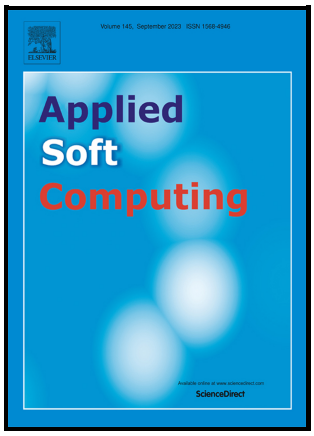
Journal Pre-proof

# Hybrid Deep Learning Model for Wave Height Prediction in Australia's Wave Energy Region

Abul Abrar Masrur Ahmed, S Janifer Jabin Jui, Mohanad S. AL-Musaylh, Nawin Raj, Reepa Saha, Ravinesh C Deo, Sanjoy Kumar Saha




PII: S1568-4946(23)01021-9

DOI: https://doi.org/10.1016/j.asoc.2023.111003

Reference: ASOC111003

To appear in: *Applied Soft Computing*

Received date: 10 June 2023
Revised date: 16 September 2023
Accepted date: 29 October 2023

Please cite this article as: Abul Abrar Masrur Ahmed, S Janifer Jabin Jui, Mohanad S. AL-Musaylh, Nawin Raj, Reepa Saha, Ravinesh C Deo and Sanjoy Kumar Saha, Hybrid Deep Learning Model for Wave Height Prediction in Australia's Wave Energy Region, *Applied Soft Computing,* (2023) doi:https://doi.org/10.1016/j.asoc.2023.111003

# Hybrid Deep Learning Model for Wave Height Prediction in Australia's Wave Energy Region

**Abul Abrar Masrur Ahmed[a], S Janifer Jabin Jui[b], Mohanad S. AL-Musaylh[c], Nawin Raj[b], Reepa Saha[d], Ravinesh C Deo[b], Sanjoy Kumar Saha[e]**

[a]Department of Infrastructure Engineering, The University of Melbourne, Victoria 3010 Australia

[b] School of Mathematics Physics and Computing, University of Southern Queensland, Springfield QLD 4300, Australia.

[c] Department of Information Technologies, Management Technical College, Southern Technical University, Basrah 61001, Iraq

[d] Department of Electrical and Computer Engineering, University of Alabama at Birmingham Birmingham, Alabama, United States

[e] LOS Cable Solutions, Bandadalen 17, 5417 Stord, Vestland, Norway.

E-mail addresses: abulmasrur.ahmed@unimelb.edu.au (A. A. Masrur Ahmed)
sjanifer.jui@usq.edu.au (S. J. J. Jui)
mohanad.al-musaylh@stu.edu.iq (M. Al-Musaylh)
nawin.raj@usq.edu.au (N. Raj)
reepa@uab.edu (R. Saha)
ravinesh.deo@usq.edu.au (R. C. Deo)
Sanjoy.saha@los-cablesolutions.no (S. K. Saha)

**Corresponding Author*:** Prof Ravinesh Deo (ravinesh.deo@usq.edu.au)



## Abstract

Waves are emerging as a renewable energy resource, but the harnessing of such energy remains among the least developed in terms of renewable energy technologies on a regional or a global basis. To generate usable energy, wave heights must be predicted in near-real-time, which is the driving force for wave energy converters. This study develops a hybrid Convolutional Neural Network-Long Short-Term Memory-Bidirectional Gated Recurrent Unit forecast system (CLSTM-BiGRU) trained to accurately predict significant wave height ($H_{sig}$) at multiple forecasting horizons (30 minutes, 0.5*H*; 2 hours, 02*H*; 3 hours, 03*H* and 6 hours, 06*H*. In this model, convolutional neural networks (CNNs), long-short-term memories (LSTMs), and bidirectional gated recurrent units (BiGRUs) are employed

to predict $H_{sig}$. To construct the proposed CLSTM-BiGRU model, historical wave properties, including maximum wave height, zero-up crossing wave period, peak energy wave period, sea surface temperature, and significant wave heights are analysed. Several wave energy generation sites in Queensland, Australia were tested using the hybrid deep learning CLSTM-BiGRU model. Based on statistical score metrics, scatterplots, and error evaluations, the hybrid CLSTM-BiGRU model generates more accurate forecasts than the benchmark models. This study established the practical utility of the hybrid CLSTM-BiGRU model for modelling $H_{sig}$ and therefore shows the model could have significant implications for wave and ocean energy generation systems, tidal or wave height monitoring as well as sustainable wave energy resource evaluation where a prediction of wave heights is required.

**Keywords** deep learning model; significant wave height; wave energy; renewable energy; sea level monitoring system

## 1.0 Introduction

Global warming has become one of the world's most critical issues today. In the last decade, the global mean surface temperature (GMST) was over 1.2 °C higher than the pre-industrial baseline [1]. Increasing temperatures worldwide have been linked to climate change, and severe and frequent extreme weather events, such as droughts and bushfires [1-5]. In Australia, the catastrophic wild bushfire in 2019-20 caused due to climate change wiped away 17 million hectares of land, 3000 houses, 33 people, and over a billion wild animals across New South Wales (NSW), Victoria, ACT, Western Australia, and South Australia [6, 7]. The use of fossil fuels to produce energy contributes to the increase in significant gases (GHG) [5], [8]. This affects not only available fossil fuel resources and the increase in global temperature but also the population's health due to breathing in poor-quality air [9, 10]. Isphording and Pestel [11] have also explored the short-term exposure to ambient air pollutants on the spread of Covid-19 and discovered a positive effect on death numbers. With the increase in urbanization, 40% of the total world’s energy consumption is for lighting, cooling, and heating [5]. Energy consumption is expected to increase

over time, as will the risk of extreme weather events. To mitigate this problem, it is essential to carry out effective fuel management and find more economical and sustainable alternatives to fossil fuels [12].

Over the last few decades, scientists and researchers have explored renewable energy sources such as solar, hydroelectric, and wind power resources. In 2020, Australia generated 24% of its total electricity from renewable sources, of which 9% came from solar, 9% from wind, and only 6% from hydropower [13]. The ocean surrounding Australia provides a potential source of wave energy, as waves travel long distances without losing energy in deep water [14], and is a cost-effective alternative to conventional energy. However, oceanic waves are variable, often unpredictable, and are affected by various natural phenomena. Therefore, understanding the correlated variables and utilizing them in a forecasting model is the key to using this energy.

One component essential for wave energy is significant wave height ($H_{sig}$). $H_{sig}$, the mean wave height of the highest third of the waves, is computed by the difference between the wave crest and the preceding wave through [15, 16]. Accurate prediction of $H_{sig}$ is also vital for safe operations in marine and offshore environments [17], installation of wind turbines, cargo transfer, rescue missions [18], marine and coastal engineering [19], and energy generation [20], to name a few. Therefore, numerous data-driven forecasting models have been developed to estimate $H_{sig}$ [20-23].

Forecasting and modelling of $H_{sig}$ are still in their infancy since coastal waves are unpredictable, non-stationary, and nonlinear [24]. Numerous studies have examined algorithms such as k-nearest neighbours (kNN), linear regressions (LR), model trees (M5), multilayer perceptron neural networks (MLP), robust regressions, and support vector regressions (SVR) [25-27]. Özger [28] has concluded wavelet fuzzy logic approach (WFL) outperforms when compared with artificial neural network (ANN) and autoregressive moving average (ARMA). Cuadra et al. [29] compared MLP and other regression models with ANN and found out ANN performs better than the regression models. In the prediction of $H_{sig}$,, Etemad-Shahidi, and Mahjoobi [19] $_g$ depicted ANN are less transparent than semi-empirical regression-based models like M5 algorithm as neural network (NN) requiring more time to find network parameters such as the number of hidden layers and neurons through trial and error. Furthermore, these studies have utilized standalone

models, which are sometimes unsuitable for analysing complex inputs and stochastic features within the data.

To resolve this issue, several studies have used hybrid forecasting models [30-34]. James et al. [35] trained a machine learning model to act for a physical-based SWAN (Simulating WAves Nearshore) model representing the significant wave height field, and an SVM model simulated the characteristic period. The hybrid machine learning (ML) model, integrated with the extreme gradient boosting model (XGBoost) and decision tree (DT), has performed significantly better than other standalone ML models [36]. Ali et al. [24] have introduced an extreme learning machine (ELM) named improved complete ensemble empirical mode decomposition with adaptive noise (ICEEMDAN) by incorporating a Gaussian white noise with ensemble-EMD to eliminate the mode mixing issue in EMD; again, it was not entirely noise-free for which complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) was proposed [37]. The data decomposition is time-consuming and, therefore, unsuitable for analysing more extended time series. Furthermore, conventional machine learning approaches can suffer from data overfitting issues, especially with large datasets [38], and do not identify the short- and long-term correlations between the predictors and the target [15], which can be overcome by using deep learning (DL) approaches.

DL hybrid models for forecasting have become extremely popular [39-44]. Deep learning models can extract deep features and multidimensional dependencies to generate better predictions [45, 46]. This technology is also easy to use, automated, and capable of analysing large amounts of data that would otherwise require computationally expensive methods [47, 48]. The study is based on a hybrid model that combines a convolutional neural network (CNN), a long short-term memory (LSTM), and a bidirectional gated recurrent unit (BiGRU). Numerous studies have demonstrated that CNN outperforms many existing machine-learning methods in forecasting applications [49, 50].

Accordingly, this study employs CNN to extract features to improve prediction accuracy [51]. LSTM and BiGRU (an improved version of LSTM requiring less training, thus timesaving) are variants of recurrent neural networks (RNN) that can avoid short-term memory issues related to gradient vanishing

and explosion revealing information in time-series data. There have been some recent studies hybridizing CNN and GRU in week-ahead evapotranspiration forecasting [52], CNN and LSTM in solar radiation and air quality forecast [51, 53], and CNN-BiGRU in load forecasting [54]. In this study, LSTM and BiGRU are employed for prediction purposes, combined with CNN for feature extraction. Notably, this hybrid approach has not been used before in any other studies, particularly in forecasting significant wave heights, $H_{sig}$.

To build a reliable model, this paper examines the complex, stochastic patterns of oceanic significant wave height ($H_{sig}$), as well as the maximum wave height ($H_{max}$), zero up crossing wave period ($T_z$), peak energy wave period ($T_p$), sea surface temperature ($SST$) to forecast $H_{sig}$ over relatively short time intervals (i.e., half-hourly, two hours, three hours, and six hours) in Queensland, Australia. The proposed hybrid model CLSTM-BiGRU has been tested against standalone models and their respective hybrid models. This study, therefore, examines the efficacy of the proposed model at four key study sites (i.e., Cairns, Emu Park, Mooloolaba, and Gold Coast) in Queensland, Australia, as these locations may be used to identify probable spots for future wave energy generation, adding more renewable energy to the main transmission systems and achieving energy sustainability.

## 2.0 Theoretical Overviews of Predictive Models

While Convolutional Neural Networks (CNN) are well-known for working on spatial or 2D image datasets, they can also extract hidden features from time-series data and generate filters capturing those features in predictive models [43]. The CNN works more like a regularized version of the feed-forward neural network (NN) for solving one-dimensional problems (time series classification and prediction). As part of the extraction process, three mapping layers are typically applied: the convolutional layer, the pooling layer, and the fully connected layer. An LSTM network has been used to interpret wave height features based on the extracted feature information from CNN across time steps. Consequently, the combination of two sub-models: CNN and LSTM, has been employed together as C-LSTM to produce better performance in time series data problems, such as wave height predictions. Convolutional filtering is generally used in the convolutional layer to extract potential features. In

addition to reducing the size of the series, the pooling layer preserves the essential characteristics identified by the convolutional layer. In this layer, the objective variables are estimated based on the features of the predictor variable. Each convolutional layer is defined as follows:

$$h_{ij}^{k} = f((W^{k} * x)_{ij}) + b_{k} \quad (1)$$

where $f$ in equation (1) denotes the activation function, $W^{k}$ is the weight of the kernel connected to $k^{th}$ feature map and the star $(*)$ is an operator of the convolutional process.

The two most popular variants of recurrent neural networks (RNN) are the long-short-term memory neural network (LSTM) and the gated recurrent unit (GRU). Both variants can capture the temporal characteristics of the prediction problem, which avoids short-term memory issues related to gradient vanishing and explosion, as well as reveal the intrinsic association between time series data [55]. This network comprises an underlying component known as the memory cell, which can memorize the temporal state using three types of gates: input, forget, and output [55]. The input gate activation tracks the input information stored in the memory cell. In contrast, the output gate can control the dissemination of the latest information to the ultimate state. The function of the forget gate is to determine unimportant information and forget that information from the training data. Additionally, there is another update gate to update the cell. These four gates together help regulate the information flow.

To implement, update the LSTM cell state, and compute the LSTM outputs, the equations (2)–(9) are required to follow [55].

$$F_t = \sigma(W_{xf}X_t + W_{hf}H_{t-1} + B_f) \quad (2)$$

$$I_t = \sigma(W_{xi}X_t + W_{hi}H_{t-1} + B_i) \quad (3)$$

$$\bar{C}_t = \sigma(W_{xc}X_t + W_{hc}H_{t-1} + B_c) \quad (4)$$

$$C_t = F_t * C_{t-1} + I_t * \bar{C}_t) \quad (5)$$

$$O_t = \sigma(W_{xo}X_t + W_{ho}H_{t-1} + B_0) \quad (6)$$

$$H_t = O_t \tanh(C_t) \quad (7)$$

$$Y_t = \sigma(W_{hy}XH_t + B_y) \quad (8)$$

$$\sigma_x = \frac{1}{1+e^{-x}} \quad (9)$$

where $X_t$, $Y_t$, $I_t, F_t$, $O_t, C_t, \overline{C_t}$, and σ represent input vector, output vector, input gate outcome, forget gate outcome, output gate outcome, finishing state in a memory block, temporary, and sigmoid function. $W_{xf}$, $W_{xi}$, $W_{xc}$, and $W_{xo}$ symbolize input weight matrices and $W_{hf}$, $W_{hi}$, $W_{hc}$, and $W_{ho}$ are recurrent weight matrices in equations (2)- (5), $W_{hy}$ is output weight matrix in equation (8) and $B_f$, $B_i$, $B_c$, $B_0$, and $B_y$ are the related bias vectors for equations (2)- (5),(6),(8).

The sigmoid function's output values range from 0 to 1, allowing the neural network to remove unrelated information. The hybrid CNN-LSTM (or CLSTM) configuration, incorporates one convolutional layer, one max pooling layer, a flattened layer, an LSTM layer, and a fully connected layer with the output to reduce raw data features using conventional filters [56]. The wave height prediction result using CLSTM has shown high accuracy and better prediction performance than the standalone LSTM or CNN network. Despite the superior performance, this hybrid CLSTM configuration is relatively complex compared to other individual configurations.

Gated Recurrent Unit (GRU), a modification of the LSTM concept, requires less training and computational time but provides improved network performance. Moreover, GRU combines the hidden and cell states into one state; therefore, it has fewer parameters. Thus, the total number of gates in GRU is half compared to the total number of LSTM gates, making GRU popular and a shortened variant of the LSTM cell. To define the relationship between predictors and predictands in a GRU Network, two input features, i.e., input vector $x(t)$ and output vector $h(t-1)$, need to be considered in each layer [57]. The outcome of each gate can be generated employing logical operation (pointwise multiplication and addition) and nonlinear transformation of predictors using equations (10)-(13), as shown in Fig 1. The equations are defined as follows:

$$r(t) = \sigma_g (W_r x(t) + U_r h(t-1) + B_r) \quad (10)$$

$$z(t) = \sigma_g (W_z x(t) + U_z h(t-1) + B_z) \quad (11)$$

$$h(t) = (1 - z(t)) o (t-1) + z(t) o \overline{h}(t) \quad (12)$$

$$\overline{h}(t) = \sigma_h (W_h x(t) + U_h (r(t)) o h(t-1)) \quad (13)$$

where r(t), z(t), W and U are defined as the reset gate vector, update gate vector, parameter metrics and vector respectively. $\sigma_h$, and $\sigma_g$ are signified as a hyperbolic tangent, and a sigmoid function.

In order to forecast the height of waves with confidence, a forecasting model must be able to extract both the implicit features and the complex variances within the sequence data. It must be noted, however, that the GRU can only extract information from the forward direction. Therefore, a model must draw valuable information from backward time series data. To extract information from both directions, the Bidirectional GRU, or BiGRU, is implemented effectively to encapsulate knowledge between production variance and input variables. The BiGRU is a sequence processing model comprising two GRUs. Out of two GRUs, one GRU takes the input in a forward direction and the other in a backward direction. It is a bidirectional recurrent neural network with only the input and forgets gates. According to the proposed Bi-GRU model, bi-directional regularities can be depicted between multiple inputs and outputs, and it could be used to investigate the mechanism of stimulation performance based on relevant production data.

In both LSTM and BiGRU algorithms, gates control the memory process; GRU uses fewer training parameters, requires less memory, and is faster than LSTM, while LSTM is more accurate on a large dataset. The BiGRU shows efficacy when past and future information is required to be incorporated into production sequences. We used a classical machine learning model as a baseline, random forest (RF) model, a popular supervised machine-learning algorithm, can accumulate predictors associated with different values of random vectors sampled independently [58]. This model trains several trees (decision tree 1, 2, …, N), in parallel and uses the majority voting/ averaging of the trees as the final prediction or results of the RF model. This model adopts a bagging-type ensemble (collection). A randomly selected sample is assigned to each split node that obtains a better prediction result with a higher accuracy rate and avoids overfitting. The individual decision tree model is easy to interpret. Still, the model is nonunique and exhibits high variance. Equation (14) calculates the predicted values for unseen complexes:

$$y = \frac{1}{B}\sum_{b=1}^{B} t_B(x) \quad (14)$$

where $B$ represents the number of data points, $t_B(x)$ portrays the result of $(f_b - y_b)^2$. $f_b$ is showing the value returned by the model and $y_b$ is the actual value for datapoint b.

The hybrid architecture of RNN and convolutional neural network (CNN) has emerged mainly to capture the temporal correlation of data along with extracting features from a given dataset, e.g., high-resolution images or tensor concurrently in addition to classifying or making predictions. Exploiting CNN and LSTM collectively, the CLSTM neural network is proposed to handle the input data containing many features efficiently. It is noteworthy that time-series data usually are lengthy due to the high sampling frequency of digital signal devices nowadays, which will be facilitated by feature extraction via convolutional layers.

Therefore, we propose a novel CLSTM-BiGRU-based deep learning hybrid model, which takes advantage of the intrinsic features of CLSTM neural networks and a bidirectional GRU, or BiGRU, to forecast significant wave height in this study. In this hybrid approach, statistical methods are combined with machine learning methods to compensate for the limitations of one approach with the strengths of the other, especially in forecasting time series data.

## 3.0 Study Area and Data Description

This study aims to develop a set of forecasting models based on the wave time-series recorded at four data collection stations in Queensland. The study used a 30-minute interval dataset of recorded wave parameters (see Table 1) from 2015 to 2021. Table 2 and Fig. 1(a) show the selected sites and their geographical locations. Since the data-driven models depend on predictive features in historical data for future forecasting, wave features are used for significant wave height forecasting.

**Table 1** The model input parameters and their descriptions including units.

| **Wave Property Parameters** | | **Descriptions** | **Unit** |
|---|---|---|---|
| **Time-Lagged Combinations of Predictor (Input) Variables** (see Fig. 2) | $H_{max}$ | Maximum Wave Height | m |
| | $T_z$ | Zero Up Crossing Wave Period | Seconds |
| | $T_p$ | Peak Energy Wave Period | Seconds |
| | *SST* | Sea Surface Temperature | $^0$C |

| | | | |
|---|---|---|---|
| | $H_{sig}$ | Significant Wave Height | m |
| **Objective Target Variable** | $H_{sig}$ | Significant Wave Height | m |

**Table 2** The geographical location of the study sites where the proposed **CLSTM-BiGRU was constructed.**

| Study Site | Geographical Location |
|---|---|
| Gold Coast | 27° 57' 53" S, 153° 20' 58" E |
| Cairns | 16° 55' 34" S, 145° 46' 27" E |
| Mooloolaba | 26° 40' 53" S, 153° 07' 09" E |
| Emu Park | 23° 15' 25" S, 150° 49' 35" E |

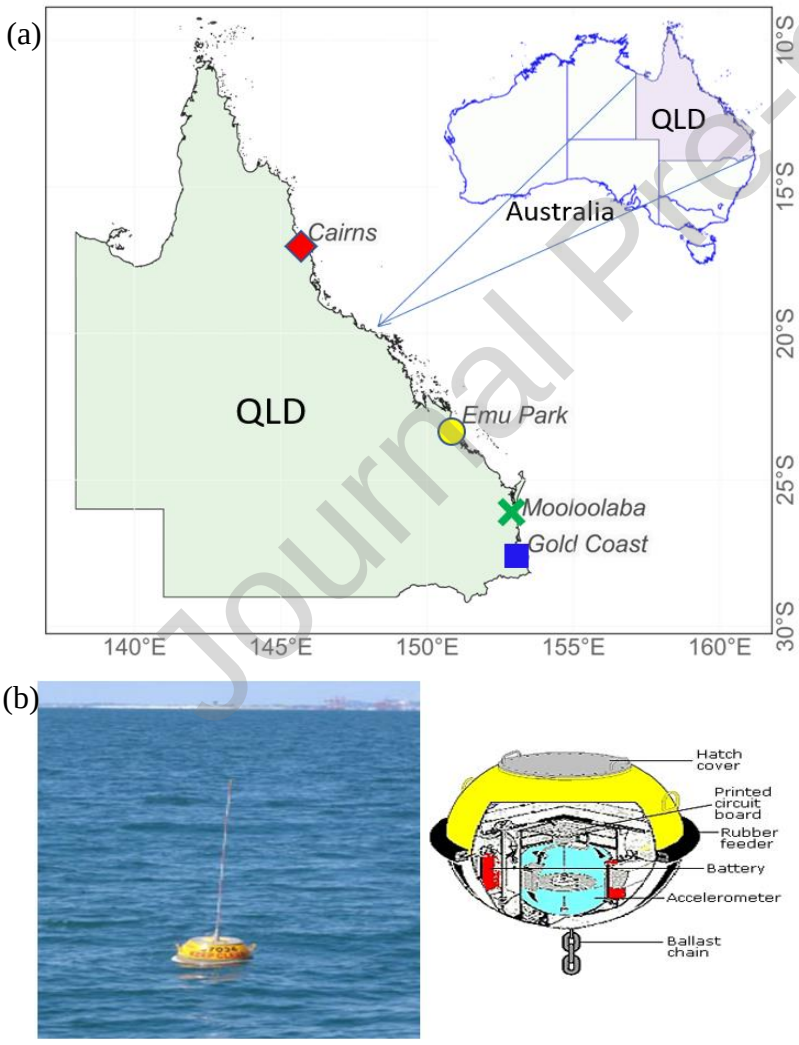



**Fig. 1. (a)** Geographic location of the present sites studied to build the CLSTM-BiGRU-based hybrid deep learning model to forecast significant wave height was developed: Cairns, Emu Park,

Mooloolaba, and Gold Coast located in Queensland, Australia. **(b)**. The monitoring buoys that are employed in the Queensland waters where data were collected.

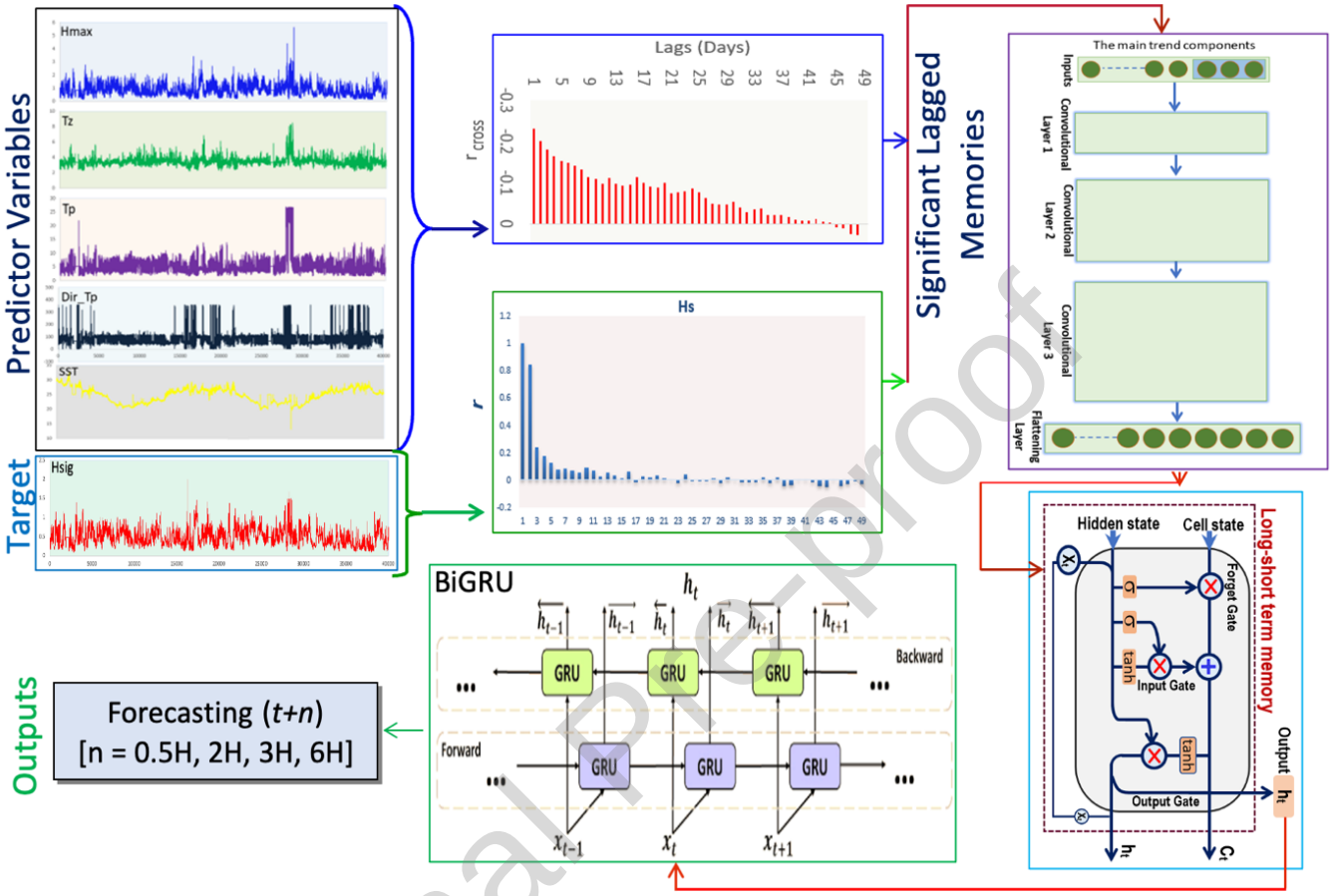



**Fig. 2** Schematic workflow of CLSTM-BiGRU model development and architecture of the convolutional neural network (CNN), long-short term memory (LSTM), and bidirectional gated recurrent unit (BiGRU) to forecast significant wave height at multi-step horizons. Note that the model is constructed using oceanic significant wave height ($H_{sig}$), as well as the maximum wave height ($H_{max}$), zero up crossing wave period ($T_z$), peak energy wave period ($T_p$), sea surface temperature ($SST$) to forecast $H_{sig}$ over relatively short time (i.e., half-hourly, two hourly, three hourly, and six hourly) intervals in Queensland, Australia.

Considering, $H_{sig}$ as the target time series variable for the 30-minute interval, the significant lags are then used with other wave features; maximum wave height ($H_{max}$), zero up crossing wave period ($T_z$), peak energy wave period ($T_p$), and sea surface temperature (SST) to predict the significant wave height $H_{sig}$. It is noted that $H_{sig}$ is generally measured as an average of the third-highest wave in the recording period, and this measurement is based on the hypothesis that smaller waves are not considered because

they are insignificant by the observer. As a rule, these smaller waves do not have much influence on the overall processes of the waves. The $H_{max}$, another property used to develop the proposed model, is defined as the distance between the top of the wave (i.e., the wave crest) and the bottom of the wave (wave trough). In a wave, the $T_z$ parameter indicates the time between two zero-level up-crossings. Basically, SST refers to the temperature close to the surface of the ocean, which is called the 'skin' temperature of the ocean. Surface temperature is generally measured from the range of 1 mm to 20 m from the top. In a wave recording, $T_p$ represents the wave period of the waves giving the most energy. In addition to ocean waves, distant disturbances such as storms can also generate these waves. Using various wave properties to construct the proposed CLSTM-BiGRU-based deep learning hybrid model was a deliberate strategy to maximize the performance of the model used to predict significant wave height.

These wave parameters are monitored continuously by floating buoys located at the study sites in Queensland. During the wave heave recording and processing, the wave heave is recorded and processed electronically. As soon as the data has been collected, it is sent to the nearby station (see Fig. 1b). The station devices, which include a computer, radio receiver, and modem, store and analyse the data. Datasets are sent to the data server for further processing. The buoys are calibrated for twelve months before being deployed in the ocean. The buoy is a stainless-steel device that can range from 0.4 m to 0.9 m and is designed to follow the movement of the wave.

## 4.0 Model Development Procedures

### 4.1 Data Normalization

To improve the model's convergence into its optimal state for best accuracy, the predictors and predictands are normalized to remove both dimensionality and variance of variables. To execute the normalization stage, the minimum and maximum values of each variable, $x_i$ was calculated. For each data sample, $x_j$, the normalization process is denoted in Eq. 15.

$$\bar{\omega}_i = \frac{\omega_i - \min\limits_{1\le j\le n}\{\omega_j\}}{\max\limits_{1\le j\le n}\{\omega_j\} - \min\limits_{1\le j\le n}\{\omega_j\}} \tag{15}$$

where $\omega_i \in \{\omega_1, \omega_2, \cdots, \omega_n\}$ is the original data and $\bar{\omega}_i \in [0,1]$ is the normalized data.

### 4.2 Data Partitioning

This study used data partitioning as a regular method of validating the deep hybrid CLSTM-BiGRU model against independent BiGRU and RF models. By using the PACF and CCF methods, the predictands ($H_{sig}$) are correlated to create the input and target data necessary to build a predictive model (see Fig. 3). It is necessary to divide the input data into training, testing, and validation sets when building predictive models. For the model to learn more about the characteristics of the data over time, it uses a training set, which consists of a collection of data that is repeatedly used during training. The validation process intends to provide information that may be used to adjust the model hyperparameters. Training sets are different from validation sets, which are used to assess and validate the model as it is being trained. The test set is used only after a model has been trained (using train and validation sets) and primarily to evaluate the model. The datasets between Jan 2015 and Aug 2021 at 30 min interval is partitioned as 70% for training, 15% for validation, and 15% for testing.

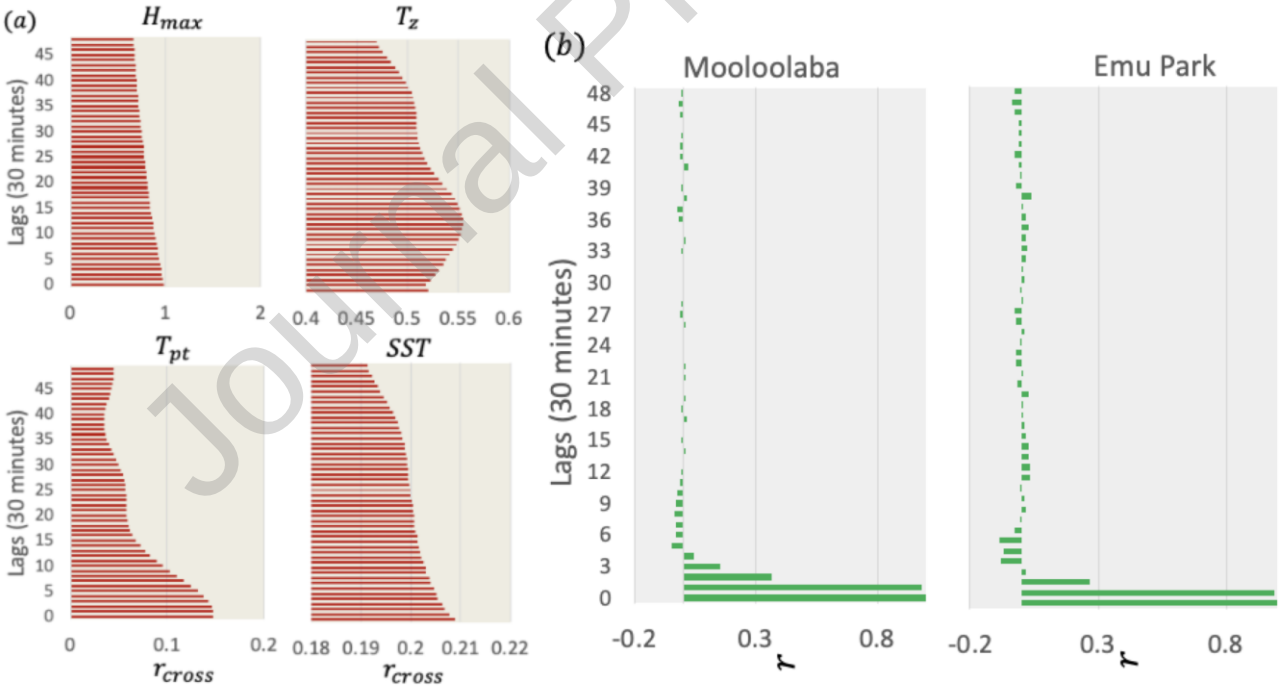



**Fig. 3** (a) The correlogram shows the covariance between the objective target ($H_{sig}$) and the predictor ($H_{max}$, $T_z$, $T_p$, $SST$) variables in terms of cross-correlation coefficients ($r_{cross}$), and (b) a partial autocorrelation function (PACF) plot of the $H_{sig}$ time series exploring the antecedent behaviour for Mooloolaba and Emu Park study sites.

### 4.3 The Development of CLSTM-BiGRU Objective Model

In this study, we developed a novel hybrid predictive model (CLSTM-BiGRU) that incorporates CNN, LSTM and BiGRU algorithms. For the prediction of $H_{sig}$ time series at multiple forecast horizons (i.e., 0.5H, 02H, 03H, and 06H), the proposed deep learning hybrid CLSTM-BiGRU model consists of three convolutional layers, a pooling layer, and the final layer, which is flattened and input to the LSTM and then to the BiGRU model.

Three crucial steps comprise the modelling process:

- To pre-train the CNN, we first input the training data into the CNN model and then compute the convolutional and fully connected layer parameters. Through the convolution layers, the features of the training data are retrieved and filtered.
- Two LSTM layers and one BiGRU layer use the extracted features as input to calculate the significant wave height ($H_{sig}$) for four hourly horizons. The model's predictive capability is increased by retrieving data from the flatten layer's output once more using the CLSTM-BiGRU model.

For a deep learning prediction model, hyperparameter optimization is essential. The optimally selected hyperparameters of deep learning models are tabulated in Table A1 which can be placed at Appendix. This should be performed to enhance the performance of the model on independent (test) datasets. Grid search has also been utilized well because they facilitate the training of deep learning models [59, 60]. We employed the Stochastic gradient descent optimization approach, which uses an iterative method for optimizing an objective function with appropriate smoothness characteristics [61]. This technique's benefits are simplicity, effectiveness, minimal memory requirements, re-scalability of the gradient's diagonal, and adaptability for massive data sets [62, 63]. With a constant learning rate of (lr) 0.001, decay rates of (1=0.9 & 2 = 0.99), and an epsilon of $10^{-8}$, we employed the Adam optimization algorithm. Additionally, every output layer was followed by the Rectified Linear Units (ReLU) activation function except the final one. ReLU, a popular activation function in DL models, is parameter-free and non-saturating, which can speed up stochastic gradient descent's convergence saturation [64]. ReLU can

greatly boost deep learning performance in terms of faster convergence and higher accuracy when compared to its saturated counterpart activation functions, such as sigmoid and tanh [63]. Additionally, the robust deep hybrid CLSTM-BiGRU model used in this study to forecast daily $H_{sig}$ has been regularised using the following techniques:

- During the model training phase: Many epochs during the model training phase can result in an overfitted model, while fewer epochs might result in an underfitted model. Early stopping (es) was implemented [65] to avoid these mistakes. The training phase is terminated when the model's performance does not improve on a validation dataset. As a result, training was stopped during model construction after 15 (patience) consecutive epochs in which the loss had ceased reducing.
- To avoid the over-fitting: The "ModelCheckpoint" call back is used in this study to preserve the version of the model with the greatest performance at the conclusion of an epoch after using Keras "ReduceLROnPlateau" function to lower the learning rate when a validation loss stops increasing [66]. With patience of 10, the learning rate (lr) is decreased by a factor of 0.2.

**4.3 Benchmark Model Development**

The proposed objective model (i.e., the deep hybrid CLSTM-BiGRU) and the benchmark deep learning models were created using TensorFlow 2.0.1 [67, 68] and Keras 2.2.4 Libraries on a Python programming environment. The training process of all the predictive models was conducted on a system with a CPU type of Intel® Core ™ i7, 16GB RAM. A classical machine learning model, i.e., Random Forest Regressor (RF), was prepared using scikit-learn to further compare the performance of proposed model.

**4.4 The performance Evaluation Metrics**

We use various visual and statistical criteria during our models' independent testing phase. The performance requirements largely concern with the model's characteristics and applicability, information about available inputs, and model-specific knowledge [69]. The link between planned and observed values determines a model's efficiency; yet these criteria are usually identified without considering the model's purposes and projects. Several scoring measures were used, including three efficiency metrics (Pearson's correlation coefficient (r), Kling-Gupta Efficiency (KGE) [70] and Nash-Sutcliffe Efficiency

(NSE) [71] and two error metrics [72] and Root Mean Square Error (RMSE; m). According to Willmott and Matsuura, MAE is a more accurate predictor of model performance than RMSE [73]. Equations (16-28) give the corresponding mathematical formulas for MAE, RMSE, NSE, MAPE, and RMAE.

**Mean Absolute Error** (MAE, m) is defined as:

$$MAE = \frac{1}{\mathrm{N}}\sum_{\mathrm{i=1}}^{\mathrm{N}}\left|H_{\mathrm{sig.for}} - H_{\mathrm{sig.obs}}\right|,\ 0 \le MAE \le \infty \tag{16}$$

**Root Mean Square Error** (RMSE; m) is given as:

$$RMSE = \sqrt{\frac{1}{\mathrm{N}}\sum_{\mathrm{i=1}}^{\mathrm{N}}(H_{\mathrm{sig.for}} - H_{\mathrm{sig.obs}})^2}, 0 \le RMSE \le \infty \tag{17}$$

**Nash – Sutcliffe Efficiency** (NSE) is expressed as:

$$NSE = 1 - \left[1 - \frac{\sum_{\mathrm{i=1}}^{\mathrm{N}} H_{\mathrm{sig.for}})^2}{\sum_{\mathrm{i=1}}^{\mathrm{N}}(H_{sig.obs} - \bar{H}_{sig.for})^2}\right], -\infty \le NSE \le 1 \tag{18}$$

**Mean Absolute Percentage Error** (MAPE, %) is expressed as:

$$MAPE = \frac{1}{\mathrm{N}}\left(\sum_{N}^{\mathrm{i=1}}\left|\frac{(H_{\mathrm{sig.for}} - H_{\mathrm{sig.obs}})}{H_{sig.obs}}\right|\right) * 100 \tag{19}$$

**Index of Agreement** (d) is stated as:

$$d = 1 - \left[\frac{\sum_{\mathrm{i=1}}^{\mathrm{N}}(H_{\mathrm{sig.for}} - H_{\mathrm{sig.obs}})^2}{\sum_{\mathrm{i=1}}^{\mathrm{N}}(|H_{\mathrm{sig.for}} - \bar{H}_{\mathrm{sig.obs}}| + |H_{\mathrm{sig.obs}} - \bar{H}_{\mathrm{sig.obs}}|)^2}\right], 0 \le WI \le 1 \tag{20}$$

**Relative Mean Absolute Error** (RMAE, %)

$$RMAE = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{|(H_{\mathrm{sig.for}} - H_{\mathrm{sig.obs}})| \times 100\%}{H_{\mathrm{sig.for}}}\right) \tag{21}$$

In the commonly used persistence model, the calculations assume that atmospheric conditions are stationary between the present and the anticipated time. In the case of a positive value, the suggested deep hybrid CLSTM-BiGRU predictive model is expected to beat the persistence, baseline model; in the case of a negative value, the persistence model is most likely superior. Moreover, this study has evaluated the performance of the proposed model using Kling-Gupta Efficiency (KGE) [70].

We also adopted the Promoting Percentage of Kling-Gupta Efficiency (KGE) [70] ($\partial_{KGE}$) and Relative Mean Absolute Error ($\partial_{RMAE}$) to compare the various models used in $H_{sig}$ prediction.

$$\partial_{KGE} = |(KGE_1 - KGE_2)/KGE_1| \tag{22}$$

$$\partial_{RMAE} = |(RMAE_1 - RMAE_2)/RMAE_1| \quad (23)$$

where,

$KGE_1$ *and* $RMAE_1$ = CLSTM-BiGRU model performance metrics

$KGE_2$ *and* $RMAE_2$ = benchmark model performance

**Kling – Gupta Efficiency** (KGE) is expressed as:

$$KGE = 1 - \sqrt{(r-1)^2 + \left(\frac{SD_{sig.for}}{SD_{sig.obs}} - 1\right)^2 + \left(\frac{\bar{H}_{sig.for}}{\bar{H}_{sig.obs}} - 1\right)^2} \quad -\infty \leq KGE \leq 1 \quad (24)$$

And r is Correlation Coefficient, which is mathematically expressed as below:

$$r = \left\{ \frac{\sum_{i=1}^{N}(H_{sig.obs} - \bar{H}_{sig.obs})(H_{sig.for} - \bar{H}_{sig.for})}{\sqrt{\sum_{i=1}^{N}(H_{sig.obs} - \bar{H}_{sig.obs})^2 \sum_{i=1}^{N}(H_{sig.for} - \bar{H}_{sig.for})^2}} \right\}^2 \quad (25)$$

Finally, we adopted the direction of movement as measured by **Expanded uncertainty ($U_{95}$)** such that:

$$U_{95} = 1.96 * (SD^2 + RMSE^2)^2 \quad (26)$$

**MAE Skill Score ($MAE_{SS}$):**

$$MAE_{SS} = \frac{MAE_{RF} - MAE_{DL}}{MAE_{RF}} \quad (27)$$

**RMSE Skill Score ($RMSE_{SS}$):**

$$RMSE_{SS} = \frac{RMSE_{RF} - RMSE_{DL}}{RMSE_{RF}} \quad (28)$$

Where $H_{sig.obs}$ and $H_{sig.for}$ denote the observed and model forecasted value from the $i^{th}$ element; $\bar{H}_{sig.obs}$ and $\bar{H}_{sig.for}$ denote their average, respectively, SD represents the standard deviation of the data and N signifies the number of observations of the $H_{sig}$. DL is referred as three deep learning models (i.e., BiGRU, CNN-LSTM and CLSTM-BiGRU).

## 5.0 Experimental Results and Discussion

This study demonstrates the effectiveness of a newly designed deep learning hybrid CLSTM-BiGRU model over the classic deep learning models of CLSTM and BiGRU, and a machine learning model RF, to forecast the significant wave height ($H_{sig}$) at four areas: Gold Coast, Cairns, Mooloolaba,

and Emu Park located in the state of Queensland, Australia. The models were developed using four-time steps of 0.5 H, 2H, 3H, and 6H. In this section, two statistical tools of mean absolute error (MAE) and Index of Agreement (d) and different schemes have been used to determine the prediction accuracy and performance of the CLSTM-BiGRU model and the comparison models. According to the description and mechanism of the MAE and d metrics, the model with the lowest MAE and highest d is elected as the best model.

Comparing the results that are demonstrated in Table 3 for the machine learning model (RF) and the deep learning models (hybrid CLSTM-BiGRU, CLSTM, and BiGRU), the machine learning RF model had the lowest accuracy with all study areas and prediction steps. On the other hand, when the comparison was made among the deep learning models, superior performance was made by the suggested study model CLSTM-BiGRU. In terms of the half hourly prediction (0.5H) with all study sites, the CLSTM-BiGRU model has made the best values for both metrices (MAE/d). For Cairns, Emu Park, Gold Coast and Mooloolaba, respectively, those values were 0.033/0.994, 0.024/0.998, 0.055/0.994 and 0.011/0.997 compared to 0.034/0.988, 0.034/0.995, 0.056/0.990 and 0.014/0.993 for CLSTM and 0.045/0.991, 0.039/0.990, 0.059/0.991 and 0.019/0.989 for BiGRU. Although by relatively small margin, the BiGRU model had the lowest MAE values [0.053 for Gold Coast (2H) and 0.060 for Emu Park (6H)], the CLSTM-BiGRU model yielded the best MAE values when the data of 2H and 6H from other sites were used as well as the highest d values for all study zones with respect to these time steps. Using 3H datasets, excluding the d value for Cairns, the CLSTM-BiGRU model achieved the best values for both metrices outperforming the CLSTM and BiGRU models.

**Table 3** Mean Absolute Error (MAE) and Index of Agreement (*d*) between the observed and forecasted $H_{sig}$ using the proposed CLSTM-BiGRU model compared to the CLSTM, BiGRU and RF models.

| **Forecast Horizon** | **Cairns** | | **Emu Park** | | **Gold Coast** | | **Mooloolaba** | |
|---|---|---|---|---|---|---|---|---|
| | *MAE* | *d* | *MAE* | *d* | *MAE* | *d* | *MAE* | *d* |
| **CLSTM-BiGRU (Proposed Model)** | | | | | | | | |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **0.5H** | **0.033** | **0.994** | **0.024** | **0.998** | **0.055** | **0.994** | **0.011** | **0.997** |
| **02H** | **0.037** | **0.986** | **0.035** | **0.995** | 0.055 | **0.993** | **0.015** | **0.995** |
| **03H** | **0.041** | 0.985 | **0.034** | **0.994** | **0.076** | **0.985** | **0.047** | **0.993** |
| **06H** | **0.053** | **0.975** | 0.062 | **0.982** | **0.110** | **0.967** | **0.054** | **0.990** |
| **CLSTM** | | | | | | | | |
| **0.5H** | 0.034 | 0.988 | 0.034 | 0.995 | 0.056 | 0.990 | 0.014 | 0.993 |
| **02H** | 0.040 | 0.984 | 0.038 | 0.992 | 0.054 | 0.990 | 0.021 | 0.989 |
| **03H** | 0.051 | **0.989** | 0.045 | 0.991 | 0.078 | 0.984 | 0.054 | 0.982 |
| **06H** | 0.058 | 0.972 | 0.062 | 0.979 | 0.129 | 0.956 | 0.058 | 0.973 |
| **BiGRU** | | | | | | | | |
| **0.5H** | 0.045 | 0.991 | 0.039 | 0.990 | 0.059 | 0.991 | 0.019 | 0.989 |
| **02H** | 0.042 | 0.983 | 0.036 | 0.993 | **0.053** | 0.984 | 0.032 | 0.975 |
| **03H** | 0.049 | 0.988 | 0.035 | 0.988 | 0.088 | 0.979 | 0.056 | 0.979 |
| **06H** | 0.058 | 0.970 | **0.060** | 0.981 | 0.123 | 0.955 | 0.057 | 0.966 |
| **RF** | | | | | | | | |
| **0.5H** | 0.061 | 0.869 | 0.089 | 0.969 | 0.139 | 0.934 | 0.052 | 0.961 |
| **02H** | 0.112 | 0.798 | 0.087 | 0.970 | 0.129 | 0.945 | 0.124 | 0.968 |
| **03H** | 0.113 | 0.766 | 0.091 | 0.968 | 0.141 | 0.933 | 0.132 | 0.961 |
| **06H** | 0.113 | 0.786 | 0.102 | 0.955 | 0.157 | 0.919 | 0.148 | 0.953 |

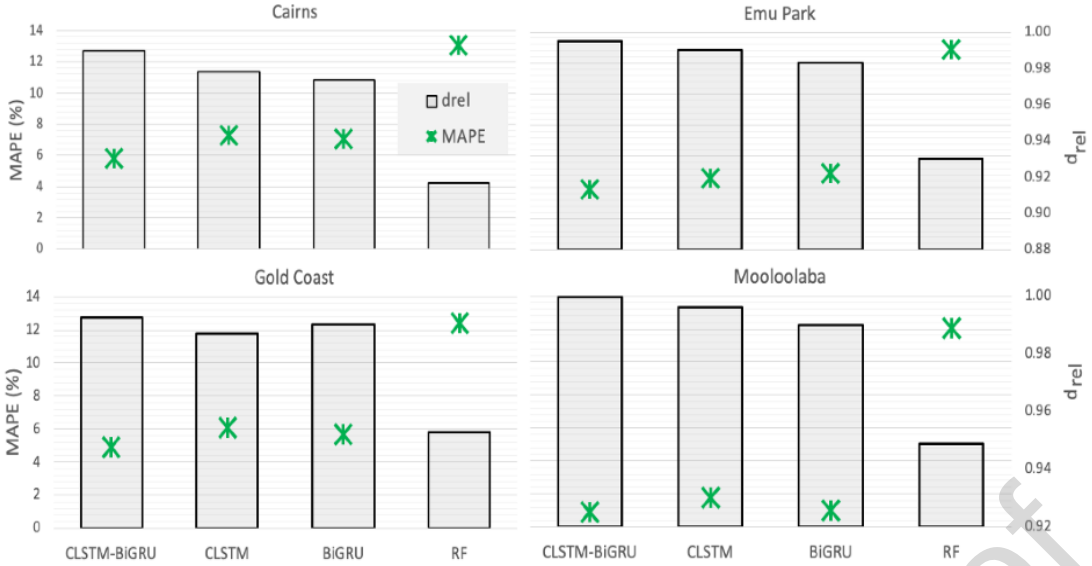



**Fig. 4** Comparison of the predictive skill of the proposed CLSTM-BiGRU model *vs*. CLSTM, BiGRU and RF (benchmark) models in terms of the Mean Absolute Percentage Error (MAPE %) and Relative Index of Agreement ($d_{rel}$) computed in the testing period for 0.5H forecasting horizon.

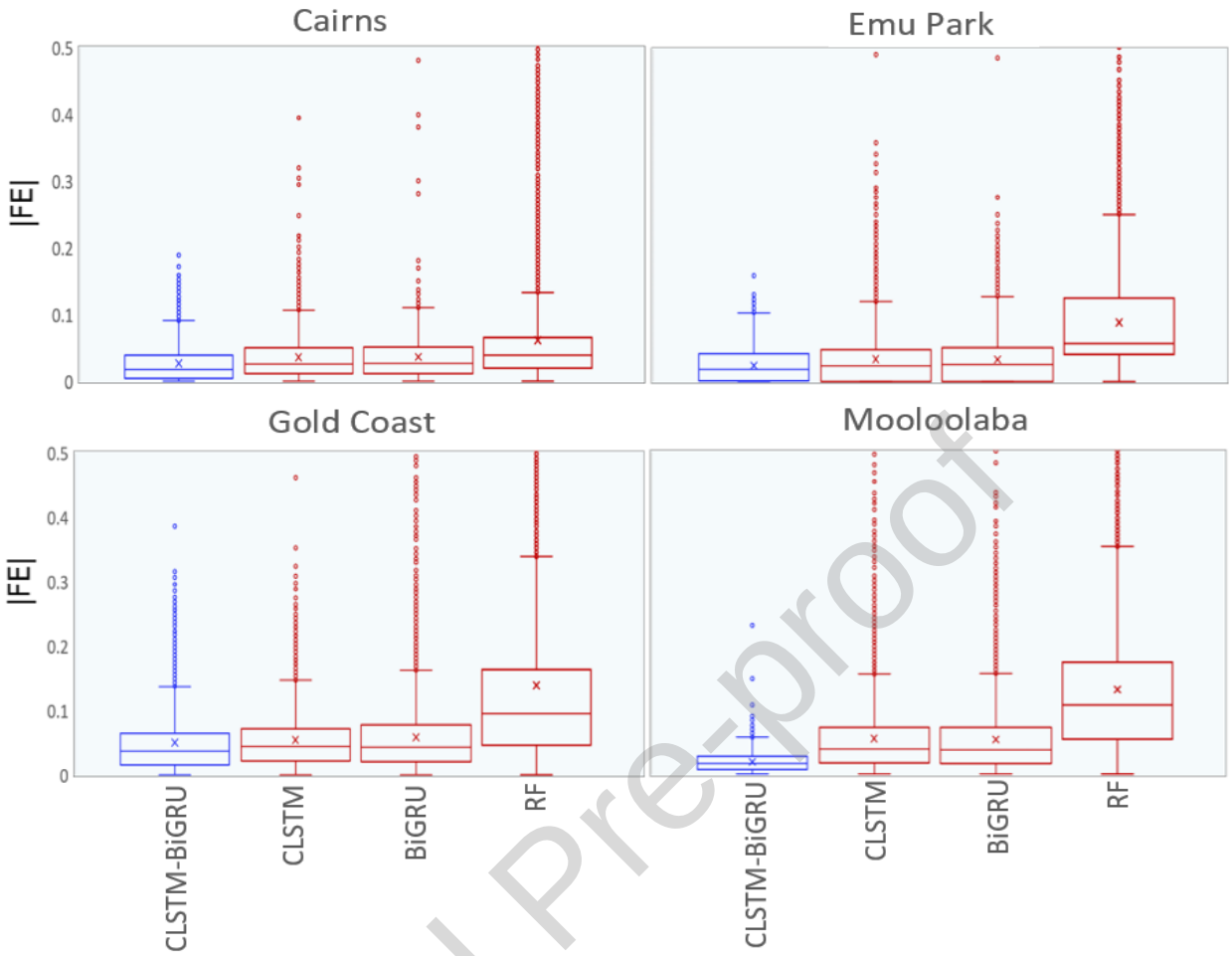



**Fig. 5** An evaluation of the proposed CLSTM-BiGRU model in respect to the benchmark models based on absolute forecasted error |FE| for 0.5H forecasting horizon.

Various graphics have been presented in this research to discuss the experimental results further to show the proposed model's ability to accurately forecast oceanic significant wave height ($H_{sig}$). Firstly, the study illustrated the Mean Absolute Percentage Error (MAPE %) and Relative Index of Agreement ($d_{rel}$) in Fig. 4 to examine the precision of the models for $H_{sig}$ prediction in 0.5H horizon. Accordingly, the best values (lowest MAPE and highest $d_{rel}$) were created by the CLSTM-BiGRU model when they were compared to the developed benchmarked from deep and machine learning models. The recommended model performed near unity $d_{rel}$ values with significantly low values of MAPE using all study sites. Furthermore, the ability of the CLSTM-BiGRU model to predict $H_{sig}$ was confirmed to be the best by presenting the boxplots in Fig. 5. Using the forecasted error |FE| for 0.5H forecasting horizon, the boxplots showed the |FE| values with respect to different statistical values of minimum, average,

maximum, first quartile (25%), second (medium) quartile (50%) and the third quartile (75%). Based on these statistical tools, hence, the study objective model has generated the lowest values due to its advantage in dealing with time-series data verifying its considerable ability to yield better estimation of $H_{sig}$ data than the other models.

To investigate the relationship between the observed and predicted data, scatterplots were employed in Fig. 6 using 0.5 h with all study regions. The regression line of y = ax + b, which is corresponding to $H_{sig.for} = a * H_{sig.obs} + b$ in this study, and the correlation of determination ($R^2$) were used to assess the deep learning model's accuracy. The values of $R^2$ were 0.998, 0.997, 0.997 and 0.997 for CLSTM-BiGRU, 0.995, 0.995, 0.995 and 0.982 for CLSTM and 0.992, 0.990, 0.993 and 0.979 for BiGRU using the 0.5H dataset for Cairns, Gold Coast, Emu Park, and Mooloolaba, respectively. Based on those values, again, the CLSTM-BiGRU model had the highest accuracy in forecasting the oceanic wave height data.

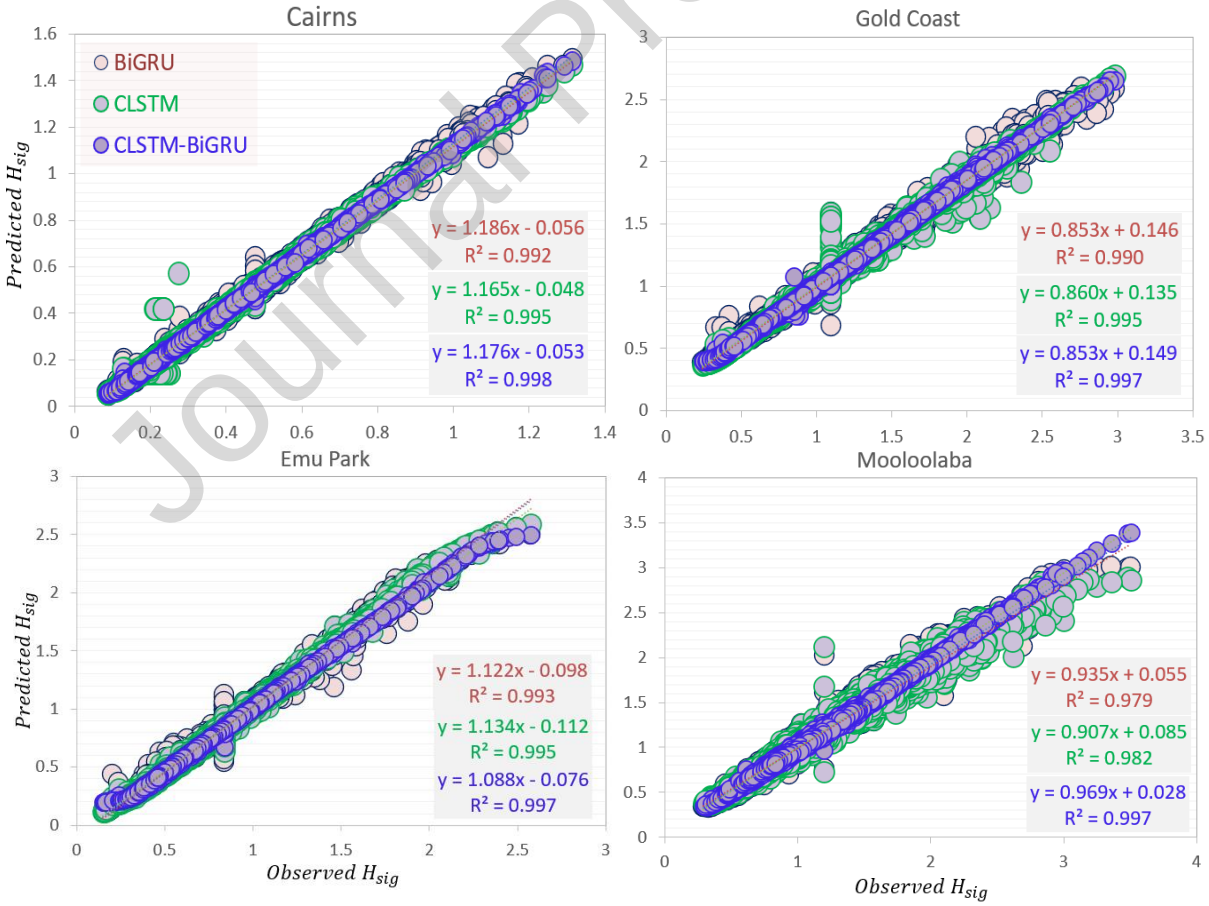



**Fig. 6** Scatter plot of forecasted and observed $H_{sig}$ in testing phase at the four stations using the proposed CLSTM-BiGRU *vs.* CLSTM and BiGRU models for 0.5 H horizon. Least

square regression line and coefficient of determination ($R^2$) with a linear fit is shown in each sub-panel.

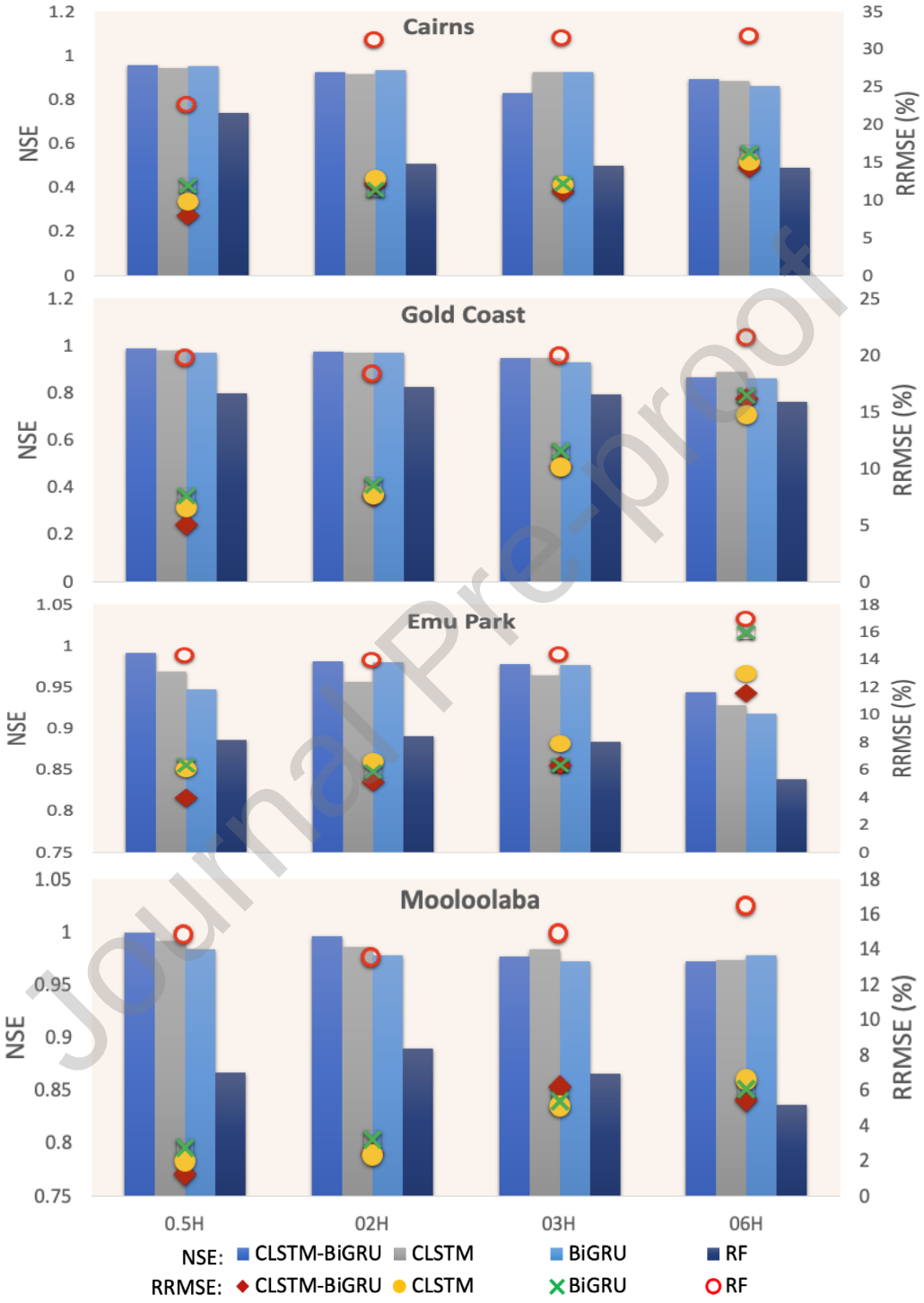



**Fig. 7** Comparison of the predictive skill for proposed CLSTM-BiGRU *vs*. CLSTM, BiGRU and RF models in terms of the relative error: RRMSE (%) and the NSE value within the testing period computed for the multi-step horizons.

Concurring with the earlier results, Fig. 7 confirms that the hybrid CLSTM-BiGRU model had the most extraordinary power compared to the CLSTM, BiGRU, and RF models to predict oceanic wave height values. In association with Fig.7, two metrics of the Relative Root Mean Square Error (RRMSE%) and the Nash–Sutcliffe Coefficients (NSE) were utilized to determine the predictive proficiency of the used models in which the model that generates the lowest percentage of RRMSE and highest value of NSE is considered the best one. Accordingly, the CLSTM-BiGRU technique has achieved these criteria, as shown in Fig. 7, presenting better forecasting values and outperforming CLSTM, BiGRU, and RF models using the multi-step horizons of 0.5H, 2H, 3H, and 6H with all study station's datasets.

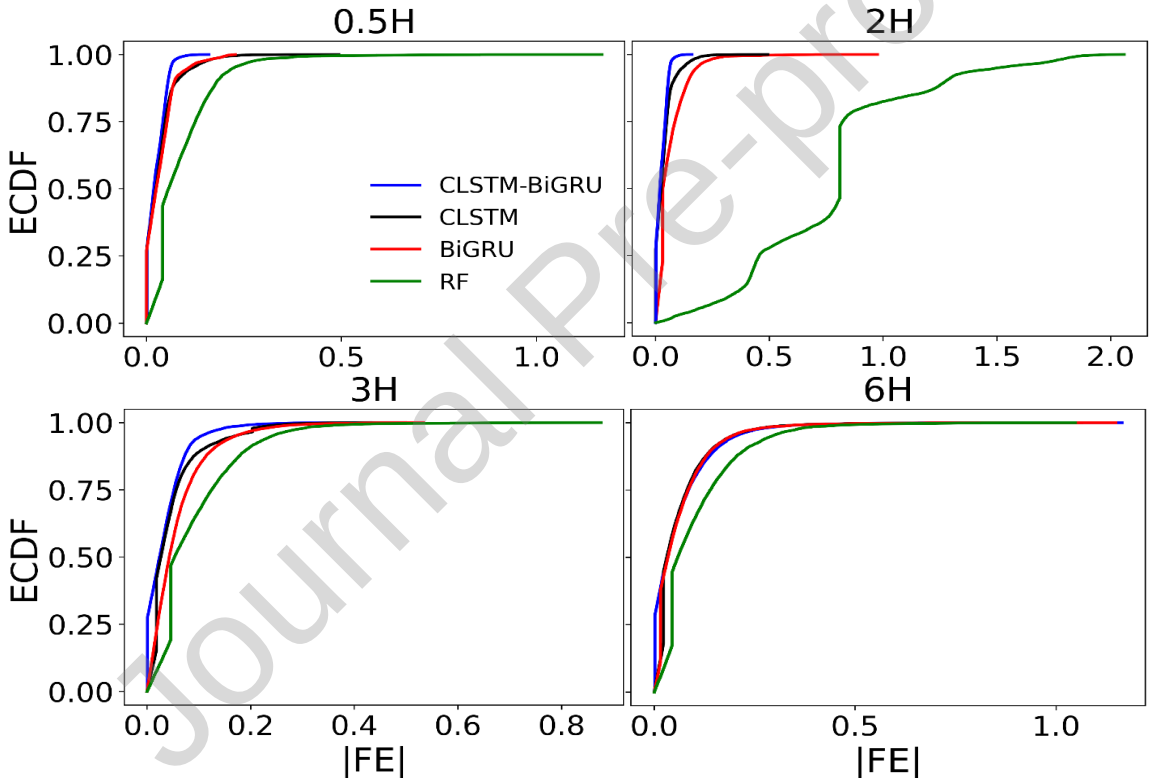



**Fig. 8** Empirical cumulative distribution function (ECDF) in absolute forecast error |FE| for the proposed CLSTM-BiGRU *vs.* CLSTM, BiGRU and RF models for Mooloolaba station presented for multi-step forecast horizons.

Another model evaluation graphical approach was implemented in Fig. 8 of this study to further check the strength of the suggested CLSTM-BiGRU model over the other tested models for forecasting $H_{sig}$. This was the Empirical cumulative distribution function (ECDF) that was plotted for the forecasted error |FE| using the multi-step horizons of the Mooloolaba site. Again, the optimal performance was made by

the CLSTM-BiGRU model due to having most of its forecasted errors |FE| in the smallest bracket of 0 to ± 0.5 for 0.5H, 2H, 3H, and 6H whereas it was 0 to ± 0.2 for the 3H horizon. Fig. 8 presents the more detail of this phenomena.

From the foregoing results and discussion, it can be concluded that the study-selected model CLSTM-BiGRU has the considerable ability to produce relatively precise prediction values of oceanic wave height. This model can highly support the Australian government by instilling an automatic high-quality early warning system that can provide different benefits, such as (1) estimating the level of the wave before it occurs, (2) offering valuable information for diverse real-world applications such as, in marine conveyance, environmental supervising, as well as coastal protection and engineering [74]. Thus, the CLSTM-BiGRU model is needed to address the practical problems that create potential risks for industries, governments, and people's daily lives.

## 6.0 Further Discussion

This study has made significant contributions in respect to developing and verifying the predictive stability and capability of the proposed hybrid CLSTM-BiGRU model. The approach integrated three phases comprised of the CNN, LSTM and BiGRU methods to predict the oceanic significant wave height ($H_{sig}$). The forecasting achievement of the preferred model has been selected as the best performing by comparing it with a two-phase model of CLSTM and two single-phase models of BiGRU and RF. Although different criteria from statistical metrices and high quality of graphical analyses have been employed in the previous section to determine the best model, this section also presented three other tools to further assess the model's performance. Having these as further results and discussion can confirm the power of the CLSTM-BiGRU model in generating optimum forecasting values.

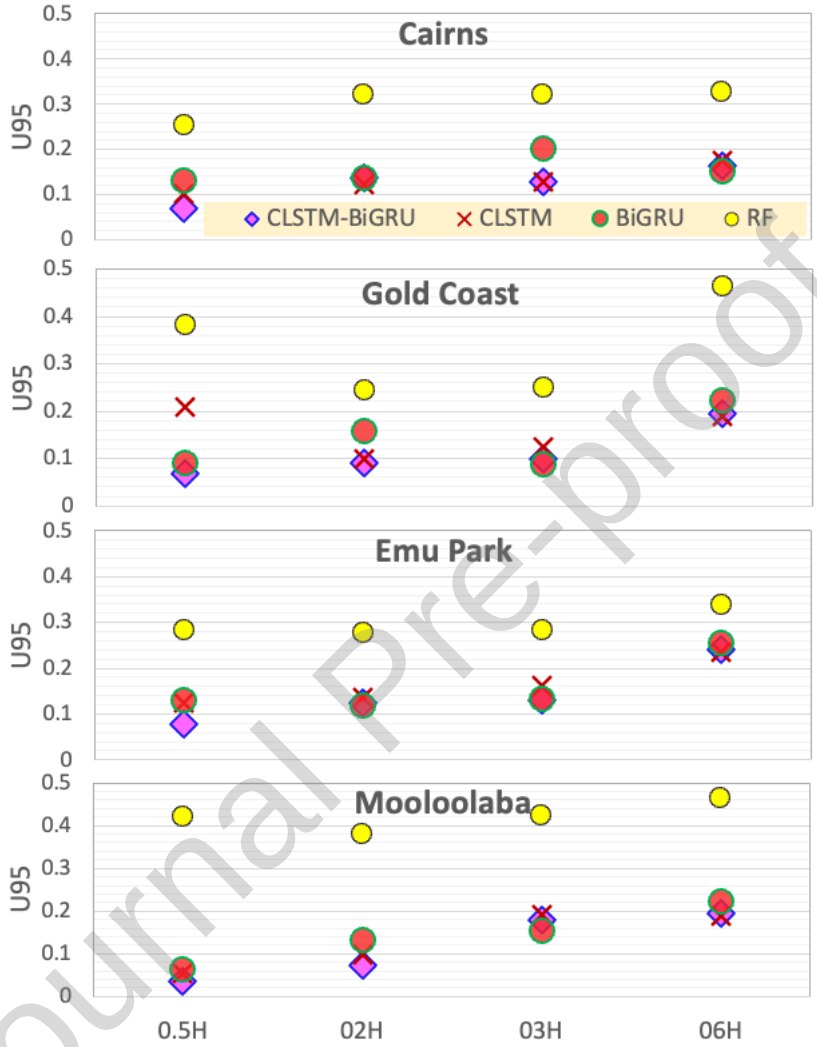





**Fig. 9** Comparison of the predictive skill of the proposed CLSTM-BiGRU model using the Expanded Uncertainty (U95) metric against the benchmark CLSTM, BiGRU and RF models for the multi-step forecast horizons.

To measure the 95% level of confidence, Fig. 9 is shown for all models, study sites, and multi-steps based on the uncertainty at 95% (U95), in which the closest value to zero is expected to indicate the best model. Using this method of evaluation, high forecast accuracy was achieved by the CLSTM-BiGRU model that outperformed other developed models presented in Fig. 9. Additionally, to find the

closest model to the CLSTM-BiGRU model, the differencing values in Kling-Gupta Efficiency (KGE) and Root Mean Absolute Error (RMAE) named here $\partial_{KGE}$ and $\partial_{RMAE}$, respectively, between the suggested model and each developed model have been separately calculated and plotted in Fig. 10.

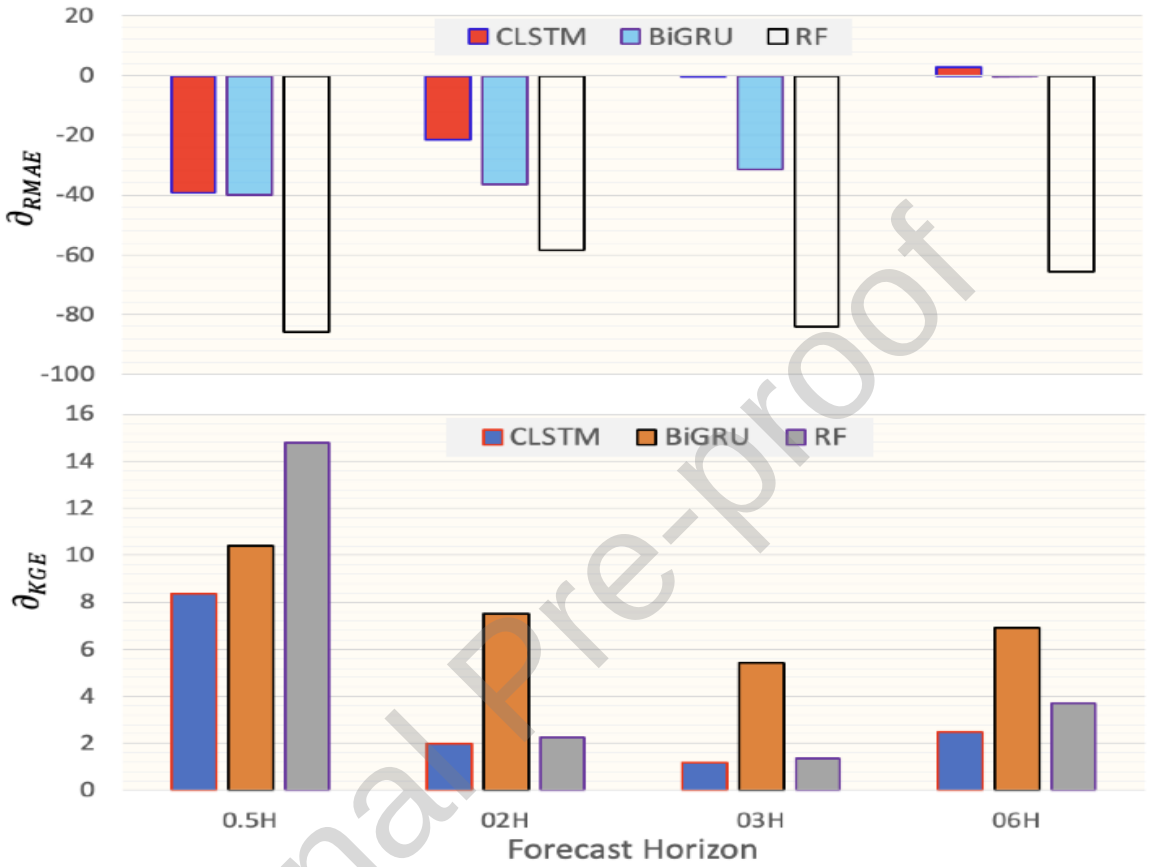



**Fig. 10** The prompting percentage change in KGE ($\partial_{KGE}$) and RMAE ($\partial_{RMAE}$) calculated with respect to the proposed CLSTM-BiGRU model for Mooloolaba stations at the multi-step forecast horizons.

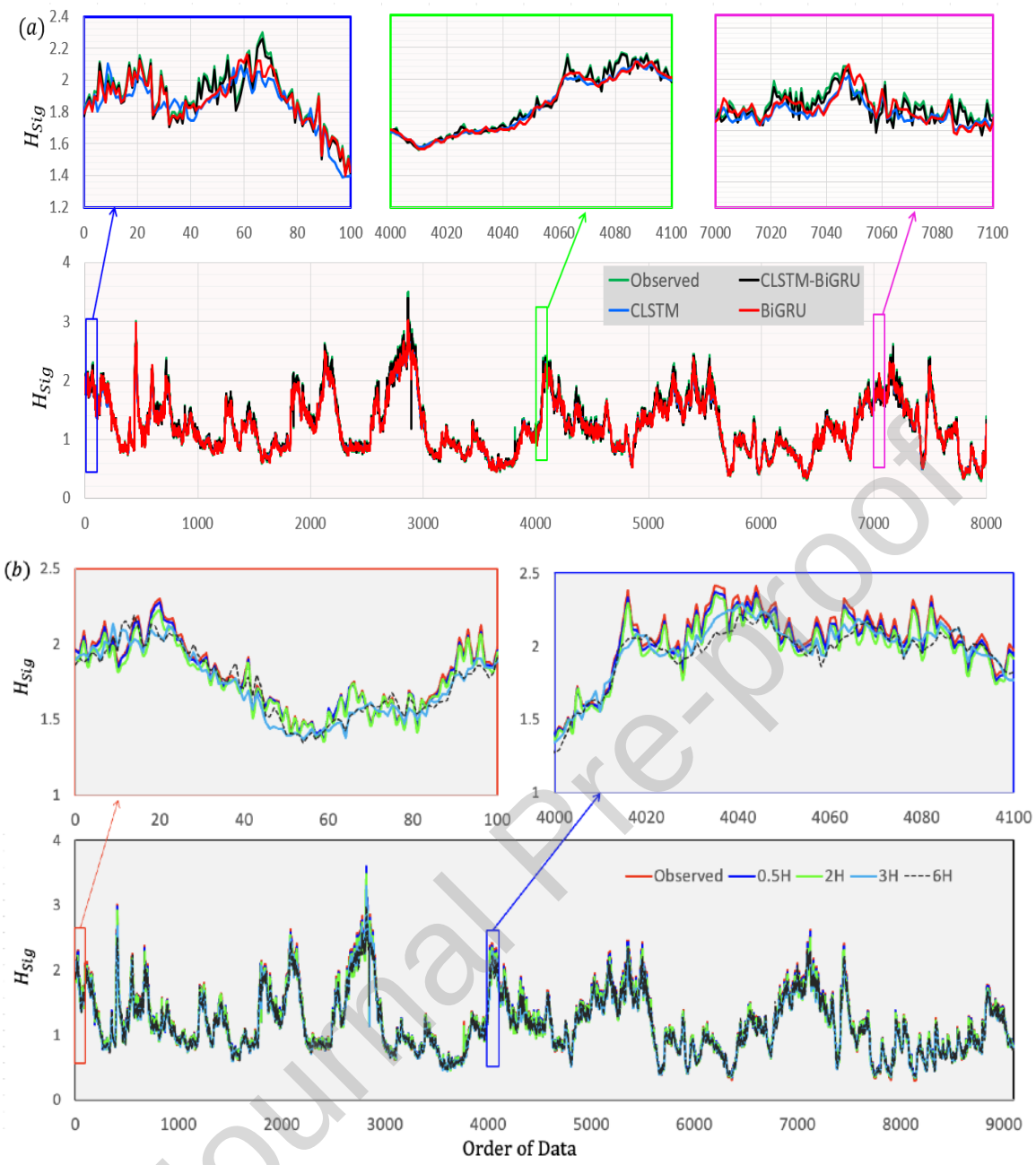



**Fig. 11** The time series plot at Mooloolaba study site of: (a) forecasted and observed $H_{sig}$ for the proposed CLSTM-BiGRU model compared with the CLSTM and BiGRU models at 0.5 h forecast horizons, (b) the forecasted and observed $H_{sig}$ generated by the proposed CLSTM-BiGRU model at multi-step forecasting horizons.

Based on the structure of these tools, the model that achieved the lowest values of $\partial_{KGE}$ and $\partial_{RMAE}$ signifies the closest model to the CLSTM-BiGRU model. Accordingly, the outcomes of Fig. 10, which have been analysed using the Mooloolaba station with multi-step datasets, have shown that the CLSTM model accomplished these conditions indicating to have accurate prediction values that are closest to the best study model of CLSTM-BiGRU. Finally, the forecasted and observed values are also presented in

Fig. 11 for the Mooloolaba station to show the consistency of the presented model over the benchmarked models. Fig. 11 (a) shows the forecasted and observed values at 0.5 h for the CLSTM-BiGRU, CLSTM, and BiGRU models, while Fig. 11 (b) presents these values for the study suggested model only with multi-step forecasting horizons. According to Fig.11, the prediction values obtained by the CLSTM-BiGRU model are the closest to the observed 0.5 h data when the comparison was made with those values generated by the CLSTM and BiGRU models. On the other hand, the forecasting values using 0.5H, 2H, 3H, and 6H of the proposed model were significantly close to the observed data.

To assess potentiality of our approach against a state-of-the-art method, we calculated skill score of MAE ($MAE_{SS}$) and RMSE ($RMSE_{SS}$) between hybrid deep learning models, namely BiGRU, CNN-LSTM, and CLSTM-BiGRU, against a baseline model represented by RF (Random Forest), which serves as a classical benchmark (**Fig. 12**). For the MAE Skill Score, the CLSTM-BiGRU model consistently demonstrates its superiority over both BiGRU and CNN-LSTM across all time horizons and locations. Specifically, at the 0.5H time horizon, the CLSTM-BiGRU achieves a skill score of 0.57, outperforming BiGRU's 0.53 and CNN-LSTM's 0.59. This trend continues as the time horizons extend, highlighting the robustness of the CLSTM-BiGRU model's predictions. Examining the RMSE Skill Score, again, the CLSTM-BiGRU exhibits remarkable proficiency. At the 02H time horizon, it attains a score of 0.45, surpassing both BiGRU's 0.56 and CNN-LSTM's 0.56. This pattern endures for subsequent time horizons and locations, reaffirming the consistency of the CLSTM-BiGRU’s performance. This comparative analysis thus underscores the potential of the CLSTM-BiGRU model as a valuable tool for accurate and reliable wave forecasting in various coastal regions, outperforming both BiGRU and CNN-LSTM models.

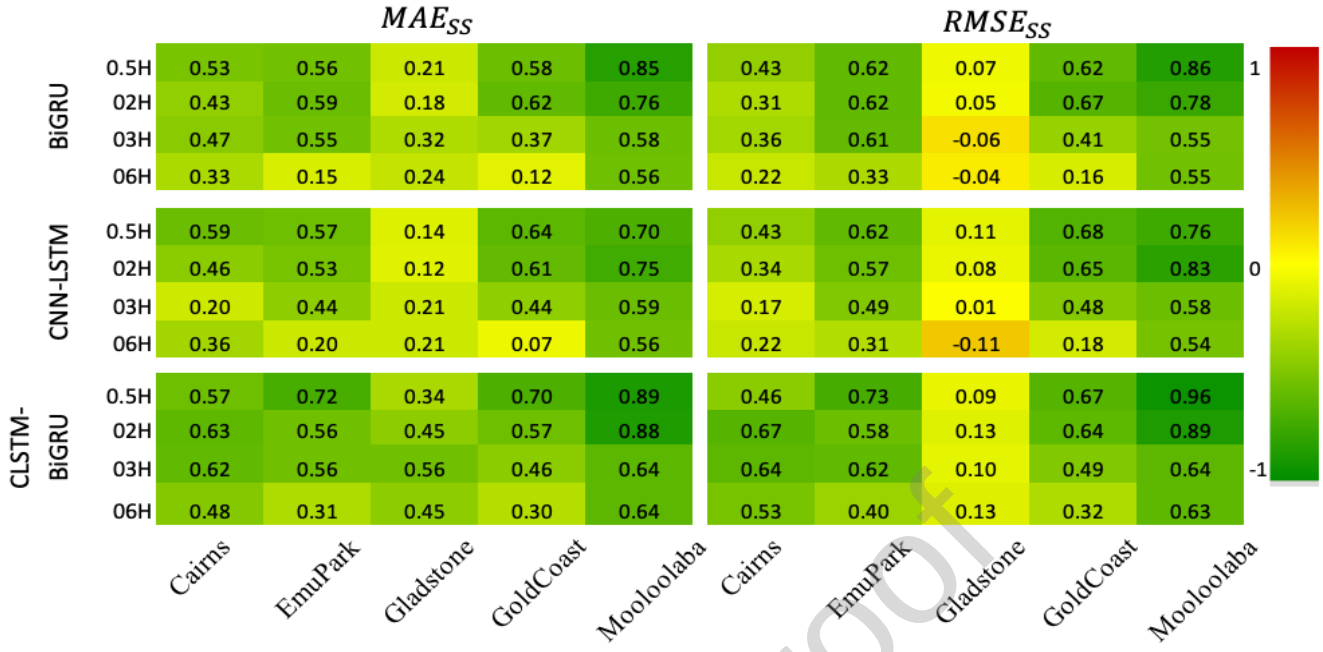



**Fig. 12** Skill Score of MAE ($MAE_{SS}$) and RMSE ($RMSE_{SS}$) between deep learning models against classical machine learning (RF) model for five study regions at multi-step forecasting horizons.

Parameter sensitivity analysis is a technique used to assess how changes in the values of input parameters affect the output of a model or a system. Figure 13 shows the % change of RMSE on the proposed CLSTM-BiGRU model by systematically altering the input features across varying percentages to comprehend the model's responsiveness to parameter changes. For instance, considering a 1% increase, the $H_{max}$ displayed a 2.0% change, while Tz exhibited a larger 4.0% alteration. Similarly, the $T_p$ underwent fluctuations of 5.0%, and the sea surface temperature (SST) showcased a distinct 12.0% variation. This comprehensive analysis provides a nuanced understanding of how different input features influence the behaviour of our CLSTM-BiGRU model. The proposed model response over different forecasting horizon shows a very positive correlation with the input variables. Lastly, the process time is important to make the model more adaptable specially when our model consists multi-layered hybrid deep learning model. The overall processing time of the proposed model was $11 \pm 3$ minutes for different region and forecasting horizon establishing the approach easily adaptable.

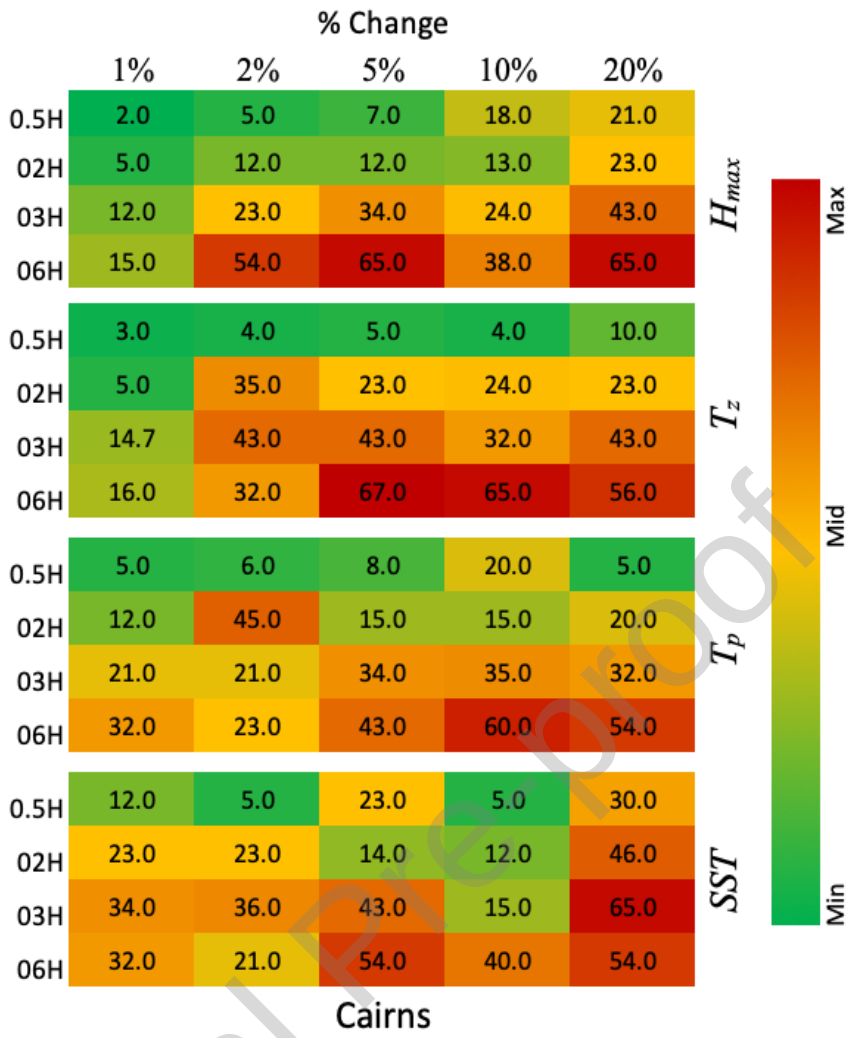



**Fig. 13** Parameter sensitivity analysis of proposed CLSTM-BiGRU model for RMSE

## 7.0 Conclusions and further research outlook

The purpose of this study was to develop an artificial intelligence methodology for forecasting significant wave heights at four stations in Queensland, Australia, using deep learning algorithms. The proposed deep hybrid CLSTM-BiGRU model was built using an innovative method that combines Convolutional Neural Networks (CNN) with Long Short-Term Memories (LSTM) and Bidirectional Gated Recurrent Unit (BiGRU) to achieve maximum accuracy. The most important features were extracted by incorporating the CNN algorithm into the proposed deep learning model. After the extraction of the features, the LSTM and BiGRU layers were used to forecast the significant wave height based on the extracted features. Based on the analyses, the deep hybrid CLSTM-BiGRU model outperformed some of the most well-known prediction models, including LSTM, BiGRU, and RF. Furthermore, the proposed

deep hybrid model was thoroughly tested, which confirmed that our modelling strategy produced a viable method of predicting $H_{sig}$ in the short term. The study shows that the proposed deep hybrid CLSTM-BiGRU model can be used to solve a variety of complicated and challenging prediction problems, including those involving the forecasting of wind speed, crude oil prices, traffic flow, the stock market, exchange rates, tidal energy, etc. Accordingly, the CLSTM-BiGRU model was highly accurate in predicting $H_{sig}$ based on the robust evaluation methods used in this study. While the method has been successful, there may be some limitations that can be addressed in future research. For example, we may improve the model's precision even further by considering other predictors, such as weather data. A second challenge is that this study did not assess long-term prediction skills, which can provide more useful information in making decisions related to tidal and wave energy systems, as well as establishing a robust prediction model for monitoring marine water during natural disasters.

The proposed CLSTM-BiGRU hybrid model for predicting the wave energy indeed holds promise for broader applications within the realm of time series forecasting. This hybrid approach's effectiveness in capturing temporal dependencies and spatial patterns within wave data suggests its potential applicability to various other time series tasks. The hybrid approaches were adopted in addressing real life problems associated with hydrological [43, 52, 75, 76], energy [51, 77-80] and medical [81] sectors. This model's ability to learn from historical data and its capacity to handle multiple input channels, as validated by its success in predicting wind energy series, hints at its potential to be employed across diverse domains, making it an exhilarating avenue for future exploration.

In this study, we have developed a multi-step model for significant wave height prediction in Australia’s wave energy region. If the model is integrated with a wave energy converter (WEC) through an appropriate modelling platform, the CLSTM-BiGRU technique can be used to monitor and predict wave energy harnessed at the sites [82]. This model can also provide early warning of energy shortfalls, through AI-based predictive methods proposed on this study. We have added a paragraph in conclusion section and cites a few references.

## Credit authorship contribution statement

**Abul A. Masrur Ahmed:** Writing - original draft, conceptualization, methodology, software, editing, proofreading, model development, and application. **S Janifer Jabin Jui:** Writing, editing, and proofreading. **Mohanad S. AL-Musaylh:** Writing, editing, and proofreading. **Nawin Raj:** Writing, editing **Reepa Saha:** Writing, and proofreading. **Ravinesh C Deo:** Writing, editing, and proofreading. **Sanjoy Kumar Saha:** proofreading

## Availability of Data and Materials

This article presents an original research work executed by the authors, so all the data presented depend on their findings and analysis techniques. The datasets used in this article are available from the corresponding author on reasonable request.

## Ethics approval

"Not applicable", research does not report on or involve the use of any animal or human data or tissue.

## Consent for publication

Not applicable.

## Consent to participate.

Not applicable.

## Conflict of interest

The authors declare no competing interests.

## Funding

Not applicable.

## Competing Interests

Not applicable.

## References:

[1] WMO, State of the Global Climate 2020, in, World Meteorological Organization, 2021.
[2] A. AghaKouchak, Chiang, F., Huning, L. S., Love, C. A., Mallakpour, I., Mazdiyasni, O., Moftakhari, H., Papalexiou, S, M., Ragno, E., and Sadegh, M.,, Climate Extremes and Compound Hazards in a Warming World, Annual Review of Earth and Planetary Sciences 48 (2020) 519-548.
[3] K.E. Trenberth, Changes in precipitation with climate change, Climate Research, 47 (2011) 123-128.
[4] M.K.V. Aalst, The impacts of climate change on the risk of natural disasters, Disasters, 30 (2006) 5-18.

[5] A.M. Omer, Energy, environment and sustainable development, Renewable and sustainable energy reviews, 12 (2008) 2265-2300.
[6] L. Richards, Brew, N., Smith, L., 2019–20 Australian bushfires—frequently asked questions: a quick guide, in, Parliamentary Library, Parliament of Australia, 2020.
[7] A.I. Filkov, Ngo, T., Matthews, S., Telfer, S., & Penman, T. D, Impact of Australia's catastrophic 2019/20 bushfire season on communities and environment. Retrospective analysis and current trends. Journal of Safety Science and Resilience
, Journal of Safety Science and Resilience, 1 (2020) 44-56.
[8] E.G. Dountio, Meukam, P., Tchaptchet, D. L. P., Ango, L. E. O., & Simo, A., Electricity generation technology options under the greenhouse gases mitigation scenario: Case study of Cameroon, Energy Strategy Reviews, 13 (2016) 191-211.
[9] J.O. Anderson, Thundiyil, J. G., & Stolbach, A. , Clearing the air: a review of the effects of particulate matter air pollution on human health, Journal of medical toxicology, 8 (2012) 166-175.
[10] S. Heft-Neal, Burney, J., Bendavid, E., & Burke, M., Robust relationship between air quality and infant mortality in Africa, Nature, 559 (2018) 254-258.
[11] I.E. Isphording, & Pestel, N., Pandemic meets pollution: poor air quality increases deaths by COVID-19, Journal of Environmental Economics and Management, 108 (2021) 102448.
[12] A. Jeanneau, Zecchin, A., van Zelden, H., McNaught, T., & Maier, H., INFLUENCE OF CLIMATE CHANGE AND FUEL MANAGEMENT ON BUSHFIRE RISK IN WESTERN AUSTRALIA, in, Bushfire and Natural Hazards CRC, 2021.
[13] energy.gov.au, Renewables, in, 2022.
[14] A.K. Pecher, J.P. , Handbook of Ocean Wave Energy, Springer Nature, UK, 2017.
[15] N. Raj, & Brown, J. , An EEMD-BiLSTM Algorithm Integrated with Boruta Random Forest Optimiser for Significant Wave Height Forecasting along Coastal Areas of Queensland, Australia, Remote Sensing, 13 (2021) 1456.
[16] K. Thomsen, Offshore Wind: A Comprehensive Guide to Successful Offshore Wind Farm Installation, 2014.
[17] W.-y. Duan, Huang, L.-m., Han, Y., and Huang, D.-t., A hybrid EMD-AR model for nonlinear and non-stationary wave forecasting, Journal of Zhejiang University-SCIENCE A, 17 (2016) 115-129.
[18] G. Tang, Du, H., Hu, X., Wang, Y., Claramunt, C., & Men, S. , An EMD-PSO-LSSVM hybrid model for significant wave height prediction, Ocean Science Discussions, (2021) 1-16.
[19] A. Etemad-Shahidi, & Mahjoobi, J., Comparison between M5′ model tree and neural networks for prediction of significant wave height in Lake Superior, Ocean Engineering, 36 (2009) 1175-1181.
[20] J.C. Fernández, Salcedo-Sanz, S., Gutiérrez, P. A., Alexandre, E., & Hervás-Martínez, C., Significant wave height and energy flux range forecast with machine learning classifiers, Engineering Applications of Artificial Intelligence, 43 (2015) 44-53.
[21] O.J. Aarnes, Abdalla, S., Bidlot, J. R., & Breivik, Ø., Marine wind and wave height trends at different ERA-Interim forecast ranges, Journal of Climate, 28 (2015) 819-837.
[22] S. Almeida, Rusu, L., & Guedes Soares, C. , Application of the Ensemble Kalman Filter to a high-resolution wave forecasting model for wave height forecast in coastal areas, Maritime technology and engineering, 2 (2015) 1349-1354.
[23] T.M. Tsai, Yen, P. H., & Huang, T. J., Wave Height Forecast Using Self-Organization Algorithm Model, in: The Nineteenth International Offshore and Polar Engineering Conference, Osaka, Japan, 2009.
[24] M. Ali, R. Prasad, Significant wave height forecasting via an extreme learning machine model integrated with improved complete ensemble empirical mode decomposition, Renewable and Sustainable Energy Reviews, 104 (2019) 281-295.
[25] C.C. Wei, Nearshore wave predictions using data mining techniques during typhoons: a case study near Taiwan's northeastern coast, Energies, 11 (2018) 11.
[26] M.R. Kaloop, Beshr, A. A., Zarzoura, F., Ban, W. H., & Hu, J. W., Predicting lake wave height based on regression classification and multi input–single output soft computing models, Arabian Journal of Geosciences, 13 (2020) 1-14.
[27] M. Ali, R. Prasad, Y. Xiang, R.C. Deo, Near real-time significant wave height forecasting with hybridized multiple linear regression algorithms, Renewable and Sustainable Energy Reviews, 132 (2020) 110003.
[28] M. Özger, Significant wave height forecasting using wavelet fuzzy logic approach, Ocean Engineering, 37 (2010) 1443-1451.

[29] L. Cuadra, Salcedo-Sanz, S., Nieto-Borge, J. C., Alexandre, E., & Rodríguez, G. , Computational intelligence in wave energy: Comprehensive review and case study, Renewable and Sustainable Energy Reviews, 58 (2016) 1223-1246.
[30] S. Ardabili, Mosavi, A., & Várkonyi-Kóczy, A. R. , Advances in machine learning modeling reviewing hybrid and ensemble methods, Springer, Cham, 2020.
[31] M.R. Kaloop, Kumar, D., Zarzoura, F., Roy, B., & Hu, J. W. , A wavelet-Particle swarm optimization-Extreme learning machine hybrid modeling for significant wave height prediction, Ocean Engineering, 213 (2020) 107777.
[32] E. Alexandre, Cuadra, L., Nieto-Borge, J. C., Candil-García, G., Del Pino, M., & Salcedo-Sanz, S., A hybrid genetic algorithm—extreme learning machine approach for accurate significant wave height reconstruction, Ocean Modelling, 92 (2015) 115-123.
[33] W.Y. Duan, Han, Y., Huang, L. M., Zhao, B. B., & Wang, M. H., A hybrid EMD-SVR model for the short-term prediction of significant wave height, Ocean Engineering, 124 (2016) 54-73.
[34] E. Androulakis, & Galanis, G., A two-step hybrid system towards optimized wave height forecasts, Stochastic Environmental Research and Risk Assessment, (2021) 1-14.
[35] S.C. James, Zhang, Y., & O'Donncha, F., A machine learning framework to forecast wave conditions, Coastal Engineering, 137 (2018) 1-10.
[36] N.K.K. Pani, Jha, V. A., Bai, L., Cheng, L., & Zhao, T., A Hybrid Machine Learning Approach to Wave Energy Forecasting, in: 2021 North American Power Symposium (NAPS), IEEE, College Station, TX, USA, 2021, pp. 1-5.
[37] M.E. Torres, M.A. Colominas, G. Schlotthauer, P. Flandrin, A complete ensemble empirical mode decomposition with adaptive noise, in: 2011 IEEE international conference on acoustics, speech and signal processing (ICASSP), IEEE, 2011, pp. 4144-4147.
[38] W. Zhang, Z. Qu, K. Zhang, W. Mao, Y. Ma, X. Fan, A combined model based on CEEMDAN and modified flower pollination algorithm for wind speed forecasting, Energy Conversion and Management, 136 (2017) 439-451.
[39] H.H. Aly, Intelligent optimized deep learning hybrid models of neuro wavelet, Fourier Series and Recurrent Kalman Filter for tidal currents constitutions forecasting, Ocean Engineering, 218 (2020) 108254.
[40] J. Zhang, Meng, Y., Wei, J., Chen, J., & Qin, J. , A novel hybrid deep learning model for sugar price forecasting based on time series decomposition, Mathematical Problems in Engineering, 2021 (2021).
[41] C. Ni, & Ma, X. , An integrated long-short term memory algorithm for predicting polar westerlies wave height, Ocean Engineering, 215 (2020) 107715.
[42] M. Ahmed, R.C. Deo, N. Raj, A. Ghahramani, Q. Feng, Z. Yin, L. Yang, Deep Learning Forecasts of Soil Moisture: Convolutional Neural Network and Gated Recurrent Unit Models Coupled with Satellite-Derived MODIS, Observations and Synoptic-Scale Climate Index Data, Remote Sensing, 13 (2021) 554.
[43] A.A.M. Ahmed, R. Deo, A. Ghahramani, N. Raj, Q. Feng, Z. Yin, L. Yang, LSTM integrated with Boruta-random forest optimiser for soil moisture estimation under RCP4. 5 and RCP8. 5 global warming scenarios, Stochastic Environmental Research and Risk Assessment, (2021) 1-31.
[44] A. Ahmed, M.H. Ahmed, S.K. Saha, O. Ahmed, A. Sutradhar, Optimization algorithms as training approach with hybrid deep learning methods to develop an ultraviolet index forecasting model, Stochastic Environmental Research and Risk Assessment, (2022) 1-29.
[45] P. Li, Zhou, K., Lu, X., & Yang, S., A hybrid deep learning model for short-term PV power forecasting, Applied Energy, 259 (2020) 114216.
[46] Y. Shi, Feng, H., Geng, X., Tang, X., & Wang, Y. , A survey of hybrid deep learning methods for traffic flow prediction, in: Proceedings of the 2019 3rd international conference on advances in image processing, 2019, pp. 133-138.
[47] U.E. Akpudo, & Hur, J. W., D-dCNN: A Novel Hybrid Deep Learning-Based Tool for Vibration-Based Diagnostics, Energies, 14 (2021) 5286.
[48] J.I.Z. Chen, & Smys, S., Social multimedia security and suspicious activity detection in SDN using hybrid deep learning technique, Journal of Information Technology, 2 (2020) 108-115.
[49] X. Dong, Qian, L., & Huang, L., Short-term load forecasting in smart grid: A combined CNN and K-means clustering approach, in: IEEE (Ed.) 2017 IEEE international conference on big data and smart computing (BigComp), 2017, pp. 119-125.
[50] K. Amarasinghe, Marino, D. L., & Manic, M. , Deep neural networks for energy load forecasting, in: IEEE (Ed.) 2017 IEEE 26th International Symposium on Industrial Electronics (ISIE) 2017, pp. 1483-1488.

[51] S. Ghimire, R.C. Deo, N. Raj, J. Mi, Deep solar radiation forecasting with convolutional neural network and long short-term memory network algorithms, Applied Energy, 253 (2019).
[52] A.A.M. Ahmed, R.C. Deo, Q. Feng, A. Ghahramani, N. Raj, Z. Yin, L. Yang, Hybrid deep learning method for a week-ahead evapotranspiration forecasting, Stochastic Environmental Research and Risk Assessment, (2021) 1-19.
[53] E. Sharma, Deo, R. C., Prasad, R., Parisi, A. V., & Raj, N. , Deep air quality forecasts: suspended particulate matter modeling with convolutional neural and long short-term memory networks, IEEE Access, 8 (2020) 209503 - 209516.
[54] L.D. Soares, & Franco, E. M. C. , BiGRU-CNN neural network applied to short-term electric load forecasting, Production, 32 (2021).
[55] S. Ghimire, R.C. Deo, H. Wang, M.S. Al-Musaylh, D. Casillas-Pérez, S. Salcedo-Sanz, Stacked LSTM Sequence-to-Sequence Autoencoder with Feature Selection for Daily Solar Radiation Prediction: A Review and New Modeling Results, Energies, 15 (2022) 1061.
[56] S.B. Boubaker, Mohamed & Mellit, A. & Lefza, Ayoub & Kahouli, A. & Kolsi, Lioua. , Deep Neural Networks for Predicting Solar Radiation at Hail Region, Saudi Arabia, (2021).
[57] A.A.M. Ahmed, R.C. Deo, N. Raj, A. Ghahramani, Q. Feng, Z. Yin, L. Yang, Deep Learning Forecasts of Soil Moisture: Convolutional Neural Network and Gated Recurrent Unit Models Coupled with Satellite-Derived MODIS, Observations and Synoptic-Scale Climate Index Data, Remote Sensing, 13 (2021) 554.
[58] S.J.J. Jui, A.A.M. Ahmed, A. Bose, N. Raj, E. Sharma, J. Soar, M.W.I. Chowdhury, Spatiotemporal Hybrid Random Forest Model for Tea Yield Prediction Using Satellite-Derived Variables, Remote Sensing, 14 (2022) 805.
[59] J. Bergstra, D. Yamins, D.D. Cox, Hyperopt: A python library for optimizing the hyperparameters of machine learning algorithms, in: Proceedings of the 12th Python in science conference, Citeseer, 2013, pp. 20.
[60] C. Thornton, F. Hutter, H.H. Hoos, K. Leyton-Brown, Auto-WEKA: Combined selection and hyperparameter optimization of classification algorithms, in: Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining, 2013, pp. 847-855.
[61] L. Bottou, Large-scale machine learning with stochastic gradient descent, in: Proceedings of COMPSTAT'2010, Springer, 2010, pp. 177-186.
[62] R. Solgi, H.A. Loáiciga, M. Kram, Long short-term memory neural network (LSTM-NN) for aquifer level time series forecasting using in-situ piezometric observations, Journal of Hydrology, 601 (2021) 126800.
[63] B. E, B. D, V.K. Elumalai, V. R, Automatic and non-invasive Parkinson's disease diagnosis and severity rating using LSTM network, Applied Soft Computing, 108 (2021) 107463.
[64] K. Andersson, C.W. Oosterlee, A deep learning approach for computations of exposure profiles for high-dimensional Bermudan options, Applied Mathematics and Computation, 408 (2021) 126332.
[65] D.C. Plaut, Experiments on Learning by Back Propagation, (1986).
[66] F. Chollet, Keras: Deep learning library for theano and tensorflow, URL: https://keras. io/k, 7 (2015) T1.
[67] P. Goldsborough, A tour of tensorflow, arXiv preprint arXiv:1610.01178, (2016).
[68] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, Tensorflow: a system for large-scale machine learning, in: OSDI, 2016, pp. 265-283.
[69] K.B. Matthews, M. Rivington, K. Blackstock, G. McCrum, K. Buchan, D.G. Miller, Raising the bar? – The challenges of evaluating the outcomes of environmental modelling and software, Environmental Modelling & Software, 26 (2011) 247-257.
[70] H.V. Gupta, H. Kling, K.K. Yilmaz, G.F. Martinez, Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, Journal of hydrology, 377 (2009) 80-91.
[71] J.E. Nash, J.V. Sutcliffe, River flow forecasting through conceptual models part I—A discussion of principles, Journal of hydrology, 10 (1970) 282-290.
[72] T. Chai, R.R. Draxler, Root mean square error (RMSE) or mean absolute error (MAE), Geoscientific Model Development Discussions, 7 (2014) 1525-1534.
[73] C.J. Willmott, K. Matsuura, Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance, Climate research, 30 (2005) 79-82.
[74] S. Shamshirband, A. Mosavi, T. Rabczuk, N. Nabipour, K.-w. Chau, Prediction of significant wave height; comparison between nested grid numerical model, and machine learning models of artificial neural networks, extreme learning and support vector machines, Engineering Applications of Computational Fluid Mechanics, 14 (2020) 805-817.

[75] A.A.M. Ahmed, Development of Deep Learning Hybrid Models for Hydrological Predictions, in, University of Southern Queensland, 2022.
[76] A.A.M. Ahmed, R.C. Deo, Q. Feng, A. Ghahramani, N. Raj, Z. Yin, L. Yang, Deep learning hybrid model with Boruta-Random forest optimiser algorithm for streamflow forecasting with climate mode indices, rainfall, and periodicity, Journal of Hydrology, 599 (2021).
[77] N. Raj, Z. Gharineiat, A.A.M. Ahmed, Y. Stepanyants, Assessment and Prediction of Sea Level Trend in the South Pacific Region, Remote Sensing, 14 (2022) 986.
[78] S. Ghimire, R.C. Deo, N.J. Downs, N. Raj, Global solar radiation prediction by ANN integrated with European Centre for medium range weather forecast fields in solar rich cities of Queensland Australia, Journal of Cleaner Production, 216 (2019) 288-310.
[79] A.A.M. Ahmed, N. Bailek, L. Abualigah, K. Bouchouicha, A. Kuriqi, A. Sharifi, P. Sareh, A.M.G. Al khatib, P. Mishra, I. Colak, E.-S.M. El-kenawy, Global control of electrical supply: A variational mode decomposition-aided deep learning model for energy consumption prediction, Energy Reports, 10 (2023) 2152-2165.
[80] R.C. Deo, A. Ahmed, D. Casillas-Perez, S.A. Pourmousavi Kani, G. Segal, Y. Yu, S. Salcedo-Sanz, Cloud Cover Bias Correction in Numerical Weather Model Simulations for Solar Energy Monitoring and Forecasting Systems: A New Kernel Ridge Regression Approach, David and Pourmousavi Kani, Seyyed Ali and Segal, Gary and Yu, Yanshan and Salcedo-Sanz, Sancho, Cloud Cover Bias Correction in Numerical Weather Model Simulations for Solar Energy Monitoring and Forecasting Systems: A New Kernel Ridge Regression Approach (September 28, 2022), (2022).
[81] S.K. Prabhakar, D.-O. Won, Medical text classification using hybrid deep learning models with multihead attention, Computational intelligence and neuroscience, 2021 (2021).
[82] S.M. Mousavi, M. Ghasemi, M. Dehghan Manshadi, A. Mosavi, Deep learning for wave energy converter modeling using long short-term memory, Mathematics, 9 (2021) 871.

Appendix:

**Table A1:** Optimally selected hyperparameters of deep learning models. ReLU stands for Rectified Linear Units, SGD stands for stochastic gradient descent optimiser

| Model Hyper-parameter Names | Optimal Hyper-Parameters | | |
|---|---|---|---|
| | **CLSTM-BiGRU** | **CLSTM** | **BiGRU** |
| Convolution Layer 1 (C1) | 70 | 70 | |
| C1- Activation function | ReLU | ReLU | |
| C1-Pooling Size | 1 | 1 | |
| Convolution Layer 2 (C2) | 60 | 60 | |
| C2- Activation function | ReLU | ReLU | |
| C2-Pooling Size | 1 | 1 | |
| Convolution Layer 3 (C3) | 80 | 50 | |
| LSTM Layer 1 (L1) | 70 | 60 | |
| L1- Activation function | ReLU | Tanh | |
| LSTM Layer 2 (L2) | 70 | 60 | |
| L2- Activation function | ReLU | ReLU | |
| BiGRU Layer 1 (L1) | 65 | | 50 |
| L1- Activation function | Tanh | | Softmax |
| Drop-out rate | 0.2 | 0.2 | 0.2 |
| Optimiser | SGD | SGD | SGD |
| Padding | Same | Same | Same |
| Batch Size | 5 | 7 | 6 |

| | | | |
|---|---|---|---|
| Epochs | 1000 | 1000 | 500 |

## Credit authorship contribution statement

**Abul A. Masrur Ahmed:** Writing - original draft, conceptualization, methodology, software, editing, proofreading, model development, and application. **S Janifer Jabin Jui:** Writing, editing, and proofreading. **Mohanad S. AL-Musaylh:** Writing, editing, and proofreading. **Nawin Raj:** Writing, editing **Reepa Saha:** Writing, and proofreading. **Ravinesh C Deo:** Editing, and proofreading. **Sanjoy Kumar Saha:** proofreading

**Declaration of interests**

☒ 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.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

## Highlights

- We propose deep learning CLSTM-BiGRU hybrid model to predict significant wave heights
- CLSTM-BiGRU model is tested at multiple forecast (30 minutes, 2 h, 3 h and 6 h) horizons
- CLSTM-BiGRU analysed at wave energy sites in Queensland show model's overall efficacy
- CLSTM-BiGRU model has positive implications in wave and ocean energy generation
- CLSTM-BiGRU model is useful for ocean monitoring and wave energy resource evaluations