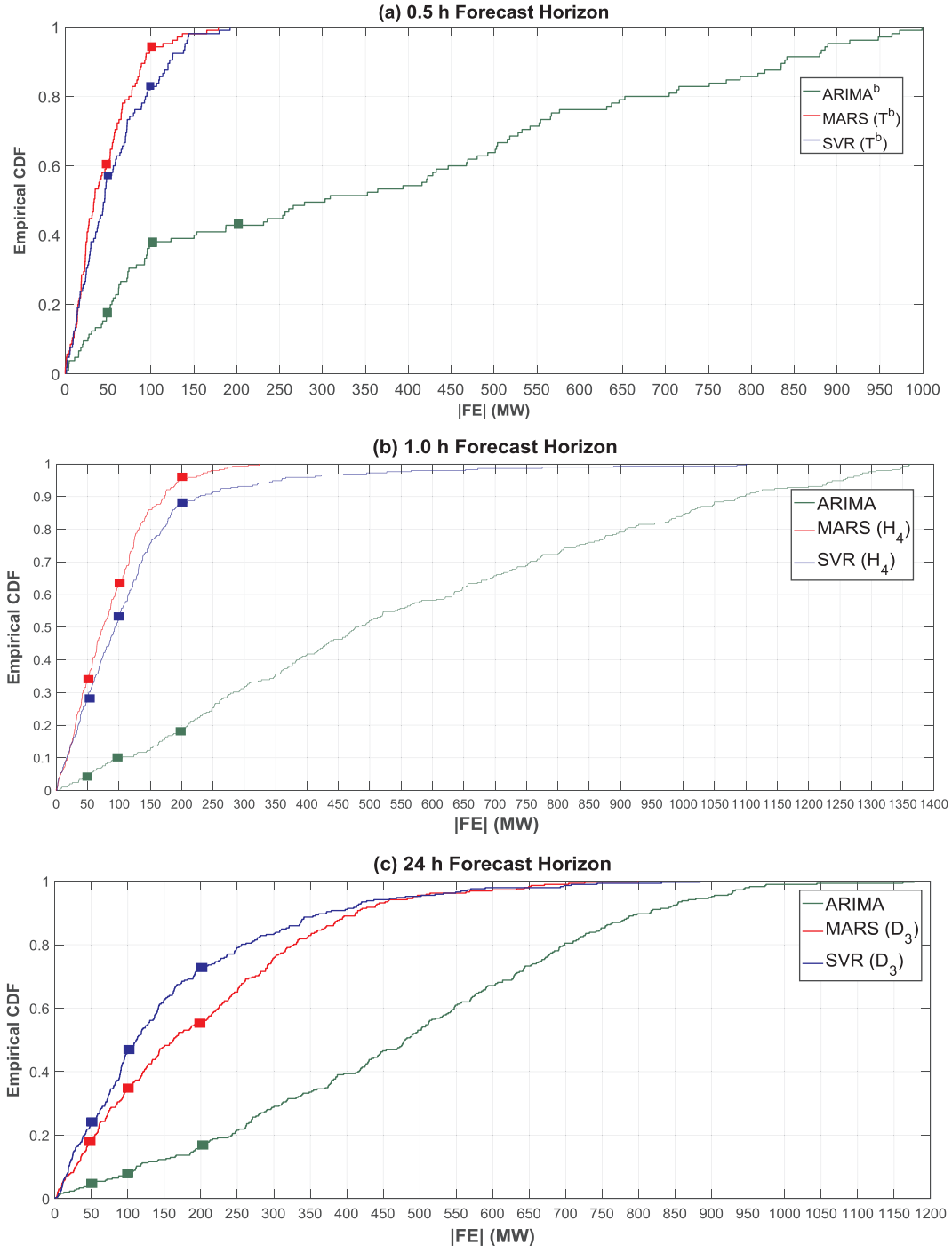

Fig. 9. Empirical cumulative distribution function (ECDF) of the forecast error, $|FE|$ for: (a) 0.5 h, (b) 1.0 h and (c) 24 h forecast horizons using the MARS, SVR and ARIMA models.

5. Further discussion, limitations and opportunities for future research

Data-driven models applied for *G* forecasting over multiple forecast horizons were evaluated. The SVR models were constructed by optimizing regulation constants (minimizing the training error) and radial basis function width (Table 3). The MARS models were tuned with a piecewise multivariate regression function based on the lowest GCV statistic, while the ARIMA models were optimised by a trial and error process (Tables 3 and 4). A comprehensive evaluation showed a greater accuracy of the MARS models when compared to the SVR and ARIMA models for 0.5 h and 1.0 h forecasting horizons. However, for 24 h forecasting horizon, the SVR performed considerably better (Tables 5 and 6).

Given the importance of accurately forecasting *G* data to meet engineering and energy demand challenges, including the sustainable operation of the NEM, this research paper has highlighted the potential utility of further exploring the MARS and SVR models to improve *G* forecasting accuracy. Particularly, this research study established the distinct advantage of the MARS model if employed in real-time *G* forecasting. In terms of greater speed, simplicity of development, and efficiency in performance, the MARS model was best adapted to such forecasts given the SVR models' requirements for tedious modelling phases (*i.e.*, identifying the regulation and kernel width parameters via

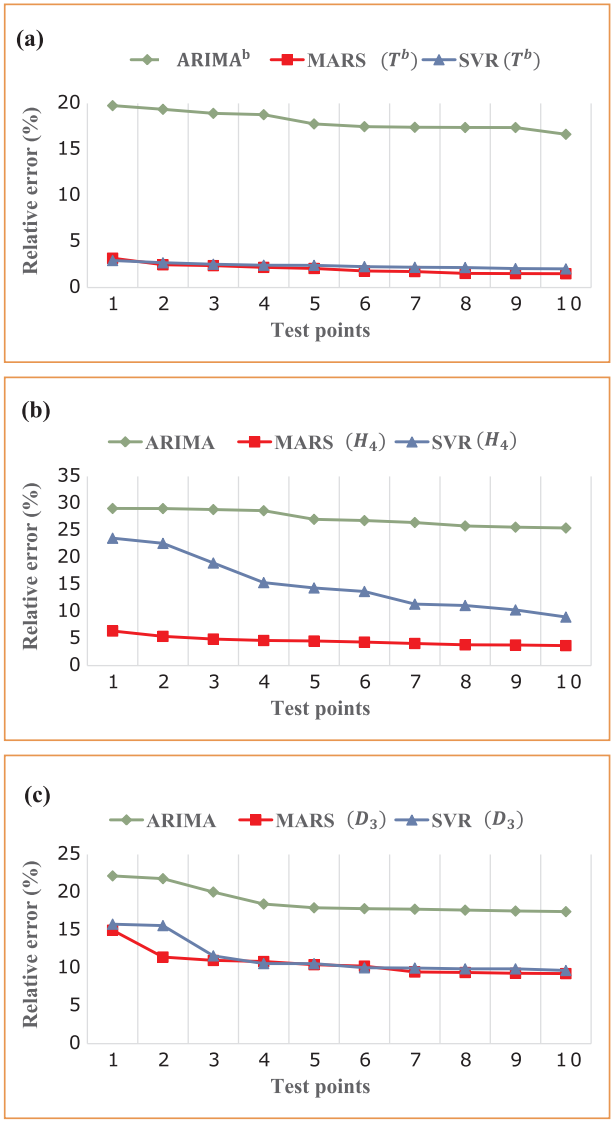

Fig. 10. The top ten peak relative forecast errors (%) generated by the MARS, SVR and ARIMA models for: (a) 0.5 h, (b) 1.0 h and (c) 24 h forecast horizons.

a grid search approach).

Comparable to existing studies in Australia (*e.g.*, [48,49,80,81]), this research has revealed the greater accuracy of the proposed models employed for forecasting. For instance, the SVR model [SVR(D_3)], applied for daily forecasting, attained an $RMSE_{\overline{G}}$ of 3.781% (Table 6), which is similar to 2.42% (whole weekly forecast) reported in [80]. Likewise, *MAE* and *RMSE* values for the weekly-average data forecasted in the same study were 224.18 MW and 311.04 MW, whereas for the SVR(D_3) model they resulted in 162.363 MW and 225.125 MW. Also, an adaptive neuro-wavelet model employed for *G* forecasting, in Queensland, showed a $0.16\% < MAE_{\overline{G}} < 0.99\%$ over 7 days in the test period [81]. Comparably, the $MAE_{\overline{G}}$ values were 0.355 for the MARS and 0.502 for the SVR models for the 0.5 h forecasting horizon in the present study. Moreover, recent studies [48,49] have adopted statistical approaches for 0.5 h forecasting to support the Australian Energy Market Operator; they have used the drivers of energy use (*e.g.*, temperatures, calendar effects, demographic and economic variables) in combination with demand and time of the year to forecast *G*. Differently to these studies, which adopted a semi-parametric additive model, the developed MARS and SVR models were an improvement as data assumptions or linear considerations were not employed. These models were guided by a fully data-driven modelling process.

Although this study was the first to evaluate the MARS and SVR models for short-term *G* forecasting in Queensland, multiple limitations should be addressed in future research. In this paper, the only predictor data used was time-lagged (historical) *G*. Alternative models for short-term horizons can also incorporate climate data (*e.g.*, temperature, rainfall, humidity and solar radiation) that modulate electricity demand influenced by consumers' needs under different conditions. According to previous work (*e.g.*, [19,82]), climatic factors can have an influence on *G*. For instance, an inverse relationship exists between electricity demand and ambient air temperature in wintertime, when lighting and heating usage are likely to increase. Similarly, this relationship can also occur in the summer when an increase in temperature can lead to increased air-conditioning demand [83]. Therefore, in a follow-up study, the MARS and SVR models could utilise seasonal data (both *G* and climatic factors). While this study provided accurate aggregated data models for Queensland, distinctive regions in the state are likely to exhibit different conditions.

In this study, a radial basis function was used to develop the SVR models employing a grid search to identify the parameters (C and σ). Despite the grid search demonstrating good performance, it is envisaged that a genetic algorithm (GA) [84] could serve to identify appropriate parameters for the model. GAs have been extensively applied to optimization problems [85–87]. According to [88], a GA-SVR was able to outperform other comparable models and yield high forecasting accuracy.

It is important to note that the MARS and SVR models could be improved by wavelet transformation (WT) and ensemble-based uncertainty testing via a bootstrapping procedure. This procedure uses a Bayesian Model Averaging (BMA) framework to assess the models' stability [52,89,90]. Many studies (*e.g.*, [17,91,92]) have suggested that WT could deliver benefits by decomposing predictor time series into time and frequency domains. Also, non-stationarity features in real data can be encapsulated by partitioning them with low and high pass filters. For example, very good results were obtained by a WT-SVR model for short and long-term solar forecasting when compared to the standard SVR model [17]. In addition, the data-driven technique of bootstrapping can also serve as an ensemble framework to reduce parametric uncertainties through resampling of inputs [93,94]. A hybrid wavelet-bootstrap-neural network model could be explored since such a model has outperformed non-WT models for water demand forecasting [51]. The use of the BMA also resulted in a better understanding of model uncertainty compared to a simple equal-weighted forecasting averaging method [95]. In addition to WT-based models, empirical model decomposition, applied for *G* forecasting in New South Wales (Australia), could similarly be employed in the present region to improve the MARS and SVR models. Considering other work [17,96–98], it is recommended that future research applies the WT, ensembles, and BMA to explore their usefulness for *G* forecasting.

6. Concluding remarks

Data-driven models based on the MARS, SVR and ARIMA algorithms were evaluated for short-term *G* forecasting using Queensland's area-aggregated data from the Australian Energy Market Operator. To demonstrate their feasibility for real-time applications, partial autocorrelation functions were applied to *G* data to identify significant inputs for three forecast horizons: 0.5 h, 1.0 h, and 24 h, with an identical number of predictive features (Table 2 and Fig. 2).

The versatility of the trained models for shorter span predictor data (31-12-2015) was investigated. Performances were assessed via correlation coefficient (r) between observed and forecasted *G* data in the testing period along with other performance metrics such as root mean square error (*RMSE*), mean absolute error (*MAE*), relative *RMSE* and *MAE* (%), Willmott's Index (*WI*), Nash–Sutcliffe coefficient (E_{NS}), and Legates and McCabe Index (E_{LM}). In terms of the statistical metrics, the

MARS model yielded the most accurate results for 0.5 h and 1.0 h forecasts, whereas the SVR models were better for a 24 h horizon. As expected, given its linear formulation in the modelling process, the ARIMA model's performance was lower for all forecasting horizons as it generated very high forecast errors.

Although this paper has advanced the work of previous studies (*e.g.*, [48,49,80,81]), it is also a pilot study in the context of the present study region (*i.e.*, Queensland). Future studies will employ Energex *G* data coupled with exogenous climate predictors for identified substations in the metropolitan Queensland area with the largest populations (*i.e.*, Brisbane, Gold Coast, Sunshine Coast, Logan, Ipswich, Redlands and Moreton Bay). The aim is to apply the MARS and SVR models via wavelet transformation and incorporate an ensemble framework and BMA to explore a coherent mechanism for uncertainty in forecasting models.

To summarize, the MARS and SVR models represent useful data-driven tools that can be used for *G* forecasting, and as such, they should be explored by forecasters working in the National Electricity Market (*e.g.*, AEMO). In particular, this study found that the MARS models provide a powerful, yet simple and fast forecasting framework when compared to the SVR models. However, the incorporation of a data pre-processing scheme (*e.g.*, wavelet transformation or empirical mode decomposition) as well as model uncertainty tools (*e.g.*, Bayesian and ensemble models) are alternative tools that could be explored for energy demand forecasting for engineering applications in an independent follow-up study.

Acknowledgments

Data were acquired from the Australian Energy Market Operator (AEMO). We thank Associate Professor Xiaohu Wen (Cold and Arid Regions Engineering and Environmental Institute, China) for his initial advice on the SVR model and the Ministry of Higher Education and Scientific Research, in the Government of Iraq for funding the first author's PhD project. We extend thanks to Dr. Georges Dodds for editing. Dr Ravinesh Deo acknowledges the Academic Development and Outside Studies Program and CAS Presidential Fellowship. We thank both Reviewers and the Handling Editor Professor Timo Hartmann for their constructive criticisms that helped improve the clarity of the paper.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.aei.2017.11.002.

References

[1] Z. Hu, Y. Bao, T. Xiong, Electricity load forecasting using support vector regression with memetic algorithms, Scient. World J. 2013 (2013).
[2] F. Kaytez, M.C. Taplamacioglu, E. Cam, F. Hardalac, Forecasting electricity consumption: a comparison of regression analysis, neural networks and least squares support vector machines, Int. J. Electr. Power Energy Syst. 67 (2015) 431–438.
[3] D. Akay, M. Atak, Grey prediction with rolling mechanism for electricity demand forecasting of Turkey, Energy 32 (2007) 1670–1675.
[4] S. Fan, L. Chen, Short-term load forecasting based on an adaptive hybrid method, IEEE Trans. Power Syst. 21 (2006) 392–401.
[5] D. Bunn, E.D. Farmer, Comparative models for electrical load forecasting, 1985.
[6] T. Haida, S. Muto, Regression based peak load forecasting using a transformation technique, IEEE Trans. Power Syst. 9 (1994) 1788–1794.
[7] E. Erdogdu, Electricity demand analysis using cointegration and ARIMA modelling: a case study of Turkey, Energy Policy 35 (2007) 1129–1146.
[8] H. Zareipour, K. Bhattacharya, C. Canizares, Forecasting the hourly Ontario energy price by multivariate adaptive regression splines, 2006 IEEE Power Engineering Society General Meeting, IEEE, 2006, p. 7.
[9] G. Nasr, E. Badr, M. Younes, Neural networks in forecasting electrical energy consumption: univariate and multivariate approaches, Int. J. Energy Res. 26 (2002) 67–78.
[10] L. Suganthi, T. Jagadeesan, Energy substitution methodology for optimum demand variation using Delphi technique, Int. J. Energy Res. 16 (1992) 917–928.
[11] N. Dalkey, O. Helmer, An experimental application of the Delphi method to the use of experts, Manage. Sci. 9 (1963) 458–467.
[12] L. Suganthi, A.A. Samuel, Energy models for demand forecasting—a review, Renew. Sustain. Energy Rev. 16 (2012) 1223–1240.
[13] E. Xydas, C. Marmaras, L.M. Cipcigan, N. Jenkins, S. Carroll, M. Barker, A data-driven approach for characterising the charging demand of electric vehicles: A UK case study, Appl. Energy 162 (2016) 763–771.
[14] M. Florens, I.H. Cairns, S. Knock, P. Robinson, Data-driven solar wind model and prediction of type II bursts, Geophys. Res. Lett. 34 (2007).
[15] J. Contreras, R. Espinola, F.J. Nogales, A.J. Conejo, ARIMA models to predict next-day electricity prices, IEEE Trans. Power Syst. 18 (2003) 1014–1020.
[16] A. Sözen, M. Ali Akçayol, Modelling (using artificial neural-networks) the performance parameters of a solar-driven ejector-absorption cycle, Appl. Energy 79 (2004) 309–325.
[17] R.C. Deo, X. Wen, F. Qi, A wavelet-coupled support vector machine model for forecasting global incident solar radiation using limited meteorological dataset, Appl. Energy 168 (2016) 568–593.
[18] A.K. Singh, S.K. Ibraheem, M. Muazzam, D. Chaturvedi, An overview of electricity demand forecasting techniques, Network Complex Syst. 3 (2013) 38–48.
[19] C. Sigauke, D. Chikobvu, Daily peak electricity load forecasting in South Africa using a multivariate non-parametric regression approach, ORiON. 26 (2010) 97.
[20] A.J. Smola, B. Schölkopf, Learning with Kernels, Citeseer, 1998.
[21] P.-S. Yu, S.-T. Chen, I.-F. Chang, Support vector regression for real-time flood stage forecasting, J. Hydrol. 328 (2006) 704–716.
[22] S. Salcedo-Sanz, J.L. Rojo-Álvarez, M. Martínez-Ramón, G. Camps-Valls, Support vector machines in engineering: an overview, Wiley Interdiscipl. Rev.: Data Min. Knowl. Discov. 4 (2014) 234–267.
[23] B.E. Türkay, D. Demren, Electrical load forecasting using support vector machines, 2011 7th International Conference on Electrical and Electronics Engineering (ELECO), IEEE, 2011, pp. I-49–I-53.
[24] M. Mohandes, Support vector machines for short-term electrical load forecasting, Int. J. Energy Res. 26 (2002) 335–345.
[25] C. Sivapragasam, S.-Y. Liong, Flow categorization model for improving forecasting, Hydrol. Res. 36 (2005) 37–48.
[26] R.C. Deo, P. Samui, D. Kim, Estimation of monthly evaporative loss using relevance vector machine, extreme learning machine and multivariate adaptive regression spline models, Stoch. Env. Res. Risk Assess. 1–16 (2015).
[27] V. Sharda, S. Prasher, R. Patel, P. Ojasvi, C. Prakash, Performance of Multivariate Adaptive Regression Splines (MARS) in predicting runoff in mid-Himalayan micro-watersheds with limited data/Performances de régressions par splines multiples et adaptives (MARS) pour la prévision d'écoulement au sein de micro-bassins versants Himalayens d'altitudes intermédiaires avec peu de données, Hydrol. Sci. J. 53 (2008) 1165–1175.
[28] J.H. Friedman, Multivariate adaptive regression splines, Annals Stat. (1991) 1–67.
[29] N.O. Attoh-Okine, K. Cooger, S. Mensah, Multivariate adaptive regression (MARS) and hinged hyperplanes (HHP) for doweled pavement performance modeling, Constr. Build. Mater. 23 (2009) 3020–3023.
[30] P. Samui, Determination of ultimate capacity of driven piles in cohesionless soil: a multivariate adaptive regression spline approach, Int. J. Numer. Anal. Meth. Geomech. 36 (2012) 1434–1439.
[31] W. Zhang, A.T.C. Goh, Multivariate adaptive regression splines for analysis of geotechnical engineering systems, Comput. Geotech.. 48 (2013) 82–95.
[32] C. Morana, A semiparametric approach to short-term oil price forecasting, Energy Econ. 23 (2001) 325–338.
[33] R.C. Deo, O. Kisi, V.P. Singh, Drought forecasting in eastern Australia using multivariate adaptive regression spline, least square support vector machine and M5Tree model, Atmos Res. 184 (2017) 149–175, http://dx.doi.org/10.1016/j.atmosres.2016.10.004.
[34] P. Sephton, Forecasting recessions: Can we do better on mars, Federal Reserve Bank of St Louis Rev. 83 (2001).
[35] A. Abraham, D. Steinberg, Is Neural Network a Reliable Forecaster on Earth? A MARS query! Bio-Inspired Applications of Connectionism, Springer, 2001, pp. 679–686.
[36] W.K. Buchanan, P. Hodges, J. Theis, Which way the natural gas price: an attempt to predict the direction of natural gas spot price movements using trader positions, Energy Econ. 23 (2001) 279–293.
[37] C. Yuan, S. Liu, Z. Fang, Comparison of China's primary energy consumption forecasting by using ARIMA (the autoregressive integrated moving average) model and GM (1, 1) model, Energy 100 (2016) 384–390.
[38] N.H. Miswan, N.H. Hussin, R.M. Said, K. Hamzah, E.Z. Ahmad, ARAR Algorithm in Forecasting Electricity Load Demand in Malaysia, Glob. J. Pure Appl. Math. 12 (2016) 361–367.
[39] AEU, Australian Energy Update. Department of Industry and Science, 2015, Australian energy update, Canberra, August, 2015.
[40] AEMO, Aggregated Price and Demand Data - Historical. Australia: The Australian Energy Market Operator, 2016. p. The Australian Energy Market Operator is responsible for operating Australia's largest gas and electricity markets and power systems.
[41] K. Mohammadi, S. Shamshirband, M.H. Anisi, K.A. Alam, D. Petković, Support vector regression based prediction of global solar radiation on a horizontal surface, Energy Convers. Manage. 91 (2015) 433–441.
[42] K. Mohammadi, S. Shamshirband, C.W. Tong, M. Arif, D. Petković, S. Ch, A new hybrid support vector machine–wavelet transform approach for estimation of horizontal global solar radiation, Energy Convers. Manage. 92 (2015) 162–171.
[43] V.N. Vapnik, Statistical Learning Theory, Wiley, New York, 1998.
[44] J.A.K. Suykens, J. De Brabanter, L. Lukas, J. Vandewalle, Weighted least squares support vector machines: robustness and sparse approximation, Neurocomputing. 48 (2002) 85–105.

[45] P.A.W. Lewis, J.G. Stevens, Nonlinear modeling of time series using multivariate adaptive regression splines (MARS), J. Am. Stat. Assoc. 86 (1991) 864–877.
[46] G.E. Box, G.M. Jenkins, Time Series Analysis: Forecasting and Control, revised ed., Holden-Day, 1976.
[47] R.E. Rosca, Stationary and non-stationary time series, USV Ann. Econ. Public Administr. 10 (2011) 177–186.
[48] S. Fan, R.J. Hyndman, Forecasting long-term peak half-hourly electricity demand for Queensland. Business & Economic Forecasting Unit (Monash University): Report for The Australian Energy Market Operator (AEMO), 2015.
[49] S. Fan, R.J. Hyndman, Monash Electricity Forecasting Model (Version 2015.1). Business & Economic Forecasting Unit (Monash University): Report for The Australian Energy Market Operator (AEMO), 2015.
[50] K. Sudheer, A. Gosain, K. Ramasastri, A data-driven algorithm for constructing artificial neural network rainfall-runoff models, Hydrol. Process. 16 (2002) 1325–1330.
[51] M.K. Tiwari, J. Adamowski, Urban water demand forecasting and uncertainty assessment using ensemble wavelet-bootstrap-neural network models, Water Resour. Res. 49 (2013) 6486–6507.
[52] M.K. Tiwari, C. Chatterjee, A new wavelet–bootstrap–ANN hybrid model for daily discharge forecasting, J. Hydroinformat. 13 (2011) 500–519.
[53] C.-W. Hsu, C.-C. Chang, C.-J. Lin, A practical guide to support vector classification, 2003.
[54] H.-T. Lin, C.-J. Lin, A study on sigmoid kernels for SVM and the training of non-PSD kernels by SMO-type methods, Neural Comput. (2003) 1–32.
[55] C.-C. Chang, C.-J. Lin, LIBSVM: a library for support vector machines, ACM Trans. Intell. Syst. Technol. (TIST). 2 (2011) 27.
[56] S.S. Keerthi, C.-J. Lin, Asymptotic behaviors of support vector machines with Gaussian kernel, Neural Comput. 15 (2003) 1667–1689.
[57] R. Maity, P.P. Bhagwat, A. Bhatnagar, Potential of support vector regression for prediction of monthly streamflow using endogenous property, Hydrol. Process. 24 (2010) 917–923.
[58] M.K. Goyal, B. Bharti, J. Quilty, J. Adamowski, A. Pandey, Modeling of daily pan evaporation in sub tropical climates using ANN, LS-SVR, Fuzzy Logic, and ANFIS, Exp. Syst. Appl. 41 (2014) 5267–5276.
[59] N.-D. Hoang, A.-D. Pham, M.-T. Cao, A novel time series prediction approach based on a hybridization of least squares support vector regression and swarm intelligence, Appl. Comput. Intell. Soft Comput. 2014 (2014) 15.
[60] G. Jekabsons, Adaptive Regression Splines toolbox for Matlab/Octave. Version. 2013;1:72.
[61] C. Kooperberg, D.B. Clarkson, Hazard regression with interval-censored data, Biometrics (1997) 1485–1494.
[62] S. Milborrow, Multivariate Adaptive Regression Splines. Package 'earth': Derived from mda:mars by Trevor Hastie and Rob Tibshirani Uses Alan Miller's Fortran utilities with Thomas Lumley's leaps wrapper, 2016 < http://www.milbo.users.sonic.net/earth > .
[63] J. Adamowski, H. Fung Chan, S.O. Prasher, B. Ozga-Zielinski, A. Sliusarieva, Comparison of multiple linear and nonlinear regression, autoregressive integrated moving average, artificial neural network, and wavelet artificial neural network methods for urban water demand forecasting in Montreal, Canada, Water Resour. Res. 48 (2012).
[64] S. Hu, Akaike Information Criterion, Center for Research in Scientific Computation, 2007.
[65] Hyndman R, Khandakar Y. Automatic Time Series Forecasting: The Forecast Package for R 7. 2008. < https://wwwjstatsoftorg/article/view/v027i03 > , 2007.
[66] P. Krause, D. Boyle, F. Bäse, Comparison of different efficiency criteria for hydrological model assessment, Adv. Geosci. 5 (2005) 89–97.
[67] C.J. Willmott, Some comments on the evaluation of model performance, Bull. Am. Meteorol. Soc. 63 (1982) 1309–1313.
[68] C.J. Willmott, On the validation of models, Phys. Geogr. 2 (1981) 184–194.
[69] C.W. Dawson, R.J. Abrahart, L.M. See, HydroTest: a web-based toolbox of evaluation metrics for the standardised assessment of hydrological forecasts, Environ. Modell. Software 22 (2007) 1034–1052.
[70] T. Chai, R.R. Draxler, Root mean square error (RMSE) or mean absolute error (MAE)?–arguments against avoiding RMSE in the literature, Geoscient. Model Develop. 7 (2014) 1247–1250.
[71] J. Nash, J. Sutcliffe, River flow forecasting through conceptual models part I—a discussion of principles, J Hydrol. 10 (1970) 282–290.
[72] C.J. Willmott, S.M. Robeson, K. Matsuura, A refined index of model performance, Int. J. Climatol. 32 (2012) 2088–2094.
[73] C.J. Willmott, On the Evaluation of Model Performance in Physical Geography. Spatial Statistics and Models, Springer, 1984, pp. 443–460.
[74] D.R. Legates, G.J. McCabe, Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation, Water Resour. Res. 35 (1999) 233–241.
[75] B.P. Wilcox, W. Rawls, D. Brakensiek, J.R. Wight, Predicting runoff from rangeland catchments: a comparison of two models, Water Resour. Res. 26 (1990) 2401–2410.
[76] M.-F. Li, X.-P. Tang, W. Wu, H.-B. Liu, General models for estimating daily global solar radiation for different solar radiation zones in mainland China, Energy Convers. Manage. 70 (2013) 139–148.
[77] A.B. Heinemann, P.A. van Oort, D.S. Fernandes, A.d.H.N. Maia, Sensitivity of APSIM/ORYZA model due to estimation errors in solar radiation, Bragantia 71 (2012) 572–582.
[78] R. Deo, N. Downs, A. Parisi, J. Adamowski, J. Quilty, Very short-term reactive forecasting of the solar ultraviolet index using an extreme learning machine integrated with the solar zenith angle, Environ. Res. 155 (2017) 141.
[79] M. Niu, S. Sun, J. Wu, L. Yu, J. Wang, An innovative integrated model using the singular spectrum analysis and nonlinear multi-layer perceptron network optimized by hybrid intelligent algorithm for short-term load forecasting, Appl. Math. Model. 40 (2016) 4079–4093.
[80] N. An, W. Zhao, J. Wang, D. Shang, E. Zhao, Using multi-output feedforward neural network with empirical mode decomposition based signal filtering for electricity demand forecasting, Energy 49 (2013) 279–288.
[81] B.-L. Zhang, Z.-Y. Dong, An adaptive neural-wavelet model for short term load forecasting, Electr. Power Syst. Res. 59 (2001) 121–129.
[82] C.-L. Hor, S.J. Watson, S. Majithia, Analyzing the impact of weather variables on monthly electricity demand, IEEE Trans. Power Syst. 20 (2005) 2078–2085.
[83] UKCIP, London's warming - The Impacts of Climate Change on London. Technical report. The UK Climate Impacts Programme (UKCIP), 2002.
[84] D. Zhang, W. Liu, A. Wang, Q. Deng, Parameter optimization for support vector regression based on genetic algorithm with simplex crossover operator, J. Inform. Comput. Sci. 8 (2011) 911–920.
[85] D.E. Golberg, Genetic algorithms in search, optimization, and machine learning. Addion wesley, 1989;1989:102.
[86] J.J. Grefenstette, Optimization of control parameters for genetic algorithms, IEEE Trans. Syst., Man, Cybernet. 16 (1986) 122–128.
[87] D.B. Fogel, An introduction to simulated evolutionary optimization, IEEE Trans. Neural Networks 5 (1994) 3–14.
[88] C.-H. Wu, G.-H. Tzeng, Y.-J. Goo, W.-C. Fang, A real-valued genetic algorithm to optimize the parameters of support vector machine for predicting bankruptcy, Exp. Syst. Appl. 32 (2007) 397–408.
[89] J.M. Sloughter, T. Gneiting, A.E. Raftery, Probabilistic wind speed forecasting using ensembles and Bayesian model averaging, J. Am. Stat. Assoc. 105 (2010) 25–35.
[90] Q. Feng, X. Wen, J. Li, Wavelet analysis-support vector machine coupled models for monthly rainfall forecasting in arid regions, Water Resour. Manage. 29 (2015) 1049–1065.
[91] P. Areekul, T. Senjyu, N. Urasaki, A. Yona, Neural-wavelet approach for short term price forecasting in deregulated power market, J. Int. Council Electr. Eng. 1 (2011) 331–338.
[92] Z. Tan, J. Zhang, J. Wang, J. Xu, Day-ahead electricity price forecasting using wavelet transform combined with ARIMA and GARCH models, Appl. Energy 87 (2010) 3606–3610.
[93] B. Efron, Bootstrap Methods: Another Look At the Jackknife. Breakthroughs in Statistics, Springer, 1992, pp. 569–593.
[94] B. Efron, R.J. Tibshirani, An Introduction to the Bootstrap, CRC Press, 1994.
[95] J.H. Wright, Forecasting US inflation by Bayesian model averaging, J. Forecast. 28 (2009) 131–144.
[96] J. Kim, B.P. Mohanty, Y. Shin, Effective soil moisture estimate and its uncertainty using multimodel simulation based on Bayesian Model Averaging, J. Geophys. Res.: Atmos. 120 (2015) 8023–8042.
[97] S. Crimp, K.S. Bakar, P. Kokic, H. Jin, N. Nicholls, M. Howden, Bayesian space–time model to analyse frost risk for agriculture in Southeast Australia, Int. J. Climatol. 35 (2015) 2092–2108.
[98] R. Maheswaran, R. Khosa, Long term forecasting of groundwater levels with evidence of non-stationary and nonlinear characteristics, Comput. Geosci. 52 (2013) 422–436.