IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 24 July 2023, accepted 18 August 2023, date of publication 29 August 2023, date of current version 11 October 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3309882

RESEARCH ARTICLE

New Improved Wave Hybrid Models for Short-Term Significant Wave Height Forecasting

PRITAM ANAND, SHANTANU JAIN, AND HARSH SAVALIYA
Dhirubhai Ambani Institute of Information and Communication Technology (DA-IICT), Gandhinagar 382007, India
Corresponding author: Pritam Anand (pritam_anand@daiict.ac.in)

ABSTRACT In this paper, we have developed a series of wave hybrid models for significant wave height prediction. Our developed hybrid models uses a triplet of signal decomposition method, regression model and meta-heuristic algorithm. We have used the ϵ-Support Vector Regression (ϵ-SVR), Least Squares Support Vector Regression (LS-SVR), Long Short-Term Memory (LSTM) and Large-margin Distribution Machine based Regression (LDMR) model for the regression task. For signal decomposition methods, we have considered the Wavelet Decomposition (WD), Empirical Mode Decomposition (EMD) and Variational Mode Decomposition (VMD) method. Apart from this, we have also used the Particle Swarm Optimization (PSO) method to tune the parameters of the used regression model in our wave hybrid models. Till now, the VMD method and LDMR model have not been used in any wave hybrid model. We have evaluated the performance of our developed wave hybrid models on time-series significant wave heights, collected from four different buoys using the different evaluation criteria. After the detailed statistical analysis of the obtained numerical results, we conclude that the VMD-PSO-LDMR based wave hybrid model obtain best performance on six datasets out of seven considered datasets. Also, the VMD based wave hybrid models can obtain better performance than other decomposition based hybrid models. Further, we also conclude from our numerical results that the LSTM model outperforms the SVR, LS-SVR and LDMR based hybrid models if we do not decompose the significant wave height signals apriori. But, when we decompose the SWH time-series signals using a particular decomposition method, then SVR, LS-SVR and LDMR based hybrid models tend to improve their prediction ability significantly.

INDEX TERMS Support vector regression, machine learning, hybrid model, renewable energy, wave height prediction, wavelet transform, empirical mode decomposition.

I. INTRODUCTION

The renewable energy technologies have obtained a rapid growth in recent years to meet the fossil fuels crisis and reduce the climate threats. The wave energy is one of the promising and reliable clean energy sources for catering the future energy demands. It has broader prospects and higher energy density compared with other clean energy sources like solar and wind [1], [2]. The Significant Wave Height (SWH) is very fundamental and necessary parameter in the determination of wave power level. It is obtained by taking the average of one-third highest ocean waves observed in a given time interval. The significant wave height prediction plays a very crucial role in wave power generation. A reliable and efficient prediction of significant wave height of few hours ahead is required for the enhancing the performance of wave energy converter [3]. Apart from this, the hourly efficient prediction of SWH can significantly help to improve the decisions in maritime and off-shore activities. But, the highly random and chaotic nature of ocean waves make the significant wave height prediction task difficult and challenging.

Traditionally, researchers have used the energy balance equations to obtain the wave forecast over a large spatial and temporal domain using other different ocean variables [4]. But, these models are computationally expensive and complex which limits their applicability specifically for hourly SWH predictions for a particular location. Researchers

The associate editor coordinating the review of this manuscript and approving it for publication was Akin Tascikaraoglu.

have preferred the data driven models for short term wave height forecasting which estimates the SWH of next few hours using the time-series data.

The parametric time-series models such as Auto-Regressive (AR), Auto-Regressive Moving Average (ARMA), Auto-Regressive Integrated Moving Average (ARIMA), etc. have been well exploited by researchers for SWH forecasting. Some important literature that makes use of these models effectively for wave prediction are ([5], [6], [7], and [8]). The ARMA model obtains good prediction only when the time series is stationary. For non-stationary time series, the ARIMA model is preferred over the ARMA model. But, all of these parametric time-series models fail to capture the complex non-linear relationship. They also assume that the data noise is from the normal distribution which may not be relevant particularly for SWH time-series data.

The non-parametric modern machine learning models have shown more promising results, especially for short-term SWH forecasting tasks as they do not make any assumption regarding the noise distribution. Researchers have used the different variants of neural architecture and obtained the reliable short-term prediction of SWH. Deo et al. have used the three-layer feed-forward ANN model for the prediction of SWH [9]. In [10], the author has improved the SWH prediction by using the ANN model. In [11], authors have used the Extreme Learning Machine (ELM) along with grouping genetic algorithm for searching the effective features set and obtaining the effective forecast of short-term SWH. The ELM is a feed-forward neural network in which hidden layer weights are selected randomly. Unlike the ANN model, it does not require the costly back-propagation method for tunning all weights. Only the output weights are obtained using a single matrix inversion. But, the solution obtained by the ELM depends upon the randomly selected hidden layer weights. For obtaining more robust prediction, authors have used an ensemble of ELM models to obtain the daily forecast of SWH in [12].

The Recurrent Neural Network (RNN) is a sequential learning model which is capable of identifying the temporal pattern in the data. RNN is a very effective and popular approach for SWH prediction among researchers [13], [14]. But, the RNN architecture suffers from vanishing gradient problems and fails to learn the long-term dependencies in data. The Gated Recurrent Unit (GRU) [15] improves the RNN by using the gate mechanism for controlling the flow of information and getting rid of the vanishing gradient problem. The different variants of GRU have been used to obtain the estimate of SWH in [14] and [16]. The Long Short Term Memory Network (LSTM) [17] is a more complex neural architecture than GRU and involves more gates for controlling the flow of information. Some important research works which make the use of LSTM model for forecasting SWH are [18], [19], [20], and [21]. But, all of these sequential learning models are complex, require a large number of parameters to be learned and are prone to over-fitting. It makes them not a suitable choice for a single time stamp ahead SWH forecasting with small or medium sample size training sets.

Support Vector Regression (SVR) models [22] are popular choice of researchers in hourly forecasting tasks due to its simplicity. These models are based on statistical learning theory [23] and can minimize an appropriate trade-off between empirical risk and model complexity in their optimization problem. Further, these models can obtain the global optimal solution which remain missing in other neural network architecture based machine learning models. The different variants of the SVR models have been used in the efficient predictions of the SWH in several research works. Some of them are [24], [25], [26], and [27].

But, there are two main challenges associated with SVR forecasting. The first one requires the efficient tuning of user-defined parameters of SVR models. The second challenge requires an informative set of features for obtaining good forecasts with SVR models. SVR models may obtain poor forecasts with raw SWH time-series signals.

To get rid of these challenges, researchers have started to use hybrid wave models for improving prediction in recent years. In wave hybrid models, a decomposition technique is used to decompose the time-series data into more informative features. Thereafter, these features have been supplied to machine learning algorithms for obtaining more accurate predictions. Researchers have proposed the variants of the hybrid models by considering the different combinations of decomposition methods and machine learning models and studied their behavior on different SWH datasets.

In SWH forecasting literature, the two decomposition method namely Wavelet Decomposition (WD) and Empirical Mode Decomposition (EMD) [28] has been used dominantly along with different machine learning algorithm to develop the efficient hybrid models.

Prahlada and Deka have proposed a hybrid model with the Wavelet Decomposition (WD) method for the prediction of SWH for the lead time up to 48 hours [29]. Other relevant hybrid models which make use of the wavelet decomposition technique are [16], [30], and [31]. Duan et al. have used the EMD [28] technique along with ϵ-SVR model [22] for obtaining the short-term forecasting of significant wave height [4]. Tang et al. have used the Least Square Support Vector Regression (LS-SVR) [32] models along with the EMD technique for short-time wave height prediction. Musaylh et al. used two phases PSO-based SVR model with an improved EMD method for multiple origin electricity demand forecasting [33]. Ali and Parasad have used an improved ensemble EMD method along with the ELM model to obtain the forecast of significant wave height 30 minutes ahead [34]. But, the wavelet and EMD decomposition method may be sensitive to noise and sampling.

Also, only ϵ-SVR and LS-SVR models are exploited in the hybrid wave model in the literature. The ϵ-SVR is preferred when noise is uniform whereas the LS-SVR model is preferred for normal noise in data.

In this paper, we have developed a sequence of hybrid models and compared their performance for hourly significant wave height predictions. These hybrid models use different regression models along with a signal decomposition method for obtaining their prediction. Our developed hybrid models include the wave hybrid models which use the LDMR model [35] for prediction. The LDMR model [35] is a more general SVR model which offers more flexibility and the ability to handle the mixture of noise. Also, some of our developed wave hybrid models use the Variational Mode Decomposition (VMD) [36] method for signal decomposition which is more robust than the wavelet and EMD method.

We summarize the contribution of this paper as follows.

1) We have designed sixteen different wave hybrid models and compared their performances for hourly SWH forecasting. Our wave hybrid models use the LDMR [35], LS-SVR [32], ϵ-SVR [22] and LSTM model [17] for prediction. We have also used the Wavelet Decomposition (WD), EMD and VMD [36] methods to decompose the time-series data into more suitable features vectors in our hybrid models.
2) The LDMR model [35] can minimize a good trade-off among the ϵ-insensitive loss function, quadratic loss function and model complexity in its optimization problem which enables it to obtain the better prediction than ϵ-SVR and LS-SVR models. To the best of our knowledge, no wave hybrid models have exploited the LDMR model for improving the prediction of SWH. In this paper, we have attempted to use the LDMR model along with WD, EMD and VMD methods.
3) Researchers have recently shown that the use of VMD [36] technique in wind hybrid models results in significant improvement in short-term wind speed prediction. Motivated by this, we have also used the VMD decomposition method in our different wave hybrid models and obtained an improvement in SWH prediction over other decomposition-based hybrid models.
4) We have also used the Particle Swarm Optimization (PSO) algorithm [37] in our wave hybrid models to obtain the appropriate parameter values for used regression models.
5) We have collected the hourly SWH time-series data from four different ocean buoys located at different geographical regions and evaluated our developed wave hybrid models namely PSO-LDMR, PSO-SVR, PSO- LS-SVR, LSTM, WD-PSO-LDMR, WD-PSO-SVR, WD-PSO-LS-SVR, WD-LSTM, EMD-PSO-LDMR, EMD-PSO-SVR, EMD-PSO-LS-SVR, EMD-LSTM, VMD-PSO-LDMR, VMD-PSO-SVR, VMD-PSO-LS-SVR and VMD-LSTM models using different evaluation criteria. After the brief analysis of the obtained numerical results, we conclude that the LDMR model based wave hybrid models can obtain better predictions than other regression model-based hybrid models. Also, the VMD method based hybrid models can obtain better prediction than other decomposition based hybrid models. Further, we also conclude from our numerical results that the LSTM model outperforms the SVR, LS-SVR and LDMR based hybrid models if we do not decompose the significant wave height signals apriori. But, when we decompose the SWH time series signals using a particular decomposition method, then SVR, LS-SVR and LDMR based hybrid models tend to improve their prediction ability significantly and outperform the LSTM model. Contrary to them, LSTM based models fail to improve their prediction ability significantly when signals are decomposed apriori.

We have organized the rest of this paper as follows. In Section-II and Section- III, we have briefly described the used signal and machine learning models and decomposition methods respectively. In Section-IV, we have developed different wave hybrid models and briefly described their methodologies and implementation details. In Section-V, we have presented the numerical results and their brief analysis. Section-VI concludes this paper.

II. MACHINE LEARNING MODELS

In this sections, we shall briefly describe the used machine learning models in our wave hybrid models.

Given the training set $T = \{(x_i, y_i) : x_i \in \mathbb{R}^n, y_i \in \mathbb{R},\ i = 1, 2, \ldots, l\ \}$, SVR models minimizes a linear combination of the loss function and regularization term to obtain a linear estimate $f(x) : w^T x + b,\ w \in \mathbb{R}^n, b \in R$. For estimating the non-linear function, it finds $f(x) : w^T\phi(x) + b = K(x^T, A^T)u + b$ in feature space, where K is an appropriate kernel satisfying Mercer condition [38] such that $\phi(x_i)^T\phi(x_j) = K(x_i, x_j)$. The matrix A is $l \times n$ data matrix containing the l data points in $\mathbb{R}^n$.

A. ϵ-SUPPORT VECTOR REGRESSION MODEL

The ϵ-SVR model [22], [23] minimizes the ϵ-insensitive loss function along with the $\frac{1}{2}w^T w$ regularization. It finds the solution of the following optimization problem

$$\min_{w,b} \frac{1}{2}w^T w + C\sum_{i=1}^{l} |y_i - (w^T\phi(x_i) + b)|_\epsilon, \tag{1}$$

where $|y_i - (w^T\phi(x_i)+b)|_\epsilon = \max(0, |y_i - (w^T\phi(x_i)+b)| - \epsilon)$ is the ϵ-insensitive loss function which can ignore an error up to ϵ. After introducing the slack variable ${\xi_1}^i$ and ${\xi_2}^i$ for $i = 1, 2, .., l$, the ϵ-SVR problem (1) is solved by converting the QPP:

$$\begin{aligned} \min_{(w,b,\xi_1,\xi_2)} \quad & \frac{1}{2}\|w\|^2 + Ce^T(\xi_1 + \xi_2) \\ \text{subject to,} \quad & Y - (\phi(A)w + eb) \leq \epsilon e + \xi_1, \\ & (\phi(A)w + eb) - Y \leq \epsilon e + \xi_2, \\ & \xi_1 \geq 0,\ \xi_2 \geq 0 \end{aligned} \tag{2}$$

Here $C > 0$ and ϵ are the user specified positive parameters that balances the trade off between the training error and the flatness of the regression function. And the vector Y and is $l \times 1$ vector containing the responses y_i. Further, $\xi_1 = [\xi_1^1, \xi_1^2, \ldots, \xi_1^l]$ and $\xi_2 = [\xi_2^1, \xi_2^2, \ldots, \xi_2^l]$ are l dimensional column vectors.

Using the Karush-Kuhn-Tucker(KKT) conditions, the Wolfe dual problem of the primal problem (2) can be obtained as

$$\begin{aligned} \min_{(\alpha_1,\alpha_2)} \ & \frac{1}{2}(\alpha_1-\alpha_2)^T K(A,A^T)(\alpha_1-\alpha_2) \\ & - Y^T(\alpha_1-\alpha_2) + \epsilon e^T(\alpha_1+\alpha_2) \\ \text{subject to,} \quad & e^T(\alpha_1-\alpha_2) = 0, \\ & 0 \le \alpha_1 \le Ce, \\ & 0 \le \alpha_2 \le Ce. \end{aligned} \tag{3}$$

where $\alpha_1 = [\alpha_1^1, \alpha_1^2, \ldots, \alpha_1^l,]$ and $\alpha_2 = [\alpha_2^1, \alpha_2^2, \ldots, \alpha_2^l]$ are l dimensional vectors for Lagrangian multipliers. After obtaining the optimal values of the Lagrangian multipliers vectors, α_1 and α_2, from (3), the estimated regressor is obtained for the given $x \in R^n$ as follow.

$$\begin{aligned} f(x) &= w^T\phi(x) + b \\ &= K(x^T, A^T)(\alpha_1-\alpha_2) + b. \end{aligned} \tag{4}$$

B. LEAST SQUARES SUPPORT VECTOR REGRESSION MODEL

The LS-SVR model [32] minimizes the quadratic loss function along with $\frac{1}{2}w^Tw$ regularization term. It minimizes

$$\min_{w,b} \frac{1}{2}w^Tw + C_1\sum_{i=1}^{l}(y_i - (w^T\phi(x_i)+b))^2, \tag{5}$$

in its optimization problem along with the regularization term $\frac{1}{2}||w||^2$. The optimization problem of the LS-SVR model can be expressed as

$$\begin{aligned} \min_{w,b,\xi} \ & \frac{1}{2}\|w\|^2 + C_1\sum_{i=1}^{l}(\xi_i^2) \\ \text{subject to,} \quad & y_i - (\phi(A_i)w + b) = \xi_i, \quad i = 1,2,\ldots,l, \end{aligned} \tag{6}$$

where $C_1 > 0$ is a user defined parameter. The solution of problem (6) can be obtained by solving the following system of equations.

$$\begin{bmatrix} 0 & e^T \\ e & K(A,A^T) + \frac{2}{C_1}I \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ Y \end{bmatrix} \tag{7}$$

After obtaining the (b, α) by solving the above system of equations, we can estimate our regression function for a given $x \in \mathbb{R}$ using

$$f(x) = w^T\phi(x) + b = K(x^T, A^T)\alpha + b. \tag{8}$$

C. LARGE-MARGIN DISTRIBUTION MACHINE REGRESSION MODEL

The least square loss function used in LS-SVR model makes it to perform optimal in case of the presence of normal noise in the data. The ϵ-insensitive loss function in the ϵ-SVR model makes it perform optimal in the presence of the uniform noise.

The LDMR model [35] minimizes the linear combination of the quadratic loss and ϵ -insensitive loss function for measuring the empirical risk along with $\frac{1}{2}w^Tw$ regularization term. It enables it to perform better than both ϵ-SVR and LS-SVR model. The LDMR model [35] minimizes

$$\begin{aligned} \min_{w,b} \frac{1}{2}w^Tw + \frac{1}{2}\sum_{i=1}^{l}(y_i - (w^T\phi(x)+b))^2 \\ + C\sum_{i=1}^{l}|y_i - (w^T\phi(x)+b)|_\epsilon, \end{aligned} \tag{9}$$

which can be converted to the following QPP

$$\begin{aligned} \min_{w,b,\xi_1,\xi_2} \ & \frac{c}{2}||w||^2 + \frac{1}{2}||Y - (Aw+eb)||^2 \\ & + Ce^T(\xi_1+\xi_2) \\ \text{subject to,} \quad & Y - (Aw+eb) \le e\epsilon + \xi_1, \\ & (Aw+eb) - Y \le e\epsilon + \xi_2, \\ & \xi_1 \ , \xi_2 \ge 0, \end{aligned} \tag{10}$$

where C, c and ϵ are the user specified positive parameters. We prefer to solve the Wolfe dual of the primal problem (10) which can be obtained as follows,

$$\begin{aligned} \min_{\alpha_1,\alpha_2} \ & \frac{1}{2}(\alpha_1-\alpha_2)^T H(c(\alpha_1-\alpha_2) \\ & + Y^T H(cI_0 + H^TH)^{-1}H^T(\alpha_1-\alpha_2) \\ & - Y^T(\alpha_1-\alpha_2) + \epsilon e^T(\alpha_1+\alpha_2) \\ \text{subject to,} \ & 0 \le \alpha_1 \le Ce, 0 \le \alpha_2 \le Ce. \end{aligned} \tag{11}$$

Here $H = [K, e]$ is a augmented matrix and $I_0 = \begin{bmatrix} I & 0 \\ & . \\ & . \\ 0 \ldots 0 \end{bmatrix}$ where I is $n \times n$ identity matrix. After obtaining the solution of problem (11), the solution vector $u = [w, \ \ b]^T$ can obtained using

$$u = [w, \ \ b]^T = (cI_0 + H^TH)^{-1}H^T(\alpha_1 - \alpha_2 + Y) \tag{12}$$

and the regression function is estimated using (8).

D. LONG SHORT-TERM MEMORY MODEL

Long Short-Term Memory (LSTM) is a recurrent neural network (RNN) architecture that can handle long-range dependencies in sequential data. Unlike feedforward neural networks, LSTM incorporates feedback loops that enable it to learn from both current and past inputs. A typical

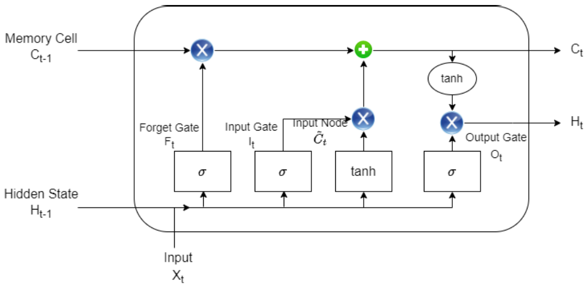


FIGURE 1. **LSTM architecture.**

LSTM unit consists of four components: a memory cell, an input gate, an output gate, and a forget gate. We can see architecture in Figure 1. These components interact to regulate the information flow within the unit and to maintain a persistent state over time.

Figure 1 illustrates the different gates that regulate the information flow in an LSTM block. The input gate determines the new information that will be incorporated into the cell state, as shown in (13). The forget gate controls the information that will be discarded from the previous cell state, as shown in (14). The cell state is the key component of the LSTM block that stores and updates the relevant information over time, as shown in (17). The output gate decides the information that will be emitted from the current LSTM block, as shown in (15). The hidden state of the LSTM block is updated according to (18).

$$I_t = \sigma(W_{xi} * X_t + W_{hi} * h_{t-1} + b_i) \tag{13}$$
$$F_t = \sigma(W_{xf} * X_t + W_{hf} * h_{t-1} + b_f) \tag{14}$$
$$O_t = \sigma(W_{xo} * X_t + W_{ho} * h_{t-1} + b_o) \tag{15}$$
$$\tilde{C}_t = I_t * (tanh(W_{xc} * X_t + W_{hc} * h_{t-1} + b_c)) \tag{16}$$
$$C_t = F_t * C_{t-1} * \tilde{C}_t \tag{17}$$
$$H_t = O_t * tanh(C_t) \tag{18}$$

III. METHODOLOGY OF PROPOSED WAVE HYBRID MODELS FOR SWH PREDICTION

In this section, we shall develop novel variants of wave hybrid model and briefly describe the methodology and implementation details of them. Our wave hybrid models involve five general successive steps which we have presented in the Figure 2. Further, we briefly describe these steps as below.

1) **Decomposition of signal:-** After receiving the time-series SWH data, we need to decompose it into more informative data signals. A standalone regression model performs poorly with raw time-series signifi. The performance of a regression model depends upon the quality of supplied feature set. We have used popular decomposition techniques to decompose the raw time-series significant wave height into more strong and informative features which is supplied to regression model for efficient prediction The decomposition of the signal is important as it helps to obtain the more informative features for our machine learning method. In our developed wave hybrid models, we have used the WD, EMD or VMD decomposition method to decompose the time series SWH $x(t)$ into $[x_1(t), x_2(t), \ldots, x_n(t)]$ component data signals.
2) **Attributes/features composition:-** After decomposing the SWH signal, we need to prepare the dataset for machine learning model. For this, we need to compose the feature set and corresponding response values. For i^{th} signal component, we compose the p-dimensional $[x_i(t), x_i(t-1), x_i(t-2), .., x_i(t-(p-1))]$ features for the prediction of $x_i(t+1)$.
3) **Parameters Tunning:-**We need to train the n machine learning models to obtain the prediction for each of $x_i(t+1)$ for $i = 1, 2, .., n$, using our composed feature sets. But, before training a machine learning model, we need to tune its parameters. We have used the PSO algorithm to tune the parameters of our machine learning models.
4) **Prediction:-** After selecting the appropriate values of parameters for our machine learning model, we finally train them for obtaining the prediction of $x_i(t+1)$ for $i = 1, 2, .., n$.
5) **Aggregation:-** After obtaining the prediction of $x_i(t+1)$ for $i = 1, 2, .., n$, we aggregate them to obtain the final prediction for $x(t+1)$.

Now we detail the working of our proposed wave hybrid models as follows.

1) **LDMR based hybrid models**:- In LDMR based hybrid model, we use the LDMR model for regression tasks. In Figure 3, we have shown the steps of training the PSO-LDMR model for a given dataset. In PSO-LDMR model, the choice of kernel parameter q, trade-off parameters C and c has been supplied by the PSO algorithm after tuning them efficiently. We have listed the flow chart of PSO algorithm in Figure6. Now we detail the working of the LDMR based hybrid model one by one.

   a) **PSO-LDMR model** :- For given time-series SWH data $(X_1, X_2, \ldots, X_t)$, the PSO-LDMR model construct the training set (A, Y) using $A = \begin{bmatrix} X_1, & X_2, & \ldots, & , X_p \\ X_2, & , X_3, & \ldots & , X_{p+1} \\ \ldots & \ldots & \ldots \\ X_{t-p}, & X_{t-p+1}, & \ldots, & X_{t-1} \end{bmatrix}$ and $Y = \begin{bmatrix} X_{p+1} \\ X_{p+2} \\ \ldots \\ X_t \end{bmatrix}$. Using the training set (A, Y), the PSO-LDMR model is trained using the steps of flowchart given in Figure 3. For the prediction of the SWH x_{t+1}, the PSO-LDMR model is estimated for the test point $[X_{t-p+1}, X_{t-p+2}, \ldots, X_t]$.

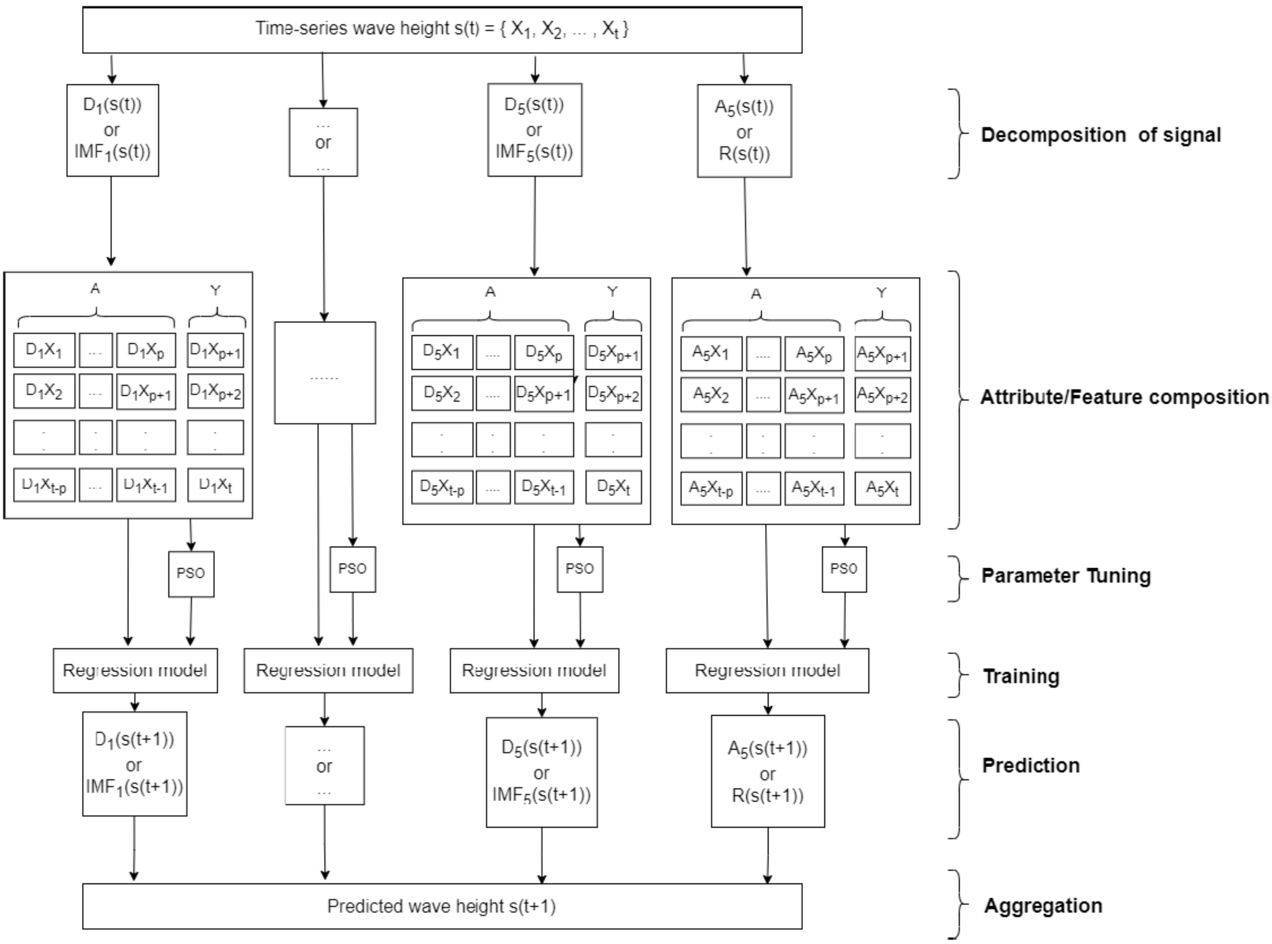


FIGURE 2. Wave hybrid model.

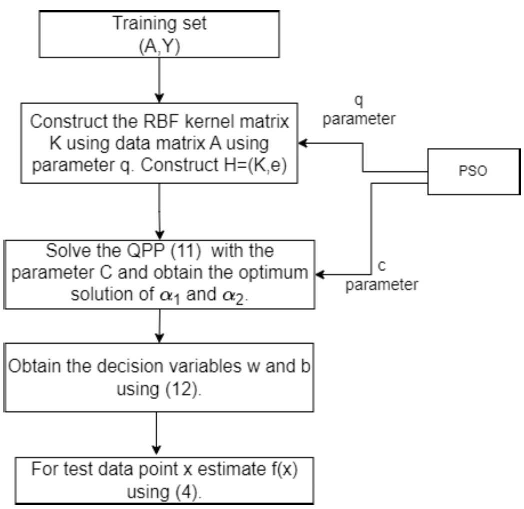


FIGURE 3. LDMR model flowchart.

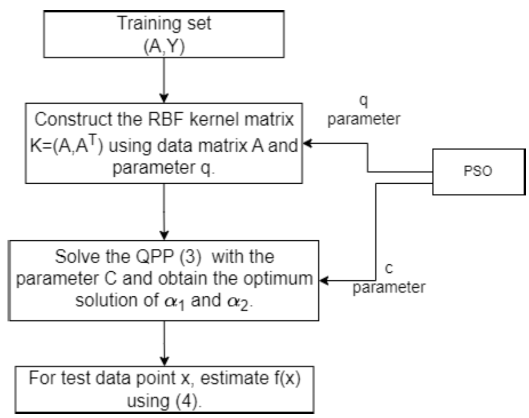


FIGURE 4. SVR model flowchart.

b) **WD-PSO-LDMR model**:- In WD-PSO-LDMR model, we first decompose the input time-series SWH signal $s(t) = (X_1, X_2, \ldots, X_t)$ into five high frequency detail signals $(D_1, D_2, D_3, D_4, D_5)$ and a low frequency approximate signal A_5. Thereafter, we construct the six different training sets by using each decomposed signal as shown in the Attribute/Feature Composition step of Figure 2. For each training set, we train a PSO-LDMR model separately. For the prediction of the SWH x_{t+1}, we consider the different decomposition of test point $[X_{t-p+1}, X_{t-p+2}, \ldots, X_t]$ similar to the training data and send the decomposed signal to the corresponding trained LDMR model. The prediction

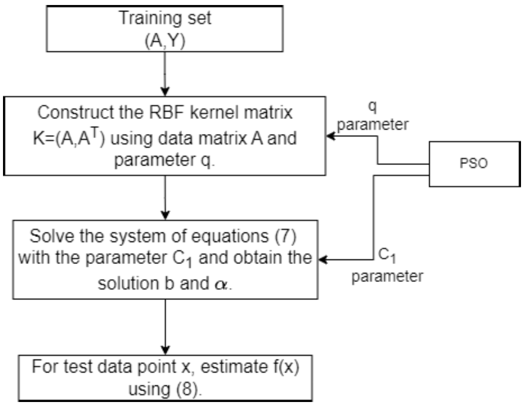


FIGURE 5. LS-SVR model flowchart.

of the all LDMR model is aggregated to obtain the final prediction of x_{t+1}.

c) **EMD-PSO-LDMR model**:- The working of the EMD-PSO-LDMR model is similar to the WD-PSO-LDMR model except that the EMD-PSO-LDMR model uses the EMD decomposition technique to decompose the input signals. In EMD-PSO-LDMR model, input time-series SWH signal $s(t) = (X_1, X_2, \ldots, X_t)$ is decomposed in five Intrinsic Mode Functions (IMFs) and one residuals function.

d) **VMD-PSO-LDMR model**:- The working of the VMD-PSO-LDMR model is similar to the WD-PSO-LDMR model except that the VMD-PSO-LDMR model uses the VMD decomposition technique to decompose the input signals. In VMD-PSO-LDMR model, input time-series SWH signal $s(t) = (X_1, X_2, \ldots, X_t)$ is decomposed in five IMFs and one residuals function using the VMD decomposition.

2) **SVR based hybrid models:-** The PSO-SVR model is a hybrid model that uses SVR for regression and PSO for parameter optimization. Figure 4 shows the steps for training the PSO-SVR model on the given SWH dataset. The PSO algorithm tunes the kernel parameter q and the regularization parameter C for the SVR model. The PSO-SVR model constructs the training set (A,Y) from the time-series SWH data $s(t) = (X_1, X_2, \ldots, X_t)$ as in the PSO-LDMR model. The PSO-SVR model then trains on the training set (A,Y) following the steps in Figure 2. To predict the SWH x_{t+1}, the PSO-SVR model takes the test point $[X_{t-p+1}, X_{t-p+2}, \ldots, X_t]$ as input. The other hybrid models that use SVR for regression, such as WD-PSO-SVR, VMD-PSO-SVR, and EMD-PSO-SVR, work similarly to the LDMR based hybrid models, except that they replace LDMR with SVR.

3) **LS-SVR based hybrid models:-** The PSO-LS-SVR model is a hybrid model that uses LS-SVR for regression and PSO for parameter optimization. Figure 5 shows the steps for training the PSO-LS-SVR model on the given SWH dataset. The PSO algorithm tunes the kernel parameter q and the regularization parameter C_1 for the LS-SVR model. The PSO-LS-SVR model constructs the training set (A,Y) from the time-series SWH data $s(t) = (X_1, X_2, \ldots, X_t)$ as in the PSO-LDMR model. The PSO-LS-SVR model then trains on the training set (A,Y) following the steps in Figure 5. To predict the SWH x_{t+1}, the PSO-SVR model takes the test point $[X_{t-p+1}, X_{t-p+2}, \ldots, X_t]$ as input. The other hybrid models that use LS-SVR for regression, such as WD-PSO-LS-SVR, VMD-PSO-LS-SVR, and EMD-PSO-LS-SVR, work similarly to the LDMR based hybrid models, except that they replace LDMR with LS-SVR model.

A. IMPLEMENTATION DETAILS

In Wavelet decomposition based hybrid models, we have decomposed the SWH data signal into five high frequency detail signals $(D_1, D_2, D_3, D_4, D_5)$ and a low frequency approximate signal A_5. For this, we have used the Daubechies 4 discrete wavelet filter available in MATLAB (www.mathworks.com). In Figure 8 (a), we have shown the wavelet decomposition for the dataset **A**. For our EMD and VMD decomposition based wave hybrid models, we have decomposed the initial data signal into six IMFs using their respective algorithms. In Figure 8 (b), we have shown the VMD decomposition for the dataset **A**. After decomposing the original signal into six different components, we have composed the feature sets and response values in each case and divided the obtained datasets into training set and testing set. The 80% of a dataset was used as training set and remaining was used as testing set.

We have used the SVR, LS-SVR, LDMR and LSTM models for the prediction of SWH data. We have implemented the SVR, LS-SVR and LDMR models in MALTAB environment by writing an appropriate function in MATLAB. The SVR and LDMR model require the solution of the dual QPPs (3) and (11) respectively. The solution of these QPPs of SVR and LDMR models have been obtained using the 'quadprog' function available in MATLAB by using the 'interior-point convex' algorithm. The LS-SVR model only requires the solution of a system of equations to obtain its solution. In SVR, LDMR and LS-SVR models, we have used the RBF kernel of the form $k(x, y) = e^{-q||x-y||}$, where q is the kernel parameter. We have implemented the LSTM model in python using Jupyter notebook.

We need to tune and obtain the appropriate parameters of our machine learning model for a given dataset before using them for obtaining the prediction. We have used the PSO algorithm to tune the parameters of our used machine learning models. We have illustrated the flow chart of used PSO algorithm in our hybrid model in Figure 6. The LDMR model involves four parameters namely ϵ, C, c and kernel

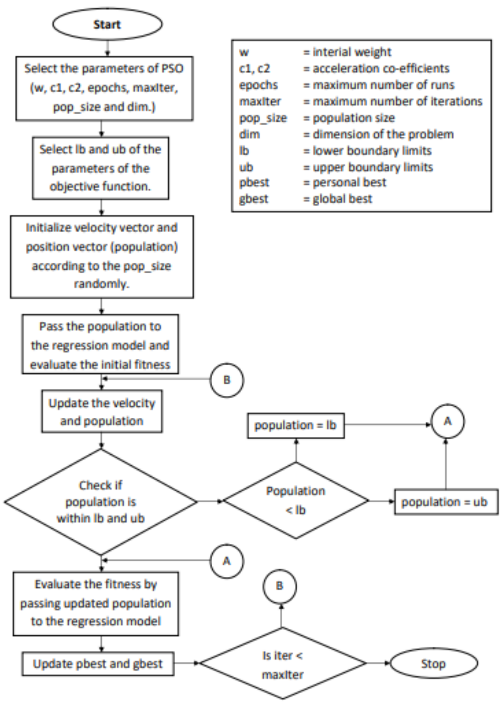


FIGURE 6. **Flow chart of implemented PSO algorithm in our wave hybrid models.**

parameter q. The ϵ-SVR involves three parameters namely ϵ, C and kernel parameter q and LS-SVR models involve two parameters C and kernel parameter q only. The ranges of parameters ϵ, C, c and q ([lb,ub] in Figure 7) have been fixed with $[0, 2]$, $[0, 1024]$, $[0, 1024]$, $[0, 1024]$ respectively for all considered SVR models. For the evaluation of fitness function in our PSO algorithm, we need to obtain the testing error by solving the dual problem (3) and (11) for ϵ-SVR and LDMR model respectively. For the LS-SVR model, we need to obtain the solution of system of equations (7). For LSTM model, we have tunned number of hidden units and batch sizes using grid search method in [1, 128].

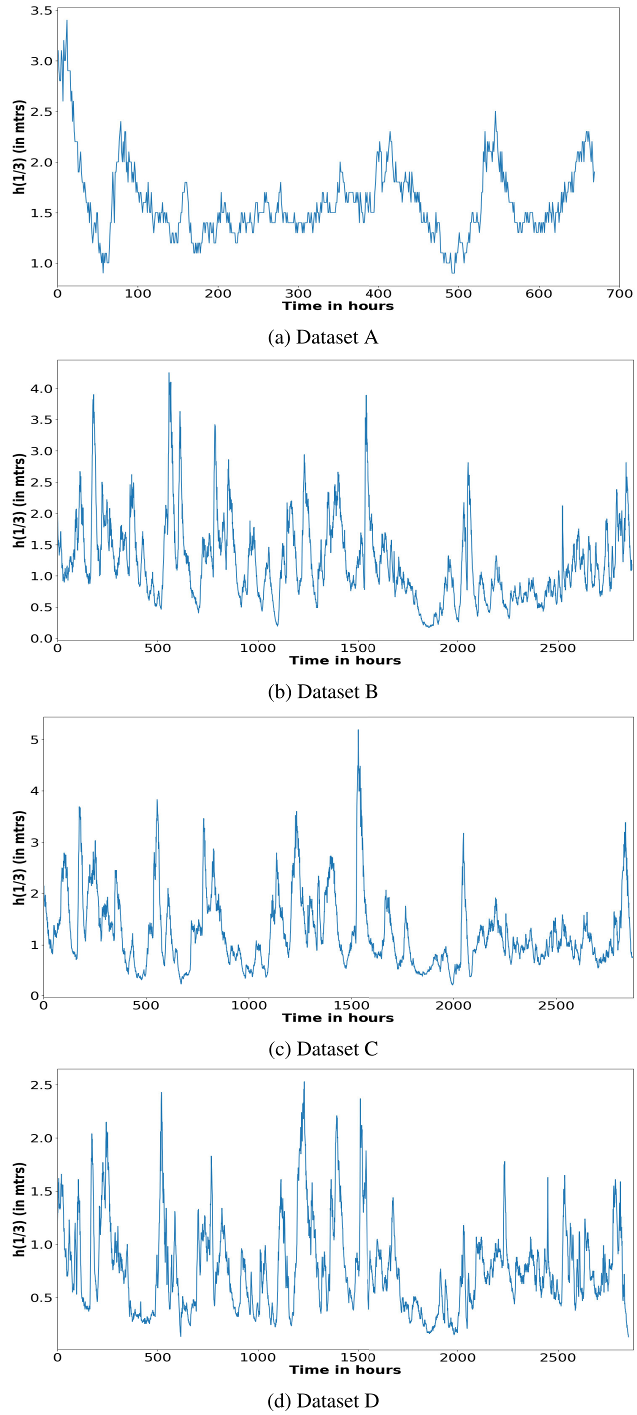
IV. EXPERIMENTAL RESULTS

In this section, we shall be comparing the performances of developed wave hybrid models along with different existing wave hybrid models on different ocean significant wave height datasets.

A. DATASETS DESCRIPTION

We have collected the time-series hourly SWH data from the different ocean buoys available on the National Data Buoy Center (NDBC) (https://www.ndbc.noaa.gov/). We have listed a brief details of chosen ocean buoys datasets at Table 1.

At first, we have considered the forecasting of one hour ahead SWH. For this, we have collected four datasets namely **A**, **B**, **C** and **D**, which contain the hourly reading of SWH from ocean buoys detailed in Table 1. The datasets **A**, **B**, **C** and **D** record the SWH for 670, 2869, 2872, and 2852 continuous hours respectively.



(a) Dataset A

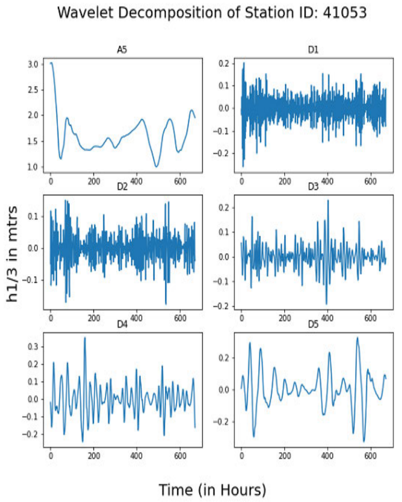
(b) Dataset B

(c) Dataset C

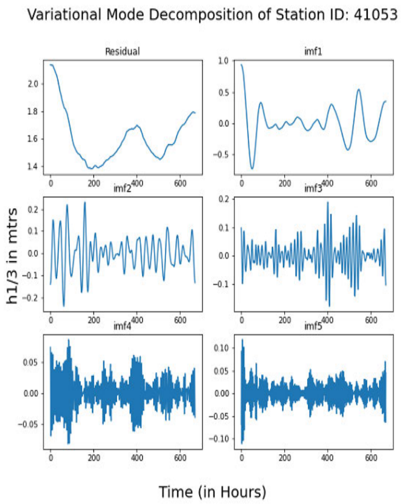
(d) Dataset D

FIGURE 7. **Plot of different time series SWH datasets.**

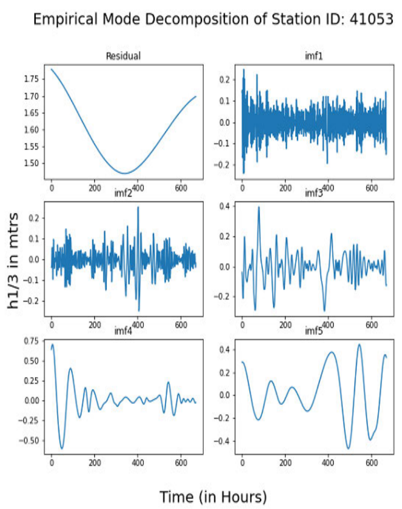
We have plotted the time series signals of SWH dataset **A**, **B**, **C** and **D** in Figure 7. The SWH ($h_{\frac{1}{3}}$) is mean of top one third of wave height. We can observe that our datsets contain highly non-linear and chaotic signals. We have shown the WD, VMD and EMD decomposition of the time-series signal for dataset **A** in the Figure 8.



(a)



(b)



(c)

FIGURE 8. Decomposition of signal for dataset A.

We have also compared the existing and proposed hybrid models for six hour ahead forecasting of SWH. For this, we have considered dataset **E**, **F**, and **G**, which records the SWH at every six hour time interval. We have also collected this datasets from buoys detailed in Table 1. The datasets **E**, **F**, and **G** contains 476, 478 and 479 readings of SWH.

TABLE 1. Description of datasets.

Station ID	Co-ordinates	Time
410532	66°5'58" W, 18°28'27" N	Feb 2021
42001	89°39'25" W, 25°56'31" N	Jan 2015
42002	93°38'46" W, 26°3'18" N	Jan 2015
42035	94°24'45" W, 29°13'54" N	Jan 2015

B. EVALUATION CRITERIA

For evaluating the different considered wave hybrid models, we need to fix the evaluation criterion first. We have used the commonly used evaluation criterion for measuring the performance of wave hybrid models. With the notations, y_i as actual SWH for i^{th} test sample, $\hat{y}_i$ as predicted SWH for i^{th} test sample, $\bar{y}_i$ as mean of actual SWH and n as total number of testing samples, we briefly describe our evaluation criteria as follows.

1) Root Mean Square of Errors (RMSE): It is obtained by $\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$.
2) Mean of Absolute Deviations (MAD): It is obtained by $\frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$.
3) Mean Absolute Percentage of Errors (MAPE):- It is obtained by $\frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i-\hat{y}_i}{y_i}\right| \times 100$.
4) Normalized Mean Squares of Errors (NMSE):- It is ratio of Sum of Squares of Errors (SSE) and Sum of Squares of Testing samples (SST) and is obtained by $\frac{\sum_{i=1}^{n}(\hat{y}_i-y_i)^2}{\sum_{i=1}^{n}(y_i-\bar{y})^2}$.
5) Accuracy (Acc): It is the measure for the accuracy of obtained predictions. It is obtained by Acc = 100 - MAPE.

C. RESULTS

After fixing the evaluation criteria, we present our numerical results obtained from the extensive experiments. We have listed the performance of proposed LDMR, SVR, LS-SVR and LSTM models with no decomposition, WD,EMD and VMD techniques on our selected SWH datasets with lead time of 1 hour and 6 hours using Accuracy, RMSE, NMSE, MAD and MAPE at Table 2 and Table 3 respectively. We have also listed the total computational time in seconds for each of our wave hybrid models in Table 2 and Table 3.

D. ANALYSIS AND INFERENCE

We shall briefly analyze the numerical results presented in Table 2 and Table 3. Out of seven considered dataset

TABLE 2. Numerical results of lead time 1 hrs.

Dataset	Method	Accuracy (%)	RMSE (mtrs)	NMSE	MAD (mtrs)	MAPE (%)	Time (sec)
A	**PSO-LDMR**	**94.89**	**0.1037**	**0.1165**	**0.0866**	**5.11**	413.38
	PSO-SVR	94.85	0.112	0.1359	0.0888	5.15	441.01
	PSO-LS-SVR	94.82	0.1039	0.1170	0.0876	5.18	**18.02**
	LSTM	92.15	0.1826	0.3613	0.1425	7.84	87.01
	WD-PSO-LDMR	**98.09**	**0.0381**	**0.0157**	**0.0307**	**1.91**	**4339.38**
	WD-PSO-SVR	95.22	0.6061	3.9803	0.5135	4.78	1431.84
	WD-PSO-LS-SVR	97.57	0.049	0.0260	0.0391	2.43	**45.87**
	WD-LSTM	77.37	0.4398	2.096	0.3714	22.62	187.66
	EMD-PSO-LDMR	89.44	0.1896	0.3895	0.1669	10.56	**3890.66**
	EMD-PSO-SVR	90.19	0.177	0.3395	0.1559	9.81	
	EMD-PSO-LS-SVR	90.31	0.1768	0.3387	0.1556	9.69	**121.84**
	EMD-LSTM	**91.91**	**0.1562**	**0.2645**	**0.137**	**8.08**	181.27
	VMD-PSO-LDMR	**98.32**	**0.0344**	**0.0128**	**0.0283**	**1.52**	3574.10
	VMD-PSO-SVR	98.28	0.0349	0.0132	0.0285	1.72	2418.09
	VMD-PSO-LS-SVR	97.92	0.042	0.0191	0.0348	2.08	**181.61**
	VMD-LSTM	85.69	0.2711	0.7968	0.2492	14.30	183.66
B	**PSO-LDMR**	82.51	0.2965	0.4835	0.2300	17.49	153.94
	PSO-SVR	82.59	0.3195	0.5614	0.2379	17.41	119.33
	PSO-LS-SVR	80.52	0.3305	0.6007	0.2554	19.48	**2.17**
	LSTM	**90.65**	**0.164**	**0.1404**	**0.1151**	**9.34**	69.88
	WD-PSO-LDMR	92.83	0.1226	0.0827	0.0908	7.17	**860.00**
	WD-PSO-SVR	92.78	0.1168	0.0750	0.0911	7.22	570.29
	WD-PSO-LS-SVR	87.99	0.1873	0.1929	0.1443	12.01	39.50
	WD-LSTM	**93.31**	**0.098**	**0.0501**	**0.0764**	**6.68**	183.78
	EMD-PSO-LDMR	80.28	0.3189	0.5593	0.2614	19.72	**823.52**
	EMD-PSO-SVR	78.59	0.3916	0.8434	0.3048	21.41	481.36
	EMD-PSO-LS-SVR	74.90	0.4153	0.9486	0.3407	25.10	28.42
	EMD-LSTM	**81.88**	**0.2797**	**0.4087**	**0.2125**	**18.11**	200.82
	VMD-PSO-LDMR	**95.77**	**0.0661**	0.0240	**0.0538**	**4.23**	1580.70
	VMD-PSO-SVR	86.49	0.2238	0.2755	0.1702	13.51	864.21
	VMD-PSO-LS-SVR	95.43	0.0616	**0.0209**	0.0584	4.57	**42.58**
	VMD-LSTM	94.11	0.1087	0.0617	0.074	5.88	182.24
C	**PSO-LDMR**	79.05	0.2812	0.2483	0.2233	20.95	163.73
	PSO-SVR	29.26	1.0469	3.4418	0.8815	70.74	114.37
	PSO-LS-SVR	70.02	0.3055	0.2931	0.2323	21.45	**8.87**
	LSTM	**90.37**	**0.1503**	**0.092**	**0.1092**	**9.62**	69.78
	WD-PSO-LDMR	92.76	0.1137	0.0406	0.0810	7.24	966.99
	WD-PSO-SVR	92.25	**0.1072**	**0.0361**	0.0817	7.75	532.57
	WD-PSO-LS-SVR	89.25	0.1344	0.0567	0.1074	10.75	**39.83**
	WD-LSTM	**93.34**	0.1092	0.0485	**0.0773**	**6.65**	183.59
	EMD-PSO-LDMR	78.43	0.3123	0.3063	0.2425	21.57	601.53
	EMD-PSO-SVR	79.72	0.2991	0.2809	0.2273	20.28	384.61
	EMD-PSO-LS-SVR	77.55	0.3202	0.3220	0.2531	22.45	**34.47**
	EMD-LSTM	**81.91**	**0.2553**	**0.2655**	**0.201**	**18.05**	183.49
	VMD-PSO-LDMR	**95.03**	0.0672	0.0142	**0.0500**	**4.97**	1091.91
	VMD-PSO-SVR	94.60	**0.0667**	**0.0140**	0.0527	5.40	797.86
	VMD-PSO-LS-SVR	94.46	0.0711	0.0159	0.0536	5.54	**40.88**
	VMD-LSTM	94.9	0.0858	0.03	0.0595	5.09	185.55
D	**PSO-LDMR**	74.84	0.241	0.5785	**0.0876**	**5.18**	18.02
	PSO-SVR	73.06	0.2475	0.6101	0.1919	26.94	112.74
	PSO-LS-SVR	70.75	0.2502	0.6235	0.1941	29.25	**8.66**
	LSTM	**86.3**	**0.1433**	**0.218**	0.1021	13.69	85.34
	WD-PSO-LDMR	**90.72**	0.1000	0.0996	0.0671	**9.28**	1286.98
	WD-PSO-SVR	90.57	**0.0955**	**0.0908**	**0.0642**	9.43	553.00
	WD-PSO-LS-SVR	81.01	0.1859	0.3442	0.1418	18.99	**44.42**
	WD-LSTM	89.55	0.0971	0.1001	0.0737	10.44	183.56
	EMD-PSO-LDMR	**80.69**	**0.1983**	**0.3917**	0.1531	**19.31**	817.20
	EMD-PSO-SVR	78.88	0.2068	0.4260	**0.1528**	21.12	397.04
	EMD-PSO-LS-SVR	78.55	0.2542	0.6436	0.1776	21.45	**34.39**
	EMD-LSTM	71.45	0.2116	0.4754	0.1912	28.54	185.99
	VMD-PSO-LDMR	93.23	**0.0592**	**0.0349**	**0.0443**	6.77	1128.97
	VMD-PSO-SVR	92.45	0.0613	0.0374	0.0486	7.55	889.45
	VMD-PSO-LS-SVR	90.38	0.083	0.0686	0.0625	9.62	**43.54**
	VMD-LSTM	**93.71**	0.0719	0.0549	0.0493	**6.28**	184.64

VMD-PSO-LDMR model obtain the best performance on six datasets among other considered wave hybrid models.For clear visualization of obtained numerical result, we have compared the accuracy obtained by different wave hybrid models using a box plot in Figure 9 for all considered seven datasets.

In box plot of Figure 9, we can compare the median of accuracy obtained by different wave hybrid models.

TABLE 3. Numerical results of lead time 6 hrs.

Dataset	Method	Accuracy (%)	RMSE (mtrs)	NMSE	MAD (mtrs)	MAPE (%)	Time (sec)
E	PSO-LDMR	**74.84**	**0.241**	**0.5785**	**0.0876**	**5.18**	18.02
	PSO-SVR	73.06	0.2475	0.6101	0.1919	26.94	112.74
	PSO-LS-SVR	70.75	0.2502	0.6235	0.1941	29.25	**8.66**
	LSTM	59.33	0.3258	1.0436	0.2489	40.66	85.34
	WD-PSO-LDMR	**90.72**	0.1	0.0996	0.0671	**9.28**	1286.98
	WD-PSO-SVR	90.57	**0.0955**	**0.0908**	**0.0642**	9.43	553.00
	WD-PSO-LS-SVR	81.01	0.1859	0.3442	0.1418	18.99	**44.42**
	WD-LSTM	56.77	0.3164	0.9846	0.2628	43.22	183.56
	EMD-PSO-LDMR	**80.69**	**0.1983**	**0.3917**	0.1531	**19.31**	817.20
	EMD-PSO-SVR	78.88	0.2068	0.4260	**0.1528**	21.12	397.04
	EMDPSO-LS-SVR	78.55	0.2542	0.6436	0.1776	21.45	**34.39**
	EMD-LSTM	75.5	0.3055	0.9175	0.219	24.49	185.99
	VMD-PSO-LDMR	**93.23**	**0.0592**	**0.0349**	**0.0443**	**6.77**	1128.97
	VMD-PSO-SVR	92.45	0.0613	0.0374	0.0486	7.55	889.45
	VMD-PSO-LS-SVR	90.38	0.083	0.0686	0.0625	9.62	**43.54**
	VMD-LSTM	76.41	0.1984	0.387	0.1508	23.58	184.64
F	PSO-LDMR	82.51	**0.2965**	**0.4835**	**0.23**	17.49	153.94
	PSO-SVR	**82.59**	0.3195	0.5614	0.2379	**17.41**	119.33
	PSO-LS-SVR	80.52	0.3305	0.6007	0.2554	19.48	**2.17**
	LSTM	77.4	0.3973	0.9175	0.3058	22.59	69.88
	WD-PSO-LDMR	**92.83**	0.1226	0.0827	**0.0908**	**7.17**	860
	WD-PSO-SVR	**92.83**	**0.1168**	**0.075**	0.0911	7.22	570.29
	WD-PSO-LS-SVR	**92.83**	0.1873	0.1929	0.1443	12.01	**39.5**
	WD-LSTM	67.66	0.4473	1.1629	0.3875	32.33	183.78
	EMD-PSO-LDMR	81.58	0.2756	0.4177	0.2206	18.42	366.81
	EMD-PSO-SVR	**84.25**	**0.2521**	**0.3495**	**0.1988**	**15.75**	281.07
	EMD-PSO-LS.SVR	81.17	0.3379	0.6279	0.2532	18.83	**16.79**
	EMD-LSTM	80.27	0.3019	0.5297	0.2418	19.72	200.82
	VMD-PSO-LDMR	**95.77**	0.0661	0.024	**0.0538**	**4.23**	1580.7
	VMD-PSO-SVR	86.49	0.2238	0.2755	0.1702	13.51	864.21
	VMD-PSO-LS-SVR	95.43	**0.0616**	**0.0209**	0.0584	4.57	**42.58**
	VMD-LSTM	76.66	0.4333	1.0912	0.3319	23.33	182.24
G	PSO-LDMR	**79.05**	**0.2812**	**0.2483**	**0.2233**	**20.95**	163.73
	PSO-SVR	29.26	1.0469	3.4418	0.8815	70.74	114.37
	PSO-LS-SVR	70.02	0.3055	0.2931	0.2323	21.45	**8.87**
	LSTM	73.13	0.5014	0.8201	0.3296	26.86	69.78
	WD-PSO-LDMR	**92.76**	0.1137	0.0406	**0.081**	**7.24**	966.99
	WD-PSO-SVR	92.25	**0.1072**	**0.0361**	0.0817	7.75	532.57
	WD-PSO-LS-SVR	89.25	0.1344	0.0567	0.1074	10.75	**39.83**
	WD-LSTM	48.22	0.6131	1.2263	0.5414	51.77	183.59
	EMD-PSO-LDMR	81.56	0.2723	0.2328	0.1856	18.44	389.97
	EMD-PSO-SVR	**82.94**	**0.2621**	**0.2157**	**0.1799**	**17.06**	205.79
	EMD-PSO-LS-SVR	80.57	0.2673	0.2244	0.1948	19.43	**17.32**
	EMD-LSTM	76.84	0.3207	0.3355	0.2431	23.15	183.49
	VMD-PSO-LDMR	**95.77**	0.0661	0.024	**0.0538**	**4.23**	1580.7
	VMD-PSO-SVR	86.49	0.2238	0.2755	0.1702	13.51	864.21
	VMD-PSO-LS-SVR	95.43	**0.0616**	**0.0209**	0.0584	4.57	**42.58**
	VMD-LSTM	80.73	0.3216	0.3374	0.227	19.26	185.55

The VMD-PSO-LDMR model obtains highest 95.77 median accuracy which is followed by the VMD-PSO-LS-SVR model with 95.43 median accuracy. The VMD-PSO-LDMR model also excels with other wave hybrid models if we consider the 25^{th} and 75^{th} percentile of accuracy values. Also, the maximum accuracy obtained by the VMD-PSO-LDMR model is 98.32, which is the highest among all obtained accuracy values by any wave hybrid model.

From the box plot, we can observe that the VMD-PSO-LDMR, VMD-PSO-LS-SVR, WD-PSO-SVR, and WD-PSO-LDMR models are four top-performing wave hybrid models on our dataset. We considered four top-performing hybrid models on SWH datasets and computed listed their obtained ranks in Table 5. The VMD-PSO-LDMR model obtains the first rank on all seven SWH datasets. The average ranks obtained by WD-PSO-LDMR, VMD-PSO-LS-SVR, and WD-PSO-SVR models are 2.57, 2.71, and 3.71, respectively.

TABLE 4. Wilcoxon test.

	EMD (R^+, R^-, p-value) Decision	WD (R^+, R^-, p-value) Decision	No decom. (R^+, R^-, p-value) Decision
VMD	(396, 10, 0.00001) REJECT	(347, 59, 0.00104) REJECT	(395, 10, 0.00001) REJECT
EMD		(101, 305, 0.0203) REJECT	(263, 143, 0.1706) ACCEPT
WD			(350, 56, 0.0008) REJECT

We have plotted the prediction obtained by VMD-PSO-LDMR model for dataset A and D in Figure 12. We have shown the prediction of VMD-PSO-SVR model in Figure 13.

TABLE 5. Rank obtained by four top-performing wave hybrid model namely VMD-LS-SVR, WD-SVR, VMD-LDMR and WD-LDMR on considered SWH datasets.

Dataset	VMD-PSO-LS-SVR	WD-PSO-SVR	VMD-PSO-LDMR	WD-PSO-LDMR
A	3	4	1	2
B	2	4	1	3
C	2	4	1	3
D	4	3	1	2
E	4	3	1	2
F	2	4	1	3
G	2	4	1	3
Average Rank	2.71	3.71	1	2.57

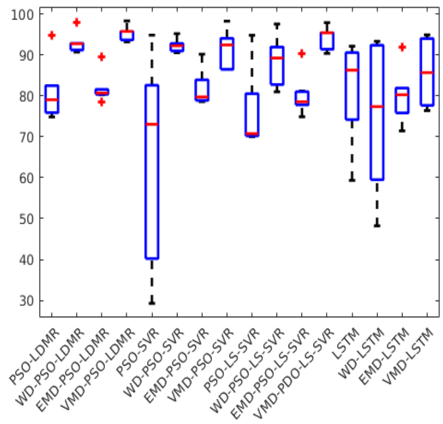


FIGURE 9. Comparison of different wave hybrid models using box plot.

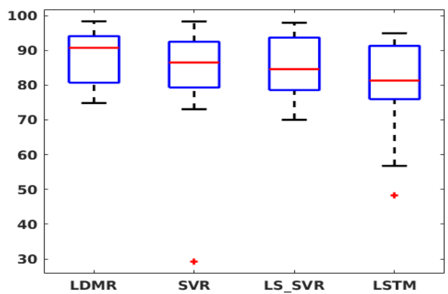


FIGURE 10. Box plot for comparison of different regression models.

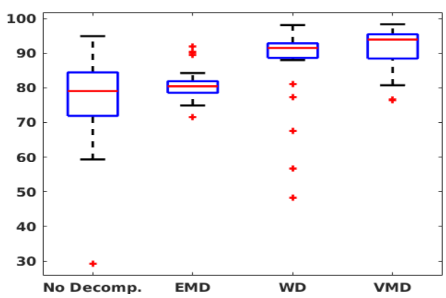
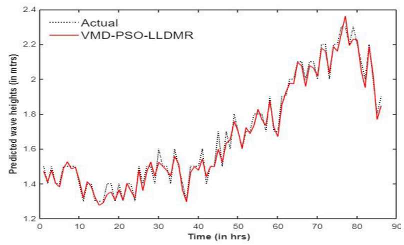
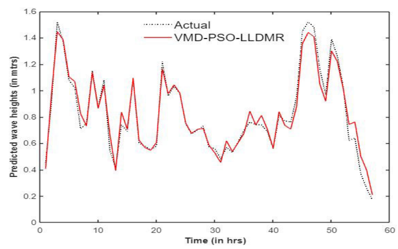


FIGURE 11. Box plot for comparison of different decomposition methods.



(a) Dataset A



(b) Dataset D

FIGURE 12. Performance of VMD-PSO-LDMR model.

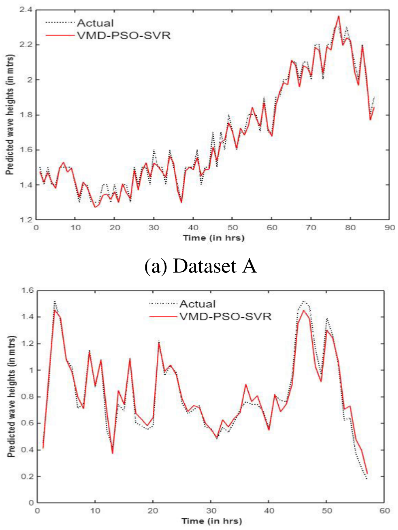
Also, we have shown the prediction of VMD-PSO-LS-SVR and VMD-PSO-LSTM model for dataset D in Figure 14.

Now, we compare the effectiveness of regression models used in considered wave hybrid models. We have used four variants of LDMR, SVR, LS-SVR and LSTM based hybrid models on seven different datasets. Figure 10 compares LDMR, SVR, LS-SVR and LSTM based hybrid models using box plot. The LDMR based hybrid models obtain highest 90.72 median accuracy which is followed by the SVR based hybrid models with 86.49 median accuracy. The LS-SVR based hybrid models and LSTM based hybrid models obtain 84.58 and 81.30 median accuracy respectively. The 75^{th} percentile accuracy values obtained by LDMR based hybrid models is 93.64 where as 75^{th} percentile accuracy values obtained by SVR, LS-SVR and LSTM based hybrid models are 92.45, 93.23 and 90.96 respectively. The 25^{th} percentile accuracy values obtained by LDMR based hybrid models is 80.69 where as 25^{th} percentile accuracy values obtained by SVR, LS-SVR and LSTM based hybrid models are 79.51, 78.55 and 76.18 respectively.

We have shown the average accuracy values obtained by different wave hybrid models in Table 6. We can observe that PSO-LDMR based wave hybrid models obtain 87.79 mean accuracy which is followed by PSO-SVR based hybrid models with 82.92 mean accuracy.

We can easily realize that LDMR based hybrid models obtain best performance on considered seven datasets for SWH prediction. We can also find that SVR based hybrid models outperform LS-SVR and LSTM based hybrid models.

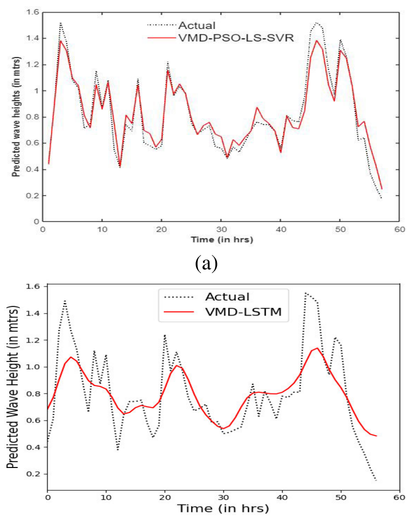
Let us compare the performances of LDMR, SVR, LS-SVR and LSTM based wave hybrid models in case of no



(a) Dataset A

(b) Dataset D

FIGURE 13. Performance of VMD-PSO-SVR model.



(a)

(b)

FIGURE 14. Performance of (a) VMD-PSO-LS-SVR and (b) VMD-PSO-LSTM model on dataset D.

TABLE 6. Average accuracy values obtain by wave hybrid models.

	No decomp	WD	EMD	VMD	Average
PSO-LDMR	81.10	92.96	81.81	95.30	87.79
PSO-SVR	66.38	92.35	81.92	91.04	82.92
PSO-LS-SVR	76.77	88.42	80.23	94.20	84.91
LSTM	81.33	75.17	79.97	86.03	80.63
Average	76.40	87.23	80.98	91.64	84.06

decomposition of SWH signals. In Table 6, we have listed the average of accuracy values obtained by LDMR, SVR,LS-SVR and LSTM models in case of different decomposition techniques. From Table 6 and Figure 9, we can note that, in the case of no decomposition, the LSTM model obtained 86.30 median accuracy and 81.33 mean accuracy, which is highest among all no decomposition based wave hybrid models. In case of no decomposition of signals, the median accuracy obtained by PSO-LDMR, PSO-SVR and PSO-LS-SVR models are 79.05, 73.06 and 70.75, respectively and mean accuracy obtained by PSO-LDMR, PSO-SVR and PSO-LS-SVR models are 81.10, 66.38 and 76.77 respectively. We can observe that the LSTM model outperforms the SVR, LS-SVR and LDMR based hybrid models if we don't decompose the SWH signals apriori.

From Table 6, we also observe that the PSO SVR, PSO LS-SVR and PSO LDMR model obtain poor performances. But, when these models are used with decomposition techniques, their performances are significantly improved. It is because of the fact that the decomposition techniques provide informative and suitable features to SVR models. Contrary to this, LSTM models fail to improve their prediction ability significantly when signals are decomposed apriori.

Further, we compare the performance of the different decomposition methods used in considered wave hybrid models. We have used four decomposition techniques with all considered wave hybrid models on seven different datasets. Figure 11 compares No decomposition, EMD, WD and VMD methods using a box plot. The VMD method obtained the highest 93.91 median accuracy, which is followed by the WD method with 91.48 median accuracy. The no decomposition and EMD method obtain 79.05 and 80.42 median accuracy, respectively. The 75^{th} percentile accuracy values obtained by VMD method is 95.43, whereas 75^{th} percentile accuracy values obtained by no decomposition, WD and EMD methods are 83.51, 92.83 and 81.88 respectively. The 25^{th} percentile accuracy values obtained by VMD methods is 89.40, whereas 25^{th} percentile accuracy values obtained by no decomposition, WD and EMD methods are 72.48, 88.93 and 78.55 respectively. From Table 6, we can also observe that mean accuracy obtained by VMD, WD, EMD and No decomposition methods based hybrid models are 91.64, 87.23, 80.98 and 76.40.

From VMD based wave hybrid models obtain the best performance among other used wave hybrid models. The WD methods based hybrid models also perform better than EMD and no decomposition methods. Researchers have shown that the EMD method may be sensitive to noise and sampling [36]. Our datasets contain the reading of sensors deployed at ocean buoys, which may be subject to noise. It causes the EMD based wave hybrid models to perform poorer than VMD and WD based wave hybrid models.

We now use the non-parametric Wilcoxon single rank test [39] for comparing the different decomposition based wave hybrid models using accuracy. Our null hypothesis is that all considered decomposition-based hybrid models perform equally. We have reported the R^+, R^-, and p values computed in the Wilcoxon single rank test for every pair of decomposition methods in Table 4. Here R^+ and R^- are the sums of positive and negative ranks respectively obtained for a pair of all the decomposition methods of wave hybrid models. We have considered the level of significance α = 0.05 in this study. If the p value is lesser than α, then we reject our null hypothesis; Otherwise, we need to accept it. We have also listed the decision of the Wilcoxon single rank test for every pair of decomposition methods in Table 4. We can observe from the Table 4 that the numerical results obtained from VMD based hybrid models are significantly different from the numerical results obtained by the WD, EMD and No-decomposition based wave hybrid models. Furhter, WD based wave hybrid models produce numerical results that are significantly different from the results obtained by EMD and no-decomposition based hybrid methods.

V. CONCLUSION

In this paper, we have designed sixteen different wave hybrid models namely PSO-LDMR, PSO-SVR, PSO- LS-SVR, LSTM, WD-PSO-LDMR, WD-PSO-SVR, WD-PSO-LS-SVR, WD-LSTM, EMD-PSO-LDMR, EMD-PSO-SVR, EMD-PSO-LS-SVR, EMD-LSTM, VMD-PSO-LDMR, VMD-PSO-SVR, VMD-PSO-LS-SVR and VMD-LSTM model for hourly forecasting of significant wave height.

We have used the Wavelet, EMD and VMD methods in our hybrid models for decomposing the time-series signal into informative features. We have first used the VMD decomposition method in wave hybrid models for significant wave height forecasting and shown that it improves performance significantly.

We have used the LDMR models ϵ-SVR, LS-SVR, LSTM models for forecasting of SWH. No wave hybrid model has ever exploited the LDMR model for SWH forecasting. The LDMR model offers more flexibility and the ability to handle the mixture of noise well. We have shown that the use of the LDMR model in wave hybrid models improves SWH forecasting. We have also used the PSO method to tune the parameters of the used regression model in our developed wave hybrid models

We have checked the performances of wave hybrid models using different evaluation criteria on seven real-world ocean wave height datasets collected from four different ocean buoys.

After the brief analysis of the numerical results, we find that the VMD-PSO-LDMR model obtains the best performance on six datasets out of seven datasets with a median accuracy of 95.77 and mean accuracy of 95.30. We find that the LDMR based wave hybrid models perform better than SVR, LS-SVR and LSTM based wave hybrid models. We have also compared decomposition methods used in our wave hybrid models. We find that VMD based wave hybrid models outperform EMD, WD and No decomposition based wave hybrid models significantly. Further, we observe that in case of no decomposition of signals, the LSTM based wave hybrid models perform better than SVR, LS-SVR, and LDMR based wave hybrid models. But, when signal decomposition techniques are used in SVR, LS-SVR, and LDMR based wave hybrid models, then they improve their performances significantly. Contrary to them, the LSTM based hybrid models do not improve their prediction significantly when signal decomposition techniques are used.

Our developed wave hybrid models are only capable of single time stamp ahead forecasting. Further, it does not involve any mechanism to exploit the information available like wind speed, wind directions, temperature and sea level and for improvement in SWH forecast. We have planned to improve our hybrid models in this sense in the future.

REFERENCES

[1] M. Lehmann, F. Karimpour, C. A. Goudey, P. T. Jacobson, and M.-R. Alam, ''Ocean wave energy in the United States: Current status and future perspectives,'' *Renew. Sustain. Energy Rev.*, vol. 74, pp. 1300–1313, Jul. 2017.
[2] I. Fairley, M. Lewis, B. Robertson, M. Hemer, I. Masters, J. Horrillo-Caraballo, H. Karunarathna, and D. E. Reeve, ''A classification system for global wave energy resources based on multivariate clustering,'' *Appl. Energy*, vol. 262, Mar. 2020, Art. no. 114515.
[3] S. Yang, Z. Deng, X. Li, C. Zheng, L. Xi, J. Zhuang, Z. Zhang, and Z. Zhang, ''A novel hybrid model based on STL decomposition and one-dimensional convolutional neural networks with positional encoding for significant wave height forecast,'' *Renew. Energy*, vol. 173, pp. 531–543, Aug. 2021.
[4] W. Y. Duan, Y. Han, L. M. Huang, B. B. Zhao, and M. H. Wang, ''A hybrid EMD-SVR model for the short-term prediction of significant wave height,'' *Ocean Eng.*, vol. 124, pp. 54–73, Sep. 2016.
[5] F. Fusco, ''Short-term wave forecasting as a univariate time series problem,'' Dept. Electron. Eng., NUI Maynooth, Kildare, Ireland, Tech. Rep. EE/2009/JVR/3, 2009.
[6] C. G. Soares and A. M. Ferreira, ''Representation of non-stationary time series of significant wave height with autoregressive models,'' *Probabilistic Eng. Mech.*, vol. 11, no. 3, pp. 139–148, Jul. 1996.
[7] M. Ge and E. C. Kerrigan, ''Short-term ocean wave forecasting using an autoregressive moving average model,'' in *Proc. UKACC 11th Int. Conf. Control (CONTROL)*, Aug. 2016, pp. 1–6.
[8] E. Vanem and S.-E. Walker, ''Identifying trends in the ocean wave climate by time series analyses of significant wave height data,'' *Ocean Eng.*, vol. 61, pp. 148–160, Mar. 2013.
[9] M. C. Deo, A. Jha, A. S. Chaphekar, and K. Ravikant, ''Neural networks for wave forecasting,'' *Ocean Eng.*, vol. 28, no. 7, pp. 889–898, Jul. 2001.
[10] O. Makarynskyy, ''Improving wave predictions with artificial neural networks,'' *Ocean Eng.*, vol. 31, nos. 5–6, pp. 709–724, Apr. 2004.
[11] L. Cornejo-Bueno, J. C. Nieto-Borge, P. García-Díaz, G. Rodríguez, and S. Salcedo-Sanz, ''Significant wave height and energy flux prediction for marine energy applications: A grouping genetic algorithm—Extreme learning machine approach,'' *Renew. Energy*, vol. 97, pp. 380–389, Nov. 2016.
[12] N. K. Kumar, R. Savitha, and A. Al Mamun, ''Ocean wave height prediction using ensemble of extreme learning machine,'' *Neurocomputing*, vol. 277, pp. 12–20, Feb. 2018.
[13] C. E. Balas, L. Koç, and L. Balas, ''Predictions of missing wave data by recurrent neuronets,'' *J. Waterway, Port, Coastal, Ocean Eng.*, vol. 130, no. 5, pp. 256–265, 2004.
[14] S. Mandal and N. Prabaharan, ''Ocean wave forecasting using recurrent neural networks,'' *Ocean Eng.*, vol. 33, no. 10, pp. 1401–1410, Jul. 2006.
[15] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, ''Empirical evaluation of gated recurrent neural networks on sequence modeling,'' 2014, *arXiv:1412.3555*.
[16] X. Chen and W. Huang, ''Spatial–temporal convolutional gated recurrent unit network for significant wave height estimation from shipborne marine radar data,'' *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 4201711.
[17] S. Hochreiter and J. Schmidhuber, ''Long short-term memory,'' *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997.
[18] S. Fan, N. Xiao, and S. Dong, ''A novel model to predict significant wave height based on long short-term memory network,'' *Ocean Eng.*, vol. 205, Jun. 2020, Art. no. 107298.

[19] M. M. Pushpam P. and F. Enigo V. S., ''Forecasting significant wave height using RNN-LSTM models,'' in *Proc. 4th Int. Conf. Intell. Comput. Control Syst. (ICICCS)*, May 2020, pp. 1141–1146.
[20] S. Li, P. Hao, C. Yu, and G. Wu, ''CLTS-Net: A more accurate and universal method for the long-term prediction of significant wave height,'' *J. Mar. Sci. Eng.*, vol. 9, no. 12, p. 1464, Dec. 2021.
[21] N. Raj and J. Brown, ''An EEMD-BiLSTM algorithm integrated with Boruta random forest optimiser for significant wave height forecasting along coastal areas of Queensland, Australia,'' *Remote Sens.*, vol. 13, no. 8, p. 1456, Apr. 2021.
[22] H. Drucker, C. J. C. Burges, L. Kaufman, A. Smola, and V. Vapnik, ''Support vector regression machines,'' in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 9, 1997, pp. 155–161.
[23] V. Vapnik, *The Nature of Statistical Learning Theory*. Cham, Switzerland: Springer, 1999.
[24] J. Berbić, E. Ocvirk, D. Carević, and G. Lončar, ''Application of neural networks and support vector machine for significant wave height prediction,'' *Oceanologia*, vol. 59, no. 3, pp. 331–349, Jul. 2017.
[25] D. Gao, Y. Liu, J. Meng, Y. Jia, and C. Fan, ''Estimating significant wave height from SAR imagery based on an SVM regression model,'' *Acta Oceanologica Sinica*, vol. 37, no. 3, pp. 103–110, Mar. 2018.
[26] K. Zhao and J. Wang, ''Significant wave height forecasting based on the hybrid EMD-SVM method,'' *Indian J. Geo Mar. Sci.*, vol. 48, no. 12, pp. 1957–1962, 2019.
[27] J. Mahjoobi and E. A. Mosabbeb, ''Prediction of significant wave height using regressive support vector machines,'' *Ocean Eng.*, vol. 36, no. 5, pp. 339–347, Apr. 2009.
[28] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, ''The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis,'' *Proc. Roy. Soc. London. Ser. A, Math., Phys. Eng. Sci.*, vol. 454, no. 1971, pp. 903–995, Mar. 1998.
[29] R. Prahlada and P. C. Deka, ''Forecasting of time series significant wave height using wavelet decomposed neural network,'' *Aquatic Proc.*, vol. 4, pp. 540–547, 2015.
[30] M. R. Kaloop, D. Kumar, F. Zarzoura, B. Roy, and J. W. Hu, ''A wavelet—Particle swarm optimization—Extreme learning machine hybrid modeling for significant wave height prediction,'' *Ocean Eng.*, vol. 213, Oct. 2020, Art. no. 107777.
[31] P. Dixit, S. Londhe, and Y. Dandawate, ''Removing prediction lag in wave height forecasting using neuro—Wavelet modeling technique,'' *Ocean Eng.*, vol. 93, pp. 74–83, Jan. 2015.
[32] J. A. K. Suykens and J. Vandewalle, ''Least squares support vector machine classifiers,'' *Neural Process. Lett.*, vol. 9, no. 3, pp. 293–300, Jun. 1999.
[33] M. S. AL-Musaylh, R. C. Deo, Y. Li, and J. F. Adamowski, ''Two-phase particle swarm optimized-support vector regression hybrid model integrated with improved empirical mode decomposition with adaptive noise for multiple-horizon electricity demand forecasting,'' *Appl. Energy*, vol. 217, pp. 422–439, May 2018.
[34] M. Ali and R. Prasad, ''Significant wave height forecasting via an extreme learning machine model integrated with improved complete ensemble empirical mode decomposition,'' *Renew. Sustain. Energy Rev.*, vol. 104, pp. 281–295, Apr. 2019.
[35] R. Rastogi, P. Anand, and S. Chandra, ''Large-margin distribution machine-based regression,'' *Neural Comput. Appl.*, vol. 32, no. 8, pp. 3633–3648, Apr. 2020.
[36] K. Dragomiretskiy and D. Zosso, ''Variational mode decomposition,'' *IEEE Trans. Signal Process.*, vol. 62, no. 3, pp. 531–544, Feb. 2014.
[37] J. Kennedy and R. Eberhart, ''Particle swarm optimization,'' in *Proc. Int. Conf. Neural Netw.*, vol. 4. Nov./Dec. 1995, pp. 1942–1948.
[38] J. Mercer, ''XVI. Functions of positive and negative type, and their connection the theory of integral equations,'' *Phil. Trans. Roy. Soc. London. Ser. A, Containing Papers A Math. Phys. Character*, vol. 209, nos. 441–458, pp. 415–446, 1909.
[39] J. Demšar, ''Statistical comparisons of classifiers over multiple data sets,'' *J. Mach. Learn. Res.*, vol. 7, pp. 1–30, Dec. 2006.

PRITAM ANAND received the Ph.D. degree from South Asian University, New Delhi. He is currently an Assistant Professor with the Dhirubhai Ambani Institute of Information and Communication Technology (DA-IICT), Gandhinagar, India. He has coauthored about 20 research articles which includes reputed journal and conference publications. His research interests include support vector machines, regression models, loss functions, quantile regression models extreme learning machines, and time series forecasting.

SHANTANU JAIN received the B.Tech. degree in electrical engineering from Pandit Deendayal Petroleum University, Gandhinagar, India, in 2019, and the M.Tech. (ICT) degree in machine learning from the Dhirubhai Ambani Institute of Information and Communication Technology (DA-IICT), Gandhinagar, in 2022. He is currently a Data Scientist/Machine Learning Engineer with InfoObjects Inc., Jaipur.

HARSH SAVALIYA received the B.Tech. degree in computer engineering from the Atmiya Institute of Technology and Science, Rajkot, in 2021. He is currently pursuing the M.Tech. (ICT) degree in software systems with the Dhirubhai Ambani Institute of Information and Communication Technology (DA-IICT), Gandhinagar, India.