

Short-term electricity demand forecasting using machine learning methods enriched with ground-based climate and ECMWF Reanalysis atmospheric predictors in southeast Queensland, Australia



Mohanad S. AL-Musaylh^{a,b,*}, Ravinesh C. Deo^{a,*}, Jan F. Adamowski^c, Yan Li^a

^a School of Agricultural, Computational and Environmental Sciences, Institute of Life Sciences and the Environment, 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

ARTICLE INFO

Keywords:

Predictive model for electricity demand
Climate and ECMWF Reanalysis variables
ANN
MARS
MLR
ARIMA
Hybrid ANN
Bootstrapping
Sustainable energy management systems

ABSTRACT

Reliable models that can forecast energy demand (G) are needed to implement affordable and sustainable energy systems that promote energy security. In particular, accurate G models are required to monitor and forecast local electricity demand. However, G forecasting is a multivariate problem, and thus models must employ robust pattern recognition algorithms that can detect subtle variations in G due to causal factors, such as climate variables. Therefore, this study developed an artificial neural network (ANN) model that used climatic variables for 6-hour (h) and daily G forecasting. The input variables included the six most relevant climate variables from Scientific Information for Land Owners (SILO) and 51 Reanalysis variables obtained from the European Centre for Medium-Range Weather Forecast (ECMWF) models. This information was used to forecast G data obtained from the energy utility (Energex) at 8 stations in southeast Queensland, Australia, by utilizing statistically significant lagged cross-correlations of G with its predictor variables. The developed ANN model was then benchmarked against multivariate adaptive regression spline (MARS), multiple linear regression (MLR), and autoregressive integrated moving average (ARIMA) models using various statistical metrics, such as relative root-mean square error (RRMSE%). Additionally, this study developed a hybrid ANN model by combining the forecasts of the ANN, MARS, and MLR models. The bootstrap (B) technique was also used with the hybrid ANN model, creating the B-hybrid ANN, to estimate the forecast uncertainty. According to both forecast horizons, the results indicated that the ANN model was more accurate than the ARIMA, MARS, and MLR models for G forecasting. Furthermore, the hybrid ANN was the most accurate model developed in this research study. For example, at the best site (Redcliffe), the hybrid ANN model generated an RRMSE of 3.85% and 4.37% for the 6-h and daily horizons, respectively. This study found that an ANN model could be used for accurately forecasting G over multiple horizons in southeast Queensland.

1. Introduction

In 2015 the 2030 Agenda for Sustainable Development was adopted by all United Nations member states, representing their commitment to the strategic implementation of 17 sustainable development goals (SDGs) [1]. In particular, Goal seven aims to improve global energy systems [1] through the achievement of five objectives by 2030. These include (1) provide global access to affordable, reliable, and modern energy, (2) greatly increase the share of renewable energy around the world, (3) double the universal rate of improving energy efficiencies,

(4) improve international collaborations and access to energy studies and skills, and (5) support technologies that supply up-to-date and sustainable energy services [1]. These objectives are important as global energy demand is expected to increase by over 50% before 2030 if worldwide demand growth rates are not curtailed [2].

An overview of projected changes in electricity demand (G) from 2020 to 2030 in the Australian Energy Market identified significant factors that affect the electricity industry [3]. These factors included potential changes in regional markets and the need for better systems to project demand accurately. It is expected that the demand of

* Corresponding authors. School of Agricultural, Computational and Environmental Sciences, Institute of Life Sciences and the Environment, 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).

Acronyms	
MW	Megawatt
G	Electricity load demand, MW
MARS	Multivariate adaptive regression splines
SVR	Support vector regression
RMSE	Root-mean square error
MAE	Mean absolute error
RRMSE	Relative root-mean square error, %
MAPE	Mean absolute percentage error, %
WI	Willmott's Index
E_{NS}	Nash-Sutcliffe efficiency coefficient
E_{LM}	Legates and McCabe's Index
SILO	Scientific Information for Land Owners
ECMWF	European Centre for Medium Range Weather Forecasts
G^{for}	G forecasted
ICEEMDAN	Improved version of empirical mode decomposition with adaptive noise
MLR	Multiple linear regression
G^{obs}	G observed
$trainbfg$	Levenberg-Marquardt
$trainlm$	Broyden-Fletcher-Goldfarb-Shanno
$tansig$	Tangent sigmoid
p	Autoregressive term in ARIMA
q	Moving average term in ARIMA
σ^2	Variance
$logsig$	Log sigmoid
ARIMA	Autoregressive integrated moving average
ANN	Artificial neural network
AEMO	Australian Energy Market Operator
$purelin$	Positive linear
L	Hidden neuron size
PACF	Partial auto-correlation function
B	Bootstrapping algorithm
R^2	Coefficient of determination
BMA	Bayesian model averaging
FA	Firefly algorithm
WT	Wavelet transforms
GCV	Generalized cross-validation
G_i^{for}	i th forecasted value of G , MW
G_i^{obs}	i th observed value of G , MW
\overline{G}^{for}	Mean of forecasted values
\overline{G}^{obs}	Mean of observed values
FE	Absolute forecasted error statistics
PSO	Particle swarm optimization
LH	Log likelihood
d	Degree of differencing in ARIMA
AIC	Akaike information criterion
SDGs	Sustainable development goals

conventional energy sources in Australia will be reduced over the next decade from increased renewable energy supply, improvements in the energy efficiency of building structures, and the use of solar panels in residential areas [3]. However, the short and long-term stochastic behaviour or subtle variation of energy demand is influenced by several exogenous factors. These factors include changes in solar radiation (*i.e.*, impacting the solar energy industry), wind regimes (*i.e.*, impacting the wind energy industry), and localized air temperature (*i.e.*, driving people to demand more energy to improve their comfort in extremely warm and cold weather). Given the potential influence of these exogenous factors on energy demand, a variety of predictor datasets are needed when developing a model for G monitoring and forecasting. Accurate knowledge of current and forecasted G could aid the design of energy devices that can support nations in achieving Goal 7 of the SDGs.

In regional and national energy grid systems, accurate G forecasts can support the development of energy security strategies. In particular, G forecasts can help in addressing issues of localized energy fluctuation in distribution networks, energy system installations, renewable energy investments, electricity demand planning, and creation of new energy management policies [4]. Currently, in Australia, forecasted energy demand is primarily managed by forecasters working in the Australian Energy Market Operator (AEMO) [5], who utilize semi-parametric-type additive models that explore multiple input variables such as air temperature and calendar effects, as well as demographic and economic datasets associated with energy demands for whole states. Additive models are useful for approximating high-dimensional regression functions and extending generalized linear models to combine statistical learning with interpretability and flexibility. However, these models do not consider demand data for smaller areas, such as substations, which can be significantly affected by climate variables.

Furthermore, these models have limitations due to assumptions of the data distribution, a tendency for overfitting, and a loss of predictability when smoothed variables have values outside of the training data range [6]. However, recent advancements in data analytics have led to improvements in machine learning methods through the black-box approach, which has subsequently become a standard methodology for G forecasting [23]. For example, a number of recent studies have

forecasted G using statistically significant lag combinations of historical demand data along with partial autocorrelation functions (*e.g.*, [7–12]), where different data-driven models for G forecasting were developed with machine learning algorithms.

Specifically, G data for the entire area of Queensland, Australia, were forecasted by Al-Musaylh and Deo [7] using multivariate adaptive regression splines (MARS), support vector regressions (SVR), and autoregressive integrated moving average (ARIMA) models. The authors concluded that the MARS algorithm performed the best for the short-term (0.5 h and 1.0 h) forecast horizons, and SVR performed the best for the daily (24 h) forecast horizon. Another study [8], by the same authors, led to an improvement in forecasting accuracy through the use of a two-phase particle swarm optimized (PSO)-SVR hybrid model, which was integrated with improved empirical mode decomposition with adaptive noise (ICEEMDAN) as a pre-processing algorithm to decompose G data (before running the forecasting model). In that study, multiple forecast horizons, including short-term (*i.e.*, weekends, working days, whole weeks, and public holidays) and long-term (*i.e.*, monthly) horizons, were employed to test the predictive ability of the ICEEMDAN-PSO-SVR hybrid approach. These studies, which specifically focused on Australia, contributed to the utilization of machine learning algorithms to forecast electrical energy demand.

Elsewhere, the performance of artificial neural network (ANN) models were tested using statistically significant combinations for different datasets including energy prediction data for buildings [13,14], electricity price data [15], and energy demand data [16,17]. Although a reliable forecasting accuracy for the energy demand model was obtained using statistically significant lagged inputs of real demand series [7,8], model performance was improved when other variables were incorporated into the model, likely because several interacting elements influence changes in electricity demand. Localized weather-related variables, such as air temperature, rainfall, wind regimes, and cloud cover, are likely to affect electrical energy usage.

A literature review revealed that a limited number of research studies (*e.g.*, [18–20]) have incorporated exogenous predictor variables from climate-based elements to forecast G data. Research undertaken by Mirasgedis et al. [18] focused on a region in Greece where a few meteorological variables, such as air temperature, humidity, wind

speed, and solar radiation were employed for medium-term G^{for} . They concluded that air temperature and humidity could be crucial predictor variables used in an energy demand model. Additionally, in South Africa, Lebotsa et al. [19] used calendar effects, air temperature, and a lagged demand dataset to train a partial linear additive quantile regression model for energy demand forecasting. Other research has indicated a possible causal relationship between climate change and energy demand without examining the feasibility of designing a demand model for G data forecasting [20]. In the context of Australia, where climate-based variables play a crucial role in forecasting solar radiation [21,22], exploratory studies that incorporate these variables into energy demand models are an interesting and novel research endeavour. Studies on multivariate methods for G forecasting can potentially inform engineers, power analysts, and climate-energy policymakers on how to improve energy demand models by considering the role of exogenous climate-based predictors in real-world energy forecasting problems.

Therefore, the current research presents a novel study that aimed to advance a new approach for energy demand forecasting by using extensive sets of climate-based predictors, including both ground-based measurements and atmospheric Reanalysis data obtained from the numerical weather forecasting model, to forecast station-based electricity energy demand. The study region was southeast Queensland, Australia, where ground-based model data were extracted from the Scientific Information for Land Owners (SILO), and numerical weather forecast data from the European Centre for Medium-Range Weather Forecasts (ECMWF). This study utilized a relatively wide range of input features and patterns from climate-based predictor variables that may be related to energy demand data in order to construct a robust model framework. A total of six ground-based variables and 51 ECMWF-based variables

were used. The energy demand data were extracted from the database of the primary energy utility, Energex, a leading power distributor in southeast Queensland.

To achieve this goal, physical and data-driven models were considered. While physically-based models are based on the physical behaviour of model parameters, data-driven models employ machine learning algorithms to detect the relationships between predictors and objective variables using historical datasets [23–25]. Recent studies (e.g., [26,27]), have found that data-driven forecasting models such as ANN led to improvements in forecasting accuracy over physical models such as the Predictive Ocean Atmosphere Model for Australia (POAMA). Five advantages of data-driven models, in comparison to physically-based models, have been identified [26,28,29]. Specifically, data-driven models (1) can be run and the output can be explained with less complexity in the future pattern's description of the climate variables, (2) are easier to assess, (3) require less data and costs, (4) require shorter training and testing periods, and (5) have high specificity. Therefore, applying data-driven techniques to create a G forecasting model can have many advantages as an alternative to physically-based models.

In this paper, an ANN model, a data-driven black-box tool that does not require detailed information about the predictor variables in comparison with physical techniques [23], was developed for G forecasting. The ANN model can predict future values of more than one variable at the same time and handle non-linear datasets [30]. The ANN model has various applications for data forecasting (e.g., [31,32]) when exogenous variables are used, and when it has been used globally, it has often achieved a relatively small margin of predictive errors for different data forecasting horizons (e.g., [16,28,32–34]). Additionally, different studies (e.g., [35–38]) have developed ensemble-based uncertainty

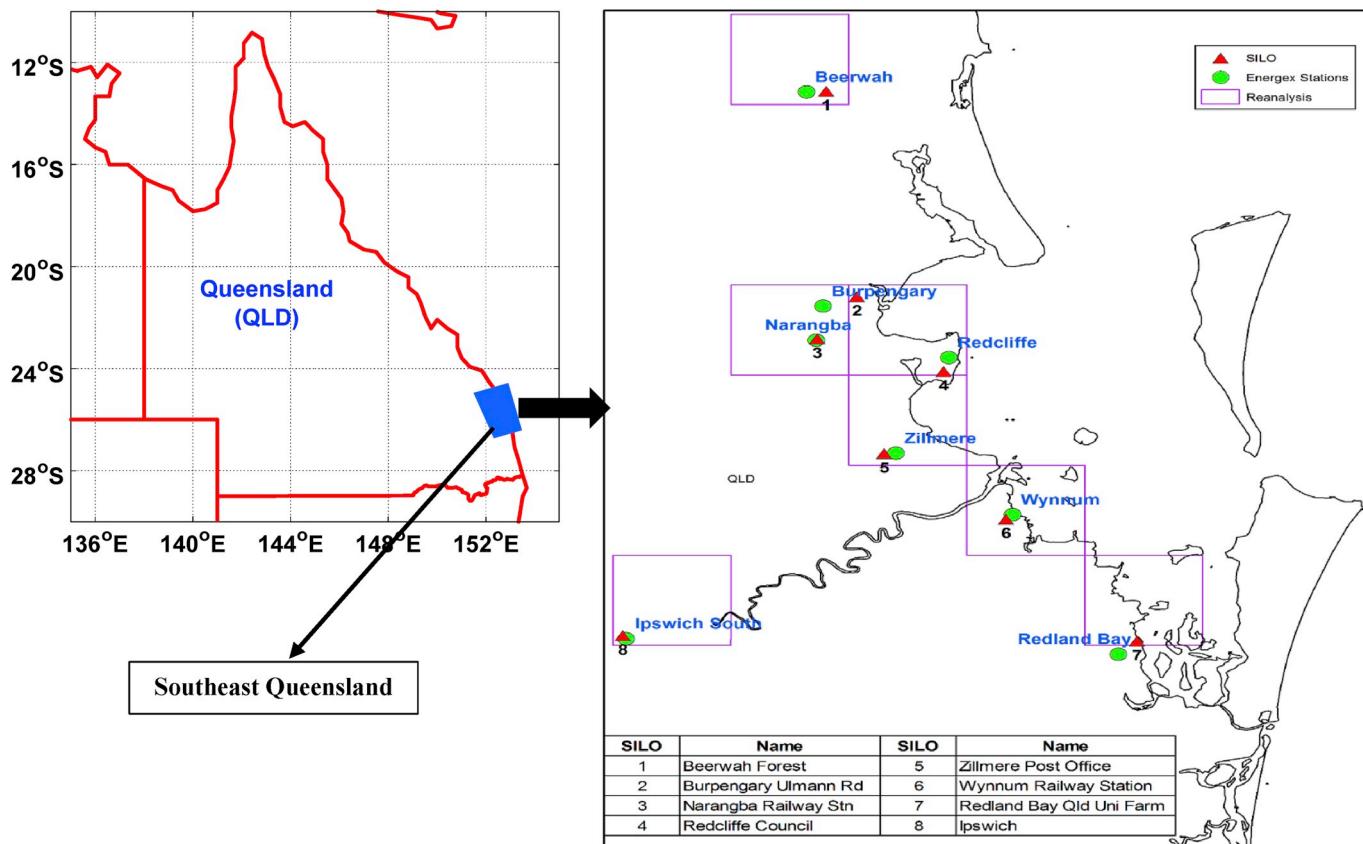


Fig. 1. Map of the present study area showing the locations in southeast Queensland where the energy demand forecasting models were developed and tested using climate-based predictor variables. (a) Energex stations are shown as green circles (b) SILO points are shown as red triangles and (c) ECMWF (ERA-Interim) Reanalysis grids are shown as pink boxes, where each grid contains four points with the mean of these points used as the data point for that grid. Information in Table 1 was used to generate the spatial map. (To view this figure in colour, the reader is referred to the Web version of this article).

Table 1

The longitudes and latitudes for data locations used in this study: (a) Energex (b) SILO (c) ECMWF (ERA-Interim) Reanalysis.

Energex Station			SILO			Reanalysis			
Name	Latitude	Longitude	Name	Latitude	Longitude	Latitude	Longitude	Longitude	
Beerwah	26.8573	152.9552	Beerwah Forest	26.8564	152.9764	26.7500	26.8750	152.8750	153.0000
Burpengary	27.1544	152.9728	Burpengary Ullmann Rd	27.1414	153.0089	27.1250	27.2500	153.0000	153.1250
Ipswich South	27.61551	152.7641	Ipswich	27.6117	152.7608	27.5000	27.6250	152.7500	152.8750
Narangba	27.2015	152.9655	Narangba Railway Stn	27.2000	152.9667	27.1250	27.2500	152.8750	153.0000
Redcliffe	27.2258	153.1063	Redcliffe Council	27.2450	153.1006	27.1250	27.2500	153.0000	153.1250
Redland Bay	27.6372	153.2857	Redland Bay Qld Uni Farm	27.6192	153.3056	27.5000	27.6250	153.2500	153.3750
Wynnum	27.4437	153.1736	Wynnum Railway Station	27.4500	153.1667	27.3750	27.5000	153.1250	153.2500
Zillmere	27.3583	153.0500	Zillmere Post Office	27.3589	153.0375	27.2500	27.3750	153.0000	153.1250

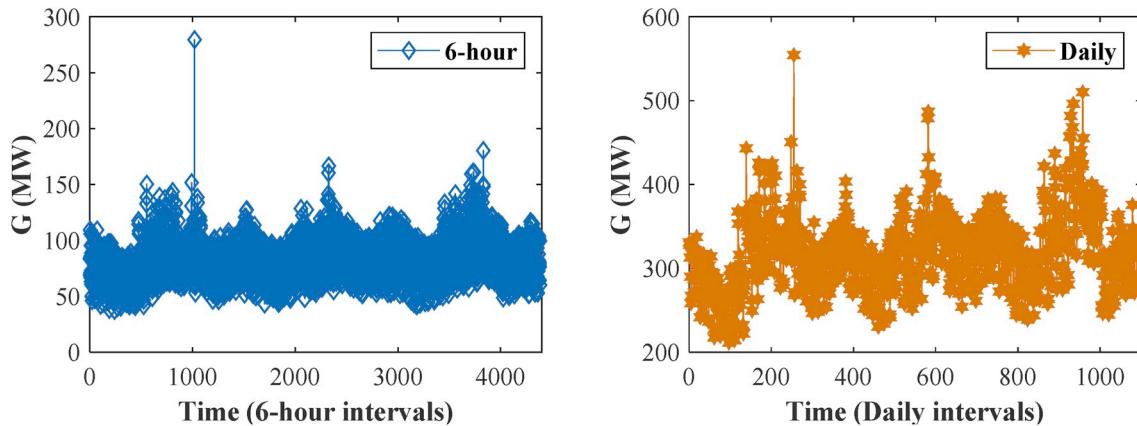


Fig. 2. Time-series of the electricity demand (G , MW) data used for the Beerwah study site for 6-h and daily periods.

Table 2

Descriptive statistics of the electricity demand (G , MW) for the eight Energex stations with 6-h and daily periods.

Forecast horizon	Data Period (dd-mm-yyyy)	Property	Beerwah	Burpengary	Ipswich South	Narangba	Redcliffe	Redland Bay	Wynnum	Zillmere
6-h	01-07-2014 to 30-06-2017	Minimum (MW)	37.40	21.55	38.81	21.27	73.74	18.09	59.40	78.39
		Maximum (MW)	279.52	258.00	347.98	251.66	235.45	359.37	266.43	600.87
		Mean (MW)	79.65	87.93	94.09	100.96	125.72	67.72	109.19	165.51
		Standard deviation (MW)	18.76	30.48	28.08	29.10	27.55	24.55	28.33	39.58
Daily	01-07-2014 to 29-06-2017	Minimum (MW)	211.54	234.71	258.18	220.63	392.69	203.88	301.63	485.10
		Maximum (MW)	554.27	678.53	696.52	683.55	718.56	557.40	724.38	1106.43
		Mean (MW)	318.59	351.76	376.39	403.85	502.89	270.90	436.76	662.01
		Standard deviation (MW)	44.68	61.77	62.90	78.78	55.90	44.62	59.46	87.58

forecasted models by using a bootstrap (B) procedure to assess the model's stability based on the prediction intervals of the ANN models. In particular, B is a data-driven technique that can utilize an ensemble framework to help reduce the prediction uncertainties through intensive resampling with replacement [39,40]. In Australia, three studies [23,31,32] have successfully applied the ANN model to forecast precipitation, solar radiation, and streamflow. However, the ability of the ANN model integrated with large datasets to forecast G has not yet been explored everywhere.

Therefore, in this study, the ANN model was evaluated against two alternative multivariate models, MARS and multiple linear regression (MLR), developed as high-dimensional models with several predictors, as well as the low dimensional (simple) model, ARIMA. Various statistical metrics were used to assess the performance of the ANN for G forecasting, in comparison to the other models employed. In addition, a hybrid ANN model was constructed by combining the three models of ANN, MARS, and MLR. For the hybrid model, the prediction interval (PI) was calculated using the B-hybrid ANN model to address uncertainty assessment for G forecasting.

There were five main objectives of this study, (1) to extract datasets from SILO and ECMWF to be used as predictor variables, and datasets

from Energex to be used as the target variable, (2) to develop an ANN model for short-term forecasting of 6-hour (h) and daily horizons, (3) to evaluate the performance of the ANN model in comparison to the MARS, MLR, and ARIMA models, (4) to construct a hybrid ANN model by combining ANN with MARS and MLR, and (5) to estimate the forecast uncertainty using the B-hybrid ANN model. These objectives are achieved in the following sections. Section 2 explains the supporting theory for the different models, while section 3 provides the materials and methods, which include the data, forecast model development, and assessment criteria. Section 4 presents the results and discussion, while section 5 provides possible solutions to address challenges in future work. Finally, section 6 summarizes the research results and contributions.

2. Conceptual framework

In this section, only the ANN, MLR, and B methods are described thoroughly. The details of the MARS, and ARIMA strategies, presented by Friedman [41] and Box and Jenkins [42], are well explained in previous works (e.g., [7,8,31,43–45]), and thus not presented here.

Table 3Model input variables from SILO and ERA-Interim Reanalysis used for 6-h and daily G^{for}

Source	No.	Variable name	Acronym	Mean (e.g. Beerwah data)	
				6-h	Daily
SILO	1	Maximum temperature	TMax	26.31	
	2	Minimum temperature	TMin	15.88	
	3	Rainfall	Rain	3.73	
	4	Evaporation	Evap	5.15	
	5	Solar radiation	Radn	18.28	
	6	Vapour pressure	VP	17.84	
Reanalysis	7	Vertical integral of mass of atmosphere	p53.162	10186.67	10186.64
	8	Vertical integral of temperature	p54.162	2601493.88	2601506.13
	9	Vertical integral of water vapour	p55.162	25.66	25.66
	10	Vertical integral of ozone	p58.162	0.01	0.01
	11	Vertical integral of kinetic energy	p59.162	2535683.21	2533465.76
	12	Vertical integral of thermal energy	p60.162	2611649009.46	2611660635.11
	13	Vertical integral of potential + internal energy	p61.162	2626949967.16	2626961540.62
	14	Vertical integral of potential + internal + latent energy	p62.162	2691110176.10	2691142610.04
	15	Vertical integral of total energy	p63.162	2693645851.60	2693676068.17
	16	Vertical integral of energy conversion	p64.162	164.85	165.06
	17	Vertical integral of eastward mass flux	p65.162	111538.38	111457.87
	18	Vertical integral of northward mass flux	p66.162	12733.86	12778.81
	19	Vertical integral of eastward kinetic energy flux	p67.162	94633097.10	94501469.70
	20	Vertical integral of northward kinetic energy flux	p68.162	-2924314.67	-2848793.90
	21	Vertical integral of eastward heat flux	p69.162	26271018743.55	26252081657.94
	22	Vertical integral of northward heat flux	p70.162	3401655077.02	3411787299.63
	23	Vertical integral of eastward water vapour flux	p71.162	10.25	10.25
	24	Vertical integral of northward water vapour flux	p72.162	26.61	26.58
	25	Vertical integral of eastward geopotential flux	p73.162	10698893652.73	10690448746.89
	26	Vertical integral of northward geopotential flux	p74.162	653385798.50	658668079.51
	27	Vertical integral of eastward total energy flux	p75.162	37090107708.45	37062579702.24
	28	Vertical integral of northward total energy flux	p76.162	4118857514.61	4134281550.79
	29	Vertical integral of eastward ozone flux	p77.162	-0.02	-0.02
	30	Vertical integral of northward ozone flux	p78.162	0.00	0.00
	31	Vertical integral of divergence of mass flux	p81.162	0.00	0.00
	32	Vertical integral of divergence of kinetic energy flux	p82.162	4.23	4.09
	33	Vertical integral of divergence of thermal energy flux	p83.162	105.85	106.17
	34	Vertical integral of divergence of moisture flux	p84.162	0.00	0.00
	35	Vertical integral of divergence of geopotential flux	p85.162	-178.25	-178.37
	36	Vertical integral of divergence of ozone flux	p87.162	0.00	0.00
	37	Vertical integral of northward cloud liquid water flux	p89.162	-0.02	-0.02
	38	Vertical integral of northward cloud frozen water flux	p91.162	-0.12	-0.12
	39	Vertical integral of mass tendency	p92.162	0.00	0.00
	40	Surface pressure	sp	99897.13	99896.77
	41	Total column water	tcw	25.73	25.74
	42	Total column water vapour	tcvv	25.66	25.66
	43	Soil temperature level 1	stl1	293.26	293.26
	44	Mean sea level pressure	msl	101702.71	101702.33
	45	Total cloud cover	tcc	0.38	0.38
	46	10 m U wind component	u10	-0.68	-0.68
	47	10 m V wind component	v10	0.49	0.49
	48	2 m temperature	t2m	293.53	293.53
	49	2 m dewpoint temperature	d2m	287.97	287.97
	50	Soil temperature level 2	stl2	292.82	292.82
	51	Albedo	al	0.15	0.15
	52	Soil temperature level 3	stl3	292.81	292.82
	53	Low cloud cover	lcc	0.15	0.15
	54	Medium cloud cover	mcc	0.15	0.15
	55	High cloud cover	hcc	0.17	0.17
	56	Total column ozone	tco3	0.01	0.01
	57	Soil temperature level 4	stl4	292.80	292.80

Table 4

Data splitting technique used for model development and testing.

Forecast horizon	Data Period (dd-mm-yyyy)	Number of data points		
		Total	Training (80%)	Testing (20%)
6-h	01-07-2014 to 30-06-2017	4383	3506	877
Daily	01-07-2014 to 29-06-2017	1095	876	219

2.1. Artificial neural network (ANN)

A non-linear regression problem with multiple inputs can be solved by an ANN technique [46] as follows:

$$y(x) = F \left(\sum_{i=1}^L w_i(t) \cdot x_i(t) + b \right) \quad (1)$$

where $x_i(t) = \{x_i\}_{i=1}^{i=N} \in R^N$ are the input variables (SILO, ECMWF Reanalysis or partial autocorrelation function (PACF) for G data), and $y(x) = \{y_i\}_{i=1}^{i=N} \in R$ is the target variable (G^{for}) in the training period. The values of b , $w_i(t)$, F , and L are the neuronal bias, the weight that

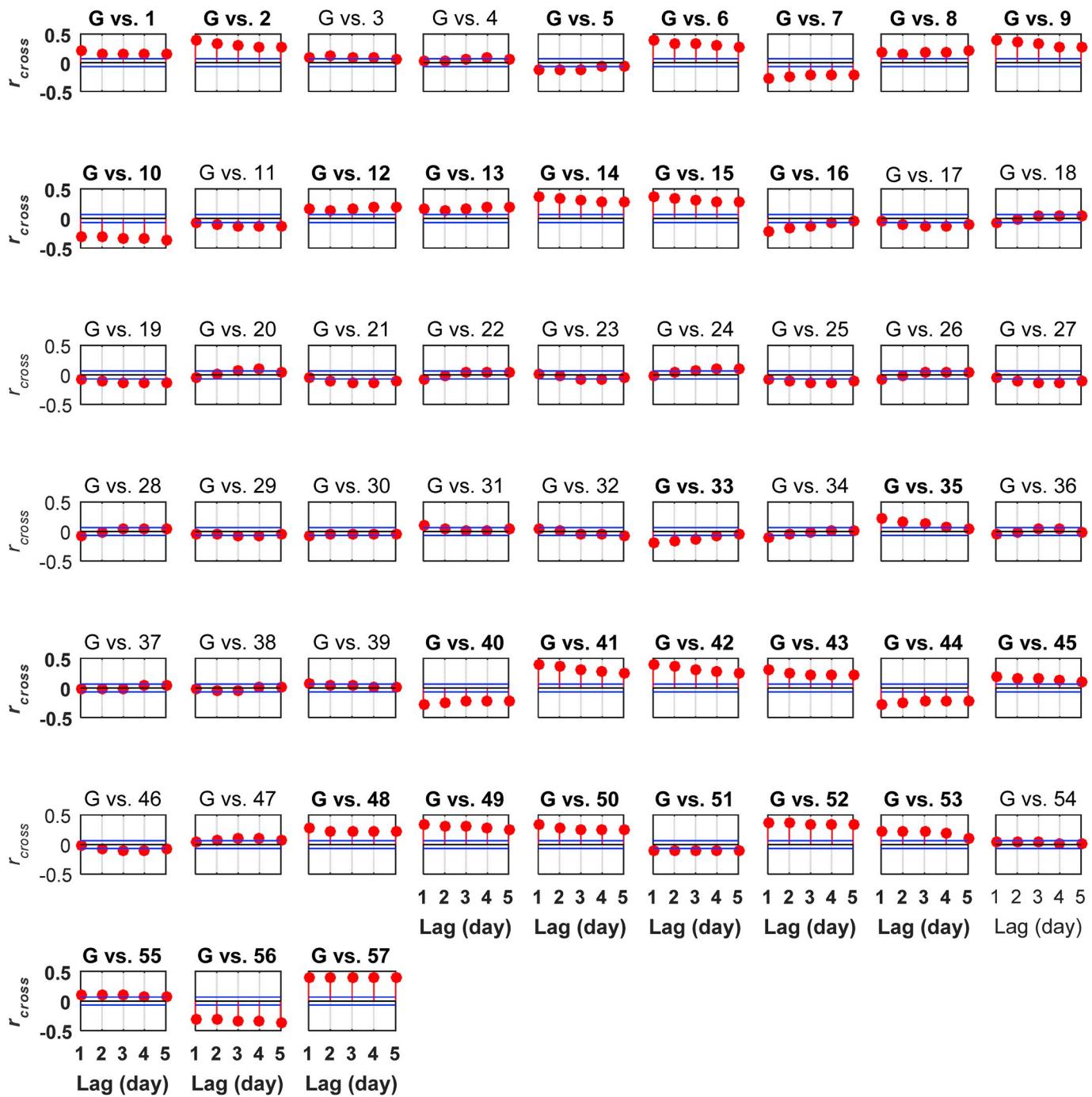


Fig. 3. Cross-correlation coefficients (r_{cross}) performed to investigate the co-variance between daily electricity demand (G) vs. the predictor variables from SILO (variables 1 to 6) and ERA-Interim Reanalysis dataset (variables 7 to 57) for the Beerwah station in the model training period. The 95% significance boundaries are shown in blue, indicating the variables selected to calculate statistically significant lagged 1 inputs matrix for the models. The selected variables are shown in boldface. Details of the variables are shown in Table 3. (To view this figure in colour, the reader is referred to the Web version of this article.)

connects the neuron in the input layer, the hidden transfer function, and the hidden neuron size determined iteratively, respectively.

The ANN algorithm is a black-box model that needs an iterative identification process to identify the training algorithm [31,32]. Therefore, two training MATLAB algorithms, Levenberg-Marquardt (LM) (*trainbfg*) and Broyden-Fletcher-GoldfarbShanno (BFGS) (*trainlm*) [47–49], were used to minimize the mean square error [50].

2.2. Multiple linear regression (MLR)

The MLR model, a statistical procedure that shows the relationship

between the inputs (SILO, ECMWF Reanalysis, and PACF G data) and the target (G) [31], was used in to evaluate the performance of the ANN model. The regression equation of the MLR can be expressed as [51,52]:

$$y(x) = c + a_1x_1 + a_2x_2 + \dots + a_kx_k \quad (2)$$

where x and y are defined in Eq. (1). The values of k , c , and a are the number of the predictor variables, the y -intercept, and the multiple regression coefficient for each regression variable, respectively [31,33,53].

Table 5

Optimum model development for the ANN, MARS, and ARIMA models, showing the models' parameters and predictor datasets for both forecast horizons in the training period.

Station	No. input variables (*) ANN				MARS				ARIMA (*)						
	Lag 1-SILO & Reanalysis	G-PACF lags	Hidden transfer function	Output transfer function	Training algorithm	Hidden neuron size (L)	Optimum number of spline function	Variable importance		p	d	q	σ^2	LH	AIC
								Lowest	Highest						
6-h forecast horizon															
Beerwah	21	5 (1–5)	logsig	tansig	trainlm	10	49	p92.162	G-lag 4	5	1	3	0.0015	6357.94	-12697.89
Burpengary	24	3 (2, 4, 5)	tansig	tansig	trainlm	8	47	u10	G-lag 4	5	1	2	0.0023	5697.31	-11378.62
Ipswich South	24	4 (1, 4–6)	logsig	tansig	trainlm	5	29	d2m	G-lag 4	5	1	3	0.0015	6456.23	-12894.46
Narangba	24	14 (1–14)	tansig	logsig	trainbfg	37	33	G-lag 11	G-lag 1	2	1	2	0.011	2895.47	-5780.95
Redcliffe	33	34 (1–34)	tansig	purelin	trainlm	6	49	G-lag 13	G-lag 4	2	1	2	0.0097	3113.22	-6216.45
Redland Bay	24	6 (1–6)	tansig	purelin	trainlm	5	49	p81.162	G-lag 4	4	1	3	0.0009	7261.34	-14506.68
Wynnum	20	9 (1–9)	tansig	purelin	trainlm	10	17	G-lag 3	G-lag 4	4	1	3	0.0022	5770.37	-11524.73
Zillmere	26	6 (1–6)	logsig	tansig	trainlm	4	17	G-lag 2	G-lag 4	4	1	3	0.0011	6992.26	-13968.53
Daily forecast horizon															
Beerwah	30	3 (1–3)	tansig	logsig	trainlm	2	27	stl3	G-lag 1	5	1	2	0.0057	1021.87	-2027.74
Burpengary	35	1	logsig	purelin	trainbfg	20	11	d2m	G-lag 1	2	1	1	0.0053	1047.28	-2086.56
Ipswich South	31	3 (1–3)	tansig	purelin	trainbfg	12	33	tcw	G-lag 1	2	1	1	0.0064	965.01	-1922.01
Narangba	24	3 (1–3)	tansig	purelin	trainbfg	5	23	TMax	G-lag 1	2	1	2	0.012	689.22	-1368.44
Redcliffe	37	8 (1–8)	tansig	purelin	trainbfg	2	41	stl3	G-lag 7	3	1	3	0.0073	906.40	-1798.80
Redland Bay	27	3 (1–3)	tansig	purelin	trainlm	1	17	G-lag 2	G-lag 1	3	0	1	0.0047	1105.41	-2198.82
Wynnum	32	3 (1–3)	tansig	logsig	trainbfg	3	31	p55.162	G-lag 1	1	0	3	0.005	1075.79	-2139.58
Zillmere	38	3 (1–3)	tansig	purelin	trainbfg	1	17	stl4	G-lag 1	2	1	3	0.0068	940.32	-1868.65

* Note that for the ANN and MARS models, the SILO and ERA-Interim Reanalysis data with the G-PACF lags were used for daily forecast horizon, whereas only ERA-Interim Reanalysis data with the G-PACF lags were used for 6-h forecast horizons due to the availability of data. However, only the G data with no lags (single input) were used for the ARIMA model for both forecasting horizons.

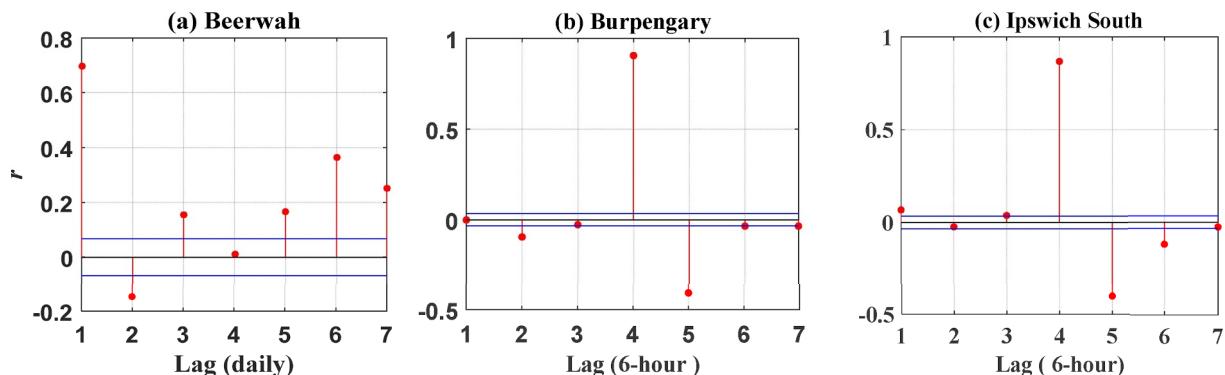


Fig. 4. Statistically significant lags of the electricity demand data (G) analysed in the training phase to construct the model's input variables. (a) The first three lags were selected for daily Beerwah data (b) Lags 2, 4 and 5 were selected for 6-h Burpengary data (c) Lags 1, 4, 5 and 6 were selected for 6-h Ipswich South data.

2.3. Bootstrapping procedures (ensemble approach)

The bootstrap is a data-driven technique that aims to reduce uncertainties through intensive resampling with replacement [39,40]. Suppose that T_n is a random sample of size n where $T_n = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ is extracted from a population with an unknown probability distribution F . Also, $t_i(x_i, y_i)$ is a realization drawn independently and identically distributed (iid) from F , and consists of an input vector x_i and the corresponding target variable y_i . For T_n , the empirical distribution function is defined as \hat{F} with a probability of $1/n$ for each $t_i(x_i, y_i)$, where $i = 1, 2, \dots, n$. The bootstrap sample of size, n , can be denoted by T_n^* , taken from iid with replacement from \hat{F} , and written as $T^1, T^2, \dots, T^S, \dots, T^S$, where S is the total number of bootstrap samples. According to Efron [39], the value of S is usually between 50 and 200 samples. A model can be developed for each T^s using all n observations in this study. The result will be considered as $f_{\text{model}}(x_i, w_s)$, where w_s is the weight vector for the developed model based on the

bootstrap sample, T^s . In the training period, the model performance will be evaluated using a set A from the observations, $t_i(x_i, y_i)$, that are not involved in the bootstrap samples. The estimation of the generalization error (\hat{E}_0) is used to calculate the model performance, as shown below [38,54]:

$$\hat{E}_0 = \frac{\sum_{s=1}^S \sum_{i \in A} (y_i - f_{\text{model}}(x_i, w_s))^2}{\sum_{s=1}^S (A)} \quad (3)$$

Jia and Culver [38] show that $\hat{y}(x)$ and $\sigma^2(x)$ are the average and the variance of the S bootstrapped estimates for every developed model, respectively:

$$\hat{y}(x) = \frac{1}{S} \sum_{s=1}^S f_{\text{model}}(x_i, w_s) \quad (4)$$

$$\sigma^2(x) = \frac{\sum_{s=1}^S (\hat{y}(x) - f_{\text{model}}(x_i, w_s))^2}{S-1} \quad (5)$$

Table 6

Optimum model development for hybrid ANN and B-hybrid ANN showing the models' parameters and predictor datasets for both forecast horizons in the training period.

Station	No. input variables	Hidden neuron size (L)	Hybrid ANN Hidden transfer function	Output transfer function	Training algorithm	B-hybrid ANN Hidden transfer function	Output transfer function	Training algorithm
6-h forecast horizon								
Beerwah	3	7	logsig	tansig	trainlm	logsig	tansig	trainbfg
Burpengary	3	10	tansig	purelin	trainlm	tansig	purelin	trainbfg
Ipswich South	3	4	logsig	tansig	trainbfg	logsig	tansig	trainbfg
Narangba	3	2	tansig	purelin	trainbfg	tansig	purelin	trainbfg
Redcliffe	3	8	tansig	tansig	trainlm	tansig	tansig	trainbfg
Redland Bay	3	4	logsig	purelin	trainbfg	logsig	purelin	trainbfg
Wynnum	3	3	tansig	tansig	trainlm	tansig	tansig	trainlm
Zillmere	3	5	tansig	tansig	trainlm	tansig	logsig	trainlm
Daily forecast horizon								
Beerwah	3	2	tansig	tansig	trainlm	tansig	tansig	trainlm
Burpengary	3	8	logsig	tansig	trainbfg	logsig	tansig	trainbfg
Ipswich South	3	4	tansig	purelin	trainlm	logsig	tansig	trainbfg
Narangba	3	4	tansig	purelin	trainbfg	logsig	tansig	trainlm
Redcliffe	3	8	logsig	logsig	trainbfg	logsig	logsig	trainbfg
Redland Bay	3	9	logsig	tansig	trainbfg	tansig	logsig	trainlm
Wynnum	3	3	logsig	purelin	trainlm	tansig	logsig	trainbfg
Zillmere	3	4	tansig	tansig	trainbfg	tansig	tansig	trainbfg

Table 7

Root-mean square error (RMSE, MW), for all the station datasets for the 6-h and daily forecast horizons in the training period using the ANN, MARS, MLR, hybrid ANN, and ARIMA models.

Station	6-h					Daily				
	ANN	MARS	MLR	Hybrid ANN	ARIMA	ANN	MARS	MLR	Hybrid ANN	ARIMA
Beerwah	5.88	7.76	9.05	5.83	9.52	24.38	25.23	26.62	23.86	25.65
Burpengary	6.49	9.37	11.45	6.31	11.22	21.89	28.51	28.94	21.53	32.37
Ipswich South	8.36	9.81	11.42	8.32	11.81	27.00	30.22	33.78	26.51	35.03
Narangba	9.08	13.18	13.96	9.04	24.31	40.71	50.15	51.57	39.81	50.74
Redcliffe	3.48	5.66	6.22	3.39	15.95	22.45	21.12	22.72	20.42	27.69
Redland Bay	6.89	8.14	9.85	6.88	10.35	21.76	20.69	22.49	19.78	24.10
Wynnum	4.80	8.66	8.90	4.78	9.60	26.20	25.31	27.26	24.67	29.81
Zillmere	11.54	14.11	16.06	11.35	17.19	48.17	46.16	48.05	44.71	51.08

Since the method contains repeated applications, the frequency (K) for the prediction interval (PI) that would contain the true value is calculated as $K = (1 - \alpha) * 100\%$. With $\alpha = 0.05$, $K = (1 - 0.05) * 100\% = 0.95\%$ represents the prediction bounds [35]. Efron and Tibshirani [40] stated that PI can be estimated using the equation below:

$$\text{PI} = [\text{UB}, \text{LB}] = [\hat{y}(x) + t_{n-p}^{\alpha/2} \sigma(x), \hat{y}(x) - t_{n-p}^{\alpha/2} \sigma(x)] \quad (6)$$

where UB and LB are the upper and lower bands, respectively, n is the total number of discharge observations, p is the total number of model parameters, and $t_{n-p}^{\alpha/2}$ is the $\alpha/2$ percentile for the Student t distribution with $n - p$ degrees of freedom.

3. Materials and methods

3.1. Electricity demand data (G)

In this study, G data were requested from Energex [55] for 01/07/2004 to 30/06/2017. The G data in megawatts (MW) were recorded every 30 min, using Brisbane (the capital city of Queensland) time, from various stations in southeast Queensland, which covered more than 200 suburbs in the cities of Brisbane, Gold Coast, Sunshine Coast, Logan, Ipswich, Redlands, and Moreton Bay. For some stations, there were data limitations, such as missing points, zeros, negative values, and absent G values for previous or subsequent years. Furthermore, at some stations, there were difficulties in matching G data with corresponding SILO and ECMWF Reanalysis datasets. However, the study examined all the stations in the different periods to obtain the best and most accurate target (G data) by avoiding all the data subject to the issues stated above. Consequently, data from eight Energex stations, which included Beerwah, Burpengary, Ipswich South, Narangba,

Redcliffe, Redland Bay, Wynnum, and Zillmere were used from 01/07/2014 to 30/06/2017 for the 6-h forecast horizon, and 01/07/2014 to 29/06/2017 for the daily forecast horizon. The locations of the Energex stations are shown in Fig. 1. The longitudes and latitudes of each station are indicated in Table 1.

The half-hour Energex data phases were converted to 6-h data phases by calculating the sum of every 12 values starting from 01/07/2014 at 4:30 a.m. until 30/06/2017 at 10:00 p.m. As a result, four points, constructed at 10:00 a.m. (from 4:30 a.m. to 10:00 a.m.), 4:00 p.m. (from 10:30 a.m. to 4:00 p.m.), 10:00 p.m. (from 4:30 p.m. to 10:00 p.m.), and 4:00 a.m. (from 10:30 p.m. to 4:00 a.m.), were obtained for each day. These four points were summed to get the daily G values. These data intervals were chosen so that the Energex G data were temporally aligned with the SILO and ECMWF Reanalysis datasets. Fig. 2 shows plots of the G data at the Beerwah site for both daily and 6-h horizons. Table 2 provides descriptive statistics for both forecast horizons at all sites.

3.2. SILO data

Historical climate datasets were extracted from SILO [56] for the period and locations corresponding to the G data retrieved from the eight Energex stations. The SILO station names, longitudes, and latitudes are shown in Table 1 and Fig. 1 in section 3.1. Table 3 shows the details of the SILO data, including the mean values for each variable, at the Beerwah site. The six SILO variables, which included the maximum and minimum air temperature (TMax and TMin), rainfall (Rain), evaporation (Evap), solar radiation (Radn), and vapour pressure (VP) time-series, were recorded every 24 h (daily) and, thus, only used to feed the daily forecasting models.

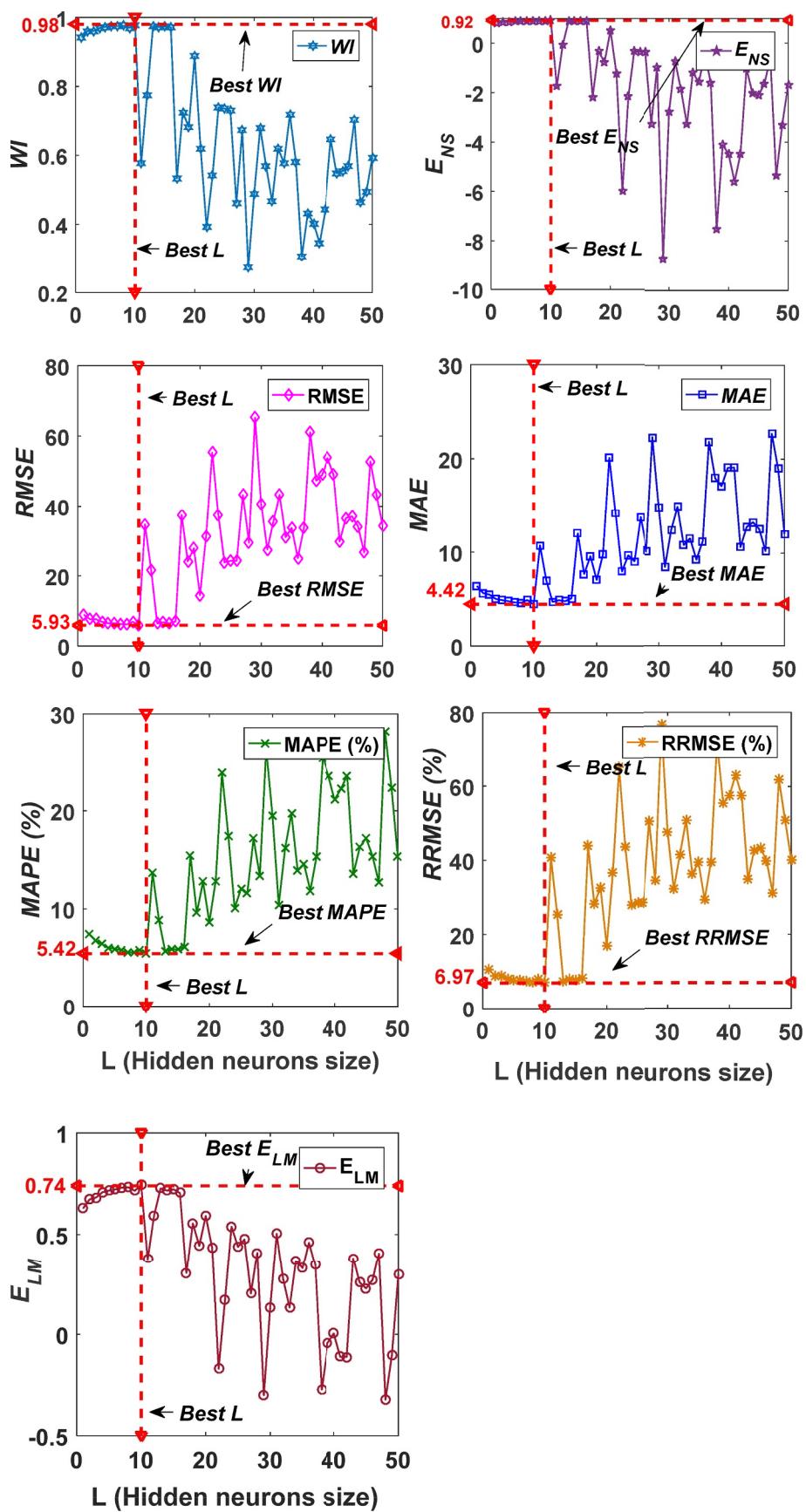


Fig. 5. Selecting the best ANN model in the test period for 6-h Beerwah forecast horizons based on the statistical indicators (Eqs. (7)–(13)). Only the 50 ANN models with the best parameters, indicated in Table 5, are shown.

Table 8

Optimum model performance in the test period for the 6-h forecast horizon as shown by Willmott's index (WI), Nash-Sutcliffe efficiency coefficient (E_{NS}), root-mean square error (RMSE, MW), and mean absolute error (MAE, MW).

Station	ANN			MARS			MLR			ARIMA						
	WI	E_{NS}	RMSE (MW)	MAE (MW)	WI	E_{NS}	RMSE (MW)	MAE (MW)	WI	E_{NS}	RMSE (MW)	MAE (MW)	WI	E_{NS}	RMSE (MW)	MAE (MW)
Beerwah	0.98	0.92	5.93	4.42	0.96	0.88	7.34	5.31	0.95	0.83	8.63	6.39	0.31	-0.06	21.49	16.75
Burpengary	0.98	0.93	8.83	6.27	0.97	0.90	10.90	7.70	0.96	0.85	13.39	9.03	0.39	-0.09	35.58	24.95
Ipswich South	0.98	0.94	7.85	5.89	0.98	0.92	9.10	6.68	0.97	0.89	10.80	7.86	0.35	-0.04	33.22	24.007
Narangba	0.97	0.89	11.10	7.99	0.95	0.81	14.56	10.83	0.93	0.77	16.06	12.07	0.20	-0.02	33.68	27.56
Redcliffe	0.99	0.97	5.03	3.44	0.99	0.96	5.68	4.04	0.99	0.95	6.40	4.58	0.14	0.01	29.34	24.02
Redland Bay	0.98	0.93	6.98	4.88	0.97	0.90	8.69	6.02	0.96	0.87	9.68	6.73	0.41	-0.13	28.99	20.26
Wynnum	0.99	0.95	7.06	5.06	0.98	0.91	9.43	6.70	0.97	0.90	10.09	7.31	0.41	-0.08	32.82	24.97
Zillmere	0.97	0.90	13.35	9.41	0.97	0.88	14.66	10.12	0.96	0.87	15.58	11.02	0.45	-0.19	47.13	35.21

Table 9

The mean absolute percentage error (MAPE, %), relative root-mean square error (RRMSE, %), and Legates & McCabe's Index (E_{LM}) for the optimum models for the 6-h forecast horizon for the testing datasets.

Station	ANN			MARS			MLR			ARIMA		
	MAPE (%)	RRMSE (%)	m	MAPE (%)	RRMSE (%)	E_{LM}	MAPE (%)	RRMSE (%)	E_{LM}	MAPE (%)	RRMSE (%)	E_{LM}
Beerwah	5.42	6.97	0.74	6.37	8.62	0.69	7.57	10.13	0.63	18.93	25.24	0.02
Burpengary	7.44	9.37	0.77	9.21	11.56	0.71	10.20	14.21	0.66	25.16	37.75	0.07
Ipswich South	6.24	7.92	0.77	6.90	9.19	0.74	8.15	10.90	0.69	22.81	33.55	0.07
Narangba	8.06	10.26	0.71	10.93	13.47	0.60	12.13	14.86	0.56	31.07	31.16	-0.008
Redcliffe	2.63	3.88	0.86	3.12	4.39	0.83	3.54	4.94	0.81	19.40	22.65	0.01
Redland Bay	7.06	9.43	0.78	8.84	11.72	0.72	9.96	13.06	0.69	25.62	39.13	0.07
Wynnum	4.45	6.07	0.81	5.86	8.09	0.74	6.41	8.67	0.72	20.29	28.18	0.04
Zillmere	5.31	7.41	0.74	5.65	8.14	0.72	6.14	8.65	0.69	17.76	26.16	0.01

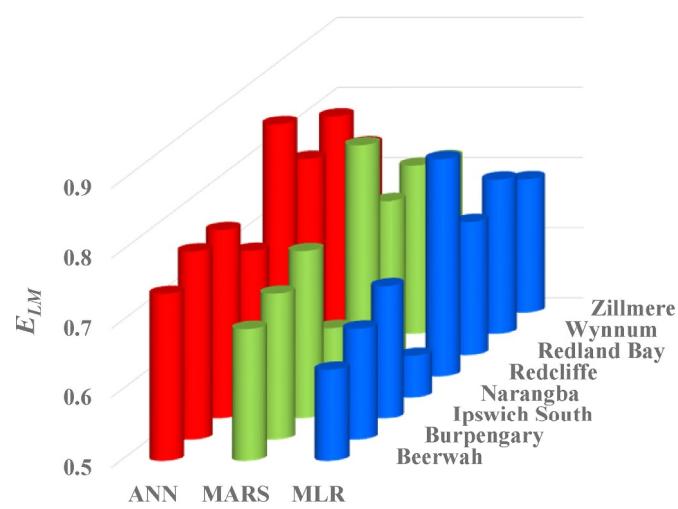


Fig. 6. Bar graphs of the Legates & McCabe's Index (ELM) showing the optimal 6-h models for the eight tested stations in southeast Queensland.

3.3. ECMWF (ERA-Interim) Reanalysis data

A variety of factors, such as longitude, latitude, step, time, date, grid, and variable selection, were carefully considered before downloading 'Interim' Reanalysis data, which are available from 1979 until the current date. The data were generated from atmospheric models and observational data based on the European Centre for Medium-Range Weather Forecasts (ECMWF) and global numerical weather prediction models [57]. The data were recorded every 6 h at 00, 06, 12, and 18 UTS time, which was equivalent to 10:00 a.m., 4:00 p.m., 10:00 p.m., and 4:00 a.m. Brisbane time, respectively. In this study, 51 predictor variables were extracted for the same dates and stations as the G and SILO datasets. These variables were used for the 6-h forecasts, as well as for the daily forecasts (averaging every four data points).

Table 3 shows all the Reanalysis variables, with their abbreviations and averages, using the Beerwah site data as an example. However, as ECMWF data was in a grid format, the locations of the ECMWF (ERA-Interim) Reanalysis data was not fully aligned with the G and SILO data. Therefore, the ECMWF data for four datasets corresponding with the grid corners of the tiles covering each G and SILO station were aggregated and averaged into one dataset based on the two longitudes and two latitudes shown in **Table 1**. The ECMWF (ERA-Interim) grids are shown as the pink boxes in **Fig. 1**.

3.4. Forecast model development

For the ANN, MARS, and MLR models, the historical SILO and ERA-Interim Reanalysis data were used to forecast daily G data, while only the Reanalysis data were used for the 6-h forecasting. The data were split into training (80%) and testing (20%) datasets (**Table 4**) [22]. Three steps were used to build the models in the training period:

1. The higher frequency data from SILO, ECMWF Reanalysis, and G datasets were normalized between zero and one using Equation (7) to avoid large numeric ranges from the values of the predictor variables [58].
2. The best input variables were selected by calculating the cross-correlation between the target (G) and the inputs (SILO and ERA-Interim Reanalysis). The variables that showed statistically significant relationships with the G data, for a 95% prediction interval in lag 1, were used as an input matrix for the models. For example, the daily forecast horizon using the Beerwah datasets is shown in **Fig. 3**. **Table 5** presents the number of selected variables for all the stations for the 6-h and daily forecast horizons.
3. The statistically significant lags were calculated relating the

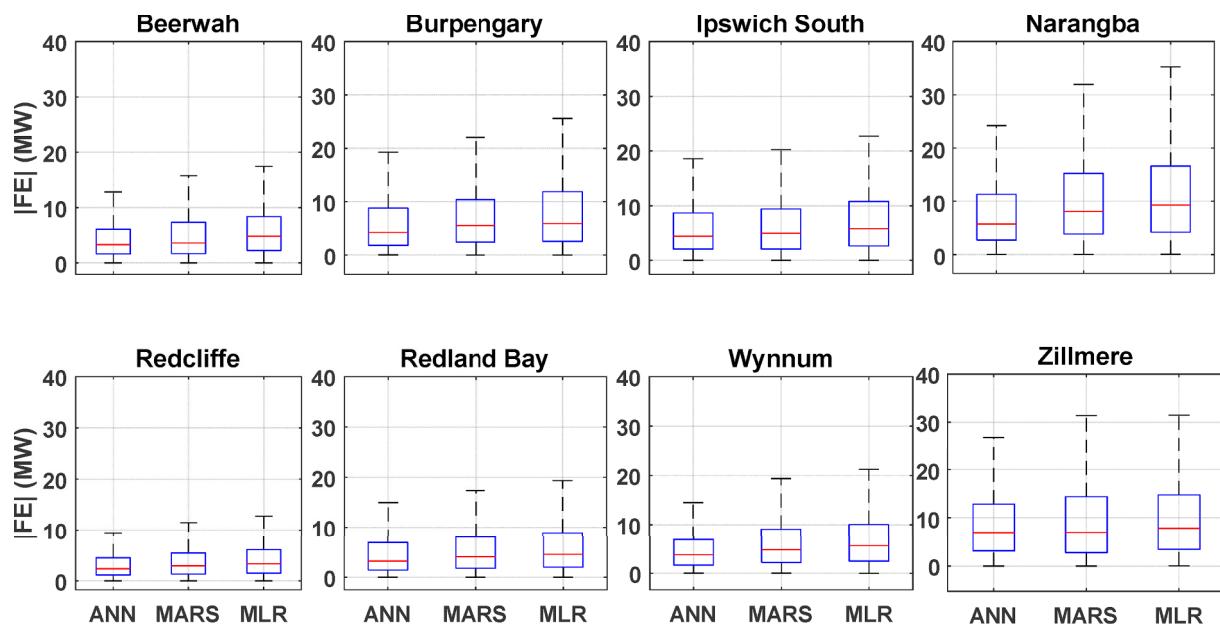


Fig. 7. Boxplots of the 6-h absolute forecasting error, $|FE| = |G_i^{for} - G_i^{obs}|$, in the testing period for the eight stations.

historical G data to the model's input variables through the use of the partial autocorrelation function (PACF). Fig. 4a uses the Beerwah datasets to show the PACF for the daily forecast horizon with the first three statistically significant lags marked. The same technique was applied to the data of the other stations for both forecast horizons. However, for the Burpengary and Ipswich South areas, the 6-h G data showed no or low significance with the first lag, respectively. This helped to identify other highly statistically significant lags, i.e., the second, fourth and fifth lags for Burpengary (Fig. 4b) and the fourth, fifth, and sixth lags, plus the first one, for Ipswich South (Fig. 4c). Table 5 lists the statistically significant lags for both the 6-h and daily forecast horizons using datasets at all the sites.

The main model (ANN) and the comparative models (MARS and MLR) were developed using MATLAB running on an Intel i7 processor at 3.60 GHz. For the ANN algorithm, previous studies (e.g., [31,32]) have indicated that the selection of the training algorithm is an important part of model development. The hidden transfer (F) and output layer ($y(x)$) functions, as well as the optimal hidden neuron size (L), must also be determined. Therefore this study tested two hidden transfer functions (*tansig* and *logsig*) and three output layer functions (*tansig*, *purelin* and *logsig*), as well as L values from 1 to 50. These functions are available in the MATLAB programming environment and are defined below [31,32] in Eqs. (8)–(10). The ANN parameters, which include the training algorithm, hidden transfer function, output layer function, and L were tested one by one, resulting in 600 ANN models with 6-h and daily forecast horizons for each dataset. The optimal ANN model parameters for G forecasting were selected during the testing phase based on the statistical indicators shown in Eqs. (11)–(17). Table 5 describes the structures used for the ANN in the training period.

$$F(x) = \frac{2}{1 + \exp(-2x)} - 1 \text{ Tangent Sigmoid (tan sig)} \quad (8)$$

$$F(x) = \frac{1}{1 + \exp(-x)} \text{ Log Sigmoid (log sig)} \quad (9)$$

$$F(x) = x \text{ Positive Linear (purelin)} \quad (10)$$

The MARS and MLR models were constructed based on the piecewise cubic and linear regression functions, respectively, to evaluate the accuracy of the ANN model [7,8,43,44]. However, in this study, 50

MARS models were developed for each forecast horizon dataset based on the number of the model spline functions (Table 5). The best MARS model was selected based on the statistical metrics (Eqs. (11)–(17)) analysed in the testing period. Additional work, which was not presented previously, showed the importance of estimating the input variables when training the MARS model. This was calculated using the scaled formula of the “square root of the Generalized Cross-Validation (GCV) of the model with all basis functions minus square root of the GCV score of the corresponding full model”, where a value of zero meant that the variable was not used during the training period, while a value close to 100 or zero indicated the variables with the highest and lowest relative importance, respectively [59,60]. Table 5 shows these two variables for the eight stations for both forecasting horizons. The MLR technique was built with the y -intercept utilizing the same input variables shown in Table 5. A single MLR forecast model was constructed in the training phase for each dataset and forecast horizon.

To show the ANN's ability to forecast G as a high-dimensional model, a low-dimensional ARIMA model was built using an R package for automatic time series forecasting [61,62]. In the training phase, the ARIMA model was constructed as a single input model using only G data with no lags, based on the model's three parameters: the auto-regressive (p), differencing (d), and moving average (q). The model was evaluated using the Akaike information criterion (AIC), log likelihood (LH), and variance (σ^2) [63]. Table 5 shows the model development for the 6-h and daily forecasting horizons.

To improve the ability of the ANN model to forecast G , a hybrid ANN model was created by combining the forecasted values of the high-dimensional ANN, MARS, and MLR models. The outputs of the three models were used as inputs to feed the hybrid ANN model. After testing this technique, an accurate groundwater risk map [64] and a forecast of soil moisture [65] were created. Additionally, the B method, a powerful pre-processing forecasting tool, was applied in this research paper to investigate the uncertainty in G forecasting using an ensemble of the hybrid ANN model. The B-hybrid ANN model was then developed using 200 bootstrap samples of the raw datasets in the training period. These 200 constructed hybrid ANN models were used to build the 95% prediction bands from the testing datasets using Eq. (6). Table 6 shows the optimum parameters for the hybrid ANN and B-hybrid ANN models in the training period for both forecasting horizons.

The performance of the models in the training period is summarized in Table 7 based on the root-mean square error (RMSE, Eq. (11)) using

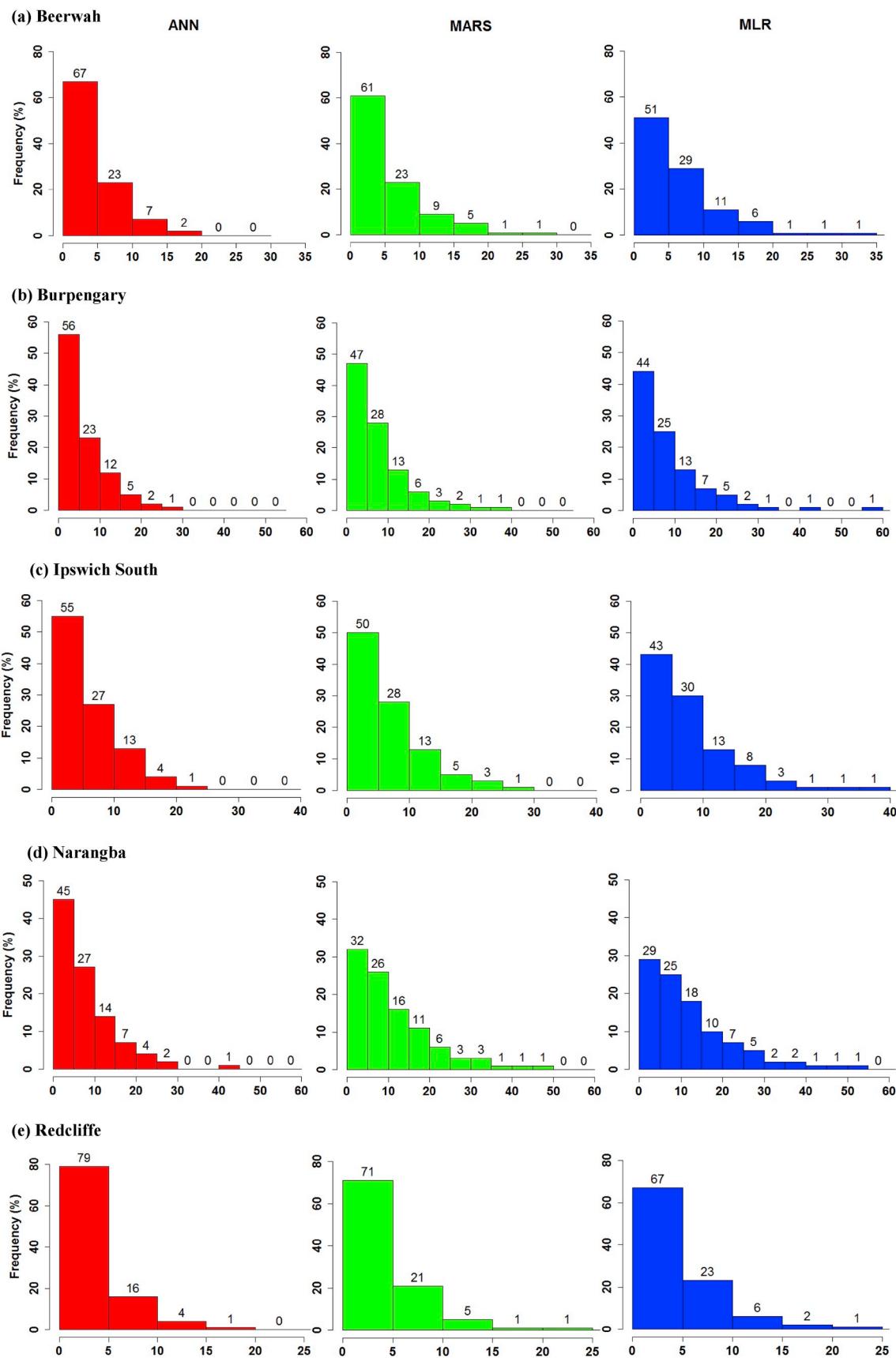


Fig. 8. The percentage frequency distribution of the 6-h forecasted error, $|FE|$, for the eight stations in the testing phase.

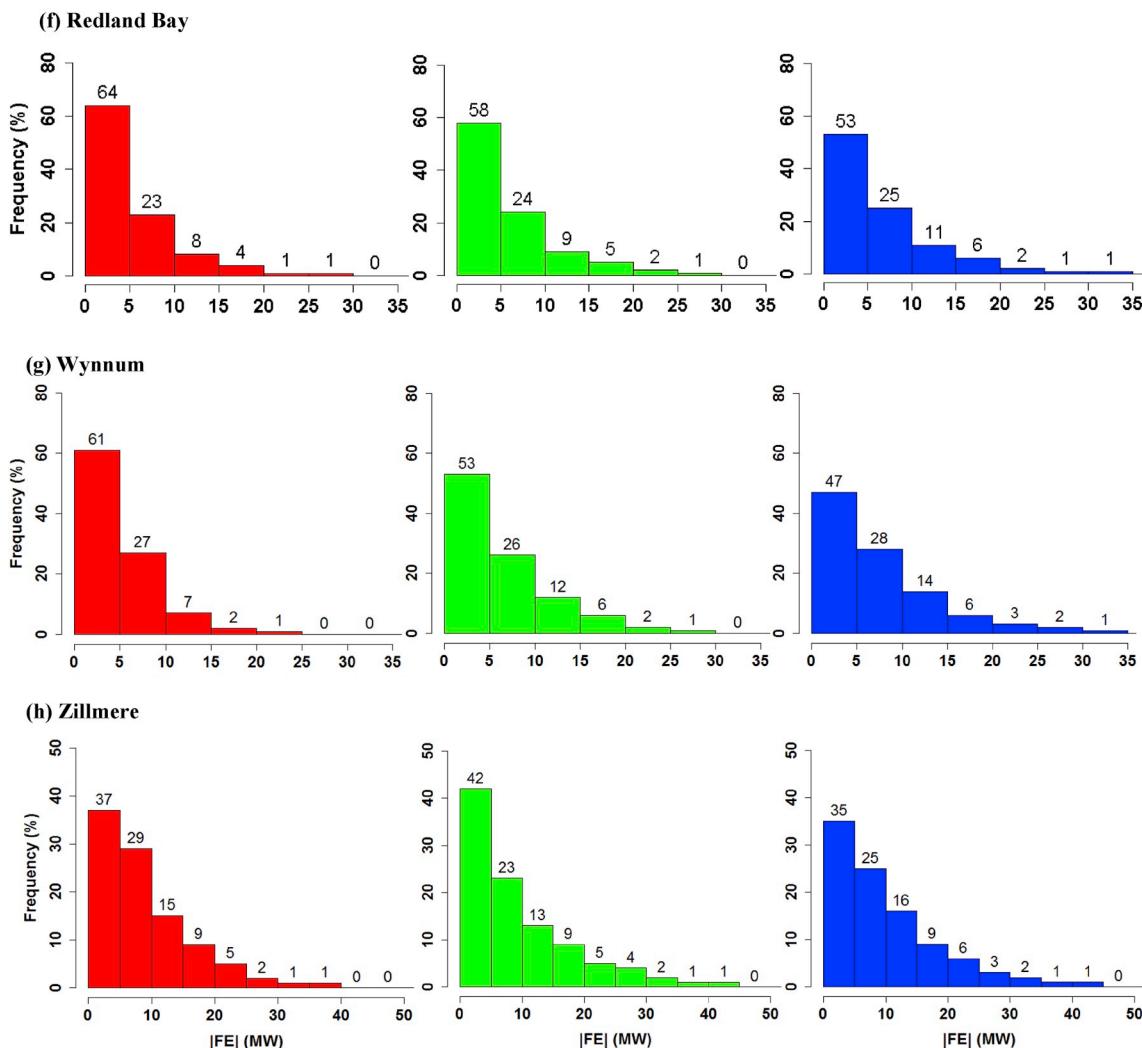


Fig. 8. (continued)

Table 10

Model performance in the test period for the daily forecast horizon by Willmott's index (*WI*), Nash–Sutcliffe efficiency coefficient (*E_{NS}*), root-mean square error (*RMSE*, MW), and mean absolute error (*MAE*, MW).

Station	ANN				MARS				MLR				ARIMA			
	WI	<i>E_{NS}</i>	RMSE (MW)	MAE (MW)	WI	<i>E_{NS}</i>	RMSE (MW)	MAE (MW)	WI	<i>E_{NS}</i>	RMSE (MW)	MAE (MW)	WI	<i>E_{NS}</i>	RMSE (MW)	MAE (MW)
Beerwah	0.89	0.68	28.47	21.20	0.88	0.67	29.20	22.33	0.87	0.65	29.78	22.90	0.44	-0.57	63.19	48.46
Burpengary	0.95	0.81	33.45	24.58	0.93	0.79	35.36	25.26	0.93	0.78	36.49	27.07	0.43	-0.54	96.27	68.19
Ipswich South	0.94	0.80	36.32	27.39	0.93	0.79	37.27	27.49	0.93	0.78	38.24	29.48	0.41	-0.35	94.34	68.79
Narangba	0.89	0.68	51.04	37.68	0.86	0.60	57.38	46.02	0.84	0.58	58.80	48.03	0.44	-0.67	116.70	97.86
Redcliffe	0.96	0.85	25.80	19.59	0.96	0.84	26.02	19.93	0.95	0.83	27.34	20.81	0.35	-0.13	69.90	56.79
Redland Bay	0.92	0.75	28.17	19.50	0.92	0.75	28.29	19.14	0.93	0.78	26.43	19.25	0.41	-0.40	67.21	47.77
Wynnum	0.94	0.80	34.34	25.50	0.92	0.77	36.23	27.04	0.93	0.78	35.98	27.11	0.39	-0.29	86.30	66.22
Zillmere	0.88	0.65	55.83	42.45	0.88	0.63	57.08	44.26	0.88	0.65	55.95	43.88	0.43	-0.97	132.17	103.17

the datasets for all the stations for both forecast horizons. For the 6-h forecasts, the hybrid ANN model yielded the highest accuracy with the lowest *RMSE* for all datasets compared with the ANN, MARS, MLR, and ARIMA models. Similarly, for the daily forecasting, the hybrid ANN model outperformed the ANN, MARS, MLR, and ARIMA models for all stations. Table 7 presents the details of the models' evaluation in the training phase.

3.5. Model prediction quality

Seven statistical metrics, including Legates and McCabe's Index (*E_{LM}*), Willmott's Index (*WI*), and Nash–Sutcliffe efficiency coefficient (*E_{NS}*), as well as mean absolute error (*MAE*) and root-mean square error (*RMSE*) together with their relative error values (*MAPE*, % and *RRMSE*, %), were employed to test the performance of the models for the 6-h and daily *G* forecast horizons in the testing period. The equations for the statistical indicators are listed below [66–75]:

Table 11

The mean absolute percentage error (*MAPE*, %), relative root-mean square error (*RRMSE*, %), and Legates & McCabe's Index (*E_{LM}*) for the optimum models with daily forecast horizon in the test datasets.

Station	ANN			MARS			MLR			ARIMA		
	<i>MAPE</i> (%)	<i>RRMSE</i> (%)	<i>E_{LM}</i>									
Beerwah	6.31	8.36	0.47	6.53	8.57	0.44	6.77	8.74	0.43	13.15	18.55	-0.20
Burpengary	6.38	8.87	0.60	6.51	9.38	0.59	7.04	9.68	0.56	15.83	25.53	-0.11
Ipswich South	6.79	9.17	0.57	6.87	9.41	0.57	7.37	9.65	0.54	15.60	23.82	-0.07
Narangba	9.01	11.80	0.47	11.13	13.26	0.35	11.67	13.59	0.32	21.40	26.98	-0.38
Redcliffe	3.88	4.98	0.65	3.89	5.02	0.65	4.12	5.28	0.63	10.53	13.49	-0.01
Redland Bay	6.19	9.51	0.57	5.99	9.55	0.58	6.20	8.92	0.58	14.31	22.69	-0.05
Wynnum	5.50	7.37	0.59	5.76	7.77	0.57	5.86	7.72	0.57	13.25	18.52	-0.05
Zillmere	5.83	7.75	0.40	6.09	7.92	0.38	6.04	7.77	0.38	13.32	18.35	-0.45

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{i=n} (G_i^{for} - G_i^{obs})^2} \quad (11)$$

$$MAE = \frac{1}{n} \sum_{i=1}^{i=n} |G_i^{for} - G_i^{obs}| \quad (12)$$

$$RRMSE = 100 \times \frac{\sqrt{\frac{1}{n} \sum_{i=1}^{i=n} (G_i^{for} - G_i^{obs})^2}}{G^{obs}} \quad (13)$$

$$MAPE = 100 \times \frac{1}{n} \sum_{i=1}^{i=n} \left| \frac{G_i^{for} - G_i^{obs}}{G_i^{obs}} \right| \quad (14)$$

$$WI = 1 - \left[\frac{\sum_{i=1}^{i=n} (G_i^{for} - G_i^{obs})^2}{\sum_{i=1}^{i=n} (|G_i^{for} - G_i^{obs}| + |G_i^{obs} - G^{obs}|)^2} \right], \text{ and } 0 \leq WI \leq 1 \quad (15)$$

$$E_{NS} = 1 - \left[\frac{\sum_{i=1}^{i=n} (G_i^{for} - G_i^{obs})^2}{\sum_{i=1}^{i=n} (G_i^{obs} - G^{obs})^2} \right], \text{ and } \infty \leq E_{NS} \leq 1 \quad (16)$$

$$E_{LM} = 1 - \left[\frac{\sum_{i=1}^{i=n} |G_i^{obs} - G_i^{for}|}{\sum_{i=1}^{i=n} |G_i^{obs} - G^{obs}|} \right], \text{ and } (\infty \leq E_{LM} \leq 1) \quad (17)$$

where G_i^{for} , G_i^{obs} , $\overline{G^{for}}$ and $\overline{G^{obs}}$ are the forecasted, observed, mean of forecasted, and mean of observed values of G , respectively, and n is the total number of G_i^{for} (or G_i^{obs}) values in the testing period.

Since a single indicator would not be able to show all shortcomings of the models used, multiple statistical criteria were used to assess each model's performance [74]. The best performing model had values of *RMSE* and *MAE* that were closest to zero and values of *WI*, *E_{LM}*, and *E_{NS}* that were closest to one. For comparative analysis, *MAPE* (%) and *RRMSE* (%) were also used to describe the models' behaviour over a range of statistically different hydrological flows [32]. Specifically, a model's performance was considered excellent with *RRMSE* < 10%, good with 10% < *RRMSE* < 20%, fair with 20% < *RRMSE* < 30%, and poor with *RRMSE* > 30% [66,76–78].

4. Results and discussion

4.1. 6-h forecast horizon

The models were assessed using the testing dataset based on the seven statistical criteria shown in Eqs. (11)–(17). The best model parameters were selected according to the lowest *RMSE*, *MAE*, *RRMSE*, and *MAPE* values, as well as the highest *WI*, *E_{NS}*, and *E_{LM}* values. The best overall model was also selected using these criteria. For the ANN, a total 600 models were created for each station's dataset with two types of training algorithms and hidden transfer functions, as well as three types of output layer functions and L between 1 and 50. For the 6-h Beerwah forecast horizon, the best ANN parameters were found to be

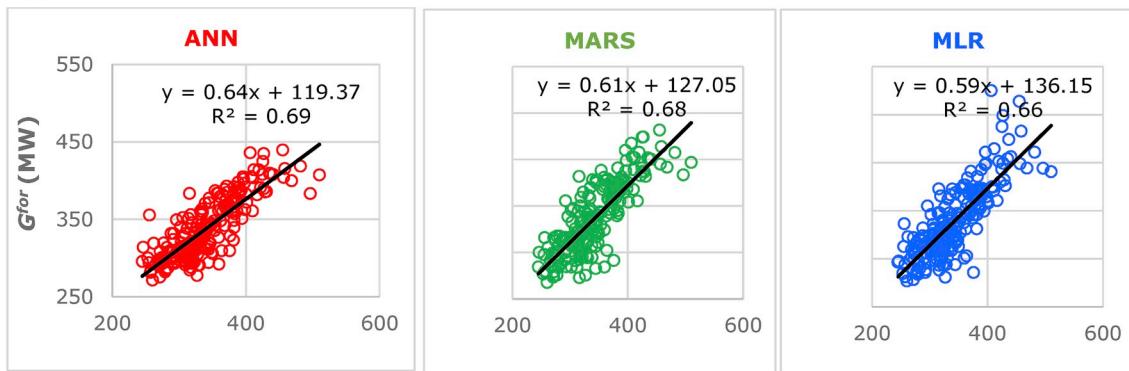
trainlm, *logsig*, *tansig*, and 10 for the training algorithm, hidden transfer function, output layer function, and L , respectively. Table 5 shows these factors in detail for all models, stations, and forecast horizons, while Fig. 5 illustrates the best ANN model with L equal to 10. Similarly, Eqs. (11)–(17) helped to select the best number of spline functions for the MARS model, whereas only one MLR model and one ARIMA model were developed in this study.

According to the results, the low-dimensional ARIMA model had the lowest accuracy (Tables 8 and 9), in contrast to the high-dimensional ANN, MARS, and MLR models, which likely performed better because of their use of the extensive data from the SILO and ERA-Interim Reanalysis datasets. When comparing the three high-dimensional models developed in this study, the ANN model outperformed the MARS and MLR models for all eight datasets (Table 8). For example, for the Beerwah dataset, *WI* = 0.98, *E_{NS}* = 0.92, *RMSE* = 5.93 MW, and *MAE* = 4.42 MW (for ANN) compared with *WI* = 0.96, *E_{NS}* = 0.88, *RMSE* = 7.34 MW, and *MAE* = 5.31 MW (for MARS) and *WI* = 0.95, *E_{NS}* = 0.83, *RMSE* = 8.63 MW, and *MAE* = 6.93 MW (for MLR). Similar results were obtained for all other stations indicating that the ANN model had the highest accuracy when forecasting G .

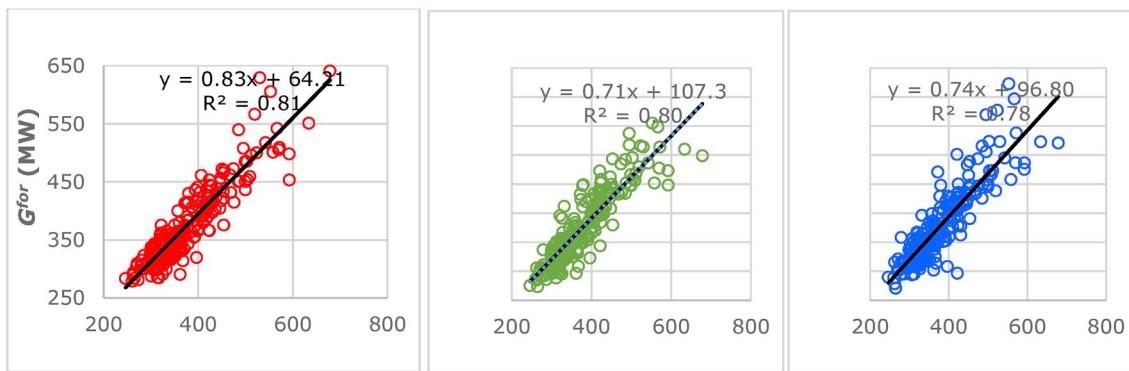
The performance of the ANN model was further evaluated using three other metrics, namely *MAPE*, *RRMSE*, and *E_{LM}* (Table 9). These statistical indicators were extracted from *MAE*, *RMSE*, and *WI* to calculate the 'goodness-of-fit' between the G_i^{for} and G_i^{obs} values (e.g., [7,8]). For all the forecasting datasets, the ANN model outperformed the MARS and MLR models. For the Beerwah dataset, the ANN model had an *MAPE* = 5.42%, *RRMSE* = 6.97%, and *E_{LM}* = 0.74, while the MARS model had an *MAPE* = 6.37%, *RRMSE* = 8.62%, and *E_{LM}* = 0.69 and the MLR model had an *MAPE* = 7.57%, *RRMSE* = 10.13% and *E_{LM}* = 0.63. Table 9 shows these results for all stations, while Fig. 6 shows bar graphs of *E_{LM}* metric for the three forecasting models and all stations.

Boxplots and histograms were plotted for the absolute forecasted error statistics, $|FE| = |G_i^{for} - G_i^{obs}|$, for the 6-h forecasts of the ANN, MARS, and MLR models to examine the forecasting datasets. The boxplots (Fig. 7) display the lower (first line), median (second line), and upper (third line) quartiles of the $|FE|$. Based on those quartiles, larger error values were evident for the MARS and MLR models, in comparison to the ANN. The histograms (Fig. 8) show the percentage of error frequency distributions for the $|FE|$ for all datasets. For the first seven stations, the ANN model had the least error, which was indicated by the highest percentage of error frequency in the smallest (0 to ± 5 MW) bracket for the 6-h G^{for} . Although the MARS model achieved better results than the ANN model (i.e., 42% vs. 37% of errors in the 0 to ± 5 MW bracket, respectively) for the Zillmere station forecasting dataset, the ANN model outperformed the MARS and MLR models in the 0 to ± 10 MW bracket (i.e., 66% vs. 65% vs. 60%, respectively). Generally, the lower prediction error for the ANN model forecasts demonstrated that this model had better accuracy for G^{for} than the MARS and MLR models.

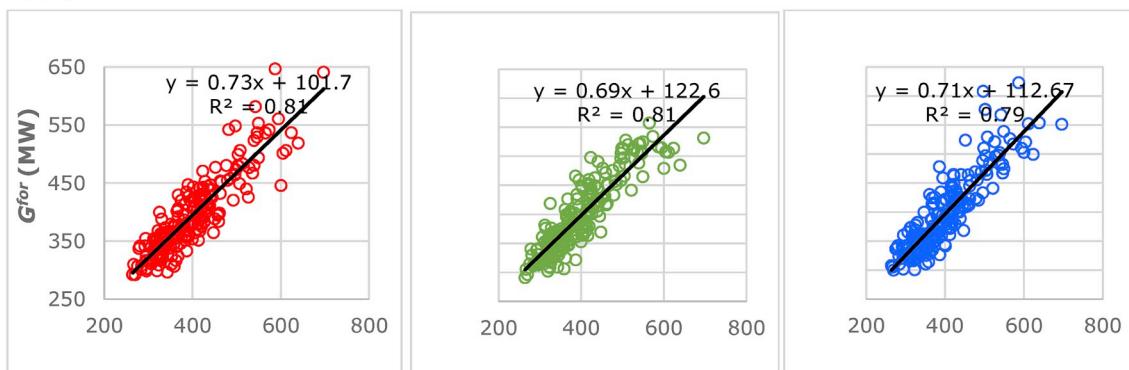
(a) Beerwah



(b) Burpengary



(c) Ipswich South



(d) Narangba

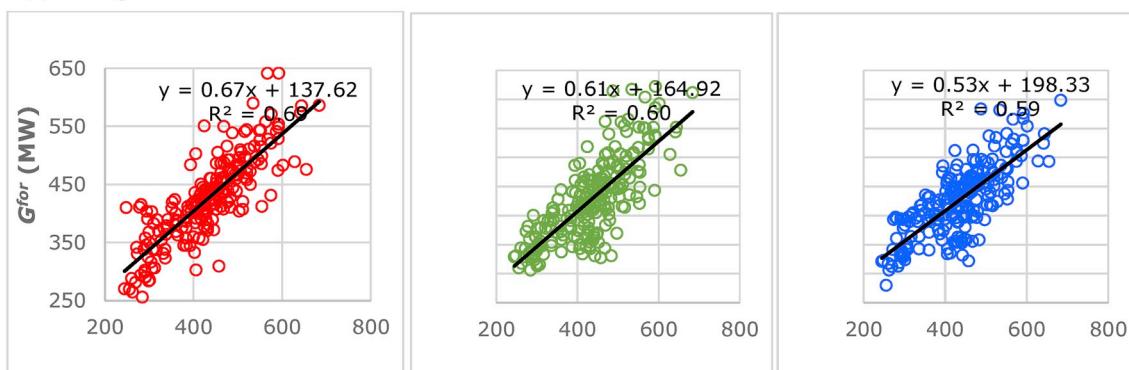


Fig. 9. Scatterplot of the daily forecasted electricity demand (G^{for}) vs. observed electricity demand (G^{obs}) in the testing phase for the eight stations.

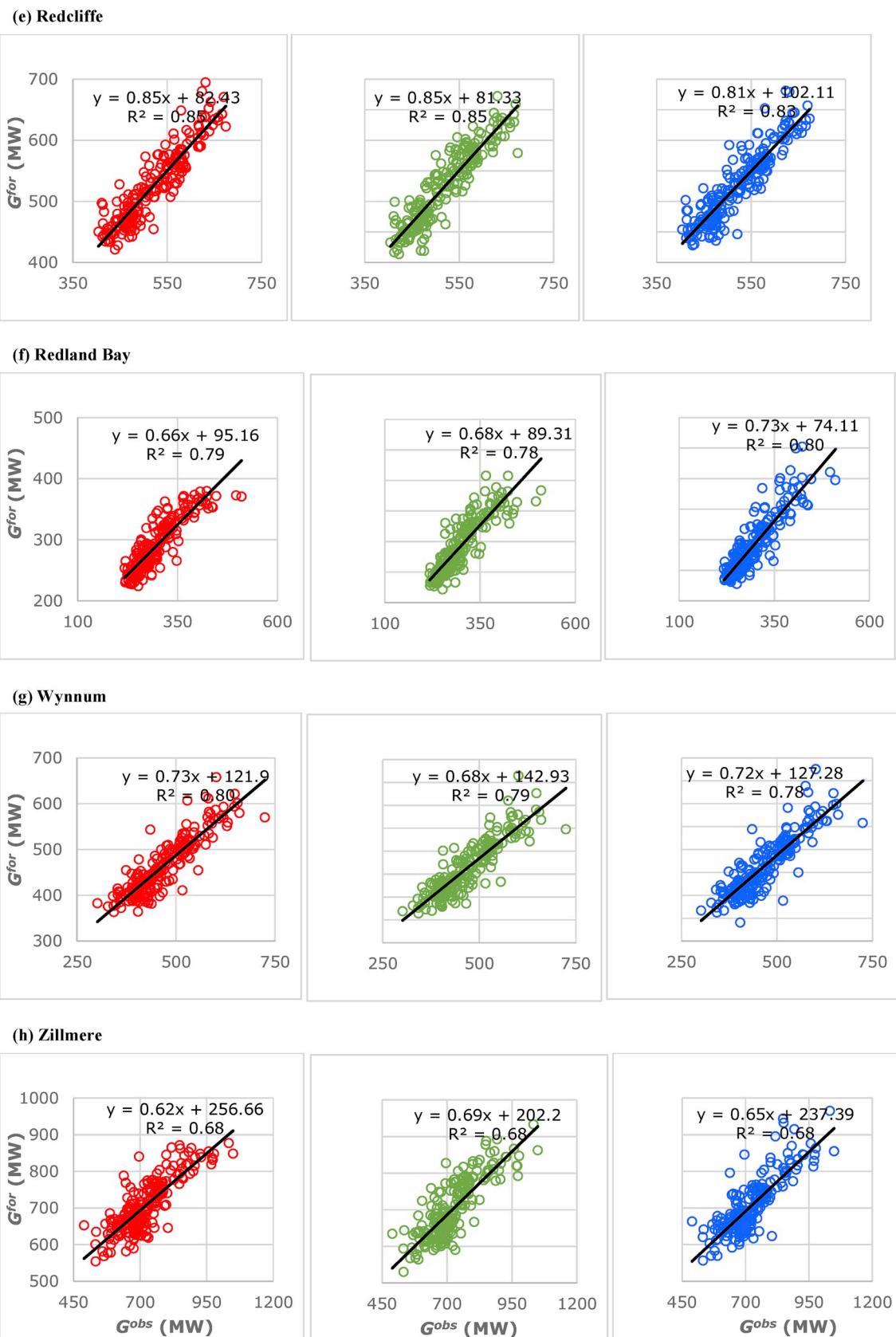


Fig. 9. (continued)

4.2. Daily forecast horizon

The ability to forecast G over a longer forecasting horizon was

evaluated for the ANN, MARS, and MLR models, using a daily time step. Again, the ARIMA model underperformed compared to the other models for daily G forecasting. Tables 10 and 11 summarize the

Table 12

Optimum hybrid ANN model performance in the test period for both forecast horizons by Willmott's index (*WI*), Nash–Sutcliffe efficiency coefficient (E_{NS}), root-mean square error ($RMSE$, MW), mean absolute error (MAE , MW), mean absolute percentage error ($MAPE$, %), relative root-mean square error ($RRMSE$, %), and Legates & McCabe's Index (E_{LM}).

Station	daily													
	6-h	<i>WI</i>	E_{NS}	$RMSE$ (MW)	MAE (MW)	$MAPE$ (%)	$RRMSE$ (%)	E_{LM}	<i>WI</i>	E_{NS}	$RMSE$ (MW)	MAE (MW)	$MAPE$ (%)	$RRMSE$ (%)
Beerwah	0.98	0.92	5.84	4.29	5.29	6.86	0.75	0.90	0.69	28.05	20.69	6.12	8.23	0.49
Burpengary	0.98	0.94	8.65	6.14	7.34	9.18	0.77	0.95	0.83	32.20	23.85	6.20	8.54	0.61
Ipswich South	0.99	0.95	7.63	5.76	6.13	7.71	0.78	0.94	0.81	35.18	25.90	6.45	8.88	0.60
Narangba	0.97	0.89	10.96	7.85	7.94	10.14	0.71	0.90	0.70	49.44	37.03	8.93	11.43	0.48
Redcliffe	0.99	0.97	4.99	3.39	2.57	3.85	0.86	0.96	0.86	24.48	18.46	3.63	4.73	0.67
Redland Bay	0.98	0.94	6.85	4.80	6.95	9.25	0.78	0.94	0.79	26.04	18.05	5.78	8.79	0.60
Wynnum	0.99	0.95	7.03	5.03	4.40	6.04	0.81	0.93	0.80	34.33	25.25	5.39	7.37	0.60
Zillmere	0.98	0.91	13.04	9.21	5.22	7.24	0.74	0.90	0.68	53.09	41.48	5.75	7.37	0.42

performance of each model. For most of the stations' datasets, the ANN model only achieved the highest forecasting accuracy by a small margin, in comparison to the MARS and MLR models. However, for the Narangba station dataset the ANN model significantly outperformed the other models, with $WI = 0.89$, $E_{NS} = 0.68$, $RMSE = 51.04$ MW, $MAE = 37.68$ MW, $MAPE = 9.01\%$, $RRMSE = 11.80\%$, and $E_{LM} = 0.47$, in comparison to the MARS model with $WI = 0.86$, $E_{NS} = 0.60$, $RMSE = 57.38$ MW, $MAE = 46.02$ MW, $MAPE = 11.13\%$, $RRMSE = 13.26\%$, and $E_{LM} = 0.35$, and the MLR model with $WI = 0.84$, $E_{NS} = 0.58$, $RMSE = 58.80$ MW, $MAE = 48.03$ MW, $MAPE = 11.67\%$, $RRMSE\% = 13.59$, and $E_{LM} = 0.32$. Overall, considering the seven statistical indicators used in this study, better G^{for} was yielded by the ANN model when compared to the MARS and MLR models.

Fig. 9 shows scatterplots with the least squares regression line, $[y(G_i^{for}) = aG_i^{obs} + b]$, and correlation of determination (R^2) used to evaluate the relationship between G_i^{for} and G_i^{obs} for all the G^{for} station datasets. For six stations' datasets (Beerwah, Burpengary, Ipswich South, Narangba, Redcliffe, and Wynnum), based on the values of a , b , and R^2 , the ANN model obtained better results than the MARS and MLR models, especially for forecasting data at the Narangba station. The outcomes were $a = 0.67$, $b = 137.62$, and $R^2 = 0.69$ (for ANN), $a = 0.61$, $b = 164.92$, and $R^2 = 0.60$ (for MARS), and $a = 0.53$, $b = 198.33$, and $R^2 = 0.59$ (for MLR). However, for Zillmere and Redland Bay stations, MARS and MLR outperformed the ANN by a small margin.

4.3. Hybrid ANN and uncertainty assessment using a B-hybrid ANN

To further enhance the accuracy of the ANN model, a hybrid ANN model was developed using the forecast values of the ANN, MARS, and MLR models as inputs. Table 12 shows the performance of the hybrid ANN model for the 6-h and daily forecasting horizons during the testing period. The daily forecasts at the Redland Bay site showed the most significant improvement of hybrid ANN ($WI = 0.94$, $E_{NS} = 0.79$, $RMSE = 26.04$ MW, $MAE = 18.05$ MW, $MAPE = 5.78\%$, $RRMSE = 8.79\%$, and $E_{LM} = 0.60$) in comparison to the standard ANN ($WI = 0.92$, $E_{NS} = 0.75$, $RMSE = 28.17$ MW, $MAE = 19.50$ MW, $MAPE = 6.19\%$, $RRMSE = 9.51\%$, and $E_{LM} = 0.57$). Consequently, G was found to be most accurately forecasted using the hybrid technique, which could support the National Electricity Market in Queensland, Australia, and potentially other areas as well.

To estimate the forecasting bands for the 6-h and daily horizons, ensemble-based uncertainty assessments via the bootstrapping procedure (B) was used in this paper in combination with the hybrid ANN model. The B-hybrid ANN model was developed using 200 B samples with replacement. The 95% prediction bands were built using Eq. (6) based on the averages and variances of the 200 forecasted values in the test datasets. As shown in Fig. 10, the lower and upper forecasted bands are close to the observed values. This demonstrates the ability and

reliability of the hybrid ANN model to forecast G by addressing some of the uncertainty associated with forecasted values.

5. Further discussion, limitations and opportunity for future work

Developing a globally trained model for electricity demand forecasting was identified as an important component for supporting policy development related to energy technology by the United Nations Development Program on sustainable energy. Recently, national electricity markets globally have reported high forecasting errors due to the use of old techniques to predict G , which has encouraged researchers to develop more accurate forecasting strategies. Since G usage is expected to increase (e.g., [2,79,80]), there is a need to develop new methods to accurately forecast G to support the transition to sustainable energy resources, as stated in Goal 7 of the UN SDGs. The aim of this study was to support the National Electricity Market in Australia, and in other areas, by constructing a highly accurate G forecasting model with an ANN algorithm using a wide range of datasets extracted from three sources: SILO and ECMWF (ERA-Interim) for the input variables, and Energex for the target variable (G). The forecasting performance of the developed ANN model was compared with the MARS, MLR, and ARIMA models based on several statistical indicators using data from southeast Queensland, Australia. Furthermore, a hybrid ANN model was designed and compared with the standard ANN model, while the PI was calculated using the B method integrated with the hybrid model to evaluate the uncertainty assessments obtained from the forecast values. This study found that the hybrid ANN model performed well and generally outperformed the ANN, MARS, MLR, and ARIMA models in forecasting G .

This study was the first to use extensive datasets from SILO and ECMWF (ERA-Interim) Reanalysis to develop an accurate ANN model for G forecasting. However, some challenges should be addressed in future works. Since the SILO and ERA-Interim Reanalysis data were not available for every single Energex station in the study area, future research should consider using satellite data. For example, satellite data can be extracted from NASA's Moderate Resolution Imaging Spectroradiometer (MODIS) and utilized as an alternative predictor dataset to develop subsequent G forecasting models. Hourly variables over a $0.05^\circ \times 0.05^\circ$ grid resolution can be generated from this source with two primary sensors (Terra and Aqua MODIS) [31,81–83]. Satellite data have successfully been used to forecast solar radiation in the past [31,57,84], which resulted in high-quality ANN and self-adaptive differential evolutionary extreme learning machines (SaDE-ELM) models. Therefore, further research could investigate creating G forecasting models by combining datasets from alternate sources, such as SILO, Reanalysis and/or satellite data.

Another challenge with the current methodology is that non-climatic factors were not considered in this study. Since energy demand is a multivariate problem, social and population variables may need to be taken into account to further improve the accuracy of G forecasts.

(a)

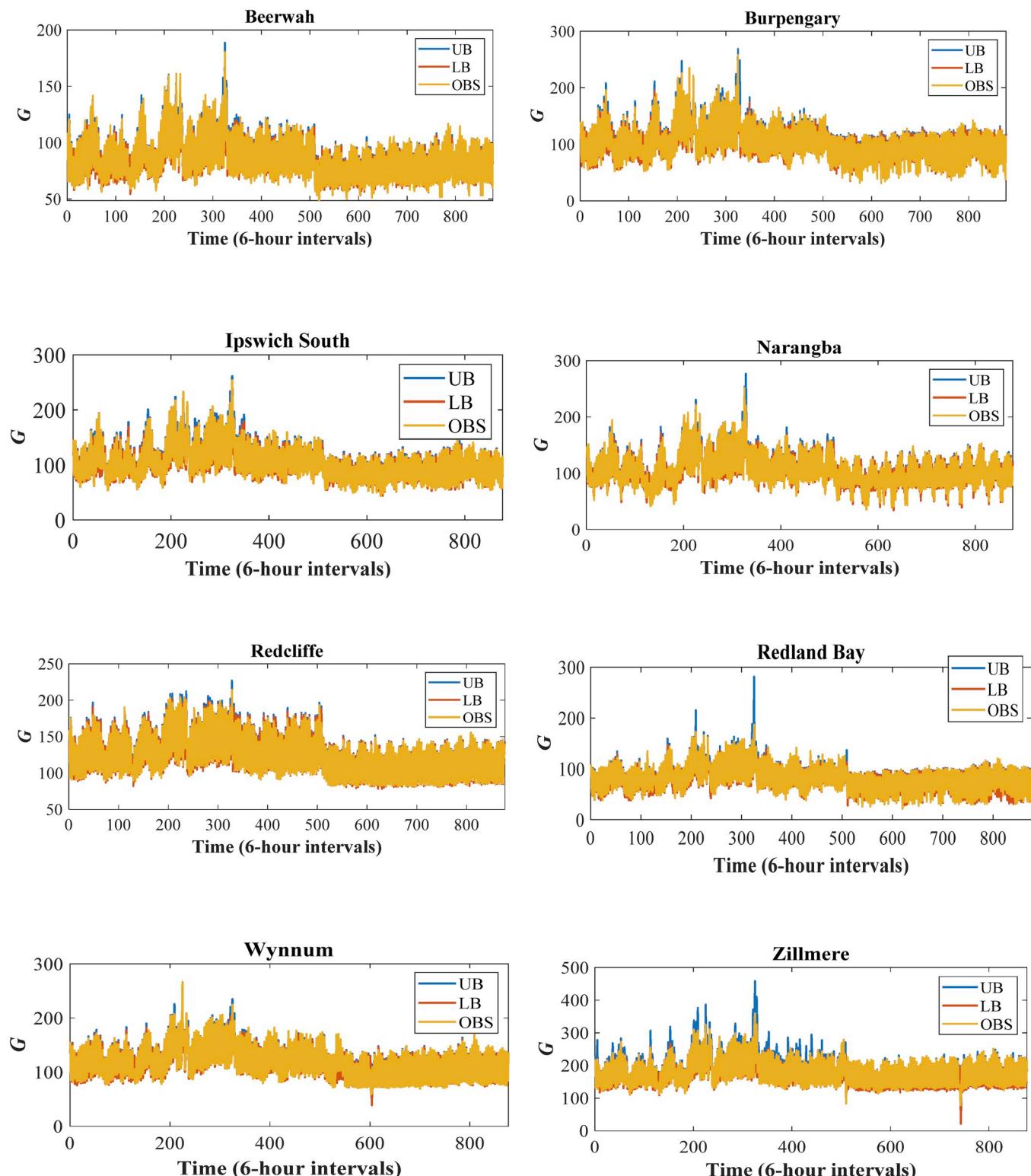


Fig. 10. Electricity demand (G) observed with 95% prediction bands forecasted using a B-hybrid-ANN for all the eight stations for the (a) 6-h and (b) daily forecasting horizons.

However, this may be addressed in an upcoming work on deep learning or long-short term memory network methods (e.g., [85]) using real-life energy predictor variables, for instance, population, that can be extracted from the Australian Bureau of Statistics [86].

Different pre-processing techniques could also be used to improve the forecasting performance of the ANN model. Firstly, a suitable feature selection method, such as iterative input selection (IIS) [87], grouping genetic algorithm (GGA) [88], or coral reef optimisation

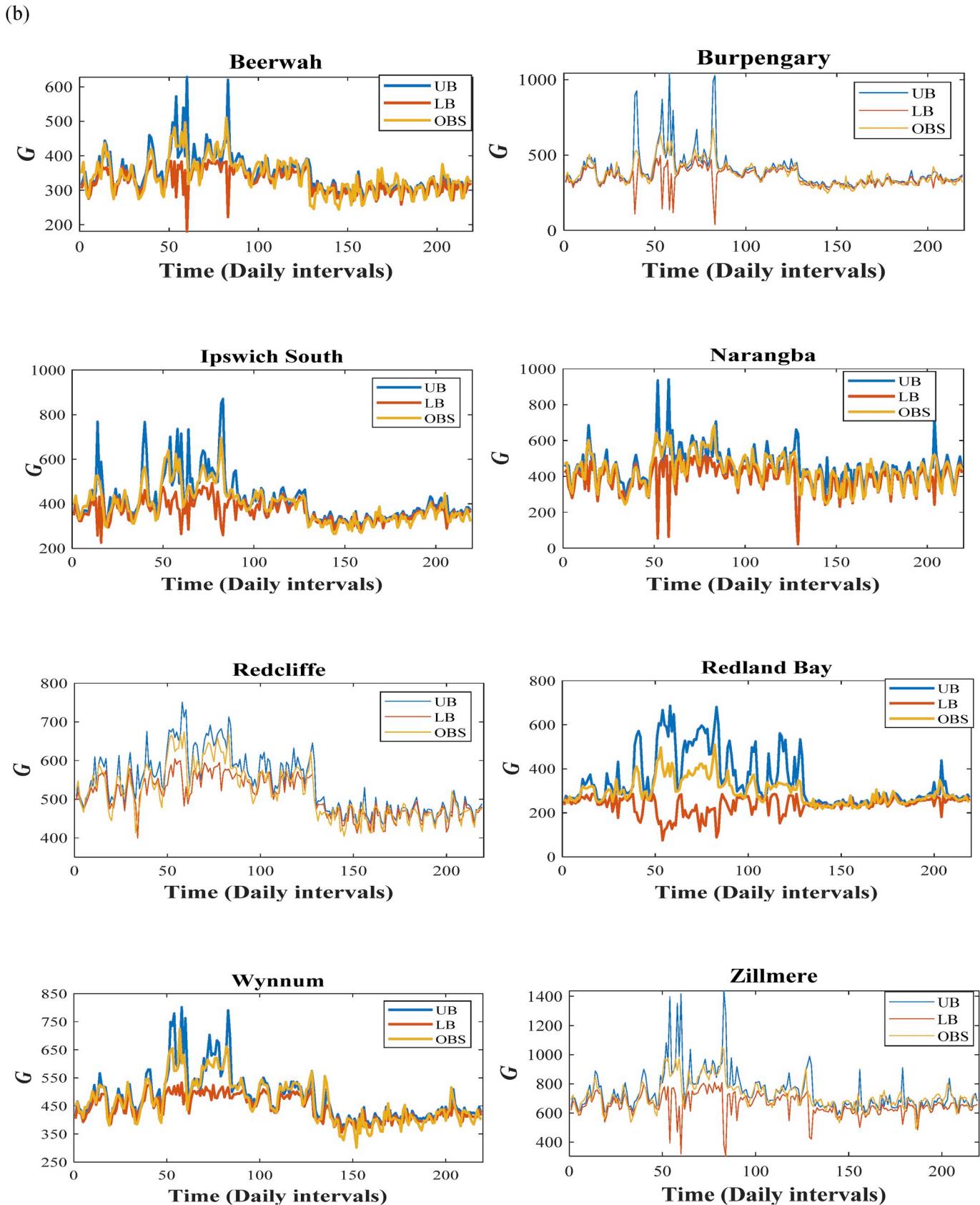


Fig. 10. (continued)

(CRO) [89,90] could be applied with a deep learning strategy or long-short term memory network to select the input variables that have significant influence on the model. This would help reduce both the size

of the predictor datasets used to forecast electricity demand and model complexity. Secondly, wavelet transformations (WT) (e.g., [22,32,35,91]) could be used to address non-stationary data, which

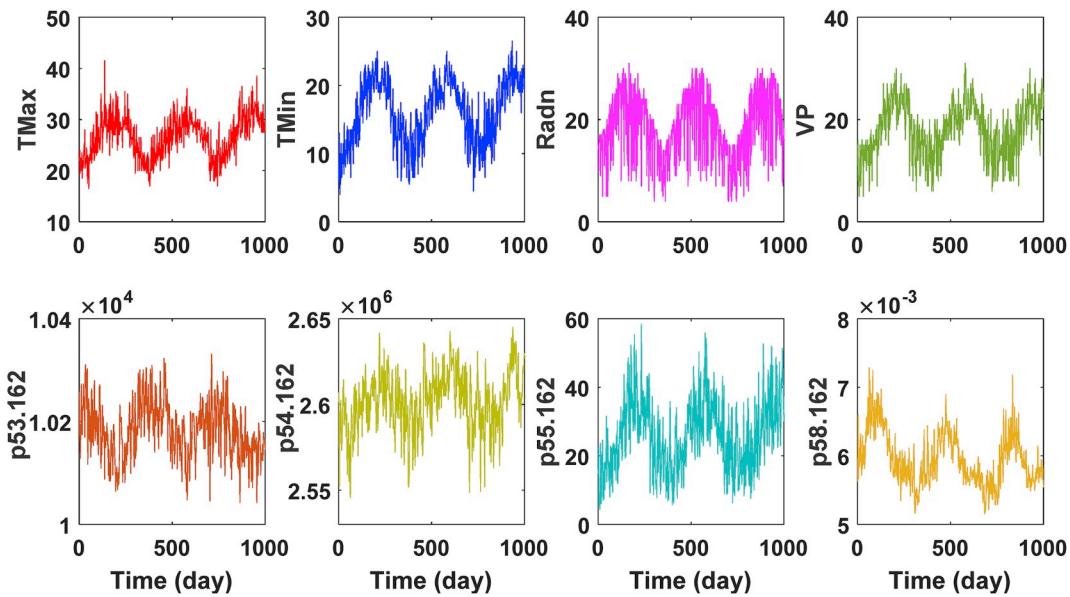


Fig. 11. Non-stationary daily time-series of the eight variables from SILO and ECMWF (ERA-Interim) Reanalysis fields.

appeared in the current SILO and ERA-Interim Reanalysis inputs variables (e.g., Fig. 11), by decomposing the data into high and low-pass filters [22]. However, Quilty and Adamowski [92] stated that WT have been incorrectly used in several research studies to forecast real-world datasets, as they created errors in the inputs of a forecasting model during wavelet decomposition, which also should be addressed in future research. Furthermore, a bootstrapping algorithm (B) or a Bayesian model averaging (BMA) technique could be used with WT and one of the feature selection techniques, such as IIS, to further improve G forecasting accuracy (e.g., [32,37,93]). Additionally, the G forecasting accuracy could be enhanced by using the Firefly algorithm (FA), which has been tested with backpropagation (BP) [94] to forecast electricity prices, to optimize the parameters of the ANN model. Hence, based on the methods suggested above, a hybrid model could be constructed as IIS-WT-B-FA-ANN or IIS-WT-BMA-FA-ANN to forecast G data.

6. Concluding remarks

In this paper, the normal and hybrid ANN models were introduced for 6-h and daily G forecasting horizons using datasets for southeast Queensland, Australia, extracted from Energex, SILO, and ECMWF. Three other models (MARS, MLR, and ARIMA) were also employed in this work to allow for a comparative evaluation of the forecasting accuracy of the ANN and hybrid ANN models. The performance accuracy of the models was evaluated using seven statistical metrics as presented in section 3.5. The results concluded that the Hybrid ANN model achieved the best forecasting accuracy in comparison to other models developed in this study.

In spite of the advantages of the ANN model for G forecasting, the ANN forecasting performance could be improved in future studies by integrating an ANN with different techniques, such as IIS for feature selection, WT-B or WT-BMA for data decomposition and uncertainty assessment, and FA for the ANN optimization problem in order to construct a hybrid model of IIS-WT-B-FA-ANN or IIS-WT-BMA-FA-ANN. Follow-up research studies could also explore the use of shrinkage methods, such as Lasso and elastic-net, including Lasso via hierarchical interactions with the ANN model to potentially further improve forecasting accuracy.

In summary, the ANN model reliably achieved an acceptable G forecasting accuracy and could be used to support the National Electricity Markets, such as Energex. In particular, this study found that the ANN model generally outperformed the MARS, MLR, and ARIMA

models. Furthermore, the hybrid ANN model outperformed the standard ANN. However, the limitations described above should be carefully addressed in future research works.

Acknowledgments

We would like to thank Energex, Scientific Information for Land Owners (SILO) and the European Centre for Medium-Range Weather Forecasts (ECMWF) for providing the required data. We are also grateful to the Ministry of Higher Education and Scientific Research in the Government of Iraq for funding the first author's Ph.D. project. Furthermore, thank you to the University of Southern Queensland for providing access to different academic services. Special thanks to Dr. Barbara Harmes (Language Centre, Open Access College, University of Southern Queensland, Australia), Ms. Mela Aryal (Ph.D. student, University of Southern Queensland, Australia), and Ms. Laura Gilbert (Ph.D. student in Natural Resource Sciences, McGill University, Canada) for providing help in proof-reading this paper. We also would like to thank Dr. Mukesh Tiwari (Department of Irrigation & Drainage Engineering, College of Agricultural Engineering and Technology, Anand Agricultural University, Gujarat, India) for his advice on the bootstrap (B) technique. Finally, we thank both reviewers and the Editor-In-Chief for their comments that have improved this paper.

References

- [1] Assembly UNG. Resolution adopted by the general assembly on 25 september 2015. Washington, United Nations; 2015.
- [2] Smith M, et al. Energy transformed: sustainable energy solutions for climate change mitigation. The Natural Edge Project, CSIRO, and Griffith University; 2007.
- [3] Brinsmead TS, et al. Australian electricity market analysis report to 2020 and 2030 2014. CSIRO Report No. EP141067.
- [4] Pai P-F, Hong W-C. Forecasting regional electricity load based on recurrent support vector machines with genetic algorithms. *Electr Power Syst Res* 2005;74(3):417–25.
- [5] AEMO. Aggregated Price and Demand Data - Historical. The Australian Energy Market Operator is responsible for operating Australia's largest gas and electricity markets and power systems]. 2018 <http://www.aemo.com.au/>.
- [6] Schimek, M.G. and B.A. Turlach, *Additive and generalized additive models: a survey*. 1998, discussion papers, interdisciplinary research project vol. 373: Quantification.
- [7] Al-Musaylh MS, et al. Short-term electricity demand forecasting with MARS, SVR and ARIMA models using aggregated demand data in Queensland, Australia. *Adv Eng Inf* 2018;35:1–16.
- [8] Al-Musaylh MS, et al. Two-phase particle swarm optimized-support vector regression hybrid model integrated with improved empirical mode decomposition with adaptive noise for multiple-horizon electricity demand forecasting. *Appl Energy* 2018;217:422–39.

- [9] Taylor JW. Short-term electricity demand forecasting using double seasonal exponential smoothing. *J Oper Res Soc* 2003;54(8):799–805.
- [10] Taylor JW. Triple seasonal methods for short-term electricity demand forecasting. *Eur J Oper Res* 2010;204(1):139–52.
- [11] Pang Y, et al. Hierarchical electricity demand forecasting by exploring the electricity consumption patterns. *ICPRAM*. 2018, 576–581.
- [12] Fiot J-B, Dinuzzo F. Electricity demand forecasting by multi-task learning. *IEEE Trans on Smart Grid* 2018;9(2):544–51.
- [13] Yang J, Rivard H, Zmeureanu R. On-line building energy prediction using adaptive artificial neural networks. *Energy Build* 2005;37(12):1250–9.
- [14] Yang J, Rivard H, Zmeureanu R. Building energy prediction with adaptive artificial neural networks. Ninth international IBPSA conference. Canada: Montréal; 2005. August.
- [15] Szkuta B, Sanabria LA, Dillon TS. Electricity price short-term forecasting using artificial neural networks. *IEEE Trans Power Syst* 1999;14(3):851–7.
- [16] Kandanamond K. Forecasting electricity demand in Thailand with an artificial neural network approach. *Energies* 2011;4(8):1246–57.
- [17] Ghiassi M, Zimbra DK, Saidane H. Medium term system load forecasting with a dynamic artificial neural network model. *Electr Power Syst Res* 2006;76(5):302–16.
- [18] Mirasgedis S, et al. Models for mid-term electricity demand forecasting incorporating weather influences. *Energy* 2006;31(2–3):208–27.
- [19] Lebotsa ME, et al. Short term electricity demand forecasting using partially linear additive quantile regression with an application to the unit commitment problem. *Appl Energy* 2018;222:104–18.
- [20] Mohanad SA-M, Ravinesh CD, Yan L. Particle swarm optimized-support vector regression hybrid model for daily horizon electricity demand forecasting using climate dataset. E3S web of conferences. EDP Sciences; 2018.
- [21] Deo R,C, Sahin M. Forecasting long-term global solar radiation with an ANN algorithm coupled with satellite-derived (MODIS) land surface temperature (LST) for regional locations in Queensland. *Renew Sustain Energy Rev* 2017;72:828–48.
- [22] Deo RC, Wen X, Qi F. A wavelet-coupled support vector machine model for forecasting global incident solar radiation using limited meteorological dataset. *Appl Energy* 2016;168:568–93.
- [23] Deo RC, Şahin M. Application of the artificial neural network model for prediction of monthly standardized precipitation and evapotranspiration index using hydro-meteorological parameters and climate indices in eastern Australia. *Atmos Res* 2015;161:65–81.
- [24] Acharya N, et al. Development of an artificial neural network based multi-model ensemble to estimate the northeast monsoon rainfall over south peninsular India: an application of extreme learning machine. *Clim Dyn* 2014;43(5–6):1303–10.
- [25] Şahin M. Modelling of air temperature using remote sensing and artificial neural network in Turkey. *Adv Space Res* 2012;50(7):973–85.
- [26] Abbot J, Marohasy J. Input selection and optimisation for monthly rainfall forecasting in Queensland, Australia, using artificial neural networks. *Atmos Res* 2014;138:166–78.
- [27] Hudson D, et al. Bridging the gap between weather and seasonal forecasting: intraseasonal forecasting for Australia. *Q J R Meteorol Soc* 2011;137(656):673–89.
- [28] Abbot J, Marohasy J. Application of artificial neural networks to rainfall forecasting in Queensland, Australia. *Adv Atmos Sci* 2012;29(4):717–30.
- [29] Ortiz-García E, Salcedo-Sanz S, Casanova-Mateo C. Accurate precipitation prediction with support vector classifiers: a study including novel predictive variables and observational data. *Atmos Res* 2014;139:128–36.
- [30] Hamzâcibî C. Forecasting of Turkey's net electricity energy consumption on sectoral bases. *Energy Policy* 2007;35(3):2009–16.
- [31] Deo RC, Şahin M. Forecasting long-term global solar radiation with an ANN algorithm coupled with satellite-derived (MODIS) land surface temperature (LST) for regional locations in Queensland. *Renew Sustain Energy Rev* 2017;72:828–48.
- [32] Prasad R, et al. Input selection and performance optimization of ANN-based streamflow forecasts in the drought-prone Murray Darling Basin region using IIS and MODWT algorithm. *Atmos Res* 2017;197:42–63.
- [33] Şahin M, Kaya Y, Uyar M. Comparison of ANN and MLR models for estimating solar radiation in Turkey using NOAA/AVHRR data. *Adv Space Res* 2013;51(5):891–904.
- [34] Nastos P, et al. Artificial neural networks modeling for forecasting the maximum daily total precipitation at Athens, Greece. *Atmos Res* 2014;144:141–50.
- [35] Tiwari MK, Adamowski J. Urban water demand forecasting and uncertainty assessment using ensemble wavelet-bootstrap-neural network models. *Water Resour Res* 2013;49(10):6486–507.
- [36] Tiwari MK, Chatterjee C. Development of an accurate and reliable hourly flood forecasting model using wavelet-bootstrap-ANN (WBANN) hybrid approach. *J Hydrol* 2010;394(3):458–70.
- [37] Tiwari MK, Chatterjee C. A new wavelet-bootstrap-ANN hybrid model for daily discharge forecasting. *J Hydroinf* 2011;13(3):500–19.
- [38] Jia Y, Culver TB. Bootstrapped artificial neural networks for synthetic flow generation with a small data sample. *J Hydrol* 2006;331(3):580–90.
- [39] Efron B. Bootstrap methods: another look at the jackknife. *Breakthroughs in statistics*. Springer; 1992. p. 569–93.
- [40] Efron B, Tibshirani RJ. An introduction to the bootstrap. CRC press; 1994.
- [41] Friedman JH. Multivariate adaptive regression splines. *The annals of statistics*; 1991. p. 1–67.
- [42] Box GE, Jenkins GM. Time series analysis: forecasting and control. revised ed Holden-Day; 1976.
- [43] Deo RC, Kisi O, Singh VP. Drought forecasting in eastern Australia using multivariate adaptive regression spline, least square support vector machine and M5Tree model. *Atmos Res* 2017;184:149–75.
- [44] Deo R, et al. Very short-term reactive forecasting of the solar ultraviolet index using an extreme learning machine integrated with the solar zenith angle. *Environ Res* 2017;155:141.
- [45] Ergodlu E. Electricity demand analysis using cointegration and ARIMA modelling: a case study of Turkey. *Energy Policy* 2007;35(2):1129–46.
- [46] McCulloch WS, Pitts W. A logical calculus of the ideas immanent in nervous activity. *Bull Math Biophys* 1943;5(4):115–33.
- [47] Huang H-Y. Unified approach to quadratically convergent algorithms for function minimization. *J Optim Theory Appl* 1970;5(6):405–23.
- [48] Dennis Jr. JE, Schnabel RB. Numerical methods for unconstrained optimization and nonlinear equations vol. 16. Siam; 1996.
- [49] Marquardt DW. An algorithm for least-squares estimation of nonlinear parameters. *J Soc Ind Appl Math* 1963;11(2):431–41.
- [50] HariKumar R, Vasanthi N, Balasubramani M. Performance analysis of artificial neural networks and statistical methods in classification of oral and breast cancer stages. *Int J Soft Comput Eng* 2009;ume. 2.
- [51] Draper NR, Smith H. Applied regression analysis vol. 326. John Wiley & Sons; 2014.
- [52] Montgomery DC, Peck EA, Vining GG. Introduction to linear regression analysis vol. 821. John Wiley & Sons; 2012.
- [53] Civelekoglu G, et al. Prediction of bromate formation using multi-linear regression and artificial neural networks. *Ozone: Sci Eng* 2007;29(5):353–62.
- [54] Twomey JM, Smith AE. Bias and variance of validation methods for function approximation neural networks under conditions of sparse data. *IEEE Trans Syst Man Cybern C Appl Rev* 1998;28(3):417–30.
- [55] Energex. Zone substation load data request form. <https://www.energex.com.au/about-us/contact-us/forms/general-forms/zone-substation-load-data-request-form;2018>.
- [56] Jeffrey SJ, et al. Using spatial interpolation to construct a comprehensive archive of Australian climate data. *Environ Model Softw* 2001;16(4):309–30.
- [57] Ghimire S, et al. Self-adaptive differential evolutionary extreme learning machines for long-term solar radiation prediction with remotely-sensed MODIS satellite and Reanalysis atmospheric products in solar-rich cities. *Remote Sens Environ* 2018;212:176–98.
- [58] Hsu C-W, Chang C-C, Lin C-J. A practical guide to support vector classification. 2003.
- [59] Friedman JH. Estimating functions of mixed ordinal and categorical variables using adaptive splines. Stanford Univ CA Lab for Computational Statistics; 1991.
- [60] Jakabson G. Adaptive regression splines toolbox for matlab/octave. Version 2013;1:72.
- [61] Hyndman R, Khandakar Y. Automatic time series forecasting: the forecast package for R 7. 2008. 2007 <https://www.jstatsoft.org/article/view/v027i03>.
- [62] Wang X, Smith K, Hyndman R. Characteristic-based clustering for time series data. *Data Min Knowl Discov* 2006;13(3):335–64.
- [63] Hu S. Akaike information criterion. Center for Research in Scientific Computation; 2007.
- [64] Barzegar R, et al. Mapping groundwater contamination risk of multiple aquifers using multi-model ensemble of machine learning algorithms. *Sci Total Environ* 2018;621:697–712.
- [65] Prasad R, et al. Ensemble committee-based data intelligent approach for generating soil moisture forecasts with multivariate hydro-meteorological predictors. *Soil Tillage Res* 2018;181:63–81.
- [66] Mohammadi K, et al. Support vector regression based prediction of global solar radiation on a horizontal surface. *Energy Convers Manag* 2015;91:433–41.
- [67] Willmott CJ, Robeson SM, Matsuura K. A refined index of model performance. *Int J Climatol* 2012;32(13):2088–94.
- [68] Willmott CJ. On the evaluation of model performance in physical geography. *Spatial statistics and models*. Springer; 1984. p. 443–60.
- [69] Willmott CJ. Some comments on the evaluation of model performance. *Bull Am Meteorol Soc* 1982;63(11):1309–13.
- [70] Willmott CJ. On the validation of models. *Phys Geogr* 1981;2(2):184–94.
- [71] Dawson CW, Abrahart RJ, See LM. HydroTest: a web-based toolbox of evaluation metrics for the standardised assessment of hydrological forecasts. *Environ Model Softw* 2007;22(7):1034–52.
- [72] Legates DR, McCabe GJ. Evaluating the use of “goodness-of-fit” measures in hydrologic and hydroclimatic model validation. *Water Resour Res* 1999;35(1):233–41.
- [73] Krause P, Boyle D, Bäse F. Comparison of different efficiency criteria for hydrological model assessment. *Adv Geosci* 2005;5:89–97.
- [74] Chai T, Draxler RR. Root mean square error (RMSE) or mean absolute error (MAE)?—Arguments against avoiding RMSE in the literature. *Geosci Model Dev (GMD)* 2014;7(3):1247–50.
- [75] Nash JE, Sutcliffe JV. River flow forecasting through conceptual models part I—a discussion of principles. *J Hydrol* 1970;10(3):282–90.
- [76] Li M-F, et al. General models for estimating daily global solar radiation for different solar radiation zones in mainland China. *Energy Convers Manag* 2013;70:139–48.
- [77] Mohammadi K, et al. A new hybrid support vector machine-wavelet transform approach for estimation of horizontal global solar radiation. *Energy Convers Manag* 2015;92:162–71.
- [78] Heinemann AB, et al. Sensitivity of APSIM/ORYZA model due to estimation errors in solar radiation. *Bragantia* 2012;71(4):572–82.
- [79] Campbell A. Price and income elasticities of electricity demand: evidence from Jamaica. *Energy Econ* 2018;69:19–32.
- [80] Akay D, Atak M. Grey prediction with rolling mechanism for electricity demand forecasting of Turkey. *Energy* 2007;32(9):1670–5.
- [81] Wan Z. MODIS land-surface temperature algorithm theoretical basis document (LST ATBD). Santa Barbara: Institute for Computational Earth System Science; 1999. p. 75.
- [82] Wan Z, et al. Quality assessment and validation of the MODIS global land surface

- temperature. *Int J Remote Sens* 2004;25(1):261–74.
- [83] MODIS. Moderate-resolution imaging spectroradiometer). https://modis.gsfc.nasa.gov/about/media/modis_brochure.pdf; 2018. *MODIS*.
- [84] Dee RC, et al. Universally deployable extreme learning machines integrated with remotely sensed MODIS satellite predictors over Australia to forecast global solar radiation: a new approach. *Renew Sustain Energy Rev* 2019;104:235–61.
- [85] Ghimire S, et al. Deep learning neural networks trained with MODIS satellite-derived predictors for long-term global solar radiation prediction. *Energies* 2019;12(12):2407.
- [86] Statistics ABO. Census. 2018 <http://www.abs.gov.au/>.
- [87] Galelli S, Castelletti A. Tree-based iterative input variable selection for hydrological modeling. *Water Resour Res* 2013;49(7):4295–310.
- [88] Cornejo-Bueno L, et al. Significant wave height and energy flux prediction for marine energy applications: a grouping genetic algorithm–Extreme Learning Machine approach. *Renew Energy* 2016;97:380–9.
- [89] Salcedo-Sanz S, et al. Feature selection in wind speed prediction systems based on a hybrid coral reefs optimization–Extreme learning machine approach. *Energy Convers Manag* 2014;87:10–8.
- [90] Salcedo-Sanz S, et al. Daily global solar radiation prediction based on a hybrid Coral Reefs Optimization–Extreme Learning Machine approach. *Sol Energy* 2014;105:91–8.
- [91] Ghimire S, et al. Wavelet-based 3-phase hybrid SVR model trained with satellite-derived predictors, particle swarm optimization and maximum overlap discrete wavelet transform for solar radiation prediction. *Renew Sustain Energy Rev* 2019;113:109247.
- [92] Quilty J, Adamowski J. Addressing the incorrect usage of wavelet-based hydrological and water resources forecasting models for real-world applications with best practices and a new forecasting framework. *J Hydrol* 563, 2018, 336–353.
- [93] Sloughter JM, Gneiting T, Raftery AE. Probabilistic wind speed forecasting using ensembles and Bayesian model averaging. *J Am Stat Assoc* 2010;105(489):25–35.
- [94] Wang D, et al. Multi-step ahead electricity price forecasting using a hybrid model based on two-layer decomposition technique and BP neural network optimized by firefly algorithm. *Appl Energy* 2017;190:390–407.