An APMLP Deep Learning Model for Bathymetry Retrieval Using Adjacent Pixels

Jinshan Zhu, Jian Qin, Fei Yin, Zhaoyu Ren, Jiawei Qi, Jingyu Zhang, and Ruifu Wang

***Abstract*—Shallow water depth plays an important role in marine development, navigation safety, and environmental protection. It is an efficient and economical way to obtain water depth by remote sensing technology. At present, most empirical models based on multispectral image usually obtain water depth by the relationship between the sea surface reflectance (SSR) (a single pixel) and *in situ* water depth, it is a one-to-one correspondence between the reflectance and depth. However, seafloor substrate and inherent optical properties (IOP) will also have contribution to the SSR. In this article, we propose an adjacent pixels multilayer perceptron model (APMLP) model using adjacent pixels to weaken the influence of seafloor substrate and IOP. Datasets on Oahu Island (Sentinel-2B, LIDAR *in situ* data) and Saint Thomas Island (Sentinel-2A, LIDAR *in situ* data) are used to establish and verify the model. The APMLP model are also compared with the multilayer perceptron model (MLP) model, BP neural network model, and Log-ratio model. The overall root-mean-square error (RMSE) of APMLP model on Oahu Island is 0.72 m, which is much better than the other three models (MLP 1.07 m, BP 1.05 m, Log-ratio 1.52 m). Similar results are obtained from the Saint Thomas Island dataset, RMSE of APMLP model is 1.56 m, better than the other three (MLP 1.91 m, BP 1.89 m, Log-ratio 2.39 m). The study confirms that considering adjacent pixels in an artificial neural network model can effectively improve the performance of water depth retrieval.**

***Index Terms*—Bathymetry, neural networks, remote sensing.**

I. Introduction

SHALLOW water depth is important marine hydrological information that plays a guiding role in chart drawing, coastal engineering, marine resource development, and management. Traditional bathymetric surveys are based on field surveys, mainly using ship-borne sonar or airborne lidar [1]. These survey methods have the advantage of high measurement accuracy, but will consume considerable manpower and material resources and may be easily affected by sea conditions [2]–[4]. In sea areas with complex terrain or disputes of interest, it is often difficult to conduct *in site* water depth detection.

Compared with traditional airborne or shipborne surveys, inversion technology based on remote sensing images (RSI) provides higher spatial coverage and temporal resolution at a lower cost and overcomes many obstacles of traditional methods [5]. Although the detectable depth and accuracy level are lower than those based on ships and lidar, remote sensing technology is still very attractive [6], [7]. In recent decades, many water depth inversion methods based on RSI have been proposed [8]–[13]. Most of these methods use the visible bands (420–780 nm) for water depth inversion. These methods can be roughly divided into two categories: physical models based on hyperspectral remote sensing images (HSIs) and empirical models based on multispectral remote sensing images (MSIs) [14].

Physical models can invert the water depth without in-situ data. Previous studies have found that the semianalytical model (a physical model) can obtain relatively accurate shallow water depths from HSIs [15]–[17]. However, there are some difficulties in the practical application. It is often hard to find an appropriate HSIs for a selected location. Additionally, the spatial resolution of HSIs is usually not very high. For example, the spatial resolution of ADEOS-2 is 250 m, that of HJ-1A is 100 m and EnMAP is 30 m. When the coastal shallow water area has complex seabed topography and large depth fluctuation, RSI with lower spatial resolution will not be suitable for high-precision water depth inversion.

These difficulties of HSIs rarely occur on MSIs. There are many satellites carrying multispectral sensors, such as Sentinel-2, QuickBird, Worldview-2, IKONOS, and Landsat-5. These data are very easy to be obtained, and the spatial resolution is usually higher than that of HSIs. For example, the spatial resolution of Sentinel-2 is 10 m and that of QuickBird is 0.6 m. Disadvantage of MSIs is that their number of bands is far less than that of HSIs, there are usually no more than five visible bands, which can be used for water depth inversion.

Compared with the physical models, the empirical models have low requirements for the number of bands. Many experts have achieved successful water depth retrieval using MSIs [18]–[20]. At present, there are several commonly used empirical models. For example, Stumpf *et al.* [21] proposed a Log-ratio model (LG) in 2003, in this model Log-ratio of the blue band and green band is used to invert water depth, the form of this model is relatively simple, and only two or three model parameters need to be tuned. This model is used widely until nowadays for its

Manuscript received September 8, 2021; revised October 28, 2021, November 18, 2021, and December 3, 2021; accepted December 6, 2021. Date of publication December 9, 2021; date of current version December 27, 2021. This work was supported by the Open Research Foundation of Funded by Key Laboratory of Ocean Geomatics, Ministry of Natural Resources, China under Grant 2021B06 and in part by the Key projects of National Natural Science Foundation of China - Research on Evolution Mechanism of islands and reefs in the South China Sea and multiscale simulation and prediction technology based on Data Mining under Grant 51839002. *(Corresponding author: Ruifu Wang.)*

Jinshan Zhu, Jian Qin, Fei Yin, Zhaoyu Ren, Jiawei Qi, and Ruifu Wang are with the College of Geodesy and Geomatics, Shandong University of Science and Technology, Qingdao 266590, China, and also with the Key Laboratory of Ocean Geomatics, Ministry of Natural Resources of China, Qingdao 266590, China (e-mail: zhujinshan@sdust.edu.cn; wrf@sdust.edu.cn).

Jingyu Zhang is with the First Institute of Oceanography MNR, Qingdao 266061, China (e-mail: zhangjingyu@fio.org.cn).

Digital Object Identifier 10.1109/JSTARS.2021.3134013

simplicity and ease of use [22]–[26]. Wang *et al.* [27] tried to use a back propagation neural model (BP) for water depth retrieval, their article proved that the BP neural network can be used to invert water depth, but there may have problems (error-large and change trend-inconsistency) when the water is deeper than 10 m. The reason for these problems may be that the amount of the training samples is not large enough. A multilayer perceptron model (MLP) is used for water depth inversion by Wang *et al.* [28]. Their article proved that the spectral features alone are insufficient for water bathymetry retrieval, except the spectral features, spatial location feature is also integrated into the model, and this can address the problems caused by heterogeneous bottom types well.

At present, most empirical models will establish a relationship between the sea surface reflectance (SSR) of the target pixel and its *in situ* water depth. The SSR and water depth has a one-to-one relationship in the model. But in fact, as proved in Wang's research [28], in most cases the relationship between them is not only one-to-one. The seafloor substrate and water column inherent optical properties (IOP) may also have contribution to the SSR, the SSR and the water depth many have a many-to-one relationship. Usually, the water depth difference between adjacent pixels is not so large, but their seafloor substrate and IOP may not be identical, so we consider to introduce the adjacent pixels into our model in this article and use the multiple SSRs from the target pixel and the adjacent pixels to correspond to one water depth, and try to improve the inversion accuracy.

In this article, an adjacent pixels multilayer perceptron model (APMLP) is proposed. In this model, the adjacent pixels are taken into account. One target pixel and the surrounding eight pixels are assigned as a small region (SR). We assume that the water depth of these nine pixels inside one SR is the same. This SR is used to instead the one pixel in the commonly used empirical models. When training the APMLP model, each water depth will correspond to nine different SSRs, which can help to solve the problem of one-to-one relationship between SSR and water depth, and weaken the influence of seafloor substrate and IOP.

The original contribution of this article is that a SR is used to instead the one single pixel in other empirical models, so that each *in situ* water depth will correspond to nine different SSRs. This will avoid the one-to-one relationship between the SSR and water depth. Information not only about the water depth, but also about seafloor substrate and IOP will be introduced into the model. According to our experiment, our strategy using a SR in a MLP machine learning model is very effective to improve the water depth inversion accuracy comparing with other models.

II. Materials

A. Study Areas

The first study area is the area surrounding Diamond Head Beach Park, located approximately 8.2 km southeast of Honolulu, Oahu Island, Hawaiian Islands, in the North Pacific [see Fig. 1(a)], hereinafter, we will call it Honolulu for short. The surrounding underwater sediment is complex, and the seawater quality and transparency are good.

Another study area is located between Saint Thomas Island and Saint John Island in the U.S. Virgin Islands [see Fig. 1(c)], hereinafter we will call it Saint Thomas Island for short. There are many small islands in the study area, and deep-water areas and shallow-water areas are staggered with each other.

B. Satellite Images and In Situ Depth Data

A Sentinel-2B multispectral image over Honolulu acquired on December 1, 2020 [see Fig. 1(b)] and A Sentinel-2A multispectral image over Saint Thomas Island acquired on January 15, 2019 [see Fig. 1(d)] were used to invert shallow water depths. They are the L2A-level (bottom-of-atmosphere corrected reflectance) data obtained from the official website of the European Space Agency.[1] This means that they have been processed by orthorectification, subpixel geometric correction, and atmospheric correction. There are 12 bands in the remote sensing image with three spatial resolutions, and only bands B1–B5 can be used for water depth inversion. The image scene quality is excellent, with no broken waves, no clouds, and minimal levels of wave-induced sun glint at the sea surface.

A total of 30 000 water depth points were selected from the scanning hydrographic operational airborne lidar survey (SHOALS) lidar bathymetry data [see Fig. 2(a)] acquired from the website of the University of Hawaii, Manoa.[2] Another 30 000 water depth points were selected from lidar bathymetry data [see Fig. 2(b)] acquired from the website of NOAA.[3] They are used as *in situ* depth data for Honolulu and Saint Thomas Island, respectively. Seventy percent of them were used for network training, and the other 30% were used to test the model.

C. Image Processing

Sentinel-2 image has 12 bands, but only the visible to near infrared bands B1–B5 (411–714 nm) are useful for water depth inversion. However, the spatial resolution of band B1 is 60 m, that of band B5 is 20 m and bands B2–B4 is 10 m. To solve the problem of mismatched spatial resolution, spatial resolution of bands B1 and B5 are improved into 10 m. ESA offers a Sen2res plug-in[4] on SNAP software,[5] which provides us a way to improve the spatial resolution.

III. Methodology

For shallow water, a simplified shallow water radiation transfer equation is proposed by Maritorena *et al.* [29], Philpot [30]. The equation is formulated as

$$R_w = R_\infty + (A_d - R_\infty)\exp(-2kH) \tag{1}$$

where R_w is SSR; R_∞ is the surface reflectance of the optically deep water; A_d is bottom albedo, which is affected by seafloor substrate; $2k$ is the effective two-ways attenuation coefficient, which is affected by IOP, and H is the depth of water.

[1] [Online]. Available: https://scihub.copernicus.eu/dhus/#/home

[2] [Online]. Available: http://www.soest.hawaii.edu/coasts/data/hawaii/shoals.html

[3] [Online]. Available: https://www.ncei.noaa.gov/maps/bathymetry/

[4] [Online]. Available: http://step.esa.int/main/third-party-plugins-2/sen2res/

[5] [Online]. Available: http://step.esa.int/main/download/

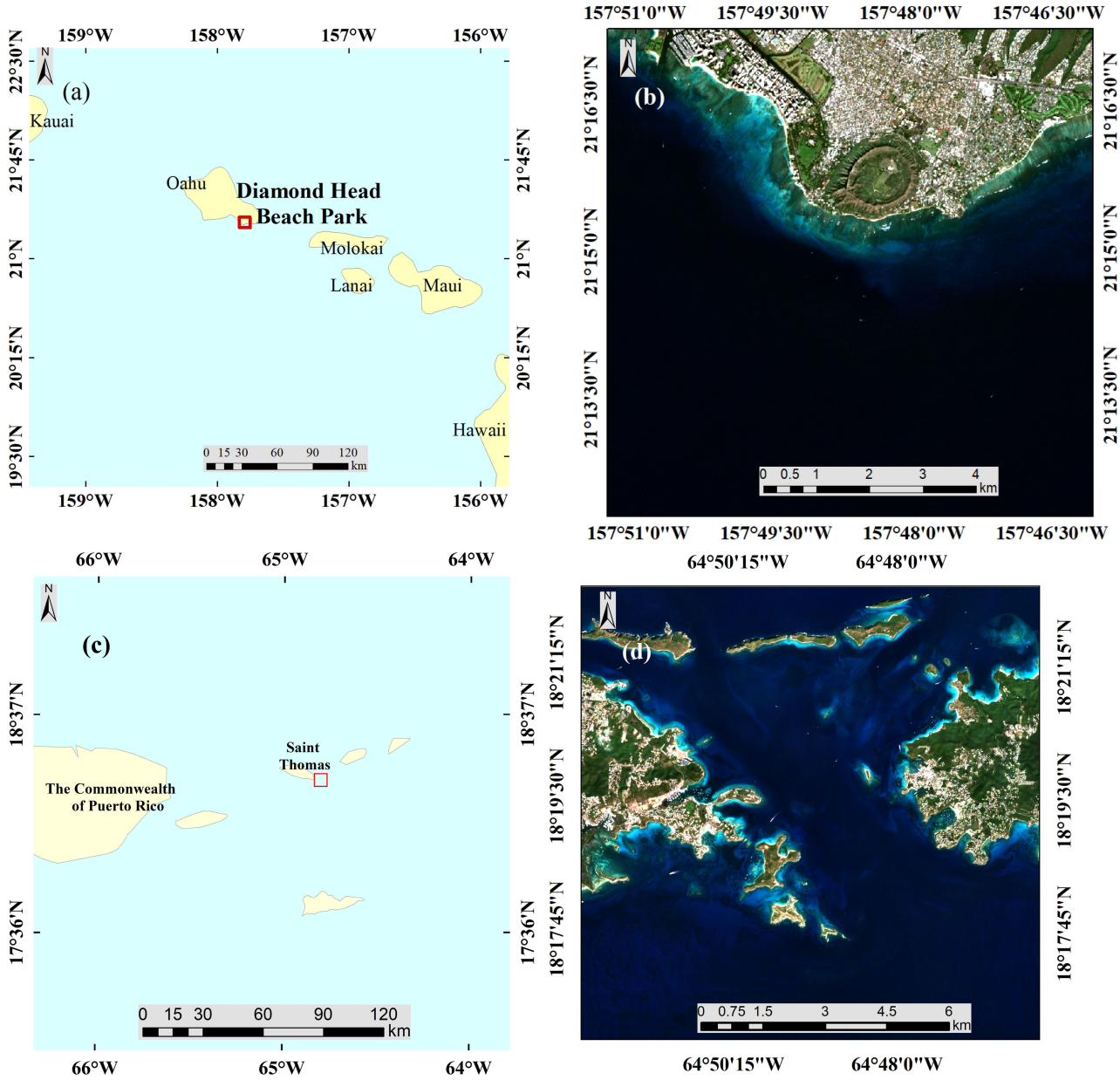
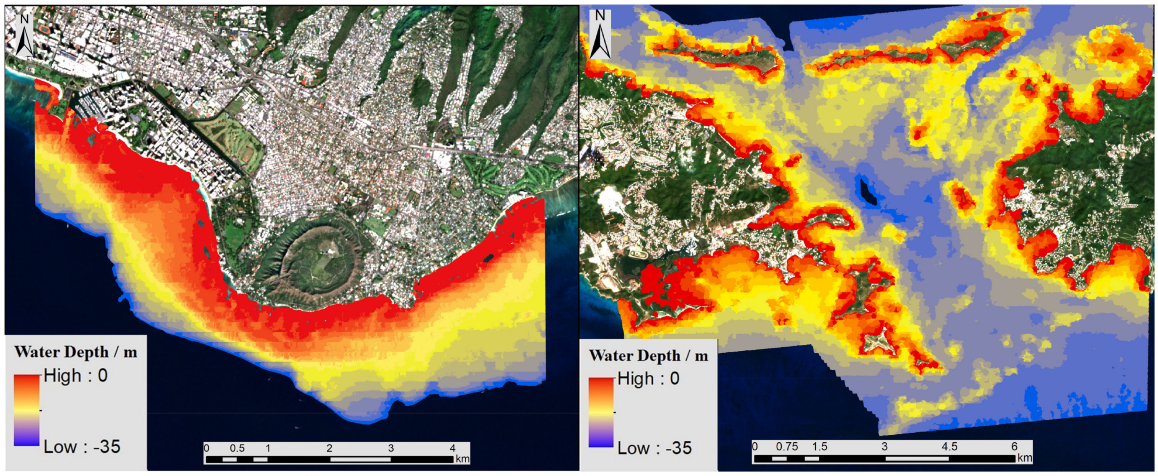


Fig. 1. (a) Study sites for diamond head beach park. (b) Saint Thomas Island. (c) Sentinel-2B multispectral image over diamond head beach park acquired on December 1, 2020. (d) Sentinel-2A multispectral image over Saint Thomas Island acquired on January 15, 2019.

According to (1), SSR is mainly determined by seafloor substrate, IOP and water depth. If we classify the seafloor substrate in advance, invert the water depth, respectively, according to the seafloor substrate type, for each training, the seafloor substrate type of the samples input to the model is the same, i.e., the change of A_d in (1) is very small. According to (1), when the change of A_d is very small, SSR is mainly determined by water IOP and depth, so the influence of seafloor substrate on SSR can be weakened. When the seafloor substrate type of samples is the same, the model does not have to learn the effects of different seafloor substrate types on SSR, which reduces the time required to train the model. So, it is necessary to conduct seafloor substrate classification in advance.

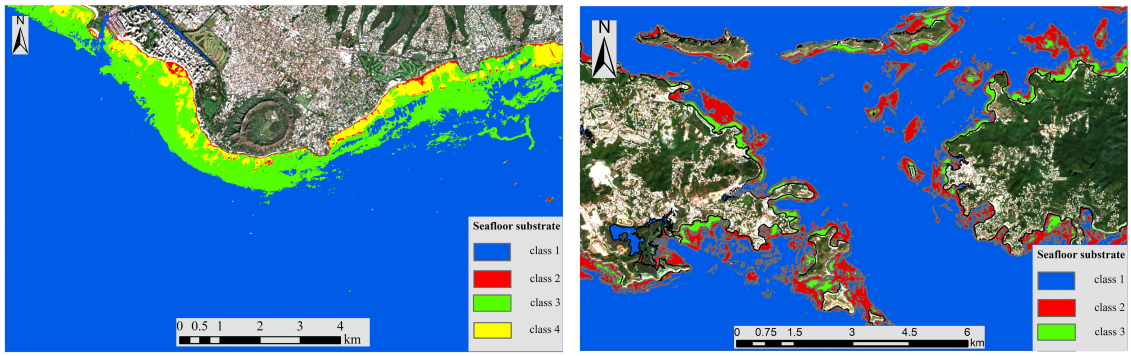
The seafloor substrate of study areas is complex and there are many mixed pixels. It is hard to know the types of seafloor substrate exactly, or even how many kinds. Compared with other algorithms, ISODATA does not need to determine the number of categories in advance, and the algorithm will automatically adjust the number of categories and category centre, which is very suitable for current situation. The parameters of ISODATA algorithm are listed in Table I. Seafloor substrate is classified into four types (Honolulu) and three types (Saint Thomas Island). The classification results are shown in Fig. 3.

As shown in Fig. 4 row A, the blue square on the left is one pixel, the line chart in the middle shows the SSR spectrum of one pixel, and H on the right is the corresponding water depth of



(a) (b)

Fig. 2. Spatial distribution of water depth points. They are used to train and test models. (a) Honolulu. (b) Saint Thomas Island.



(a) (b)

Fig. 3. Classification results of seafloor substrate in study areas. (a) Honolulu and (b) Saint Thomas Island.

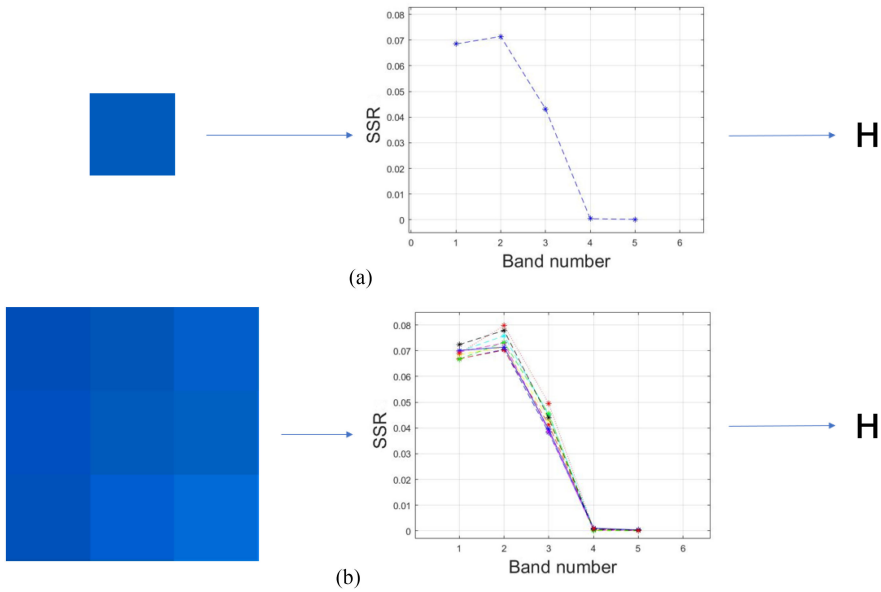


Fig. 4. Main principles of several current empirical models and APMLP model. Models using one pixel (row A); APMLP (row B).

TABLE I
PARAMETER SETTINGS AND VALUES OF THE MODEL

Parameter name	Parameter setting or value
Max number of classes	7
Min number of classes	3
Max iterations	10
Min pixel in class	1
Max class stdv	1
Min class distance	5
Max merge pairs	2

TABLE II
MSES OF TRAINING DATA AND TEST DATA IN DIFFERENT STRUCTURES

Layers of model structure	6	7	**8**	9	10
MSE of training data	0.255	0.250	**0.238**	0.244	0.249
MSE of test data	0.289	0.311	**0.258**	0.293	0.275

left pixel. For a commonly used empirical model, in the training dataset, each sample consists of an SSR spectrum and a water depth, this means it is a one-to-one correspondence relationship. When the model is trained by this kind of dataset, the model will learn a one-to-one relationship between the SSR and *in situ* depth only. However, according to (1), the SSR is also related to seafloor substrate and IOP. So, in our APMLP model, we take the seafloor substrate and water column IOP into account, and established a new dataset. As shown in Fig. 4 row B, on the left is pixels in one SR, in the middle is their SSRs. For our APMLP model, in the new dataset each sample consists of these nine SSRs and one water depth.

It can be observed through visual interpretation that the colors of these nine pixels are not exactly the same, and they have nine different SSRs. According to (1), there are three factors, water depth, seafloor substrate, and IOP, can influent the shape of the SSR spectrum. So, these nine different SSRs mean at least one of the three factors is different. In our APMLP model, we assume that the water depth of these nine pixels inside SR is the same, and replace the SSR of the target pixel (commonly used empirical model) with the nine SSRs in one SR (APMLP model). This makes nine different SSRs correspond to one water depth, and this is a many-to-one relationship between SSRs and water depth. At the same time, when training the model using this kind of dataset, the model will learn the following fact: if the water depth is the same, the changes of seafloor substrate, and IOP will also have contribution to the SSR.

We use MLP as the main network structure of APMLP. MLP simulates the working process of brain nerve tissue, and it comprises an input layer, hidden layers, and an output layer [31]–[33]. In order to determine the appropriate number of network layers, we used some data to test network with different layers, and some of the results are listed in Table II. According to this table, model with eight layers performed best. So, a network structure with eight layers is adopted for our APMLP.

APMLP is built under Pytorch deep learning framework, and the structure is shown as Fig. 5. The number of nodes in the input

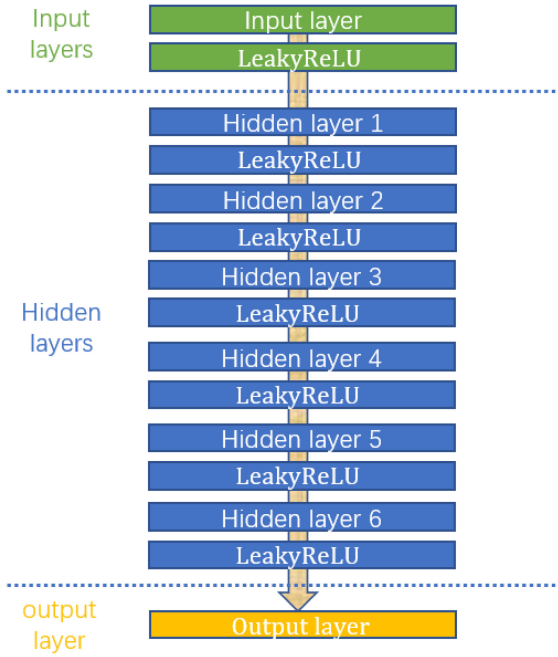


Fig. 5. Network structure of the APMLP. It contains an input layer, six hidden layers, and an output layer.

layer is 45, for the hidden layers, the number of nodes is 180, 180, 60, 30, 30, 10, respectively, and the output layer has only 1 node. The nodes in each layer (except output layer) is connected to all nodes in the next layer. Activation functions are added between all layers to ensure network has enough nonlinear fitting ability. Each SR contains nine pixels, each pixel corresponds to one SSR spectrum, each SSR spectrum has five bands. So, each sample of our training dataset has 45 SSR values, and they correspond to 45 nodes of input layer. There is only one node in the output layer, it's value equal to the predicted water depth.

To balance performance and time costs, we chose Leakyrelu as the activation function, which can be expressed as (2). Leakyrelu uses a α parameter to give all negative values a nonzero slope so that the network can transfer the gradient of the negative part and avoid the death of neuron nodes. Because it does not include an exponential operation, it converges faster than the sigmoid and tanh activation function

$$\text{LeakyReLU (x)} = \begin{cases} x, & x \geq 0 \\ \alpha_{\text{negative}_{\text{slope}}} x, & x < 0 \end{cases}. \tag{2}$$

The mean square error is used as the loss function of APMLP and it can be expressed as follows:

$$\text{loss} = \frac{1}{N} * \sum_{i=1}^{N} (y_i - \widehat{y_i})^2 \tag{3}$$

where N is the number of training samples, y_i is the *in situ* data, and $\widehat{y_i}$ is the predicted value.

We choose Adam as the optimization algorithm. It combines the advantages of Adagrad and RMSprop and assigns independent adaptive learning rates for different parameters by

TABLE III
PARAMETER SETTINGS AND VALUES OF THE MODEL

Parameter name	Parameter setting or value
Activation function	LeakyReLU
α	0.01
Learning rate	0.0001
Iterations	3000
Optimization algorithm	Adam
Beta1	0.9
Beta2	0.999

comprehensively considering the first moment estimation and second moment estimation of the gradient [34].

During training, APMLP will receive 45 SSR values from one SR and output a predicted water depth value. The loss function will calculate the difference between the predicted and *in situ* water depth and feed it to optimization function. The optimization function will continuously adjust the weight and bias of each layer, so as to keep the results of the model close to *in situ* value gradually. All parameter settings or values of APMLP are shown in Table III.

IV. RESULTS

In this article, two datasets, Honolulu and Saint Thomas Island, were used to train and verify our APMLP model. The result was also compared with BP, MLP, and LG model. The performance of each model was assessed by five statistical parameters, i.e., the correlation coefficient (r), mean relative error (MRE) root-mean-square error (RMSE), bias ($\bar{\varepsilon}$) and absolute percentage difference ($|\bar{\varepsilon}|$), which can be expressed as

$$r = \frac{1}{N-1} * \sum_{i=1}^{N} \left(\frac{A_i - \mu_A}{\sigma_A} \right) * \left(\left(\frac{B_i - \mu_B}{\sigma_B} \right) \right) \quad (4)$$

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{N} (A_i - B_i)^2}{N}} \quad (5)$$

$$\mathrm{MRE} = \frac{1}{N} * \sum_{i=1}^{N} \frac{|A_i - B_i|}{A_i} \quad (6)$$

$$\bar{\varepsilon} = \mathrm{median} \left(\frac{B_i - A_i}{A_i} \right) * 100\% \quad (7)$$

$$|\bar{\varepsilon}| = \mathrm{median} \left(\left| \frac{B_i - A_i}{A_i} \right| \right) * 100\% \quad (8)$$

where A_i and B_i refer to the *in situ* and estimated water depth; μ_A and σ_A are the mean and standard deviation of A and μ_B and σ_B are that of B, respectively; N is the number of test data points.

A. Results of Honolulu

Fig. 6 is compares of the results from the four models (Honolulu) and the *in situ* measured water depth, from Fig. 6(a) to (d) is for APMLP, MLP, BP, and LG, respectively. According to Fig. 6, when the water depth is lower than 7 m, for all of the four models, the inversion results fit the measured depth very well; when the water depth is 7–22 m, the dispersion degree has a slightly increase; but when the water depth is more than 22 m, the dispersion degree increases. By comparing the four models, APMLP is the best, it has the highest correlation coefficient (0.99573), which indicating that the predicted depth of APMLP is very close to the measured depth. The $\bar{\varepsilon}$ and $|\bar{\varepsilon}|$ of all of the four models are quite small, which means small bias. We can also find that, in Fig. 6(b) to (d), there are many outliers far from diagonal line, for example, in the lower right corner of Fig. 6(c), a point with a measured water depth about 34 m is incorrectly estimated as 6 m, the error is very large. However, outliers with so large error do not appear in Fig. 6(a), which proves once again that by adding adjacent pixels, the performance of the model can be improved significantly.

In order to further compare the performance of the models at different depth ranges, we divided water depths into three intervals: 0–7 m, 7–22 m, and 22–35 m according to the distribution of points in Fig. 6. The results are summarized in Table IV. As Table IV, when water depth increases, the accuracy of the four models decreases obviously. For example, when water depth exceeds 22 m, the RMSE of APMLP is 2.04 m, compare to that in 0–7 m (0.39 m), it has an increase of 423%. The overall RMSE of APMLP (0.72 m) is the lowest of the four models, which is 0.8 m lower than the LG model (the worst performing, 1.52 m). In the water depth range of 0–7 m, the MRE of APMLP is 24.05%, which is 8.01% lower than that of the MLP model (the second place, 32.06%). The performance of APMLP is still the best in the range of 7 to 22 m. In the range of 22–35 m, although the RMSE of APMLP exceeds 2 m, it is still the lowest among the four models.

The accuracy of water depth inversion can be affected by many factors, usually, the spatial distribution of the absolute of prediction residuals(APR) is uneven. The spatial distribution of APR for the Honolulu dataset is plotted In Fig. 8, from Fig. 8(a) to (d) is for the APMLP, MLP, BP, and LG, respectively. The APR is divided into nine segments, and each segment is represented by small dots with different sizes and colors. The smaller and bluer the dot is, the smaller the APR. In contrast, the larger and redder the dot is the larger the APR. As shown from Fig. 8(a)–(d), there are many blue or light blue small dots near the coastline. With the increase of offshore distance, yellow dots begin to appear, and there are a few red dots far away from the coastline. All of these cases suggests that the farther away from the coast, the larger the APR. The number of yellow dots near the coastline in Fig. 8(a) is significantly less than that in Fig. 8(b)–(d), this phenomenon is more pronounced on the left and right sides of Fig. 8(a)–(d). But in the area far from the coastline, the change between them is not obvious. The excellent performance of APMLP in shallow water area shows that adding adjacent pixels into the model can effectively improve the accuracy.

B. Results of Saint Thomas Island

Fig. 7 is compares of the results from the four models (Saint Thomas Island) and the *in situ* measured water depth, from

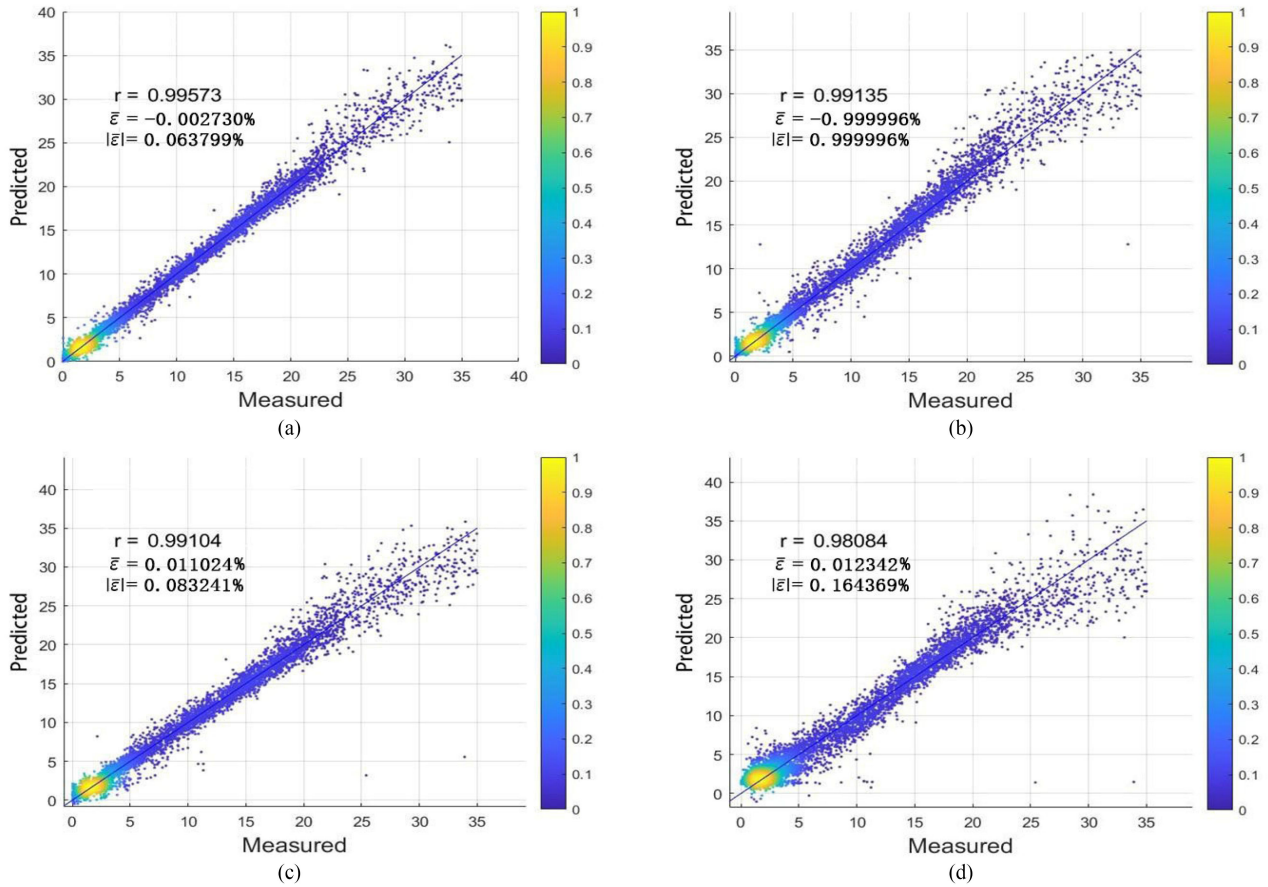


Fig. 6. Maps of estimated model values versus measured values for Honolulu. (a) APMLP model. (b) MLP model. (c) BP neural network model. (d) Log-ratio model.

TABLE IV
COMPARISON OF MODEL ACCURACY (HONOLULU)

	RMSE (m)				MRE			
	Overall	0–7 m	7–22 m	22–35 m	Overall	0–7 m	7–22 m	22–35 m
APMLP model	0.72	0.39	0.79	2.04	17.22%	24.05%	4.29%	5.77%
MLP model	1.07	0.53	1.34	2.73	23.25%	32.06%	6.74%	7.75%
BP neural network model	1.05	0.54	1.15	2.97	29.04%	41.31%	6.02%	7.63%
Log-ratio model	1.52	1.01	1.59	3.97	55.26%	80.14%	8.86%	10.08%

Fig. 7(a) to (d) is for APMLP, MLP, BP, and LG, respectively. By comparing, we can find the APMLP is the best, it has the highest correlation coefficient (0.98468), which indicating that the predicted depth is still the closest to the measured depth.

In Fig. 7, for the results of the four model, they all some outliers far from diagonal line, for example, in Fig. 7(b), a point with a measured water depth of about 3 m in the upper left corner is wrongly estimated as 33 m; in Fig. 7(c), a point with a measured value about 25 m at the bottom of is wrongly estimated as -30 m; while outliers with so large error appeared few in Fig. 7(a).

Here, the water depth is also divided into three intervals: 0-7 m, 7–22 m, and 22–35 m to compare the performance of models at different depth ranges. The results are summarized in Table V. Compared with Honolulu, the overall accuracy of the four models has decreased significantly. For example, the overall RMSE of APMLP in Saint Thomas Island (1.56 m) has increased 0.84 m comparing the same result in Honolulu (0.72 m). Although the APMLP model does not perform so well as in Honolulu, it still has the most accurate result of the four models. In the water depth range of 0–7 m, the MRE of APMLP is 37.34%, which is 10.77% lower than that of BP model (the second place, 48.11%), but it won out of four models only a in the range of 7–35 m.

The spatial distribution of APR for the Saint Thomas Island dataset is plotted In Fig. 9, from Fig. 9(a) to (d) is for the APMLP, MLP, BP, and LG, respectively. As shown in Fig. 9(b) and (c), there are a large number of light blue points in the shallow water

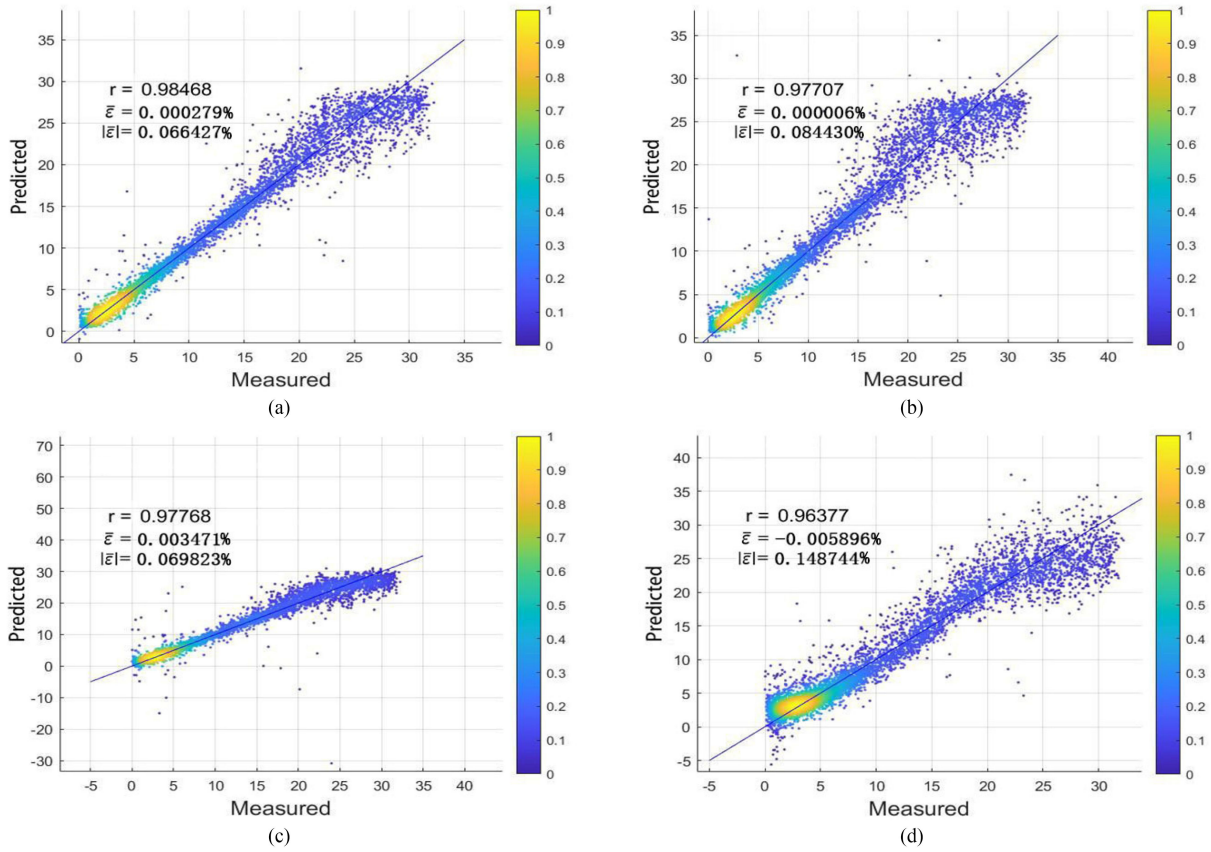


Fig. 7. Maps of estimated model values versus measured values for Saint Thomas Island. (a) APMLP model. (b) MLP model. (c) BP neural network model. (d) Log-ratio model.

TABLE V
COMPARISON OF MODEL ACCURACY (SAINT THOMAS ISLAND)

	RMSE (m)				MRE			
	Overall	0–7 m	7–22 m	22–35 m	Overall	0–7 m	7–22 m	22–35 m
APMLP model	1.56	0.67	1.51	2.93	21.88%	37.34%	6.54%	8.78%
MLP model	1.91	0.93	2.11	3.18	29.11%	50.00%	9.36%	9.63%
BP neural network model	1.89	1.12	1.75	3.36	27.33%	48.11%	7.27%	8.83%
Log-ratio model	2.39	1.59	2.34	3.95	58.26%	106.81%	12.98%	11.66%

area along the coast and in Fig. 9(d), almost all points are yellow and red. However, at the same position in Fig. 9(a), the light blue dots in coastal areas have become blue dots or dark blue dots, and the yellow dots also decreased, but the situation of red dots has hardly changed comparing with the other three Figures.

Experiment results of Saint Thomas Island proved again that adding adjacent pixels into a machine learning model is helpful to improve the inversion accuracy.

V. DISCUSSION

A. Selection Range of Adjacent Pixels

SR in the previous article is the green and blue pixels in Fig. 10, the blue pixel is the target pixel, and green pixels are called the adjacent pixels of the blue pixel. But adjacent pixels are a broad concept, all pixels surrounding blue pixel can be called adjacent pixels (even if they do not touch each other). For example, yellow and red pixels are also adjacent pixels of blue pixels. Although these adjacent pixels are farther from the blue pixel, they also have similar depths, which means that increasing the number of adjacent pixels may further improve accuracy. However, as the distance between adjacent pixels and target pixel increases, the water depth difference between them will also increase, so we cannot increase the number of adjacent pixels infinitely. How to determine a suitable number of adjacent pixels is a problem to be solved in future research.

B. Weight of Adjacent Pixels

We hope that the model can output the water depth of the target pixel, so the water depth of target pixel is used as the water

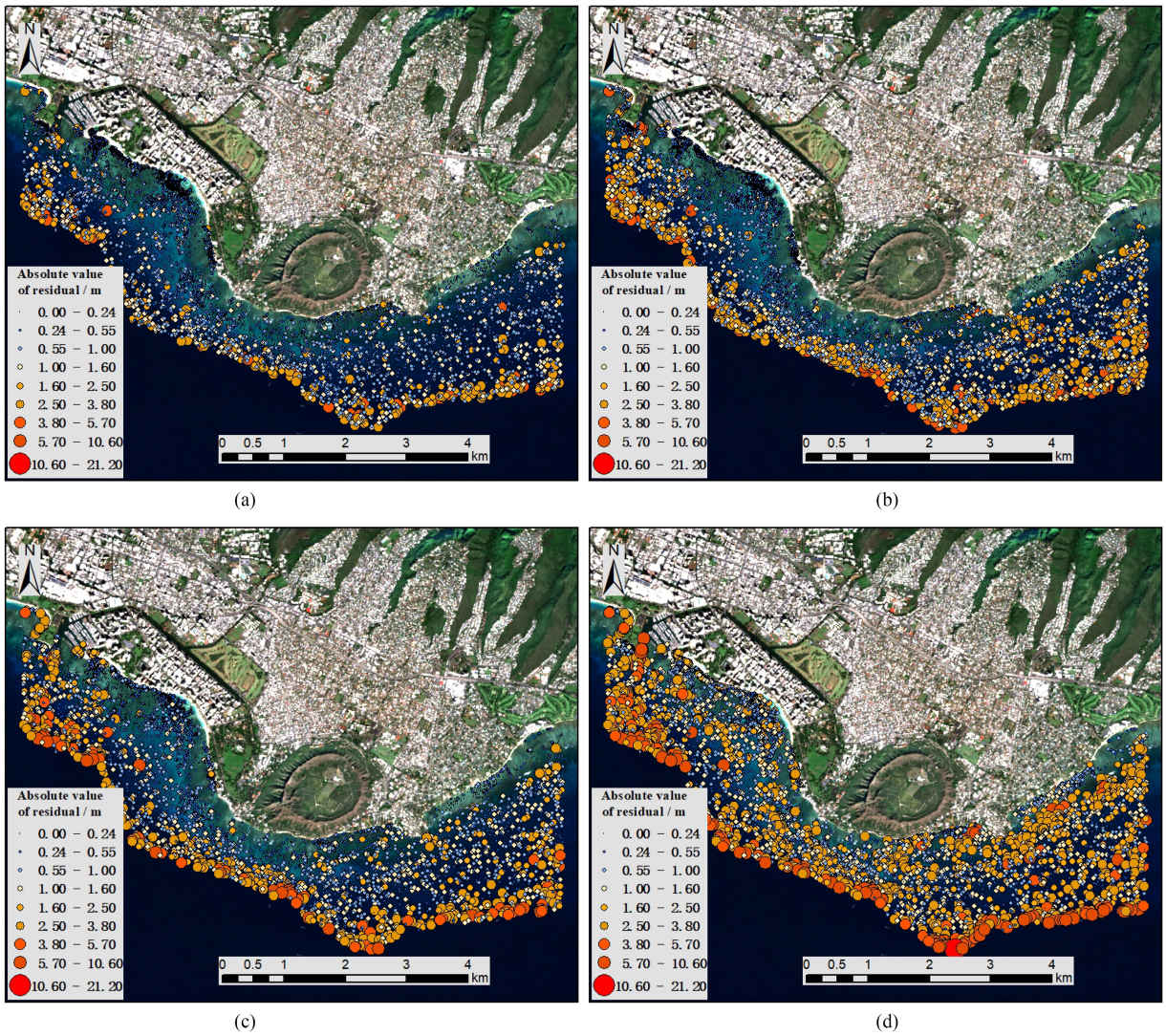


Fig. 8. Spatial distribution of APR in Honolulu. (a) APMLP model. (b) MLP model. (c) BP neural network model. (d) Log-ratio model.

depth of SR during training. However, as the distance between adjacent pixels and target pixel gradually increases, the water depth difference between them will also increase. For example, in Fig. 10, there is a difference between *in situ* depth of the blue pixel and green pixel, and the difference will increase as the distance increases. Therefore, it is unreasonable not to set weights (all pixels have the same weight as) during training.

Obviously, the weight of the blue pixel should be the highest, as the distance increases, the weight of adjacent pixels gradually decreases, and the pixels of the same color have the same weight (they have the same distance to the blue pixel). How to determine the weight of each pixel is also a problem to be solved in the future.

C. Limitation of the APMLP

In APMLP, there is an assumption: the nine pixels in one SR have the same water depth. When the spatial resolution of the image is high, the water depth change between adjacent pixels is usually very small, in this situation this assumption is reasonable. However, when the spatial resolution of the image is not high, for example, the resolution is 1 km, the water depth change between adjacent pixels may be very large, so it will be unreasonable to make this assumption. Therefore, when using our proposed APMLP model, high-resolution remote sensing images is recommended.

D. Edge Pixels in the APMLP

In APMLP, one target pixel and the surrounding eight pixels are assigned as a SR. However, when the target pixel is an edge pixel, there are less than eight adjacent pixels around the target pixel, so a SR cannot be formed. In this article, we use different processing methods to solve this problem according to the position of the target pixel.

1) Target Pixel at the Edge of the Image: As shown in Fig. 11, the blue pixel is the target pixel, and the red and green pixels are adjacent pixels. Assuming that the yellow line is the image

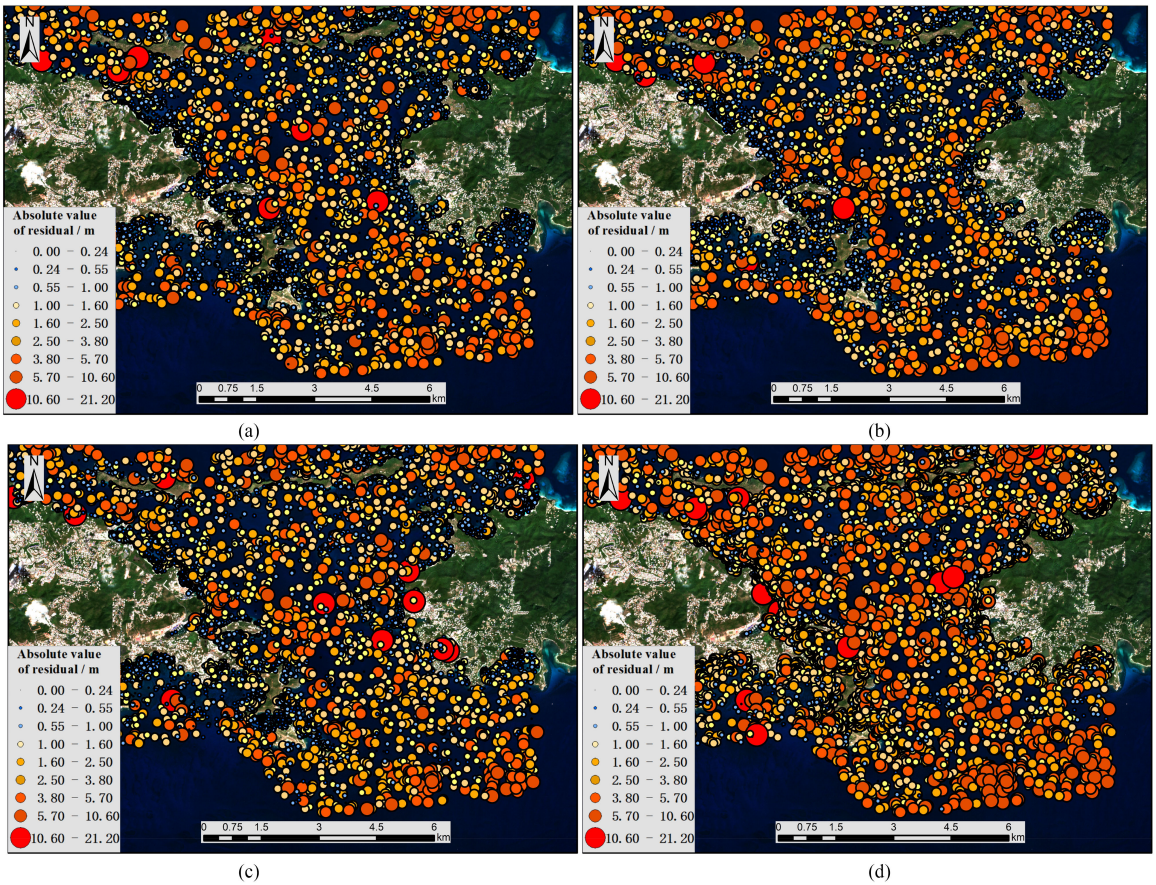


Fig. 9. Spatial distribution of APR in Saint Thomas Island. (a) APMLP model. (b) MLP model. (c) BP neural network model. (d) Log-ratio model.

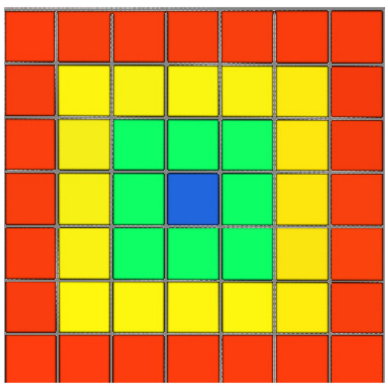


Fig. 10. Spatial position between the target pixel and adjacent pixels.

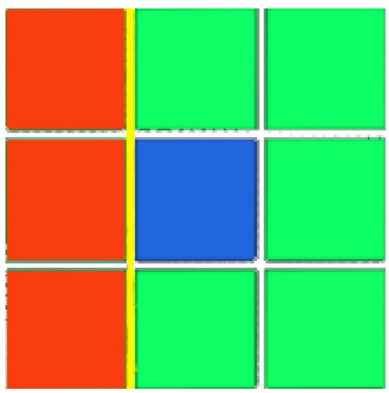


Fig. 11 Schematic diagram of target pixels at the edge in various states. Blue pixel is target pixel, red and green pixels are adjacent pixels, and the yellow line is the dividing line.

boundary, and the target pixel is located at the edge of the image, the red pixel does not actually exist. Therefore, we cannot obtain a SR centered on the blue pixel. In this article, these kinds of pixels are given up, water depth of the target pixels are not inverted.

2) Target Pixel at the Edge of Land and Water or Seafloor Substrate Types: As shown in Fig. 11, assuming the yellow line is the water-land boundary or the boundary of seafloor substrate types, the red pixels are land pixels or pixels from other seafloor substrate types, blue pixel is the target pixel. In this case, the target pixel has eight adjacent pixels, the water depth is inverted by these nine pixels in our this article. In fact, this method will cause errors, in future research, appropriate ways should be used to solve this problem.

VI. Conclusion

At present, most empirical models based on multispectral image usually obtain water depth by the relationship between the

SSR (a single pixel) and *in situ* water depth. The model will find a one-to-one relationship between the SSR and *in situ* depth. But, in fact, SSR is also related to seafloor substrate and IOP. Adjacent pixels may include the contribution of seafloor substrate and IOP to the SSR. In this article, an APMLP model including contribution of adjacent pixels is proposed. In this model, instead of one single pixel in the commonly used empirical models, a SR is defined. The water depth of the nine pixels in one SR is assumed to the same. Under this circumstance, when training the model, it can obtain a many-to-one relationship between SSRs and a water depth. This many-to-one relationship may also include contribution of seafloor substrate and IOP from the adjacent pixels.

Two datasets, Honolulu and Saint Thomas Island, were used to train and verify our APMLP model. The result was also compared with the BP, MLP, and LG model. In Honolulu, its correlation coefficient is 0.99573, which is the highest among the four models. There are some outliers in the other three models, but in APMLP, there are few, which proves that by adding adjacent pixels, the performance of the model can be improved significantly. In order to further compare the performance of the models at different depth ranges, we divided water depths into three intervals. In each interval, the APMLP model is always the one with the best performance. According to Fig. 8, the APR of APMLP near the coastline is significantly reduced, but in the area far from the coastline, the performance of APMLP has little different from that of other models. Similar performance was also achieved in Saint Thomas Island, the correlation coefficient of APMLP (0.98468) is still the highest among the four models. Compared with results of Honolulu, the overall accuracy of the four models has decreased significantly, but the accuracy of APMLP in each interval is still higher than that of the other three models. Compare Fig. 9(a)and (b), the light blue dots [see Fig. 9(b)] in shallow water areas mostly become blue or dark blue dots [see Fig. 9(a)], and the yellow dots disappeared significantly. This shows that introducing adjacent pixels into the model can effectively improve the inversion accuracy, especially in shallow water area. According to our experiment, the APMLP demonstrates its superiority.

The article confirms that considering adjacent pixels in an artificial neural network model can effectively improve the performance of water depth retrieval.

REFERENCES

[1] H. Xia *et al.*, "A bathymetry mapping approach combining log-ratio and semianalytical models using four-band multispectral imagery without ground data," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 4, pp. 2695–2709, Apr. 2020.

[2] A. Kanno and Y. Tanaka, "Modified lyzenga's method for estimating generalized coefficients of satellite-based predictor of shallow water depth," *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 4, pp. 715–719, Jul. 2012.

[3] I. Caballero, R. Stumpf, and A. Meredith, "Preliminary assessment of turbidity and chlorophyll impact on bathymetry derived from Sentinel-2A and Sentinel-3A satellites in South Florida," *Remote Sens.*, vol. 11, no. 6, p. 645, 2019.

[4] J. S. Bird and G. K. Mullins, "Analysis of swath bathymetry sonar accuracy," *IEEE J. Ocean. Eng.*, vol. 30, no. 2, pp. 372–390, Apr. 2005.

[5] A. Pacheco *et al.*, "Retrieval of nearshore bathymetry from landsat 8 images: A tool for coastal monitoring in shallow waters," *Remote Sens. Environ.*, vol. 159, pp. 102–116, 2015.

[6] D. Kohler, W. D. Philpot, and D. C.. Mobley, "Derivative based hyperspecral algorithm for bathymetric mapping," in *Proc. Ocean Opt. XIV*, 1998, pp. 10–13.

[7] I. N. Figueiredo, L. Pinto, and G. Gonçalves, "A modified lyzenga's model for multispectral bathymetry using tikhonov regularization," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 1, pp. 53–57, Jan. 2016.

[8] N. S. Hashim *et al.*, "Shallow-Water bathymetry estimation at pantai tok jembal, terengganu, Malaysia using landsat 8 (OLI)," in *Proc. IOP Conf. Ser., Earth Environ. Sci.*, 2021, Art. no. 012008.

[9] K. Zhang *et al.*, "Improving statistical uncertainty estimate of satellite-derived bathymetry by accounting for depth-dependent uncertainty," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–9, Apr. 2021, Art. no. 5401309.

[10] N. Xu, X. Ma, Y. Ma, P. Zhao, J. Yang, and X. H. Wang, "Deriving highly accurate shallow water bathymetry from sentinel-2 and ICESat-2 datasets by a multi-temporal stacking method," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 6677–6685, 2021.

[11] H. Li, P. Cheng, and X. Huang, "Hyperspectral bathymetry retrieval using a newly developed normalized algorithm in shallow water," *J. Indian Soc. Remote Sens.*, vol. 49, no. 10, pp. 2425–2436, 2021.

[12] E. Ariyasu, S. Kakuta, and T. Takeda, "Bathymetry mapping using hyperspectral data: A case study of, and N. J. YAMADA BAY," *ISPRS Int. Archives Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. XLI-B7, pp. 61–64, 2016.

[13] Z. Liu and Z. Yan, "Direct inversion of shallow-water bathymetry from EO-1 hyperspectral remote sensing data," *Chin. Opt. Lett.* vol. 9, no. 6, pp. 60102–60105, 2011.

[14] B. Chen *et al.*, "A dual band algorithm for shallow water depth retrieval from high spatial resolution imagery with no ground truth," *ISPRS J. Photogrammetry Remote Sens.*, vol. 151, pp. 1–13, May 2019.

[15] L. Zhang *et al.*, "Bathymetry using hyperspectral imagery based on Semi-analytical model," in *Hydrographic Surveying and Charting*, vol. 31, no. 4, pp. 17–21, 2011.

[16] H. Xia *et al.*, "A bathymetry mapping approach combining log-ratio and semianalytical models using four-band multispectral imagery without ground data," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 4, pp. 2695–2709, Apr. 2020.

[17] Z. Lee *et al.*, "Hyperspectral remote sensing for shallow waters. I. A semianalytical model," *Appl. Opt.*, vol. 37, no. 27, pp. 6329–6338, 1998.

[18] W. Shen *et al.*, "The GF-2 capability analysis in shallow water remote sensing bathymetry," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2019, pp. 8253–8256.

[19] C. Antoine and J. Hench, "Extracting shallow bathymetry from very high resolution satellite spectral bands and a machine learning algorithm," *Int. Council Exploration Sea*, no. 24, Jan. 2015.

[20] W. Shen *et al.*, "Comparative study of different machine learning models for remote sensing bathymetry inversion," in *Proc. Int. Conf. Intell. Fuzzy Syst.*, 2020, pp. 1140–1148.

[21] R. P. Stumpf, K. Holderied, and M. Sinclair, "Determination of water depth with high-resolution satellite imagery over variable bottom types," *Limnol. Oceanogr.*, vol. 48, no. 1, pp. 547–556, 2003.

[22] Z. Deng, M. Ji, and Z. Zhang, "Mapping bathymetry from multi-source remote sensing images: A case study in the Beilun Estuary, Guangxi, China. The international archives of the photogr ammetry," *Remote Sens. Spatial Inf. Sci.*, vol. 37, pp. 1321–1326, 2008.

[23] J. M. Kerr and S. Purkis, "An algorithm for optically-deriving water depth from multispectral imagery in coral reef landscapes in the absence of ground-truth data," *Remote Sens. Environ.*, vol. 210, pp. 307–324, 2018.

[24] J. X. Leon and T. J. Cohen, "An improved bathymetric model for the modern and palaeo Lake Eyre," *Geomorphology*, vol. 173/174, pp. 69–79, 2012.

[25] V. Mateo-Pérez, M. Corral-Bobadilla, F. Ortega-Fernández, and E. P. Vergara-González, "Port bathymetry mapping using support vector machine technique sentinel-2 Satellite Imagery," *Remote Sens.*, vol. 12, 2020, Art. no. 2069.

[26] X. Monteys, P. Harris, S. Caloca, and C. Cahalane, "Spatial prediction of coastal bathymetry based on multispectral satellite imagery and multibeam data," *Remote Sens.*, vol. 7, pp. 13782–13806, 2015.

[27] Z. Wang *et al.*, "Comparative study on water depth remote sensing in yellow river estuary based on BP ANN method and bottom Albedo-independent bathymetry algorithm," *J. Central China Normal Univ. (Natural Sci.)*, vol. 50, pp. 112–119, 2016.

[28] Y. Wang, X. Zhou, C. Li, Y. Chen, and L. Yang, "Bathymetry model based on spectral and spatial multifeatures of remote sensing image," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 1, pp. 37–41, Jan. 2020.

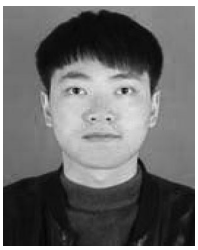
[29] S. Maritorena, D. A. Siegel, and A. R. Peterson. "Optimization of a semianalytical ocean color model for global-scale applications," *Appl. Opt.*, vol. 41, no. 15, pp. 2705–2714, 2002.
[30] W. D. Philpot, "Bathymetric mapping with passive multispectral imagery," *Appl. Opt.*, vol. 28, no. 8, pp. 1569–1578, 1989.
[31] R. Lippmann, "An introduction to computing with neural nets," *IEEE ASSP Mag.*, vol. 4, no. 2, pp. 4–22, Apr. 1987.
[32] J. M. Zurada, *Introduction to Artificial Neuralsystems*. Eagan, MN, USA: West Pub., 1992.
[33] K. Shibata and Y. Ikeda, "Effect of number of hidden neurons on learning in large-scale layered neural networks," in *Proc. Int. Joint Conf.*, 2009, pp. 5008–5013.
[34] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *Proc. 3rd Int. Conf. Learn. Representations*, San Diego, 2015, *arXiv:1412.6980v9*.

Jinshan Zhu received the Ph.D. degree in cartography and geographical information system from the Ocean University of China, Qingdao, China, in 2008.

He is currently with the College of Geomatics, Shandong University of Science and Technology, Qingdao. His main research interests include satellite-derived bathymetry data for shallow marine water from optical sensors.

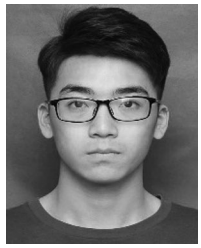
Jian Qin received the B.S. degree from Jiangsu Ocean University, Lianyungang, China, in 2020. He is currently working toward the M.S. degree with the College of Geodesy and Geomatics, Shandong University of Science and Technology, Qingdao, China, both in surveying and mapping engineering.

His research interests include ocean remote sensing.



Fei Yin received the B.S. degree from Jiangsu Ocean University, Lianyungang, China, in 2020. He is currently working toward the M.S. degree from the College of Geodesy and Geomatics, Shandong University of Science and Technology, Qingdao, China, both in surveying and mapping engineering.

His research interests include on ocean remote sensing.



Zhaoyu Ren received the B.S. degree in surveying and mapping engineering from the Shandong University of Science and Technology, Qingdao, China, in 2019. He is currently working toward the M.S. degree in surveying and mapping engineering with the College of Geodesy and Geomatics, Shandong University of Science and Technology, Qingdao, China.

His research interests include the application of atmospheric correction for ocean data.

Jiawei Qi received the B.S. degree in hydrographic surveying and charting from the Shandong University of Science and Technology, Qingdao, China, in 2019. He is currently working toward the M.S. degree in hydrographic surveying and charting with the College of Geodesy and Geomatics, Shandong University of Science and Technology, Qingdao, China.

His research interests include remote sensing bathymetry retrieval.

Jingyu Zhang received the B.S. degree in ocean technology from Shanghai Ocean University, Shanghai, China, in 2011, and the M.S. degree in physical oceanography from the First Institute of Oceanography, MNR, Qingdao, China, in 2015.

She is currently an Engineer with the First Institute of Oceanography. Her research interests include remote sensing image processing techniques (classification and feature extraction) and coastal monitoring by high resolution remote sensing images (bathymetry and benthic classification).

Ruifu Wang received the Ph.D. degree in GIS and remote sensing from the Ocean University of China, Qingdao, China, in 2006.

He is currently an Associate Professor with the College of Geomatics, SDUST, Qingdao. His research interests include on marine GIS and remote sensing of sea ice.