Renewable Energy 177 (2021) 1031–1044

Contents lists available at ScienceDirect

Renewable Energy

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

Advanced extreme learning machines *vs.* deep learning models for peak wave energy period forecasting: A case study in Queensland, Australia

Mumtaz Ali [a, *], Ramendra Prasad [b], Yong Xiang [a], Adarsh Sankaran [c], Ravinesh C. Deo [d], Fuyuan Xiao [e], Shuyu Zhu [a]

[a] Deakin-SWU Joint Research Centre on Big Data, School of Information Technology, Deakin University, VIC, 3125, Australia
[b] Department of Science, School of Science and Technology, The University of Fiji, Saweni, Lautoka, Fiji
[c] Department of Civil Engineering, TKM College of Engineering Kollam, Kerala, India
[d] School of Sciences, Centre for Applied Climate Sciences & Centre for Sustainable Agricultural Systems, University of Southern Queensland, Springfield, QLD, 4300, Australia
[e] College of Computer and Information Science, Southwest University, China

ARTICLE INFO

Article history:
Received 17 March 2020
Received in revised form
23 May 2021
Accepted 10 June 2021
Available online 12 June 2021

Keywords:
Deep learning
RNN
CNN
ELM
Peak wave energy period
Coastal waves

ABSTRACT

The peak period of an energy-generating wave is one of the most important parameters that describe the spectral shape of the oceanic wave, as this indicates the duration for which the waves prevail with respect to their maximum extractable energy. In this paper, a half-hourly peak wave energy period (T_P) forecast model is constructed using a suite of statistically significant lagged inputs based on the partial auto-correlation function with an extreme learning machine model developed and its predictive utility is benchmarked against deep learning models, i.e., convolutional neural network (CNN/CovNet) and recurrent neural network (RNN) models and other traditional M5tree, Conditional Maximization based Multiple Linear Regression (MLR-ECM) and MLR models. The objective model (ELM) *vs.* the comparison models (CNN, RNN, M5tree, MLR-ECM, and MLR) were trained and validated independently on the test dataset obtained from coastal zones of eastern Australia that have a high potential for implementation of wave energy generation systems. The outcomes ascertain that the ELM model can generate significantly accurate predictions of the half-hourly peak wave energy period, providing a good level of accuracy relative to deep learning models in selected coastal study zones. The study establishes the practical usefulness of the ELM model as being a noteworthy methodology for the applications in renewable and sustainable energy resource management systems.

1. Introduction

The future planning, design, and construction of any coastal infrastructure require profound knowledge on the oceanic wave behaviour, which are turbulent, chaotic and largely indeterministic over short-term or real-time intervals. The operation of coastal infrastructures, including ships, harbours, inshore and offshore structures, dynamic risers and position mooring systems together with related hydrogeological activities such as coastal erosion, sediment transport, coastal inundation, saltwater intrusion and ocean wave energy production is contingent upon this important information [1–3]. Two key oceanic parameters that are commonly considered in monitoring ocean waves include the significant wave height (denoted as Hs) that describes the magnitude of bulk wave energy present at respective locations and the peak wave period (denoted as Tp) that represents the time accompanying with the most energetic wave [4]. In particular, the ocean wave period remains the same because of shallow water transformations, and therefore, prior knowledge of Tp becomes increasingly important in designing any coastal infrastructure. This is because structural failures are likely to result if the peak wave period acts to resonate with the natural period of these structures [5]. In addition, the

* Corresponding author.
E-mail addresses: mumtaz.ali@deakin.edu.au (M. Ali), ramendrap23@gmail.com (R. Prasad), yong.xiang@deakin.edu.au (Y. Xiang), adarsh1982@tkmce.ac.in (A. Sankaran), ravinesh.deo@usq.edu.au (R.C. Deo), xiaofuyuan@swu.edu.cn (F. Xiao), zhushuyu@deakin.edu.au (S. Zhu).

https://doi.org/10.1016/j.renene.2021.06.052
0960-1481/

identification of extreme wave period conditions is also critical for the proper operation of wave energy converters (WECs) [3,6]. Considering this issue, the design of weather resilient structures that are placed within the vicinity of coastal zones and the proper management of WECs is vital through modelling systems. Therefore, future knowledge on the scale and the characteristics of the peaks of wave energy and its period is important.

For operational and design purposes, the conventional methods of estimating wave energy parameters are undertaken by measurements and analysis of the actual wave behaviour at respective locations [2,3]. However, large-scale spatial wave data are required for this purpose, which can be relatively costly and is somewhat difficult to collate. The readily available wind data is largely used to estimate the corresponding wave information that lacks the required accuracy and validity [2]. Often, the numerical wave models developed on the basis of deterministic equations are employed. These models are based upon oceanic bathymetry and wave generation mechanisms such as topographical interactions (shoaling, refraction, wave trapping, refraction and diffraction) and dissipation (wave breaking and bottom friction) processes [7,8]. In postulating wave equations the assumption is that water is homogeneous, incompressible, with no surface tension. In addition, the bottom layer is assumed to be horizontal, stationary, and impermeable, while the air-sea interface pressure is regarded to be constant and during wave shoals, no energy loss occurs. Further assumptions are required for storm-time simulations. The complexities arising from these assumptions in boundary conditions and nonlinear wave-wave interactions (i.e., triad and quartet resonances), however, induce innate uncertainties. Additionally, by virtue of their nature, the ocean waves are highly irregular and as evident from their non-uniform time series, together with irregular heights and lengths that continue in unsteady and unreliable sequences [9]. To compound the problem, wind-driven waves also become highly randomized in terms of their direction, amplitude and frequency, and thus, cannot be aptly described using deterministic modelling approaches that come with its own assumptions of boundary conditions and nonlinearity behaviour. This issue becomes more complicated when dealing with tasks associated with real-time monitoring of ocean waves, such as those performed at short-term, half-hourly intervals as studied in this paper.

Due to the advent of big data and predictive modelling techniques, artificial intelligence-based machine learning (ML) algorithms are now showing significant capabilities to approximate the nonlinear behaviour of interrelated elements (i.e., inputs and targets) whilst emulating a practical predictive system without the need for any *a priori* knowledge and oceanic boundary condition assumptions. Such characteristics of big data and predictive modelling techniques have made ML algorithms significantly useful in wave energy-related applications. In modelling the oceanic energy and related wave energy parameters, artificial neural network (ANN) have been one of the widely used ML models since they were seen to perform relatively better in comparison to the conventional regression-based models [10–13]. Specifically, the study of [7] forecasted Hs via ANN models, while [2] used 1 week's data as input to forecast Hs using a recurrent network called Non-linear Auto Regressive with exogenous input (NARX) network. In terms of wave energy prediction studies within the Caspian Sea, forecasting models based on ANNs were undertaken by Hadadpour et al. [14]. In addition, a cascaded feed-forward neural network (CFNN) with gene expression programming (GEP) was employed in a wave energy flux forecasting study [15]. Classification and regression tree type C5 algorithms [10], M5 model tree [16], Adaptive Network-Based Fuzzy Inference Systems (ANFIS) [11,17] and support vector machines (SVM) [11,13] have also been trailed in predicting various oceanic wave parameters such as significant wave heights and peak spectral period. Conversely, very few studies have been conducted pertaining to the forecasting of the peak period, Tp of an energy-generating wave. In one study [7] applied an ANN in modelling Tp within the Sines Harbour region in Portugal, while in another study [5] carried out an estimation of wave energy periods from the Erlang distribution numerical coastal wave model. However, the ML models based on ANN are considered classical or standalone approaches that can often lack sufficient predictive skills in capturing the non-linear dynamics of randomized oceanic waves.

Considering the shortfalls of prior studies on the peak period of an energy-generating wave and its important role in wave energy monitoring systems, further studies that specifically develop the Tp forecasting models are necessary. However, such models must also develop advanced ML techniques that are of a low-cost investment (i.e., time and labour savings) and are computationally efficient in providing a robust and practical implementation platform. Taking this primary objective into consideration, the present study focuses on designing an Extreme Learning Machine (ELM) model tailored for peak wave energy period forecasting at a near real-time, i.e., half-hourly forecast horizons. The forecasting performance of this fast and efficient modelling framework is then compared with a set of competing methods such as the M5 Model Tree (M5tree), Expectation Conditional Maximization based Multiple Linear Regression (MLR-ECM) Convolutional Neural Networks (CNN), and Recurrent Neural Networks (RNN).

The Single Layer Feedforward Neural Network (SLFN) structured ELM model developed in this study is able to acquire the predictive features from the dataset in tuning the SLFN parameters [18]. The ELM has been found to have a faster execution speed in comparison to conventional neural networks counterparts and has better accuracy [18–20]. This has led to many applications of the ELM model and their variants in the forecasting of Hs and related energy parameters. For example, the study of [21] developed an ensemble-ELM model [22], incorporated grouping genetic algorithm into an ELM model [23], designed a feature selection method using coral reef optimization embedded into an ELM model [24], built a universally deployable ELM model in forecasting global solar radiation using MODIS satellite data of Australia, and recently [25], coupled improved complete ensemble empirical mode decomposition with ELM. These all studies focussed on energy and environment-related research areas (e.g., Refs. [26–30]), where the ELM model has been found to significantly outperform the conventional models such as the ANN, SVM, MLR, random forest and other machine learning algorithms.

Other than the extensive application of the ELM model, two other significant comparative models have evolved quite recently, and these are currently being experimented within a number of research areas. These models include the Convolutional Neural Network (CNN) and the Recurrent Neural Network (RNN), both of which are spin-offs from the recent advancements in machine learning termed as deep learning approaches. Indeed, CNNs and their related deep learning algorithms have recently been attractive in the energy and hydrology and health sectors (e.g., Refs. [31–34]) but its application in the forecasting peak period of an energy-generating wave and the other ocean energy conversion systems is yet to be made. Other than deep learning models, the M5 Tree model, which is a regression tree-based approach, has wide applications in hydrology [35–42], solar ultraviolet (UV) radiation [26], photosynthetically active solar radiation [24], yet, no applications in forecasting of wave parameters has been noted. In addition, the Expectation Conditional Maximization based Multiple Linear Regression (MLR-ECM) model, which is a variant of multiple linear regression, has been applied for the comparison of the ELM (and related models) applied in a problem of half-hourly prediction of peak wave energy periods.

The performance of the computationally efficient ELM modelling framework is evaluated with four of the other ML models (i.e., two deep learning (CNN, RNN) and three conventional (MLR, MLR-ECM and M5 Tree models) and tested for its ability in forecasting half-hourly peak wave energy periods using antecedent lagged inputs based on Tp, in this study. In doing so, the salient lagged inputs based on wave height time series data were determined using partial autocorrelation function and the performance of the prescribed ELM (and related benchmark) model was then evaluated using statistical scoring metrics and diagnostic plots of data generated by the four benchmark models.

2. Theoretical framework of machine learning techniques

To construct and evaluate the half-hourly based peak wave energy period (T_P) forecasting model, the ELM as the objective method *vs* CNN, RNN, M5tree, MLR-ECM, and MLR models as the benchmark comparison techniques are described in the following section.

2.1. Extreme learning machine (ELM)

Designed by Huang et al. [43]; the ELM is believed to be a fast machine learning model consisting of SLFN making ELM computationally efficient [19]. The ELM is expressed as:

$$\sum_{i=1}^{M} \rho_i f(T_{Pk},; c_i, w_i) = T_{Pfor} \tag{1}$$

Where $i, k = 1, 2, 3, ..., M$ with $c_i \in \Gamma$ is the arbitrarily allotted bias *ith* node and $w_i \in \Gamma$ describes the weight of the predictor. Here M is the hidden layer of M nodes to the m $\geq$ 1 output nodes, and $\rho = [\rho_1, \rho_2, \cdots, \rho_M]^T$ is the output weight vector. The function $f(T_{Pk},; c_i, w_i)$ denotes the forecasted peak wave energy period (T_P) which is ELM nonlinear feature mapping of the *ith* hidden node at a lag (*t-j*) where *j = 1, 2, ..., 24*. Thus Eq. (1) can be written in the matrix as:

$$H\beta = Y \tag{2}$$

where $H = \begin{bmatrix} f(T_{P1},; c_1, w_1) & \cdots & f(T_{P1},; c_M, w_M) \\ \vdots & \cdots & \vdots \\ f(T_{PN},; c_1, w_1) & \cdots & f(T_{PN},; c_M, w_M) \end{bmatrix}_{N \times M}$,

$\beta = ({\beta_1}^T {\beta_2}^T, \cdots, {\beta_L}^T)^T{}_{m \times M}$ and $Y = ({t_1}^T {t_2}^T, \cdots, {t_L}^T)^T{}_{m \times M}$. Where H is an N-by-M matrix N being the number of rows while M being the number of columns. This linear system reduces to the following solution:

$$\Xi = H^{+} \Lambda \tag{3}$$

H^+ represents the Moore–Penrose generalized inverse of H. The SLFNs successfully obtain training with a low possibility of errors [44,45].

2.2. Convolutional neural network (CNN)

CNN is considered to be the earliest deep neural network utilized in computer vision [46]. The CNN model is largely used in graphics and video processing problems [47–49]. A typical CNN architecture is basically described as:

a) Input layer: The input layers of CNN are commonly organized in height, width, and depth structure. The T_P predictor data is paralleled to height, width, and depth coordinates to reserve the space-time association between respective datasets.
b) Convolutional layer: The kernels convolve on the significant PACF based lag input data mainly used in this layer. A neuron is linked with a preceding layer to estimate the inner-product of the area and weight. In broad, the CNN picks distinct convolution kernels based on particular circumstances.
c) Activation layer: The rectifier excitation function in this layer was adopted here and widely applied in deep learning. This function possesses independent suppression, large excitement boundary and scarce activation features.
d) Pooling layers: Down-sampling of the information attained in the activation layer is carried out in this layer. Here, the maximum value pooling layer is widely recommended.
e) Fully connected layer: This layer has different conventions conditional upon the connection of the activation function. In classification, SoftMax while for fixed values, the *relu* and sigmoid are used. Every connected neuron interacts in this layer.

It is worth mentioning that the convolutional and fully connected neurons retain weight and offset parameters trained by data [50].

2.3. Recurrent neural network (RNN)

The RNN also known as the long short-term memory (LSTM) neural network model has the capacity to study long-term dependencies among consecutive events on a relevant timestamp. The RNN has been found in several application areas including solar radiation [51,52], droughts [53], weather forecasting [54]. The RNN abstains long-term dependencies [55] which retains data in a controller exterior to the normal flow by a new state unit. This memory state block also takes place in a traditional neuron.

There are units known as input, output and forget gates in the hidden state that allows RNN (i.e. LSTM) to appraise and control data flow in isolated chunks which, then divides into two portions called memory cells and working memory. The memory cell is responsible for predecessor data from the previous hidden state (i.e., h_{t-1}) and a new input data, x_t, by the forget gate (g):

$$g_t = \sigma\left(w_g \times [h_{t-1}, x_t] + b_g\right) \tag{4}$$

Where as the working memory (h_t) is utilized as an output gate to govern the proportion of the current memoryc_t. $\tilde{c}_t$is an updated state created by $x_t$and h_{t-1} with the help oftanhlayer. It can be written as:

$$j_t = \rho\left(\nu_j \times [h_{t-1}, x_t] + \varphi_j\right) \tag{5}$$

$$\tilde{c}_t = \tanh(\nu_c \times [h_{t-1}, x_t] + \varphi_c) \tag{6}$$

Both C_{t-1} and C_t is determined in the forget gate (g) and the input gate (j) as follows:

$$c_t = g_t \times c_{t-1} + i_t \times \tilde{c}_t \tag{7}$$

There are two steps to process the output. The output gate is known as a new gate, which is responsible for deciding appropriate parts. Thec_t state is activated by tanh which is then clarified with multiplying by$o_t$to acquire the desired outputh_t:

$$o_t = \sigma(\nu_o \times [h_{t-1}, x_t] + b_o) \tag{8}$$

$$h_t = o_t \times \tanh(c_t) \tag{9}$$

Where $\nu_g, \nu_j, \nu_c, \nu_o$ are weight matrices, $\varphi_g, \varphi_j, \varphi_c, \varphi_o$ are bias vectors and $\rho()$ is the sigmoid activation function.

2.4. Expectation Conditional Maximization based multiple linear regression (MLR-ECM)

The MLR is a linear regression model that capitalizes on the causativeness among target and input variables utilizing the extreme deviations of the data to estimate their analogous regression coefficients. Another major advantage of MLR is that it can minimize the variations due to the unexplained "noise". Mathematically, MLR is expressed [56,57] with n being the number of observations of k input variables:

$$\Pi = \theta + \varpi_1 \nu_1 + \varpi_2 \nu_2 + \ldots + \varpi_k \nu_k \tag{10}$$

Where the forecasted H_{sig} is represented by $\Pi(n+1)$, $\nu(n \times k)$ denotes the input vectors, θ and β denotes the y-intercept and coefficient respectively [58,59]. Using the least-squares technique, the magnitude of β is calculated (*e.g.*, Refs. [60,61]). During the training process, a set of Π and Δ matrix are fitted following Eq. (10) to model the causality. The coefficients together with the y-intercept in the validation period give the forecast in terms of fitted MLR. The MLR model is optimized by the ECM algorithm to get the optimum set of parameters. The ECM algorithm is explained below.

The ECM is an extension of the expectation-maximization (EM) technique to identify the parameters of MLR following the data augmentation principle [62]. The observed or actual data is augmented with unobserved variables to expedite the maximum-likelihood exploration [63].

The architecture of ECM involves augmenting the observed and unobserved data to ease the maximum-likelihood search. Then using the present estimate of parameters in maximizing the expected log-likelihood, the algorithm constructs a task for log-likelihood expectation [63,64]. By conditioning on a subset of parameters, the ECM replaces the maximization step over one's parameters of interest [63]. For more details on ECM, please refer to Refs. [63,64].

2.5. M5 tree (M5tree) model

The M5tree model [65] is mainly constructed on a tree structure. The linear regression establishes an association among predictors and response variables [66] where these variables are partitioned into subgroups at further stages [67]. An N-dimensional sampling matrix of training data of input predictors (H_{max}, T_z, T_p, Dirr, and SST) with respect to the response data (H_{sig}) is constructed by M5tree [39] algorithm. The whole process is recursively constructed on the divide and conquer principle splitting the N-data samples by creating similar subsets to test using standard deviation and the decrease of $\pi_{\Re}$ [39,68] as in the following:

$$\pi_{\Re} = \pi(\Phi) - \sum\left(\frac{\Phi_k}{\Phi}\pi(\Phi_k)\right) \tag{11}$$

Here π indicates the set of examples and Φ_k is the respective *jth* result. The best split of data sample containing patterns and attributes acquired to optimize the M5tree model with a smaller $\pi_{\Re}$ until all the sample data reaches a stationary node. The pruning is established due to a complex network structure formed as a result of splitting the input data. A smoothing procedure is required to avoid the unexpected breaks arising in the adjacent data [39,68] which in turn increases the precision to update the linear equation of the model [65].

3. Description of case study and data

3.1. Study location

Australia is the largest island-continent, hence all states are bordered by the ocean. The southern region is surrounded by the Southern Ocean, while the west is bordered by the Indian Ocean and the east touches the Pacific Ocean [69]. Further, the Australian coast is also surrounded by several seas. They include the Arafura Sea, the Coral Sea and the Tasman Sea [70]. The vast and diverse coastal region comprised approximately 85% of the populations living within 50 km of these coasts. The diverse oceanographic conditions provide Australia a great potential for wave energy

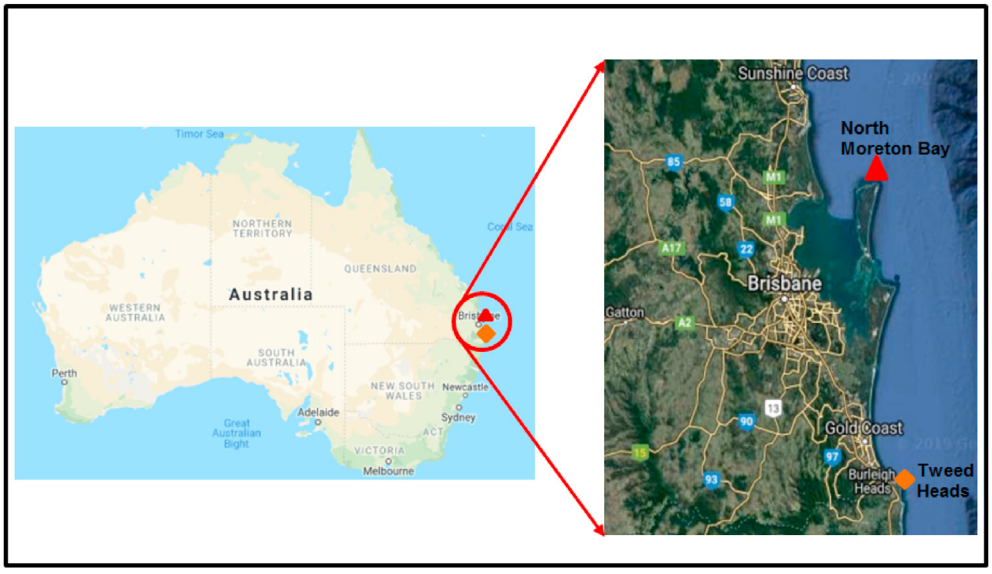

Fig. 1. Map of Australia showing the exploded section of the study sites.

Table 1
Geographic coordinates of the present study sites.

Study Sites	Long. (°S)	Lat. (°E)	Mean	Minimum	Maximum	Standard Deviation	Skewness	Kurtosis
North Moreton Bay	152.81°	25.25°	7.62	1.95	19.25	2.68	0.49	−0.42
Tweed Head	153.34°	28.10^0	8.60	2.57	21.18	2.64	0.29	−0.52

derived from coastal waves. In this paper, Queensland and New South Wales are the states where the coastal stations are selected carefully that which represents the diverse geophysical and bathymetric surroundings. The North Moreton Bay is counted as a significant coastal resource in Queensland.

The peak wave energy period, T_P, data is gathered by oceanographic wave measuring buoys anchored within the ocean. The Buoy Type is Datawell WR Mk3 0.9 Cu and the water depth is 16 m. The highest wave, as recorded at this site, was 7.50 m at 03:00 p.m. on 01/05/2015 [71]. The second station Tweed Heads is situated on the Tweed River in north-eastern New South Wales which also adjacent next to the border with Queensland. The Tweed head coastal station is located in New South Wales next to the border of the Gold Coast. Here the Buoy type is Datawell WR Mk4 0.9 Cu with a water depth of 24 m. In May 1996, the highest recorded wave was 13.1 m [71]. Fig. 1 displays a diagram of Australia together with the exploded sections showing the study sites.

3.2. Data collection and preparation

The peak wave energy period (T_P) is characterized as the period associated with the peak of the wave energy spectrum [72,73]. The peak wave period (in seconds) is defined as the wave period associated with the most energetic waves in the total wave spectrum at a specific point [72,73]. Wave regimes that are dominated by wind waves tend to have smaller peak wave periods, and regimes that are dominated by swell tend to have larger peak wave periods [72,73]. Therefore, the modelling of this parameter is extremely important for future expansion of the wave energy apparatus. The coastal waves possess a periodic phenomenon where crests and troughs persistently alternate and repeat during the propagation [74]. T_P is the time associated with a wave full of energy at a particular point in the total wave spectrum. Table 1 presents a basic and some descriptive analysis of the T_P used in this paper.

3.3. Model forecasting accuracy measures

The preciseness of ELM *vs*. CNN, RNN, M5tree, MLR-ECM and MLR models were measured via some statistical metrics to forecast the peak wave energy period (T_P). The mathematical representation of these metrics [75–83] is given below.

I Coefficient of Correlation (R):

$$R=\left(\frac{\sum_{i=1}^{N}\left(T_P^{Act,i}-\overline{T_P^{Act,i}}\right)\left(T_P^{For,i}-\overline{T_P^{For,i}}\right)}{\sqrt{\sum_{i=1}^{N}\left(T_P^{Act,i}-\overline{T_P^{Act,i}}\right)^2}\sqrt{\sum_{i=1}^{N}\left(T_P^{For,i}-\overline{T_P^{For,i}}\right)^2}}\right) \tag{12}$$

II Willmott's Index (E_{WI}) defined as:

$$E_{WI}=1-\left[\frac{\sum_{i=1}^{N}\left(T_P^{For,i}-T_P^{Act,i}\right)^2}{\sum_{i=1}^{N}\left(\left|T_P^{For,i}-\overline{T_P^{Act,i}}\right|+\left|T_P^{Act,i}-\overline{T_P^{Act,i}}\right|\right)^2}\right],\quad 0\le E_{WI}\le 1 \tag{13}$$

III Nash-Sutcliffe efficiency (E_{NS}) agreement:

$$E_{NS}=1-\left[\frac{\sum_{i=1}^{N}\left(T_P^{Act,i}-T_P^{For,i}\right)^2}{\sum_{i=1}^{N}\left(\overline{T_P^{Act,i}}-\overline{T_P^{For,i}}\right)^2}\right],\ -\infty\le E_{NS}\le 1 \tag{14}$$

IV Root mean square error ($RMSE$) is mathematically derived as:

$$RMSE=\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(T_P^{For,i}-T_P^{Act,i}\right)^2} \tag{15}$$

V Mean absolute error (MAE):

$$MAE=\frac{1}{N}\sum_{i=1}^{N}\left|\left(T_P^{For,i}-T_P^{Act,i}\right)\right| \tag{16}$$

VI Legates and McCabe's (E_{LM}):

$$E_{LM}=1-\left[\frac{\sum_{i=1}^{N}\left|T_P^{For,i}-T_P^{Act,i}\right|}{\sum_{i=1}^{N}\left|T_P^{Act,i}-\overline{T_P^{Act,i}}\right|}\right],0\le E_{LM}\le 1 \tag{17}$$

VII Relative root mean squared percentage error ($RRMSPE$; %), is stated as

$$RRMSPE=\frac{1}{N}\sum_{i=1}^{N}\left|\frac{\left(T_P^{For,i}-T_P^{Act,i}\right)}{T_P^{Act,i}}\right|\times 100 \tag{18}$$

VIII Relative mean absolute percentage error ($RMAPE$; %):

$$RMAPE=\frac{1}{N}\sum_{i=1}^{N}\left|\frac{\left(T_P^{For,i}-T_P^{Act,i}\right)}{T_P^{Act,i}}\right|\times 100 \tag{19}$$

In Eq. (12)- (19), $T_P^{Act,i}$ is the actual and $T_P^{For,i}$ is the forecasted ith magnitudes of peak wave energy period, while the actual and forecasted mean T_P values are represented by $\overline{T_P^{Act,i}}$ and $\overline{T_P^{For,i}}$,

where N represents the entire value of testing data records.

3.4. Modelling framework to forecast peak wave energy period

The advanced data intelligent models utilizing the ELM and others, i.e., CNN, RNN, M5tree, MLR-ECM and MLR models have been constructed in the MATLAB R2018(b) programming platform (The Math Works Inc. USA). All these models were accomplished on a 2.93 GHz dual-core personal computer (PC) running a Pentium 4 operating system. The construction of the ELM versus the benchmarking models, CNN, RNN, M5tree, MLR-ECM and MLR consisted of the following important steps:

Determining statistically significant lagged inputs: A set of correlograms based on partial auto-correlation coefficient (PACF) was performed to examine the relationship between T_P *at (t)-time* and *(t-1)*-time interval to generate the predictor variables. Fig. 2 shows the statistically significant lagged data of each coastal station. The time-lagged inputs at $(t-1)$ were used to develop the data intelligent models to forecast the half-hour lead-time peak wave energy period (T_P) values.

Data partitioning: The datasets were subdivided into 75% - training and 25% - testing phases, prior to model development from 1st January 2011 to 06th May 2017. The total number of training data was 111243 and testing data was 36943 at 30-min intervals. The purpose of data partitioning is to validate the ELM model in comparison with CNN, RNN, M5tree, MLR-ECM, and MLR models independently from training.

Normalization procedure: The normalization was performed via. Eq. (15) to smooth the data [84] whereby all data were restrained between [0, 1]:

$$T_{P.NORM} = \frac{(T_P - T_{P.MIN})}{(T_{P.MAX} - T_{P.MIN})} \tag{20}$$

In Eq. (20), T_P denotes the input/output data, $T_{P.MIN}$ = the lowest, $T_{P.MAX}$ = the largest and $T_{P.NORM}$ = the required normalized value.

Construction of ELM: In this stage, the ELM model is developed for half-hourly peak wave energy period forecasts. After feeding the statistically significant PACF lags at $(t-1)$, $(t-2)$, …., $(t-10)$ for North Moreton Bay station and $(t-1)$, $(t-2)$, …., $(t-11)$ lags for Tweed Head station into the ELM model, the trial-and-error procedure was used to fix the best-hidden neurons and activation functions. Different combinations of hidden neurons represented by n ranging from 1 to 60 were tried with several activation functions such as sine (sin), sigmoidal (sig), hard limit (hardlim), radial basis (radbas) and triangular basis (tribas) functions. For North Moreton Bay, the optimum number of hidden neurons n is 11 with sigmoidal (sig) as an activation function, whereas for Tweed Head station, the optimum hidden neurons are 10 and activation function is again sigmoid (sig). After training, the ELM model is validated independently on testing data.

Benchmark comparing Models: The performance of the advanced ELM is compared with CNN, RNN, M5tree, MLR-ECM, and MLR models. The ELM, M5tree, MLR-ECM and MLR models were implemented and executed in the Matlab R2018b programming platform whereas the CNN (*i.e.* LSTM) and RNN networks were executed in Python with the Keras deep learning package. The input layer of RNN consists of no. of features and time-steps. Further, hidden neurons were 80 while relu as an output activation function with one neuron in this layer. The highest number of epochs is 50. For constructing the CNN model, the input nodes were 514, activation function was relu, batch size = 65 and so on. Other parameters of CNN and RNN can be seen in Tables 2 and 3.

While constructing the M5tree model, attention was paid to the parameters including model Tree = 1, the minimum number of training data cases = 4, pruning = 1, smoothing = 0, smoothing coefficient = 15 and split threshold = 0.05. On the other hand, the MLR-ECM and MLR models utilize several types of parameters such as regression coefficients, F-value and covariance matrix as illustrated in Tables 4 and 5. The architecture of ECM consists of augmenting the actual data with the predictor data to enable the successive maximum-likelihood parameter search. In order to maximize the expected log-likelihood, based on the existing estimations, the ECM develops a function for this evaluation [63,64]. Finally, the maximization step over one's parameters of interest is replaced by conditioning on a subset of these parameters [63].

Fig. 3 shows the schematic view of the modelling strategy developed in this paper.

4. Experimental results

The proposed data-driven models, including the ELM, and its counterpart models, are undertaken to predict T_P for two highly rich wave energy zones in Queensland, Australia. The forecasting accuracy of the ELM model is evaluated against the CNN, RNN, M5 Model Tree, MLR-ECM and the MLR models based on *R*, *RMSE*, and *MAE* within the testing dataset.

The ELM model established for North Moreton Bay region was able to accomplish a high *R* value and generally lower *RMSE* and *MAE* values (i.e., $R = 0.963$, $RMSE = 0.52$ min, $MAE = 0.27$ min) in terms of predicting the T_P values. This contrasted the results of the CNN model (i.e., $R = 0.932$, $RMSE = 0.98$ min, $MAE = 0.72$ min), the

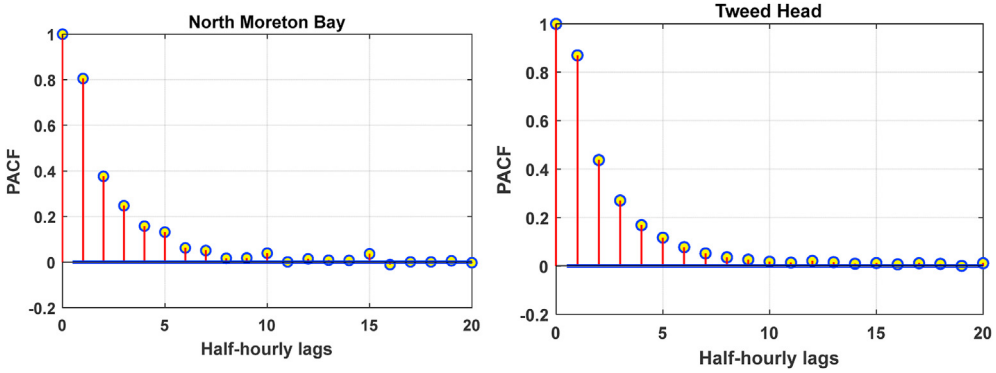

Fig. 2. Correlograms based partial auto-correlation function (PACF) performed on each of the study sites examining covariance between the peak wave energy period (T_P) and the predictor (or input) variables for study site 1: North Moreton Bay; and study site 2: Tweed Head. Blue lines indicate the significance of r at the 95% confidence interval.

Table 2
The architecture and design parameters of the optimized CNN model. Note: 'relu' = rectified linear units and 'RMSprop' optimizer restricts the oscillation of features in the vertical directions.

Study Sites	Learning Rate	Dense Layers	Input Nodes	Dense Nodes	Activation function	Batch size	Adam decay	optimizer	Dropout rate	epochs	CNN layers
North Moreton Bay	0.001	1	512	13	relu	64	1e-3	RMSprop	0.2	1	2
Tweed Heads	0.001	1	512	13	relu	64	1e-3	RMSprop	0.2	1	2

Table 3
The architecture and design parameters of the optimized RNN model.

Study Sites	Learning Rate	Hidden Neurons	Activation Function	Decay Rate	Optimizer	Dropout Rate	epochs
North Moreton Bay	0.01	80	relu	0.01	opt	0.1	50
Tweed Heads	0.01	80	relu	0.01	opt	0.1	50

Table 4
The design parameters and the coefficients of the hybridized MLR-ECM model where c, α1, …α10 are the model parameters.

Study Sites	c	α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8	α_9	α_{10}
North Moreton Bay	0.321	0.216	0.150	0.092	0.062	0.039	0.022	0.031	0.031	0.032	–
Tweed Head	0.291	0.213	0.154	0.096	0.074	0.051	0.035	0.023	0.022	0.017	0.020

Table 5
Design parameters and coefficients of the standalone MLR model where c (y-intercept) b1 … b11 are the model parameters.

Study Sites	c	b_1	b_2	b_3	b_4	b_5	b_6	b_7	b_8	b_9	b_{10}	b_{11}
North Moreton Bay	0.019	0.313	0.211	0.146	0.088	0.059	0.036	0.018	0.027	0.025	0.024	–
Tweed Head	0.015	0.285	0.209	0.151	0.093	0.072	0.049	0.032	0.020	0.018	0.013	0.013

RNN model ($R = 0.928$, $RMSE = 1.24$ min, $MAE = 1.08$ min), the M5 Model Tree ($R = 0.872$, $RMSE = 0.94$ min, $MAE = 0.47$ min), the ML model ($R = 0.932$, $RMSE = 0.70$ min, $MAE = 0.37$ min) and the MLR-ECM model ($R = 0.931$, $RMSE = 0.69$ min, $MAE = 0.35$ min). Equally, the ELM was able to generate considerably healthier forecasts of T_p for the Tweed Head case study station (see Table 6). This result thus established that the ELM is a possible candidate model to be implemented as a well-designed forecasting approach for T_P predictions in contrast to the other benchmarking models based on deep learning (CNN and RNN), the M5 Model Tree, MLR and MLR-ECM models.

In Table 5, we employed multiple criteria based on E_{WI}, E_{NS} and E_{LM} to examine the ELM *vs.* CNN, RNN, M5tree, MLR and MLR-ECM models. Based on to these model penalization measures, the results produced by the ELM model for the case of North Moreton Bay yielded a value of ($E_{WI} = 0.933$, $E_{NS} = 0.926$ and $E_{LM} = 0.809$), followed by those for an MLR-ECM model ($E_{WI} = 0.889$, $E_{NS} = 0.867$ and $E_{LM} = 0.757$), the MLR model ($E_{WI} = 0.873$, $E_{NS} = 0.865$ and $E_{LM} = 0.741$), the M5 Model Tree ($E_{WI} = 0.790$, $E_{NS} = 0.751$ and $E_{LM} = 0.674$), the CNN model ($E_{WI} = 0.657$, $E_{NS} = 0.733$ and $E_{LM} = 0.498$) and the RNN model ($E_{WI} = 0.334$, $E_{NS} = 0.571$ and $E_{LM} = 0.245$) models. Considering these higher-order metrics compared to R, RMSE and MAE it is evident that the ELM model has advanced predictive capabilities to provide accurate forecasts of T_p.

For the Tweed Head station, again, the ELM method acts as a more superior model that exhibits ($E_{WI} = 0.950$, $E_{NS} = 0.937$ and $E_{LM} = 0.822$). The lower statistical score metrics attained by the benchmarking models (as shown in Table 7) are backing the supremacy of the prescribed ELM model. Hence, according to these criteria, the prescribed ELM model could be considered 'very satisfactory' at all sites in predicting T_P. This is verified by the fact that predictive models that registered an E_{NS} below 0.800 are considered as 'unsatisfactory', while those with E_{NS} ranging 0.800–0.900 are deemed 'fairly good' while the ones with E_{NS} exceeding 0.900 are 'very satisfactory' [85].

Fig. 4 plots the kernel smoothing density function (KSDF) of the forecasted and actual peak wave energy period generated by the ELM *vs.* CNN, RNN, M5tree, MLR-ECM, and MLR models. Fig. 4 displays the ELM model at all sites registered equally the same KSDF distributions, enlightening its higher accuracy. The KSDF-plots provide a clear difference of accuracy since the distributions of ELM for all sites (North Moreton Bay and Tweed Head) fairly followed the same path as actual T_P in the testing period. Furthermore, the empirical cumulative distribution function (ECDF) of forecasted and actual T_P was also drawn in Fig. 4 to catch a clear representation. For all two locations, the ECDF line plots of ELM displayed a very close profile against CNN, RNN, M5tree, MLR-ECM and MLR models. Hence, the KSDF-plots (Fig. 4) in combination with the ECDF (Fig. 5) further ascertains improved efficiency of the ELM model in predicting half-hourly T_P compared to the benchmarking deep learning and other data intelligent models.

Fig. 6 displays a scatterplot between the forecasted and the actual values of T_P where the coefficient of determination (r^2) is also computed. The ELM was undoubtedly well accurate than the benchmark counterparts based on deep learning (i.e., CNN, RNN) and the other simplistic models (i.e., M5 Model Tree, MLR, MLR-ECM)) in terms of r^2 (ELM = 0.928, CNN = 0.869, RNN = 0.861, M5 Model Tree = 0.761, MLR-ECM = 0.867, MLR = 0.861) for North Moreton Bay coast. The proposed ELM model for the Tweed Head region is also better (Fig. 5) against benchmark approaches.

To compare directly the proposed ELM model *vs.* the CNN, RNN, M5 Model Tree, MLR and the MLR-ECM models, Fig. 7 plots the forecasted and the actual T_P for the tested dataset. For all two coastal stations, the time-series graph of forecasted T_P generated by the proposed ELM model appears to be more stable with actual T_P. Contrary to that, the forecasted T_P generated by M5 Model Tree and the RNN models are more unstable showing the high fluctuation of forecasted T_P in comparison with actual T_P. Overall, the proposed ELM model seems to perform very well for all study regions while the deep learning and the M5 Model Tree, MLR and MLR-ECM

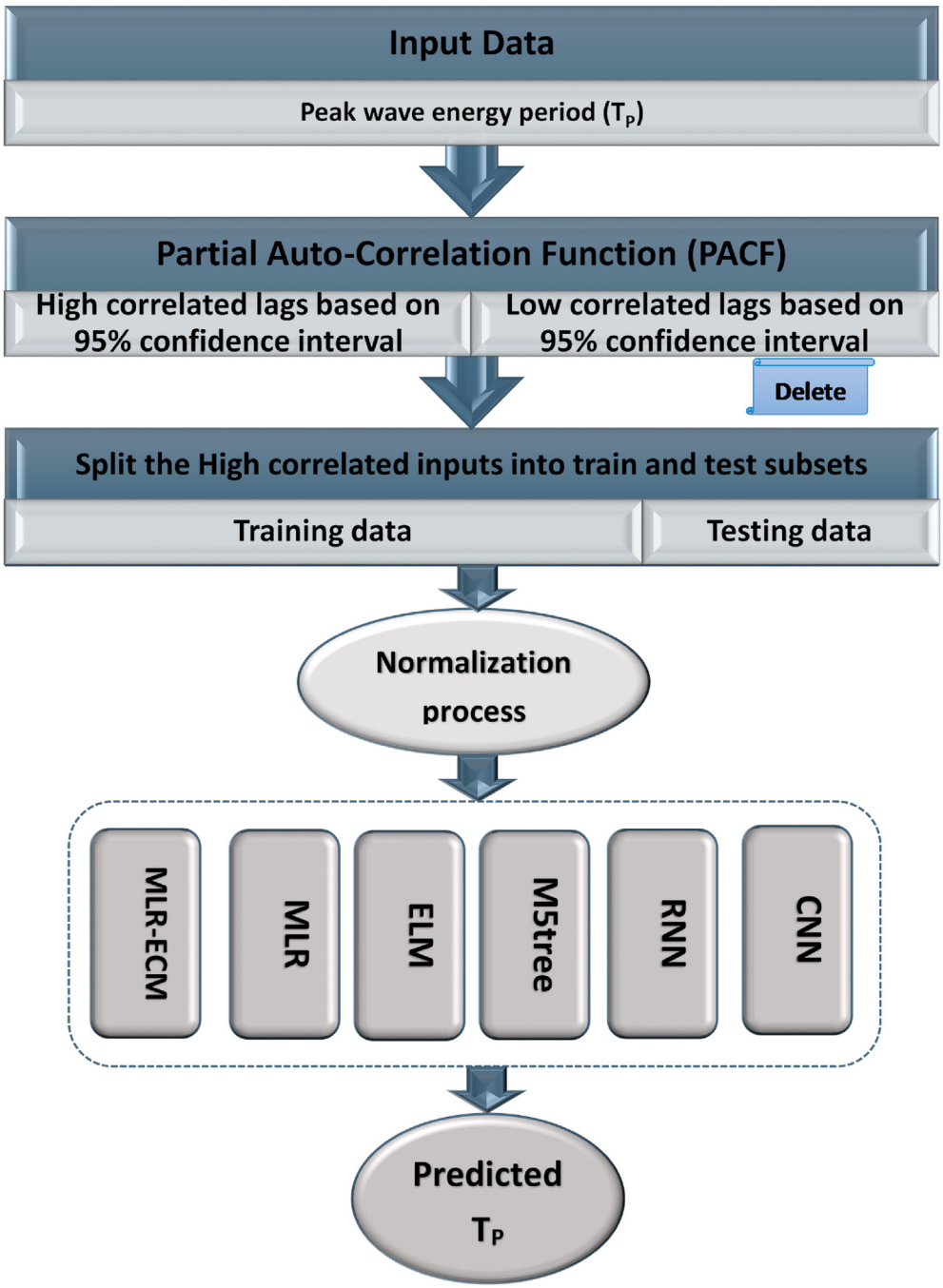

Fig. 3. Schematic view of the modelling strategy designed in this study.

models do not provide reasonably good results in the testing phase.

From the modelling outcomes enumerated in Tables 6 and 7, a discernible difference in the performance of these models is apparent. However, an important restriction of the above measures (Eq. (12)–(17)) is their failure to match models that operate at physically different locations. The comparative accuracy (Table 8) therefore appropriately revealed that the ELM model exhibited the lowest *RRMSPE* and *RMAPE* (i.e., normalized percentage errors) in comparison with the CNN, RNN, M5 Model Tree, MLR and MLR-ECM approaches for these study sites. More accurately, the magnitudes of *RRMSPE* and *RMAPE* when comparing the ELM model with the subsequent best forecasting technique, MLR-ECM, in the grouping [ELM: MLR-ECM] were as follows: North Moreton Bay: [8.94%, 4.63%: 11.95%, 5.77%] and Tweed Head: [7.66%, 4.05%: 10.02%, 5.09%]. The deep learning counterpart models (i.e., CNN and RNN) also appeared to be relatively poor models in this study. Consequently, the *RRMSPE* and *RMAPE* show clearly that the ELM accomplished the best predictive accuracy for the Tweed Head case study site, followed by North Moreton Bay.

To provide further interpretation of the predicting capacity of the designed modelling systems for the prediction of T_p, Fig. 8 plots a frequency distribution of the forecasted T_P generated by the ELM *vs.* the other benchmarking techniques including the actual T_P. The attained frequency distribution of the ELM model and the other

Table 6
The testing performance of the ELM model vs. CNN, RNN, M5tree, MLR, and MLR-ECM models as measured by correlation coefficient (R), root mean square error (RMSE) and mean absolute error (MAE).

Study Site 1: North Moreton Bay			
Forecast Models	*R*	*RMSE* (min)	*MAE* (min)
CNN	0.932	0.98	0.72
RNN	0.928	1.24	1.08
M5tree	0.872	0.94	0.47
MLR	0.932	0.70	0.37
MLR-ECM	0.931	0.69	0.35
ELM	0.963	0.52	0.27
Study Site 2: Tweed Head			
CNN	0.942	0.94	0.75
RNN	0.939	2.06	1.59
M5tree	0.864	1.06	0.57
MLR	0.945	0.67	0.39
MLR-ECM	0.945	0.67	0.36
ELM	0.968	0.51	0.28

Table 7
The performance of the ELM model vs. CNN, RNN, M5tree, MLR and MLR-ECM models using the Willmott's Index (EWI), Nash-Sutcliffe Coefficient (ENS) and the Legates & McCabe's (ELM) Index of Agreement. Note that the best model is boldfaced (blue).

Study Site 1: North Moreton Bay			
Forecast Model	E_{WI}	E_{NS}	E_{LM}
CNN	0.657	0.733	0.498
RNN	0.334	0.571	0.245
M5tree	0.790	0.751	0.674
MLR	0.873	0.865	0.741
MLR-ECM	0.889	0.867	0.757
ELM	0.933	0.926	0.809
Study Site 2: Tweed Head			
CNN	0.752	0.783	0.524
RNN	0.621	−0.037	−0.017
M5tree	0.787	0.726	0.636
MLR	0.909	0.889	0.751
MLR-ECM	0.920	0.892	0.771
ELM	0.950	0.937	0.822

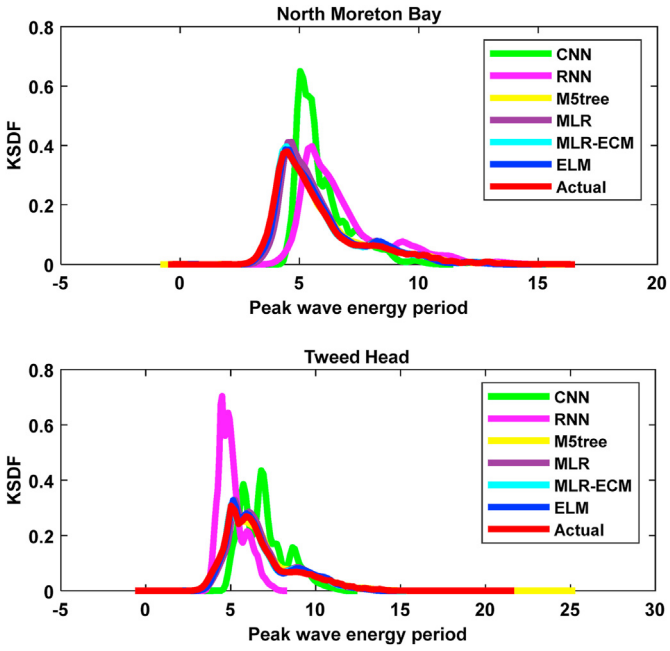

Fig. 4. Kernel smoothing density function (KSDF) of the forecasted and actual peak wave energy period generated by the ELM model *vs.* CNN, RNN, M5tree, MLR-ECM and MLR models applied at the study sites North Moreton Bay and Tweed Head.

benchmarking models against the actual T_P values showed the range to be $0 \leq T_P \leq 15$ for both sites. By comparing the frequency of the over and under forecasted T_P, the major variance was apparent in CNN and RNN, followed by M5tree models.

To broadly gauge the efficiency, a Taylor graph that delivers a thorough and comprehensive assessment of the models is plotted [86]. Fig. 9 depicts a more concrete and conclusive statistical association of the forecasted and actual T_P reliant on R and standard deviations. It is demonstrated that the CNN, RNN, and M5tree are not proper as the R to standard deviation was extremely farther from the observed T_P. The ELM model was closely positioned to the actual T_P approving the prediction was considerably better at both locations.

A further evaluation of the forecasts of maximum and minimum cases was also carried out to determine the model efficacy for extreme events as in Fig. 10. In both the maximum cases, the ELM model followed the trend closely, yet was unable to get to the peak values. The RNN model at the maximum case at Site 1 performed well in capturing the maximum values. For minimum instances, the ELM model tried to capture the trend and at the minimum case at Site 2, the ELM models almost reached the observed minimum values, while the other deep learning models (RNN and CNN) over-predicted these events. the ELM model has been seen to perform well in extreme events, however, further refining or hybrid methodologies would be required in order for the model to fully capture the required observed maximum and minimum magnitudes.

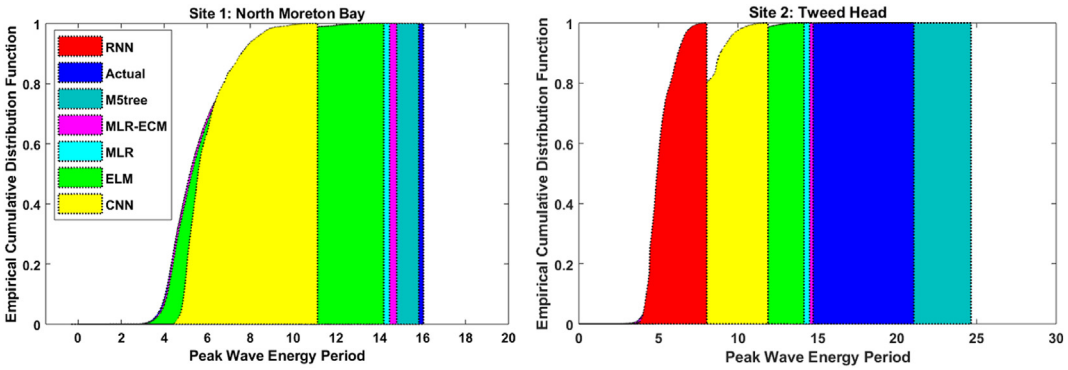

Fig. 5. Empirical cumulative distribution function (ECDF) of the forecasted and the actual peak wave energy period generated by the ELM model *vs.* CNN, RNN, M5tree, MLR-ECM, and MLR models.

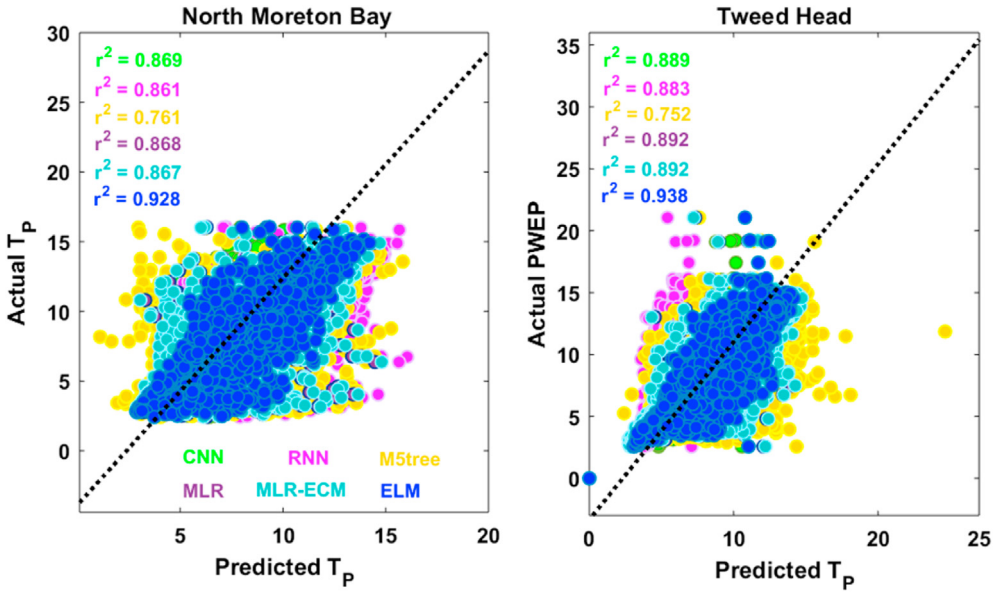

Fig. 6. Scatterplots of the forecasted *vs.* the actual peak wave energy period (T_P) for the two tested sites in Australia's wave energy region. A least-squares regression line and coefficient of determination (r^2) are also inserted in each sub-panel.

5. Discussion with further analysis

The aptness of deep learning models (i.e., RNN and CNN), as well as a number of classical machine learning models (i.e. ELM, M5tree, MLR-ECM and MLR), have been explored here to forecast peak wave energy period (T_P). The accuracy of the ELM models was substantially better than RNN (*i.e.* LSTM) and CNN models to confirm that the deep learning models were not performing well. Further, the ELM model's performance was reasonably good as compared to M5tree, MLR-ECM and standalone MLR models in this paper (Tables 4–6) for all locations, illustrating that the ELM model was a well-established and optimized algorithm to extract pertinent features to simulate T_P using significant PACF lags. The developed ELM model was well assessed to produce lower errors based on RMSPE, MAPE and larger magnitudes of R, E_{WI}, E_{NS} and E_{LM} (Tables 4and5) which confirms better efficiency based on these evaluation criteria (Eqs. (12)–(19)).

The forecasted peak wave energy period (T_P) generated by the ELM model can help in describing the future spectral shape of ocean waves [87,88]. Further, the forecasted T_P indicates the duration for which the waves carry maximum energy. Hence, accurate T_P forecasting of the whole spectrum is important for wave energy estimations as well [89]. The precise prediction of T_P is also important in assessing the interaction of waves and human-built structures such as drilling platforms [90].

Data partitioning can be very crucial in improving the performance of machine learning models. It is to be noted that the size of the dataset is important for a good generalization of the model [91]. The different data partitioning used in data intelligent techniques has different performance accuracy. The dataset from 1st January 2011 to 06th May 2017 was partitioned straight into 75% - training and 25% - testing phases, following [92] as it is the most common approach for data partitioning [93]. It is important to note that the cross-validation method or any random sampling procedure cannot be adopted here as time-series data by definition occur in a temporal order/sequence and this order or sequence must be preserved in order to keep the structure of the series intact [94].

Accurate T_P predictions are imperative for the efficient operation of the spectral shape of ocean waves and energy [87–89]. From a practical viewpoint, the characterization of the T_P forecast horizon could be extended to intermediate or short-term periods together with real-time forecasting. Forecasts for sustainable

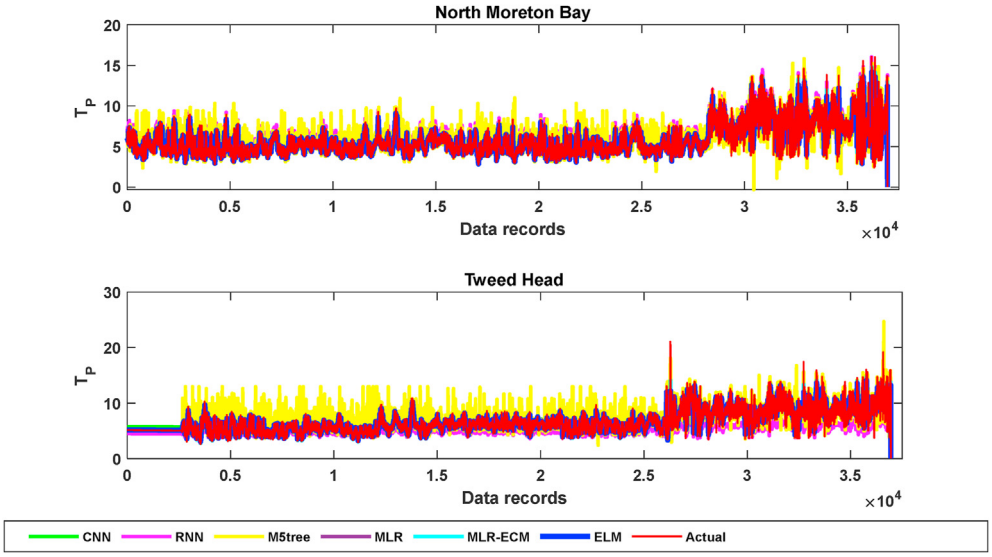

Fig. 7. Time-series plot of the actual and the forecasted peak wave energy period (PWEP) using the ELM model *vs* the CNN, RNN, M5tree, MLR-ECM and MLR models for all study sites.

Table 8
The geographic comparison of the accuracy of the ELM model vs. other comparative models in terms of the relative errors (RMSPE, %) and (MAPE, %) computed within the test sites. Note that the best model is boldfaced (blue).

Forecast Model	North Moreton Bay		Tweed Head	
	RMSPE, %	MAPE, %	RMSPE, %	MAPE, %
CNN	16.95	12.97	14.21	12.08
RNN	21.48	20.77	31.03	20.94
M5tree	16.34	7.92	15.96	8.79
MLR	12.04	6.36	10.15	5.79
MLR-ECM	11.95	5.77	10.02	5.09
ELM	8.94	4.63	7.66	4.05

energy via coastal waves can strategize active usage of energy, ranging from clean energy production to energy management early warning systems [95,96].

The scope of this study was short-term or near real-time forecasting and was focused on a 30-min forecast horizon. Moreover, long-term forecasting will be conducted in future follow-up studies with a comparison with numerical and other mathematical models. When forecasts differ widely from actuals, it creates problems for the oceanic waves. Essentially, short-term forecasts are bound to be more accurate than long-term forecasts, since a longer forecasting horizon significantly increases the chance of changes not known yet that can impact the future magnitudes.

Since the present work employed the historical significant lags based on PACF to forecast T_P, it has some limitations. To broaden the input variable pool, other oceanic parameters (e.g., maximum wave height, significant wave height, zero up-crossing wave period, direction, and sea surface temperature, etc.) and climatological data such as wind-speed, sunshine and air temperatures, solar radiation etc. could also be incorporated to improve the T_P forecast. The remotely sensed parameters via satellites or other atmospheric now-cast simulation models) (*e.g.*, Refs. [97–101]) could probably be incorporated to significantly improve the prediction of T_P in remotely situated oceanic positions. Furthermore, as process-based working systems are resource-demanding and cost-prohibitive, the advanced ELM is considered a feasible solution, that is lucid and easy to comprehend. Moreover, medium and long term (hourly, daily, weekly and monthly) T_P at multiple forecast-horizon can be performed in further independent work.

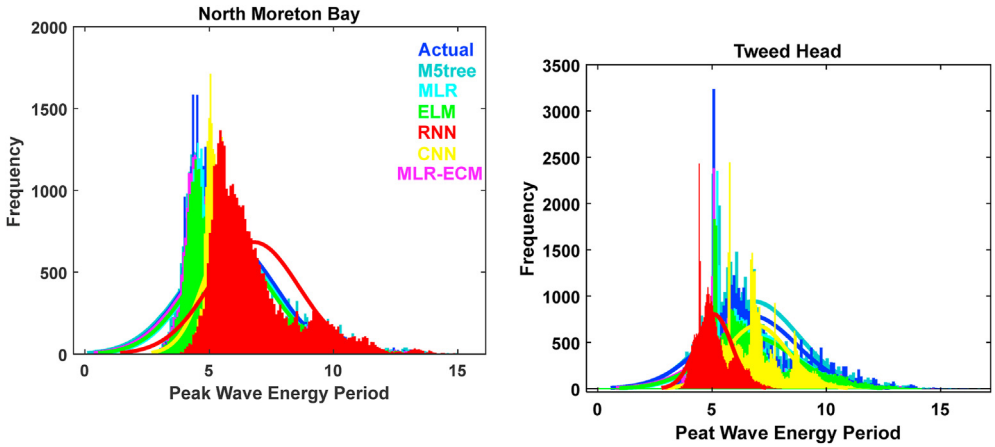

Fig. 8. Histogram of the frequency distribution of forecasted and actual peak wave energy period calculated for the ELM model *vs.* CNN, RNN, M5tree, MLR-ECM, and MLR models.

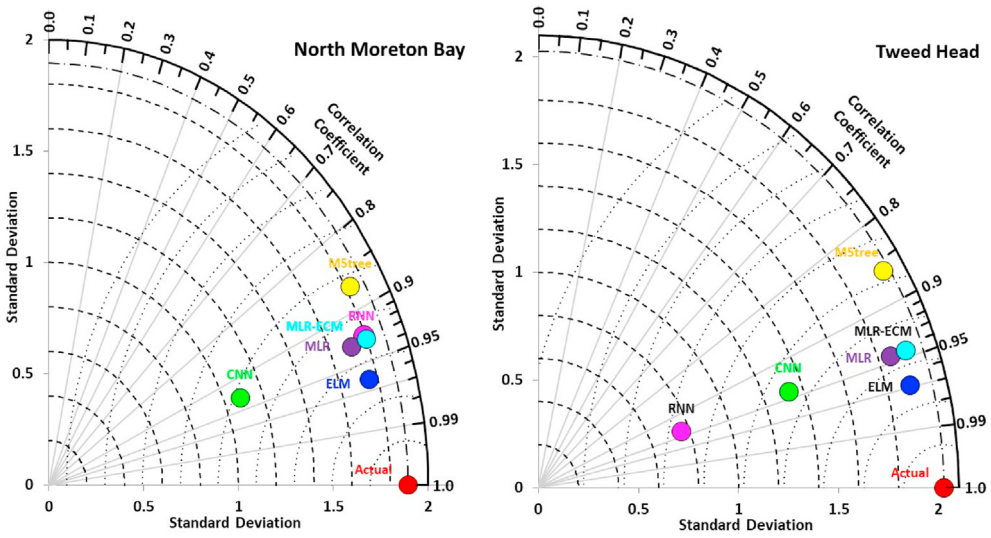

Fig. 9. Taylor diagram depicting the correlation coefficient for the ELM model *vs.* CNN, RNN, M5tree, MLR-ECM and MLR models at the two tested stations in Queensland, Australia.

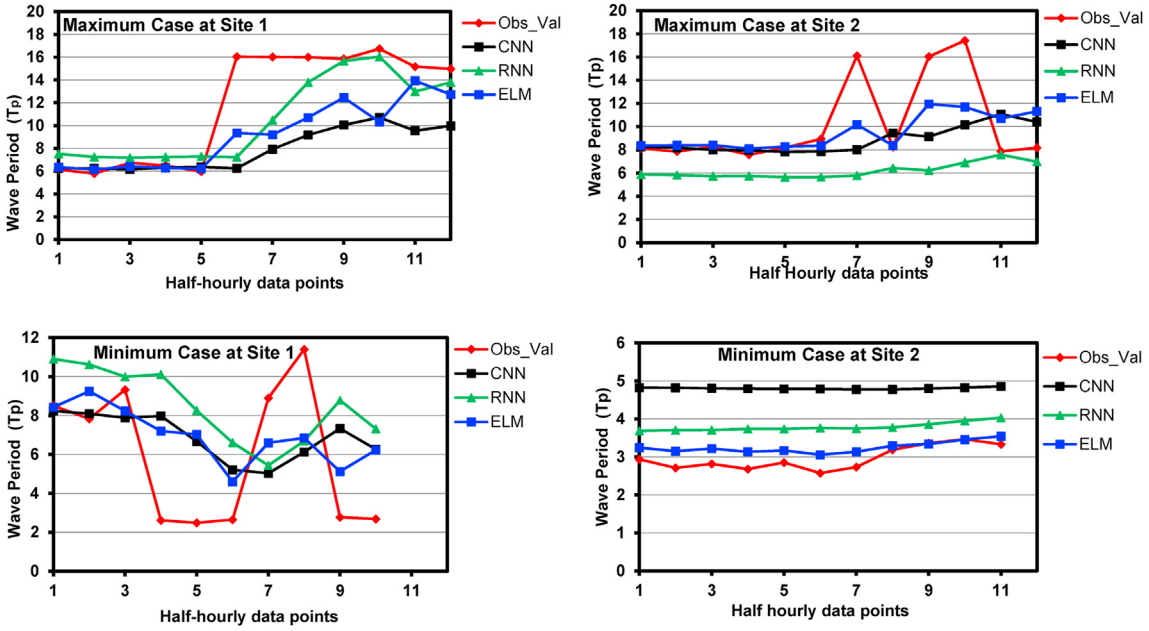

Fig. 10. Forecasting maximum and minimum cases at the two test sites with the proposed ELM model.

6. Conclusion

A robust extreme learning machine (ELM) model is designed for near real-time forecasting of T_P, which is the peak period of an energy-generating wave. The accuracy of ELM models was benchmarked with deep learning models (CNN and RNN) as well as M5tree, MLR-ECM and MLR models using several types of evaluation criteria and analytical plots. The constructed ELM model used the statistically significant lags based on PACF of T_P at the 30-min interval to forecast future T_P values. The proposed ELM model is validated against benchmark models at two high wave energy zones North Moreton Bay and Tweed Heads in Queensland, Australia.

This proposed modelling framework implementing ELM modelling was evaluated against a suite of deep machine learning models in forecasting T_P. Extending the scope of the proposed ELM model, we can aver that new work can validate the model in other emerging areas of interest such as solar radiation, wind energy, rainfall patterns, drought, hydrology, agriculture crops, and energy demand, to enable the government and policymakers to manage the climate change scenarios, agriculture crops, and energy-related matters with clearer models.

The proposed ELM-based peak wave energy period forecast can enable Governments and investors in the renewable and sustainable energy sector for better decision making (*e.g.*, smart grids, efficient and economic integration of wave energy and energy management systems). Moreover, the modelling strategy can be beneficial to other applications in physical modelling, especially climate change scenarios on energy utilization and monitoring where advanced artificial intelligence models provide better predictions.

CRediT authorship contribution statement

Mumtaz Ali: Writing – original draft, preparation, Conceptualization, Methodology, Software, Visualization. **Ramendra Prasad:** Conceptualization, Writing – review & editing, Investigation. **Yong Xiang:** Writing – review & editing, Supervision. **Adarsh Sankaran:** Writing – review & editing. **Ravinesh C. Deo:** Writing – review & editing, Supervision. **Fuyuan Xiao:** Writing – review & editing. **Shuyu Zhu:** Tables.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors are thankful to the Environment and Science, Queensland Government, Coastal Data System, Queensland for providing the peak wave energy data.

References

[1] P.V. Guimarães, L. Farina, E.E. Toldo Jr., Analysis of extreme wave events on the southern coast of Brazil, Nat. Hazards Earth Syst. Sci. 14 (12) (2014) 3195–3205.
[2] D.I. Gopinath, G.S. Dwarakish, Wave prediction using neural networks at new mangalore port along west coast of India, Aquatic Procedia 4 (2015) 143–150.
[3] A.C. Eckert-Gallup, C.J. Sallaberry, A.R. Dallman, V.S. Neary, Application of principal component analysis (PCA) and improved joint probability distributions to the inverse first-order reliability method (I-FORM) for predicting extreme sea states, Ocean Eng. 112 (2016) 307–319.
[4] NOAA, Peak wave period and direction forecasts, 2008.
[5] G. Muraleedharan, M. Sinha, A.D. Rao, N.U. Nair, P.G. Kurup, Estimation of wave period statistics using numerical coastal wave model, Nat. Hazards 49 (2) (2008) 165–186.
[6] W. Sheng, H. Li, A method for energy and resource assessment of waves in finite water depths, Energies 10 (4) (2017) 460.
[7] O. Makarynskyy, A.A. Pires-Silva, D. Makarynska, C. Ventura-Soares, Artificial Neural Networks in the Forecasting of Wave Parameters, 7th International workshop on wave hindcasting and forecasting. Banff, Alberta, Canada, 2002.
[8] L. Gorrell, B. Raubenheimer, S. Elgar, R.T. Guza, SWAN predictions of waves observed in shallow water onshore of complex bathymetry, Coast. Eng. 58 (6) (2011) 510–516.
[9] M. Raza Ul Mustafa, S.Z.A.S. Ahmad, M.K.A. Husain, N.I.M. Zaki, M.H. Mohd, G. Najafian, I. Bin Othman, M. Latheef, D. Bayu Endrayana, N. Zulaikha Bt Yusof, Comparison of various spectral models for the prediction of the 100-year design wave height, MATEC Web Conf. 203 (2018), 01020.
[10] J. Mahjoobi, A. Etemad-Shahidi, An alternative approach for the prediction of significant wave heights based on classification and regression trees, Appl. Ocean Res. 30 (3) (2008) 172–177.
[11] I. Malekmohamadi, M.R. Bazargan-Lari, R. Kerachian, M.R. Nikoo, M. Fallahnia, Evaluating the efficacy of SVMs, BNs, ANNs and ANFIS in wave height prediction, Ocean Eng. 38 (2–3) (2011) 487–497.
[12] L. Cuadra, S. Salcedo-Sanz, J.C. Nieto-Borge, E. Alexandre, G. Rodríguez, Computational intelligence in wave energy: comprehensive review and case study, Renew. Sustain. Energy Rev. 58 (2016) 1223–1246.
[13] J. Berbić, E. Ocvirk, D. Carević, G. Lončar, Application of neural networks and support vector machine for significant wave height prediction, Oceanologia 59 (3) (2017) 331–349.
[14] S. Hadadpour, A. Etemad-Shahidi, B. Kamranzad, Wave energy forecasting using artificial neural networks in the Caspian Sea, Proc. Inst. Civil Eng. Maritime Eng. 167 (1) (2014) 42–52.
[15] K. Mahmoodi, H. Ghassemi, H. Nowruzi, Data mining models to predict ocean wave energy flux in the absence of wave records, Sci. J. Maritime Univ. Szczecin 49 (121) (2017) 119–129.
[16] A. Etemad-Shahidi, J. Mahjoobi, Comparison between M5′ model tree and neural networks for prediction of significant wave height in Lake Superior, Ocean Eng. 36 (15–16) (2009) 1175–1181.
[17] M.H. Kazeminezhad, A. Etemad-Shahidi, S.J. Mousavi, Application of fuzzy inference system in the prediction of wave parameters, Ocean Eng. 32 (14–15) (2005) 1709–1725.
[18] G.-B. Huang, Q.-Y. Zhu, C.-K. Siew, Extreme Learning Machine: a New Learning Scheme of Feedforward Neural Networks, 2004 IEEE International Joint Conference on Neural Networks (IEEE Cat. No.04CH37541), 2004, pp. 985–990.
[19] R. Rajesh, J.S. Prakash, Extreme learning machines-a review and state-of-the-art, Int. J. Web Based Communities 1 (1) (2011) 35–49.
[20] G.-B. Huang, What are extreme learning machines? Filling the gap between frank rosenblatt's dream and john von Neumann's puzzle, Cogn. Comput. 7 (3) (2015) 263–278.
[21] N.K. Kumar, R. Savitha, A. Al Mamun, Ocean wave height prediction using ensemble of Extreme Learning Machine, Neurocomputing 277 (2018) 12–20.
[22] L. Cornejo-Bueno, J.C. Nieto-Borge, P. García-Díaz, G. Rodríguez, S. Salcedo-Sanz, Significant wave height and energy flux prediction for marine energy applications: a grouping genetic algorithm – extreme Learning Machine approach, Renew. Energy 97 (2016) 380–389.
[23] S. Salcedo-Sanz, R.C. Deo, L. Cornejo-Bueno, C. Camacho-Gómez, S. Ghimire, An efficient neuro-evolutionary hybrid modelling mechanism for the estimation of daily global solar radiation in the Sunshine State of Australia, Appl. Energy 209 (2018) 79–94.
[24] R.C. Deo, N.J. Downs, J.F. Adamowski, A.V. Parisi, Adaptive neuro-fuzzy inference system integrated with solar zenith angle for forecasting sub-tropical photosynthetically active radiation, Food Energy Sec. 8 (1) (2019), e00151.
[25] M. Ali, R. Prasad, Significant wave height forecasting via an extreme learning machine model integrated with improved complete ensemble empirical mode decomposition, Renew. Sustain. Energy Rev. 104 (2019) 281–295.
[26] R.C. 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) 141–166.
[27] M. Ali, R.C. Deo, N.J. Downs, T. Maraseni, Multi-stage committee based extreme learning machine model incorporating the influence of climate parameters and seasonality on drought forecasting, Comput. Electron. Agric. 152 (2018a) 149–165.
[28] M. Ali, R.C. Deo, N.J. Downs, T. Maraseni, Multi-stage hybridized online sequential extreme learning machine integrated with Markov chain Monte Carlo copula-Bat algorithm for rainfall forecasting, Atmos. Res. 213 (2018b) 450–464.
[29] S. Ghimire, R.C. Deo, N.J. Downs, N. Raj, 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, Rem. Sens. Environ. 212 (2018) 176–198.
[30] R. Prasad, R.C. Deo, Y. Li, T. Maraseni, Soil moisture forecasting by a hybrid machine learning technique: ELM integrated with ensemble empirical mode decomposition, Geoderma 330 (2018) 136–161.
[31] S. Ghimire, R.C. Deo, N. Raj, J. Mi, Deep learning neural networks trained with MODIS satellite-derived predictors for long-term global solar radiation prediction, Energies 12 (12) (2019) 2407.
[32] G. Bargshady, X. Zhou, R.C. Deo, J. Soar, F. Whittaker, H. Wang, Enhanced deep learning algorithm development to detect pain intensity from facial expression images, Expert Syst. Appl. 149 (2020), https://doi.org/10.1016/j.eswa.2020.113305.
[33] M. Moishin, R.C. Deo, R. Prasad, N. Raj, S. Abdulla, Designing deep-based learning flood forecast model with ConvLSTM hybrid algorithm, IEEE Access 9 (2021) 50982–50993.
[34] M.T. Sattari, H. Apaydin, S.S. Band, A. Mosavi, R. Prasad, Comparative analysis of kernel-based versus ANN and deep learning methods in monthly reference evapotranspiration estimation, Hydrol. Earth Syst. Sci. 25 (2) (2021) 603–618.
[35] B. Bhattacharya, D.P. Solomatine, Neural Networks and M5 Model Trees in Modelling Water Level Discharge Relationship for an Indian River, European Symposium on Artificial Neural Networks, Bruges (Belgium), ESANN'2003 proceedings, 2003.
[36] D.P. Solomatine, K.N. Dulal, Model trees as an alternative to neural networks in rainfall—runoff modelling, Hydrol. Sci. J. 48 (3) (2003) 399–411.
[37] D.P. Solomatine, M.B.L.A. Siek, Flexible and Optimal M5 Model Trees with Applications to Flow Predictions, 6th International Conference on Hydro-informatics. P. B. Liong, World Scientific Publishing Company, 2004.
[38] D.P. Solomatine, Y. Xue, M5 model trees and neural networks Application to flood forecasting in the upper reach of the Huai River in China, J. Hydrol. Eng. 9 (6) (2004) 491–501.
[39] B. Bhattacharya, D.P. Solomatine, Neural networks and M5 model trees in modelling water level–discharge relationship, Neurocomputing 63 (2005) 381–396.
[40] S.N. Londhe, P.R. Dixit, Forecasting stream flow using support vector regression and M5 model trees, Int. J. Eng. Res. Dev. 2 (5) (2012) 1–12.
[41] E.K. Onyari, F.M. Ilunga, Application of MLP neural network and M5P model tree in predicting streamflow_South Africa, Int. J. Innov. Manag. Technol. 4 (1) (2013) 11–15.
[42] M.T. Sattari, M. Pal, H. Apaydin, F. Ozturk, M5 model tree application in daily river flow forecasting in Sohu Stream, Turkey, Water Resour. 40 (3) (2013) 233–242.
[43] G.-B. Huang, Q.-Y. Zhu, C.-K. Siew, Extreme learning machine: theory and applications, Neurocomputing 70 (1–3) (2006) 489–501.
[44] S.i. Tamura, M. Tateishi, Capabilities of a four-layered feedforward neural network: four layers versus three, IEEE Trans. Neural Network. 8 (2) (1997) 251–255.
[45] G.-B. Huang, Learning capability and storage capacity of two-hidden-layer feedforward networks, IEEE Trans. Neural Network. 14 (2) (2003) 274–281.

[46] K. Gopalakrishnan, S.K. Khaitan, A. Choudhary, A.J.C. Agrawal, B. Materials, Deep Convolutional Neural Networks with transfer learning for computer vision-based data-driven pavement distress detection 157 (2017) 322–330.
[47] A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, Commun. ACM 60 (6) (June 2017) 84–90, https://doi.org/10.1145/3065386.
[48] T. Chen, D. Borth, T. Darrell, S. Chang, Deepsentibank: visual sentiment concept classification with deep convolutional neural networks, 2014.
[49] W. Hu, Y. Huang, L. Wei, F. Zhang, H.J. J.o.S. Li, Deep Convolutional Neural Networks for Hyperspectral Image Classification, 2015, 2015.
[50] J. Dou, C. Liu, B. Wang, Short-term wind power forecasting based on convolutional neural networks. IOP Conference Series: Earth and Environmental Science, IOP Publishing, 2018.
[51] Y. Yu, J. Cao, J. Zhu, An LSTM short-term solar irradiance forecasting under complicated weather conditions, IEEE Access 7 (2019) 145651–145666.
[52] H. Zhou, Y. Zhang, L. Yang, Q. Liu, K. Yan, Y. Du, Short-term photovoltaic power forecasting based on long short term memory neural network and attention mechanism, IEEE Access 7 (2019) 78063–78074.
[53] J. Le, H. El-Askary, M. Allali, D.C.J.A.R. Struppa, Application of recurrent neural networks for drought projections in California 188 (2017) 100–106.
[54] A.G. Salman, B. Kanigoro, Y. Heryadi, Weather Forecasting Using Deep Learning Techniques, 2015 international conference on advanced computer science and information systems (ICACSIS), IEEE, 2015.
[55] S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural Comput. 9 (8) (1997) 1735–1780.
[56] N. Draper, H. Smith, Applied Regression Analysis, John Wiley, New York, 1981, p. 709.
[57] D.C. Montgomery, E.A. Peck, G.G. Vining, Introduction to Linear Regression Analysis, John Wiley & Sons, 2012.
[58] G. Civelekoglu, N. Yigit, E. Diamadopoulos, M. Kitis, Prediction of bromate formation using multi-linear regression and artificial neural networks, Ozone Sci. Eng. 29 (5) (2007) 353–362.
[59] M. Şahin, Y. Kaya, M. Uyar, Comparison of ANN and MLR models for estimating solar radiation in Turkey using NOAA/AVHRR data, Adv. Space Res. 51 (5) (2013) 891–904.
[60] A. Apaydın, A. Kutsal, C. Atakan, Ankara, The Statistics in Practice, Hacettepe Pub, 1994.
[61] K. Ozdamar, The Statistical Data Analysis with Software Packages, Kaan press, Eskis ehir, 2004.
[62] C. Liu, D.B.J.B. Rubin, The ECME algorithm: a simple extension of EM and ECM with faster monotone convergence 81 (4) (1994) 633–648.
[63] X.-L. Meng, D.B.J.B. Rubin, Maximum likelihood estimation via the ECM algorithm: a general framework 80 (2) (1993) 267–278.
[64] A.P. Dempster, N.M. Laird, D.B. Rubin, Maximum likelihood from incomplete data via the EM algorithm, J. Roy. Stat. Soc. B 39 (1) (1977) 1–22. https://www.jstor.org/stable/2984875.
[65] J.R. Quinlan, Learning with Continuous Classes, 5th Australian joint conference on artificial intelligence, Singapore, 1992.
[66] T.M. Mitchell, Machine Learning, ser." Computer Science Series. Singapure, McGraw-Hill Companies, Inc, 1997.
[67] A. Rahimikhoob, M. Asadi, M. Mashal, A comparison between conventional and M5 model tree methods for converting pan evaporation to reference evapotranspiration for semi-arid region, Water Resour. Manag. 27 (14) (2013) 4815–4826.
[68] O. Kisi, Pan evaporation modeling using least square support vector machine, multivariate adaptive regression splines and M5 model tree, J. Hydrol. 528 (2015) 312–320.
[69] A. Schiller, P.R. Oke, Development of operational oceanography and servicing in Australia, J. Res. Pract. Inf. Technol. 39 (2) (2007) 151–164.
[70] T.H. van Andel, G.R. Heath, T.C. Moore, D.F.R. McGeary, Late Quaternary history, climate, and oceanography of the Timor Sea, northwestern Australia, Am. J. Sci. 265 (9) (November 1967) 737–758, https://doi.org/10.2475/ajs.265.9.737.
[71] Queensland, Environment and science. https://data.qld.gov.au/dataset/coastal-data-system-waves-gold-coast, 2018.
[72] J.A. Battjes, H.W. J.C.e. Groenendijk, Wave height distributions on shallow foreshores 40 (3) (2000) 161–182.
[73] I.J. Moon, I. Ginis, T. Hara, H.L. Tolman, C.W. Wright, E.J. Walsh, Numerical Simulation of Sea Surface Directional Wave Spectra under Hurricane Wind Forcing, J. Phys. Oceanogr. 33 (8) (2003) 1680–1706, https://doi.org/10.1175/2410.1.
[74] Y.Y. Chao, L.D. Burroughs, H.L. Tolman, The North Atlantic Hurricane Wind Wave Forecasting System (NAH), National Weather ServiceOffice of MeteorologyTechnical Procedures BulletinSeries No. 478, 2004. https://polar.ncep.noaa.gov/mmab/tpbs/operational.tpbs/tpb478/tpb478.htm.
[75] ASCE, Criteria for evaluation of watershed models, J. Irrigat. Drain. Eng. 119 (3) (1993) 429–442.
[76] B.C. Yen, Discussion and closure: criteria for evaluation of watershed models, J. Irrigat. Drain. Eng. 121 (1) (1995) 130–132.
[77] ASCE, Artificial neural networks in hydrology. II: hydrologic applications, J. Hydrol. Eng. 5 (2) (2000) 124–137.
[78] C.J. Willmott, On the validation of models, Phys. Geogr. 2 (2) (1981) 184–194.
[79] C.J. Willmott, Some comments on the evaluation of model performance, Bull. Am. Meteorol. Soc. 63 (11) (1982) 1309–1313.
[80] C.J. Willmott, On the Evaluation of Model Performance in Physical Geography, Spatial statistics and models, Springer, 1984, pp. 443–460.
[81] D.R. Legates, G.J. McCabe, Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation, Water Resour. Res. 35 (1) (1999) 233–241.
[82] 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. Model. Software 22 (7) (2007) 1034–1052.
[83] 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.
[84] C.-W. Hsu, C.-C. Chang, C.-J. Lin, A practical guide to support vector classification, 2003.
[85] A.Y. Shamseldin, Application of a neural network technique to rainfall runoff, J. Hydrol. 199 (1997) 272–294.
[86] Z. Xu, Z. Hou, Y. Han, W. Guo, A diagram for evaluating multiple aspects of model performance in simulating vector fields, Geosci. Model Dev. (GMD) 9 (12) (2016) 4365–4380.
[87] G. Kim, W.M. Jeong, K.S. Lee, K. Jun, M.E.J.E. Lee, Offshore and nearshore wave energy assessment around the Korean Peninsula 36 (3) (2011) 1460–1469.
[88] B. Cahill, T. Lewis, Wave periods and the calculation of wave power, 2014.
[89] H. Yavuz, T.J. Stallard, A.P. McCabe, G.A. Aggidis, Time series analysis-based adaptive tuning techniques for a heaving wave energy converter in irregular seas, Proc. IME J. Power Energy (2007), https://doi.org/10.1243/09576509JPE291.
[90] A. Pecher, J.P. Kofoed, Handbook of Ocean Wave Energy, Springer, London, 2017.
[91] D.E. Birba, A Comparative Study of Data Splitting Algorithms for Machine Learning Model Selection, 2020.
[92] J. Quilty, J.J. J.o. h. Adamowski, 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 563 (2018) 336–353.
[93] B. Cannas, A. Fanni, L. See, G. Sias, Data preprocessing for river flow forecasting using neural networks: wavelet transforms and data partitioning, Phys. Chem. Earth, Parts A/B/C 31 (18) (2006) 1164–1171.
[94] C. Bergmeir, J.M. Benítez, On the use of cross-validation for time series predictor evaluation, Inf. Sci. 191 (2012) 192–213.
[95] E.M. Smith, D.R. Sewell, P.T. Golden, System and Method for Energy Management, Google Patents, 2004.
[96] B.L. Capehart, W.C. Turner, W.J. Kennedy, Guide to Energy Management, The Fairmont Press, Inc, 2006.
[97] M.E. Bauer, The role of remote sensing in determining the distribution and yield of crops, Adv. Agron. 27 (1975) 271–304.
[98] C. Chen, H. McNairn, A neural network integrated approach for rice crop monitoring, Int. J. Rem. Sens. 27 (7) (2006) 1367–1393.
[99] D. Stathakis, I. Savin, T. Nègre, Neuro-fuzzy modeling for crop yield prediction, Int. Arch. Photogram. Rem. Sens. Spatial Inf. Sci. 34 (XXX) (2006) p1–4.
[100] J. Dempewolf, B. Adusei, I. Becker-Reshef, M. Hansen, P. Potapov, A. Khan, B. Barker, Wheat yield forecasting for Punjab Province from vegetation index time series and historic crop statistics, Rem. Sens. 6 (10) (2014) 9653–9675.
[101] P. Kumar, D.K. Gupta, V.N. Mishra, R. Prasad, Comparison of support vector machine, artificial neural network, and spectral angle mapper algorithms for crop classification using LISS IV data, Int. J. Rem. Sens. 36 (6) (2015) 1604–1617.