A Machine-Learning Model for Forecasting Internal Wave Propagation in the Andaman Sea

Xudong Zhang, *Member, IEEE*, Xiaofeng Li, *Fellow, IEEE*, and Quanan Zheng

***Abstract*—Internal waves (IWs) are broadly distributed globally and have significant impacts on offshore engineering and underwater navigation. The prediction of IW propagation is a challenging task because of the complex factors involved. In this study, a machine-learning model was developed to predict IW propagation in the Andaman Sea. The model is based on a back-propagation neural network trained by 1189 IW samples, including the crest length and the peak-to-peak distance of IWs, extracted from 123 Moderate-Resolution Imaging Spectroradiometer (MODIS) images and 33 Ocean Land Color Instrument (OLCI) images acquired from 2015 to 2019 and corresponding ocean environment parameters. Using the leading wave crest within an IW packet as input, we ran the model to forecast the IW locations and compare them with satellite observations. The average root-mean-square difference between the model-forecasted and satellite-observed IW leading crest after one tidal cycle was 3.21 km. The corresponding averaged correlation coefficient was 0.95 and the average Fréchet Distance was 11.46 km. We reiterated the model run over two tidal periods and obtained similar statistical results, indicating the robustness of forecasting IW packets. We find that reducing the time step helped to improve forecasting accuracy. The influence of input errors and seasonal variations on model results are discussed and an analysis shows that the initial propagation direction introduced to the model is necessary for cross-propagating IW patterns. Comparisons with the Korteweg-de Vries equation results show that the developed model has better performance and is more robust.**

***Index Terms*—Andaman sea, internal wave (IW) forecast, machine learning.**

I. Introduction

INTERNAL waves (IWs) are frequently observed in the global ocean [1]–[5]. Large-amplitude IWs can travel for hundreds of kilometers and carry a significant volume of energy toward the coastline. The IWs threaten the safety of underwater navigation [6], enhance ocean mixing [7], affect sediment resuspension [8], and have a significant impact on offshore facilities [9].

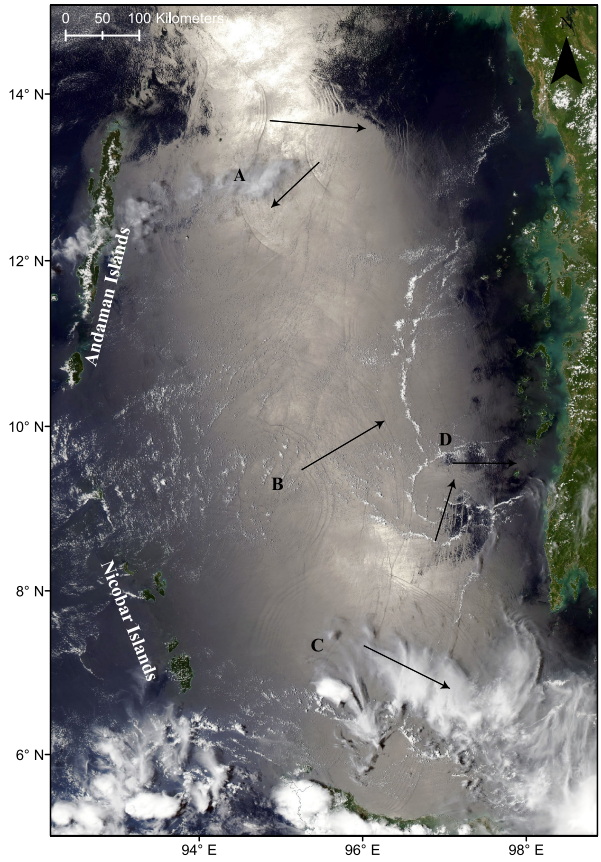

Fig. 1. MODIS true-color image of the Andaman Sea on 17 April 2003. Four IW packets are labeled by A–D. Black arrows represent IW PDs.

Remote sensing has been an effective method for oceanography research in recent years. Both synthetic aperture radar (SAR) and optical satellite images are used for IW studies [10]–[12]. Fig. 1 shows a Moderate-Resolution Imaging Spectroradiometer (MODIS) true-color image covering the Andaman Sea, a marginal sea in the northeastern Indian Ocean with a water depth over 2000 m in the central region. The IWs, which manifest as bright and dark bands on the MODIS images, are generated at the Nicobar Islands and the Andaman Islands and propagate eastward into the deep ocean with long wave crests and large amplitudes [13]–[15]. The imaging mechanism, generation sites, generation mechanisms, and spatial distribution of IWs have

Manuscript received December 15, 2020; revised February 1, 2021; accepted February 22, 2021. Date of publication March 3, 2021; date of current version March 24, 2021. This work was supported in part by the Strategic Priority Research Program of the Chinese Academy of Sciences under Grant XDA19090103, Grant XDA19060101, and Grant XDB42000000, in part by the National Natural Science Foundation for Young Scientists of China under Grant 41906157 and Grant 41604200, in part by the National Natural Science Foundation of China under Grant 41776183, in part by the Major Scientific, and Technological Innovation Projects in Shandong Province under Grant 2019JZZY010102, in part by the Key Project of Center for Ocean Mega-Science, Chinese Academy of Sciences under Grant COMS2019R02, and in part by the CAS Program under Grant Y9KY04101 L. *(Corresponding author: Xiaofeng Li.)*

The authors are with the CAS Key Laboratory of Ocean Circulation, and Waves, Institute of Oceanology, Chinese Academy of Sciences, and Center for Ocean Mega-Science, Chinese Academy of Sciences, Qingdao 266071, China (e-mail: zhangxd@qdio.ac.cn; xiaofeng.li@ieee.org; qzheng2@umd.edu).

Digital Object Identifier 10.1109/JSTARS.2021.3063529

been investigated with field data [15] and satellite images [16]–[18]. The previous investigators have found that the IWs in the Andaman Sea are mainly distributed in four regions in the northern, central, southern, and eastern Andaman Sea and propagate in different directions. We labeled them with arrows from A to D in Fig. 1. In Region A, the IWs originate from various sources and propagate southeastward and southwestward. Opposite propagating IW crests interact with each other. In Region B, the IWs have long wave crests and propagate eastward into the central Andaman Sea. The eastward propagating IWs are observed in Regions B and C. The long-crest IWs originating from the two regions collide with each other when propagating into Region D.

From satellite images like Fig. 1, the distribution [16], generation mechanism [19], propagation characteristics [20], such as reflection [21] and refraction of the IWs [22] have been reported in other ocean areas, such as the South China Sea (SCS). Since the forecast of IWs is significant for operational uses, the IWs propagation is mainly forecasted using numerical [23] and empirical models with considerable uncertainty. In the SCS, the locations of IWs have been predicted using the relation between the IWs and the westward tidal current at the Luzon Strait [24]. Based on the correlation between the IW occurrence and the tidal range, Hyder *et al.* [25] forecasted the possible IW occurrence focusing on one generation site in the northern Andaman Sea. As shown in Fig. 1, the IWs in the Andaman Sea have complicated propagation patterns and seven generation sites [16]. This complexity makes the prediction of IWs more challenging. To our knowledge, no numerical model simulation has been carried out to predict IW propagation in the Andaman Sea due to complicated IW propagation directions (PDs) and generation sites. To overcome the difficulty, we developed a machine learning method to predict IW propagation in the Andaman Sea in this study.

Machine learning techniques have shown great potential in oceanographic research [26], [27] and remote sensing data interpretation [28]. Krestenitis *et al.* [29] showed that machine learning is an efficient approach to identify an oil spill. Liu *et al.* [30] explored mining SAR imagery for coastal inundation mapping based on deep convolutional neural networks. Machine learning also provides the opportunity for building relationships among multidimensional information. Lee *et al.* [31] retrieved total perceptible water from Himawari-8 data. The relation of total perceptible water and observation data in different channels was built using machine learning techniques. Pan *et al.* [32] applied one of the machine learning approaches, the back-propagation (BP) algorithm, to retrieve the amplitude of IWs using texture information of satellite images and ocean environmental parameters. Li *et al.* [33] applied deep learning techniques to extract IW signals from satellite images and systematically reviewed the application of machine learning techniques to the information mining of remote sensing imagery. Zheng *et al.* [34] achieved the forecast of tropical instability waves based on a purely satellite data-driven deep learning model. The forecast results agree well with the satellite observations. These studies have shown that machine learning techniques have the advantages of a strong nonlinear mapping and multidimensional information processing. Zhang and Li [35] proposed a machine learning based forecast model (ZL model) for IWs in the Sulu-Celebes Sea, which shows close results with satellite observations. The ZL model neglects the transmeridional propagating IWs in the Celebes Sea, and is mainly suitable for noncross-propagating IW patterns. A major difference for IWs in the Andaman Sea and Sulu-Celebes Sea is that cross-propagating IWs are observed in the Andaman Sea. Complicated IW propagation patterns limited the direct application of the ZL model to the Andaman Sea. To handle the cross-propagating IWs, this article proposes a new machine-learning model that can be applied to other areas in the global ocean with cross-propagating IWs.

The propagation of IWs is affected by many factors ranging from the properties of IWs to ocean environments. It is conceivable to use a machine learning technique to find the relationship of IW propagation with other oceanic environmental variables. To find such a link, we built a BP neural network to predict locations of IWs when given an initial IW position. Training samples were extracted from collected MODIS images and ocean land color instrument (OLCI) images to build the BP model. The model was then used to predict the propagation characteristics of IWs in the Andaman Sea. We organized this article as follows: the machine learning IW prediction model and satellite observations of IWs in the Andaman Sea are described in Section II. The model results are shown in Section III and an analysis is given in Section IV. Section V summarizes the results.

II. Data and Method

A. Machine Learning IW Forecast Model

An IW forecast model was developed using a fully connected BP neural network. The forecast model includes two modules: the phase speed (PS) and the PD modules. We implemented the BP neural network with satellite-derived IW parameters, ocean stratification, locations, and the water depth as inputs. The BP model is a multilayer feed-forward, fully connected network trained according to the error BP mechanism, which is ideal in finding complex relationships in multidimensional information. The flowchart of the forecast model is shown in Fig. 2. The model input is an initial IW crest extracted from satellite images. The input parameters can be extracted based on the input wave crest and additional external data source. The BP neural network forecast (BPNNF) model is composed of one input layer, multiple hidden layers, and one output layer [32], [36]. The training process consists of forwarding propagation of IW PS and direction calculations and BP of errors, which can automatically change the weights of the neural network. The neural network automatically adjusted weights to minimize the error between the model outputs and expected outputs. In training, every training sample contributes to the adjusting of the neural weights. A two-hidden-layer structure was used for two modules.

The seven inputs of the PS module are latitude, longitude, water depth, mixed layer depth, density difference of the upper and lower layers, length of wave crest (LWC), and peak-to-peak (PP) distance of IWs [37]. The PD module inputs have an

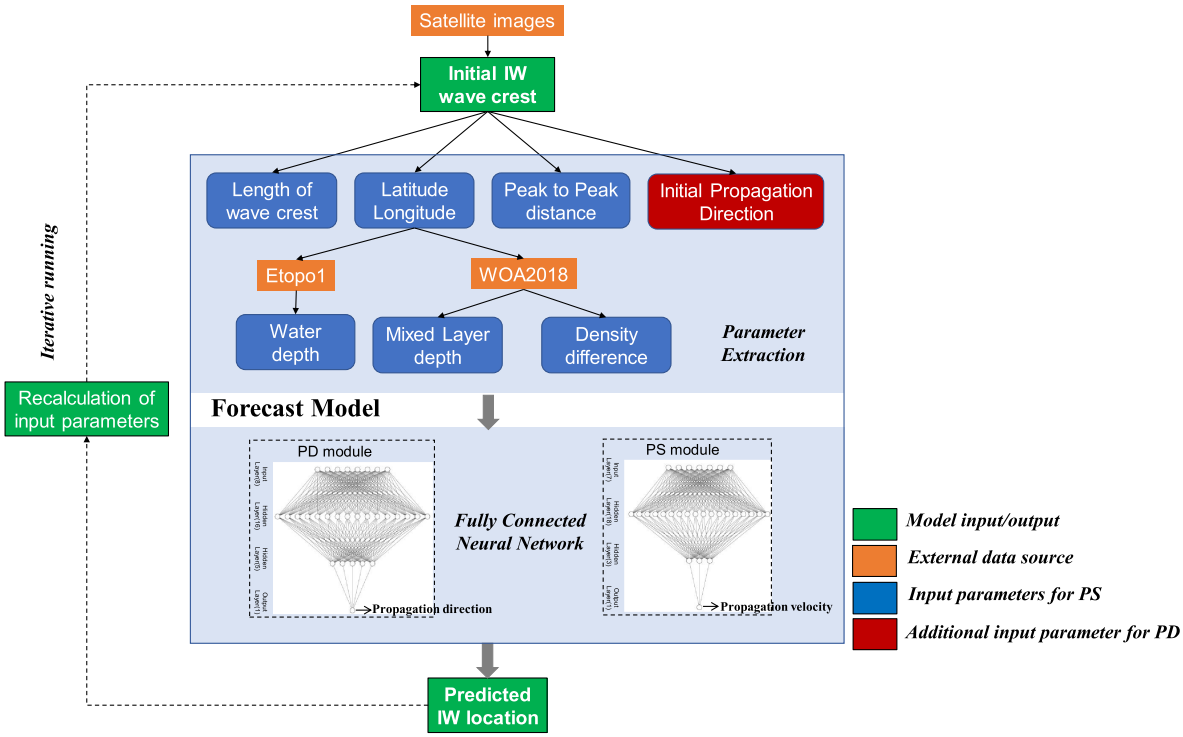

Fig. 2. Flowchart of the forecast model. The nodes of the input layer, hidden layers, and output layer for the PS module are empirically set to 7, 18, 3, and 1, and for the PD module are 8, 16, 5, and 1. The activation function between hidden layers is the tansig function. The activation function between the hidden layer and the output layer is the pureline function. The dark blue boxes indicate the input parameters for the PS module while the red box serves as an additional input parameter for the PD module.

extra parameter of the initial PD used to resolve locations with cross PDs of IWs, as shown in Region D in Fig. 1. The output parameters for the PS and PD modules are the propagation speeds and directions of IWs. The nodes for the PS module's two hidden layers were empirically set to 18 and 3, while the PD module was 16 and 5. The training function, 'trainlm' based on the Levenberg–Marquardt algorithm [38], was adopted for its fast convergence rate. The activation function among the three hidden layers was the hyperbolic tangent sigmoid transfer (tansig) function, and the activation function between the hidden and output layer was the linear transfer (pureline) function. The model initialized each node with random weights, and then the model was trained for many epochs (one epoch means the model was trained once using the training dataset). The model automatically adjusted the weights of each node to find the best weights. The trained model was evaluated using the training dataset and independent test dataset. The training dataset was divided into the training group and validation group. The training group was used for model training and the validation group for monitoring the training process. After the model reached its best validation (best performance on the validation group), the model training was stopped and further tested using an independent test dataset. Excellent performance on the training dataset and poor performance on the test dataset means the model is overfitted, and an overfitted model does not generalize well from the training dataset to an unseen dataset. To overcome the overfitting problem, we adopted the early stopping method [39]. We set the validation check as six, implying that the model will stop training if the error on the validation group does not decrease for six consecutive training steps.

The model could then be trained using an established dataset based on the fully connected (BP) network. When it comes to model validation and future applications, a satellite image containing IW signatures is needed. The input information was extracted from the satellite images and publicly available data sources (ETOPO1 and WOA 2018 datasets). The extracted IW wave crest was saved as a series of longitude and latitude points, and these points were input into the model one by one. The PS and PD modules predicted the propagation speed and direction. After integrating on the time step, the established forecast model output the locations of the IW after one or several time steps (12.42 h, for example) for each point. The forecasted points were connected together to get the forecast IW locations. When running the model iteratively, the LWC can be recalculated using the distances between adjacent points.

B. *Extraction of IW Signatures From Satellite Images*

Two types of satellite images, MODIS and OLCI images, were used in our study. The MODIS sensors onboard the National Aeronautics and Space Administration (NASA) satellites Terra and Aqua acquire data in 36 spectral bands with a swath of 2330 km and a spatial resolution between 250 and 1000 m. The OLCI is a push-broom instrument with five camera modules onboard the European Space Agency (ESA) Sentinel-3 satellite. The OLCI acquires images in 21 spectral bands with a swath of

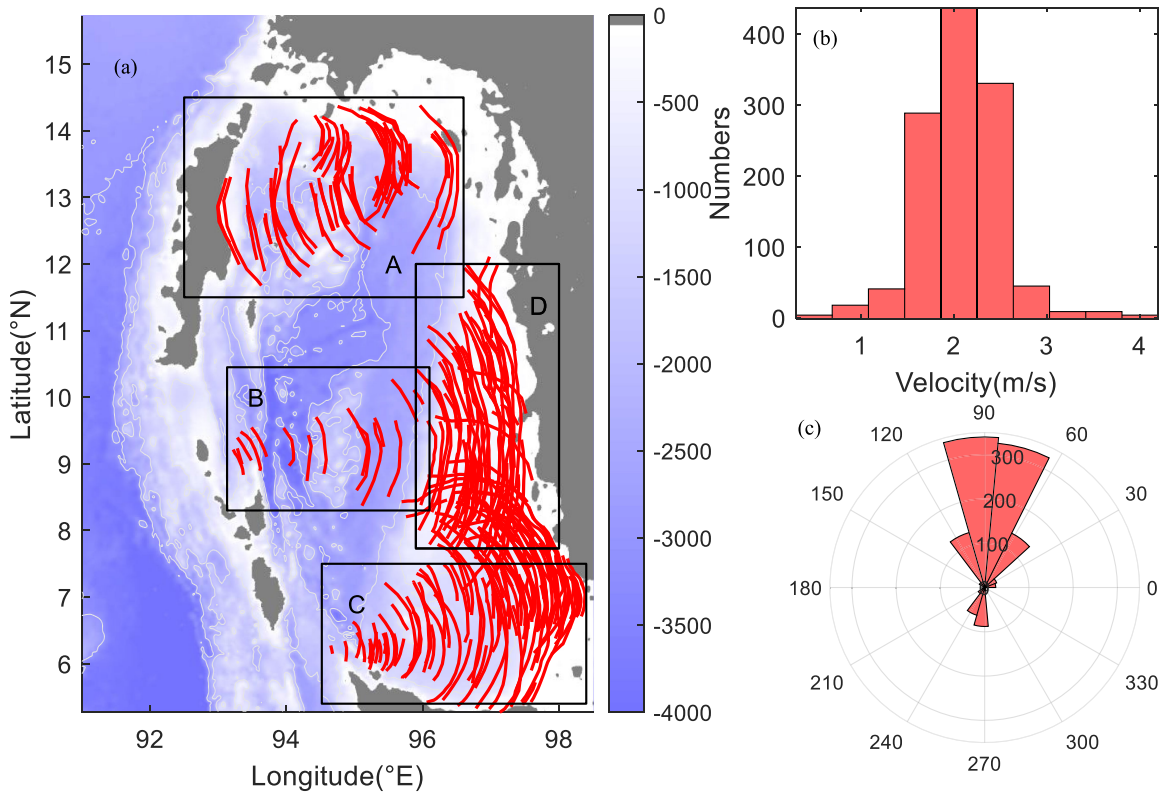

Fig. 3. (a) Leading IWs in the Andaman Sea extracted from satellite images. (b) Histogram of propagation speeds of IWs. (c) Rose map of PDs of IWs. The north corresponds to 0 degrees and follows a clockwise direction.

1440 km and a spatial resolution of 300 m for full-resolution and 1200 m for reduced-resolutions. In this study, 123 MODIS images with a spatial resolution of 250 m (Band 1 and 2) and 33 full-resolution OLCI images (Band 3) containing clear IWs signatures were used.

IWs manifest as bright and dark bands on satellite images, as shown in Fig. 1. The latitude and longitude information for the Single IW or leading wave of an IW packet was labeled manually and output to a geographic information system (GIS) file. The LWC was then measured with the created IW labels. The PP distances were obtained by measuring the positive and negative peaks of the IW profiles along the PD [37]. Propagating IWs will induce convergence and divergence regions on the surface and modulate the capillary surface waves. The modulated ocean surface will modulate the energy received by the radar and hence make the IW manifest as bright or dark bands. The PP distance is closely related to the amplitudes of IWs, which makes it a good indicator of IW scales across wave crest direction. The extracted wave crests of IWs contain the longitude and latitude information, so the water depth where IW is observed can be interpolated from the ETOPO1 Global Relief Model with a spatial resolution of 1 arc-minute. The ocean stratification data was extracted from the World Ocean Atlas (WOA) 2018 dataset, where the monthly averaged temperature and salinity data are provided. The mixed layer depth corresponding to the maximum buoyancy frequency level and stratification data were derived from the WOA 2018. The density difference was the averaged density difference between the upper and lower layers under the two-layer ocean assumptions. The initial PD was needed to resolve cross-propagation IWs, and it was determined based on the extracted wave crests.

When building the training dataset for the forecast model, we also estimated the propagation speeds of IWs from the locations and time differences of IWs. If we tracked an IW in a pair of images, we used the displacement of the IWs and the acquisition time difference to calculate the IW PS. If multiple IWs were in a single satellite image, we assumed the time difference between the two IWs was the semidiurnal tide period. Although IWs generated by different tidal cycles may not be quite the same, the extracted samples show that IW propagation speeds show no obvious variations during the spring and neap tide, mainly ranged between 2.0 and 2.5 m/s. IWs generated at the spring/neap tide had different characteristics, so the difference may be considered implicitly to some extent by including the PP distance and LWC as input parameters. As a result, we collected 1189 samples and further divided them into an 80%/20% training/test dataset.

We overlaid the IW leading wave crests on the topographic map in Fig. 3(a), which shows the spatial distribution of IWs. IW propagation characteristics show that IW activities were mainly in four regions of the Andaman Sea. In Region A, the IWs had two opposite PDs, and the LWC of the IWs ranged from tens of kilometers to over one hundred kilometers. IWs in Region B had short LWCs, which grew to several hundred kilometers when propagating to Region D. Longer IWs were mainly located in the eastern Andaman Sea because of the eastward propagating IWs and deep water. Based on the extracted IW information,

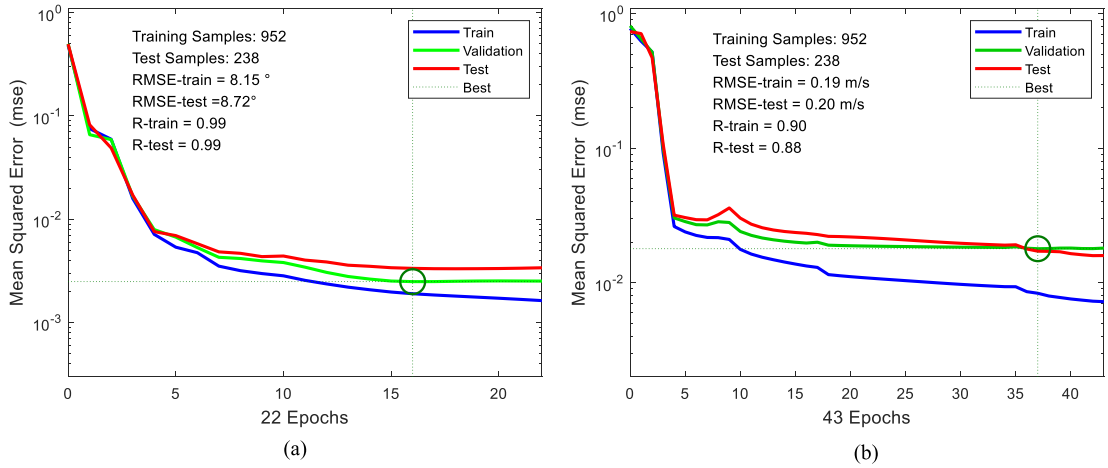

Fig. 4. Loss plot of the PD (a) and PS (b) modules. The insert texts show the root mean squared error and correlation coefficients in the training and test dataset.

the propagation speeds of the IWs in the Andaman Sea range from 0.5 to 4.1 m/s, with a mean value of 2.0 m/s [see Fig. 3(b)]. Fig. 3(c) shows that the IWs in the Andaman Sea had a significant variation of PDs, with a dominant PD of 60°–90°.

III. Results

A. Model Results

The BPNNF model was trained using the training dataset described in Section III. The model results for the training dataset and test dataset for the PS and PD modules are shown in Fig. 4. For the PS module, the root-mean-square error (RMSE) of the training dataset and test dataset were 0.19 and 0.20 m/s, respectively, and the correlation coefficients were 0.90 and 0.88, respectively. Since the mean IW propagation, PS was around 2.0 m/s, as shown in Fig. 3(b), the PS module gave a PS error of 10%. The training performance in Fig. 4 shows that the model's error decreased gradually, and the mean-square error (MSE) of the PS module reached the best validation performance at epoch 37 [see Fig. 4(b)]. Similarly, Fig. 4(a) shows that the RMSE of the PD module was less than 10°, and the correlation coefficients were both up to 0.99. Fig. 4 shows that the gradient decreased gradually, and the MSE of the PD module reached the best validation performance at epoch 16. When the MSE was not reduced, and validation checks increased in epoch 16, the model stopped training in epoch 22 (validation check reaches six). This result shows the self-adjusting ability of BP neural networks.

B. IW Forecast After One Tidal Cycle

To test the forecast model, we analyzed four MODIS images containing clear IWs signatures, as shown in Fig. 5. The solid color lines in Fig. 5 represent the locations of the IWs generated from different generation sites at the different tidal cycles. In Fig. 5(a), five IWs propagate in opposite directions in Region A. Two IWs with shorter wave crests propagate eastward, while the other three IWs with longer wave crests propagate westward. Three large IWs propagating eastward are shown in Fig. 5(b), propagating from Region B to D. Fig. 5(c) shows three IWs propagating in the northeast direction in Region C. In Region D, Fig. 5(d) shows four IWs with cross PDs. The LWC, PP distances, water depth, mixed layer depth, density difference, and locations, were extracted along the IW wave crest as model inputs. Using the leading wave crest within an IW packet as input, we ran the model to forecast the IW locations. The model-forecasted IW locations after one tidal cycle were compared with IW wave crest locations extracted from satellite observations. IWs in Fig. 5(a)–(d) are labeled as Case 1, 2, 3, and 4.

TABLE I
Biases of the Model Results and Satellite Observations

	RMSE (km)	Fréchet Distance (km)	Correlation Coefficient
Case 1-1	2.69	13.90	0.95
Case 1-2	2.99	10.66	0.97
Case 2-1	6.10	18.28	0.96
Case 2-2	2.50	9.06	0.89
Case 3-1	3.15	10.43	0.93
Case 3-2	2.37	6.42	0.93
Case 4-1	4.12	12.66	0.97
Case 4-2	1.79	10.28	0.99
Average	3.21	11.46	0.95

In each case, more than two wave crests of IWs generated at different semidiurnal tidal cycles are shown. The machine-learning model results are shown as solid black lines in Fig. 6. The red lines in Fig. 6 represent the locations of IWs in Fig. 5. The machine-learning model results agreed well with satellite observations. Since the RMSE of the PS module is about 0.20 m/s (see Fig. 4), this propagation speed error led to an average distance error of 8.94 km after one semidiurnal tidal cycle (12.42 h). Thus, the most significant distance error was 17.66 km for the northern wave crests in Case 2-1, as shown in Fig. 6(c). The IWs in the same location with cross PDs are also well forecasted, as shown in Fig. 6(g) and (h).

The comparison between machine-learning model results and satellite observations are listed in Table I. Case 2-1 has the largest RMSE, while Case 4-2 has the smallest RMSE, as

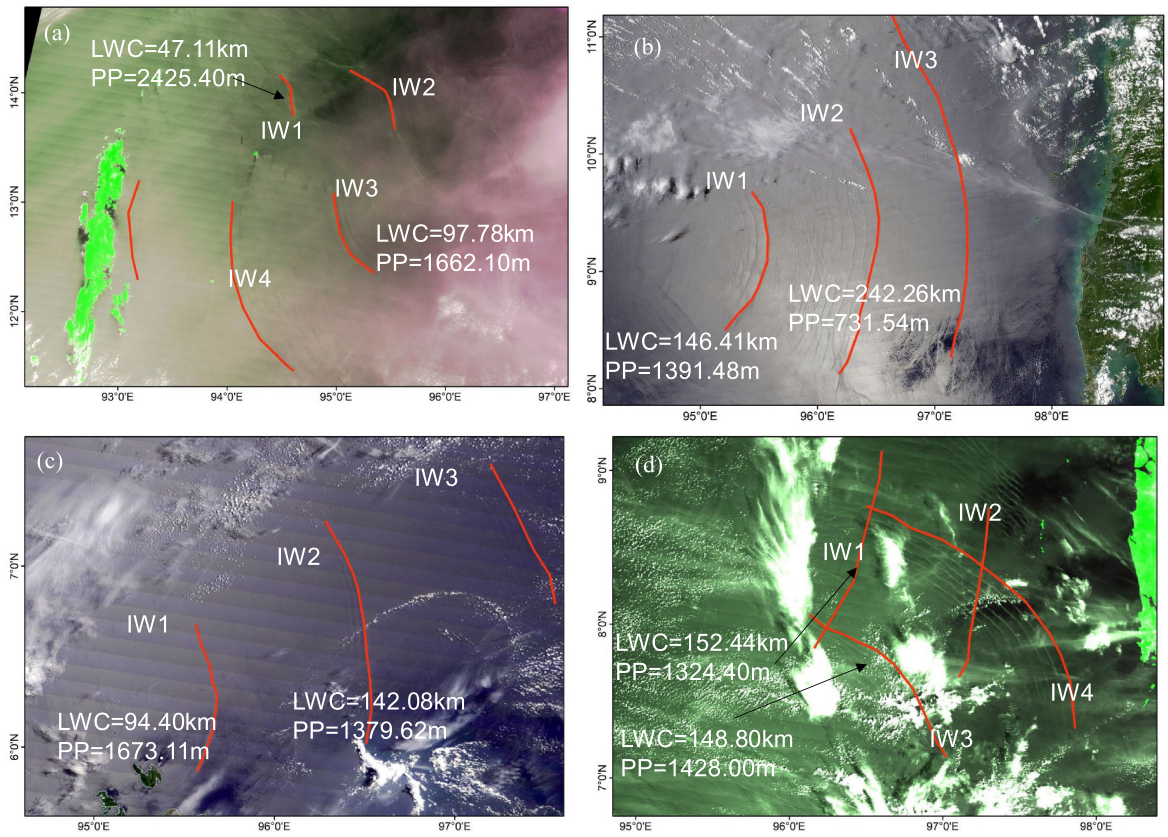

Fig. 5. Satellite images were used to test the machine-learning models. MODIS images were acquired at 04:35 on 9 February 2019 (a), 04:02 on 9 May 2017 (b), 04:15 on 30 July 2019 (c), and 04:02 on 9 May 2017 (d). Solid red lines represent locations of IWs generated at different tidal phases.

shown in Fig. 6(h). IWs propagating in regions with significant water depth changes may result in larger RMSE, as shown in Fig. 6(b) and (c), and vice versa [see Fig. 6(d) and (h)]. The total average RMSE of the machine-learning model results and satellite observations was 3.21 km. Two factors were adopted to evaluate the results, i.e., the correlation coefficient and the Fréchet Distance were used to assess the similarity of two curves by considering their distance and shape [40]. A higher similarity between two curves gives a lower Fréchet Distance. If two curves have the same location and shape, the Fréchet Distance is zero. As shown in Table I, the average correlation coefficient is 0.95, and the average Fréchet Distance is 11.46 km, indicating high agreement between the machine-learning model and satellite observations.

C. IW Forecast After Multiple Tidal Cycles

The machine-learning model forecast results of two semidiurnal tidal cycles were tested using MODIS images in Regions A–C. Fig. 7(b) shows that the IWs propagate from Region B to Region D. The forecasted locations of IWs for one and two semidiurnal tidal cycles are shown as solid black lines. The model predicted IW locations after one semidiurnal tidal cycle served as the input for the model to predict IW locations after two semidiurnal tidal cycles. Thus, the forecast model could repeatedly run to predict the IW locations for multiple tidal cycles.

The forecasted results of the machine-learning model after two semidiurnal tidal cycles are shown in Fig. 7. Fig. 7(a) and (b) shows that the machine-learning model's predicted IW locations were consistent with cloud-free satellite observations. However, Fig. 7(c) shows more significant location discrepancy. We believe that the more complicated topography in the northern Andaman Sea may lead to considerable changes in the parameters of IWs, such as the PP distance.

IV. Discussion

A. Influence of Different Model Run Time Steps

Since the machine-learning forecast model can run iteratively to predict IW locations, the time difference between the input and output of IW locations is defined as the time step. The time step of the model affects the forecasted results, especially in the IWs propagating across the isobaths. An example showing the influence of running the model with different time steps is shown in Fig. 8. We ran the model with time steps set to 1/4, 1/3, 1/2, and the full semidiurnal tidal cycle, i.e., 3.11, 4.14, 6.21, and 12.42 h, and compare the model results with satellite observations. When IW1 propagates to IW2 in Fig. 8, the water depth changes from more than 2000 m to less than 1000 m. The more substantial time step of 12.42 h shows worse results than

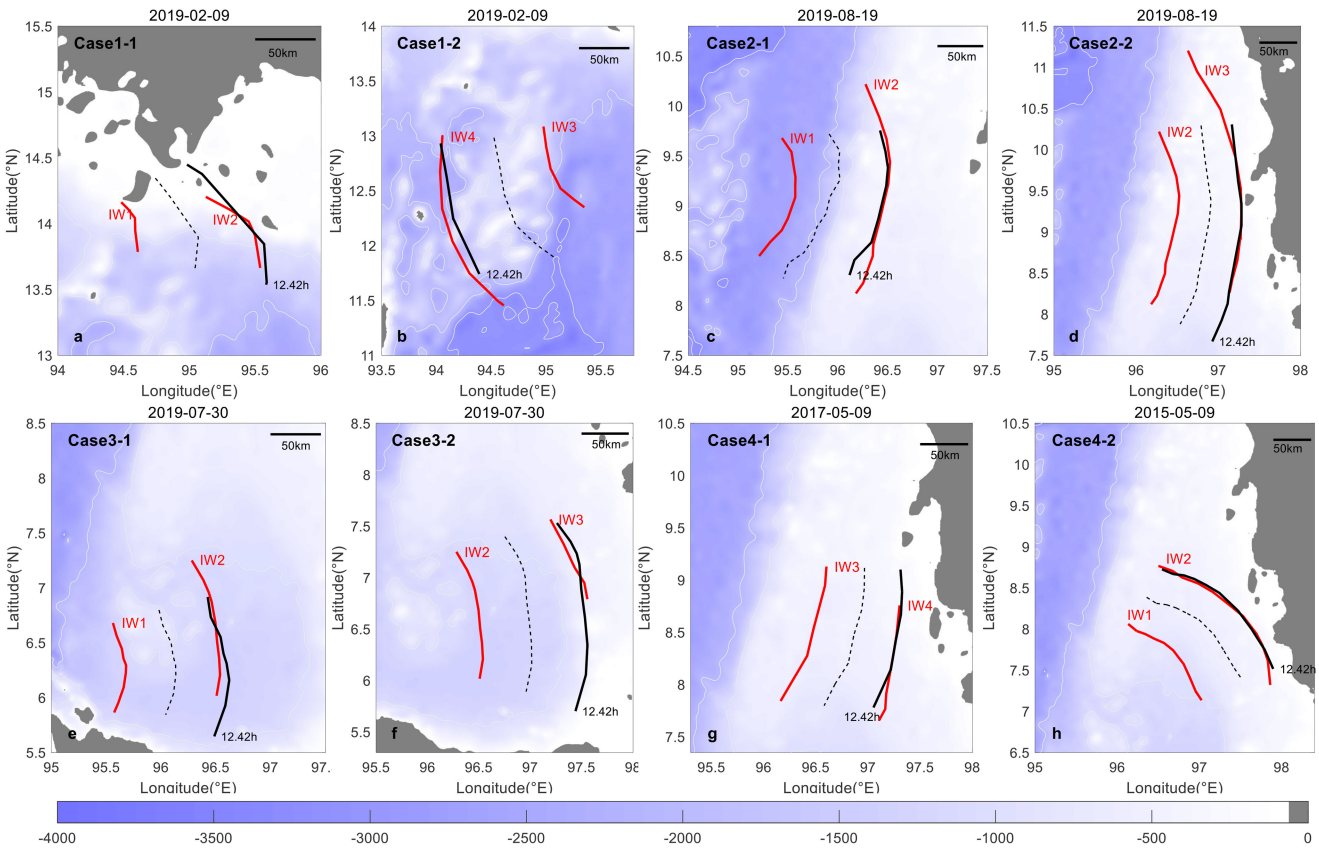

Fig. 6. Comparison of the machine-learning model's forecasted IWs to those observed in MODIS images in Fig. 5. Labels of IWs are the same as Fig. 5. Cases 1, 2, 3, and 4 correspond to Fig. 5(a)– (d). Solid red lines represent satellite observations of IWs and black dashed lines represent model-predicted locations of IWs after 6.71 h. Solid black lines represent machine-learning model predictions after one semidiurnal tidal cycle (12.42 h).

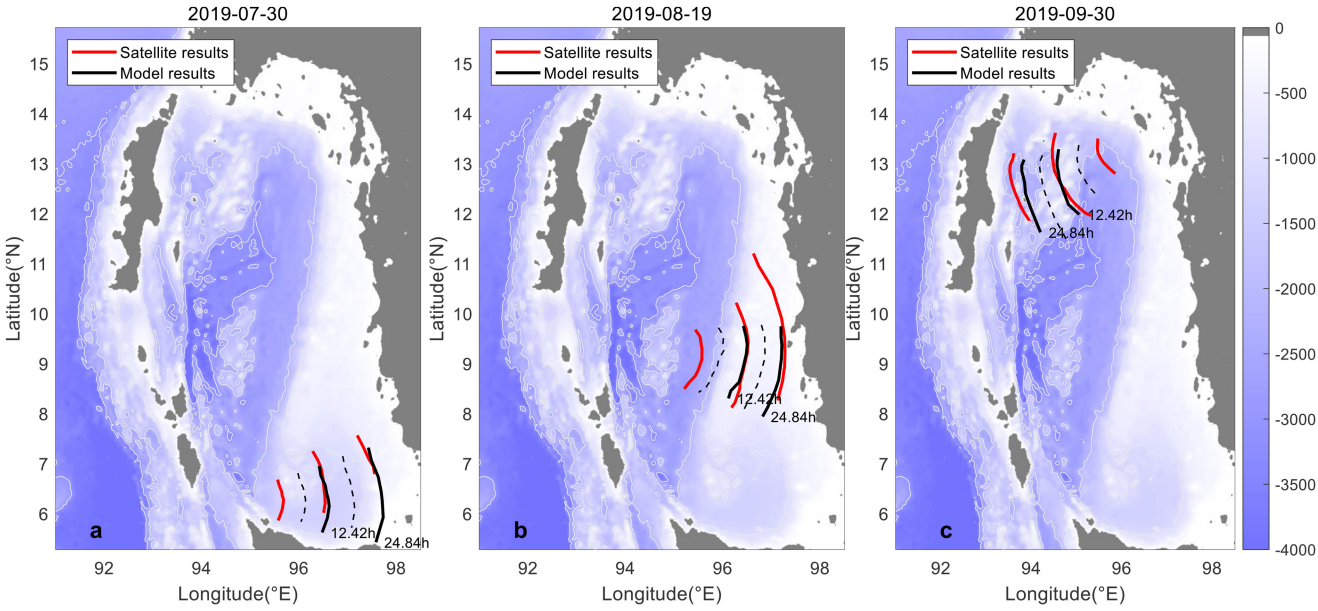

Fig. 7. Same as Fig. 6 but for the machine-learning model's forecasted results after two semi-diurnal tidal cycles. MODIS images were collected on 30 July 2019 (a), 19 August 2019 (b), and 30 September 2019 (c), respectively.

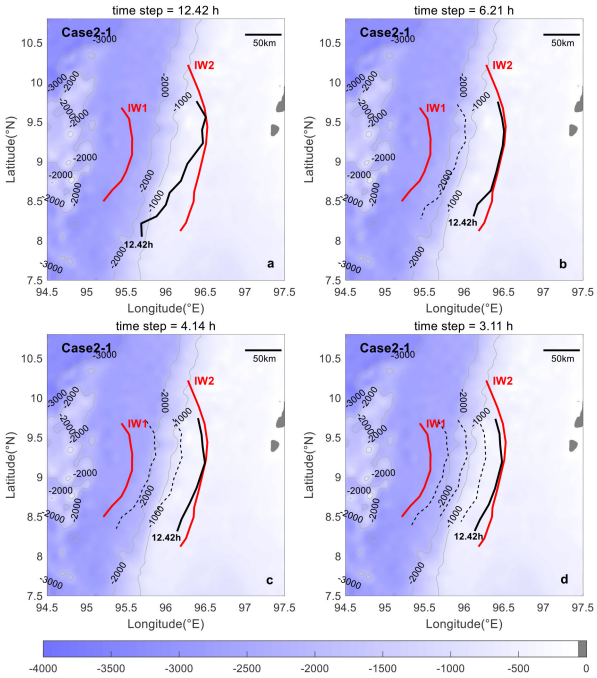

Fig. 8. Same as Fig. 6 but for different time steps. Time steps are set to 12.42 (a), 6.21 (b), 4.14 (c), and 3.11 h (d).

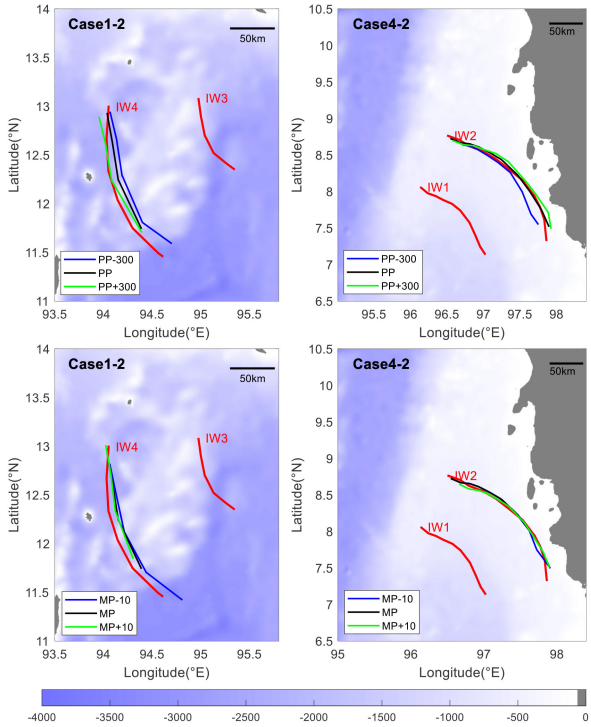

Fig. 9. Same as Fig. 6 but for different PP distances and initial main PDs. Solid black lines represent the results with normal PP distances (initial main PDs) and solid blue (green) lines represent the results with smaller (larger) PP distances (initial main PDs).

those with smaller time steps. Using a more modest time step can be considered as the bathymetry changes more gradually. When time steps were 4.14 and 3.11 h, the model results were almost the same, and the discrepancy with the time step of 6.21 h was also minimal. Thus, when the topography changes significantly, the IWs propagate across the steep isobaths; a smaller time step may improve the forecast result.

B. *Influence of Errors in Input Parameters*

Eight parameters, which served as inputs of the machine-learning forecast model (see Fig. 2), were extracted from labeled IW crests. The machine learning forecast model built the relationship between the input parameters and predicted IW locations. The error introduced in the input parameters affected the model's predicted results. When we ran the forecast model iteratively to predict IW locations after multiple tidal cycles, all the input parameters were updated except the PP distance and initial PD. The initial PP distance and PD always served as inputs in the following tidal cycles. When IWs evolved in the ocean, their properties, such as the LWC and the PP distance, changed. Errors were also introduced to the PP distance when IWs were not observed clearly due to the high wind or waves, clouds or the geometry of observations. The influences of errors in the PP distances and initial PDs to the forecasted results were tested in four regions of the Andaman Sea. The results are shown in Fig. 9.

The spatial resolutions of MODIS and OLCI images used in this study are 250 and 300 m. An error of one pixel leads to a PP distance error of about +/− 300 m. Thus, the PP distances were set to three values: PP-300 m, PP (the actual PP distance), and PP+300 m to test the influences on the machine-learning model's forecasted results. The time step of the machine-learning forecast model was set to 6.21 h. Two regions with large and small water depth changes in the Andaman Sea were tested, and the results are shown in Fig. 9. Although different errors are included in the PP distance parameters, the forecasted IW propagations after one tidal cycle were still close to each other and satellite observations.

The errors in IW's initial main PDs were tested like the PP distance cases. To evaluate the influences on the forecast results, we used three PDs as inputs: MP-10°, MP (the real initial PDs), and MP+10°. The forecast model's time step was set to 6.21 h, and the results are shown in Fig. 9. With different initial PDs, the machine-learning model's forecasted IW locations were close to each other and satellite observations.

The results shown in Fig. 9 verify that we can run the model iteratively over multiple tidal cycles to predict IW positions even without updating these two input parameters. The results indicate that the machine learning forecast model has a enormous tolerance for errors included in input parameters, such as the PP distance and initial PD. Since eight parameters determined the IW propagation (see Fig. 2), small errors in some input parameters did not deteriorate the model performance. The model still obtained results close to the satellite observations. One should realize that errors associated with the input parameters will accumulate with the iterative running of a model. However, test results shown in Fig. 7 indicate that the model's forecasted results were still reliable after two tidal cycles. We believe that

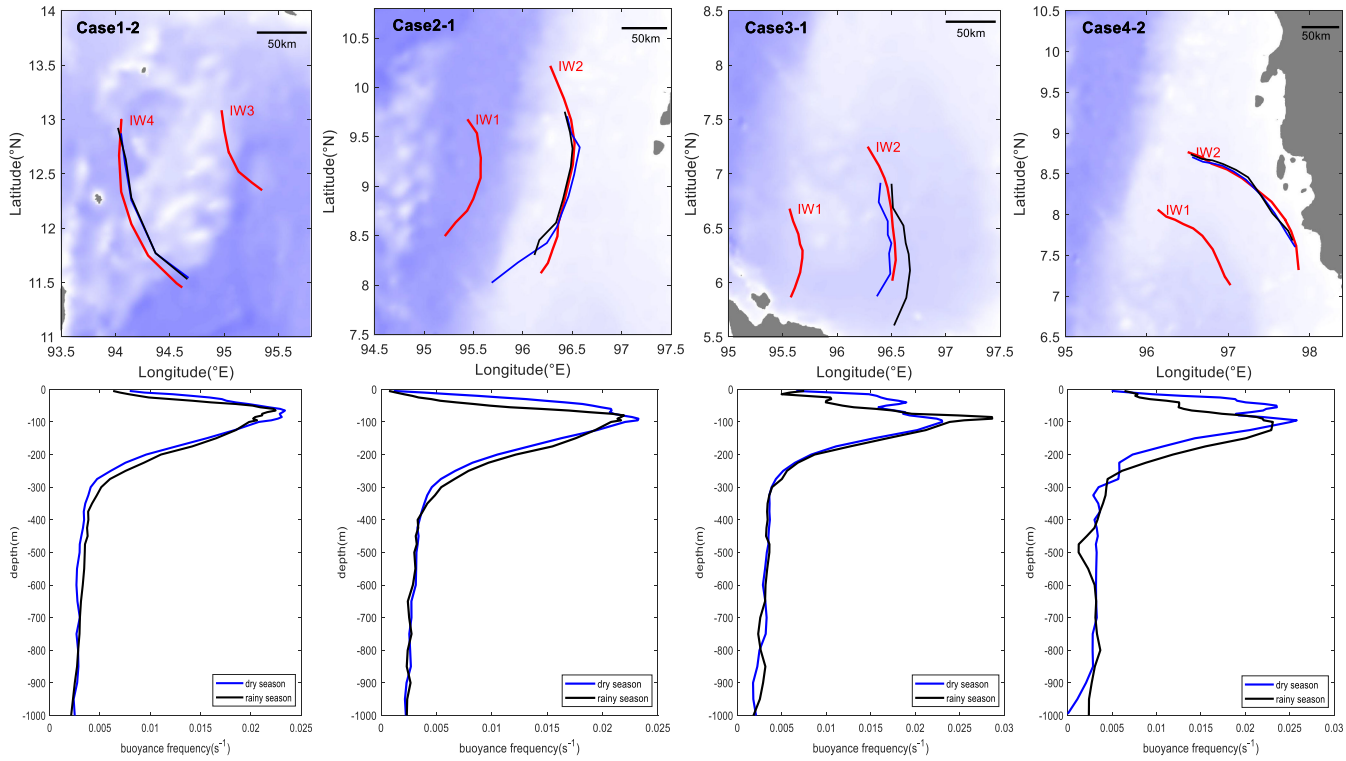

Fig. 10. Same as Fig. 6 but for IW propagation during different seasons. The upper panels show forecast results in different seasons, the corresponding buoyancy frequency is shown in lower panels. Blue lines represent the forecast (buoyancy frequency) results of the dry season and black lines represent the results of the rainy season.

two tidal cycles is long enough to predict trans-meridional IW propagation characteristics in the Andaman Sea.

C. *Influence of Seasonal Variations*

IW propagations have seasonal variations because the ocean stratification changes dues to solar heating, ocean mixing, precipitation, and other factors. The Andaman Sea has a dry season from January to April and a rainy season from May to November controlled by the Indian monsoon. To examine the influences of seasonal variations on IW propagation, we ran the machine-learning forecast model to predict IW propagation during the dry and rainy seasons in four regions of the Andaman Sea.

The four cases shown in Fig. 6 were selected to examine the IW propagation after one tidal cycle in different seasons. These cases cover the four regions labeled in Fig. 3. We chose March and August to represent the dry season and rainy season, respectively. The input parameters are the same as those in Fig. 6, with the density profiles derived from the climatology data of March and August and a time step of 6.21 h. Variations of the mixed layer depth and density difference were apparent in the two seasons. The machine-learning model's forecasted results are shown in Fig. 10. In Cases 1-2, 2-1, and 4-2, the model-predicted IW locations after one tidal cycle were close to each other and the satellite observations. IW predictions in Case 3-1 had more substantial discrepancies between the results in the dry and rainy seasons. The results indicate that IW propagations in Regions A, B, and D show little seasonal variations, while in Region C, discrepancies in the dry and rainy seasons exist. The IWs propagate faster in the rainy season.

The buoyancy frequency values of the four regions in the dry and rainy seasons are shown in Fig. 10. The distribution of buoyancy frequency in Cases 3-1 and 4-2 had more substantial discrepancies than that in Cases 1-2 and 2-1. In the dry season, Cases 3-1 and 4-2 had one more significant peak and one smaller peak, while there was only one peak in the rainy season. The peak values of the buoyancy frequency of Case 3-1 had the most significant difference, which indicates that the most considerable variations of ocean stratification were in Case 3-1 (Region C). A higher peak value of the buoyancy frequency indicates stronger ocean stratification and a more considerable density difference between the upper and lower layers of the ocean. The propagation speed of IWs is proportional to the density difference [41]. This result explains why IW propagated faster in the rainy season in Case 3-1, while the peak values in Case 4-2 were close in the dry and rainy seasons.

D. *Comparison With the Classic Korteweg–De Vries Equation Results*

The Korteweg-de Vries (KdV) equation is widely used in the estimation of IW propagation velocities. We compared the propagation velocities from the BPNN model and the KdV equation. The PD module was used to provide the PD for the KdV equation forecast. The nonlinear propagation velocity described by the

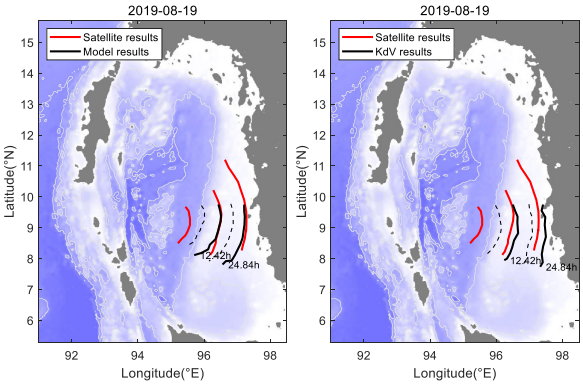

Fig. 11. Forecast of IW propagations in the Andaman Sea using the developed model and the KdV equation after two semidiurnal tidal cycles. The left image shows the results of the developed model, and the right image shows the results of the KdV equation.

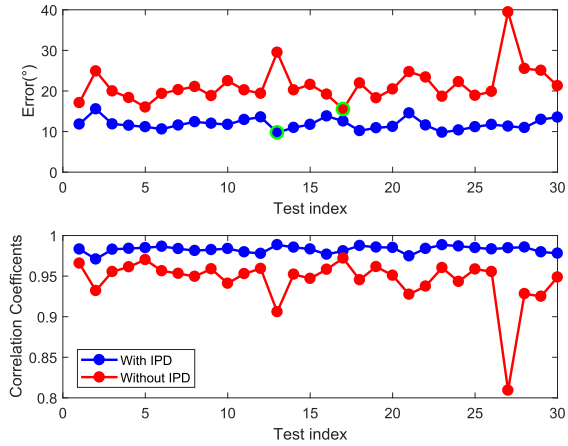

Fig. 12. Model test results with (blue lines) or without (red lines) the initial main PDs for 30 tests. The green circles indicate the best performance in the 30 tests. The upper panel shows the errors on the test datasets, and the lower panel shows the correlation coefficients during the 30 tests.

KdV equation is expressed as [15], [42]

$$C_p = C_0 + \frac{\alpha}{3} A_0 \tag{1}$$

$$\alpha = \frac{3C_0}{2} \frac{h_2 - h_1}{h_1 h_2} \tag{2}$$

where C_0 is the linear PS, α is the nonlinear coefficient, and A_0 is the amplitude of IW. The amplitude of IW in the Andaman Sea was set to 60 m in this study [15]. IW propagation after two semidiurnal tidal cycles was also tested; the results are shown in Fig. 11. We found that the KdV results had more significant deviations after one and two semidiurnal tidal cycles.

The amplitudes of the IWs ranged from tens to hundreds of meters. Amplitude retrieval from satellite images has been studied in the literature [43], but the amplitude retrieval algorithm's accuracy still needs further investigation. We can briefly evaluate the influence of amplitude error on the forecast results. Here we considered an average water depth of 2000 m (see Fig. 3), a mixed layer depth of 100 m (see Fig. 10), average densities of the upper and lower layer of 1021 and 1025 kg/m3, and an average linear propagation speed is estimated to be 1.90 m/s. If we assumed an amplitude error of 20 m, it led to a nonlinear propagation velocity error of 0.18 m/s. During one semidiurnal tidal cycle, the phase velocity error will lead to a location error of 8.10 km. If significant errors are introduced to the IW amplitude, the forecast IW locations may have large deviations. Compared to the KdV equation, the developed model had better performance and was more robust when errors were introduced to the input parameters.

E. Influence of Initial Main PDs

Cross-propagating IWs have been observed in the Andaman Sea, which is different from IW characteristics shown in the Sulu-Celebes Sea [35]. To find the influence of initial main PDs on the forecast results, we trained the model 30 times. As shown in Fig. 12, models with an initial main PD had better performance. The RMSE was smaller and more stable on an independent test dataset for models with an initial main PD than models without an initial main PD. On tests 13 and 27, we found extra-large errors and lower correlation coefficients for the model without an initial main PD. This fact shows the lower generalizability of the model. When applied to the IW forecast, models without an initial main PD are more likely to produce results with large deviations. Therefore, the introduction of the initial main PD to the forecast model is necessary for situations with cross-propagating IW patterns. Previous discussions have also shown that the model has a large tolerance for errors introduced to the initial main PD, which makes the proposed forecast model robust.

V. Conclusion

In this article, we developed a BPNNF model for the analysis of IWs in the Andaman Sea. The model was trained with IW samples extracted from MODIS and OLCI images. The IW propagation speed and direction samples were determined using the locations and time differences of IWs on different satellite images or the location differences of IWs on a single satellite image under the assumption that the semidiurnal tide primarily generated them. The model includes two modules to estimate the propagation speeds and directions of IWs. The RMSE of the PS and PD module were 0.20 m/s and 8.72°. Based on satellite images, propagation characteristics of IWs in four subregions of the Andaman Sea were analyzed. The propagation speeds of IWs ranged from 0.50 to 4.10 m/s, and the dominant PD of IWs was 60°-90° with respect to north.

Test cases covering different regions and water depths in the Andaman Sea were used to test the developed machine-learning model. The model-predicted results of IWs after one and two semidiurnal tidal cycles are compared with satellite observations. The model results generally agreed well with the satellite observations. For one semidiurnal tidal cycle forecast, the largest RMSE was 6.10 km, the smallest RMSE was 1.79 km, and the total average RMSE was 3.21 km. Two factors, the correlation coefficient, and the Fréchet Distance were adopted to evaluate

the model results. The overall average correlation coefficient was 0.95, and the average Fréchet Distance was 11.46 km. The results indicate that the developed machine-learning model is accurate enough for the IW forecast in the Andaman Sea. An IW propagating in the near-shore area is shown in Case 1-1 of Fig. 6, where the northern part of the IW was locked by the shallower water depth while the southern part kept propagating east. This interesting phenomenon is also predicted well by the forecast model.

The period of the semidiurnal tidal cycle was assumed to be 12.42 h. Although the semidiurnal tidal period may be slightly different in different oceans, the difference did not affect the accuracy of the developed model. Because the distance between IWs was extracted from satellite images, the IW propagation speed was calculated using the derived distance divided by the semidiurnal tidal period. When forecasting IW propagations, the calculated propagation speed was integrated over the assumed semidiurnal tidal period. Therefore, the assumption did not affect the accuracy of the forecast results.

The influence of the time step of the model on the forecast results was analyzed. When IWs propagate across isobaths, a smaller time step is preferred. Errors are most likely introduced to the measurements of the PP distances. When one runs the forecast model iteratively, only the PP distance and initial PD are not updated in the following forecasts. The influence of errors in the PP distance and the initial PD on model forecast results were also analyzed finding that the developed machine-learning model had a large tolerance for errors included in the input parameters. Errors of $-/+300$ m introduced to the PP distance and $-/+10^{\circ}$ introduced to the initial PD did not influence the model results significantly (as shown in Fig. 10), implying that the developed forecast model is robust. Because the proposed forecast model has eight input parameters, some errors introduced to one input parameter did not affect the results greatly. Finally, the influences of seasonal variation on the IW propagation were analyzed. The results indicate that the IWs in Regions A, B, and D had little seasonal variation. Simultaneously, in Region C, discrepancies were found due to the seasonal variations of ocean stratification. The PS module of the developed model was also compared with the KdV equation; the results show that the BPNNF model had better forecast results and was more robust.

The machine learning forecast model does not require information on the generation location of IWs. In the Andaman Sea, IWs are generated at seven sites and cross propagate each other, limiting the empirical model's applicability. The earlier developed ZL model [35] is also not applicable for studying cross-propagating IWs in the Andaman Sea. The proposed forecast model in this study can be applied to other ocean areas where IWs show cross-propagating patterns. The only information needed to run our forecast model is the initial location of an IW wave crest.

Considering the complicated propagation characteristics of IWs in the Andaman Sea, we used parameters from IWs and ocean environments to establish the forecast model for predicting IW locations. Machine-learning techniques adopted in this study show excellent potential for multi-dimensional information processing and forecasting in oceanographic research.

ACKNOWLEDGMENT

The authors would like to acknowledge the use of MODIS and VIIRS imagery from the NASA Worldview application (https://worldview.earthdata.nasa.gov), part of the NASA Earth Observing System Data and Information System (EOSDIS). The water depth data were downloaded from the ETOPO1 Global Relief Model (https://maps.ngdc.noaa.gov/viewers/wcs-client/). The stratification data were obtained from World Ocean Atlas 2018 (https://www.nodc.noaa.gov/OC5/woa18/). The TPXO data was downloaded from OSU TPXO tide models (https://www.tpxo.net/global/tpxo9-atlas).

REFERENCES

[1] X. H. Xie, Y. Cuypers, P. Bouruet-Aubertot, A. Pichon, A. Lourenco, and B. Ferron, "Generation and propagation of internal tides and solitary waves at the shelf edge of the bay of Biscay," *J. Geophys. Res.-Oceans*, vol. 120, no. 10, pp. 6603–6621, Oct. 2015.

[2] I. Kozlov, D. Romanenkov, A. Zimin, and B. Chapron, "SAR observing large-scale nonlinear internal waves in the White Sea," *Remote Sens. Environ.*, vol. 147, pp. 99–107, 2014.

[3] M. H. Alford *et al.*, "The formation and fate of internal waves in the south China sea," *Nature*, vol. 521, no. 7550, pp. 65–9, 2015.

[4] J. Pan, D. A. Jay, and P. M. Orton, "Analyses of internal solitary waves generated at the Columbia river plume front using SAR imagery," *J. Geophys. Res.-Oceans*, vol. 112, no. C7, 2007.

[5] C. Garrett and W. Munk, "Internal waves in the ocean," *Ann. Fluid Dyn.*, vol. 11, no. 1, pp. 339–369, 1979.

[6] X. Li, Z. Zhao, and W. G. Pichel, "Internal solitary waves in the northwestern South China Sea inferred from satellite images," *Geophys. Res. Lett.*, vol. 35, no. 13, pp. 1–7, Jul. 2008.

[7] Y. Jia *et al.*, "Deep-sea sediment resuspension by internal solitary waves in the northern South China Sea," *Sci. Rep.*, vol. 9, no. 1, 2019, Art. no. 12137.

[8] M. Villamaña *et al.*, "Role of internal waves on mixing, nutrient supply and phytoplankton community structure during spring and neap tides in the upwelling ecosystem of Ría De Vigo (NW Iberian Peninsula)," *Limnol. Oceanogr.*, vol. 62, no. 3, pp. 1014–1030, 2017.

[9] H. Lü, J. Xie, J. Xu, Z. Chen, T. Liu, and S. Cai, "Force and torque exerted by internal solitary waves in background parabolic current on cylindrical tendon leg by numerical simulation," *Ocean Eng.*, vol. 114, pp. 250–258, 2016.

[10] F.-C. Su, C.-R. Ho, Q. Zheng, and N.-J. Kuo, "Estimating amplitudes of internal waves using satellite ocean colour imagery of the South China Sea," *Int. J. Remote Sens.*, vol. 29, no. 21, pp. 6373–6380, 2008.

[11] D. Di, X. Yang, X. Li, and Z. Li, "SAR observation of eddy-induced mode-2 internal solitary waves in the South China Sea," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 11, pp. 6674–6686, Nov. 2016.

[12] Q. Zheng, R. D. Susanto, C. Ho, Y. T. Song, and Q. Xu, "Statistical and dynamical analyses of generation mechanisms of solitary internal waves in the northern South China Sea," *J. Geophys. Res.-Oceans*, vol. 112, no. C3, pp. 1–16, Mar. 2007.

[13] R. B. Perry and G. R. Schimke, "Large-amplitude internal waves observed off the northwest coast of sumatra," *J. Geophys. Res.*, vol. 70, no. 10, pp. 2319–2324, 1965.

[14] V. Vlasenko and N. Stashchuk, "Three-dimensional shoaling of large-amplitude internal waves," *J. Geophys. Res.-Oceans*, vol. 112, no. C11, pp. 1–15, Nov. 2007.

[15] A. R. Osborne and T. L. Burch, "Internal solitons in the Andaman Sea," *Science*, vol. 208, no. 4443, pp. 451–460, 1980.

[16] L. Sun, J. Zhang, and J. Meng, "A study of the spatial-temporal distribution and propagation characteristics of internal waves in the Andaman Sea using MODIS," *Acta Oceanolog. Sin.*, vol. 38, no. 7, pp. 121–128, 2019.

[17] J. Magalhaes and J. da Silva, "Internal solitary waves in the Andaman Sea: New insights from SAR imagery," *Remote Sens.*, vol. 10, no. 6, 2018, Art. no. 861.

[18] W. Alpers and W. Huang, "On the discrimination of radar signatures of atmospheric gravity waves and oceanic internal waves on synthetic aperture radar images of the sea surface," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 3, pp. 1114–1126, Mar. 2010.

[19] J. Magalhaes, J. da Silva, and M. C. Buijsman, "Long lived second mode internal solitary waves in the Andaman Sea," *Sci. Rep.*, vol. 10, no. 1, pp. 1–10, 2020.
[20] X. Zhang, X. Li, and T. Zhang, "Characteristics and generations of internal wave in the Sulu sea inferred from optical satellite images," *J. Oceanol. Limnol.*, vol. 38, no. 5, pp. 1435–1444, 2020.
[21] X. L. Bai, X. F. Li, K. G. Lamb, and J. Y. Hu, "Internal solitary wave reflection near Dongsha Atoll, the South China Sea," *J. Geophys. Res.-Oceans*, vol. 122, no. 10, pp. 7978–7991, 2017.
[22] X. Li, C. R. Jackson, and W. G. Pichel, "Internal solitary wave refraction at Dongsha Atoll, South China Sea," *Geophys. Res. Lett.*, vol. 40, no. 12, pp. 3128–3132, 2013.
[23] Z. Zhang, O. B. Fringer, and S. R. Ramp, "Three-dimensional, nonhydrostatic numerical simulation of nonlinear internal wave generation and propagation in the South China Sea," *J. Geophys. Res.-Oceans*, vol. 116, no. C5, pp. 1–26, May 2011.
[24] Z. Zhao and M. H. Alford, "Source and propagation of internal solitary waves in the northeastern South China Sea," *J. Geophys. Res.-Oceans*, vol. 111, no. C11, pp. 1–14, Nov. 2006.
[25] P. Hyder, D. Jeans, E. Cauquil, and R. Nerzic, "Observations and predictability of internal solitons in the northern Andaman Sea," *Appl. Ocean Res.*, vol. 27, no. 1, pp. 1–11, 2005.
[26] Y. G. Ham, J. H. Kim, and J. J. Luo, "Deep learning for multi-year ENSO forecasts," *Nature*, vol. 573, no. 7775, pp. 568–572, 2019.
[27] G. Xu *et al.*, "Oceanic eddy identification using an AI scheme," *Remote Sens.*, vol. 11, no. 11, 2019, Art. no. 1349.
[28] L. Zhang, L. Zhang, and B. Du, "Deep learning for remote sensing data: A technical tutorial on the state-of-the-art," *IEEE Trans. Geosci. Remote Sens.*, vol. 4, no. 2, pp. 22–40, Jun. 2016.
[29] M. Krestenitis, G. Orfanidis, K. Ioannidis, K. Avgerinakis, S. Vrochidis, and I. Kompatsiaris, "Oil spill identification from satellite images using deep neural networks," *Remote Sens.*, vol. 11, no. 15, pp. 1–22, Jul. 2019, Art. no. 1762.
[30] B. Liu, X. Li, and G. Zheng, "Coastal inundation mapping from bitemporal and dual-polarization SAR imagery based on deep convolutional neural networks," *J. Geophys. Res.-Oceans*, vol. 124, no. 12, pp. 9101–9113, 2019.
[31] Y. Lee, D. Han, M. Ahn, J. Im, and S. Lee, "Retrieval of total precipitable water from Himawari-8 AHI data: A comparison of random forest, extreme gradient boosting, and deep neural network," *Remote Sens.*, vol. 11, no. 15, 2019, Art. no. 1741.
[32] X. Pan, J. Wang, X. Zhang, Y. Mei, L. Shi, and G. Zhong, "A deep-learning model for the amplitude inversion of internal waves based on optical remote-sensing images," *Int. J. Remote Sens.*, vol. 39, no. 3, pp. 607–618, 2018.
[33] X. Li *et al.*, "Deep learning-based information mining from ocean remote sensing imagery," *Nat. Sci. Rev.*, vol. 7, pp. 1584–1605, 2020.
[34] G. Zheng, X. Li, R.-H. Zhang, and B. Liu, "Purely satellite datadriven deep learning forecast of complicated tropical instability waves," *Sci. Adv.*, vol. 6, no. 29, pp. 1–9, Jul. 2020. [Online]. Available: https://advances.sciencemag.org/content/6/29/eaba1482
[35] X. Zhang and X. Li, "Combination of satellite observations and machine learning method for internal wave forecast in the Sulu and Celebes seas," *IEEE Trans. Geosci. Remote Sensing*, to be published, doi: 10.1109/TGRS.2020.3008067.
[36] T. Lee, "Back-propagation neural network for long-term tidal predictions," *Ocean Eng.*, vol. 31, no. 2, pp. 225–238, 2004.
[37] Q. Zheng, Y. Yuan, V. Klemas, and X. Yan, "Theoretical expression for an ocean internal soliton synthetic aperture radar image and determination of the soliton characteristic half width," *J. Geophys. Res.-Oceans*, vol. 106, no. C12, pp. 31 415–31 423, 2001.
[38] B. M. Wilamowski and H. Yu, "Improved computation for Levenberg-Marquardt training," *IEEE T. Neural Networ.*, vol. 21, no. 6, pp. 930–937, Jun. 2010.
[39] L. Prechelt, "Automatic early stopping using cross validation: Quantifying the criteria," *Neural Netw.*, vol. 11, no. 4, pp. 761–767, 1998.
[40] H. Alt and M. Godau, "Computing the Fréchet distance between two polygonal curves," *Int. J. Comput. Geom. Ap.*, vol. 5, no. 01–02, pp. 75–91, 1995.
[41] X. Zhang, Z. Jie, C. Fan, J. Wang, and J. Meng, "Observations of internal waves with high sampling data of radar altimetry and MODIS images," *Int. J. Remote Sens.*, vol. 39, no. 21, pp. 7405–7416, 2018.
[42] B. Q. Liu, H. Yang, Z. X. Zhao, and X. F. Li, "Internal solitary wave propagation observed by tandem satellites," *Geophys. Res. Lett.*, vol. 41, no. 6, pp. 2077–2085, 2014.
[43] P. Wang, X. Wang, J. Chong, and Y. Lu, "Optimal parameter estimation method of internal solitary waves in SAR images and the Cramér-Rao bound," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 6, pp. 3143–3150, Jan. 2016.

Xudong Zhang (Member, IEEE) received the B.S. degree in optical science and technology from the Ocean University of China, Qingdao, China, in 2013, and the Ph.D. degree in marine technology from the joint program of the Ocean University of China and the First Institute of Oceanology, Ministry of Natural Resources, Qingdao, China, in 2018.

He has been holding a Postdoctoral position at the Institute of Oceanology, Chinese Academy of Sciences, Qingdao, China, since 2019. His research interests include internal waves, remote sensing, and machine learning.

Xiaofeng Li (Fellow, IEEE) received the B.S. degree in optical engineering from Zhejiang University, Hangzhou, China, in 1985, the M.S. degree in physical oceanography from the First Institute of Oceanography, Qingdao, China, in 1992, and the Ph.D. degree in physical oceanography from North Carolina State University, Raleigh, NC, USA, in 1997.

He had been with the National Environmental Satellite, Data, and Information Service, National Ocean and Atmospheric Administration, USA, from 1997 to 2019. His research interests include synthetic aperture radar (SAR) applications in oceanography and marine meteorology, artificial intelligence oceanography, big data, and satellite image processing.

Dr. Li is an Associate Editor for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING and the *International Journal of Remote Sensing*. He is an Editorial Board Member of the *International Journal of Digital Earth*, *Big Earth Data*, and the *Journal of Oceanology and Limnology*. He is also the Executive Editor-in-Chief of the *Journal of Remote Sensing* (a science partner journal).

Quanan Zheng received the diploma in physics from Jilin University, Changchun, China, in 1966, and the Ph.D. degree in physical oceanography from Institute of Oceanology, Chinese Academy of Sciences, in 1987.

He is a Research Scientist Emeritus with the Department of Atmospheric and Oceanic Science, University of Maryland, College Park, MD, USA. His research interests include physical oceanography, satellite oceanography, and radar remote sensing.

He is an Associate Editor-in-Chief of the *Acta Oceanonogica Sinica* and the *Journal of Ocean University of China*, and an Editorial Board Member of *Scientific Reports*. He is a member of the American Geophysical Union, and a life member of the Chinese-American Oceanic and Atmospheric Association (COAA).