Acta Oceanol. Sin., 2024, Vol. 43, No. 1, P. 112–122
https://doi.org/10.1007/s13131-023-2203-9
http://www.aosocean.com
E-mail: ocean2@hyxb.org.cn

A VGGNet-based correction for satellite altimetry-derived gravity anomalies to improve the accuracy of bathymetry to depths of 6 500 m

Xiaolun Chen[1], Xiaowen Luo[1, 2, 3*], Ziyin Wu[1, 2, 4*], Xiaoming Qin[1, 2], Jihong Shang[1], Huajun Xu[5], Bin Li[6], Mingwei Wang[1], Hongyang Wan[1]

[1] Key Laboratory of Submarine Geosciences, Second Institute of Oceanography, Ministry of Natural Resources, Hangzhou 310012, China

[2] Ocean College, Zhejiang University, Zhoushan 316021, China

[3] Key Laboratory of Ocean Space Resources Management Technology, Marine Academy of Zhejiang, Hangzhou 310012, China

[4] School of Oceanography, Shanghai Jiao Tong University, Shanghai 200240, China

[5] School of Civil Engineering and Architectures, Zhejiang University of Science and Technology, Hangzhou 310023, China

[6] National Center for Archaeology, Beijing 100013, China

Received 28 December 2022; accepted 10 April 2023

Abstract

Understanding the topographic patterns of the seafloor is a very important part of understanding our planet. Although the science involved in bathymetric surveying has advanced much over the decades, less than 20% of the seafloor has been precisely modeled to date, and there is an urgent need to improve the accuracy and reduce the uncertainty of underwater survey data. In this study, we introduce a pretrained visual geometry group network (VGGNet) method based on deep learning. To apply this method, we input gravity anomaly data derived from ship measurements and satellite altimetry into the model and correct the latter, which has a larger spatial coverage, based on the former, which is considered the true value and is more accurate. After obtaining the corrected high-precision gravity model, it is inverted to the corresponding bathymetric model by applying the gravity-depth correlation. We choose four data pairs collected from different environments, i.e., the Southern Ocean, Pacific Ocean, Atlantic Ocean and Caribbean Sea, to evaluate the topographic correction results of the model. The experiments show that the coefficient of determination (R^2) reaches 0.834 among the results of the four experimental groups, signifying a high correlation. The standard deviation and normalized root mean square error are also evaluated, and the accuracy of their performance improved by up to 24.2% compared with similar research done in recent years. The evaluation of the R^2 values at different water depths shows that our model can achieve performance results above 0.90 at certain water depths and can also significantly improve results from mid-water depths when compared to previous research. Finally, the bathymetry corrected by our model is able to show an accuracy improvement level of more than 21% within 1% of the total water depths, which is sufficient to prove that the VGGNet-based method has the ability to perform a gravity-bathymetry correction and achieve outstanding results.

Key words: gravity anomaly, bathymetry inversion, VGGNet, multibeam sonar, satellite altimetry

Citation: Chen Xiaolun, Luo Xiaowen, Wu Ziyin, Qin Xiaoming, Shang Jihong, Xu Huajun, Li Bin, Wang Mingwei, Wan Hongyang. 2024. A VGGNet-based correction for satellite altimetry-derived gravity anomalies to improve the accuracy of bathymetry to depths of 6 500 m. Acta Oceanologica Sinica, 43(1): 112–122, doi: 10.1007/s13131-023-2203-9

1 Introduction

Submarine topographic surveys are a fundamental aspect of marine mapping, the purpose of which is to obtain the three-dimensional coordinates of submarine topographic features, including information such as position, water depth, water level, sound velocity and attitude bearing; the central component of these is bathymetry (Zhao et al., 2017). Shipborne sounding surveys are the most direct and primitive method used to determine the topography of the seafloor and the methods have undergone an iterative evolution from single-beam to multibeam echosounders. At present, multibeam echosounder sonar systems are still one of the main methods for mapping the topography of the

Foundation item: The National Key R&D Program of China under contract Nos 2022YFC3003800, 2020YFC1521700 and 2020YFC1521705; the National Natural Science Foundation of China under contract No. 41830540; the Open Fund of the East China Coastal Field Scientific Observation and Research Station of the Ministry of Natural Resources under contract No. OR-SECCZ2022104; the Deep Blue Project of Shanghai Jiao Tong University under contract No. SL2020ZD204; the Special Funding Project for the Basic Scientific Research Operation Expenses of the Central Government-Level Research Institutes of Public Interest of China under contract No. SZ2102; the Zhejiang Provincial Project under contract No. 330000210130313013006.
*Corresponding author, E-mail: cdslxw@163.com; zywu@vip.163.com

seafloor (Wu et al., 2020). Multibeam sonar has the advantage of high spatial accuracy, which has enabled the collection of underwater sounding data to make a large leap in quality from the acquisition of point measurements to continuous lines and from lines to surfaces (Li, 1999). However, the disadvantages of low efficiency, high cost and the long measurement time required make it difficult to conduct frequent submarine surveys over a wide range of marine regions (Colbo et al., 2014; Hughes Clarke, 2018). As a result, the area covered by shipborne surveys remains very sparse at present. It is estimated that less than 20% of the global seafloor has been surveyed by shipborne systems, and a significant proportion of this area, especially in deep-sea areas, has relatively low accuracy (Coley, 2022).

Satellite altimetry is a spatial measurement technique that uses artificial satellites as carriers to measure the height of satellites relative to the Earth's surface using radar, lasers and other ranging techniques to obtain the topography of the Earth's surface (Zwally et al., 2002). Satellites carrying radar altimeters measure the shape of the global sea surface along their orbits, from which maps of the ocean gravity field can be produced. The satellite gravity field and existing depth measurements are used to determine the relationship between the gravity field and seafloor topography, and by applying this relationship to the gravity field, the seafloor topography can be predicted with certain radar bands (Smith and Sandwell, 1994, 1997). In most cases, the seafloor topographic model obtained via gravity inversion is accurate to within ±100 m. Although this model is a predicted dataset with errors, it is still the best estimate currently available for the large-scale and even global submarine topography.

Parker (1973) derived a detailed expression for gravity in the frequency domain and proposed a material interface to describe variations in the gravity anomaly model caused by fluctuations that could be applied to the frequency domain model, which laid the foundation for the development of submarine inversion topographic. Ibrahim and Hinze (1972) proposed the gravity geological method (GGM) to measure rock heights under submarine terrestrial sediments, which has since been widely used (and continuously improved) for the inversion of seafloor topography in various seas (Oldenburg, 1974; Braitenberg et al., 2006; Ouyang et al., 2014; Wei et al., 2021; An et al., 2022). The admittance theory of gravity anomalies is based on the spectral relationship between the oceanic gravity field and seafloor topography and was proposed in conjunction with the equilibrium assumption of Airy (1855) and Parker (1973) to estimate seafloor topography. This was applied to the study of modeling submarine mountains (Watts et al., 2006; Gong et al., 2021) and to the inversion of seafloor topography (Ouyang et al., 2015; Fan et al., 2021). Over the years, researchers have used these methods to produce a variety of global seafloor topographic models, such as the Earth Topography (ETOPO), General Bathymetric Chart of the Oceans (GEBCO), Danish Technical College (DTU), Scripps Institution of Oceanography (SIO) series.

Despite the increasing need for accuracy in marine scientific research in recent years, there is limited room for improving the accuracy of submarine topographic models obtained by the traditional inversion of gravity anomalies derived from satellite altimetry; and shipborne multibeam sonar and gravimetry, which have high accuracy characteristics, are often confined to areas of relatively limited coverage (Yale et al., 1998; Shi et al., 2017). Therefore, there is an urgent need to combine the advantages of their respective spatial coverages and spatial resolutions to develop a submarine topographic inversion method that covers a large range, is highly accurate and has high flexibility.

In recent years, deep learning has gradually become an important scientific computing tool in digital image processing. Deep learning uses a layer-by-layer abstraction approach to efficiently and accurately achieve the extraction of bottom- to top-level attributes and has made great contributions in the fields of natural language processing (Otter et al., 2021), image and speech recognition (Wu and Chen, 2015), and remote sensing data fusion (Benedetti et al., 2018; Yuan et al., 2020). Jena et al. (2012) developed an artificial neural network (ANN) model based on the radial basis function (RBF) to predict undetected seafloor topographic features using satellite altimetry-derived gravity data, with the results showing better accuracy than global bathymetry models. Moran (2020) provided an experiment-based evaluation of a variety of existing machine learning methods applied to global bathymetric inversions and gave recommendations for model selection based on different regional conditions. Fan et al. (2021) introduced a method for predicting seafloor topography that involved vertical gravity gradients and used a nonlinear least squares algorithm, with the accuracy of its results demonstrating the feasibility of this algorithm compared with global seafloor topographic models. Zhu et al. (2021) used a multilayer perception (MLP)-based neural network to correct satellite altimeter gravity by shipborne data. A comparative analysis was found to achieve an effective and large-scale refinement of the accuracy of the satellite-derived gravity anomaly model. Seoane et al. (2022) used an extended Kalman filter (EKF) method that iteratively integrated continuous input gravity anomaly and geoid height data to simulate regional topographic models of the seafloor in areas lacking ship-measured bathymetry. Annan and Wan (2022) introduced the vertical deflections in bathymetric inversion, which was combined with gravity anomalies and vertical gravity gradients by an outperforming convolutional neural network (CNN), demonstrating the usability of vertical deflections in bathymetric inversion. Chen et al. (2022) applied a pretrained VGGNet model to apply corrections to satellite-derived bathymetry based on multibeam sonar data, and the results showed that this method could significantly enhance the accuracy of the satellite data.

Previous studies have often noted that the main problem associated with algorithm development or model formulation for the estimation of seafloor topographic parameters is the complexity of the physical processes involved and the uncertainties associated with them. Therefore, the selection of a suitable model is a critical step before starting the inversion. Gatys et al. (2016, 2017) proposed a method based on visual geometry group network (VGGNet), a CNN commonly used for high-precision visual recognition, to computationally extract the style and content information of a set of input images and output a new fused image with both style and content information from the input ones. Inspired by this, we can treat the shipborne gravity data and satellite altimeter-derived gravity data as style and content information, respectively, to synthesize new image data through the VGGNet.

In this study, a CNN-based optimization algorithm is proposed to correct satellite altimetry-derived gravity data that contains errors through the use of a pretrained VGGNet network with shipborne marine gravity data applied as the correct value. Structurally, the model input is the data to be corrected, and the output is the data with the true values; and the aim is to identify the high-order image features in the intermediate layer. The role of the intermediate layer is to establish the complex relationship between the input and output values, i.e., to derive a complex understanding of the image features from the original image pixels and to construct a representation of the internal mapping. After

calibrating the satellite-derived gravity and inverting the seafloor topographic model of the corresponding area from it, a more accurate inversion of the bathymetry can be obtained.

The main contributions of this study are as follows.

(1) A method is proposed that uses deep learning to correct the gravity anomaly data derived from satellite altimetry with shipborne data used as the true value and to derive a high-precision seafloor topographic model accordingly.

(2) The CNN-based VGGNet model is first introduced to generate satellite-derived data with improved accuracy by calculating and minimizing the distance (loss) between the shipborne true data and the satellite data (which contains errors). The new fused data will have the advantages of both the high accuracy of the former and the high coverage of the latter.

(3) The performance of the model is validated by experiments on datasets from the Southern Ocean, Pacific Ocean and Atlantic Ocean, with the results showing that this method is effective for improving the computational accuracy of seafloor topographic inversion.

2 Methodology and data

The research structure of this paper is shown in Fig. 1. Overall, relying on gravity data pairs derived from ship measurements and satellite altimetry, the pretrained VGGNet model is able to correct the satellite data with ship measurements (the true value in the framework of a neural network). After obtaining the corrected satellite altimetry-derived gravity data, the corresponding bathymetric maps can be obtained by inversion, and the accuracy can be verified with the help of shipborne multibeam sonar bathymetry data, which can be considered measured values.

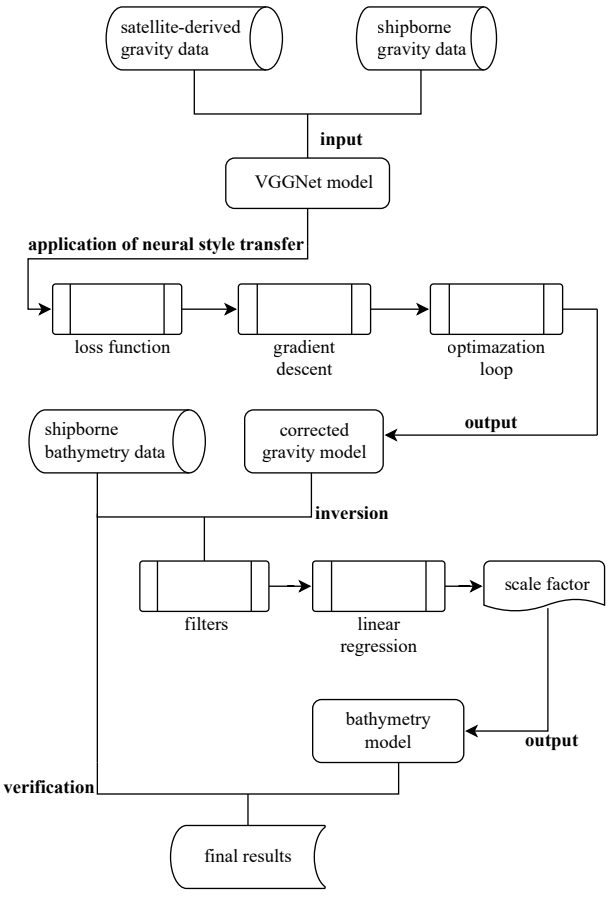

Fig. 1. The research structure.

2.1 *Framework of VGGNet*

In recent years, convolutional neural networks (CNNs) have made great progress in the field of pattern recognition in digital images, and some excellent algorithmic models, such as AlexNet (Krizhevsky et al., 2017), CaffeNet (Jia et al., 2014) and VGGNet (Simonyan and Zisserman, 2015), have emerged. VGGNet applies a very small field of view rather than a large field and considers 3 × 3 pixels with a step size of 1. The decision function is more discriminative because there are 3 ReLU units. There are also fewer parameters, with 27 times the number of channels. Without modifying the perceptual field, VGGNet uses a 1 × 1 convolutional layer to make the decision function more nonlinear. Due to the tiny size of the convolutional filter, the VGGNet model can have a considerable number of weight layers, which means better performance potential. There are two general forms of VGGNet, VGG-16, and VGG-19, which are not fundamentally different but only differ in the depth of the network, with VGG-19 having three more convolution layers. Considering the experimental requirements and data parameters, VGG-19, which can offer deeper information, is chosen in the experiment. Compared with most previous CNN-originated models that have 4–7 layers, VGG-19 is constructed with 19 layers, including 16 convolutional layers and 3 fully connected layers, enabling it to extract the more abstract and more profound image features and reduce the number of parameters while being able to retain the same receptive field. Thus, it has improved the efficiency and accuracy of image computing (Huo et al., 2020; Islam et al., 2020; Schulz et al., 2020).

The structure of VGG-19 is displayed in Fig. 2. The entire network uses the exact size of the convolution kernels (3 × 3) and the maximum pooling kernels (2 × 2). The combination of several small filter (3 × 3) convolutional layers is better than a large filter (5 × 5 or 7 × 7), as in the previous models. Since the convolution kernel focuses on expanding the number of channels and the pooling kernel focuses on reducing the width and height, the architecture is more profound and broader. Additionally, the increase in the number of calculations slows the process down, showing the network a larger receptive field. Furthermore, the network parameters are reduced, and the rectified linear unit (ReLU) activation function is used in convolutional layers and fully connected layers to create more linear transformations to enhance the learning ability.

2.2 *Model training steps*

We correct ocean gravity data using the VGG-19 model. The structural framework is to take the satellite altimetric gravity data to be corrected as the input values and the shipborne gravity data to be treated as true values as output. These procedures allow the machine to train a rationalization model in the intermediate layer that minimizes the loss (distance) of this pair of datasets. The principle of the correction model is to define a distance function that describes how different the two input images are. The gridded shipborne data image and the satellite altimetry-derived data image that covers the same area are passed to the model, which is supposed to return the intermediate layer outputs from the model. The distance function L that we use is shown as follows:

$$L^l(x,p) = \sum_{i,j}\left[F_{ij}^l(x) - P_{ij}^l(p)\right]^2, \tag{1}$$

where x stands for the shipborne gravity image, p stands for the satellite altimetry-derived gravity image, and i and j stands for the serial number of pixel points of the input images, and n is the

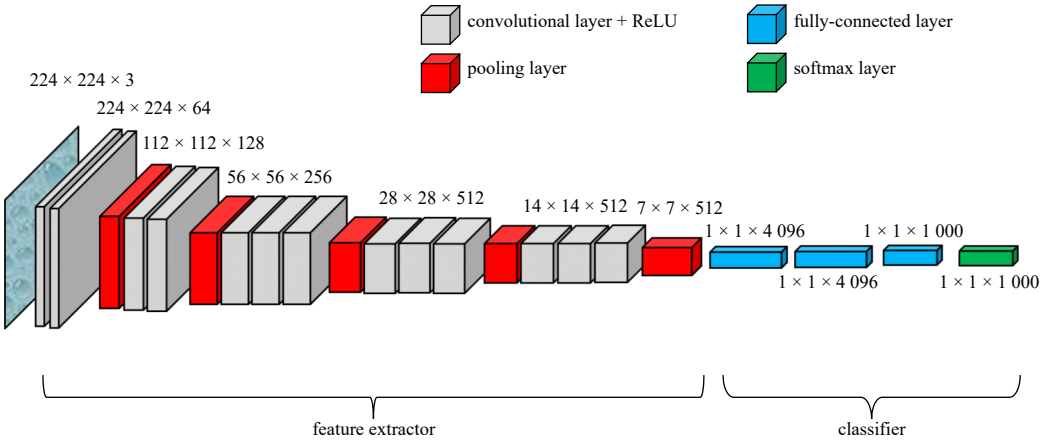

Fig. 2. Architecture of the VGG-19 model. The boxes represent the size of each layer.

exponential power, taken as a positive integer (Parker, 1973). Let V_{nn} be a pretrained VGG-19 network and X be any image; then, $V_{nn}(X)$ is the network fed by X, where X denotes the set of points on each grid point $(x_1, x_2, \cdots, x_n)$ of the set. Let $F_{ij}^{l}(x) \in V_{nn}(x)$ and $P_{ij}^{l}(p) \in V_{nn}(p)$ describe the respective intermediate feature representation of the network with the inputs x and p at layer l. Finally, optimizers' rules are applied to iteratively update the images, which minimizes a given loss to the inputs.

The evaluation of the corrected precision is based on comparisons with a previous study. To quantify the differences and connections between the predicted value and true value, we choose four evaluation measurements, root-mean-square error (RMSE), normalized RMSE (NRMSE), coefficient of determination (R^2), and standard deviation (SD), as follows:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}\left(\hat{f}_i - y_i\right)^2}{n}}, \tag{2}$$

$$\mathrm{NRMSE} = \frac{\mathrm{RMSE}}{y_{\max} - y_{\min}}, \tag{3}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - f_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}, \tag{4}$$

$$\mathrm{SD} = \sqrt{\frac{\sum_{i=1}^{n}(f_i - \bar{f})^2}{n}}, \tag{5}$$

where n represents the number of values from the dataset, i represents the serial number of the value from the dataset, f represents the predicted values, and y represents the true values. As a standardization process, NRMSE is able to limit the value of RMSE to a certain range, eliminate the influence of the scale between indicators, and solve the comparability among data indicators. NRMSE and R^2 usually range from 0 to 1. A smaller RMSE, NRMSE, SD and larger R^2 mean that there is a higher correlation and more stable distribution among the datasets.

The default hyperparameters used in the VGG-19 training model are shown in Table 1. In the experiments, the weights between the content layer, style layer, content, style and total variance loss are set to maintain approximately the same order of magnitude. The learning rate is set so that the iterations occur in a convergent process and that the loss curves decrease smoothly.

Table 1. The default hyperparameters of the training model

Hyperparameter	Setting
Content layer	"conv4_2"
Style layer	"conv1_1", "conv2_1", "conv3_1", "conv4_1", "conv5_1"
Weight of loss at content layer	1
Weight of loss at style layer	1, 1, 1, 1
Weights among content, style, and total variation loss	1×10^{-4}, 1, 1×10^{-5}
Learning rate	starts at 10 and linear decays over 100 iterations to 1

2.3 *Bathymetric inversion*

According to the flexural isostasy theory proposed and summarized by Watts (1978, 2001), a certain degree of flexural isostasy (also known as equilibrium compensation) $r(x)$ occurs in the Moho below the corresponding oceanic crust when topographic undulations $h(x)$ are applied to the seafloor. Thus, the main anomalous field sources that generate the sea surface gravity anomaly are the seafloor topography and its equilibrium compensation material. The gravity anomaly produced by a certain seafloor topography can be expressed as

$$\Delta G(k) = 2\pi G(\rho_c - \rho_w)\exp(-kd)\sum_{n=1}^{\infty} H^n(k) + 2\pi G(\rho_m - \rho_c)\exp[-(k+d)]\sum_{n=1}^{\infty} R^n(k), \tag{6}$$

where $\Delta G(k)$ is the Fourier transform of the gravity anomaly, G stands for the gravity constant, and ρ_c, ρ_w, and ρ_m represent the crust, seawater, and mantle densities, respectively. $k = \frac{2\pi}{\lambda}$, where λ represents the topographic wavelength. d represents the average water depth. $R(k)$ and $H(k)$ represent the Fourier transforms of $r(x)$ and $h(x)$, respectively. In the frequency domain,

$$R(k) = -H(k)\frac{\rho_c - \rho_w}{\rho_m - \rho_c}\Phi_e'(k), \tag{7}$$

where $\Phi_e'(k)$ is called the flexural response equation and ex-

pressed as

$$\Phi'_{\mathrm{e}}(k)=\left[\frac{Dk^{4}}{(\rho_{\mathrm{m}}-\rho_{\mathrm{c}})g}+1\right]^{-1}, \tag{8}$$

where D represents the lithosphere flexural stiffness and $D=\dfrac{ET_{\mathrm{e}}^{3}}{12(1-\nu)^{2}}$. While E represents Young's modulus, T_{e} represents the effective elastic thickness, ν represents Poisson's ratio, and g stands for gravitational acceleration (Walcott, 1970). Empirically, Eq. (6) converges quickly, and when only the linear term of this equation ($n=1$) is taken into account, Eqs (7) and (8), the gravity anomaly generated by the topography of the seafloor can be expressed as

$$\Delta G(k)=2\pi G(\rho_{\mathrm{c}}-\rho_{\mathrm{w}})\exp(-kd)H(k)\left[1-\Phi'_{\mathrm{e}}(k)\exp(-kt)\right], \tag{9}$$

where t represents the average oceanic crust thickness. That is, the admittance function $Z(k)$ from the seafloor topography to the sea surface gravity anomaly can be expressed as

$$Z(k)=2\pi G(\rho_{\mathrm{c}}-\rho_{\mathrm{w}})\exp(-kd)\left[1-\Phi'_{\mathrm{e}}(k)\exp(-kt)\right]. \tag{10}$$

The theoretical admittance function indicates that the sea surface gravity anomaly is strongly correlated with the seafloor topography only in the short and medium wavelengths due to the effects of equilibrium compensation and water depth extension (Hu et al., 2020). Therefore, based on the flexural isostasy theory, gravity anomalies are usually used only for the inversion of seafloor topography in the short and medium wavelengths (10–200 km), while shipborne bathymetry is used to construct long wavelength (>200 km) seafloor topography.

From the above, the flexural response equation part in Eq. (9) can be neglected in the short wavelength band (<200 km). The inversion of the seafloor topography from the sea surface gravity anomaly is given by

$$H(k)=\frac{\exp(kd)}{2\pi G(\rho_{\mathrm{c}}-\rho_{\mathrm{w}})}\Delta G(k). \tag{11}$$

For the corrected satellite-derived gravity data, we inverted them using the filter method (Smith and Sandwell, 1994) to obtain the corresponding bathymetric data. The filter method combines flexural isostasy theory, admittance theory and the gravity noise signal ratio to propose a low-pass and bandpass filter-based method to predict bathymetry using gravity data.

The short- and mid-wavelength bathymetry is calculated using a scaling factor $S(x)$, which is determined empirically by regression and reflects the correlation between the gravity data and the bathymetry data in the corresponding band. Eventually, the seafloor topographic model will consist of two parts: (1) a long-wave (>200 km) model $h_{\mathrm{long}}(x)$ constructed by shipborne bathymetry and (2) a shortwave and medium wave (10–200 km) model inversed by band-limited gravity anomaly $\delta g(x)$, as shown in Eq. (12), with the processing flow shown in Fig. 3.

$$h(x)=h_{\mathrm{long}}(x)+S(x)\cdot\delta g(x). \tag{12}$$

2.4 *Experimental data*

In the experiment, the satellite altimeter-derived gravity anomaly data used are acquired from the Bureau Gravimetrique International (BGI) (Bonvalot et al., 2012). The shipborne gravity anomaly data and multibeam sonar bathymetry data used are acquired from the National Geophysical Data Center (NGDC) of the National Oceanic and Atmospheric Administration (NOAA) (NOAA National Centers for Environmental Information, 2004, 2015). For data from different sources, rasters were resampled to the same resolution for subsequent analysis.

A total of four pairs of shipborne gravity-satellite gravity-multibeam bathymetry datasets from the Southern Ocean, Pacific Ocean, Atlantic Ocean, and Caribbean Sea were selected, with detailed information listed in Table 2. We have not included the temporal resolution of these data in the table, taking into account the temporal scale of ocean variability.

In the correction process, we randomly split the gravity datasets from each region in two, with 50% used as the training set to debug the parameters in the network and the remaining 50% used as the validation set to tune the hyperparameters of the model and to conduct a preliminary evaluation of the performance of the model.

3 Experimental results

3.1 *Model performance evaluation*

A loss function is applied to estimate the degree of inconsistency between the predicted (corrected) and true values of the model, as shown in Fig. 4. It is a nonnegative real-valued function, with the smaller the loss function, the better the robustness of the model, and vice versa. As seen from the results, in the experiments conducted in the four regions, a sudden drop in loss occurred in the first 20 epochs. After dropping to between 0.1 and 0.2, the loss starts to decline gently, with this phenomenon being

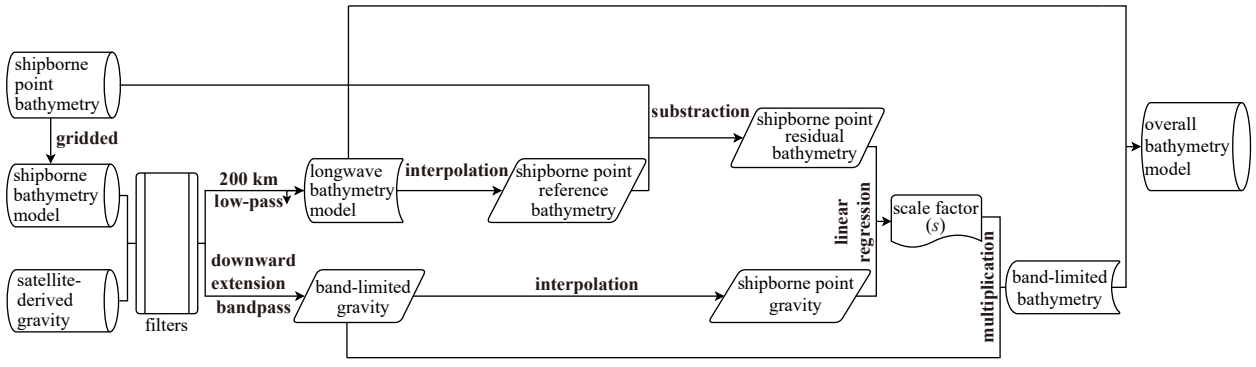

Fig. 3. The calculation and processing flow of the filter method.

Table 2. The parameters of the datasets

Location	Center point coordinate	Spatial resolution/m	Data size	Area/km^2	Bathymetry/m
Southern Ocean	71°S, 173°E	93	5 097 104	43 700	211–4 077
Pacific Ocean	9°S, 140°W	93	33 048 000	283 337	113–4 992
Atlantic Ocean	32°N, 65°W	93	3 240 000	27 778	58–4 920
Caribbean Sea	18°N, 82°W	123	10 614 363	150 310	1–6 580

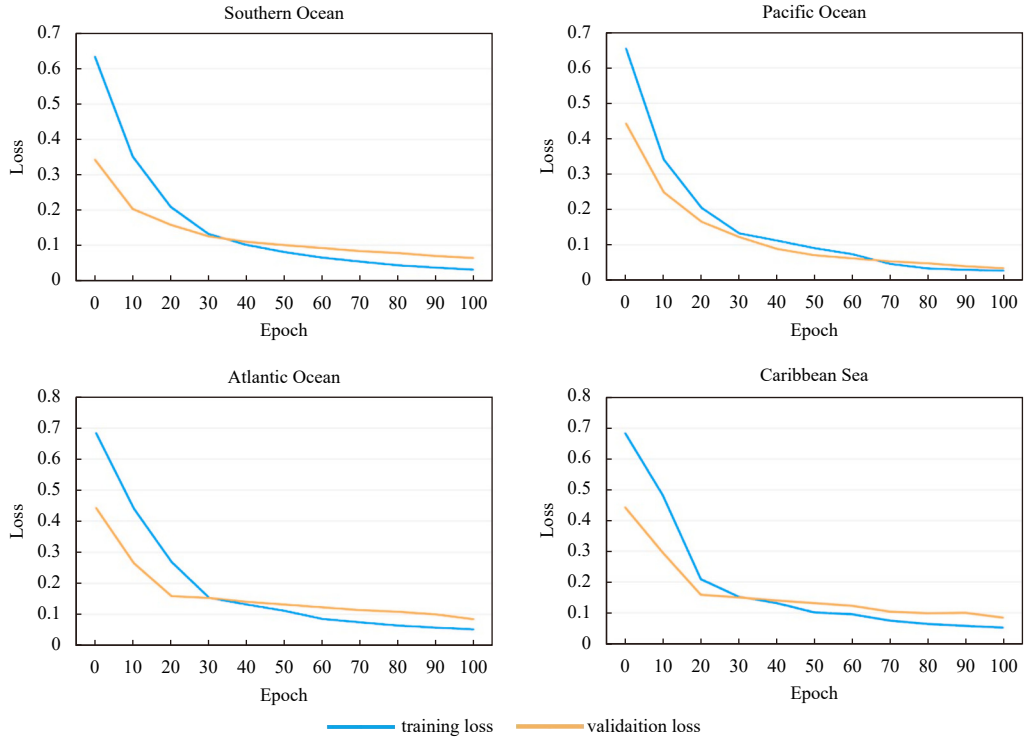

Fig. 4. The training and validation loss from the experiments in the Southern Ocean, Pacific Ocean, Atlantic Ocean, and Caribbean Sea.

more pronounced after 70 epochs. No significant overfitting is found from the curves. This performance of the loss function demonstrates the usefulness of the VGG-19 model for gravity correction.

3.2 *Overall accuracy analysis*

The output of the model results for a corrected gravity anomaly, with its verification accuracy listed in Table 3. As seen from the table, the R^2 of the network-corrected gravity anomaly model can generally reach a high level of 0.90 or more, indicating that it matches the true values from the shipborne measurements quite well. The SD is similar to the RMSE in terms of performance. The results of the normalized RMSE can provide the opportunity to conduct a comparison with previous research, and it is found that the accuracy of the output from our method can reach a level of accuracy that is indistinguishable from results obtained when using the multilayer perceptron (MLP) method, with respective performances (Zhu et al., 2021). These results provide a solid basis for the next bathymetry inversion step.

To visualize the correction results, we arbitrarily select three sites with visual differences to represent each of the four experimental areas for comparative display, as shown in Fig. 5. As seen from the figure, generally speaking, the texture of the corrected data is much finer and more consistent with the actual gravity field distribution on the seafloor. This provides a foundation for the following work on bathymetry inversion.

The overall accuracy of the bathymetry correction process has been verified with the multibeam sonar bathymetry used as the true value, and the performance of the parameters is presented in Table 4. The four datasets all have similar R^2 values of approximately 0.80, indicating a high correlation between the corrected bathymetry and the true values. There is some degree of positive correlation between the differences in the performance of the SD and RMSE for the four datasets, with values relating to the range of water depths in the experimental area. A comparison of NRMSE using the same criteria shows that the Atlantic data have the best correction accuracy, followed by the Southern Ocean and Pacific Ocean data, while the Caribbean Sea data rank last. In comparison with similar studies in recent years, we have found that this method is effective for improving accuracy by up to 15.4% (gravity-bathymetry correction) or 24.2% (bathymetry-only correction), demonstrating that this VGGNet method enables a new level of accuracy in seafloor topography correction (Annan and Wan, 2022; Chen et al., 2022).

Table 3. Overall accuracy of the gravity correction

Location	R^2	SD/mGal	RMSE/mGal	NRMSE
Southern Ocean	0.902	18.333	13.630	0.113
Pacific Ocean	0.955	17.892	10.050	0.118
Atlantic Ocean	0.930	19.567	21.549	0.114
Caribbean Sea	0.919	21.051	16.485	0.113

By comparing the variation in R^2 values at different water depths, we plot the relationship between the two parameters, as shown in Fig. 6. For the results of all four experimental areas, the

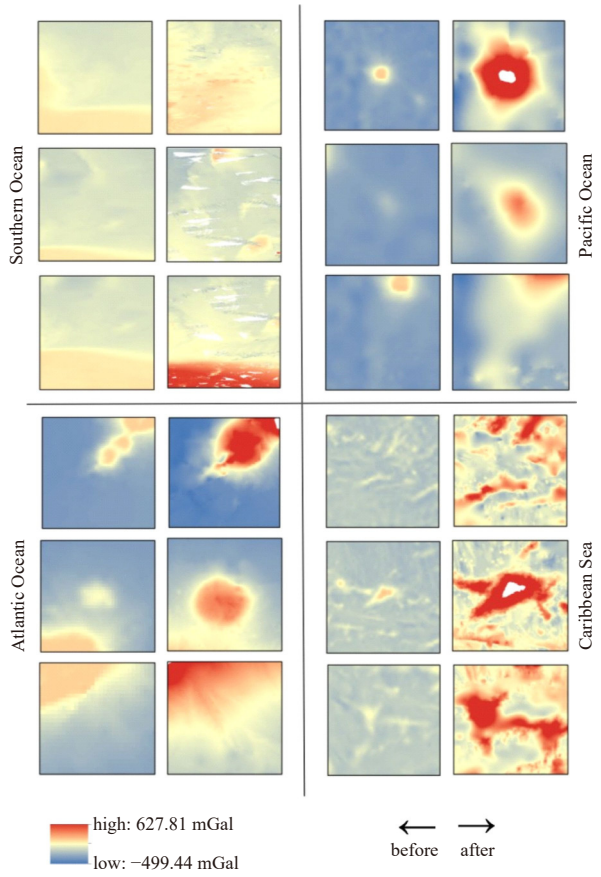

Fig. 5. Comparison of local details of gravity anomalies before and after correction from the Southern Ocean, Pacific Ocean, Atlantic Ocean, and Caribbean Sea. Each subfigure shows three visualizations of randomly selected locations.

Table 4. The overall accuracy of the bathymetry correction

Location	R^2	SD/m	RMSE/m	NRMSE
Southern Ocean	0.822	104.790	107.024	0.027
Pacific Ocean	0.834	117.630	126.366	0.026
Atlantic Ocean	0.833	124.847	136.622	0.028
Caribbean Sea	0.783	139.583	164.475	0.025

minimum values of R^2 are above 0.2 and all occur at the extremes of the water depth. The maximum values of R^2 are all above 0.9 and depth values where those values are found vary from one area to another, indicating that the correction model is able to invert bathymetry with high accuracy over a range of water depths.

3.3 *Intercomparison of accuracy*

Looking at the data separately for each region, we see that in the Southern Ocean, the accuracy reaches a maximum value in the bathymetry range of approximately 2 000 m to 1 600 m and then begins to slowly decline, with a steep drop occurring at approximately 200 m. In the Pacific Ocean, the accuracy increases more rapidly from the beginning, reaching an extreme value near 4 000 m, and then begins a gentle decline to near 0.5. The curve for the Atlantic data generally behaves similarly to the former dataset, with accuracies rising rapidly to a maximum near 4 000 m before starting a gentle decline to approximately 0.4, however, there is a large downward trend in the last 100 m of the gradient. Finally, in the Caribbean Sea data, the curve shows a different trend, with more isolated points at the water depth extremes and a more rugged overall curve trend. The fit lines show that the accuracy of the data in this region has local peaks at 5 000 m, 4 000 m to 3 000 m, and 2 500 m, with the last one being the highest, and the changes at the depth extremes are both steeper.

As a rule of thumb, the accuracy of deep learning models tends to be positively correlated with sample size. With other parameters held constant, the larger the data sample is, the higher the training accuracy tends to be, and vice versa. In the experiment, the analysis combined with the histogram of the data point distribution shows that the lowest R^2 tends to be found where the number of data points is the lowest or sparsest, such as at the water depth extremes. In contrast, the depths where higher levels of R^2 occur tend to be the depths with the largest and most concentrated distribution of data points. Then, the underwater morphological specificity in the experimental areas can also have some impact on the training accuracy. For example, in the Caribbean Sea dataset, the large and fragmented distribution of local islands can lead greater irregularity in the gravity and bathymetric undulations within some regions, which may become a factor affecting the output accuracy. In addition, since the spatial resolution of the Caribbean Sea data is approximately a quarter the size of the other three, it may also be one of the reasons that the accuracy results from this dataset were less satisfactory.

To make the accuracy validation more comparable, we reworked each of the four datasets according to the bathymetry refinement method presented in Chen et al. (2022) and overlay the fitted trend lines of the accuracy results shown in Fig. 6 for comparison. It can be summarized that, in general, our gravity-bathymetry correction method can effectively improve the accuracy compared to the bathymetry-only method. When considering the maximum accuracy within the datasets from each experimental area, we see that there is not much difference between the two methods. However, the gravity-bathymetry method is able to improve the accuracy mainly in the intermediate water depth range. Specifically, we see when we compare the performance results of the Southern Ocean dataset that the R^2 value of the bathymetry-only method starts to decline near 2 500 m, while the R^2 values from our method begin to surpass those values at that depth and reach a stable plateau until approximately 1 000 m, which is when the two methods align. In the comparison of the Pacific dataset, the performance of our results exceeds the other set of results from the maximum water depths until approximately 4 000 m. Between approximately 3 000 m and 1 000 m, our method also appears to perform significantly better. However, in shallow waters up to 1 000 m, our method underperforms the bathymetry-only method. In the Atlantic dataset, our method performs better for a long time within the depth range of approximately 4 500 m to 1 000 m but is again overtaken in shallow waters. In the Caribbean Sea dataset, the two methods perform similarly, with alternating leads occurring from time to time. Our method outperforms the bathymetry-only method mostly between depths below 5 000 m, between approximately 4 500 m and 3 500 m and between 2 000 m and 1 000 m.

The gravity-bathymetry correction method has a better R^2 values in the intermediate water depth range and can improve the overall accuracy as well; however, if we focus only on depths above approximately 1 000 m, its performance is inferior to that of the bathymetry-only method. This is presumed to be related to the band-limited correlation of the satellite gravity-bathymetry signals within the corresponding water depths (Smith and Sandwell, 1994, 1997; Sandwell et al., 2022).

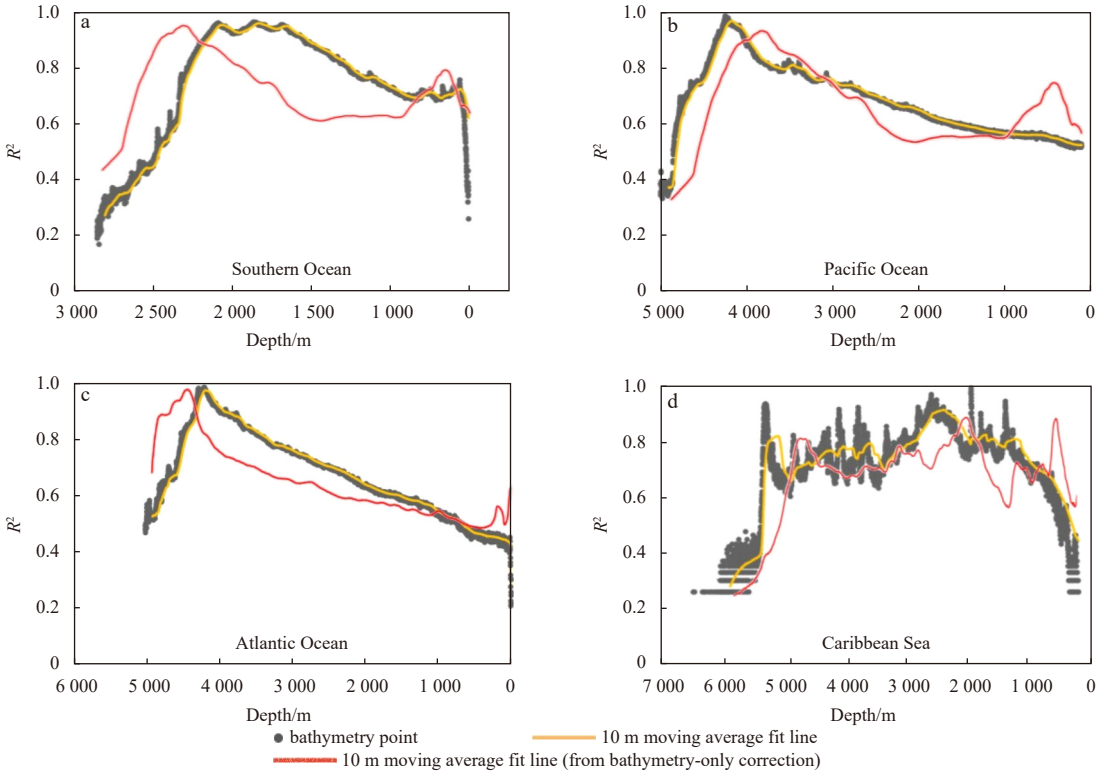

Fig. 6. R^2 values at different water depths compared with the bathymetry-only correction methos in the Southern Ocean (a), Pacific Ocean (b), Atlantic Ocean (c), and Caribbean Sea (d).

We also apply NRMSE, the normalized metric of RMSE, for the comparative evaluation of the correction effect in the four experimental areas, as plotted in Fig. 7 (the values on the horizontal axis are magnified by a factor of 1 000 for ease of display). From the figure, it can be found that the NRMSE performance curves of the four regions almost all follow the form of high in the middle-low at the edges; thus, it can be concluded that the majority of the data points are distributed in the interval of moderate performance. The extreme values of the performance of the four regions are almost identical and are distributed between 0.010 and 0.035. No obvious prominent peak areas can be seen in the performance curves of the Southern Ocean data; instead, a relatively stable plateau distribution with localized ups and downs is observed. The peak in the Pacific data occurs at approximately 0.017 and then begins to decline, but there is a small local peak at approximately 0.032, indicating a less desirable sudden change in accuracy at certain data points. The Atlantic dataset behaves similarly to that of the Pacific, with the difference that it has a bimodal structure with two local peaks between 0.015 and 0.020 to 0.025, with the former being higher than the latter. It is worth mentioning that the percentage of data points corresponding to the maximum peak in the curve of the Atlantic Ocean is the highest among the four datasets, which proves the more concentrated distribution of its high-precision data points. Compared with the right-skewed distribution exhibited by the upper two datasets, the curve of the Caribbean Sea dataset shows a clear left-skewed distribution, i.e., a larger distribution of lower precision data points, symbolizing a relatively unsatisfactory precision, despite its better-than-average peak share. It is also evident from Table 4 that the accuracy of the Caribbean data is relatively the lowest.

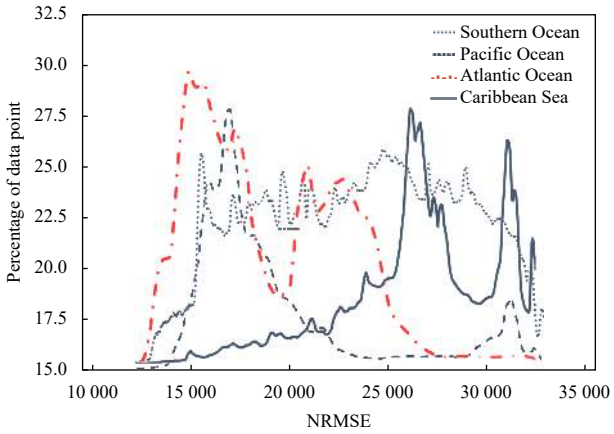

Fig. 7. Percentage distributions of NRMSE performance in datasets from the Southern Ocean, Pacific Ocean, Atlantic Ocean, and Caribbean Sea.

To demonstrate the effect of the correction more intuitively, we then subtract the corrected bathymetric values from the true values. The absolute values of the obtained differences are calculated as a percentage of the total water depth, with the results shown in Table 5. As seen from the table, the number of data points decreases as the margin of error narrows. The mean percentage of data points in the four experimental areas is approximately 71.62% within a 2% water depth range error and approximately 59.55% within a 1% error. A comparison with the results of the bathymetry-only correction method proposed in Chen et al. (2022) shows that our gravimetry-bathymetry combined correction method, while improving the accuracy by a mere 1.47% in the 2% water depth range error, improves by a larger 21.01% in the 1% error range. In this comparison, it can be concluded that the accuracy improvement provided by our novel

Table 5. Proportion of corrected errors from true values within 2% and 1% of the depth range

Location	Proportion of corrected errors/%	
	1% of depth	2% of depth
Southern Ocean	68.05	75.69
Pacific Ocean	61.53	77.52
Atlantic Ocean	57.58	68.69
Caribbean Sea	51.03	64.58

method lies to a large extent within the 1% water depth range error, despite its lack of significance in the 1% to 2% interval.

In summary, among these evaluation indicators, the performance of this correction method in the Southern Ocean is relatively better than in the other basins, while the Pacific Ocean and Atlantic Ocean performances are similar, and each has its own leading edge, and the performance of the method in Caribbean Sea ranks lowest amongst most cases.

4 Conclusions

In this study, we propose a deep learning-based inversion solution to address the current problem of the lack of high-precision topographic seafloor modeling. We output the refines satellite-derived gravity anomaly model, which was based on a pretrained VGGNet model, by training the shipborne gravity anomaly data while using a neural style transfer algorithm to minimize the distance (loss) between the two datasets. For the obtained corrected gravity anomaly model, we use the conventional filtering method to invert it and derive what can be considered a corrected bathymetry model.

Four gravity-bathymetry data pairs from the Southern Ocean, Pacific Ocean, Atlantic Ocean, and Caribbean Sea are selected to evaluate the accuracy of the corrected bathymetry model. Through the evaluation based on R^2, SD, RMSE and NRMSE metrics, we find that it can improve the accuracy by up to over 15% compared with the previous gravity-bathymetry correction method or over 24% compared with the previous bathymetry-only correction method. By plotting the R^2 and RMSE proportion as a function of water depth, we conclude that this method can effectively improve the fineness of the bathymetry. More quantitatively, our method offers the best correction for the top 1% of the total depth range of data points, which is more than 21% better than previous research. This demonstrates that the proposed model is useful and efficient as a novel approach to modeling submarine topography on a large scale of refinement.

For the factors affecting the differences in accuracy in different experimental areas, as mentioned before, the specificity of the seafloor topography can have an important effect. For instance, the scale of the water depth, the complexity of the geomorphology and the interference of land (islands) factors are all possibilities that can cause errors in the signals of the ship measurement instruments (Smith and Sandwell, 1994, 1997; Koh et al., 2022). For signals transmitted by satellites, the latitude of the operational orbit is a factor that affects their performance. In addition, the band limitations of the gravity-depth signal for altimetry and its relation to the geological environment also exert an influence (Scharroo and Visser, 1998). Combining the R^2 and NRMSE in Tables 3 and 4, we can reason that there is some degree of accuracy loss occurring during the inversion of the gravity model to the bathymetry model. The causes and mechanisms of this loss may require detailed analysis of the control variables. These systematic errors are all factors that must be taken into account.

For the factors of the model itself, the generalization can have an impact on its ability to predict new unknown data. Currently, only the Adam optimizer is applied in this pretrained model. A comparison of different types of optimizers should be discussed before future experiments are undertaken in order to select the one that can balance the loss and generalization performance (Reddi et al., 2018). Furthermore, with larger data involved, instead of dividing the training and validation sets within the dataset itself, it is also worth discussing whether selecting a completely new dataset as the validation set to be applied in the training process will improve the correction accuracy of the model.

Acknowledgements

The authors would like to express gratitude to Bureau Gravimetrique International (BGI) and National Geophysical Data Center (NGDC) of National Oceanic and Atmospheric Administration (NOAA) for open-source data services. In the meantime, the authors would like to express their gratitude to Google AI Hub for providing open-source pretrained model algorithms.

References

Airy G B. 1855. III. On the computation of the effect of the attraction of mountain-masses, as disturbing the apparent astronomical latitude of stations in geodetic surveys. Philosophical Transactions of the Royal Society of London, 145: 101–104, doi: 10.1098/rstl.1855.0003

An Dechao, Guo Jinyun, Li Zhen, et al. 2022. Improved gravity-geologic method reliably removing the long-wavelength gravity effect of regional seafloor topography: a case of bathymetric prediction in the South China Sea. IEEE Transactions on Geoscience and Remote Sensing, 60: 4211912, doi: 10.1109/TGRS.2022.3223047

Annan R F, Wan Xiaoyun. 2022. Recovering bathymetry of the gulf of guinea using altimetry-derived gravity field products combined via convolutional neural network. Surveys in Geophysics, 43(5): 1541–1561, doi: 10.1007/s10712-022-09720-5

Benedetti P, Ienco D, Gaetano R, et al. 2018. M^3Fusion: a deep learning architecture for multiscale multimodal multitemporal satellite data fusion. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 11(12): 4939–4949, doi: 10.1109/JSTARS.2018.2876357

Bonvalot S, Balmino G, Briais A, et al. 2012. World gravity map. commission for the geological map of the world. Paris, France: BGI–CGMW-CNES-IRD

Braitenberg C, Wienecke S, Wang Yong. 2006. Basement structures from satellite-derived gravity field: South China Sea ridge. Journal of Geophysical Research: Solid Earth, 111(B5): B05407, doi: 10.1029/2005JB003938

Chen Xiaolun, Luo Xiaowen, Wu Ziyin, et al. 2022. A VGGNet-based method for refined bathymetry from satellite altimetry to reduce errors. Remote Sensing, 14(23): 5939, doi: 10.3390/rs14235939

Colbo K, Ross T, Brown C, et al. 2014. A review of oceanographic applications of water column data from multibeam echosounders. Estuarine, Coastal and Shelf Science, 145: 41–56, doi: 10.1016/j.ecss.2014.04.002

Coley K. 2022. A global ocean map is not an ambition, but a necessity to support the ocean decade. Marine Technology Society Journal, 56(3): 9–12, doi: 10.4031/MTSJ.56.3.3

Fan Diao, Li Shanshan, Li Xinxing, et al. 2021. Seafloor topography estimation from gravity anomaly and vertical gravity gradient using nonlinear iterative least square method. Remote Sensing, 13(1): 64, doi: 10.3390/rs13010064

Fan Diao, Li Shanshan, Meng Shuyu, et al. 2020. Applying iterative method to solving high-order terms of seafloor topography. Marine Geodesy, 43(1): 63–85, doi: 10.1080/01490419.2019.

1670298
Gatys L A, Ecker A S, Bethge M. 2016. A neural algorithm of artistic style. Journal of Vision, 16(12): 326, doi: 10.1167/16.12.326
Gatys L A, Ecker A S, Bethge M, et al. 2017. Controlling perceptual factors in neural style transfer. In: Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA: IEEE, 3730–3738, doi: 10.1109/CVPR.2017.397
Gong Zheng, Zhang Peizhen, Zheng Wenjun, et al. 2021. The effect of altimetry data in estimating the elastic thickness of the lithosphere in the western Pacific Ocean. Geodesy and Geodynamics, 12(5): 315–322, doi: 10.1016/j.geog.2021.07.001
Hu Minzhang, Zhang Shengjun, Jin Taoyong, et al. 2020. A new generation of global bathymetry model BAT_WHU2020. Acta Geodaetica et Cartographica Sinica (in Chinese), 49(8): 939–954, doi: 10.11947/j.AGCS.2020.20190526
Hughes Clarke J E. 2018. Multibeam echosounders. In: Micallef A, Krastel S, Savini A, eds. Submarine Geomorphology. Cham: Springer, 25–41, doi: 10.1007/978-3-319-57852-1_3
Huo Guanying, Wu Ziyin, Li Jiabiao. 2020. Underwater object classification in sidescan sonar images using deep transfer learning and semisynthetic training data. IEEE Access, 8: 47407–47418, doi: 10.1109/ACCESS.2020.2978880
Ibrahim A, Hinze W J. 1972. Mapping buried bedrock topography with gravity. Groundwater, 10(3): 18–23, doi: 10.1111/j.1745-6584.1972.tb02921.x
Islam M J, Xia Youya, Sattar J. 2020. Fast underwater image enhancement for improved visual perception. IEEE Robotics and Automation Letters, 5(2): 3227–3234, doi: 10.1109/LRA.2020.2974710
Jena B, Kurian P J, Swain D, et al. 2012. Prediction of bathymetry from satellite altimeter based gravity in the Arabian Sea: mapping of two unnamed deep seamounts. International Journal of Applied Earth Observation and Geoinformation, 16: 1–4, doi: 10.1016/j.jag.2011.11.008
Jia Yangqing, Shelhamer E, Donahue J, et al. 2014. Caffe: Convolutional architecture for fast feature embedding. In: Proceedings of the 22nd ACM International Conference on Multimedia. Orlando, USA: Association for Computing Machinery, 675–678, doi: 10.1145/2647868.2654889
Koh Z W, Nimmo F, Lunine J I, et al. 2022. Assessing the detectability of Europa's seafloor topography from Europa clipper's gravity data. The Planetary Science Journal, 3(8): 197, doi: 10.3847/PSJ/ac82aa
Krizhevsky A, Sutskever I, Hinton G E. 2017. ImageNet classification with deep convolutional neural networks. Communications of the ACM, 60(6): 84–90, doi: 10.1145/3065386
Li Jiabiao. 1999. Multibeam Survey Principles, Techniques and Methods (in Chinese). Beijing: China Ocean Press
Moran N P. 2020. Machine learning model selection for predicting global bathymetry [dissertation]. New Orleans, USA: University of New Orleans
NOAA National Centers for Environmental Information. 2004. Multibeam bathymetry database (MBBDB). NOAA National Centers for Environmental Information. https://www.ncei.noaa.gov/maps/bathymetry/[2022-6-14]
NOAA National Centers for Environmental Information. 2015. Marine trackline geophysical database. NOAA National Centers for Environmental Information. https://www.ncei.noaa.gov/maps/geophysics/[2022-06-14]
Oldenburg D W. 1974. The inversion and interpretation of gravity anomalies. Geophysics, 39(4): 526–536, doi: 10.1190/1.1440444
Otter D W, Medina J R, Kalita J K. 2021. A survey of the usages of deep learning for natural language processing. IEEE Transactions on Neural Networks and Learning Systems, 32(2): 604–624, doi: 10.1109/TNNLS.2020.2979670
Ouyang Mingda, Sun Zhongmiao, Zhai Zhenhe. 2014. Predicting bathymetry in South China Sea using the gravity-geologic method. Chinese Journal of Geophysics (in Chinese), 57(9): 2756–2765, doi: 10.6038/cjg20140903
Ouyang Mingda, Sun Zhongmiao, Zhai Zhenhe, et al. 2015. Bathymetry prediction based on the admittance theory of gravity anomalies. Acta Geodaetica et Cartographica Sinica (in Chinese), 44(10): 1092–1099, doi: 10.11947/j.AGCS.2015.20140427
Parker R L. 1973. The rapid calculation of potential anomalies. Geophysical Journal International, 31(4): 447–455, doi: 10.1111/j.1365-246X.1973.tb06513.x
Reddi S J, Kale S, Kumar S. 2018. On the convergence of Adam and beyond. In: Proceedings of the 6th International Conference on Learning Representations. Vancouver, Canada: OpenReview.net
Sandwell D T, Goff J A, Gevorgian J, et al. 2022. Improved bathymetric prediction using geological information: SYNBATH. Earth and Space Science, 9(2): e2021EA002069, doi: 10.1029/2021EA002069
Scharroo R, Visser P. 1998. Precise orbit determination and gravity field improvement for the ERS satellites. Journal of Geophysical Research: Oceans, 103(C4): 8113–8127, doi: 10.1029/97JC03179
Schulz M A, Yeo B T T, Vogelstein J T, et al. 2020. Different scaling of linear models and deep learning in UKBiobank brain images versus machine-learning datasets. Nature Communications, 11(1): 4238, doi: 10.1038/s41467-020-18037-z
Seoane L, Ramillien G, Beirens B, et al. 2022. Regional seafloor topography by extended Kalman filtering of marine gravity data without ship-track information. Remote Sensing, 14(1): 169, doi: 10.3390/rs14010169
Shi Bo, Lu Xiushan, Yang Fanlin, et al. 2017. Shipborne over- and under-water integrated mobile mapping system and its seamless integration of point clouds. Marine Geodesy, 40(2/3): 104–122, doi: 10.1080/01490419.2016.1272510
Simonyan K, Zisserman A. 2015. Very deep convolutional networks for large-scale image recognition. In: Proceedings of the 3rd International Conference on Learning Representations. San Diego, USA: Association for Computing Machinery
Smith W H F, Sandwell D T. 1994. Bathymetric prediction from dense satellite altimetry and sparse shipboard bathymetry. Journal of Geophysical Research: Solid Earth, 99(B11): 21803–21824, doi: 10.1029/94JB00988
Smith W H F, Sandwell D T. 1997. Global sea floor topography from satellite altimetry and ship depth soundings. Science, 277(5334): 1956–1962, doi: 10.1126/science.277.5334.1956
Walcott R I. 1970. Flexural rigidity, thickness, and viscosity of the lithosphere. Journal of Geophysical Research, 75(20): 3941–3954, doi: 10.1029/JB075i020p03941
Watts A B. 1978. An analysis of isostasy in the world's oceans 1. Hawaiian-Emperor Seamount Chain. Journal of Geophysical Research: Solid Earth, 83(B12): 5989–6004, doi: 10.1029/JB083iB12p05989
Watts A B. 2001. Isostasy and Flexure of the Lithosphere. Cambridge, UK: Cambridge University Press
Watts A B, Sandwell D T, Smith W H F, et al. 2006. Global gravity, bathymetry, and the distribution of submarine volcanism through space and time. Journal of Geophysical Research: Solid Earth, 111(B8): B08408, doi: 10.1029/2005JB004083
Wei Zhijie, Guo Jinyun, Zhu Chengcheng, et al. 2021. Evaluating accuracy of HY-2A/GM-derived gravity data with the gravity-geologic method to predict bathymetry. Frontiers in Earth Science, 9: 636246, doi: 10.3389/feart.2021.636246
Wu Meiyin, Chen Li. 2015. Image recognition based on deep learning. In: Proceedings of 2015 Chinese Automation Congress (CAC). Wuhan, China: IEEE, 542–546, doi: 10.1109/CAC.2015.7382560
Wu Ziyin, Yang Fanlin, Tang Yong, et al. 2020. High-Resolution Seafloor Survey and Applications. Beijing: Science Press
Yale M M, Sandwell D T, Herring A T. 1998. What are the limitations of satellite altimetry?. The Leading Edge, 17(1): 73–76, doi: 10.1190/1.1437832
Yuan Qiangqiang, Shen Huanfeng, Li Tongwen, et al. 2020. Deep learning in environmental remote sensing: achievements and challenges. Remote Sensing of Environment, 241: 111716, doi:

10.1016/j.rse.2020.111716

Zhao Jianhu, Ouyang Yongzhong, Wang Aixue. 2017. Status and development tendency for seafloor terrain measurement technology. Acta Geodaetica et Cartographica Sinica (in Chinese), 46(10): 1786–1794, doi: 10.11947/j.AGCS.2017.20170276

Zhu Chengcheng, Guo Jinyun, Yuan Jiajia, et al. 2021. Refining altimeter-derived gravity anomaly model from shipborne gravity by multi-layer perceptron neural network: a case in the South China Sea. Remote Sensing, 13(4): 607, doi: 10.3390/rs13040607

Zwally H J, Schutz B, Abdalati W, et al. 2002. ICESat's laser measurements of polar ice, atmosphere, ocean, and land. Journal of Geodynamics, 34(3/4): 405–445, doi: 10.1016/S0264-3707(02)00042-X