water

MDPI

Article

Modelling of Deep Learning-Based Downscaling for Wave Forecasting in Coastal Area

Didit Adytia [1,*], Deni Saepudin [1], Dede Tarwidi [1], Sri Redjeki Pudjaprasetya [2], Semeidi Husrin [3], Ardhasena Sopaheluwakan [4] and Gegar Prasetya [5]

[1] School of Computing, Telkom University, Jalan Telekomunikasi No. 1 Terusan Buah Batu, Bandung 40257, Indonesia
[2] Industrial and Financial Mathematics Research Group, Faculty of Mathematics and Natural Sciences, Institut Teknologi Bandung, Jalan Ganesha 10, Bandung 40132, Indonesia
[3] Marine Research Centre, Ministry of Marine Affairs and Fisheries of Indonesia, Jakarta 14430, Indonesia
[4] Center for Applied Climate Services, Agency for Meteorology, Climatology, and Geophysics, Jakarta 10720, Indonesia
[5] Indonesian Tsunami Scientific Community (IATsI), Jakarta 10110, Indonesia
* Correspondence: adytia@telkomuniversity.ac.id

Abstract: Wave prediction in a coastal area, especially with complex geometry, requires a numerical simulation with a high-resolution grid to capture wave propagation accurately. The resolution of the grid from global wave forecasting systems is usually too coarse to capture wave propagation in the coastal area. This problem is usually resolved by performing dynamic downscaling that simulates the global wave condition into a smaller domain with a high-resolution grid, which requires a high computational cost. This paper proposes a deep learning-based downscaling method for predicting a significant wave height in the coastal area from global wave forecasting data. We obtain high-resolution wave data by performing a continuous wave simulation using the SWAN model via nested simulations. The dataset is then used as the training data for the deep learning model. Here, we use the Long Short-Term Memory (LSTM) and Bidirectional LSTM (BiLSTM) as the deep learning models. We choose two study areas, an open sea with a swell-dominated area and a rather close sea with a wind-wave-dominated area. We validate the results of the downscaling with a wave observation, which shows good results.

Keywords: wave forecasting; downscaling; LSTM; BiLSTM

Citation: Adytia, D.; Saepudin, D.; Tarwidi, D.; Pudjaprasetya, S.R.; Husrin, S.; Sopaheluwakan, A.; Prasetya, G. Modelling of Deep Learning-Based Downscaling for Wave Forecasting in Coastal Area. *Water* **2023**, *15*, 204. https://doi.org/10.3390/w15010204

Academic Editors: Giuseppe Pezzinga and Rafael J. Bergillos

Received: 19 November 2022
Revised: 19 December 2022
Accepted: 29 December 2022
Published: 3 January 2023

1. Introduction

Waves are a frequent natural occurrence in the ocean. The wave height analysis is required to ensure safety in offshore and coastal areas. Shipping and other maritime operations are heavily dependent on wave conditions. Numerous methodologies, algorithms, and historical data are utilised to forecast future wave conditions [1]. High-sea waves can cause significant damage to ships and fatalities. Wave forecasting can also help optimise ship scheduling and route optimisation. Predicting wave height can boost productivity, save fuel, and prevent dangerous circumstances [2]. Those who operate in the middle of the ocean and coastal communities could benefit from creating a system capable of predicting waves in time series. However, the random and nonlinear character of ocean waves makes it challenging to estimate their height [3].

Especially in a coastal area with complex geometry, wave prediction requires a numerical simulation with a high-resolution grid to capture wave propagation accurately. The resolution of the global wave forecasting system grid is usually too coarse to capture wave propagation in coastal areas. The spatial resolution of global wave forecasting systems such as the GFS of the NOAA [4] and ERA5 of the ECMWF [5] are 0.25°, see Figure 1 as an illustration of the GFS global grid. This global grid is too coarse to predict the wave

information in intricate coastal regions accurately. This problem is usually resolved by performing dynamic downscaling that simulates the global wave condition into a smaller domain with a high-resolution grid. By performing dynamic downscaling, it is possible to produce high-resolution wave simulations in a coastal region. This method uses the global domain wave information for a high-resolution grid local simulation as the initial and boundary conditions.

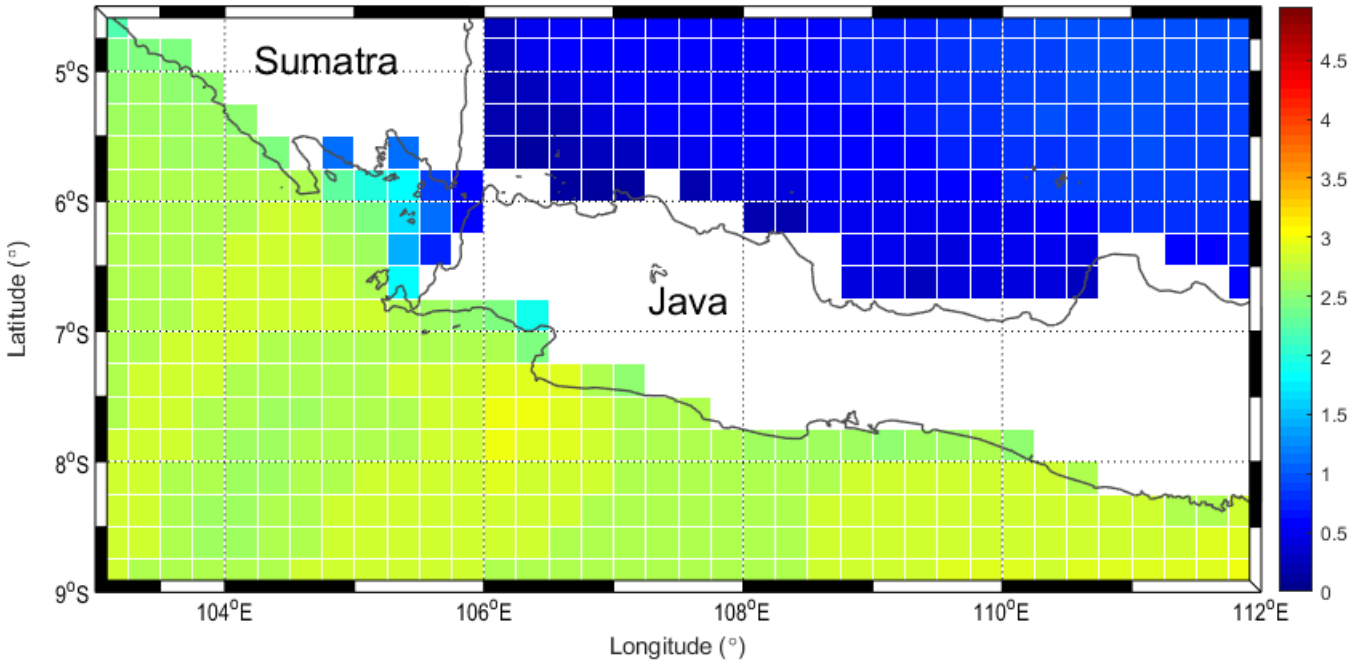

Figure 1. Snapshot of significant wave height from global forecast model GFS, with grid size of 0.25°.

High-resolution wave prediction systems rely on detailed wind fields to run accurately. Boundary conditions for the high-resolution model are taken from a coarser wave model, which provides a complete two-dimensional wave spectrum. The dynamical downscaling refers to this nested configuration, which includes complete two-dimensional wave spectrum information on the open boundaries. The downscaling approach for waves is similar to the nested method used in climate modelling. The key distinction is that the wave field utilised here is simply dependent on open boundaries, wind fields, and bathymetry [6]. Statistical downscaling is another method to predict high-resolution waves. Statistical downscaling for wave forecasting uses statistical techniques to adjust data from a coarser resolution or global models to a finer resolution or regional areas. It is used to enhance wave forecasts in areas where observational data are sparse. This is performed by applying statistical relationships between the coarse model data and observed wave data to adjust the coarse model forecasts to a finer resolution or regional area [7]. Both the dynamical and statistical downscaling of wave forecasts can provide more detailed information, which can be used to make more reliable decisions.

A machine learning approach is an alternative to numerical-based dynamical and statistical downscaling. Machine learning has gained increasing traction in coastal and ocean engineering in recent years as the need for high-resolution wave data has grown [8]. Downscaling techniques are thus utilised to compensate for the limited spatial resolution. Machine learning techniques can be used to perform downscaling. However, only a few researchers still employ this approach for wave forecasting. Michel [9] used convolutional neural network (CNN)-based deep learning for the statistical downscaling of significant wave height prediction. Their observed results were better than those of the other statistical downscaling approaches investigated but not as good as the physical models. This method's strength lies in its cheap computing cost and simple implementation. Recently,

Adytia [10] developed a significant wave height forecasting system in an environment with a complex geometry based on a combined high-resolution numerical wave simulation and Bidirectional Long Short-Term Memory (BiLSTM) deep learning method. The forecasting results using the BiLSTM had a correlation coefficient of 0.96 and a root mean squared error of 0.06.

This paper proposes a deep learning-based downscaling method for predicting significant wave height in the coastal area from global wave forecasting data. First, we obtain high-resolution wave data by performing a continuous wave simulation using the phase-averaged wave model SWAN [11] via nested simulations. The dataset is then used as the training data for the deep learning model. Here, we use the Long Short-Term Memory (LSTM) and Bidirectional LSTM (BiLSTM) as the deep learning models. We selected two study areas, an open sea with a swell-dominated area in Meulaboh, Aceh, Indonesia, and a relatively close sea with a wind-wave-dominated area in Jakarta Bay, Indonesia. As a feature selection method for the deep learning model, we selected the best locations of global wave data that have high spatially correlated with targeted local wave data to be input for the machine learning model. To test the performance of the machine learning model, we compare the prediction results with the wave observation data in Jakarta Bay and offshore Meulaboh.

The following sections are included in the core of this paper. In Section 2, we briefly present the literature review on downscaling, the wave model SWAN, and briefly, the architecture of the LSTM and BiLSTM, as they directly contribute to the wave forecasting problem. The details of the research technique, including the data generation and deep learning-based downscaling of wave data, are presented in Section 3. Section 4 compares the results of the significant wave height (Hs) in Jakarta Bay and offshore Meulaboh, Aceh, with observation data. In addition, we test the sensitivity of the training data and vary the length of the day to forecast, concerning our deep learning model's forecasting. In addition, we provide an exhaustive analysis of the experiments. Finally, Section 5 concludes the research and analyses the potential development directions.

2. Literature Review

2.1. Downscaling

In general, downscaling techniques are used to calibrate a large grid scale into a small space that is more specific to the research area. Downscaling techniques are thus utilised to compensate for the limited spatial resolution. The process of converting low-resolution meteorological data, for a given area, into high-resolution data is known as "downscaling" [12]. For example, the future local wave trend was produced using the downscaling technique. In addition, the downscaling method has been frequently employed to capture physical characteristics that require higher-resolution modelling.

Scaling methods can often be divided into statistical and dynamic scaling [13]. Statistical scaling is known to use empirical models to predict resolution variables from coarser resolution data with more precision. Artificial neural networks have recently demonstrated tremendous potential in this area due to their ability to simulate nonlinear connections [14]. Furthermore, dynamic scaling [15] uses high-resolution regional simulations to dynamically model relevant local or regional physical processes. Of course, both widely used strategies, statistical downscaling and dynamical downscaling, have advantages and disadvantages. Machine learning techniques can be used to perform downscaling.

The general flow diagram for dynamic and statistical downscaling from the global wave model (GWM) [16] is shown in Figure 2. Here, the GWM and historical data are the two inputs for the statistical downscaling. The bias correction and statistical downscaler are the main process (code) for performing the statistical downscaling. For dynamical downscaling, the output from the GWM is the only boundary condition, and the main part is to perform the local wave model (LWM). However, running an LWM is computationally expensive. In this paper, we propose deep learning-based downscaling for wave forecasting.

We use spatially correlated global wave data as the input to predict the wave conditions in the coastal area.

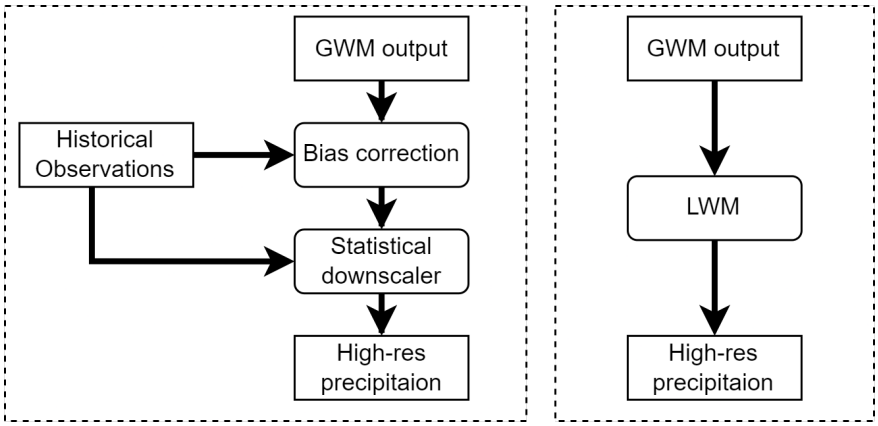

Figure 2. Flowchart of statistical downscaling (**left**) and dynamical downscaling (**right**).

2.2. Wave Model SWAN

To obtain high-resolution wave data for the deep learning model, we choose a wave model that can accurately simulate wave propagation in deep and shallow water. We can use the so-called phase-averaged wave model to simulate wave generation and propagation in large-scale areas, such as the open sea and coastal regions. Here, each individual wave, with its phases, is calculated in an "averaged" way. The most commonly used phase-averaged wave models are wave models, which are called the third-generation wave models. Three models are widely used around the world, i.e., Wave Watch III (WW3) by [17], the WAM model [18], and SWAN [11]. The third-generation wave model describes the propagation of the spatial–temporal wave spectrum using the wave energy density in the conservation balance equation, including the wave generation process by the wind. Among these three models, the SWAN model is the only one specifically designed to capture wave propagation in deep and shallow water, especially for coastal areas. In addition, the model takes into account physical phenomena such as the dissipation by bottom friction, white capping, wave breaking, and nonlinear wave–wave interactions.

As described in [11], the SWAN model is based on the action-balanced equation that can be described briefly in the following equation:

$$\frac{\partial N}{\partial t} + \frac{\partial_{cx} N}{\partial x} + \frac{\partial_{cy} N}{\partial y} + \frac{\partial_{c\sigma} N}{\partial \sigma} + \frac{\partial_{c\theta} N}{\partial \theta} = \frac{S}{\sigma} \tag{1}$$

Here, t represents time, and x and y represent the horizontal and vertical coordinates, respectively. N represents the action density. The first three terms on the left-hand side reflect the N propagation/evolution in time t and the propagation in horizontal x- and y-directions, respectively. The other three terms on the left-hand side denote the relative frequency shift due to the depth and current, and the depth- and current-induced refraction in θ space, respectively. Finally, on the right-hand side, the source term S represents the entire source and sink terms for the governing equation, which is given by the following formula:

$$S = S_{in} + S_{nl3} + S_{nl4} + S_{ds,w} + S_{ds,b} + S_{ds,br} \tag{2}$$

Here, S_{in} is the source term from the wind input, which acts as the primary driving factor of the governing equation. S_{ln3} and S_{ln4} are nonlinear triad (three waves) and quadruplet (four waves) wave interactions in shallow and deep water, respectively. $S_{ds,w}$, $S_{ds,b}$, and $S_{ds,br}$ denote sink terms by dissipation as a result of white capping, bottom roughness, and depth-induced wave breaking, see [11] for a more detailed description.

This paper uses the SWAN model to generate wave data by performing a continuous hindcasting wave simulation from the global domain to a nested local domain using

nested techniques. The wave simulation results are then used as the training data for the machine learning model, i.e., the Long Short-Term Memory (LSTM) and Bidirectional LSTM (BiLSTM). We briefly describe these two deep learning models' basic ideas in the following subsection.

2.3. Long Short-Term Memory

The Long Short-Term Memory, or LSTM, is a modified version of the recurrent neural network (RNN) introduced by Hochreiter and Schmidhuber [19] by adding a memory cell that can store information for a long time. For example, the vanishing gradient is a problem in the RNN model that fails to capture long-term dependencies, thereby reducing the accuracy of a prediction in an RNN [20]. The LSTM solved the vanishing-gradient problem because its architecture could store or discard data because each neuron has several gates regulating each neuron's memory. Figure 3 shows the gates' structure of the LSTM.

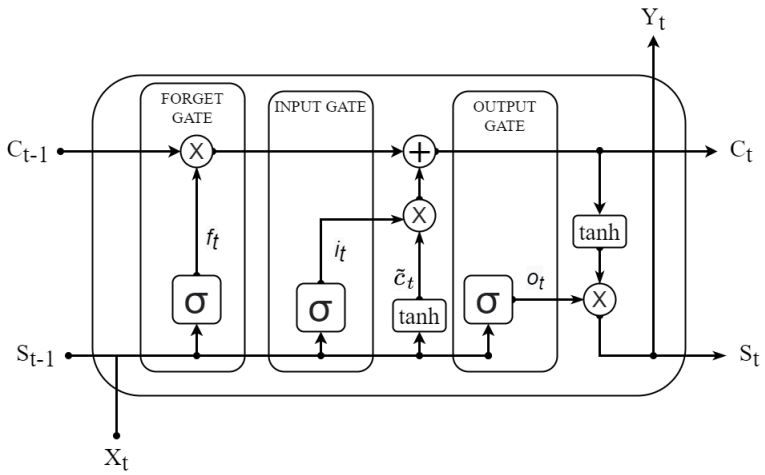

Figure 3. Illustration of Long Short-Term Memory's architecture.

The Long Short-Term Memory, or LSTM, gained its popularity because of its capability to deal with sequential data. The LSTM is used in many applications, including wave forecasting. The LSTM models can capture long-term dependencies in sequence data, making them suitable for a time-series analysis. Because LSTM models can remember information for extended periods of time, they are well-suited for tasks such as wave prediction. The LSTM models are capable of learning complicated data patterns, making them suitable for anomaly detection and forecasting [21].

In the LSTM, there are three gates, namely f_t, i_t, and o_t, as shown in Figure 3. Gate f_t is the forget gate, i_t is the input gate, and o_t is the output gate. The first step in assembling the LSTM is to differentiate the necessary and unnecessary data. A sigmoid function defines this process. This step is followed by saving and updating the data in cells from new inputs. There are two procedures in this step: the sigmoid function, which decides whether new information should be updated or discarded in the numerical value forms 0 and 1, and the tanh function, which assigns a value to each passed data, determines the value of the data in the numbers -1 to 1 [22]. The equations for each gate are given by Equations (4)–(7).

$$i_t = \sigma(W_{ix}X_t + W_{is}S_{t-1} + b_i) \tag{3}$$

$$f_t = \sigma(W_{fx}X_t + W_{fs}S_{t-1} + b_f) \tag{4}$$

$$c_t = f_t \times c_{t-1} + i_t \times \tilde{c}_t \tag{5}$$

$$\tilde{c}_t = \tanh(W_{cx}X_t + W_{cs}S_{t-1} + b_c) \tag{6}$$

$$s_t = o_t \times \tanh(c_t). \tag{7}$$

$$o_t = \sigma(W_{ox}X_t + W_{os}S_{t-1} + b_o) \tag{8}$$

Symbol W_{fx}, W_{fs}, W_{ix}, W_{is}, W_{cx}, W_{cs}, W_{ox}, W_{os} is the weight; b_f, b_i, b_c, b_o is the bias; X_t is the input; S_{t-1} is the previous state; c_t is the cell state or memory cell; and σ is the activation function sigmoid. The forget gate determines the information stored or discarded in the previous state. The input gate regulates how many states the current input passes through. The output gate decides the internal state to forward and the cell state or memory cell to forward old information with additional new information to the next cell state.

2.4. Bidirectional Long Short-Term Memory

The Bidirectional Long Short-Term Memory, or BiLSTM, is a variant of the LSTM developed by Alex Graves and Jurgen Schmidhuber [23]. The fundamental principle of the BiLSTM is that the performance is determined not only by the previous state $(t-1)$ but also by the following state $(t+1)$. This process implies that there are forward and backward states in the BiLSTM, as seen in Figure 4; therefore, the BiLSTM can improve the model's accuracy.

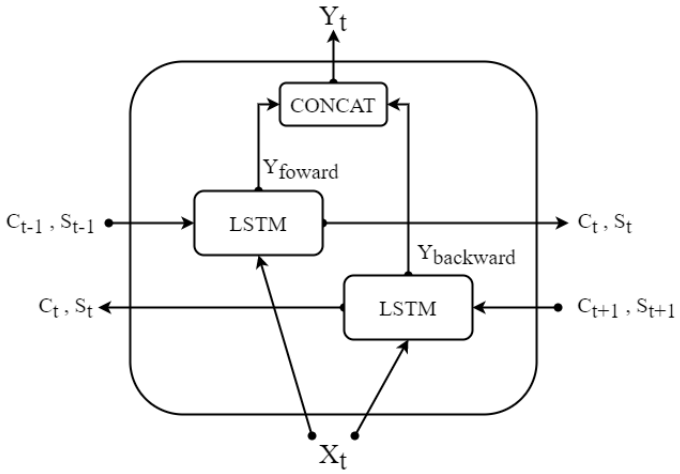

Figure 4. Illustration of the architecture of Bidirectional Long Short-Term Memory.

The advantage of the BiLSTM models is that they capture the long-term dependencies. As the model analyses the sequence in both directions, it may capture both past and future information. This enables the model to discover more intricate correlations between input and output sequences. BiLSTM models are more robust to noise and can detect subtle data patterns. Because the model processes the sequence in both directions, it is less susceptible to noise and outliers. BiLSTM models provide more precision than conventional recurrent neural networks. The model can capture more intricate links between the input and output sequences because it processes the sequence in both ways. This results in increased precision and an enhanced performance [24].

Note that, as a result of these two directions (forward and backward), the computational time for the BiLSTM may be at least twice that of the LSTM. This paper uses these two deep learning models to perform the deep learning-based downscaling for wave forecasting.

3. Methodology

In this research, we aim to create a high-resolution wave forecasting system that is accurate for wave prediction in coastal areas with complex geometry. We propose a machine learning-based procedure to downscale global wave forecasting data to local high-resolution wave data in a coastal area. We first build a high-resolution wave dataset to train the machine learning model by performing nested numerical wave simulations that simulate wave propagation from the global domain to the intermediate and local (high-resolution) domains. The wave dataset obtained is then used to train the machine

learning model. The trained machine learning algorithm is designed to mimic the downscaling process from global wave data to local high-resolution wave data.

In this paper, we use the LSTM and BiLSTM models. In the design of machine learning-based downscaling, we propose to use global wave data as a feature for machine learning to forecast wave conditions in the local coastal area. Here, we solve a supervised learning problem, a regression problem, in which machine learning calculates high-resolution local wave data from global wave data. This paper proposes two main steps to develop machine learning-based downscaling for wave forecasting. The first step is to develop a wave dataset via wave simulation in nested numerical simulations to obtain global wave data and local (high-resolution) wave data. The second step is to use the wave data obtained as features and training data to optimise machine learning for wave forecasting. The details of these two steps are described in the following subsections.

3.1. Wave Data Generation

Here, we build a high-resolution wave dataset by performing high-resolution nested wave simulation using a numerical wave model SWAN. To obtain high-resolution wave simulation in a coastal area, we perform three nested domains that simulate wave propagation from global to intermediate and finally to the local domain. As input for wave simulation, we use the ERA5 wind dataset from the European Centre for Medium-Range Weather Forecasts (ECMWF) with a resolution of 0.25° [5]. Using the SWAN model, we perform continuous wave simulation for 40 years (1980–2020). The flow chart for this step is illustrated in Figure 5.

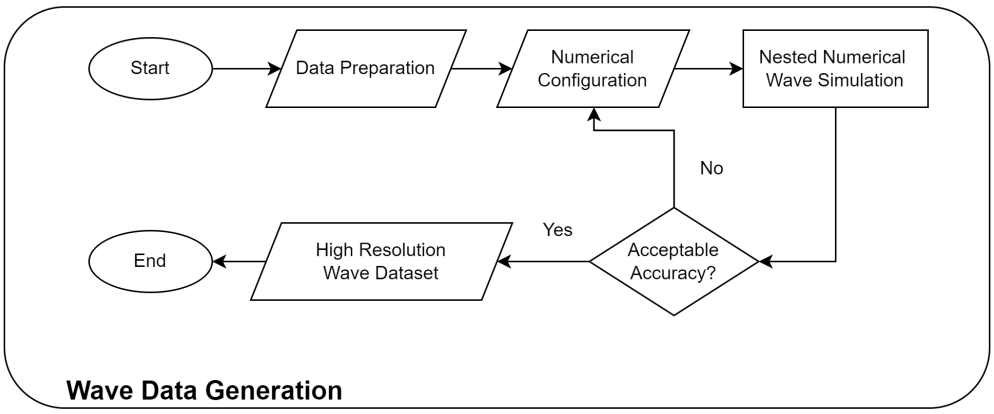

Figure 5. Flowchart of wave data generation. The wave dataset is obtained by performing continuous wave simulation using phase-averaged wave model SWAN.

To test our proposed machine learning-based wave downscaling methodology, we selected two locations of coastal areas with two different wave characteristics. The first location is in Jakarta Bay, Indonesia, a rather closed coastal area in the Java Sea, where wind-generated waves mostly dominate its waves. The other location is offshore of Meulaboh, West Aceh Regency, Indonesia. Here, waves are mostly generated by swell-dominated swells because they are directly connected to the Indian Ocean. For both areas, we describe the numerical setting for the SWAN model and wave data observation to validate the result of the SWAN simulation, as follows.

3.1.1. Data Generation for Jakarta Bay Case

As mentioned previously, wave simulation is performed in three nested domains, namely the global domain (domain I), the intermediate domain (domain II), and the local domain (domain III). This nested simulation aims to capture the wave transformation from the global grid in domain I to the local high-resolution grid in domain III. Simulation from the global domain is also necessary to capture accurate generation and propagation of

swell. The snapshot of the significant wave height Hs resulting from the Jakarta Bay case is shown in Figure 6 for domain I and Figure 7 for domains II and III.

The SWAN model's numerical configuration for the Jakarta Bay simulation is described in Table 1 for domains I, II, and III. Here, Δx and Δy denote the spatial grid in the horizontal and vertical directions, respectively. Nx and Ny represent the number of grid partitions in the horizontal and vertical directions, respectively. Note that we set the spatial grid size of domain II the same as that of global wave forecasting, such as GFS by NOAA and ERA5 by ECMWF. For domain III, we set the spatial grid size to $0.0027^\circ \approx 299$ m, which means we perform downscaling from 0.25° to 0.0027° or 92.5 times.

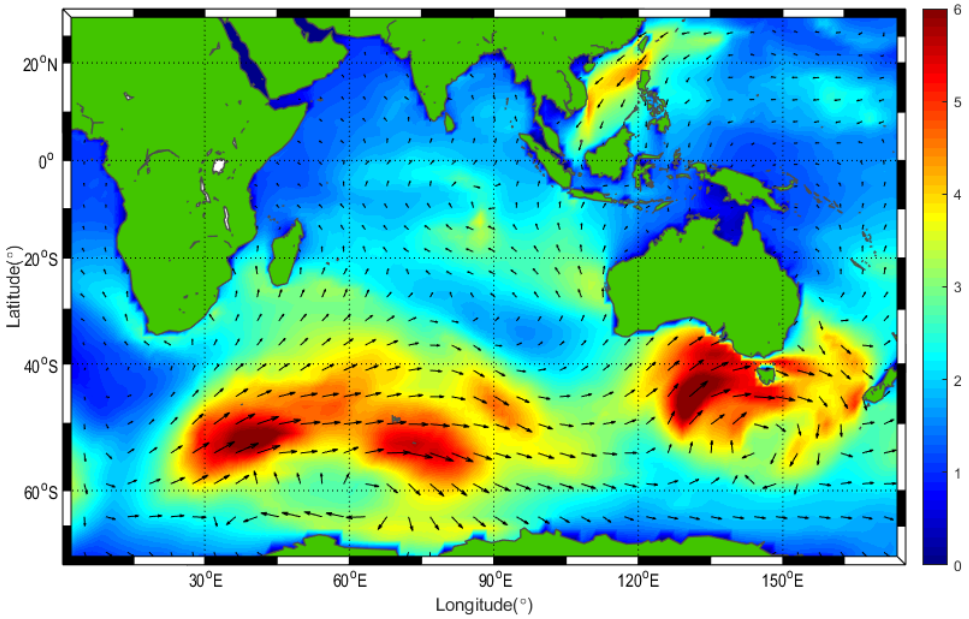

Figure 6. Snapshot of significant wave height on 6 December 2020, at 06:00 UTC, from SWAN simulation in domain I.

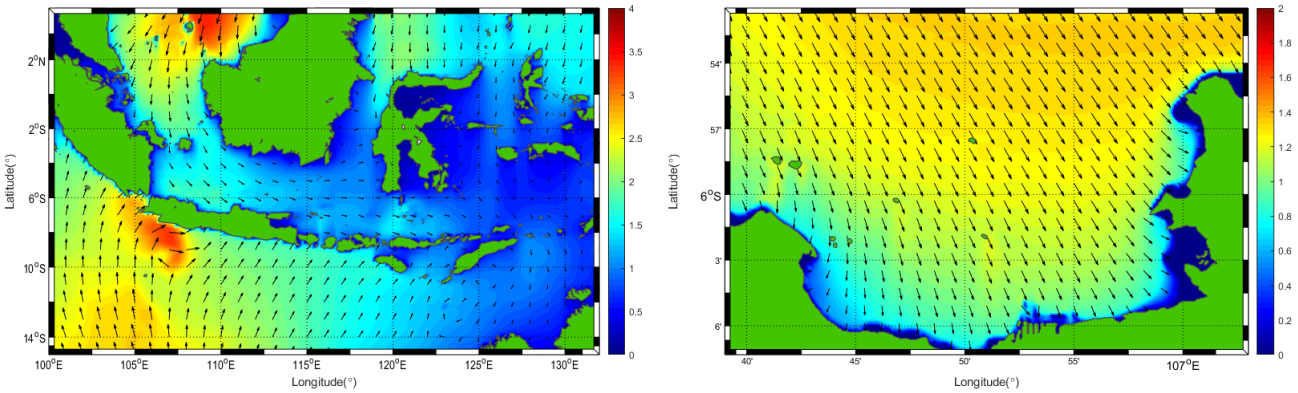

Figure 7. As in Figure 6, for domains II (**left plot**) and III (**right plot**) for Jakarta Bay area.

Table 1. Numerical configuration for SWAN model for the Jakarta Bay case.

Domain	Lon (°)		Lat (°)		Δx	Δy	Nx	Ny
	West	East	South	North				
1	0.5	175.5	−69.5	30.5	1.4957	1.4925	117	67
2	100	132	−15	5	0.25	0.25	128	80
3	106.65	107.05	−6.122	−5.858	0.0027	0.0027	150	99

3.1.2. Data Generation for Meulaboh Case

We chose the second study area offshore of Meulaboh, West Aceh Regency, Indonesia. Here, the waves are mainly dominated by swells generated from the Indian Ocean. As in the Jakarta Bay case, we perform nested simulations in three domains. Domain I, or global domain, is as in the case of Jakarta Bay, as shown in Figure 6. Domains II and III of Meulaboh's case are shown in Figure 8.

The SWAN model's numerical configuration for the simulation in the Meulaboh is described in Table 2, for domains I, II, and III. As in the Jakarta Bay case, we also set the spatial grid size of domain II the same as the grid size of global wave forecasting GFS and ERA5. For domain III, we set the spatial grid size to be $0.002^\circ \approx 222$ m, which means that we perform downscaling from 0.25° to 0.002° or 125 times finer grid.

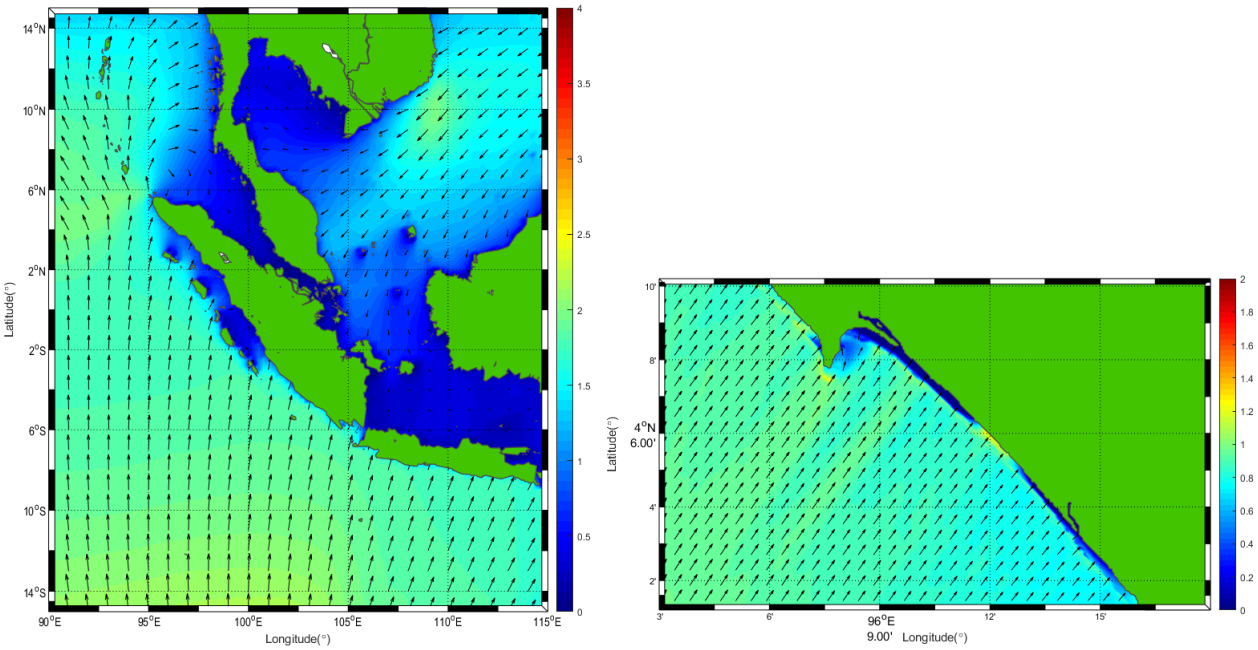

Figure 8. Snapshot of significant wave height on 1 March 2020, at 00:00 UTC, from wave simulation using SWAN model for domains II (**left plot**) and III (**right plot**) for Meulaboh area.

Table 2. Numerical configuration for the SWAN model for the Meulaboh case.

Domain	Lon (°)		Lat (°)		Δx	Δy	Nx	Ny
	West	East	South	North				
1	0.5	175.5	−69.5	30.5	1.4957	1.4925	117	67
2	85	107	−14	14	0.25	0.25	88	112
3	96	96.25	4	4.15	0.002	0.002	125	75

3.2. Deep Learning Approach for Wave Downscaling

In this second main step, we design machine learning based on the downscaling process that calculates local (high-resolution) wave data from global (low-resolution) wave datasets. In this step, we used the calculated high-resolution wave dataset obtained in the previous subsection as feature and training data for machine learning models. The main aim of machine learning here is to mimic the downscaling process performed by numerical simulation, especially from domain II to domain III. In this case, the machine learning algorithm solves the supervised problem of performing regression from a global to a local (high-resolution) grid.

In this paper, we use two machine learning models, i.e., the Long Short-Term Memory (LSTM) and Bidirectional LSTM (BiLSTM). These two models are chosen because they perform better for sequential data, such as time-series data, as in this research. The procedure

for this second main step is illustrated in the flow chart in Figure 9. The wave dataset we obtained from the previous step analyses the best feature of the machine learning input. In Figure 9, we perform feature engineering by selecting the best locations of global wave data that are highly correlated with the targeted wave data in the local domain. To this end, we calculate the correlation coefficient (CC) between each global wave data grid location and with corresponding target wave data in the local domain. The machine learning algorithm will take only the global location of the grid with a high CC value as input.

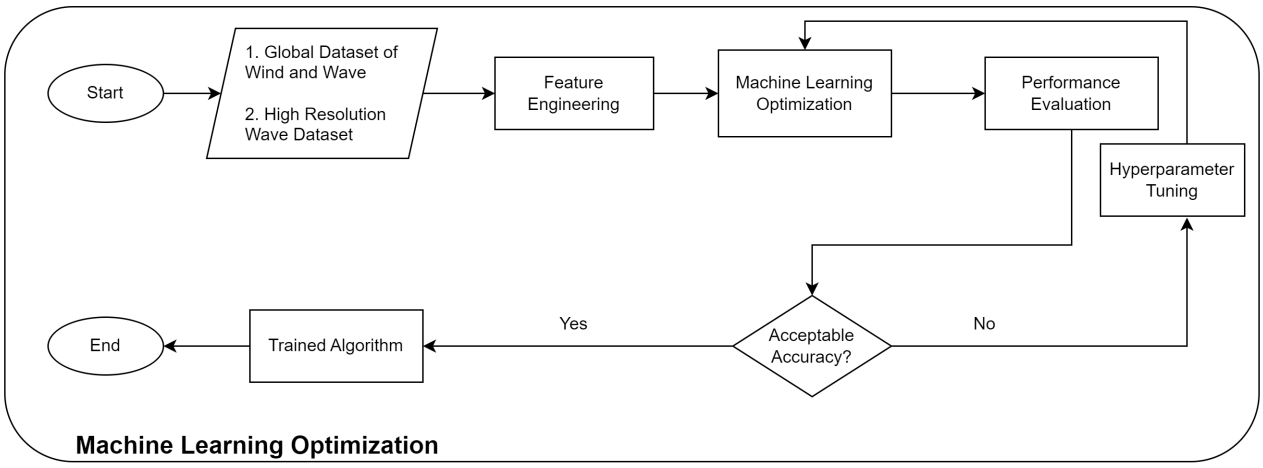

Figure 9. Flowchart of machine learning optimisation. The wave dataset from the previous step is used as training data for machine learning.

The correlation coefficient (*CC*) is defined as follows:

$$CC = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2}\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \overline{y})^2}} \tag{9}$$

where n denotes the amount of data to be compared, x_i is the values of the first variable, and $\overline{x}$ is the average of the values of the first variable. At the same time, y_i is the value of the second variable, and $\overline{y}$ is the average of the values of the second variable. The CC evaluates how well the predicted values correspond to the actual values.

We also perform hyperparameter tuning in optimising machine learning models to obtain the best results. The performance evaluation of both machine learning models is evaluated by calculating the CC, the root mean square error (*RMSE*), and the mean averaged percent error (*MAPE*). The RMSE is the square root of the average squared error between observed x_i and predicted values y_i while the MAPE is the mean of absolute percentage deviations between predicted and observed values, which are defined as follows:

$$RMSE = \left(\frac{1}{n}\sum_{i=1}^{n}(x_i - y_i)^2\right)^2, \tag{10}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{x_i - y_i}{x_i}\right|. \tag{11}$$

4. Results and Discussion

As mentioned in the previous section, we study the implementation of our proposed deep learning-based downscaling for wave forecasting in two locations, i.e., in Jakarta Bay and the Meulaboh offshore area. In this section, we first discuss the wave prediction results in Jakarta Bay, followed by the prediction results in the offshore Meulaboh region. For each case, as illustrated in Figure 9, we performed the feature selection step, selecting the best feature to be taken as the input for the machine learning models.

4.1. Wave Downscaling in Jakarta Bay

Jakarta Bay is located in the Java Sea, a relatively closed sea surrounded by the Java, Sumatra, and Borneo islands. Wind waves mostly dominate the location. As described in Section 3.1.1, we generate the wave datasets by performing the simulations in nested domains. We perform continuous hindcasting simulations for the Jakarta Bay case for 20 years (2001–2020).

Note that, as described in Section 3.1.1, in Table 1, the spatial grid resolution of domain II is 0.25°, which is chosen to be the same as the global grid wave forecasting, such as the GFS by the NOAA and ERA5 by the ECMWF. Our proposed method here aims to mimic the downscaling process by numerical simulation into a deep learning algorithm. In this case, the downscaling approach is applied from domain II to the location domain, or domain III.

To perform the downscaling, we first select which locations of the global wave data (domain II) are to be chosen as the input to predict the wave-targeted location in the local domain (domain III). We calculate the spatially correlated wave data in domain II with a target wave location in domain III. This step is performed by calculating the correlation coefficient (CC) between the significant wave height (Hs) at each grid in domain II and the Hs at a targeted location in the local domain.

For the Jakarta Bay case, we chose a target wave location located in the western part of Jakarta Bay, 106.7654° E, 6.0108° S, as shown in Figure 10. We also have a wave observation at this location that measures the significant wave height. As a part of the feature selection process, we select the locations of the global significant wave height Hs data in domain II that have a high correlation with the Hs in the observation data by using a spatial correlation as introduced in [10]. The spatial correlation here is simply by calculating the correlation coefficient (CC) defined in Equation (9) between the Hs at all the locations in domain II and the targeted Hs at domain III. The locations with high CC values correlate highly with the targeted Hs in the local domain.

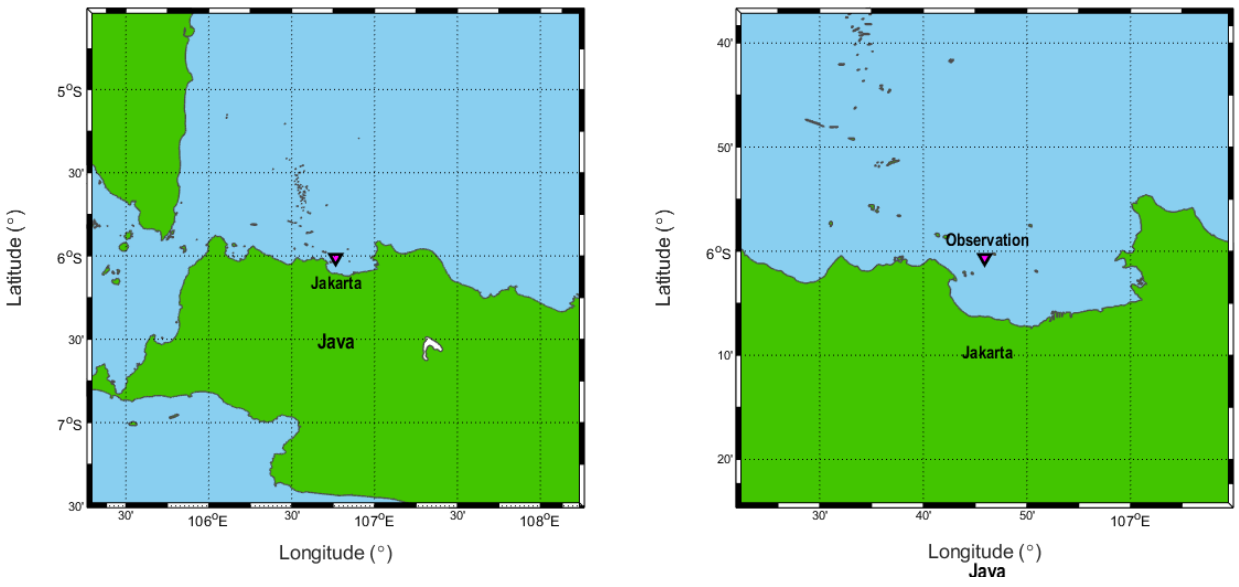

Figure 10. Location of wave observation at Jakarta Bay.

The spatial correlation results between the Hs from the global wave data (or domain II) with the Hs at the observation location in Figure 10 are shown in Figure 11. We use one-year Hs time-series data to calculate the spatial correlation map. On the left plot of Figure 11 is the spatial correlation map, with the big dots denoting the CC values ≥ 0.70, while the right plot is for the CC values ≥ 0.80, and the lower plot for the CC values ≥ 0.90. For the case of Jakarta Bay, it turns out that only six locations of the global wave data grid have a correlation coefficient value ≥ 0.90.

Using the spatial correlation map results as shown in Figure 11, we performed a sensitivity test to determine which spatial correlation map (with corresponding locations) gives the best result. In Table 3, we show the results of the downscaling performance of the BiLSTM model with the selected CC in the spatial correlation, in terms of the correlation coefficient and RMSE, for predicting 14 days. Here, the spatial correlation with a CC $\geq$ 0.90 gives the best results, with a CC of 0.87 and an RMSE of 0.07. The spatial correlation with a CC $\geq$ 0.90 gives only six point locations in the global wave data as input for the BiLSTM model. This number is significantly small compared to the spatial correlation with a CC $\geq$ 0.80, which has 23 points. From this table, we conclude that the downscaling model with the spatial correlations of a CC $\geq$ 0.90 gives the best result.

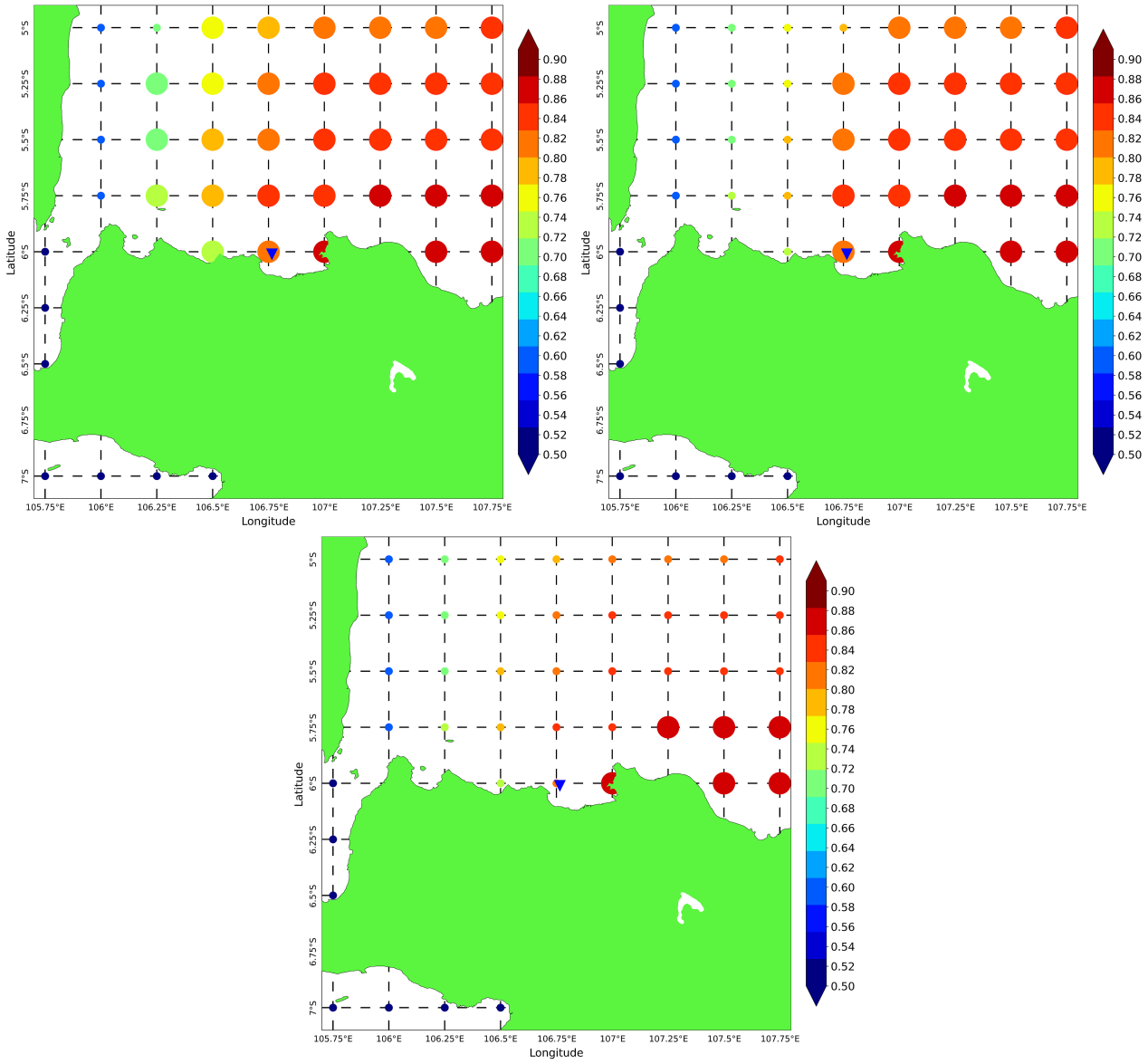

Figure 11. The spatial correlation map at Jakarta Bay was obtained by calculating the correlation coefficient (CC) between Hs at the global grid and Hs at the targeted local domain. Big dots denote CC values: upper left plot for CC values $\geq$0.70, upper right plot for $\geq$0.80, and lower plot for $\geq$0.90.

Table 3. Comparison between selected spatial correlation with results of downscaling performance for prediction 14 days ahead in Jakarta Bay area.

Area	Selected Spatial Correlation	Number of Wave Point Input	CC	RMSE
Jakarta Bay	CC > 0.70	32	0.84	0.07
	CC > 0.80	23	0.85	0.08
	CC > 0.90	6	0.87	0.07

Based on the spatial correlation map in Figure 10 and Table 3, we use the spatial correlation map with a CC > 0.90 to perform the comparison test with the observation data in Jakarta Bay. We use six grid locations in the global wave data as the input for the machine learning models. We compare the results of our deep learning-based downscaling with the wave observation in Figure 10. A qualitative comparison of the significant wave height of the wave observation with the prediction using the BiLSTM is shown in Figure 12, indicating a relatively good agreement. From the figure, the results from the BiLSTM model provide a relatively close prediction before 26 January. However, after that, it remains steady at approximately 0.5 m, whereas the observation data are slightly increasing. Moreover, it is also noticeable that the observation data oscillate for 12 h, or half a day. We suspect that is a frequency that is tidal related that cannot be captured by the BiLSTM model with input only from the global wave data.

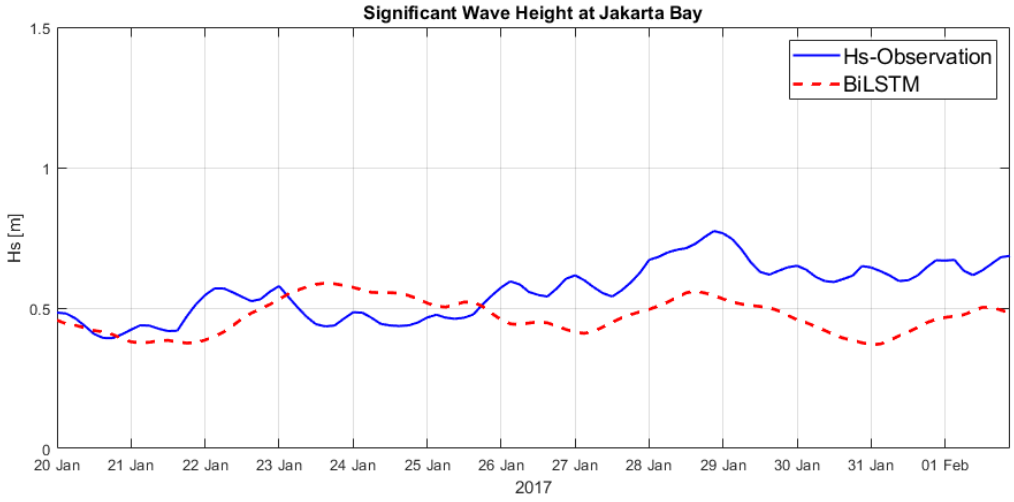

Figure 12. Comparison of significant wave height from wave observation with result of prediction by using BiLSTM at Jakarta Bay.

In Table 4, we compare the error quantitatively, in terms of the RMSE, between the Hs from the observation with the results of the NOAA GFS Forecast, downscaling using the LSTM and the BiLSTM. The prediction using the BiLSTM shows the best result. From the table, there are improvements in the accuracy from the downscaling method by the BiLSTM model of 26.3% compared to the global wave model, the GFS Forecast.

Table 4. Comparison between the significant wave height Hs from wave observation at Jakarta Bay with the results of GFS Forecast, downscaling using LSTM and BiLSTM.

Model	RMSE
LSTM	0.15
BiLSTM	0.14
GFS Forecast	0.19

4.2. Wave Downscaling in Meulaboh

Meulaboh is located in the West Aceh Region on Sumatra Island, Indonesia. The offshore of Meulaboh, located at 96.1831° E, 4.1019° N, directly faces an open sea, the Indian Ocean. Therefore, the offshore of Meulaboh is dominated by swells generated from the Indian Ocean. In this subsection, we test the validity of our proposed downscaling method against the wave observation in the Meulaboh offshore.

As in the Jakarta Bay case, we perform a continuous wave simulation in nested domains as described in Section 3.1.2. For this case, we perform 40 years of continuous wave simulations (1981–2020). Just as in the Jakarta Bay case, in Table 2, the spatial grid resolution of domain II is 0.25°, which is chosen to be the same as the global grid wave forecasting, such as the GFS by the NOAA and ERA5 by the ECMWF. In this case, we also aim to decrease the downscaling process by the machine learning algorithm by downscaling from domain II (with a grid size of 0.25°) to domain III, with a grid size of 0.002°.

After performing a continuous hindcasting wave simulation for 40 years, we built a spatial correlation map, as we did in the Jakarta Bay case. To this aim, we calculate the spatially correlated wave data in domain II of Meulaboh with a targeted wave location in domain III. For the Meulaboh case, we chose a target wave location offshore of Meulaboh, located at 96.1831° E, 4.1019° N. Moreover, at this location, we have wave observation data to compare with the results of the wave downscaling. The location of the observation point at Meulaboh is shown in Figure 13.

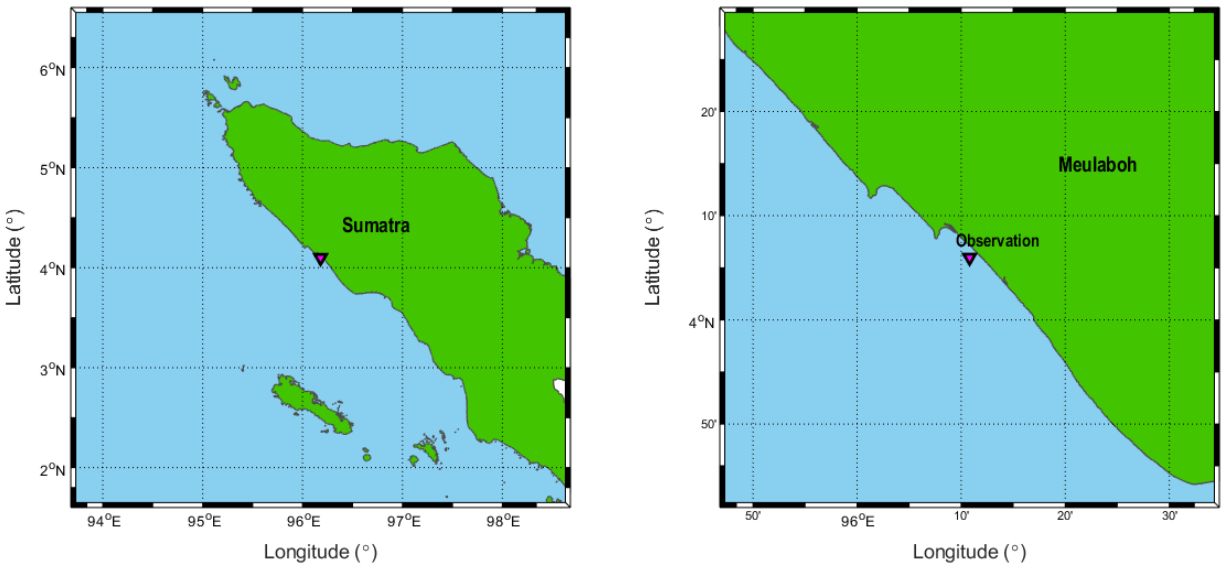

Figure 13. Location of wave observation at Meulaboh, West Aceh Regency, Indonesia.

The spatial correlation results for the Meulaboh case, between the Hs from the global wave data (or domain II) with the Hs at the observation location in Figure 13, are shown in Figure 14. We use one-year (during 2020) Hs time-series data to calculate the spatial correlation map. On the upper left plot of Figure 14 is the spatial correlation maps, where the big dots represent the CC values ≥ 0.70, while on the upper right plot is for the CC values ≥ 0.80 and in the lower plot for the CC values ≥ 0.90. Just as in the Jakarta Bay case, by using the results of the spatial correlation map as shown in Figure 14, we performed a sensitivity test to determine which spatial correlation map (with corresponding locations) gives the best result. Here, we performed the downscaling by using the BiLSTM to calculate a 14-days prediction.

In Table 5, we show the results of the downscaling performance of the BiLSTM model with the selected CC in the spatial correlation, in terms of the correlation coefficient and RMSE, for predicting 14 days. As expected, the spatial correlation with a CC ≥ 0.90 gives the best results, with a CC of 0.97 and an RMSE of 0.15. The spatial correlation with a

CC $\geq$ 0.90 gives only four point locations in the global wave data as the input for the BiLSTM model. As in the Jakarta bay case, this number is significantly small compared to the spatial correlation with a CC $\geq$ 0.80, which has 50 points. From this table, we conclude that the downscaling model with the spatial correlations of a CC $\geq$ 0.90 gives the best result.

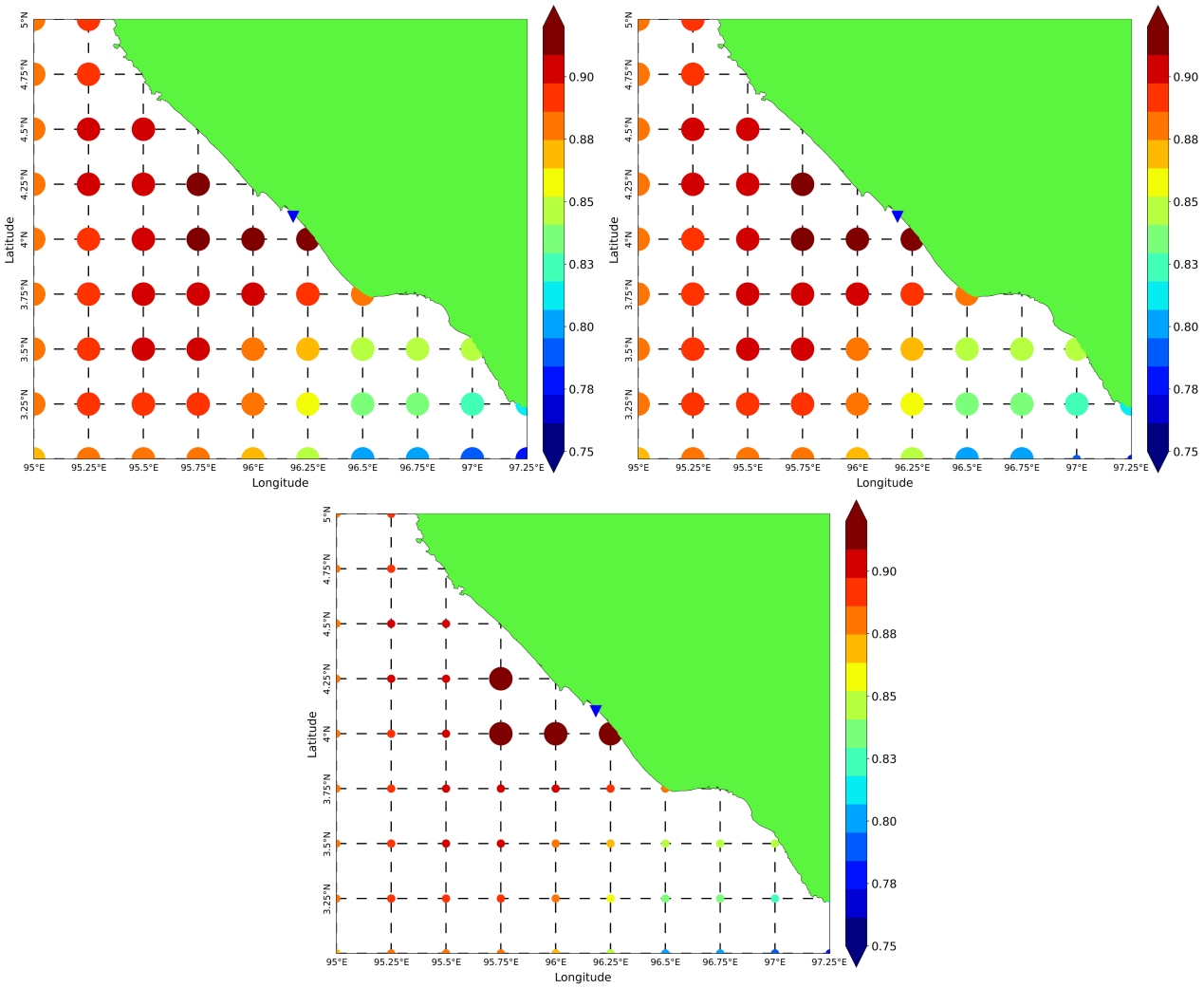

Figure 14. Spatial correlation maps at Meulaboh offshore, obtained by calculating the correlation coefficient (CC) between Hs at the global grid with Hs at a targeted local domain. Big dots denote CC values: in the upper plot for CC values $\geq$ 0.70, the lower plot for CC values $\geq$ 0.80, and the lower plot for CC values $\geq$ 0.90.

Table 5. Comparison between selected spatial correlation with results of downscaling performance for prediction 14 days ahead in Meulaboh area.

Area	Selected Spatial Correlation	Number of Wave Point Input	CC	RMSE
Meulaboh	CC > 0.70	52	0.95	0.16
	CC > 0.80	50	0.96	0.16
	CC > 0.90	4	0.97	0.15

For the case of Meulaboh, based on Figure 14 and Table 5, we use only four locations in the global wave data grid, which have correlation coefficient values $\geq$ 0.90, as the features for our deep learning model. Before we compare the simulation results with the wave observation, we will investigate the sensitivity of the training data length with the resulting

downscaling accuracy. Moreover, we also investigate the sensitivity of the downscaling prediction length with the resulting downscaling accuracy.

4.2.1. Sensitivity of Length of Training Data

In the Meulaboh case, we obtain 40 years of continuous hindcasting wave simulations for the dataset to train the machine learning algorithm. Indeed, for a practical application, this may not be efficient for the designer of a wave forecasting system that will use this methodology. In this subsection, we investigate the sensitivity of the training data with the resulting downscaling accuracy. Here, we set several scenarios of training data lengths, i.e., 1, 5, 10, 15, 20, 30, and 40 years, to test the LSTM and BiLSTM models' performances in downscaling. In Table 6, we show the downscaling results using the LSTM and BiLSTM to perform the downscaling 14 days ahead. In general, the BiLSTM performs better than the LSTM. For the length of the training data, there are two best results, namely 15 and 40 years, which give good and relatively similar results, with a correlation coefficient of 0.97 and an RMSE of 0.15. Furthermore, the training data over 15 years does not significantly increase the downscaling results.

Table 6. Sensitivity of the training data length with the accuracy of the prediction using LSTM and BiLSTM.

Length (Year)	LSTM			BILSTM		
	CC	RMSE	MAPE	CC	RMSE	MAPE
1	0.91	0.22	15.06	0.93	0.21	14.53
5	0.95	0.17	13.77	0.95	0.19	13.70
10	0.96	0.18	12.9	0.96	0.17	12.11
15	0.97	0.17	12.41	0.97	0.16	11.79
20	0.96	0.18	12.23	0.97	0.17	12.46
25	0.96	0.18	12.13	0.97	0.17	11.44
30	0.96	0.16	11.29	0.97	0.16	10.65
40	0.97	0.17	11.8	0.97	0.15	11.40

4.2.2. Sensitivity of Length of Downscaling

In this subsection, we investigate the sensitivity of the length of the downscaling. The global wave forecasting system GFS by the NOAA produces up to 16 days of prediction, which is available four times a day, or every 6 h. In forecasting, the length of prediction indeed affects the prediction's accuracy. In this paper, we performed downscaling from global wave forecasting data to local wave data. To see the performance of the resulting downscaling, we propose scenarios for a downscaling length of 1, 3, 5, 7, and 14 days. In Table 7, we show the results using the BiLSTM to calculate the downscaling durations of 1, 3, 5, 7, and 14 days in terms of the correlation coefficient (CC), RMSE, and MAPE. The table shows that the longer prediction horizon results in a lower accuracy. Note that the one-day downscaling gives an almost perfect prediction with a correlation coefficient of 0.99 and an RMSE value of 0.01.

Table 7. Sensitivity of the downscaling prediction length with the resulting accuracy of downscaling.

Downscaling Length (Day)	CC	RMSE	MAPE
1	0.99	0.01	0.72
3	0.99	0.06	8.46
5	0.99	0.09	13.42
7	0.99	0.08	10.47
14	0.97	0.16	11.79

4.2.3. Comparison with Wave Observation

To validate our deep learning-based downscaling model, we compare the result of the downscaling by using the LSTM and BiLSTM with the wave observation data in the Meulaboh offshore area. To perform the downscaling by using deep learning, we use the global GFS forecast wave data as an input for our deep learning-based downscaling models, i.e., the LSTM and BiLSTM. As mentioned, the wave observation is located on the west coast of Meulaboh, as shown in Figure 13. The wave observation data are available for 30 days. We compare the Hs from the observation with the results of the Hs prediction using the BiLSM model as shown in Figure 15. As shown in the figure, the results of the BiLSTM downscaling can follow relatively accurate observation data. In Table 8, we compare the quantitative errors, in terms of the RMSE, between the Hs from the observation with the global forecast model, the GFS forecast model by the NOAA, and the results of the prediction obtained by using the LSTM and BiLSTM models. Here, it shows that the BiLSTM gives the best results compared to the LSTM and GFS forecast (global wave forecasting model), even for a wave prediction of 30 days.

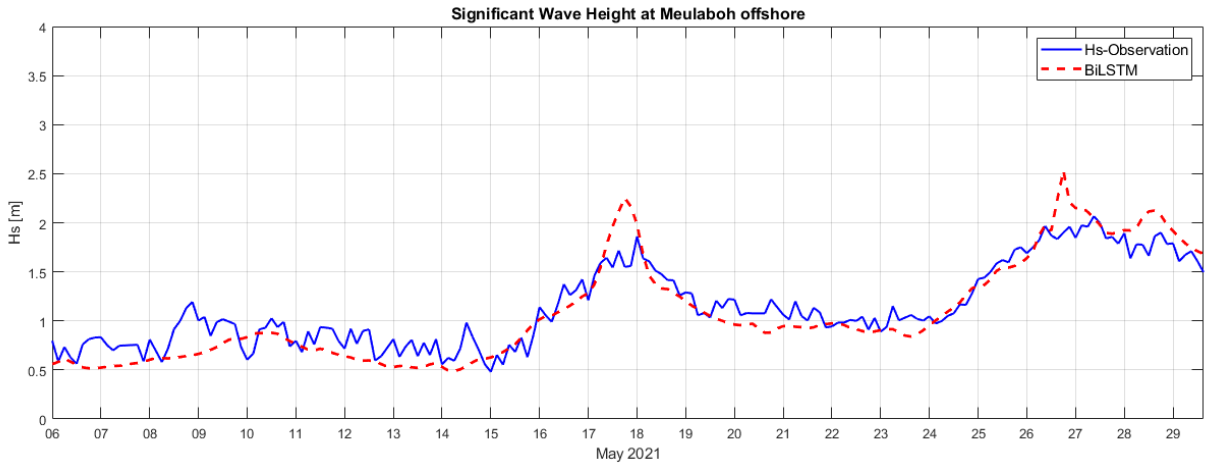

Figure 15. Comparison of significant wave height from wave observation with result of prediction by using BiLSTM at offshore of Meulaboh.

Table 8. Comparison between the significant wave height Hs from wave observation at Meulaboh with GFS forecast results, downscaling using LSTM and BiLSTM.

Model	RMSE
LSTM	0.19
BiLSTM	0.16
GFS Forecast	0.23

5. Conclusions

In coastal areas, especially with complex geometry, wave prediction requires a high-resolution grid simulation to capture the wave propagation accurately. Global wave forecasters, such as the GFS by the NOAA and ERA5 by the ECMWF, provide relatively low-resolution wave forecasting. This paper proposes an alternative way to perform downscaling using a deep learning approach, especially for wave forecasting. We approach this idea by constructing a high-resolution wave dataset using a numerical simulation with the SWAN model in nested simulations. As input for the deep learning models to perform the downscaling, the selection of grid locations of the global wave data to be downscaled is made by calculating the spatially correlated wave data in the global wave data with a targeted wave location. Here, machine learning takes only highly correlated grid locations as the input. In this paper, we study two coastal locations, i.e., Jakarta Bay and the Meulaboh offshore area. Based on these two study areas, the BiLSTM performed better than the LSTM.

From the Meulaboh case, we found that the length of the training data affects the accuracy of the downscaling prediction, especially for the length of training data of 1 to 15 years. Increasing the length of training by more than 15 years only slightly increases the accuracy of the prediction. We also investigated whether the downscaling prediction horizon for the downscaling greatly affects the accuracy of the prediction. To predict 14 days, the BiLSTM results in an accuracy with a CC value of 0.97, an RMSE value of 0.16, and a MAPE value of 11.79, while for predicting 1 day, the model can produce almost perfect accuracy, with a CC value of 0.99, an RMSE of 0.01, and a MAPE value of 0.72. The comparison between the wave observation with the prediction results using the BiLSTM in Jakarta Bay and Meulaboh's offshore shows that our deep learning-based downscaling improves the prediction of the global wave forecasting GFS.

We conclude that this paper's proposed deep learning-based downscaling method is more suitable for downscaling waves in an open area, such as in Meulaboh, than in a closed area, such as Jakarta Bay. By analysing the results from the comparison with the observation data, especially in the Jakarta Bay area, the global wave prediction such as the GFS forecast by the NOAA may not be accurate because the area is a relatively closed area with a wind-wave-dominated area. As a result, because the downscaling model takes the GFS forecast as input, the downscaling may inherit the errors from the global wave model. For the Jakarta case, there are improvements in the accuracy from the downscaling method of 26.3% compared to the global wave model. From a comparison with the observation data in the Meulaboh case, the results of the BiLSTM overestimated the wave peaks of the observation data. In contrast, at other times, the downscaling results can follow the observation data relatively well. These overestimated peaks may also be inherited from the global wave data input. Here, there are two possible errors, i.e., the errors from the GFS forecast that are taken as the input from the deep learning model and possible errors from the SWAN model. Nevertheless, the results of the prediction using the BiLSTM show relatively good results. For the case of Meulaboh, the BiLSTM results in a 30.4% improvement over the GFS in the RMSE.

Author Contributions: Conceptualisation, D.A., S.R.P. and D.T.; methodology, D.A. and D.T.; software, D.A.; validation, S.H. and G.P.; formal analysis, D.S.; writing—original draft preparation, D.A., S.R.P. and D.S.; writing—review and editing, A.S., G.P. and S.H.; funding acquisition, D.A. and G.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Kementerian Pendidikan, Kebudayaan Riset dan Teknologi, Republik Indonesia, research grant number 019/SP2H/RT-JAMAK/LL4/2022 and 369/PNLT3/PPM/2022.

Data Availability Statement: The data of the compounds are available from the authors.

Conflicts of Interest: The authors declare that there are no conflict of interest.

Sample Availability: The samples of the compounds, etc., are available from the authors.

References

1. Wu, M.; Stefanakos, C.; Gao, Z. Multi-step-ahead forecasting of wave conditions based on a physics-based machine learning (PBML) model for marine operations. *J. Mar. Sci. Eng.* **2020**, *8*, 992. [CrossRef]
2. Chen, D.; Liu, F.; Zhang, Z.; Lu, X.; Li, Z. Significant wave height prediction based on wavelet graph neural network. In Proceedings of the 2021 IEEE 4th International Conference on Big Data and Artificial Intelligence (BDAI), Qingdao, China, 2–4 July 2021; pp. 80–85. [CrossRef]
3. Tür, R. Maximum wave height hindcasting using ensemble linear-nonlinear models. *Theor. Appl. Climatol.* **2020**, *141*, 1151–1163. [CrossRef]
4. Rutledge, G.K.; Alpert, J.; Ebisuzaki, W. NOMADS: A climate and weather model archive at the National Oceanic and Atmospheric Administration. *Bull. Am. Meteorol. Soc.* **2006**, *87*, 327–342. [CrossRef]
5. Hersbach, H.; Bell, B.; Berrisford, P.; Hirahara, S.; Horányi, A.; Muñoz-Sabater, J.; Nicolas, J.; Peubey, C.; Radu, R.; Schepers, D.; et al. The ERA5 global reanalysis. *Q. J. R. Meteorol. Soc.* **2020**, *146*, 1999–2049. [CrossRef]
6. Breivik, Ø.; Gusdal, Y.; Furevik, B.R.; Aarnes, O.J.; Reistad, M. Nearshore wave forecasting and hindcasting by dynamical and statistical downscaling, *J. Mar. Syst.* **2009**, *78*, S235–S243. [CrossRef]

7. Camus, P., Herrera, S.; Gutiérrez, J.M.; Losada, I.J. Statistical downscaling of seasonal wave forecasts. *Ocean. Model.* **2019**, *138*, 1–12. [CrossRef]
8. Chen, J.; Pillai, A.C.; Johanning, L; Ashton, I. Using machine learning to derive spatial wave data: A case study for a marine energy site. *Environ. Model. Softw.* **2021**, *142*, 105066. [CrossRef]
9. Michel, M.; Obakrim, S.; Raillard, N.; Ailliot, P; Monbet, V. Deep learning for statistical downscaling of sea states. *Adv. Stat. Climatol. Meteorol. Oceanogr.* **2022**, *8*, 83–95. [CrossRef]
10. Adytia, D.; Saepudin, D.; Pudjaprasetya, S.R.; Husrin, S.; Sopaheluwakan, A. A Deep Learning Approach for Wave Forecasting Based on a Spatially Correlated Wind Feature, with a Case Study in the Java Sea, Indonesia. *Fluids* **2022**, *7*, 39. [CrossRef]
11. Ris, RC.; Holthuijsen, LH.; Booij, N. A third-generation wave model for coastal regions: 2. Verification. *J. Geophys. Res. Ocean.* **1999**, *104*, 7667–7681. [CrossRef]
12. Maraun, D.; Widmann, M. *Statistical Downscaling and Bias Correction for Climate Research*; Cambridge University Press: Cambridge, UK, 2018.
13. Willard, J.; Jia, X.; Xu, S.; Steinbach, M.; Kumar, V. Integrating scientific knowledge with machine learning for engineering and environmental systems. *Acm Comput. Surv. (CSUR)* **2021**, *55*, 1–37. [CrossRef]
14. Sharifi, E.; aghafian, B.; Steinacker, R. Downscaling satellite precipitation estimates with multiple linear regression, artificial neural networks, and spline interpolation techniques. *J. Geophys. Res. Atmos.* **2019**, *124*, 789–805. [CrossRef]
15. Srivastava, P.K.; Han, D.; Ramirez, M.R.; Islam, T. Machine learning techniques for downscaling SMOS satellite soil moisture using MODIS land surface temperature for hydrological application. *Water Resour. Manag.* **2013**, *27*, 3127–3144. [CrossRef]
16. Wang, J.; Liu, Z.; Foster, I.; Chang, W.; Kettimuthu, R.; Kotamarthi, V. R. Fast and accurate learned multiresolution dynamical downscaling for precipitation. *Geosci. Model Dev.* **2021**, *14*, 6355–6372. [CrossRef]
17. Tolman, H.L. A third-generation model for wind waves on slowly varying, unsteady, and inhomogeneous depths and currents. *J. Phys. Oceanogr.* **1991**, *21*, 782–797. [CrossRef]
18. The WAMDI Group. The WAM model—A third generation ocean wave prediction model. *J. Phys. Oceanogr.* **1988**, *18*, 1775–1810. [CrossRef]
19. Hochreiter, S.; Schmidhuber, J. Long Short-Term Memory. *Neural Comput.* **1997**, *9*, 1735–1780. [CrossRef]
20. Zhao, Z.; Chen, W.; Wu, X.; Chen, P.C.Y.; Liu, J. LSTM network: A deep learning approach for short-term traffic forecast. *Iet Intell. Transp. Syst.* **2017**, *11*, 68–75. [CrossRef]
21. Song, X.; Liu, Y.; Xue, L.; Wang, J.; Zhang, J.; Wang, J.; Jiang, L; Cheng, Z. Time-series well performance prediction based on Long Short-Term Memory (LSTM) neural network model. *J. Pet. Sci. Eng.* **2020**, *186*, 106682. [CrossRef]
22. Le X.H.; Ho, H.V.; Lee, G.; Jung, S. Application of long short-term memory (LSTM) neural network for flood forecasting. *Water* **2019**, *11*, 1387. [CrossRef]
23. Graves, A.; Schmidhuber, J. Framewise phoneme classification with bidirectional LSTM and other neural network architectures. *Neural Netw.* **2005**, *18*, 602–610. [CrossRef] [PubMed]
24. Peng, T.; Zhang, C.; Zhou, J.; Nazir, M.S. An integrated framework of Bi-directional long-short term memory (BiLSTM) based on sine cosine algorithm for hourly solar radiation forecasting. *Energy* **2021**, *221*, 119887. [CrossRef]