Received: 4 April 2023 | Revised: 28 December 2023 | Accepted: 24 January 2024

DOI: 10.1049/rpg2.12961

IET Renewable Power Generation

The Institution of Engineering and Technology WILEY

ORIGINAL RESEARCH

An integrated methodology for significant wave height forecasting based on multi-strategy random weighted grey wolf optimizer with swarm intelligence

Emrah Dokur[1] | Nuh Erdogan[2] | Mahdi Ebrahimi Salari[3] | Ugur Yuzgec[1] | Jimmy Murphy[3]

[1]School of Engineering, Bilecik Seyh Edebali University, Bilecik, Turkey

[2]School of Science and Technology, Nottingham Trent University, Nottingham, UK

[3]MaREI, The SFI Research Centre for Energy, Climate and Marine, University of College Cork, Cork, Ireland

Correspondence
Nuh Erdogan, Department of Engineering, School of Science and Technology, Nottingham Trent University, Nottingham, NG11 8NS, U.K.
Email: nuh.erdogan@ntu.ac.uk

Funding information
Türkiye Bilimsel ve Teknolojik Araştirma Kurumu, Grant/Award Number: 1059B192001283

Abstract
While wave energy is regarded as one of the prominent renewable energy resources to diversify global low-carbon generation capacity, operational reliability is the main impediment to the wide deployment of the related technology. Current experience in wave energy systems demonstrates that operation and maintenance costs are dominant in their cost structure due to unplanned maintenance resulting in energy production loss. Accurate and high performance simulation forecasting tools are required to improve the efficiency and safety of wave converters. This paper proposes a new methodology for significant wave height forecasting. It is based on incorporating swarm decomposition (SWD) and multi-strategy random weighted grey wolf optimizer (MsRwGWO) into a multi-layer perceptron (MLP) forecasting model. This approach takes advantage of the SWD approach to generate more stable, stationary, and regular patterns of the original signal, while the MsRwGWO optimizes the MLP model parameters efficiently. As such, forecasting accuracy has improved. Real wave datasets from three buoys in the North Atlantic Sea are used to test and validate the forecasting performance of the proposed model. Furthermore, the performance is evaluated through a comparison analysis against deep-learning based state-of-the-art forecasting models. The results show that the proposed approach significantly enhances the model's accuracy.

1 | INTRODUCTION

Ocean energy, derived from abundant and geographically diverse renewable resources such as waves, tidal streams, ocean thermal energy conversion, and salinity gradients [1], plays a key role in the global transition towards sustainable and clean energy production. Projections indicate that it could contribute to 10% of Europe's electricity needs, representing a significant share of total renewable generation [2]. The development of wave and tidal energy technology has gained momentum globally, with particular emphasis in Europe, as the global wave and tidal power capacity reached 66.1 MW in 2022 [3]. This is expected to have a 337 GW installation capacity by 2050 [2]. Despite these advancements, the diversity of ocean energy devices, particularly wave energy converters, lacks a standardized design, hindering their widespread adoption compared to other renewable technologies like wind energy [4]. Realizing the full potential of ocean energy requires optimizing capacity factors and enhancing the entire life cycle of processes, with a strategic focus on reducing the high levelized cost of electricity through reliable and flexible wave renewable energy devices [5, 6].

In marine science research, the significant wave height (SWH) stands as a critical parameter, representing the power of ocean waves and playing a key role in the optimal design of wave energy converters [7]. Accurate SWH forecasting tools are imperative for ensuring smooth and controllable power, minimizing the need for storage devices, and reducing overall system costs. The short-term forecasting of SWH, particularly over a few hours, is of vital interest in various ocean engineering operations [8]. Decision-makers in maritime traffic and marine structure construction rely on accurate SWH predictions for efficient planning [9]. However, the inherently non-linear and non-stationary nature of ocean waves poses a challenge to short-term SWH forecasting. Existing numerical models like WAM [10], simulating waves nearshore [11], and WAVEWATCH-III [12], while effective, often demand excessive computational resources, limiting their application to medium- and long-term forecasting [13]. In response, alternative approaches such as autoregressive moving averages [14] and seasonal autoregressive integrated moving averages [15] have been explored. Nonetheless, the linear and stationary assumptions of time series models result in lower performance for marine wave forecasting [16].

The advancement of computer science and technology has enabled the application of advanced artificial intelligence (AI) models to the stochastic nature of marine wave data. AI-based models outperform numerical models since they do not rely on precise meteorological data [17]. Using past SWH values, correlation can be achieved and used to extrapolate the future. Several studies applied artificial neural networks to wave forecasting [18–21]. Deep learning methods such as convolutional neural networks [22], the gate recurrent unit (GRU) network [23], and long short-term memory (LSTM) [24, 25] have become widespread. In Wang et al. [26], the GRU has been shown to have the ability to produce better SWH forecasting performance and capture the general data trend as compared to support vector machine (SVM) and extreme learning machine (ELM) models. It does not, however, achieve a completely satisfactory forecasting performance for extreme event peak wave heights. While machine learning models are widely used today, hybrid models have been developed to improve model performance.

Hybrid AI-based models utilize preprocessing or optimization algorithms to enhance forecasting performance [27, 28]. Due to the non-stationary and non-linear properties of ocean waves, a single resolution component is not sufficient to estimate wave heights. Herein, some preprocessing approaches can be used to decompose wave time series data. The use of decomposition methods such as empirical mode decomposition (EMD) [29], a multi-stage multivariate variational mode decomposition (VMD) [30], complete ensemble empirical mode decomposition (CEEMDAN) [31], and discrete wavelet transformation [32] has been shown to be effective in various forecasting applications. To reduce non-linearity and non-stationarity in wave data, Hao et al. [33] proposed a hybrid method based on EMD-LSTM. The EMD-LSTM model is shown to have higher accuracy than the single LSTM model without adding any preprocessing steps. Therefore, the EMD process can effectively deal with the non-stationary characteristics of the wave data and develop the prediction performance of the LSTM model. As each decomposition method has its own strengths and constraints, oscillations of different amplitudes in one mode or similar oscillations in different modes might occur in an EMD-based preprocessing step. Moreover, selecting a main wavelet function in wavelet-based decomposition methods is still fairly challenging. Swarm decomposition (SWD) [34] has been shown to be effective in dealing with the mode mixing problem of EMD for synthetic and real-time applications in several recent studies, including [17, 35, 36]. Given these advantages, the SWD was selected to decompose wave time series data in this study.

The optimal choice of parameters for machine learning models is critical since they have a significant impact on forecasting results. In this regard, meta-heuristic approaches have grown in popularity for tuning the parameters [37]. In [38], particle swarm optimization (PSO) was integrated into the optimization of input weights and hidden biases of ELM for the SWH forecasting problem. Wavelet analysis was also applied to reduce signal non-stationarity and nonlinearity. Hybrid PSO-ELM with wavelet analysis outperforms the ELM-based models. Similarly, in [39], a back propagation (BP) neural network was optimized to forecast wave heights using the mind evolutionary algorithm (MEA). Following the optimization process, it is shown that the MEA-BP hybrid model achieves higher accuracy than the standard BP neural network model. Thus, the use of meta-heuristic approaches increases model performance in terms of accuracy, as do decomposition methods. Among the meta-heuristic approaches, the Gray-Wolf Optimization algorithm (GWO) is prominent thanks to its advantages embedded within the search mechanism [40–42]. It is based on mimicking the hunting strategy and social leadership of gray wolves. However, it challenges some issues such as premature convergence, being stuck at the local minima, and over-fitting that are mostly encountered in the implementation. To address this deficiency, Lu et al. in [41] incorporated the concept of the cellular automatic approach into the GWO. As such, the model performance is enhanced. The authors in [42] proposed a multi-strategy random weighted grey wolf optimizer (MsRwGWO) to improve the search performance of GWO by implementing new mechanisms such as a transition mechanism, a random weighted updating mechanism, a mutation operator, and a boundary checking mechanism. The MsRwGWO's performance has been proven with benchmark functions known as CEC 2014 in terms of convergence, search history, trajectory, and average distance. Considering all these advantages, the MsRwGWO may show a superior approach in determining the optimal parameters of forecasting models such as the multi-layer perceptron (MLP), ELM, the adaptive neural fuzzy inference system (ANFIS), etc. The MsRwGWO model outperformed its counterparts in a wind speed forecasting application for the analysis considered in [42]. However, MsRwGWO has yet to be investigated with any decomposition methods or applied to wave data forecasting in the literature.

The main purpose of this study is to develop a forecasting tool that provides accurate SWH. It is expected to assist in the design optimization of wave converter devices. As such, planned and unplanned maintenance events can be reduced, decreasing the cost of operations while maximizing energy production. Hence, an improved methodology is proposed for short-term SWH. The proposed methodology integrates a signal processing module (i.e., SWD) and a meta-heuristic optimization algorithm (i.e., MsRwGWO) into the conventional forecasting models. The SWD is performed to reduce wave data nonlinearity and non-stationarity, while MsRwGWO is employed to tune the neural network model parameters. To test and validate the proposed hybrid model, its performance is compared with that of deep-learning and meta-heuristic-based forecasting models, namely, bidirectional LSTM (Bi-LSTM), GWO, and standalone MsRwGWO. The original SWH data from three regions, namely, AMETS Berth A Buoy, Clew Bay, and Smart Bay, are used.

In the remainder of this paper, Section 2 presents the features of the proposed forecasting model following the descriptions of the SWD, MsRwGWO, and MLP models. Section 3 introduces the characteristics of the collected SWH data sets, their decomposition, and experimental results with a thorough performance analysis. Section 5 concludes with concluding remarks and addresses future research directions.

2 | METHODOLOGY

2.1 | The swarm decomposition method

The SWD is an intelligent signal analysis method to refine non-stationary and multi-component signals. The main structure of this method is based on the swarm filtering concept. It extracts oscillatory components (OC) from signals using a virtual swarm-prey hunting algorithm. First, the dominant OC is estimated for each iteration. The estimated component is compared to the OC obtained by using SWF. The process is terminated if any OC does not have sufficient energy. The SWD consists of two aspects, namely, the swarming model and swarm-prey hunting. The swarming model uses two interaction forces: driving and cohesion. If i and n are the number of members and steps, respectively, the driving force $F_{\mathrm{dr}}(n, i)$ is defined as follows:

$$F_{\mathrm{dr}}(n, i) = P_{\mathrm{prey}}(n) - P_i(n-1), \tag{1}$$

where the position of the prey is represented by P_{prey}. Unlike the driving force, the cohesion force $F_{\mathrm{Coh},i}$ represents the interactions of all members and can be calculated as follows:

$$F^n_{\mathrm{Coh},i} = \frac{1}{M-1} \cdot \sum_{j=1, j\neq i}^{M} f(P_i[n-1] - P_j(n-1)), \tag{2}$$

$$f(d) = -sgn(d) \cdot \ln\left(\frac{|d|}{d_{\mathrm{cr}}}\right), \tag{3}$$

where M denotes the number of swarms and the $f(.)$ function is determined by the members' distance d and critical distance d_{cr}. Here, $sgn(.)$ and $ln(.)$ are the sign and logarithmic functions, respectively. During the swarm-prey hunting process, members must update their positions, P_i, and velocities, V_i, as follows:

$$V_i[n] = V_i[n-1] + \delta \cdot \left(F^n_{Dr,i} + F^n_{\mathrm{Coh},i}\right), \tag{4}$$

$$P_i[n] = P_i[n-1] + \delta \cdot (V_i[n]), \tag{5}$$

where the parameter, δ, controls the swarm's flexibility. After the hunt is completed, the output of SWF can be represented as follows:

$$y[n] = \beta \cdot \sum_{i=1}^{M} P_i[n], \tag{6}$$

where β is the scale parameter. For β, a low value, such as 0.005, is chosen [43]. Two critical parameters influence SWF output: δ and M. The following is the determination of these parameters based on normalized frequency, $\hat{\omega}$:

$$\arg_{\delta,M} min \sum_k \left\{\left|Y_{\delta,M}[k] - |S[k]|\right|\right\}^2, \tag{7}$$

$$M(\hat{\omega}) = \left[33.46\hat{\omega}^{-0.735} - 29.1\right], \tag{8}$$

$$\delta(\hat{\omega}) = -1.5\hat{\omega}^2 + 3.454\hat{\omega} - 0.01. \tag{9}$$

The SWD's main idea is to iteratively perform filter-like operations on the original signal. Detailed information about the SWD process can be found in [34].

While wavelet-based decomposition [44], EMD [17], ensemble empirical mode decomposition (EEMD) [45], and complete ensemble empirical mode decomposition adaptive noise approaches (CEEMDAN) [35] are widely used in the literature, they have some limitations. One of the shortcomings of the EMD technique is the occurrence of mode mixing, as well as end effects and a lack of an appropriate mathematical foundation. The EEMD approach has constraints in terms of computing complexity, number of ensemble trials, and determining the additional noise amplitude. The difficulty with wavelet-based decompositions is determining the appropriate wavelet base function and the number of decomposition layers. A further limitation is the necessity of predetermining the key parameters of the VMD method. Considering all these approaches, SWD outperforms other signal decomposition systems in terms of enhancing decomposition adaptability and eliminating mode aliasing. The SWD method also allows for the efficient decomposition of a signal into components while preserving its physical meaning [17]. It has proven efficient in various research areas, such as renewable energy [36] and biomedical signals [34]. To the best of our knowledge, this study is the first attempt to apply the SWD to wave forecasting.

2.2 | Multi-strategy random weighted grey wolf optimizer (MsRwGWO)

The GWO is a swarm intelligence meta-heuristic algorithm that is inspired by the leadership hierarchy and the hunting process of grey wolves, as proposed by [40]. The hierarchy of grey wolves consists of four types of wolves: alpha, beta, delta, and omega. These wolves are represented hierarchically by the best solution relative to the rest of the candidate solutions in the algorithmic process. The mathematical structure of the algorithm is also inspired by three main hunting processes, namely searching, encircling, and attacking prey. A mathematical model for encircling prey is defined by

$$\begin{cases} \vec{A} = 2\vec{b} \times \vec{r_1} - \vec{b} \\ \vec{C} = 2 \cdot \vec{r_2} \\ D = \left| \vec{C} \cdot X_p(\vec{t})) - X(\vec{t})) \right| \\ \vec{X}(t+1) = \vec{X}_p(t) - D \cdot \vec{A} \end{cases} \tag{10}$$

Herein, the current iteration and the distance between grey wolves are represented by t and $\vec{D}$, respectively. The prey's position vector is $\vec{X_p}$. The cooperative coefficient vectors for alpha $(\vec{C_1}, \vec{A_1})$, beta $(\vec{C_2}, \vec{A_2})$, delta $(\vec{C_3}, \vec{A_3})$ wolves are $\vec{A}$ and $\vec{C}$. $\vec{X}$ is the position vector of a grey wolf. $\vec{b}$ is a vector with a magnitude of linearly decreased values from 2 to 0. $\vec{r_1}$ and $\vec{r_2}$ are vectors chosen at random from the vectors in [0,1].

Omega wolves must update their positions in the grey wolf hunting process based on the positions of alpha, beta, and gamma wolves $(\vec{X_\alpha}, \vec{X_\beta}, \vec{X_\gamma})$. The whole hunting process is given as follows:

$$\begin{cases} D_\alpha = \left|\vec{C_1} \cdot \vec{X_\alpha} - \vec{X}\right|, D_\beta = \left|\vec{C_2} \cdot \vec{X_\beta} - \vec{X}\right|, D_\gamma = \left|\vec{C_3} \cdot \vec{X_\gamma} - \vec{X}\right| \\ \vec{X_1} = \vec{X_\alpha} - D_\alpha \cdot \vec{A_1}, \vec{X_2} = \vec{X_\beta} - D_\beta \cdot \vec{A_2}, \vec{X_3} = \vec{X_\gamma} - D_\gamma \cdot \vec{A_3}, \\ \vec{X}(t+1) = \frac{\vec{X_1} + \vec{X_2} + \vec{X_3}}{3} \end{cases} \tag{11}$$

Until the termination condition is satisfied, the iterative loop continues. Finally, the last position of alpha is the optimal solution in the source space.

To improve the search performance of the algorithm, a new GWO variant, namely MsRwGWO, has been proposed by the authors' previous study in [42] in which some effective and novel mechanisms were added to the original GWO. A transition mechanism was used to update parameter $\vec{b}$ in Equation (10), a transition mechanism was used. While this parameter decreases linearly for GWO, many problems need non-linear variation. In this way, local optimal solutions are avoided. The transition mechanism of the MsRwGWO is expressed as:

$$\vec{b} = 2 \cdot sin\left(\left(1 - \frac{iter}{Max_{iter}}\right) \cdot \frac{\pi}{2}\right) \cdot ones[size(Z)], \tag{12}$$

where Z is the dimension of the MLP problem. In [42], a mutation operator has also been adapted. According to a weighted updating mechanism, the fitness score determined the new position of the wolves. The update mechanism of the MsRwGWO is given as follows:

$$S = \sum_{i=\alpha,\beta,\gamma} \frac{1}{f(\vec{X_i})}, \tag{13}$$

$$w_\alpha = \frac{f(\vec{X_\alpha})^{-1}}{S}, w_\beta = \frac{f(\vec{X_\beta})^{-1}}{S}, w_\gamma = \frac{f(\vec{X_\gamma})^{-1}}{S}, \tag{14}$$

$$\vec{X_i} = w_\alpha \vec{U_\alpha} + w_\beta \vec{U_\beta} + w_\gamma \vec{U_\gamma}. \tag{15}$$

Herein, the fitness value of the positions is represented by $f(\vec{X_i})$. S and $\vec{X_i}$ denote the sum scores and positions of the alpha, beta, and delta wolves. The positions of grey wolves are updated by averaging the trial vectors $(\vec{U_\alpha}, \vec{U_\beta}, \vec{U_\gamma})$. To improve the position of wolves, a mutation operator is adopted as

$$\vec{X_i}(t+1) = \vec{X_i}(t) + 0.1(\vec{U_b} - \vec{L_b}) \cdot r_m, \tag{16}$$

where U_b and L_b are the upper and lower boundaries of the search agent, and r_m is a normally distributed random number. In addition to all these mechanisms, for the leader three wolves, a boundary checking mechanism, a greedy selection mechanism, and an updating mechanism are added to the original GWO algorithm. The performance of GWO in the search space has been shown to improve in [42] with all six different update mechanisms. Moreover, the performance of the MsRwGWO has been investigated to tune the parameters of MLP in a wind speed forecasting problem. With the motivations stated above, the MsRwGWO is performed to forecast SWH data in this study. Thus, this paper aims to utilize the advantages of the MsRwGWO with six distinct improved mechanisms to determine the optimal forecasting model parameters, thereby enhancing the forecasting model's performance.

Parameter Settings: To mitigate the risk of converging to local optima in the search space, the MsRwGWO employed in this study involves three fundamental parameters. These are the number of search agents (population size), the maximum number of iterations, and the mutation probability. Fine-tuning these parameters is essential for optimizing algorithm performance. In this study, experiments were conducted to test various parameter configurations for the MsRwGWO algorithm. The specific parameter settings considered are listed below:

(1) The optimization process of the MsRwGWO algorithm in the search space utilizes reaching the maximum number of iterations as the stopping criterion. The maximum number of iterations is set to 1000.
(2) The number of search agents is determined by the problem dimension. It is set to ten times the problem dimension (10 × D).
(3) To mitigate the risk of converging to local optima in the search space, the probability value of the mutation operator in MsRwGWO is set to a small value of 5×10^{-3}.

2.3 | Multi-layer perceptron

Neural network architectures include many types of algorithms with different topologies to suit specific applications. These algorithms differ mostly in terms of the information processing methods used [46]. The MLP model is often used for forecasting problems. In [47], the MLP model is found to be appropriate for non-stationary objectives and problems with very noisy and/or sparse gradients. This can make it suitable for wave forecasting. In order to effectively intervene between external input and network output, the MLP operates computational nodes known as hidden neurons. The input layer, hidden layer, and

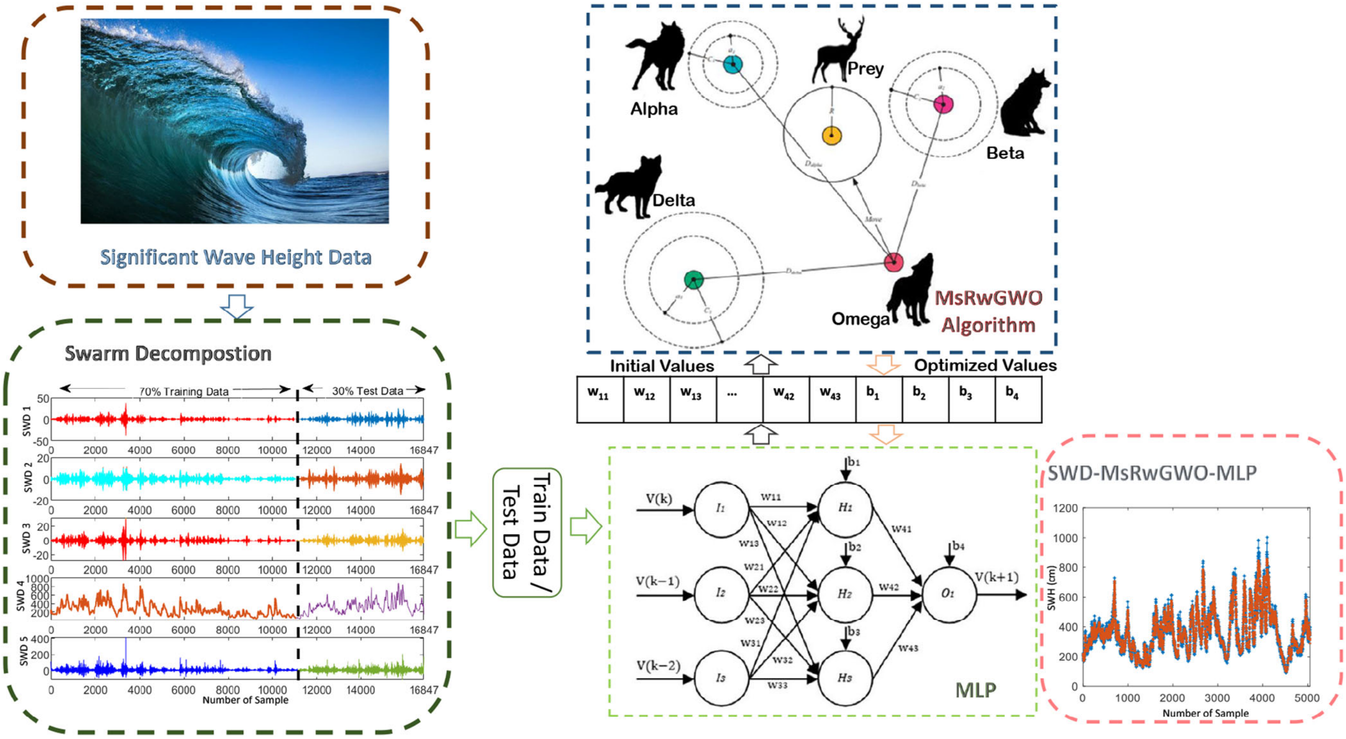

FIGURE 1 The main structure of the proposed forecasting methodology.

output layer form the main components of MLP. Each input v_k to a neuron, m, is multiplied by an adaptive coefficient, w_{mk}, called weight, and then the weighted sum of the inputs is calculated using a nonlinear activation function (φ) such as a sigmoid, hyperbolic tangent, etc. as follows:

$$y_m = \varphi\left(\sum_{k=1}^{n} w_{mk}.v_k + b_m\right), \tag{17}$$

where n and y_m represent the number of the inputs and the symbolic function of the predicted result, respectively. The network can map an input to an output thanks to an activation function, and it also gains the ability to learn complex data. In other words, the MLP model performs non-linear regression from a statistical perspective. To improve the model's performance, its parameters can be tuned by meta-heuristic algorithms.

2.4 | Framework of the proposed model: SWD-MsRwGWO-MLP

The main structure of the proposed model architecture is presented in Figure 1. To reduce the fluctuation and instability of the raw wave data, original time series signals are preprocessed by the SWD. In other words, the SWD decomposes the collected data to obtain more linear data as a preprocessing step. In the data segmentation, all decomposed signals are divided into a training set and a test set in a ratio of 70%:30%. Input data for the training and test sets is created with historical data by using the sliding window technique. In this technique, the choice of window width is critical, affecting the optimal model structure, and it is determined based on the association between current and past series values. In this study, three steps of previous data were used to forecast one step ahead. Then, the decomposed signals are fed into an MLP model by employing the sliding window technique. In this step, first, the training wave data set is input to the MLP model. In order to optimize the MLP model parameters, the proposed MsRwGWO is integrated into the model. Once the parameters along with their weights are optimized through the MsRwGWO algorithm, the MLP model is run with the test data set input. In the last stage, the forecasting results are found by taking the sum of all the individual results of the decomposed signals. As a result, the SWH of wave data is forecasted. To increase the training and test performance, normalization and denormalization operations are performed at the input and output of both processes.

The following advantages of the proposed model can be mentioned: Firstly, the SWD algorithm, which is based on a swarm-prey hunting approach, can decompose the SWH data intelligently and without loss. While it provides more stable, stationary, and regular features of the SWH time-series data, it not only effectively solves non-linearity challenges but also captures the main characteristics of the original data. Secondly, the MsRwGWO algorithm, which has been shown to be effective in forecasting some renewable applications [42] optimizes the model parameters. The MsRwGWO includes six different mechanisms that improve the model's performance. The transition mechanism in updating $\vec{b}$ parameter, for example, ensures that local optimal solutions are successfully avoided. Moreover, the other five mechanisms help to improve the exploration and

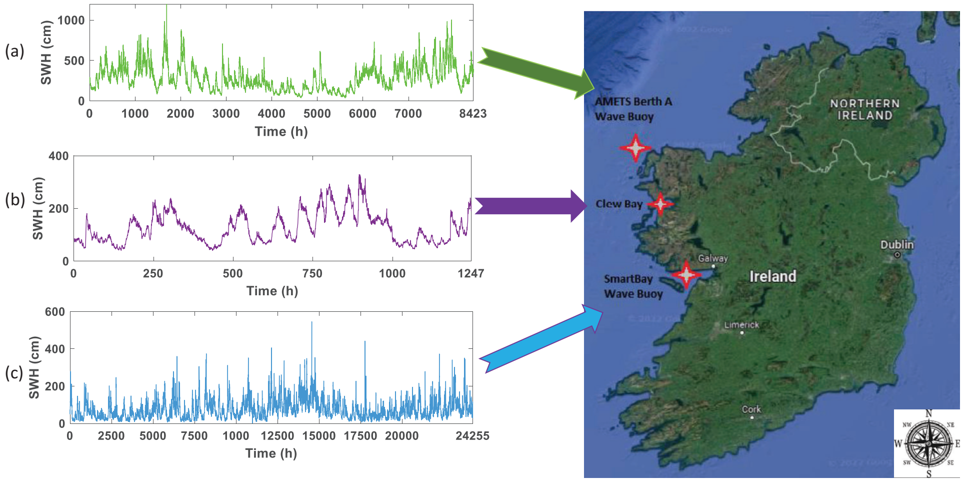

FIGURE 2 Input to the decomposition module: Original SWH datasets collected: (a) the Amets Bert A buoy, (b) the Clew Bay and (c) SmartBay Wave buoy.

exploitation abilities of the GWO algorithm. As a result, the contribution to the enhancement of forecasting accuracy of the integrated SWD and MsRwGWO-based MLP model will be assessed for the SWH application below. Furthermore, the other five mechanisms help to improve the exploration and exploitation abilities of the GWO algorithm. As a result, the contribution to the enhancement of forecasting accuracy of the integrated SWD and MsRwGWO-based MLP model will be assessed for the SWH application below.

3 | EXPERIMENTAL RESULTS AND FORECASTING ANALYSIS

3.1 | Data description

To evaluate the proposed model, the SWH data set used in this study is provided by the Marine Institute in Ireland [48]. It includes three wave datasets with different characteristics from three buoy sites. The original data and buoy locations are shown in Figure 2. Herein, all SWH data has a 30 min resolution. Amets Bert A buoy data spans the years 1 January 2021, to 31 December 2021; Clew Bay data spans the years 30 October 2021, to 1 January 2022; and SmartBay Wave buoy data spans the years 5 December 2009, to 25 February 2013. It is noted that this paper enables us to investigate the proposed model's performance in terms of different types of datasets.

The statistical descriptions of marine data, including the mean, standard deviation, minimum, maximum, kurtosis, and skewness, are reported in Table 1. As can be seen in Table 1, the fluctuation of Amets Berth A buoy data is the highest. Considering mean, maximum, and minimum values, all stations display different amplitude characteristics. While the mean of the Amets Berth A buoy data is 285.36, the SmartBay Wave buoy has the lowest mean value of 73.75. Besides, the mean value of Clew buoy data is higher than that of SmartBay Wave buoy, while the maximum value of Clew buoy data is lower. The kurtosis value and skewness value give information about the height of the distribution of marine data. In terms of skewness, the SWH is rightward, with values that are greater than zero. Compared to the kurtosis, all sites show fat tails because of values greater than 3. Considering all statistical indicators, the SWH displays remarkable differences among the sites considered.

3.2 | Forecasting results and analysis

In this section, the forecasting results of all implemented models, namely, the proposed model (SWD-MsRwGWO-MLP), the Bi-LSTM, the GWO-based MLP model, and the MsRwGWO-based MLP model, are discussed in detail. Some experiments are carried out in order to investigate the various characteristics of wave data. Herein, the purpose of the first experiment is to demonstrate the performance of the MsRwGWO data algorithm over the GWO. Secondly, the SWH forecasting results are to be presented with time series graphs, Taylor diagrams, and performance metrics in detail. One-step-ahead forecasting (30 min) is performed for all analyses.

To decompose the original data, the SWD approach is implemented first. The SWD provides a way to acquire the full details of the wave signal and improve the quality of the forecasting model. As an example, the decomposed signal of Amets Berth A buoy is shown in Figure 3. As can be seen, the original SWH data is decomposed into five sub-components. SWD4 is the main characteristic of the data, while SWD1, SWD2, SWD3, and SWD5 are residual signals. Thanks to its ability to extract the main signal, it provides an advantage in the forecasting stage. In the SWD process, there is no loss of data for the original signal. In other words, the original data can be recreated by summing the SWD decomposed signal. A similar process is realized for the entire wave dataset, and the decomposed signals are obtained.

Before assessing the contribution of the decomposition to forecasting performance, the impact of the used meta-heuristic

TABLE 1 The collected wave data statistical information.

Station	Dataset	Size	Mean (cm)	Standard deviation	Min (cm)	Max (cm)	Kurtosis	Skewness
Amets Berth A Buoy	All Data	16847	285.36	155.24	35	1197	3.79	0.91
	Training Sample	11794	253.17	148.64	35	1197	4.28	1.13
	Test Sample	5053	360.51	143.94	93	1003	4.0	0.91
Clew Bay	All Data	2494	126.29	57.98	40	329	2.97	0.75
	Training Sample	1747	127.99	56.83	40	293	2.51	0.58
	Test Sample	747	122.31	60.45	46	329	3.95	1.10
SmartBay Wave Buoy	All Data	48511	73.75	51.17	1	546	5.90	1.45
	Training Sample	33959	75.19	53.36	1	546	5.54	1.38
	Test Sample	14552	70.39	45.48	7	442	6.79	1.57

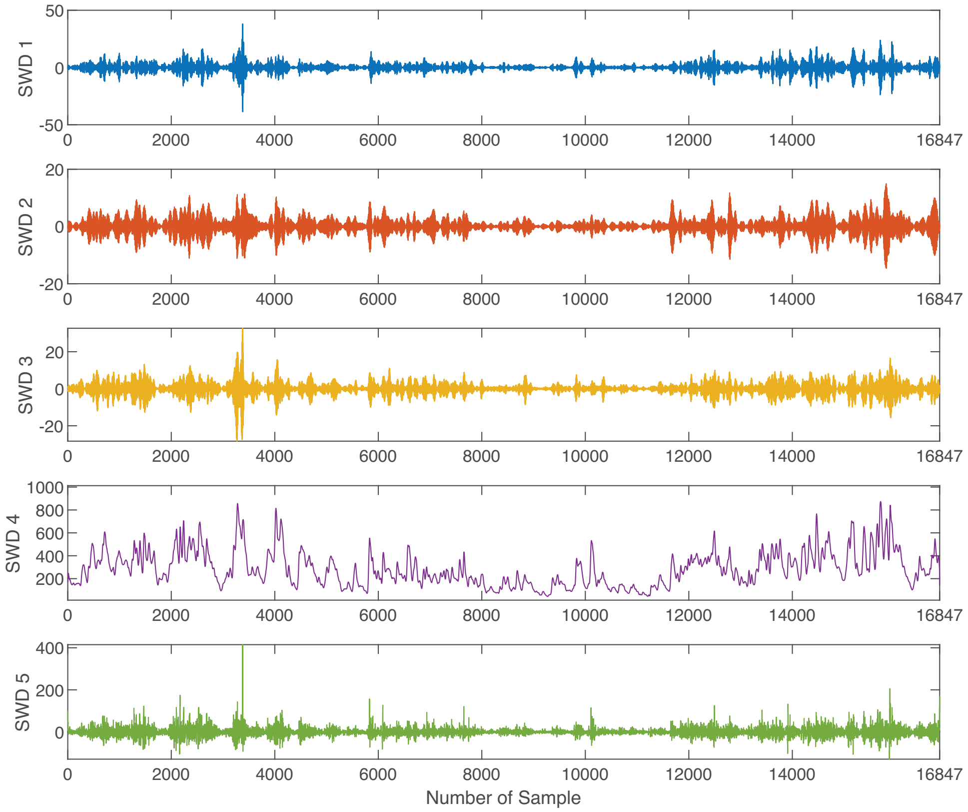

FIGURE 3 Decomposition of the SWH wave data from the Amets Berth A buoy based on the swarm decomposition approach.

optimizer (i.e. the MsRwGWO) is investigated by comparing the forecasting results with those of the GWO. For the training stage, the models' performances of GWO and MsRwGWO are shown in Figure 4. While each model follows the trend of the real signal, MsRwGWO-MLP outperforms the forecast at the local minimum points. As such, it is observed that the MLP parameters are better optimized with the MsRwGWO algorithm. The search space for optimal parameters is improved thanks to the transition mechanism in the MsRwGWO. As can be seen in Figure 4, the error values are also gathered on the zero axis. As in the author's earlier study in [42], a single MsRwGWO algorithm performs better than GWO for wind speed forecasting. Following this experiment, the proposed model focuses on adapting the decomposition process to the MsRwGWO algorithm, which has proven superior to its GWO counterpart.

The forecasting performance of the proposed model is compared with that of a deep learning model, namely, Bi-LSTM. The number of hidden layers in the Bi-LSTM model was increased by 2 for SWH data. The number of Bi-LSTM units required for four layers is calculated as 100, 100, 75, and 75. A batch size of 16 and a maximum number of training epochs of 100 are chosen. The implemented Bi-LSTM model uses the Adam

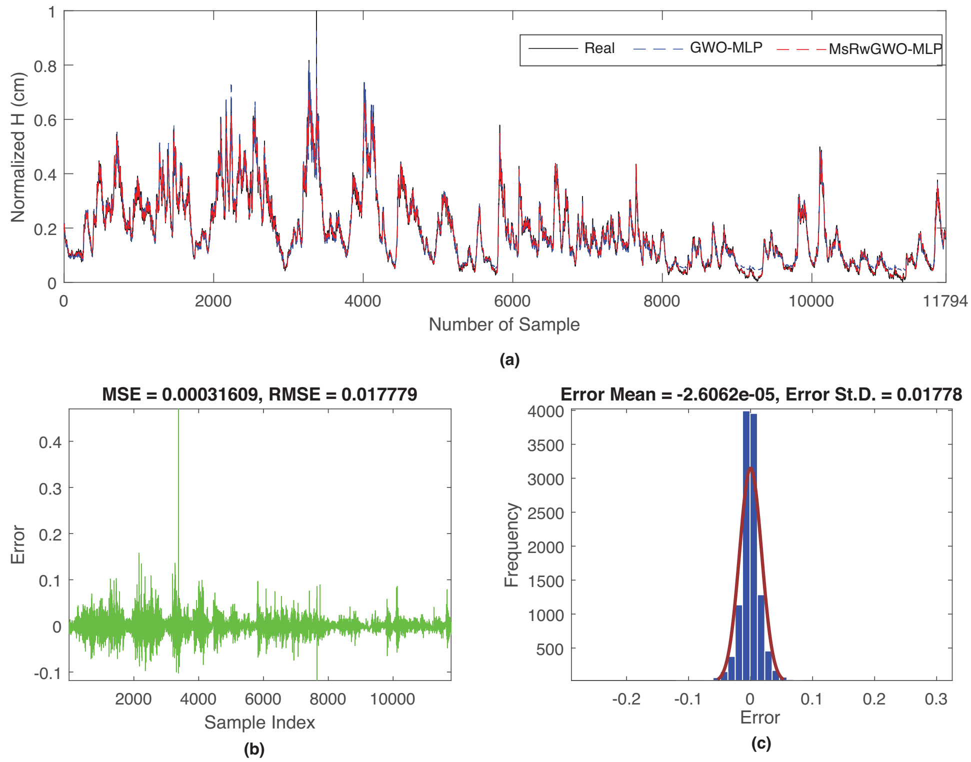

FIGURE 4 Comparison of the model results (a) train performance (b) Multi-strategy random weighted grey wolf optimizer (MsRwGWO) model error for all samples (c) error frequency of MsRwGWO model.

optimizer [49]. The learning rate is 0.005. Due to the randomness of the model parameters, all Bi-LSTM, GWO-MLP, and MsRwGWO-MLP models were run 1000 times to minimize errors. All analyses are investigated in Matlab R2020b on Windows 10 with a 2.5-GHz Intel Core i5 7200U processor and a 64-bit operating system with 8 GB of RAM. Some performance metrics, such as the root mean square error (RMSE), mean absolute error (MAE), coefficient of variation (CV), and R^2, are used to compare the performance of all implemented models. Their descriptions are given as follows:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}(y_i - \tilde{y}_i)^2}{N}}, \tag{18}$$

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}|y_i - \tilde{y}_i|, \tag{19}$$

$$\mathrm{CV} = \frac{\frac{1}{N}\sqrt{\sum_{i=1}^{N}(y_i - \tilde{y}_i)^2}}{\bar{y}_i}, \tag{20}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \tilde{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y}_i)^2}. \tag{21}$$

Figures 5 and 6 present the forecasting time series results of the implemented models for the SWH data collected from the Amets Bert A buoy and the SmartBay Wave buoy, respectively. It is seen that all models roughly captured the trend of SWH at the time of forecasting. To clearly compare the performance of the models, error metrics can be used. Table 2 reports the error metrics values for all the implemented models. In terms of performance metrics, the proposed model outperforms all others, with RMSE, MSE, CV, and R^2 values of 19.97, 14.30, 0.07, and 0.9807, respectively. In terms of accuracy, the implemented models can be sorted from highest to lowest as SWD-MsRwGWO-MLP, MsRwGWO-MLP, BiLSTM, and GWO-MLP, with R^2 values of 0.9807, 0.9660, 0.9595, and 0.9397, respectively. Similar SWH forecasting results were also

TABLE 2 Performance metric results for the forecasting analysis.

Location	Methods	RMSE	MAE	CV (%)	R^2
Amets Berth A buoy	BiLSTM	28.97	22.17	0.11	0.9595
	GWO-MLP	35.34	24.15	0.14	0.9397
	MsRwGWO-MLP	26.55	19.04	0.10	0.9660
	Proposed	19.97	14.30	0.07	0.9807
Clew Bay	BiLSTM	11.20	6.99	0.34	0.9656
	GWO-MLP	12.79	7.90	0.38	0.9551
	MsRwGWO-MLP	11.07	7.31	0.33	0.9664
	Proposed	9.93	6.53	0.30	0.9730
SmartBay Wave buoy	BiLSTM	9.31	5.50	0.11	0.9580
	GWO-MLP	11.51	7.10	0.14	0.9360
	MsRwGWO-MLP	7.90	5.10	0.09	0.9680
	Proposed	5.97	3.61	0.07	0.9820

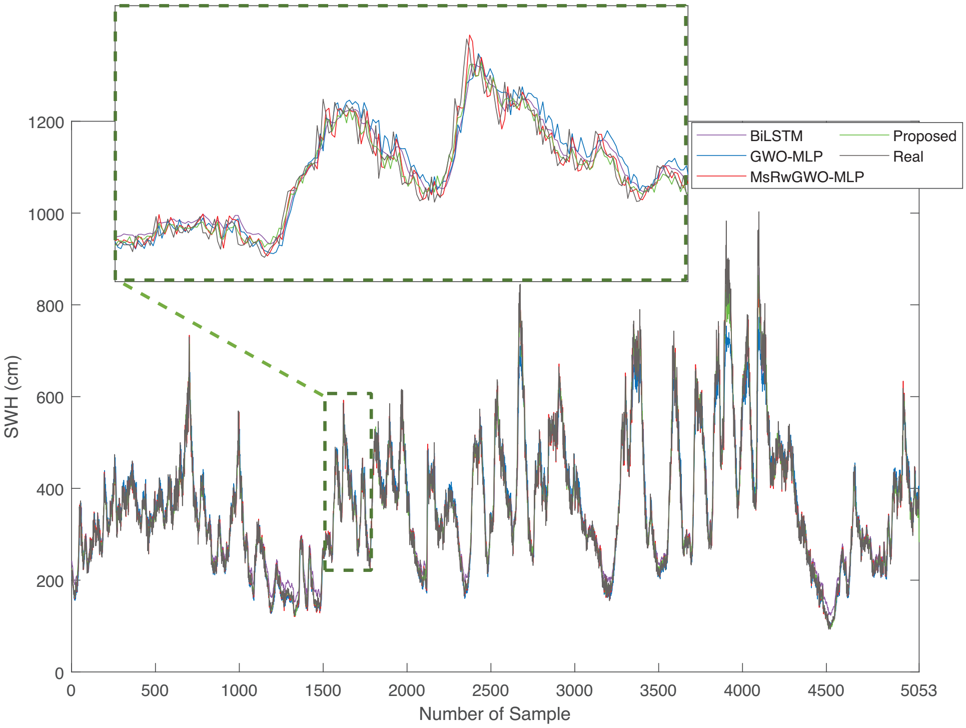

FIGURE 5 Forecasting test results of the implemented models for Amets Berth A Buoy significant wave height.

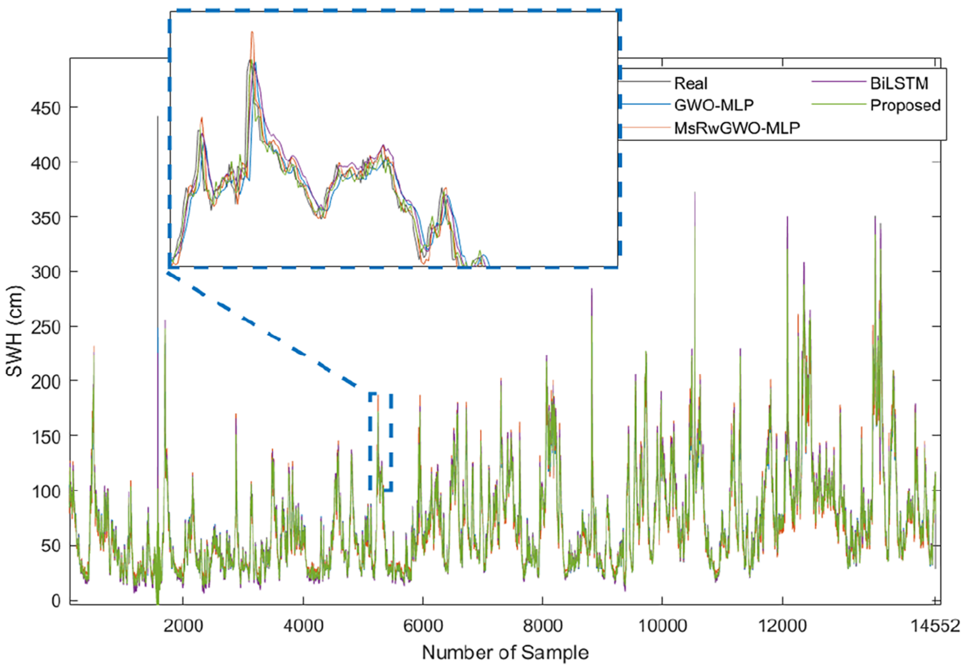

FIGURE 6 Forecasting test results of the implemented models for SmartBay Wave Buoy significant wave height.

obtained for other stations. Moreover, the R^2 of SWH data is too close to 1. As such, the SWH forecasting results obtained can be said to be more reliable and accurate. Considering the results for Clews station, it is observed that the model's performance slightly decreases due to the availability of limited training data.

Furthermore, the data set from the National Oceanic and Atmospheric Administration (NOAA) was used to demonstrate

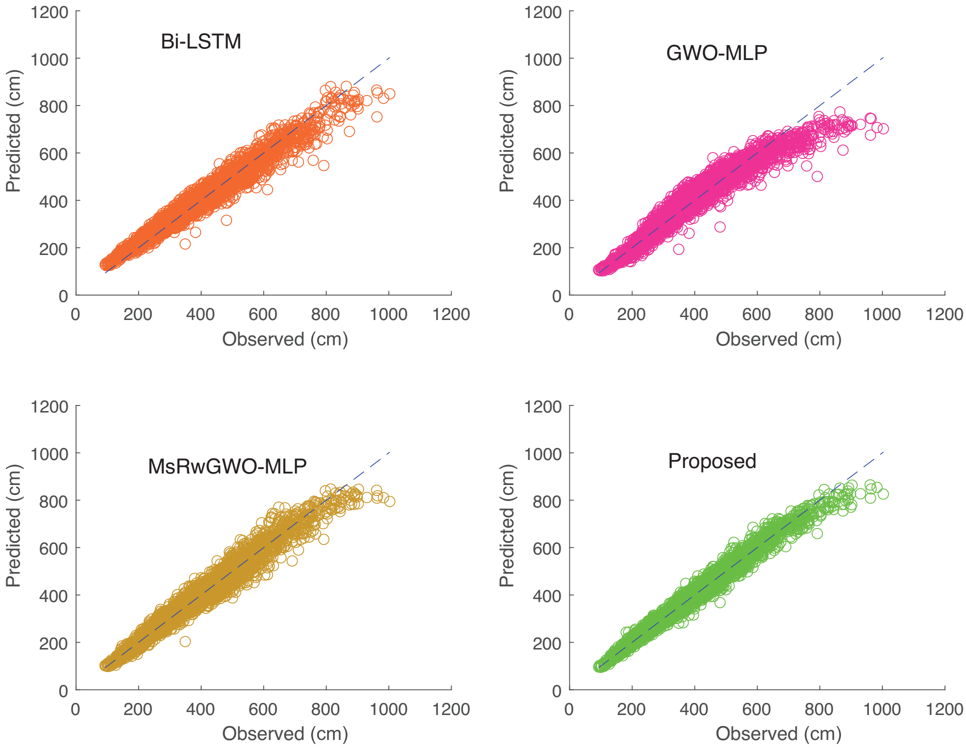

FIGURE 7 Scatter plots between original and forecasted values for the implemented models.

the superiority of the proposed methodology. The findings were compared with those of the study in [47]. The NOAA 45002 buoy station SWH data from 2014 was used for this analysis. It has been found that the MsRwGWO-MLP and the proposed models achieved R^2 values of 0.9394 and 0.9662, respectively, which are significantly higher than the R^2 value of 0.81 reported in [47]. As a result, both models displayed more precise results.

To demonstrate the superiority of the proposed methodology over a deep-learning model (e.g., Bi-LSTM) and other derived models, the scatter plots between observed and forecasted values are displayed in Figure 7 where the higher the scattering values are, the better the model's performance is. It is observed that the proposed model, SWD-MsRwGWO-MLP, has the best correlation. The findings confirm that the SWD process improved the performance of the MsRwGWO model. Moreover, to display the difference among the implemented models, their Taylor diagrams are drawn as shown in Figure 8. Thus, the link between the correlation coefficient, the root mean square deviation (RMSD), and the standard deviation is shown in Taylor diagrams. In this regard, all the models can be compared on the basis of how well they predict the target data. The triangular sign in red indicates the proposed method. As can be seen, the RMSD and standard deviation of the proposed method are lower. Furthermore, the coefficient value is close to 1. Considering all the analyses, the proposed model's performance can be said to be reliable and effective for SWH forecasting.

Although the proposed model has provided significant accuracy in SWH forecasting, there are still some limitations that require additional research and model enhancement. One of the most significant limitations is the amount of data (e.g., size) collected. Furthermore, during data collection, measurement errors may occur. During the training phase, missing or nonsense data affects the optimization of model parameters and weights. Another significant limitation is the additional computational time required for the decomposition and meta-heuristic steps. Each sub-component of the decomposition phase requires a certain amount of time to obtain. Similarly, the optimization step increases the run time for each iteration. Despite this, given the 30-min forecast horizon, these constraints have no direct impact on the SWH forecasting problem.

4 | CONCLUSIONS

This study proposed an integrated methodology for short-term SWH forecasting. The methodology incorporates a signal processing module into the forecasting model. Following the signal decomposition using the SWD, an improved meta-heuristic algorithm, namely MsRwGWO, is included to optimize the parameters of the MLP-based forecasting model. Thanks to the swarm decomposition process, the non-linearity and non-stationarity of wave data have been reduced. Furthermore, the MsRwGWO has optimized the parameters of the conventional forecasting model efficiently. Original wave data from three buoys in the North Atlantic Ocean was used to test the proposed model. For validation of the proposed model, a comparison against three deep learning (i.e. Bi-LSTM) and meta-heuristic models (i.e. GWO-MLP, MsRwGWO-MLP) has been performed.

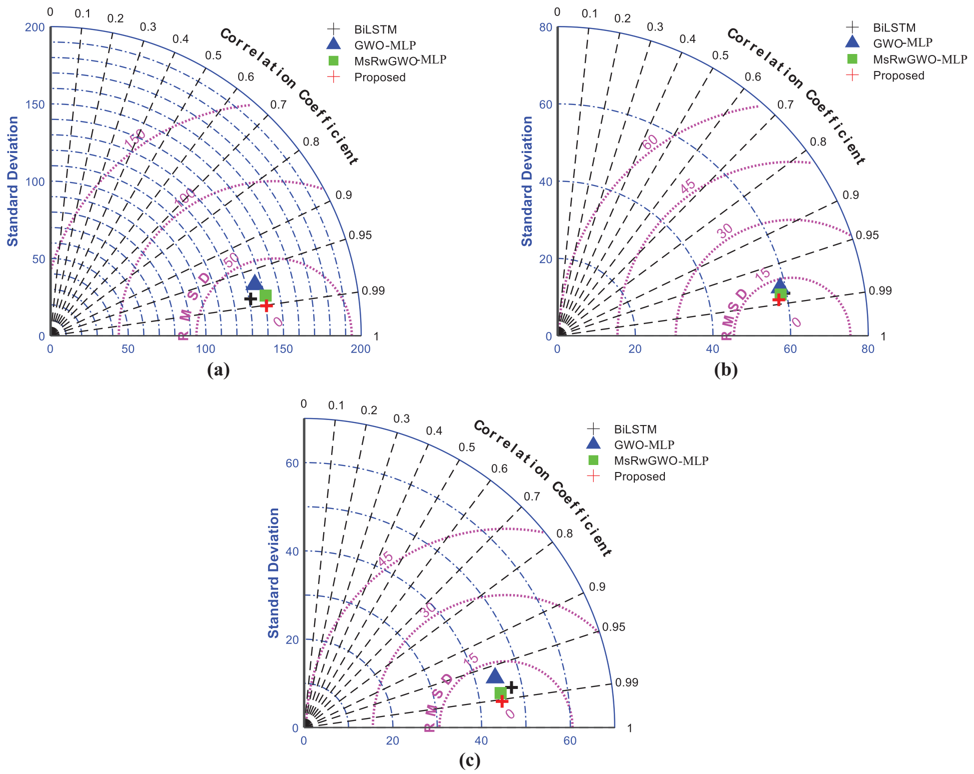

FIGURE 8 Taylor diagrams for (a) Amets Barth A buoy, (b) Clews Bay, and (c) SmartBay Wave buoy.

This proposed framework is unique in that it is a new hybrid model based on a meta-heuristic optimizer and a signal decomposition module that has significantly improved the accuracy of SWH forecasting. The proposed model has reduced forecasting errors (e.g. MAE) ranging from 17.3%–49.1% and 10.6%–29.2% depending on the wave data characteristics considered and the models compared. The comparison analysis has revealed consistency among the results for the other performance metrics. The improved optimizer with SWD has decreased forecasting errors (e.g. RMSE) by 11.3%–35.7% when compared to the deep learning approach, BiLSTM. In addition, the superiority of the proposed model over its counterpart in the literature has been shown using the well-known NOAA data set.

This model can be used to forecast other marine renewable energy problems, such as tidal currents. Moreover, the proposed approach can be investigated for problems with wind speed forecasting, electrical load forecasting, econometric forecasting etc. Future research will look into the use of a multistage decomposition process with the SWD.

AUTHOR CONTRIBUTIONS

Emrah Dokur: Conceptualization; data curation; formal analysis; investigation; methodology; software; validation; writing—original draft; writing—review and editing. **Nuh Erdogan**: Conceptualization; investigation; methodology; supervision; writing—original draft; writing—review and editing. **Mahdi Ebrahimi Salari**: Data curation; investigation; resources. **Ugur Yuzgec**: Methodology; software. **Jimmy Murphy**: Project administration; resources; supervision; writing—review and editing.

ACKNOWLEDGEMENTS

The authors wish to acknowledge the Marine Institute, Galway, Ireland, for providing a real-wave dataset.

This work was supported in part by the Scientific and Technological Research Council of Turkey through the International Postdoctoral Fellowship Program under Grant No: 1059B192001283.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

Data subject to third party restrictions.

ORCID

Nuh Erdogan https://orcid.org/0000-0003-1621-2748
Ugur Yuzgec https://orcid.org/0000-0002-5364-6265

REFERENCES

1. Borthwick, A.G.: Marine renewable energy seascape. Engineering 2(1), 69–78 (2016)
2. Ocean Energy Forum. Scaling up investments in ocean energy technologies. http://www.oceanenergy-europe.eu/ocean-energy/ (2023). Accessed 2 Feb 2024
3. Ocean Energy Europe. Ocean energy key trends and statistics-2022. https://www.oceanenergy-europe.eu/ocean-energy/ (2023). Accessed 2 Feb 2024

4. Hand, B.P., Erdogan, N., Murray, D., Cronin, P., Doran, J., Murphy, J.: Experimental testing on the influence of shaft rotary lip seal misalignment for a marine hydro-kinetic turbine. Sustainable Energy Technol. Assess 50, 101874 (2022)
5. Hafner, M.: The Palgrave Handbook of International Energy Economics. Springer Nature, Berlin, Heidelberg (2022)
6. Martinez, A., Iglesias, G.: Site selection of floating offshore wind through the levelised cost of energy: A case study in ireland. Energy Convers. Manage. 266, 115802 (2022)
7. Robertson, B., Bailey, H., Leary, M., Buckham, B.: A methodology for architecture agnostic and time flexible representations of wave energy converter performance. Appl. Energy 287, 116588 (2021)
8. Jörges, C., Berkenbrink, C., Stumpe, B.: Prediction and reconstruction of ocean wave heights based on bathymetric data using LSTM neural networks. Ocean Eng. 232, 109046 (2021)
9. Huang, W., Dong, S.: Improved short-term prediction of significant wave height by decomposing deterministic and stochastic components. Renewable Energy 177, 743–758 (2021)
10. Group, T.W.: The wam model–A third generation ocean wave prediction model. J. Phys. Oceanogr. 18(12), 1775–1810 (1988)
11. Booij, N., Ris, R.C., Holthuijsen, L.H.: A third-generation wave model for coastal regions: 1. model description and validation. J. Geophys. Res. Oceans 104(C4), 7649–7666 (1999)
12. Tolman, H.L.: The numerical model WAVEWATCH: a third generation model for hindcasting of wind waves on tides in shelf seas. Department of Civil Engineering, Delft University of Technology, Netherlands (1989)
13. Abdullah, F., Ningsih, N., Al-Khan, T.: Significant wave height forecasting using long short-term memory neural network in indonesian waters. J. Ocean Eng. Mar. Energy 8(2), 183–192 (2022)
14. Ge, M., Kerrigan, E.C.: Short-term ocean wave forecasting using an autoregressive moving average model. In: Proceedings of the 2016 UKACC 11th International Conference on Control (CONTROL), pp. 1–6. IEEE, Piscataway, NJ (2016)
15. Yang, S., Zhang, Z., Fan, L., Xia, T., Duan, S., Zheng, C., et al.: Long-term prediction of significant wave height based on sarima model in the south china sea and adjacent waters. IEEE Access 7, 88082–88092 (2019)
16. Duan, W., Han, Y., Huang, L., Zhao, B., Wang, M.: A hybrid EMD-SVR model for the short-term prediction of significant wave height. Ocean Eng. 124, 54–73 (2016)
17. Dokur, E., Erdogan, N., Salari, M.E., Karakuzu, C., Murphy, J.: Offshore wind speed short-term forecasting based on a hybrid method: Swarm decomposition and meta-extreme learning machine. Energy 248, 123595 (2022)
18. Makarynskyy, O.: Improving wave predictions with artificial neural networks. Ocean Eng. 31(5-6), 709–724 (2004)
19. Deo, M.C., Jha, A., Chaphekar, A., Ravikant, K.: Neural networks for wave forecasting. Ocean Eng. 28(7), 889–898 (2001)
20. Wei, Z.: Forecasting wind waves in the US atlantic coast using an artificial neural network model: Towards an AI-based storm forecast system. Ocean Eng. 237, 109646 (2021)
21. Bento, P., Pombo, J., Mendes, R., Calado, M., Mariano, S.: Ocean wave energy forecasting using optimised deep learning neural networks. Ocean Eng. 219, 108372 (2021)
22. Yang, S., Deng, Z., Li, X., Zheng, C., Xi, L., Zhuang, J., et al.: A novel hybrid model based on STL decomposition and one-dimensional convolutional neural networks with positional encoding for significant wave height forecast. Renewable Energy 173, 531–543 (2021)
23. Li, X., Cao, J., Guo, J., Liu, C., Wang, W., Jia, Z., et al.: Multi-step forecasting of ocean wave height using gate recurrent unit networks with multivariate time series. Ocean Eng. 248, 110689 (2022)
24. Mousavi, S.M., Ghasemi, M., Dehghan Manshadi, M., Mosavi, A.: Deep learning for wave energy converter modeling using long short-term memory. Mathematics 9(8), 871 (2021)
25. Hu, H., van der Westhuysen, A.J., Chu, P., Fujisaki-Manome, A.: Predicting lake erie wave heights and periods using XGBOOST and LSTM. Ocean Modell. 164, 101832 (2021)
26. Wang, J., Wang, Y., Yang, J.: Forecasting of significant wave height based on gated recurrent unit network in the taiwan strait and its adjacent waters. Water 13(1), 86 (2021)
27. Parmaksiz, H., Yuzgec, U., Dokur, E., Erdogan, N.: Mutation based improved dragonfly optimization algorithm for a neuro-fuzzy system in short term wind speed forecasting. Knowledge-Based Syst. 268, 110472 (2023)
28. Qian, Z., Pei, Y., Zareipour, H., Chen, N.: A review and discussion of decomposition-based hybrid models for wind energy forecasting applications. Appl. Energy 235, 939–953 (2019)
29. Zhou, S., Bethel, B.J., Sun, W., Zhao, Y., Xie, W., Dong, C.: Improving significant wave height forecasts using a joint empirical mode decomposition–long short-term memory network. J. Mar. Sci. Eng. 9(7), 744 (2021)
30. Jamei, M., Ali, M., Karbasi, M., Xiang, Y., Ahmadianfar, I., Yaseen, Z.M.: Designing a multi-stage expert system for daily ocean wave energy forecasting: A multivariate data decomposition-based approach. Appl. Energy 326, 119925 (2022)
31. Ali, M., Prasad, R.: Significant wave height forecasting via an extreme learning machine model integrated with improved complete ensemble empirical mode decomposition. Renewable Sustainable Energy Rev. 104, 281–295 (2019)
32. Prahlada, R., Deka, P.C.: Forecasting of time series significant wave height using wavelet decomposed neural network. Aquat. Procedia 4, 540–547 (2015)
33. Hao, W., Sun, X., Wang, C., Chen, H., Huang, L.: A hybrid EMD-LSTM model for non-stationary wave prediction in offshore china. Ocean Eng. 246, 110566 (2022)
34. Apostolidis, G.K., Hadjileontiadis, L.J.: Swarm decomposition: A novel signal analysis using swarm intelligence. Signal Process. 132, 40–50 (2017)
35. Dokur, E., Erdogan, N., Kucuksari, S.: EV fleet charging load forecasting based on multiple decomposition with ceemdan and swarm decomposition. IEEE Access 10, 62330–62340 (2022)
36. Dokur, E.: Swarm decomposition technique based hybrid model for very short-term solar PV power generation forecast. Elektron. Elektrotech. 26(3), 79–83 (2020)
37. Nguyen, T.H.T., Phan, Q.B.: Hourly day ahead wind speed forecasting based on a hybrid model of EEMD, CNN-BI-LSTM embedded with ga optimization. Energy Rep. 8, 53–60 (2022)
38. Kaloop, M.R., 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 Eng. 213, 107777 (2020)
39. Wang, W., Tang, R., Li, C., Liu, P., Luo, L.: A BP neural network model optimized by mind evolutionary algorithm for predicting the ocean wave heights. Ocean Eng. 162, 98–107 (2018)
40. Mirjalili, S., Mirjalili, S.M., Lewis, A.: Grey wolf optimizer. Adv. Eng. Software 69, 46–61 (2014)
41. Lu, C., Gao, L., Pan, Q., Li, X., Zheng, J.: A multi-objective cellular grey wolf optimizer for hybrid flowshop scheduling problem considering noise pollution. Appl. Soft Comput. 75, 728–749 (2019)
42. İnaç, T., Dokur, E., Yüzgeç, U.: A multi-strategy random weighted gray wolf optimizer-based multi-layer perceptron model for short-term wind speed forecasting. Neural Comput. Appl. 34(17), 14627–14657 (2022)
43. Miao, Y., Zhao, M., Makis, V., Lin, J.: Optimal swarm decomposition with whale optimization algorithm for weak feature extraction from multicomponent modulation signal. Case Stud. Mech. Syst. Signal Process. 122, 673–691 (2019)
44. Hajiabotorabi, Z., Kazemi, A., Samavati, F.F., Ghaini, F.M.M.: Improving DWT-RNN model via B-spline wavelet multiresolution to forecast a high-frequency time series. Expert Syst. Appl. 138, 112842 (2019)
45. Santhosh, M., Venkaiah, C., Kumar, D.V.: Short-term wind speed forecasting approach using ensemble empirical mode decomposition and deep boltzmann machine. Sustainable Energy, Grids Networks 19, 100242 (2019)

46. Massaoudi, M., Refaat, S.S., Chihi, I., Trabelsi, M., Oueslati, F.S., Abu-Rub, H.: A novel stacked generalization ensemble-based hybrid LGBM-XGB-MLP model for short-term load forecasting. Energy 214, 118874 (2021)
47. Feng, X., Ma, G., Su, S.F., Huang, C., Boswell, M.K., Xue, P.: A multi-layer perceptron approach for accelerated wave forecasting in lake michigan. Ocean Eng. 211, 107526 (2020)
48. Marine Institute. https://www.marine.ie/Home/home (2022). Accessed 28 Dec 2023
49. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. arXiv:14126980 (2014)

How to cite this article: Dokur, E., Erdogan, N., Salari, M.E., Yuzgec, U., Murphy, J.: An integrated methodology for significant wave height forecasting based on multi-strategy random weighted grey wolf optimizer with swarm intelligence. IET Renew. Power Gener. 18, 348–360 (2024). https://doi.org/10.1049/rpg2.12961