frontiers | Research Topics

Spatiotemporal modeling and analysis in marine science

Edited by
Junyu He, Xi Xiao and Zhenhong Du

Published in
Frontiers in Marine Science

ISSN 1664-8714
ISBN 978-2-8325-3744-2
DOI 10.3389/978-2-8325-3744-2

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Spatiotemporal modeling and analysis in marine science

Topic editors

Junyu He — Zhejiang University, China

Xi Xiao — Zhejiang University, China

Zhenhong Du — Zhejiang University, China

Citation

He, J., Xiao, X., Du, Z., eds. (2023). *Spatiotemporal modeling and analysis in marine science*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-3744-2

Table of contents

frontiers | Frontiers in Marine Science

TYPE Editorial
PUBLISHED 05 October 2023
DOI 10.3389/fmars.2023.1296334

OPEN ACCESS

EDITED AND REVIEWED BY
Hervé Claustre,
Centre National de la Recherche Scientifique (CNRS), France

*CORRESPONDENCE
Junyu He
jxgzhejunyu@163.com

RECEIVED 18 September 2023
ACCEPTED 29 September 2023
PUBLISHED 05 October 2023

CITATION
He J, Du Z and Xiao X (2023) Editorial: Spatiotemporal modeling and analysis in marine science. *Front. Mar. Sci.* 10:1296334. doi: 10.3389/fmars.2023.1296334

Editorial: Spatiotemporal modeling and analysis in marine science

Junyu He[1,2,3]*, Zhenhong Du[2,4] and Xi Xiao[1,2]

[1]Ocean College, Zhejiang University, Zhoushan, China, [2]Ocean Academy, Zhejiang University, Zhoushan, China, [3]Donghai Laboratory, Zhoushan, China, [4]School of Earth Sciences, Zhejiang University, Hangzhou, China

KEYWORDS

spatiotemporal, modeling, marine science, machine learning, time-frequency

Editorial on the Research Topic
Spatiotemporal modeling and analysis in marine science

Recent advancements in monitoring technologies have ushered in an era characterized by an explosion of spatiotemporal data across diverse fields, notably within global oceans. This surge of data, sourced from satellites, unmanned aerial vehicles, buoys, and unmanned underwater vehicles, poses challenges in extracting meaningful insights and bridging the gap between raw data and scientific understanding (Frankel and Reid, 2008; Wu et al., 2020). Data-driven spatiotemporal modeling and analysis present potential solutions, offering avenues to uncover inherent features within the data, delineate characteristics of natural phenomena, refine general knowledge or theories, and bolster the management and conservation of oceans. Consequently, based on these spatiotemporal, spatial or temporal data, specific methodologies are tailored and employed for analysis. For instance, prevalent machine learning and deep learning techniques excel in modeling the nonlinear and non-stationary aspects of spatiotemporal, spatial, or temporal data (Ham et al., 2019; Reichstein et al., 2019; Runge et al., 2019; Callaghan et al., 2021). Geostatistical methods adeptly capture spatial or spatiotemporal correlations within such data (He and Kolovos, 2018; He and Christakos, 2021; Wu et al., 2021). Additionally, time-frequency methods decompose temporal data into distinct series for nuanced variation detection (Cazelles et al., 2008; Xiao et al., 2019). Cutting-edge frameworks, such as those integrating geostatistical and machine learning methods, have been proposed to more accurately reflect the natural phenomena (Du et al., 2021). The primary objective of this Research Topic is to furnish a platform for scholars to disseminate novel methodologies or insights within the spatiotemporal context of marine or coastal regions.

Chen et al. posited that integrating artificial intelligence (AI) technology with extensive datasets from ocean observations can significantly address challenges in advancing marine science or theory, especially in the intermediate ocean depths ranging from approximately 100 to 1,000 meters. In their review titled “Deep Blue AI for Knowledge Discovery of Intermediate Ocean,” they synthesized findings from satellite remote sensing at around 100 meters deep and *in situ* observations at approximately 1,000 meters. Additionally, they

discussed three distinct AI methodologies: associative statistical, physically informed, and mathematically driven neural networks. The application of these methodologies was reviewed in context, covering areas such as the 3-D identification and trajectory prediction of oceanic eddies, the vertical reconstruction of Ekman drift, internal wave forecasting, and subsurface chlorophyll maxima prediction.

Wang et al. centered their research on predicting typhoon activity to bolster decision-making aimed at mitigating associated risks. A pivotal challenge in modeling arose when selecting the optimal number of historical satellite images for future predictions. Using too few images might not offer adequate information on typhoon trajectories, while an excessive number could diminish the size of the training dataset. To address this, the team devised a feature enhancement module paired with a channel attention module. This combination was designed to amplify the intrinsic characteristics of typhoons and determine the appropriate number of images for subsequent modeling. The spatiotemporal attributes of typhoons were then modeled using a symmetrical encode-decode module comprising convolutional long short-term memory networks (ConvLSTM). Furthermore, a multi-scale strategy was implemented to curtail information loss during the ConvLSTM process. In their quest to design the optimal structure for the Enhanced Multi-Scale Deep Neural Network (EMSN), the team explored multi-scale components, channel attention modules, and spatiotemporal capture units. By adjusting various parameters within these modules, they assessed the accuracy of typhoon predictions. Notably, the EMSN outperformed both the MSCIP satellite image predictor and the conventional U-net.

While hyperspectral remote sensing data are limited and costly, Hu et al. investigated the potential of reconstructing hyperspectral images using economical RGB images, a development that could enhance marine observations. By replacing the standard convolution kernel with atrous convolution and incorporating a multi-scale atrous convolution residual block, the image's multi-scale spatial features were more effectively extracted by integrating images across different scales into a multi-scale feature layer without increasing computational costs. The introduced Multi-Scale Atrous Convolution Residual Network (MACRN) comprises three segments: low-level feature extraction, high-level feature extraction, and feature transformation. MACRN's efficacy was evaluated using clean and real-world datasets and benchmarked against exsiting algorithms like HSCNN-R, HSCNN-D, HRNet, AWAN, and MST++.

Li devoted significant effort to applying the greedy strategy for mode identification with significantly crossed frequencies and overlapped component separation. Consequently, the Spatio-Temporal Nonconvex Penalty Adaptive Chirp Mode Decomposition (STNP-ACMD) algorithm was proposed. It addresses the limitations of traditional algorithms, such as ACMD and Variational Mode Decomposition (VMD), which primarily prioritize channel-wise processing without accommodating the coupled nature and spatio-temporal characteristics. The STNP-ACMD employs a recursive mode extraction approach to segregate overlapped components or crossed intrinsic functions and accentuates the spatio-temporal relationship within the coupled nature by refining the spatial and temporal matrices. Comparative analysis, using both numerical and real-world case studies, demonstrated that STNP-ACMD surpassed the Ensemble Empirical Mode Decomposition and VMD.

Chiu et al. leveraged the spatial correlation traits found within observed data to estimate blue carbon (BC) stocks in marsh soils. A Bayesian linear mixed-effects model was devised to assimilate auxiliary information from variables such as marsh type, soil category, soil depth, and marsh site. This model assumed that site effects adhered to the intrinsic conditional autoregressive (ICAR) spatial dependence. By incorporating the ICAR marsh site effects, an inherent spatial clustering of sites became discernible, a pattern not reflected in the primary auxiliary predictors. Notably, the ICAR model produced a narrower confidence interval for the marsh site effects' coefficient compared to its non-ICAR counterpart, highlighting the spatial correlation of BC stock and underscoring the importance of including these characteristics in BC assessment models.

Author contributions

JH: Funding acquisition, Writing – original draft, Writing – review & editing. ZD: Writing – review & editing. XX: Writing – review & editing.

Funding

This work is supported by National Natural Science Foundation of China (Nos. 42301374 and 42171398), the Major Program of Natural Science Foundation of Zhejiang for Creative Research Group (Grant No. LDT23D06024D06) and Science Foundation of Donghai Laboratory (Grant No. 2022KF01012).

Acknowledgments

We thank Dr. Dongxiao Zhang for his help on handling one manuscript submitted to this Research Topic. We also thank all authors of the articles published in this Research Topic for their significant contributions and the reviewers for their valuable comments. We appreciate the kindly help and supports from Rui Fernandes, Ronan Conlon, Alice Lickley, Joshua Chew, Minea Kaartinen, Francesco Rossetti, and Emily Keynton.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

References

Callaghan, M., Schleussner, C.-F., Nath, S., Lejeune, Q., Knutson, T. R., Reichstein, M., et al. (2021). Machine-learning-based evidence and attribution mapping of 100,000 climate impact studies. *Nat. Clim. Change* 11, 966–972. doi: 10.1038/s41558-021-01168-6

Cazelles, B., Chavez, M., Berteaux, D., Ménard, F., Vik, J. O., Jenouvrier, S., et al. (2008). Wavelet analysis of ecological time series. *Oecologia* 156, 287–304. doi: 10.1007/s00442-008-0993-2

Du, Z., Qi, J., Wu, S., Zhang, F., and Liu, R. (2021). A spatially weighted neural network based water quality assessment method for large-scale coastal areas. *Environ. Sci. Technol.* 55, 2553–2563. doi: 10.1021/acs.est.0c05928

Frankel, F., and Reid, R. (2008). Big data: Distilling meaning from data. *Nature* 455, 30–30. doi: 10.1038/455030a

Ham, Y.-G., Kim, J.-H., and Luo, J.-J. (2019). Deep learning for multi-year ENSO forecasts. *Nature* 573, 568–572. doi: 10.1038/s41586-019-1559-7

He, J., and Christakos, G. (2021). "Bayesian maximum entropy," in *Encyclopedia of mathematical geosciences, encyclopedia of earth sciences series*. Eds. B. S. Daya Sagar, Q. Cheng, J. McKinley and F. Agterberg (Cham: Springer International Publishing), 1–9. doi: 10.1007/978-3-030-26050-7_50-1

He, J., and Kolovos, A. (2018). Bayesian maximum entropy approach and its applications: a review. *Stoch Environ. Res. Risk Assess.* 32, 859–877. doi: 10.1007/s00477-017-1419-7

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., et al. (2019). Deep learning and process understanding for data-driven Earth system science. *Nature* 566, 195–204. doi: 10.1038/s41586-019-0912-1

Runge, J., Bathiany, S., Bollt, E., Camps-Valls, G., Coumou, D., Deyle, E., et al. (2019). Inferring causation from time series in Earth system sciences. *Nat. Commun.* 10, 2553. doi: 10.1038/s41467-019-10105-3

Wu, L., Chen, Z., Lin, X., and Liu, Y. (2020). Building the integrated observational network of "Transparent Ocean." *CSB* 65 (25), 2654–2661. doi: 10.1360/TB-2020-0558

Wu, J., He, J., and Christakos, G. (2021). *Quantitative analysis and modeling of earth and environmental data* (Amsterdam, Netherlands: Elsevier).

Xiao, X., He, J., Yu, Y., Cazelles, B., Li, M., Jiang, Q., et al. (2019). Teleconnection between phytoplankton dynamics in north temperate lakes and global climatic oscillation by time-frequency analysis. *Water Res.* 154, 267–276. doi: 10.1016/j.watres.2019.01.056

frontiers | Frontiers in Marine Science

TYPE Original Research
PUBLISHED 15 September 2022
DOI 10.3389/fmars.2022.945835

OPEN ACCESS

EDITED BY
Junyu He,
Zhejiang University, China

REVIEWED BY
Gael Alory,
UMR5566 Laboratoire d'études en géophysique et océanographie spatiales (LEGOS), France
Qingxuan Yang,
Ocean University of China, China

*CORRESPONDENCE
Dandan Zhao
zhdd82@163.com

SPECIALTY SECTION
This article was submitted to Ocean Observation, a section of the journal Frontiers in Marine Science

RECEIVED 09 June 2022
ACCEPTED 25 August 2022
PUBLISHED 15 September 2022

CITATION
Tang B, Zhao D, Cui C and Zhao X (2022) Reconstruction of ocean temperature and salinity profiles in the Northern South China Sea using satellite observations.
Front. Mar. Sci. 9:945835.
doi: 10.3389/fmars.2022.945835

Reconstruction of ocean temperature and salinity profiles in the Northern South China Sea using satellite observations

Bo Tang[1,2], Dandan Zhao[1,2*], Chaoran Cui[1,2] and Xinhua Zhao[3]

[1]School of Ocean Technology Science, Qilu University of Technology, Jinan, China,
[2]Institute of Oceanographic Instrumentation, Qilu University of Technology, Qingdao, China,
[3]Jiangsu Institute of Marine Resources Development, Jiangsu Ocean University, Lianyungang, China

Based on historical temperature and salinity (T–S) profiles, the correlation between the sea surface temperature (SST) anomaly, sea surface dynamic height (SSH) anomaly, and temperature profile anomaly is constructed by regression analysis. A three-dimensional temperature field is reconstructed in the northern South China Sea by satellite SST and SSH, with daily temporal and 0.25°×0.25° spatial resolutions. The three-dimensional salinity field is also reconstructed based on the correlation between salinity and temperature. Compared with the observational T–S profiles, the reconstructed T–S field reflects the characteristics and structure and accurately describes the mesoscale variability of the ocean temperature field. The new expanded T–S field can be used as the initial field in numerical models and be assimilated into numerical reanalysis and prediction systems, improving their output.

KEYWORDS

northern south china sea, reconstruction of t–s profiles, regression analysis, satellite sea surface temperature, satallite sea surface dynamic height

1 Introduction

The northern South China Sea is a large marginal sea in the tropics with a maximum depth of over 5000 m and a variable temperature and salinity (T–S) structure with many mesoscale eddies, which mainly come from the unstable changes in the process of Kuroshio intrusion and wind stress curl changes (Wang et al., 2003; Su, 2004; Wu et al., 2007; Yuan et al., 2007; Hu and Hou, 2010). The T-S profiles of this region is the key database to study the eddy variation and other ocean processes.

At present, the availability of T–S observation profiles from Argo (Array for Real-time Geostrophic Oceanography), CTD (conductivity-temperature-depth), XBT (expendable bathythermograph), and ocean stations is increasing, especially the massive Argo observation profiles have played an essential role in the ocean

observation system. For example, Nan et al. (2011) analysed the vertical structures of three long-lived anticyclonic eddies in the northern South China Sea by *in situ* measurements. Hu et al. (2012) demonstrated the penetration of nonlinear Rossby eddies into the South China Sea with cruise data. But real-time observation is still challenging, and the spatial resolution is low.

However, with the continued development of satellite remote sensing technology, sea surface temperature (SST) and sea surface dynamic height (SSH) datasets provide many real-time observations with high spatial resolution on the sea surface. For example, based on these satellite observations, detailed statistical characteristics of mesoscale eddies, including eddy census statistics, kinematic properties, shapes, nonlinearity and propagation characteristics, have been provided in many studies of the global ocean (Chaigneau et al., 2009; Chelton et al., 2011; Li et al., 2014; Li et al., 2016), especially for eddy-active regions like the South China Sea (Chen et al., 2011; Hu et al., 2011; Chen et al., 2012; Hu et al., 2012). However, satellite observations cannot provide any subsurface information. Therefore, the combination of *in situ* and satellite observation has become an effective method to reconstruct a more detailed ocean T–S field.

As early as the end of the 1980s, several methods for inverting T–S profiles by mapping sea surface information (SST or SSH) have been proposed. Hurlburt (1986) has constructed a numerical ocean model to transform simulated altimeter data into subsurface information dynamically. Studies and applications of satellite observations for inferring sea subsurface information matured using statistical methods (Carnes et al., 1990; Hurlburt et al., 1990; Carnes et al., 1994; Gavart and Mey, 1997; Pascual and Gomis, 2003), which are mainly based on empirical orthogonal functions (EOFs). These researchers showed that the deep sea or oceanic temperature and salinity profiles obtained by the combined inversion of SST and SSH is much better than the results using SST or SSH alone. Furthermore, Bruno and Santoleri (2004) used coupled pattern analysis to build the relationship between temperature profiles and sea surface information to rebuild the temperature field. The EOF and coupled pattern methods are simple and clearly reflect the physical concepts, but both require observation data with a certain continuity in time and space, which is limited in the actual ocean. Regression analysis is a data analysis method based on mathematical and statistical principles (Fox et al., 2002; Guinehut et al., 2004). First, a lot of statistics are mathematically processed, then a mathematical function expression (regression equation) with strong correlation is established by determining the correlation between the target variable and some independent variables, finally it is generalized to predict the possible changes of the dependent variable in the future. In this study, the regression analysis method was used to establish the mapping relationship between SST, SSH and temperature and salinity profiles using historical T–S profile observation data after strict quality control and fine processing for the northern South China Sea over a recent 30-year period. Then, the T–S profiles were reconstructed and examined. Compared with EOF and coupled pattern methods, althothg the regression analysis method lacks clear physical concept, it does not require time-space continuity of historical observation data, and the calculation method is simpler and more maneuverable (Fox et al., 2002).

This study is organised as follows. The historical and satellite observations used in this study are introduced in Section 2. In Section 3, the reconstruction of T–S profiles in the northern South China Sea based on the regression analysis method is described in detail. The results are examined using observed data, both in general and in case analyses. Finally, conclusions are summarised in Section 4.

2 Data and methodology

2.1 Historical argo profiles observations

The historical T–S profile data used in this study are based on the WOD18 (World Ocean Database, 2018) produced by the NODC (National Oceanographic Data Center), which mainly includes high-resolution CTD, XBT, DRB (drifting buoy), PFL (profiled buoy) and MRB (anchor buoy) data (Boyer et al., 2018). These data were supplemented by Argo float data obtained from the China Argo Real-time Data Center. The spatial range is from 110°E to 125°E, 15°N to 25°N, and the time range is from 1 January 1993 to 1 January 2018. The distribution of the Argo stations is shown in Figure 1, and they are spread throughout the northern South China Sea. Furthermore, quality control for this extensive profile dataset is essential to improve the accuracy of the results (Chaigneau et al., 2011). Thus, a valid database is constructed in several steps as follows:

1. Extract the required data including temperature, salinity, longitude, latitude, depth and date in the defined spatial and time range;
2. Delete data with NaN (not a number) values;
3. Select the data with a temperature between 0°C and 35° C and salinity between 31 psu and 36 psu;
4. Delete stations with less than three values in the vertical direction;
5. Construct a database with the standard vertical layer by linear interpolation for depths of 0, 5, 15, 20, 25, 30, 35, 50, 75, 100, 125, 150, 175, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400 and 1500 m.

Although the WOD18 data provides very high vertical resolution data, it was found in the process of data screening that the profiles change very little at depths below 500m and the profiles above 500m are highly variable. Furthermore, most ocean processes and phenomena are concentrated above 1500m, such as eddies, material transport. Therefore, we divided a total of 20

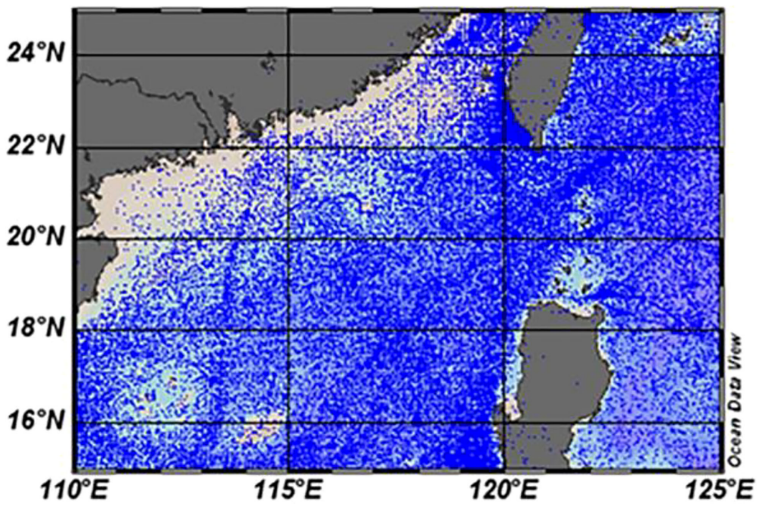


FIGURE 1
Distribution of observed T–S profile stations in the northern South China Sea.

layers between the sea surface and 500m, which is basically the same as WOD18 data. However, the deep water layer with no obvious change were divided into 10 layers from 500m to 1500m, which would not affect the overall research results, and reduce a lot of calculation in the regression statistical analysis.

2.2 Climatic T–S data

The climatic T–S data based on the WOA18 (World Ocean Atlas, 2018) from NODC were regarded as an initial field and compared with the reconstructed results (Garcia et al., 2019). The WOA18 dataset provides standard layer data that have been objectively analysed. The spatial resolution is 1°, and the temporal sampling includes both annual average and monthly average versions. Comparing the reconstructed fields to these data can indicate T–S anomalies caused by eddies in the three-dimensional structure because the WOA18 data are excessively smoothed, and the eddy signals are significantly suppressed.

2.3 Test T–S data

The test T–S data were obtained from GTSPP(Global Temperature and Salinity Profile Plan) and Argo datasets. A total of 18.174 million T-S observation profiles have been provided by GTSPP from 1990 to 2014. The datasets mainly include XBT, which has strong timeliness and can be downloaded from NODC website. The Argo dataset includes the global Argo dataset and the Chinese Argo dataset, which collectively provide more than 1.4 million T-S observation profiles and can be downloaded from the China Argo Real-time Data Center website. These test data were subjected to strict quality control and then interpolated to the standard layer for the test. The observation profiles used for the test were not used during the reconstruction process and were strictly independent.

2.4 Satellite observations

The SST data used in this study are from a dataset provided by American Remote Sensing Systems (ARSS) that combines several observations from the optimal interpolation MW (microwave) and IR (infrared) sensors (Gentemann et al., 2010), including the TRMM (Tropical Rainfall Measuring Mission) Microwave Imager, AMSR-E, AMSR-2 (Advanced Microwave Scanning Radiometer), WindSat and MODIS (Moderate Resolution Imaging Spectroradiometer). The spatial resolution is 9 km, and the temporal sampling is 1 day. Moreover, the SSH data from DUACS (Data Unification and Altimeter Combination System) DT2014 from 1993 to 2018 with a spatial resolution of 0.25° and temporal sampling of 1 day (Pujol et al., 2016) were used in this study. It is provided by AVISO (Archiving, Validation and Interpretation of Satellite Oceanographic) and the datasets combined data from several altimeters, including T/P (Topex/Poseidon Satellite), ERS-1&2 (European Remote Sensing satellites 1&2), Jason-1&2, EnviSat (Environmental Satellite) and GFO (Geo Follow-On), which could be downloaded from the Copernicus Marine Environment Monitoring Service (CMEMS). And the sea level anomaly (SLA) data used in this study from multi-satellite altimeter is also provided by AVISO and has the same resolution as the SSH data, which could be downloaded from the CMEMS (Taburet et al., 2019).

2.5 Reconstruction of temperature profiles by SST and SSH

At present, there are three regression models for retrieving three-dimensional temperature field from sea surface information: The polynomial of SST, the polynomial of SSH, and the polynomial of the combination of SST and SSH are used. The third polynomial is a good combination of the two satellite observation data of SST and SSH. Previous studies have pointed out that the three-dimensional ocean temperature profile obtained by the joint inversion of SST and SSH is better than the profile obtained by the inversion of SST or SSH alone (Guinehut et al., 2004), so the polynomial combined with SST and SSH is directly used in this paper.

First, the sea surface dynamic height at each location can be calculated from its T–S profile as:

$$h = \int_0^H \frac{[v(T,S,P) - v(0,35,P)]}{v(0,35,P)} dz \quad (1)$$

where v is the seawater specific volume, $v(0,35,P)$ is the specific volume of seawater at 0°C temperature and 35 psu salinity, and H is the sea depth. Then, the correlation between the SST, SSH and the temperature profiles are established by regression analysis, expressed as:

$$T_{i,k}(SST,h) = \overline{T_{i,k}} + a_{i,k}^{T3}(SST - \overline{T_{i,1}}) + a_{i,k}^{T4}(h - \overline{h_i}) + a_{i,k}^{T5}\left[(SST - \overline{T_{i,1}})(h - \overline{h_i}) - \overline{hSST_i}\right] \quad (2)$$

Where $T_{i,k}(SST,h)$ are the values of the reconstructed temperature for extended grid point i and depth k, $\overline{hSST_i}$ is the weighted mean of the product, $(sst_j - \overline{T_{i,1}})$ $(h_j - \overline{h_i})$, j is the observation index, i is the location index, and $a_{i,k}^{T3}$, $a_{i,k}^{T4}$ and $a_{i,k}^{T5}$ are regression coefficients, they can be calculated by statistical method (Fox et al., 2002). The regression coefficients in equation (2) are obtained at monthly timescale, by comparing the monthly static climatic fields calculated from WOD18 to monthly-averaged satellite observations, the 3D T/S fields at daily time scale can be reconstructed by satellite daily resolution.

It is important to note that we only took one kind of polynomial of the structure, but the regression coefficients in each layer were determined by the plenty of T–S data, which are different in each layer. Therefore, the temperature field extended from Equation (2) is relatively independent in each vertical layer, which can better reflect the different variation characteristics of the actual temperature profile.

2.6 Reconstruction of salinity profiles by temperature

The correlation between the temperature and salinity can be expressed using Equation (3) by regression analysis:

$$S_{i,k}(T) = \overline{S_{i,k}} + a_{i,k}^{S1}(T - \overline{T_{i,k}}) \quad (3)$$

where $\overline{S_{i,k}}$ is the average salinity, $\overline{T_{i,k}}$ is the average temperature and $S_{i,k}(T)$ are the values of reconstructed salinity for grid point i and depth k , and $a_{i,k}^{S1}$ are the regression coefficients. Equation (3) was used to reconstruct the daily salinity profile data.

3. Results and discussion

3.1 Construction of the static climatic T–S field

Based on the WOD18 data, the historical temperature profile observations in the northern South China Sea were gridded and pre-treated using Equation (4) to form a static climate temperature field product with a horizontal resolution of 0.25° and a time scale of 1 month. It should be noted that the spatial resolution of 0.25° was seleted to be consistent with the satellite data, so as to complete the reconstruction.

$$T_{i,k}^c = T_{i,k}^{WOA18} + \sum_{j=1}^{j=N} w_{i,j}\left(T_{j,k}^0 - T_{j,k}^{WOA18}\right) \quad (4)$$

where $T_{j,k}^{WOA18}$ are the WOA18 temperature interpolated to the required position and depth, $T_{j,k}^0$ are the observation data, $T_{i,k}^c$ are the results at the new grid i and depth k . The $w_{i,j}$, N weights are calculated using Equation (5),

$$C_i W_i = F_i \quad (5)$$

where $w_{i,j}$ are weight coefficients, C_i is the covariance matrix and F_i is the matrix of initial errors between the grid point and the observation point. The calculation of Equations (5) is explained very clearly by Fox et al. (2002), which determined the $w_{i,j}$ and then completed calculation of $T_{i,k}^c$.

The monthly vertical distribution of the northern South China Sea temperature profiles using these static climate field products is shown in Figure 2. Compared with the results from WOA18, the obtained temperature profiles are slightly different, especially at depths below 600 meters, where temperatures are about 0.3°C higher than the WOA18 temperatures. However, for depths above 600 meters, the temperature of the static climatic field is similar to WOA18. Figure 3 shows the vertical distribution of the temperature profile standard deviations. The values increase from the sea surface with increasing depth and reached a maximum of about 1.8 to 2.2°C near the thermocline in different months. Then temperature standard deviations gradually decrease with increasing depth and approach the minimum of about 0.2°C when the water depth exceeds 1000 meters.

Static climatic salinity fields were generated at each grid point based on Equation (3) for different water depths in the

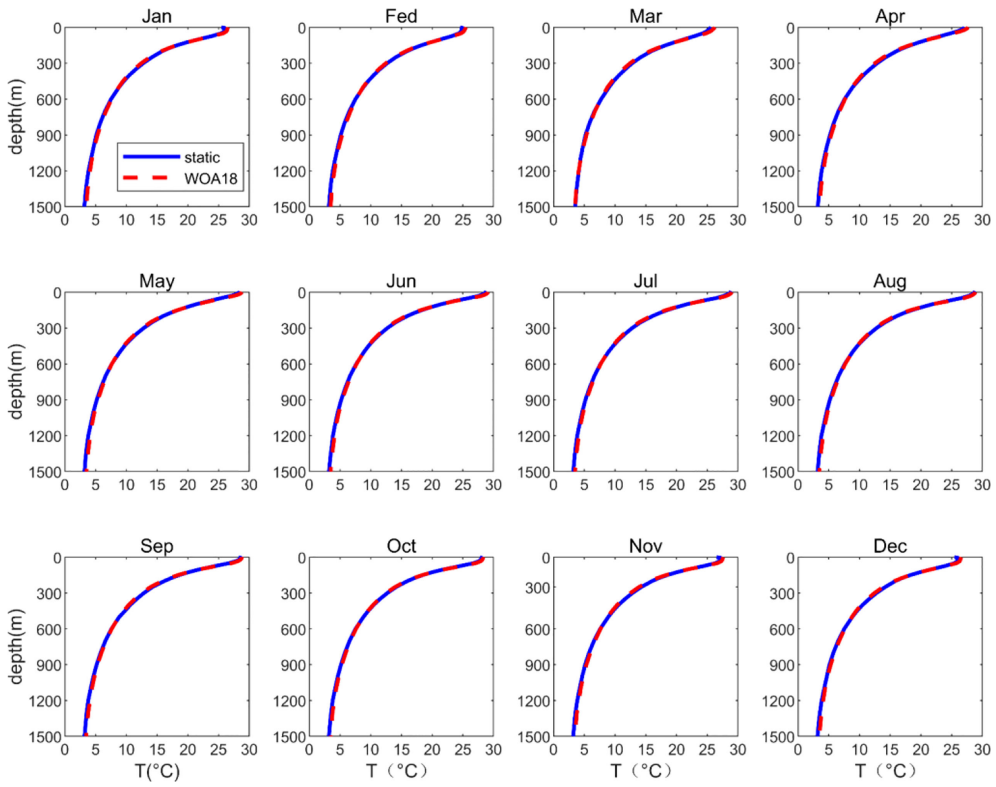


FIGURE 2
Monthly temperature profiles in the northern South China Sea. The blue and red lines represent the static climatic field and the WOA18 data, respectively.

northern South China Sea. Similar to the temperature profiles, Figure 4 shows the monthly vertical distribution of the salinity profiles. The results were about 0.1 psu higher than the WOA18 data above 1000-meter depths and are highest near 400 meters. The two datasets are fairly consistent when the water depth exceeds 1000 meters. Figure 5 shows the vertical distribution of the salinity profile monthly standard deviations. The maximum standard deviation is about 0.3 psu at the surface. The salinity standard deviation drops sharply from the sea surface to about 400 meters depth and then maintains a low value of about 0.02 psu to the deepest data points.

3.2 Reconstruction of an expanded three-dimensional T–S field

3.2.1 Expanded three-dimensional T–S field and statistical tests

Based on Equation (2), the temperature static climate field in Section 3.1 was expanded to a three-dimensional temperature field with a spatial resolution of 0.25° and a daily time scale using the daily SST and SSH datasets. The three-dimensional salinity field with the same time and space resolution was also formed using Equation (3). Finally, the T–S observation data from GTSPP and Argo datasets in the northern South China Sea for a recent 20-year period were collated as a historical observation dataset for statistical testing.

The daily expanded three-dimensional T–S field data were interpolated to each observation position to obtain the corresponding expanded profiles, and the root–mean–square errors between the expanded results and the observation profiles were calculated month by month. Figures 6, 7 show the vertical distribution of the monthly errors of the expanded temperature and salinity profiles, respectively. The results show that the errors in monthly temperature are larger in the upper ocean and are over 1.0°C near the thermocline (100–200 meters), decreasing sharply with increasing depth below the thermocline. The errors reached their largest values (1.55°C) at a depth of 75 meters in July. Furthermore, the errors in monthly salinity are larger near the sea surface, reaching about 0.70 psu from June through October. With increasing depth, the values gradually decrease and then remain small below 200 meters. The T-S errors are close to the existing products called CORA (China ocean reanalysis), which was devoloped by similar method and its average errors of temperature (salinity) are 1.14°C (0.18 psu) (Han et al., 2011). In the future work, the integration of other

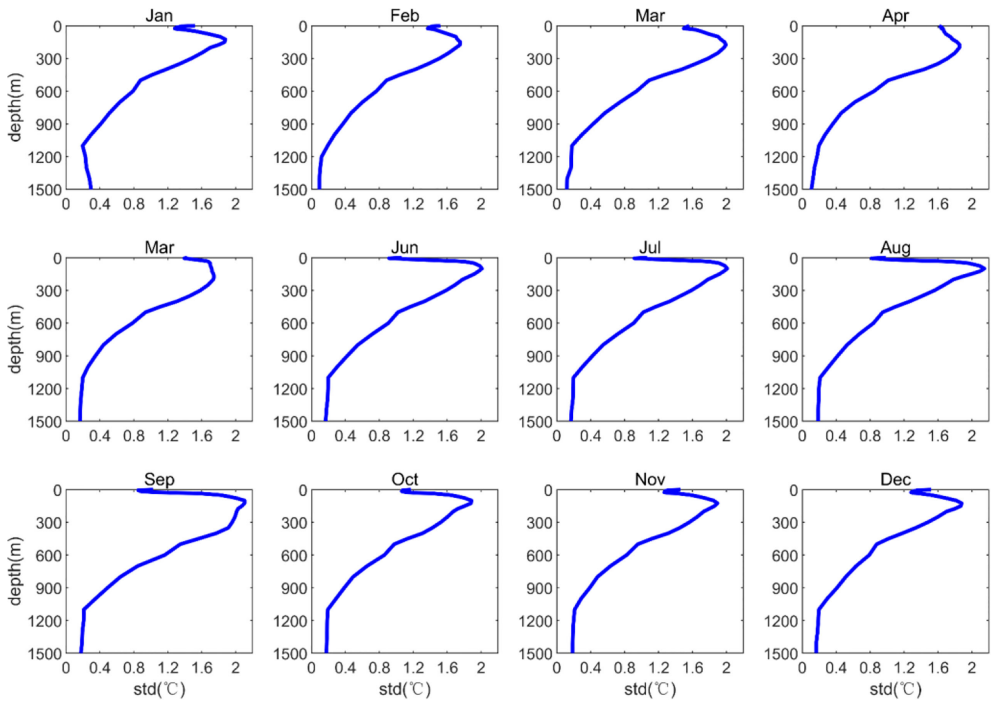


FIGURE 3
Vertical distribution of monthly static climatic temperature field standard deviations in the northern South China Sea.

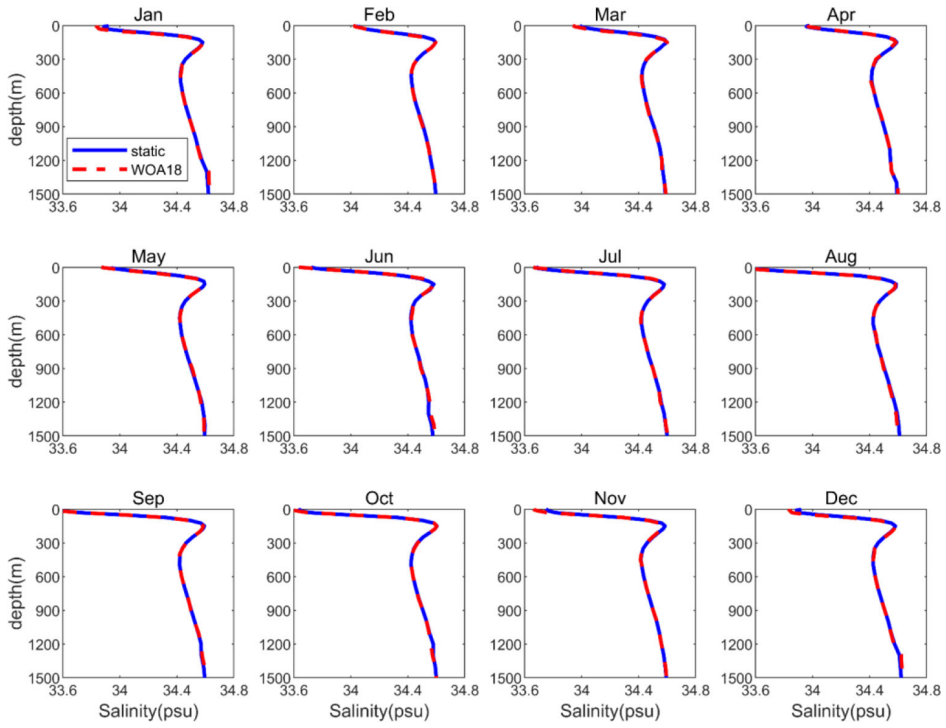


FIGURE 4
Monthly salinity profiles in the northern South China Sea. The blue and red lines represent the static climatic field and WOA18 data, respectively.

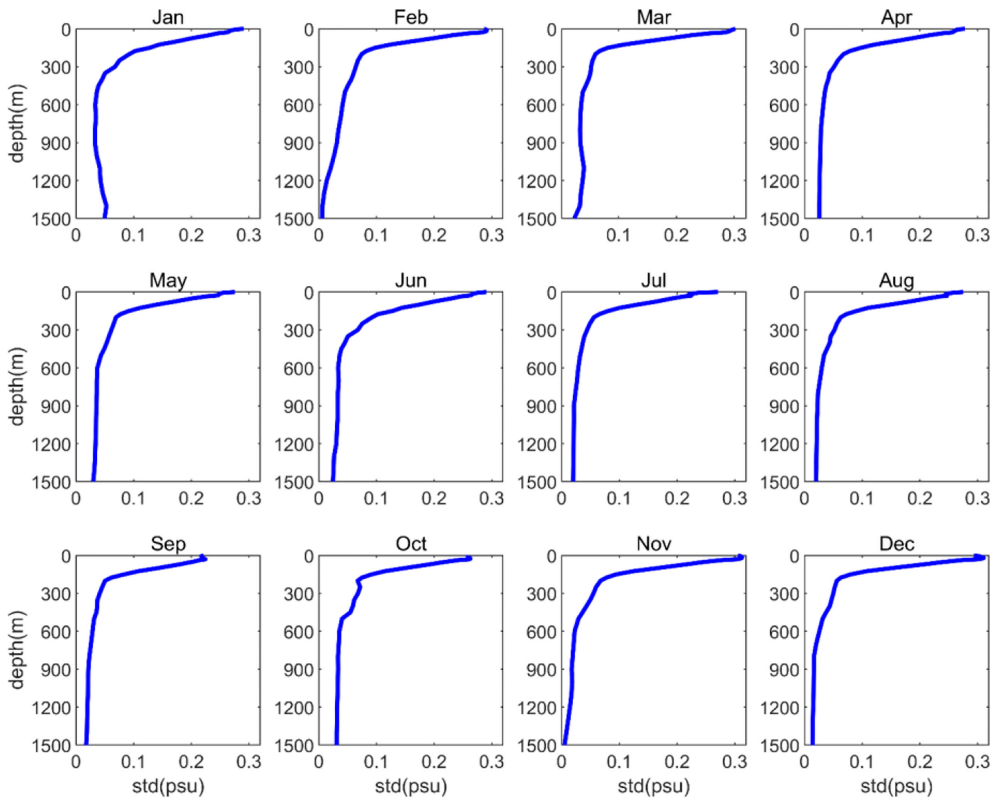


FIGURE 5
Vertical distribution of monthly static climatic salinity field standard deviations in the northern South China Sea.

data products may be an effective method to reduce the salinity profile errors at the sea surface (Yang et al., 2015; Bao et al., 2019).

3.2.2 Test of single stations

Stations 20110303 and 20110822 in Figure 8 were selected for further study. The expanded T–S profiles are compared with the observed data in Figure 9. The results show that the expanded profiles are very close to the observed profiles, proving that the expanded field can be effectively applied in this region. Furthermore, the spatial distribution of maximum temperature errors along the vertical showed in Figure 8, which revealed that the maximum errors increase near the coast of South China Sea.

3.2.3 Cross-validation tests

The cross-validation tests were used for this work to strengthen the validation of the reconstruction. The WOD18 database was randomly divided into two parts in a ratio of 3 to 1, three quarters of them were used for the reconstruction method, and anothor part was kept for validation of the method.

Figures 10, 11 show the vertical distribution of the monthly errors of the expanded temperature and salinity profiles, respectively. The distribution of the errors results is similar to Figures 6, 7. The difference is that the errors obtained by cross-validation tests are generally small. For example, the temperature errors are generally less than 1.2°C (0.8°C) in summer (winter). The salinity errors are less than 0.2 psu overall, except in summer when it can reach 0.6 psu at the sea surface. The smaller errors may be due to the different databases used for validation.

Stations 20160808 and 20130216 in Figure 8 were selected for further study. The expanded T–S profiles are compared with the observed data in Figure 12. The results show that the expanded profiles are close to the observed profiles, proving that the expanded field can be verified by cross-validation tests.

3.2.4 Comparison analysis of a temperature cross-section

An observation section between Taiwan Island and Luzon Island was selected to test the expansion results. The position of the observation section is shown in Figure 8. The starting position of the observation section is (118.0°E, 21.5°N), and the end position is (124.0°E, 16.5°N). The observation date is 19 March 2000. In the observed temperature section (Figure 13A), there is a cold eddy causing the isotherm to bulge upward near

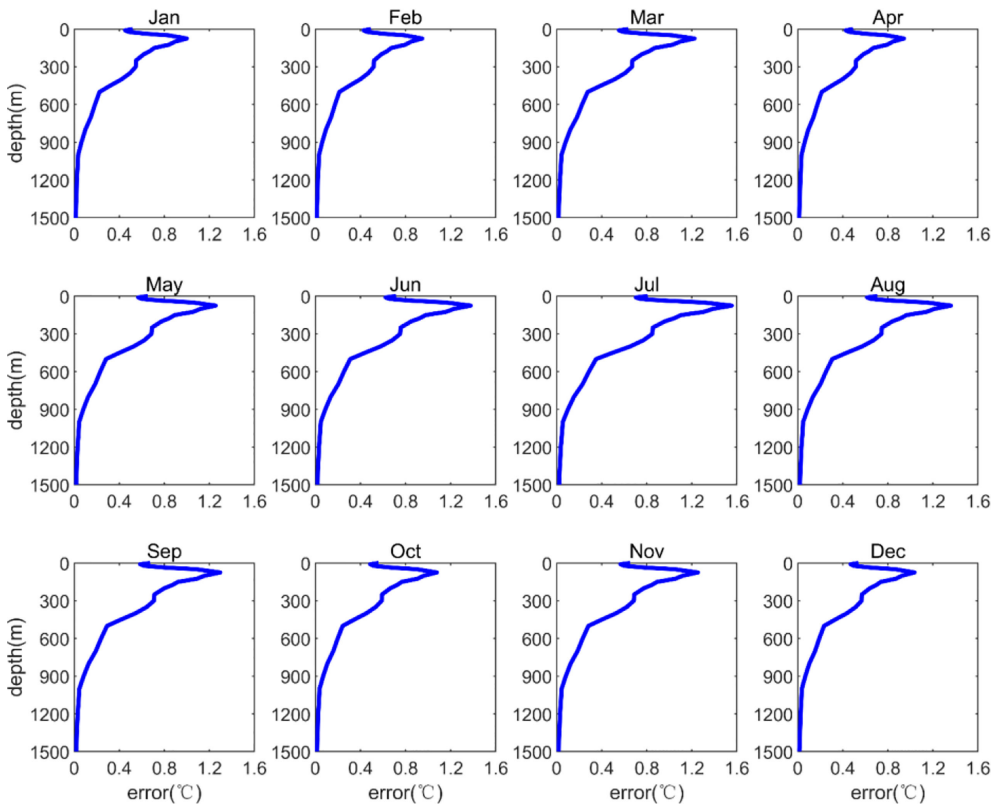


FIGURE 6
Vertical distribution of expanded temperature field monthly errors in the northern South China Sea.

122.0°E and a sinking isotherm caused by a warm eddy near 120.8°E and 123.3°E. Compared to the WOA18 climatic temperature section in March (Figure 13C), the expanded temperature section (Figure 13B) better reflects the internal variation characteristics of the mesoscale eddy. The two warm eddies in the observation section are well described, and the cold eddy is also shown. In addition, the structure of the upper mixed layer is further resolved.

3.2.5 Structure and variation of the temperature profiles affected by the luzon warm eddy

The Luzon Warm Eddy (LWE) was first recognised as an anticyclonic ring centred at about 117.5°E, 21°N in summer (Li and Pohlmann, 2002). Yuan et al. (2007) used altimeter data to identify the ring as an anticyclonic eddy generated off the northwestern coast of Luzon Island and gave the LWE its name. The LWE is a seasonal phenomenon closely related to sea level anomalies (SLA). Figure 14 describes the variation of SLA during the LWE, and a black square was seleted as the target area. The results show that an anticyclonic eddy gradually formed a ring from 15 September 2006 through 5 November 2006, with high SLA in the target area. It then moved westward and finally disappeared on January 15, 2007 and the SLA in the target area decreased accordingly.

Figure 15 shows the time series of temperature profiles corresponding to the target area in Figure 14 based on the observation data and expanded temperature field, respectively. Similar to the observation data, the results for the expanded temperature field clearly show the variability of these temperature profiles. Because of the significant influence of the LWE, the upper-ocean temperatures were up to 28°C, and the thermocline depths were about 150 meters before November 2006. After the LWE had disappeared, the thermocline uplifted gradually, reaching 100 meters, and the upper-ocean temperatures were reduced to 22°C. In addition, the subtle time-scale changes were only reflected by observation data.

4 Conclusions

The mapping relationship between sea surface and subsurface temperature information was established by the linear regression analysis method using the historical Argo profile observations (WOD18). Daily three-dimensional temperature fields with a spatial resolution of 0.25° in the

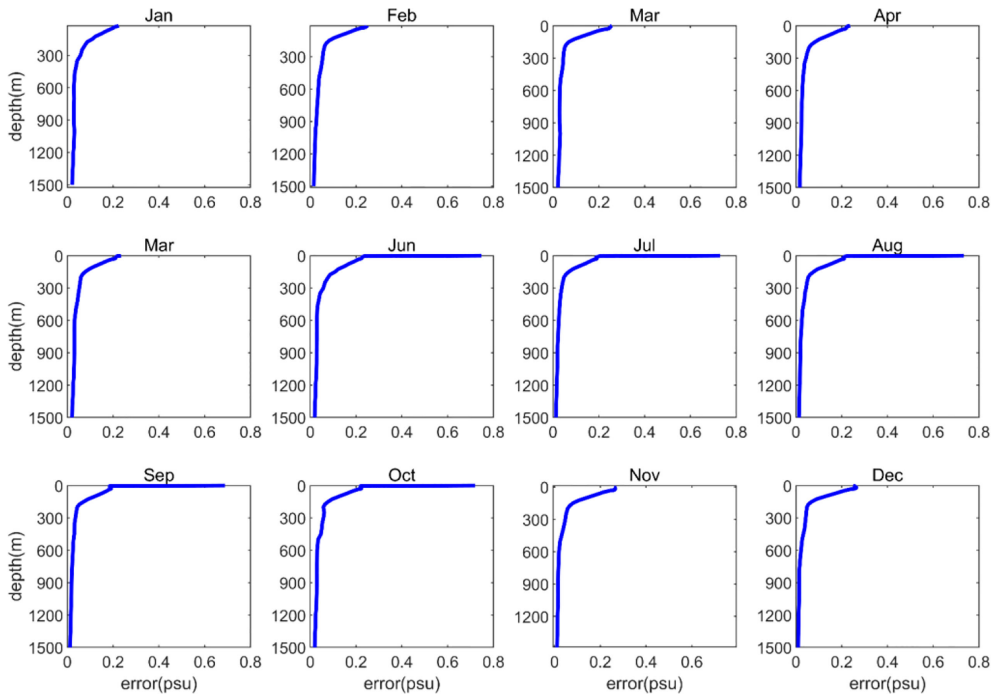


FIGURE 7
Vertical distribution of expanded salinity field monthly errors in the northern South China Sea.

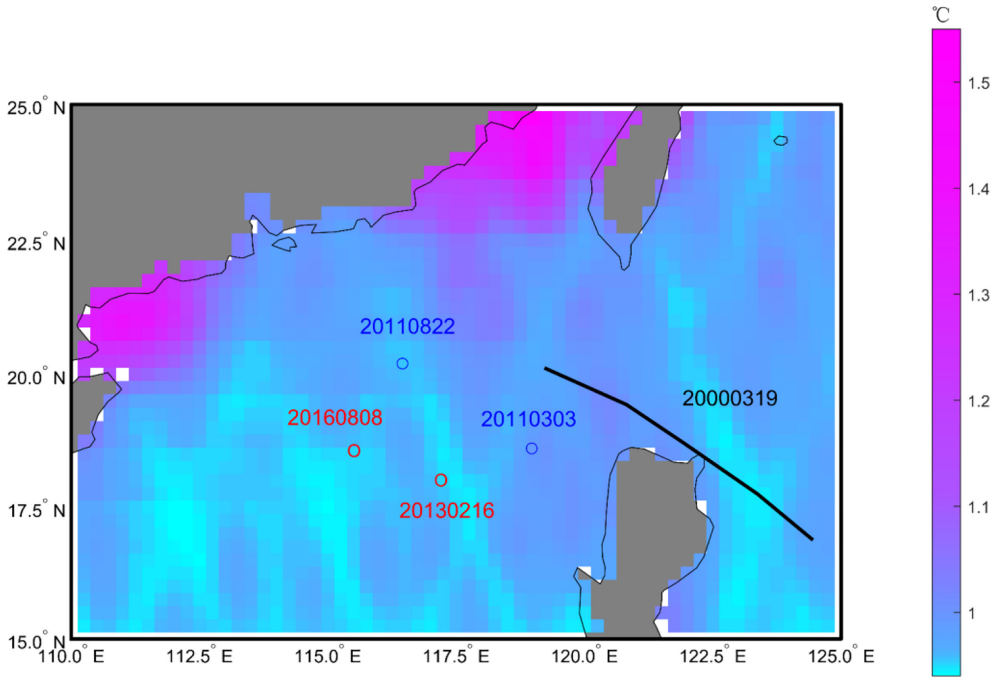


FIGURE 8
Two observation stations and a temperature cross-section. The numbers are dates and the background color is the spatial distribution of maximum temperature errors along the vertical.

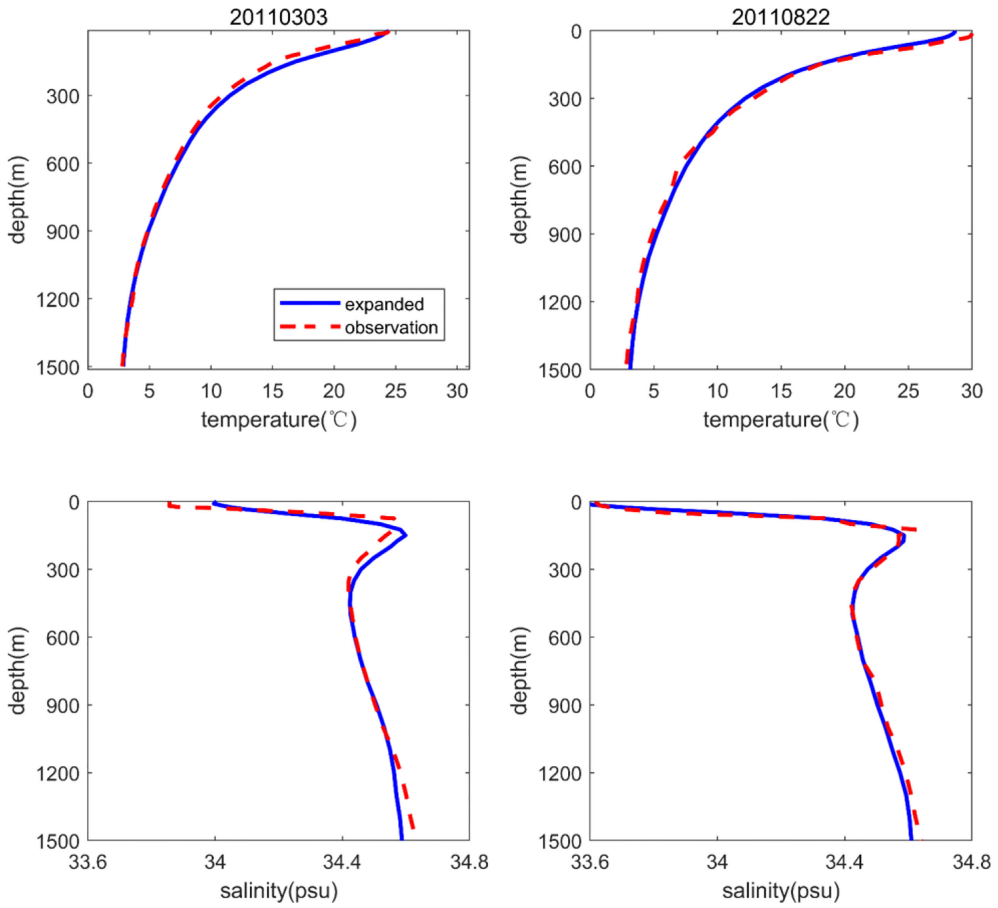


FIGURE 9
T–S profiles for two observation stations in the northern South China Sea. The blue and red lines represent the expanded T–S field results and observations (WOD18), respectively.

northern South China Sea were reconstructed based on the SST and SSH datasets derived from ARSS and CMEMS, respectively.

First, we selected valid data from WOD18 to construct a static climatic temperature field in the northern South China Sea and described their monthly vertical distributions. The results showed that these constructed temperature profiles were similar to the results from WOA18 and the standard deviations between them were relatively small, reaching a maximum of about 1.8°C near the thermocline and decreasing in other layers. The static climatic salinity fields were established next and were also close to the results from WOA18 with maximum standard deviations of about 0.3 psu at the surface.

An expanded three-dimensional T–S field was also reconstructed in the northern South China Sea. The T–S observation data from GTSPP and Argo datasets were selected as a statistical test field. The results showed that the monthly errors in the expanded temperature (salinity) profiles remained small and the maximum was less than 1.6°C (0.8 psu) near the thermocline (surface). Then, the errors of T-S profiles obtained by cross-validation were also smaller, and the maximun was less than 1.2°C (0.6 psu) near the thermocline (surface). Because of these small monthly errors, the expanded T–S fields are considered to provide a realistic representation of the northern South China Sea, and the expanded T–S profiles at two single stations were also confirmed to be similar to the observed results. Furthermore, the expanded T–S field clearly described the vertical temperature structure at a cross-section through the Luzon Strait and accurately simulated the structure and variation of the temperature profile associated with the LWE. These results prove that the ex-panded T–S field can accurately reflect the actual ocean field's vertical structure and internal variation and the mesoscale eddies inside the ocean.

Obviously, the reconstructed salinity profiles have relatively large errors on the sea surface, especially in summer. Although this does not affect the overall results of this work, we hope to reduce the errors in future work. The use of satellite SSS (Sea Surface Salinity) database may be an effective method, which can provide large-scale and continuous data (Yang et al., 2015; Bao et al., 2019).

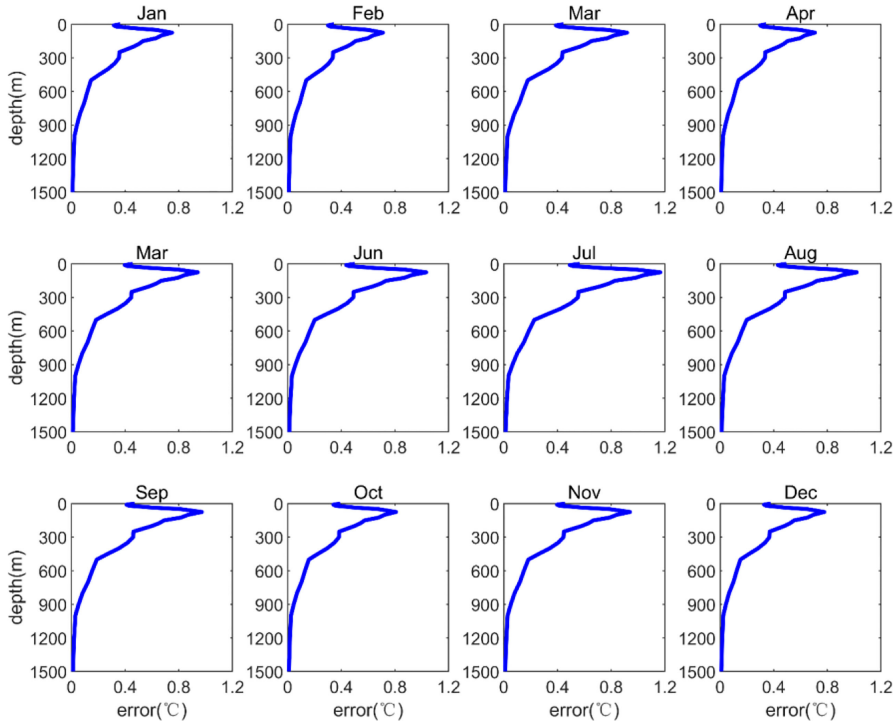


FIGURE 10
Vertical distribution of expanded temperature field monthly errors in the northern South China Sea (cross-validation tests).

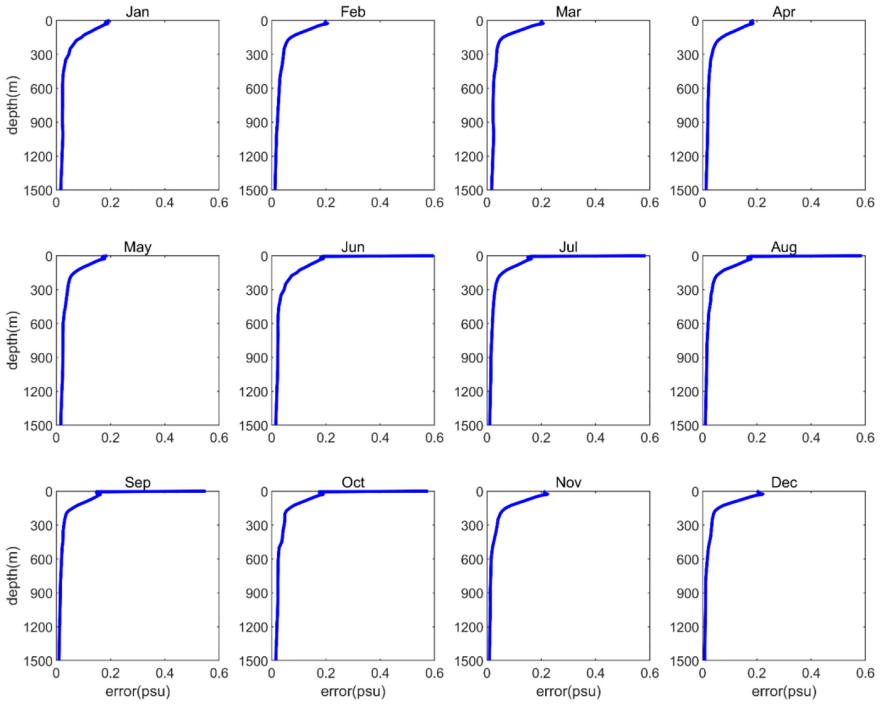


FIGURE 11
Vertical distribution of expanded salinity field monthly errors in the northern South China Sea (cross-validation tests).

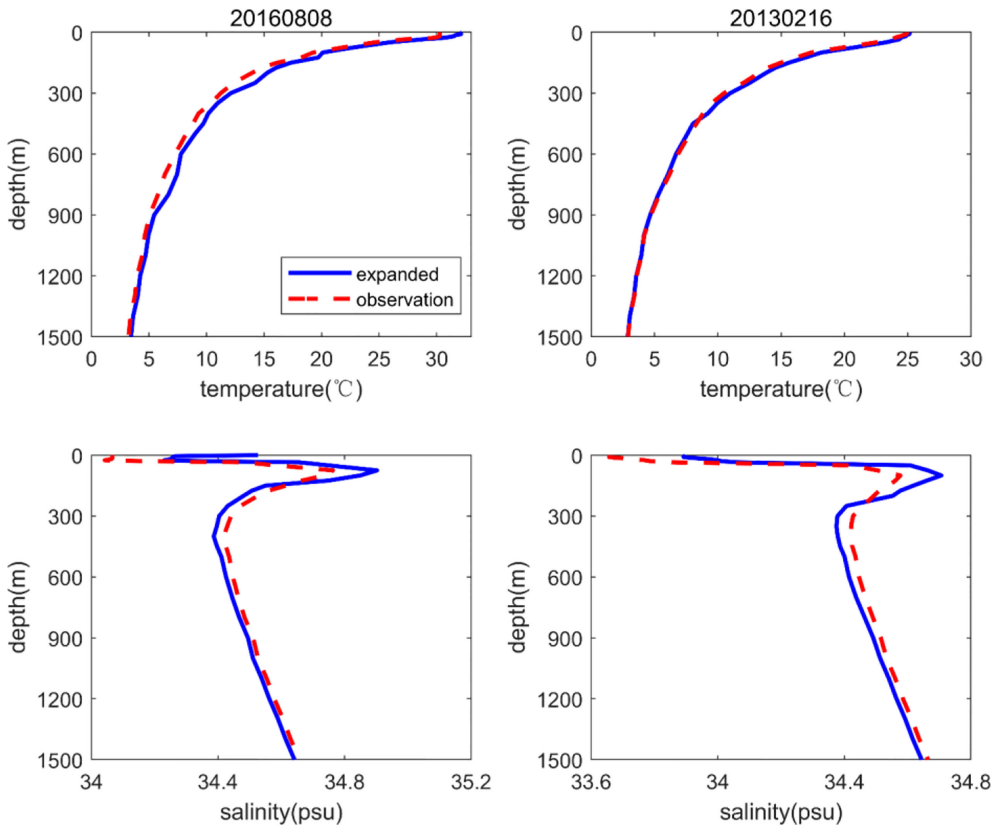


FIGURE 12
T–S profiles for two observation stations in the northern South China Sea (cross-validation tests). The blue and red lines represent the expanded T–S field results and observations (WOD18), respectively.

Recently available SSS data mainly come from SMOS (Soil Moisture and Ocean Salinity), Aquarius and SMAP (Soil Moisture Active Passive) satellites, of which Aquarius stopped working on June 7, 2015 due to power supply problems (Tang et al., 2017). However, the SSS database is also required strict quality assessment and error analysis, and the data mainly focus on the last 10 years, so the use of SSS to complete the reconstructed salinity profile needs further research.

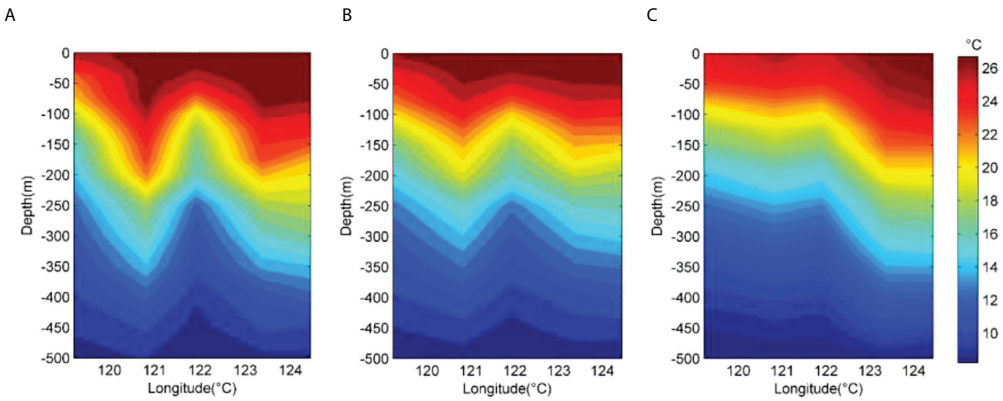


FIGURE 13
Temperature profiles of the cross-section (Figure 8) based on the observations **(A)**, expanded temperature field **(B)** and WOA18 **(C)**, respectively.

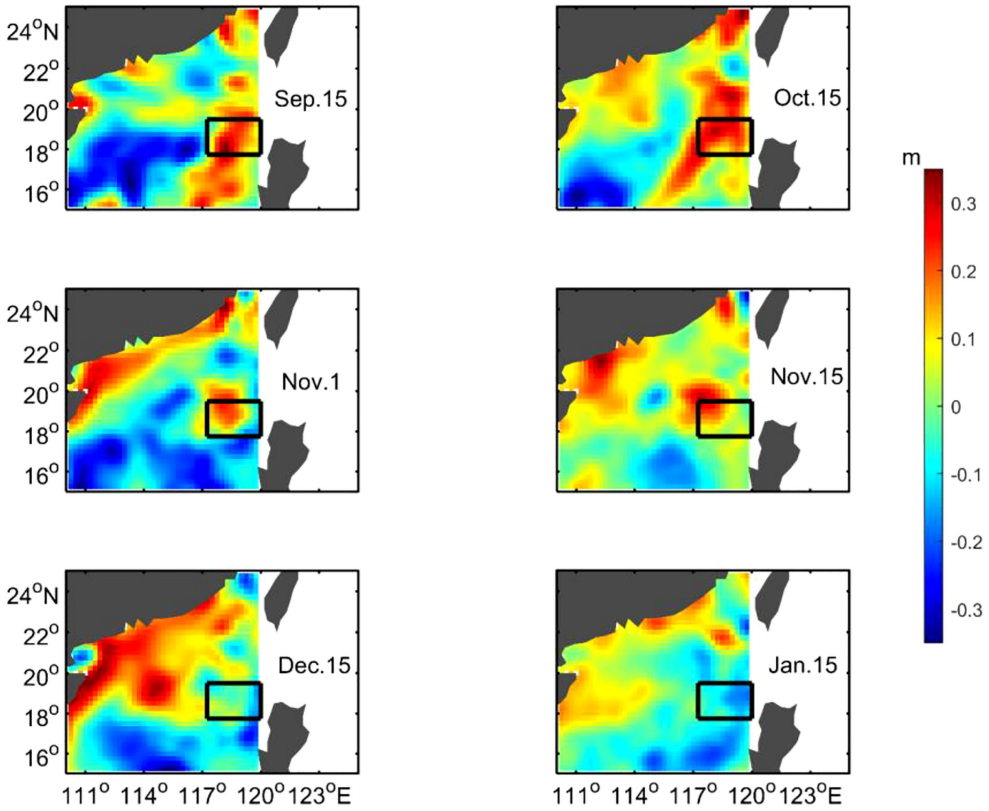


FIGURE 14
SLA (colours) from 15 September 2006 through 15 January 2007 in the western Luzon Strait. Black squares denote the locations of observed profiles from WOD18.

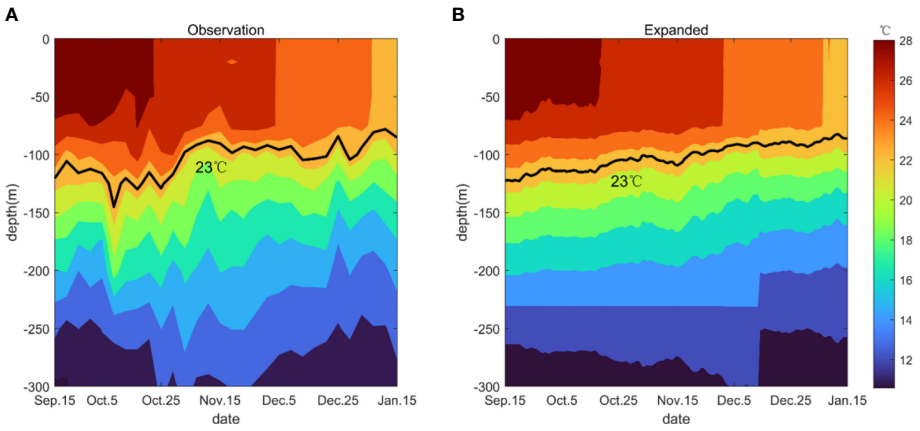


FIGURE 15
Temperature profiles corresponding to the black squares in Figure 14 based on the observation data **(A)** and the expanded temperature field **(B)**, respectively. The black lines are isotherms of 23°C used as a thermocline proxy and the interval bewewen color contours is 2°C.

The T–S information inside the ocean is a basic component of oceanographic research data, playing a crucial role in describing the nature of the ocean. For example, Han et al. (2011) developed a regional ocean reanalysis system for the coastal waters of China and adjacent seas by similar method, and the evaluations show that a good representation of the processes and phenomena were produced. Wang et al. (2012) reconstructed the T-S profiles from 1993 to 2008 near the Luson Strait, which was used to estimate the heat, salt and volume transports during mesoscale eddies movement. This study's expanded T–S field is expected to be used as the initial field for an ocean numerical model or for pseudo-observation assimilation into an ocean numerical reanalysis and prediction system to improve the three-dimensional T–S field results, which could delineate ocean processes and phenomena such as mesoscale eddies more clearly.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: NODC: https://www.nodc.noaa.gov/OC5/indprod.html; https://www.nodc.noaa.gov/gtspp/CARDC: ftp://data.argo.org.cn/pub/ARGO/raw_argo_data/; ftp://ftp.argo.org.cn/pub/ARGO/china/ARSS: http://www.remss.com/measurements/sea-surface-temperature CMEMS: http://marine.copernicus.eu.

Author contributions

Conceptualization, BT and DZ. Methodology, BT. Software, BT. Formal analysis, BT. Investigation, BT and DZ. Resources, BT. Data curation, DZ. Writing—original draft preparation, BT. Writing—review and editing, BT, DZ, CC and XZ. Visualization, BT, XZ and CC. All authors have read and agreed to the published version of the manuscript.

Funding

This research was supported by Research start-up fund of Qilu University of Technology, 81110727 and Shandong Provincial Natural Science Foundation, China, ZR202102240074.

Acknowledgments

The authors thank all agencies that provided the data. The authors are grateful for all the constructive comments from anonymous reviewers. The authors also thank the Editor for the kind assistances and beneficial comments.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Bao, S. L., Zhang, R., Wang, H. Z., Yan, H. Q., Yu, Y., and Chen, J. (2019). Salinity profile estimation in the pacific ocean from satellite surface salinity observation. *J. Atmospheric Oceanic Technology.* 36 (1), 53–68. doi: 10.1175/JTECH-D-17-0226.1

Boyer, T. P., Baranova, O. K., Coleman, C., Garcia, H. E., Grodsky, A., Locarnini, R. A., et al. (2018). *World ocean database 2018*. A. V. Mishonov (Technical Editor, NOAA Atlas NESDIS 87). Available at: https://www.ncei.noaa.gov/sites/default/files/2020-04/wod_intro_0.pdf.

Bruno, B. N., and Santoleri, R. (2004). Reconstructing synthetic profiles from surface data. *J. Atmospheric Oceanic Technology.* 21, 693–703. doi: 10.1175/1520-0426(2004)021<0693:RSPFSD>2.0.CO,2

Carnes, M. R., Mitchell, J. L., and Dewitt, P. W. (1990). Synthetic temperature profiles derived from geosat altimetry: Comparison with air-dropped expendable bathythermograph profiles. *J. Geophysical Res. Atmospheres* 95 (C10), 17979–17992. doi: 10.1029/JC095iC10p17979

Carnes, M. R., Teague, W. J., and Mitchell, J. L. (1994). Inference of subsurface thermohlaline structure from fields measurable by satellite. *J. Atmospheric Oceanic Technology.* 11, 551–566. doi: 10.1175/1520-0426(1994)011<0551:IOSTSF>2.0.CO;2

Chaigneau, A., Eldin, G., and Dewitte, B. (2009). Eddy activity in the four major upwelling systems from satellite altimetry(1992-2007). *Prog. Oceanography.* 83 (1), 117–123. doi: 10.1016/j.pocean.2009.07.012

Chaigneau, A., Texier, M. L., Eldin, G., Grados, C., and Pizarro, O. (2011). Vertical structure of mesoscale eddies in the eastern south pacific ocean: A composite analysis from altimetry and argo profiling floats. *J. Geophysical Res.* 116, C11025. doi: 10.1029/2011JC007134

Chelton, D. B., Schlax, M. G., and Samelson, R. M. (2011). Global observations of nonlinear mesoscale eddies. *Prog. Oceanography.* 91 (2), 167–216. doi: 10.1016/j.pocean.2011.01.002

Chen, G., Gan, J., Xie, Q., Chu, X., Wang, D., and Hou, Y. (2012). Eddy heat and salt transports in the south China Sea and their seasonal modulation. *J. Geophysical Research-oceans* 117, C05021. doi: 10.1029/2011JC007724

Chen, G., Hou, Y., and Chu, X. (2011). Mesoscale eddies in the south China Sea: Mean properties, spatiotemporal variability, and impact on thermohaline structure. *J. Geophysical Research-oceans* 116, C06018. doi: 10.1029/2010JC006716

Fox, D. N., Teague, W. J., Berron, C. N., Carnes, M. R., and Lee, C. M. (2002). The modular ocean data assimilation system. *J. Atmospheric Oceanic Technology.* 19, 240–252. doi: 10.1175/1520-0426(2002)019%3C0240:TMODAS%3E2.0.CO;2

Garcia, H. E., Boyer, T. P., Baranova, O. K., Locarnini, R. A., Mishonov, A. V., Grodsky, A., et al. (2019) *World ocean atlas 2018: Product documentation*. Available at: https://data.nodc.noaa.gov/woa/WOA18/DOC/woa18documentation.pdf.

Gavart, M., and Mey, P. (1997). Isopycnal EOFs in the Azores current region: A statistical tool for dynamical analysis and data assimilation. *J. Phys. Oceanography.* 27, 2146–2157. doi: 10.1175/1520-0485(0)027<2146:IEITAC>2.0.CO;2

Gentemann, C. L., Meissner, T., and Wentz, F. J. (2010). Accuracy of satellite sea surface temperatures at 7 and 11 GHz. *IEEE Trans. Geosci. Remote Sensing.* 48, 1009–1018. doi: 10.1109/TGRS.2009.2030322

Guinehut, S., Traon, P. Y. L., Larnicol, G., and Philipps, S. (2004). Combining argo and remote-sensing data to estimate the ocean three dimensional temperature fields-a first approach based on simulated observations. *J. Mar. Systems.* 46 (1-4), 85–98. doi: 10.1016/j.jmarsys.2003.11.022

Han, G. J., Li, W., Zhang, X. F., Li, D., He, Z. J., Wang, X. D., et al. (2011). A regional ocean reanalysis syestem for China coastal waters and adjacent seas. *Adv. Atmospheric Sci.* 28 (3), 682–690. doi: 10.1007/s00376-010-9184-2

Hu, J., Gan, J., Sun, Z., Zhu, J., and Dai, M. (2011). Observed three-dimensional structure of a cold eddy in the southwestern south China Sea. *J. Geophysical Research-Oceans* 116, C05016. doi: 10.1029/2010JC006810

Hu, P., and Hou, Y. (2010). Path transition of the Western boundary current with a gap due to mesoscale eddies: A 1.5-layer, wind-driven experiment. *Chin. J. Oceanology Limnology.* 28, 364–370. doi: 10.1007/s00343-010-9293-x

Hurlburt, H. E. (1986). Dynamic transfer of simulated altimeter data into subsurface information by a numerical ocean model. *J. Geophysical Research-Atmospheres* 91 (C2), 2372–2400. doi: 10.1029/JC091iC02p02372

Hurlburt, H. E., Fox, D. N., and Metzger, J. (1990). Statistical inference of weekly correlated subthermocline fields from satellite altimeter data. *J. Geophysical Research-Atmospheres* 95 (C7), 11375–11409. doi: 10.1029/JC095iC07p11375

Hu, J., Zheng, Q., Sun, Z., and Tai, C. K. (2012). Penetration of nonlinear rossby eddies into south China Sea evidenced by cruise data. *J. Geophysical Research-Oceans* 117, C03010. doi: 10.1029/2011JC007525

Li, L., and Pohlmann, T. (2002). The south China Sea warm-core ring 94S and its influence on the distribution of chemical tracers. *Ocean Dynamics.* 52 (3), 116–122. doi: 10.1007/s10236-001-0009-9

Li, Q. Y., Sun, L., and Lin, S. F. (2016). GEM: A dynamic tracking model for mesoscale eddies in the ocean. *Ocean Science.* 12 (6), 1249–1267. doi: 10.5194/os-12-1249-2016

Li, Q. Y., Sun, L., Liu, S. S., Xian, T., and Yan, Y. F. (2014). A new mononuclear eddy identification method with simple splitting strategies. *Remote Sens. Letters.* 5 (1), 65–72. doi: 10.1080/2150704X.2013.872814

Nan, F., He, Z., Zhou, H., and Wang, D. (2011). Three long-lived anticyclonic eddies in the northern south China Sea. *J. Geophysical Research-Oceans* 116, C5. doi: 10.1029/2010JC006790

Pascual, A., and Gomis, D. (2003). Use of surface data to estimate geostrophic transport. *J. Atmospheric Oceanic Technology.* 20, 912–926. doi: 10.1175/1520-0426(2003)020%3C0912:UOSDTE%3E2.0.CO;2

Pujol, M. I., Faugère, Y., Taburet, G., Dupuy, S., Pelloquin, C., Ablain, M., et al. (2016). DUACS DT2014: The new multi-mission altimeter data set reprocessed over 20 years. *Ocean Science.* 12, 1067–1090. doi: 10.5194/os-12-1067-2016

Su, J. (2004). Overview of the south China Sea cuirculation and its influene on the coatal physical oceanography outside the pearl river estuary. *Continental Shelf Res.* 24, 1745–1760. doi: 10.1016/j.csr.2004.06.005

Taburet, G., Roman, A. S., Ballarotta, M., Pujol, M. I., and Dibarboure, G. (2019). DUACS DT2018: 25 years of reprocessed sea level altimetry products. *Ocean Science.* 15, 1207–1224. doi: 10.5194/os-15-1207-2019

Tang, W. Q., Fore, A., Yush, S., Lee, T., Hayashi, A., Sanchez, F. A., et al. (2017). Validating SWAP SSS with *in situ* measurements. *Remote Sens. Environment.* 200, 326–340. doi: 10.1109/IGARSS.2017.8127518

Wang, X. D., Li, W., Qi, Y. Q., and Han, G. J. (2012). Heat, salt and volume transports by eddies in the vicinity of the Luzon strait. *Deep Sea Research Part I* 61, 21–33. doi: 10.1016/j.dsr.2011.11.006

Wang, G. H., Su, J. L., and Chu, P. C. (2003). Mesoscale eddies in the south China Sea observed with altimeter data. *Geophysical Res. Letters.* 30 (21), 2121. doi: 10.1029/2003GL018532

Wu, C. C., Lee, C. Y., and Lin, I. I. (2007). The effect of the ocean eddy on tropical cyclone intensity. *J. Atmospheric Sci.* 64, 3562–3578. doi: 10.1175/JAS4051.1

Yang, T. T., Cheng, Z. B., and He, Y. J. (2015). A new method to retrieve salinity profiles from sea surface salinity observed by SMOS satellite. *Acta Oceanologica Sinica.* 34, 85–93. doi: 10.1007/s13131-015-0735-3

Yuan, D. L., Han, W. Q., and Hu, D. X. (2007). Anti-cyclonic eddies northwest of Luzon in summer-fall observed by satellite altimeters. *Geophysical Res. Lett.* 34, L13610. doi: 10.1029/2007GL029401

frontiers | Frontiers in Marine Science

TYPE Original Research
PUBLISHED 03 October 2022
DOI 10.3389/fmars.2022.983007

OPEN ACCESS

EDITED BY
Junyu He,
Zhejiang University, China

REVIEWED BY
Pushpa Dissanayake,
University of Kiel, Germany
Didit Adytia,
Telkom University, Indonesia
Gengkun Wu,
Shandong University of Science and Technology, Qingdao, China

*CORRESPONDENCE
Shiqiu Peng
speng@scsio.ac.cn
Fan Meng
vanmeng@163.com

SPECIALTY SECTION
This article was submitted to Ocean Observation, a section of the journal Frontiers in Marine Science

RECEIVED 30 June 2022
ACCEPTED 12 September 2022
PUBLISHED 03 October 2022

CITATION
Song T, Han R, Meng F, Wang J, Wei W and Peng S (2022) A significant wave height prediction method based on deep learning combining the correlation between wind and wind waves.
Front. Mar. Sci. 9:983007.
doi: 10.3389/fmars.2022.983007

A significant wave height prediction method based on deep learning combining the correlation between wind and wind waves

Tao Song[1,2], Runsheng Han[1], Fan Meng[1,3,4]*, Jiarong Wang[1], Wei Wei[1] and Shiqiu Peng[5]*

[1]College of Computer Science and Technology, China University of Petroleum, Qingdao, China, [2]Department of Artificial Intelligence, Faculty of Computer Science, Polytechnical University of Madrid, Madrid, Spain, [3]DAMO Academy, Alibaba Group, Hangzhou, China, [4]Key Laboratory of Environmental Change and Natural Disaster of Ministry of Education, Beijing Normal University, Beijing, China, [5]State Key Laboratory of Tropical Oceanography, South China Sea Institute of Oceanology, Chinese Academy of Sciences, Guangzhou, China

Accurate wave height prediction is significant in ports, energy, fisheries, and other offshore operations. In this study, a regional significant wave height prediction model with a high spatial and temporal resolution is proposed based on the ConvLSTM algorithm. The model learns the intrinsic correlations of the data generated by the numerical model, making it possible to combine the correlations between wind and wind waves to improve the predictions. In addition, this study also optimizes the long-term prediction ability of the model through the proposed Mask method and Replace mechanism. The experimental results show that the introduction of the wind field can significantly improve the significant wave height prediction results. The research on the prediction effect of the entire study area and two separate stations shows that the prediction performance of the proposed model is better than the existing methods. The model makes full use of the physical correlation between wind and wind waves, and the validity is up to 24 hours. The 24-hour forecast R^2 reached 0.69.

KEYWORDS

wave height forecast, deep learning, high spatial and temporal resolution, new mechanism, long time prediction

1 Introduction

Wind waves are waves generated by and influenced by the local wind (Barnett and Kenyon, 1975). It is characterized by often sharp wave crests, very irregular distribution on the sea surface, short crest lines, and minor periods. When the wind is strong, the phenomenon of breaking waves often occurs, and water splashes are formed. In general, wind disturbance of the sea surface causes capillary waves (ripples) so that the wind further provides the necessary roughness for delivering energy to the sea surface. Then, the waves continue to be fueled by the pressure of the wind on its surface (Longuet-Higgins, 1963; Kirby, 1985), causing the wind waves to grow (Phillips, 1957). Wind waves dominate the motion of the sea for a short period. Therefore the study of wind waves has implications for many applications such as navigation safety and coastal engineering. Waves also define air-sea fluxes and interact strongly with surface currents, upper ocean turbulence, and sea ice. Understanding and accurately predicting waves are very beneficial to humans.

In the past few decades, researchers have made great strides in studying the causes of wind waves and the correlation between wind and waves (Barnett, 1968). In order to analyze the wind waves field, Sverdrup and Munk (1947) first used an empirical or semi-analytical approach. However, the method has obvious limitations (Kamranzad et al., 2011). Hasselmann (1968) has also studied the evolution of the wind waves' power spectrum in-depth and demonstrated a strong correlation between both wind and wind waves.

At this stage, the mainstream forecasting idea for oceanographers to forecast wind waves is to use numerical models. The numerical model uses oceanic elements such as wind as input and solves complex equations to produce wave forecasts. The most widely used models include the National Weather Service's (NWS) WaveWatch III (WW3) (Tolman et al., 2009), Simulating Waves Nearshore (SWAN) (Booij et al., 1999) developed by the Delft University of Technology, etc (Zheng et al., 2016). Traditional numerical model forecasting methods combine the advantages of physical simulation and data-driven approaches to make forecasts with high spatial and temporal resolution (Wei et al., 2013). This hybrid approach of physical simulation and data-driven prediction is theoretically sound. However, it has significant limitations in practical offshore industry applications: the time lag and its accuracy cannot be guaranteed. In addition, the expensive computational and maintenance costs of the numerical model make it a prudent consideration as an operational application (Song et al., 2022).

In recent years, the application of Artificial Intelligence (AI) in marine and atmospheric sciences has developed rapidly (Van Aartrijk et al., 2002; Bolton and Zanna, 2019). AI can naturally process many data sources, such as numerical forecast results, radar, satellite, station observations, and even decision data (natural language), which is almost impossible for existing coupled sea-air numerical models. Some studies have even found that AI models outperform existing numerical models for short-term wind waves prediction (James et al., 2018). Berbić et al. (2017); Callens et al. (2020) predicted the significant wave height within 3-hour accurately using Random Forest (RF) and Support Vector Machine (SVM), respectively. Fan et al. (2020) used the Long Short Term Memory Network(LSTM) algorithm to predict the significant wave height of several stations for 6-hour, and the results were satisfactory. Song et al. (2020) uses merged-LSTM to mine the hidden patterns in short time series to solve the long-term dependence of series variability and to make compelling predictions of sea surface height anomaly (SSHA). Meng et al. (2021) proposes a bi-directional gated recurrent unit (BiGRU) network for predicting wave heights during tropical cyclones (TCs). Artificial intelligence has the advantage of solid data drive and a high potential for model optimization, which can theoretically solve the "costly" problem of numerical forecast models while improving the "accuracy" of forecasts.

Although the application of AI in wave height prediction is becoming more and more widespread, most of them are limited to single-site forecasting. However, wind wave fields are two-dimensional fields, so predicting wave height at a point is not only a matter of time series but should also consider the spatial correlation with other surrounding points (Jönsson et al., 2003; Gavrikov et al., 2016). In addition, most current AI applications for predicting wave height use single-factor forecasting, treating each ocean variable individually, which ignores the correlation between different ocean elements and lacks physical meaning (Fu et al., 2019). Only the wave field factor is applied to forecast the wave field. The physical correlation between wind and wind waves is ignored. Zhou et al. (2021) established A two-dimensional SWH prediction model based on convolution Long and Short-term memory (ConvLSTM). However, the model only considers the wave and ignores the influence of wind. The mean absolute percentage errors of 6-hour, 12-hour, and 24-hour advance are 15%, 29%, and 61%, respectively. Moreover, the spatial and temporal resolution of the data should also be considered if the deep learning approach is to be truly applied to the problem of forecasting ocean elements. With the deep development of ocean research, human production life increasingly needs to understand the ocean elements with high spatial and temporal resolution. Most of the current deep learning wave forecasting methods are limited to low spatial and temporal resolution conditions, and such research can no longer meet the practical needs of society.

In this study, a deep learning model based on ConvLSTM was developed to combine the correlation between wind and wind waves to predict significant wave heights with high spatial and temporal resolution in the Beibu Gulf. ConvLSTM has been successfully applied to 2D precipitation prediction (Shi et al., 2015). It enables the model to learn the spatial correlation of elements through a unique convolution method, which solves

the problem of spatial information loss in traditional LSTM and improves the accuracy of 2D predictions. Specific modifications to the model were made in this study to enable the model to be adapted to the study sea area and to learn the correlation between wind and significant wave heights in the numerical model data. We then set up a series of experiments to evaluate the performance and accuracy of the model. The model successfully predicts the hourly significant wave height of 1/40° and has an excellent long-term prediction ability. After the model is trained, it is only necessary to provide the model with the corresponding wind speed and significant wave height data to obtain the required predicted significant wave height.

The rest of the paper is organized as follows: Section.2 describes the data and research area we used, and Section.3 describes the method used and the construction and evaluation metrics of the proposed model. Section.4 shows the predictive performance of the proposed model and corrects the problems in the prediction process. Finally, we conclude and discuss future research recommendations in Section.5.

2 Study area and data

2.1 Study area

In this work, the study area is the Beibu Gulf and its adjacent waters in the South China Sea (16°N - 23°N, 105°E - 113°E), as shown in the black box in Figure 1. It includes the shelf waters as well as other waters around Hainan Island; the water depth gradually deepens from the shore to the central part, with an average depth of 42 meters and a maximum depth of more than 100 meters (Gao et al., 2015). The study area is mainly surrounded by some cities in Guangxi, Guangdong, Hainan Province (China), and Vietnam, which are important ports and good fishing grounds (Koongolla et al., 2020). The Beibu Gulf is located in tropical and subtropical areas (Cooke et al., 2011). In winter, it is influenced by cold air from mainland China, with northeast winds and sea surface temperature of about 20C. In summer, the wind comes from the tropical ocean, mainly from the southwest, and the sea surface temperature is as high as 30C. It is often attacked by typhoons. Generally, about five typhoons (Shao et al., 2018) pass here every year.

2.2 Data

The data used in this study are the significant wave height (SWH) and wind speed (WS) data. It is worth noting that the significant wave height data we use refers specifically to the significant wave height of wind waves. These data were provided by the South China Sea Institute of Oceanography, Chinese Academy of Sciences. These data are the products of the WAVEWATCH III and COAM models. The researchers involved have adopted the latest wind stress calculation scheme based on the third-generation wave model WW3, which improves the model's prediction of wind and waves generated by different wind speeds and wind field variations. The model allows a better simulation of the temporal variation of the waves, and the values obtained are closer to the observed values than in the ERA5 reanalysis. The model can provide hourly forecasts with a spatial resolution of 1/40°*1/40°. A more detailed description of the data is available here (Li et al., 2021). Due to the high accuracy of these data, they can be used as an approximation of the observed data in the case of insufficient actual measurement data. Since it is difficult to obtain actual measurement data with high accuracy in the study area, we used

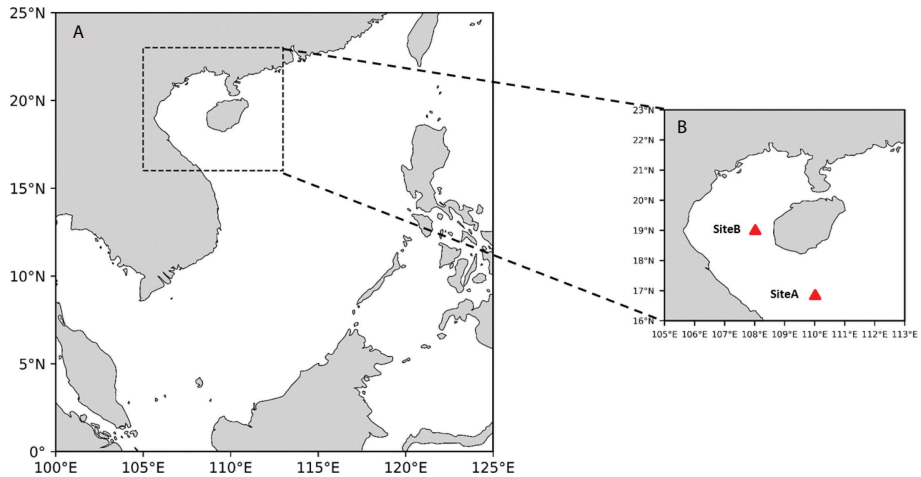


FIGURE 1
(A) South China Sea and the Beibu Gulf, and **(B)** Beibu Gulf and its adjacent waters.

the above data as a comparison value in our study. In this study, SWH and WS data with Spatio-temporal resolution of 1h and 1/40°*1/40° for two years from 2018-2019 were selected, with 80% of the data used as the training set, 10% for validation, and 10% for testing. The maximum significant wave height in the data is 7.34m and the top wind speed is 17m/s. In Section 4, we also compare the predictions with the ERA5 reanalysis information used, which can be found here (www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5). It is worth stating that the high-resolution data used in this study will be open-sourced to facilitate researchers in studying important wave height issues at high resolution. These data are available here: citep https://doi.org/10.5281/zenodo.6402321.

3 Methods

This study focuses on the significant wave height variation over the entire study area rather than on specific stations. Therefore, for each point in the study area, we need to consider it in terms of time series and spatial relationships. This study proposes a novel prediction method based on Mask-ConvLSTM deep learning network and Replace mechanism. In this study, specific modifications are made to the ConvLSTM model, which allows the model to be adapted to our study sea area and learn the physical correlation between wind speed and significant wave height. The model was used to predict the SWH conditions after a few hours. The number of layers of the network is three, containing 6,12,2 convolutional kernels, respectively, the size of these convolutional kernels is set to 3*3, and the step size of each move is 1. Our experiments were conducted on a cluster of computers. This study used an NVIDIA TeslaV100S and Intel(R) Xeon(R) Silver 4214R CPU in terms of hardware. Regarding software, this study used Tensorflow-2.4.1 and CentOS 7.6.

3.1 Convolutional LSTM network

The LSTM algorithm, known as Long short-term memory, was first proposed by Hochreiter and Jürgen Schmidhuber in 1997 and is a particular form of RNN (Recurrent neural network), while RNN is a general term for a series of neural networks capable of processing sequential data (Hochreiter and Schmidhuber, 1997). In 2005, Alex Graves and Jürgen Schmidhuber proposed a bidirectional long short-term memory neural network (BLSTM) based on LSTM, also known as vanilla LSTM (Graves and Schmidhuber, 2005). It is one of the most widely used LSTM models at present. The ability of LSTM to remove or add information to nodes to change the information flow state relies on the careful regulation of the gate structure (Meng et al., 2022). Gates are nodes that can be selected to pass information, and they consist of Sigmoid complexes and point-by-point multiplication operations. The ingenuity of the LSTM lies in the addition of input gates, forgetting gates, and output gates for protecting and controlling the information flow vector states. In this way, the scale of integration can be changed dynamically at different moments with fixed model parameters, thus avoiding the problem of gradient disappearance or gradient expansion (Hochreiter et al., 2001). The input gate determines how much of the input data of the network at the current moment needs to be saved to the cell state. The forgetting gate determines how much of the cell state needs to be preserved in the current moment from the last moment. The output gate controls how much of the current cell state needs to be output to the current output value. The computation of the LSTM layer can be expressed as follows.

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci} \circ c_{t-1} + b_i) \quad (1)$$

$$f_t = \sigma\left(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf} \circ c_{t-1} + b_f\right) \quad (2)$$

$$c_t = f_t \circ c_{t-1} + i_t \circ \tanh\left(W_{xc}x_t + W_{hc}h_{t-1} + b_c\right) \quad (3)$$

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + W_{co} \circ c_t + b_o) \quad (4)$$

$$h_t = o_t \circ \tanh\left(c_t\right) \quad (5)$$

where t_i denotes the input gate, ft denotes the forget gate, O_t denotes the output gate. C_t and C_{t-1}denote the state at the current and previous moments, respectively. W is the assigned weight for each layer, x_t is the input time step at the current moment, and b is the bias. σ denotes the sigma operation. X denotes the Hadamard product.

The internal structure of the hidden layer of the LSTM is shown in Figure 2. The forgetting gate f_t determines which information coming from the information state h_{t-1} at the previous time node needs to be discarded and which needs to be retained. The input information x_t from the current moment and h_{t-1} from the previous moment are simultaneously fed into the sigmoid activation function, and the output value is the value of the forgetting gate f_t. The value range of f_t is between (0, 1), and the closer the value is to 1 means that the information passing through the forgetting gate should be retained, and vice versa, it should be discarded. The input gate i_t controls which new inputs will be kept in the cell state. The current moment's input x_t and the previous moment's information state h_{t-1} are first fed to the sigmoid activation function, which adjusts the value of the input gate i_t to a value between (0, 1). Then x_i

and h_{t-1} are jointly delivered to the tanh function to create a new candidate cell state c_t for the current moment, which is followed by the LSTM layer back to update the cell state c_t for the current moment. The forgetting gate $f_t$is used to control which information in the previous moment's cell state c_{t-1} needs to be discarded, and then the input gate i_t is used to determine which

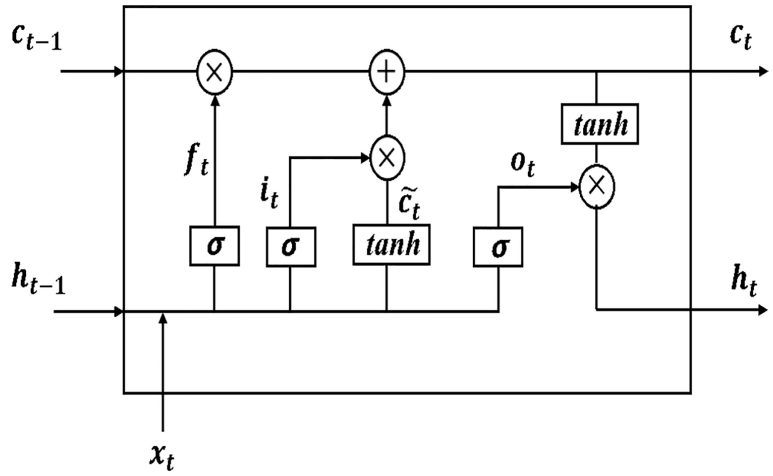


FIGURE 2
Internal structure of LSTM hidden layer.

information in the current moment's candidate state $c_t$will be retained in the new cell state, respectively, using the product calculation. Finally, the product of the two is summed to obtain the cell state c_t at the current moment. output gate o_t controls the output of the current information state, i.e., the information state h_t input to the next time node, which is jointly determined by x_1, h_{t-1}, and c_t.

The limitation of the LSTM application in the ocean domain is that it can only handle time-series data from a single location. It is well known that the ocean is a dynamically changing whole, and different points are temporally and spatially correlated with each other (Magdalena Matulka and Redondo, 2010). Although, researchers can divide the complete ocean into multiple points and use LSTM to process them one by one. However, this approach ignores the regional characteristics of different oceans and the interactions between neighboring points of the same ocean. To address this problem. Shi et al. (2015) improved the LSTM and firstly proposed the Convolutional LSTM Network (ConvLSTM). He and his team use ConvLSTM for rainfall forecasting. They have collected many radar plots which give the distribution of clouds in a given region. Moreover, these maps are changing along the time axis. So with the past timeline and cloud cover maps, it is possible to predict where the clouds should go at future points in time, weather changes, and the chances of future rainfall in an area. Using traditional LSTM models leads to the loss of geolocation information in the cloud cover map, and therefore it is difficult to predict where the clouds will move. The contribution of the original paper is to add the convolution operation that can extract spatial features to the LSTM network that can extract temporal features and propose the architecture of ConvLSTM. ConvLSTM inherits the advantages of traditional LSTM and makes it well suited for Spatio-temporal data due to its internal convolutional structure.

The computation of the ConvLSTM layer can be expressed as follows.

$$i_t = \sigma(W_{xi} * X_t + W_{hi} * H_{t-1} + W_{ci} \circ C_{t-1} + b_i) \quad (6)$$

$$f_t = \sigma(W_{xf} * X_t + W_{hf} * H_{t-1} + W_{cf} \circ C_{t-1} + b_f) \quad (7)$$

$$C_t = f_t \circ C_{t-1} + i_t \circ \tanh(W_{xc} * X_t + W_{hc} * H_{t-1} + b_c) \quad (8)$$

$$o_t = \sigma(W_{xo} * X_t + W_{ho} * H_{t-1} + W_{co} \circ C_t + b_o) \quad (9)$$

$$H_t = o_t \circ \tanh(C_t) \quad (10)$$

where * is convolution operator.

The most important feature of the ConvLSTM algorithm is that it replaces the matrix multiplication in the LSTM with convolution operations. However, its essence is still the same as LSTM, using the previous layer's output as the input of the next layer. The difference is that with the addition of the convolution operation, the temporal relationships can be obtained, and the spatial features can be extracted like the convolution layer. In this way, Spatio-temporal features can be obtained. Regional wave height forecasting is a typical Spatio-temporal problem. Therefore, the proposed model uses the ConvLSTM algorithm.

3.2 Forecasting method

For the wind wave prediction problem with high spatial and temporal resolution, we would like the proposed model to make longer time predictions. However, if we perform multi-step prediction directly, the error of the results may be unstable. Therefore, the proposed model adopted a different approach

from most current forecasting methods that directly establish correlations between specific future moments and historical data. Instead, this study used a more appropriate forecasting strategy to improve the long-term predictive capability of the model. According to previous studies, the Rolling Mechanism (RM) is more suitable for dealing with high-frequency and long-time forecasts (Akay and Atak, 2007; Kumar and Jain, 2010).

The main idea of the RM method is to use the obtained forecast data as the latest data and add it to the future forecast. Figure 3 shows the process of forecasting. The small boxes represent the data for each moment. The numbers in the small boxes represent the moments of the data, and the large boxes represent the historical data used for each forecast (time window), which is of length *N*. Our forecast is a single-step forecast. First, we use the historical data from *T-N* to the moment *T*

to forecast the data for the future moment *T+1*. Immediately after that, the time window is shifted down by one step. We treat the data just obtained for *T+N* as known data and use the *N* data from *T-N+1* to *T+1* to forecast the data at the moment *T+2*, repeating the above process n times. We then get the data from *T+1* to *T+n* moments. In this way, the prediction process of using historical data continuously to predict the next n moments is completed.

It is worth noting that the standard convolution operations in deep learning algorithms such as CNN and ConvLSTM can only act on standard rectangular areas. In our study, the study area is not a pure sea area, but a land-sea combined area, including Hainan Island and parts of the eastern Indo-China Peninsula. Due to the existence of land points, the error calculated during model training will be affected by these land points, thus affecting the effectiveness of the model. This is not what we expected. To solve this problem, we propose the Mask method described below. "Mask" is an idea in deep learning. In simple terms, it is equivalent to masking or selecting some specific elements by putting a mask over the original tensor. This study proposes the Mask method by combining the "Mask" idea.

A brief description of the Mask method is given in Figure 4. This study generates a matrix of size equal to the input data so that the points in the matrix can correspond to the points in the input data one by one. It will be referred to as the Mask matrix. The antique-white part in Figure 4 represents the land area, and we set the value of the Mask matrix corresponding to this area to 0 and the value of the Mask matrix for the ocean part represented by the blue part to 1. It is worth noting that, in practice, the land-sea distribution of the study area is much more complicated than in Figure 4. But because of the specificity of the right-angle grid data. We can still build the corresponding Mask matrix according to the idea in Figure 4. During the model's training, the model determines the direction of the subsequent gradient descent by calculating the average error between the results and the labels. To implement our Mask method, we rewrite the loss function for network training according to Eqs.11 and 12 during network training.

$$\Phi_{ocean} = U - \Phi_{land} \tag{11}$$

$$loss = \frac{\sum_0^n ((X(t) - Y_t) \circ Mask)^2}{N} (N \in \Phi_{ocean}) \tag{12}$$

where *U* denotes the ensemble of points in the study region, Φ_{land}denotes points in the land region, and Φ_{ocean} denotes points in the ocean region. *X(t)* denotes the output matrix of the prediction at this time. Y_t denotes the corresponding factual matrix, and N is the number of points in the ocean area.

In this way, during the training process of the network, the result $X_{(t)}$ of each prediction is subtracted from the control value Y_t and then dotted multiplied with the Mask matrix. Since the result is dot multiplied by Mask, and the value of the land part in the Mask matrix is zero, the error of the corresponding land part in the error matrix will also be zero. Therefore, the network only considers the error value of the ocean part in the loss value

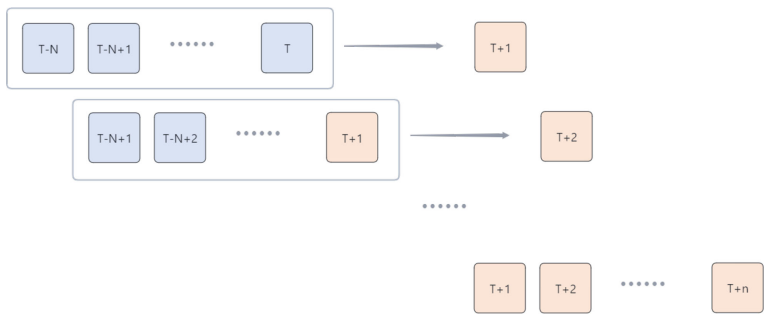


FIGURE 3
The prediction method integrated with RM.

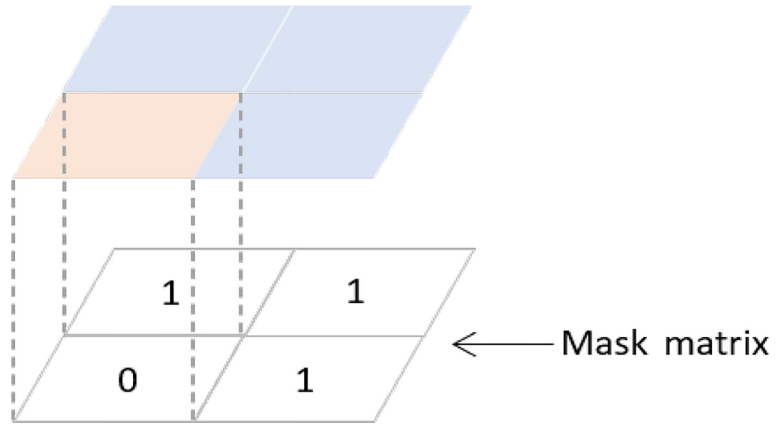


FIGURE 4
Idea of Mask method. The antique-white areas represent land areas and the blue areas represent ocean areas.

calculated in each iteration. In this way, the influence of the land region on our experiments is eliminated.

3.3 The proposed model

The wind and wind wave data used in this study have a spatial resolution of 1/40°*1/40° and a temporal resolution of 1 hour. Such high spatial and temporal resolution data means that the degree and speed of sea state variability are much more significant than other slightly lower resolution data. It increases the difficulty of forecasting. As mentioned earlier, there is a strong correlation between wind and wind waves in physical oceanography. If we want to exploit this correlation, we need to have both wind speed and significant wave height as inputs to the model and to have the two dependent on each other. So we need a multi-input network structure to capture this physical correlation and the subtle sea state variations. In this study, we combine the Mask method with ConvLSTM and change the number of input channels of the model to dual channels. It enables the model to meet our needs. The structure of the model proposed in this study is given in Figure 5.

The input data for the model are $X_{t\text{-}N}$ to X_t. Each input data consists of the wind speed (blue quadrilateral in Figure 5) and significant wave height data (green quadrilateral in Figure 5) at that moment. The data N is the length of the time window we choose. The size of these input data is (321*281) and after combining them into dual-channel data, each input X has a dimension of (321*281*2). These input data enter the regularization layer for regularization and are then fed to the three Mask-ConvLSTM layers. Between each Mask-ConvLSTM layer, model use relu as the activation function. During the model training process, the network then learns the Spatio-temporal correlation of the input data and the physical correlation between SWH and WS. Then, the size of the output data that we need to obtain is controlled by the convolution layer(Conv3D). In this way, we obtain the predicted data at $X_{t\text{-}N}$ moments and then add the obtained predicted data at X_{t+N} moments to the RM module to achieve rolling forecasts.

3.4 Evaluation metrics

In order to evaluate our model reasonably, this study selected a variety of evaluation metrics commonly used to evaluate the significant wave height prediction problem, including Root Mean Square Error (RMSE), Scatter Index (SI), and R Square (R^2). In this, the SI can measure the percentage of RMSE relative to the average actual value. However, due to the specificity of the study sea area, we make some modifications to these standard evaluation metrics to match our problem. We combine the indicator RMSE,SI with our Mask method to make it possible to focus only on the error situation in the marine area. It ensures no disturbances from the terrestrial values in the area and that there are no erroneous undercounts due to incorrect, missing point counts. The mathematical equation for these evaluation indicators is as follows:

$$RMSE = \sqrt{\frac{\sum_0^N (X - Y)^2}{N}} \tag{13}$$

$$SI = \frac{RMSE}{\bar{Y}} \tag{14}$$

$$Mask - RMSE = \sqrt{\frac{\sum_0^N ((X - Y) \circ Mask)^2}{N(N \in \boldsymbol{\Phi}_{ocean})}} \tag{15}$$

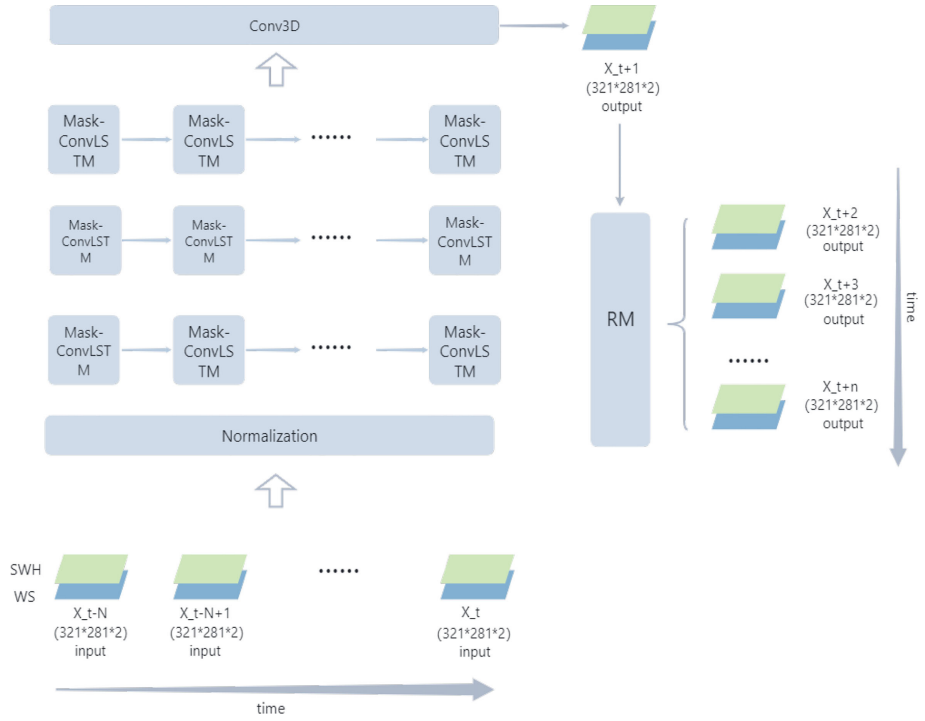


FIGURE 5
Architecture of the proposed model.

$$Mask - SI = \frac{Mask - RMSE}{\bar{Y}} \tag{16}$$

$$R^2 = 1 - \frac{\sum_0^N (X - Y)^2}{\sum_0^N (Y - \bar{Y})^2} \tag{17}$$

where X represents the predicted value, Y represents the corresponding value, and N is the number of points in the taken region.

4 Results

The effect experiments are based on the significant wave height and wind speed data for 2018-01 to 2019-12 mentioned in Section 2. In order to test the performance of the model, this study conducted multiple sets of controlled experiments. The forecast tests were conducted in the validation set that did not participate in the training.

4.1 Performance study

In order to verify the superiority of the proposed model for the high-resolution significant wave height prediction problem, this study compares several published significant wave height prediction methods. The compared methods include the traditional machine learning methods RF (Callens et al., 2020) and SVM (Berbić et al., 2017)mentioned in Section 1 and the LSTM (Fan et al., 2020), GRU (Meng et al., 2021) algorithms in deep learning. Five sets of experiments, including the model proposed in this study, used the same training data, and their performance is shown in Figure 6.

Firstly, this study conducted effect experiments for three different forecast lengths of 1 hour, 3 hours, and 6 hours. To more visually show the comparison of the effects between different methods, we chose two sites, siteA(17°N, 110°E) and siteB(19°N,108°E)(shown in Figure 2), to conduct our experiments. Figures 6A, D) show that at a forecast length of 1-h, there is little difference between the predicted and comparison values of the five methods. When the time grows to 3-h, the results of the RF methods show significant differences from the comparison values, especially in the case of low wind waves (Figure 6E). Although the SVM algorithm can predict the trend of data variation, the difference in values is significant. When it comes to 6-h, the RF and SVM methods have completely lost their forecasting ability. The effect of the three groups of deep learning algorithms also appears to be very different. Although LSTM, GRU, and ConvLSTM all capture the change in significant wave height, ConvLSTM has more accurate numerical magnitude predictions than the other two.

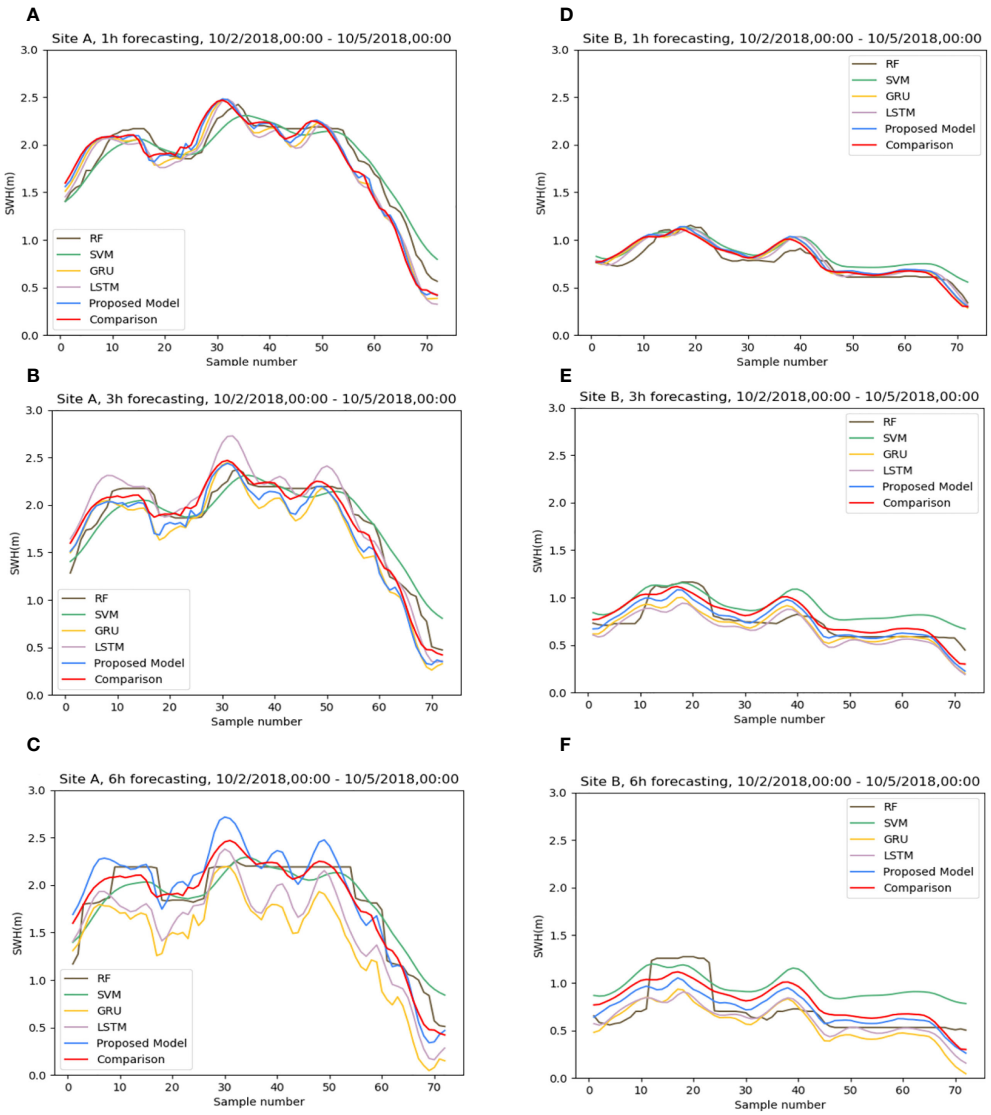


FIGURE 6
The differences in prediction effects of different algorithms at sites **(A, B)**: **(A, D)** 1h, **(B, E)** 3h, **(C, F)** 6h. Sample number indicates the time sample number.

To avoid the effect of chance, we performed the above experiments for all points in the sea area and averaged the results of the comparisons, which are shown in Figures 7, 8. Figure 7A shows the variation of RMSE for the five groups of algorithms. It can be seen that the RMSE of the SVM method has reached about 0.18 at a prediction time of 1 hour. We can also visualize it in Figure 8B. The SVM algorithm lost effectiveness when the predicted large wave height was below 0.5 m. The remaining groups of algorithms do not differ much. The prediction effectiveness of the conventional machine method decays severely with time. Especially for smaller apparent wave heights, both SVM and RF show different degrees of inaccuracy. The RMSE of the SVM algorithm reaches about 0.3 for 3 hours and even 0.7 for 6-hour forecasts, while the RMSE of RF also reaches about 0.6. Two groups of deep learning methods, LSTM and GRU, have acceptable performance in the early stage (Figure 8A, C. However, the error after 6 hours is also much higher than that of the ConvLSTM method, with the RMSE increasing to more than 0.4 and the SI reaching 0.2. The comparison of the SI of different experimental groups in Figure 7B also shows the superiority of the ConvLSTM algorithm in this study. As we mentioned in Section 3, when dealing with such high spatial and temporal resolution data if only the temporal correlation of one point is considered without considering the spatial relationship of each point. It would be difficult for the model to predict the sea state changes within the

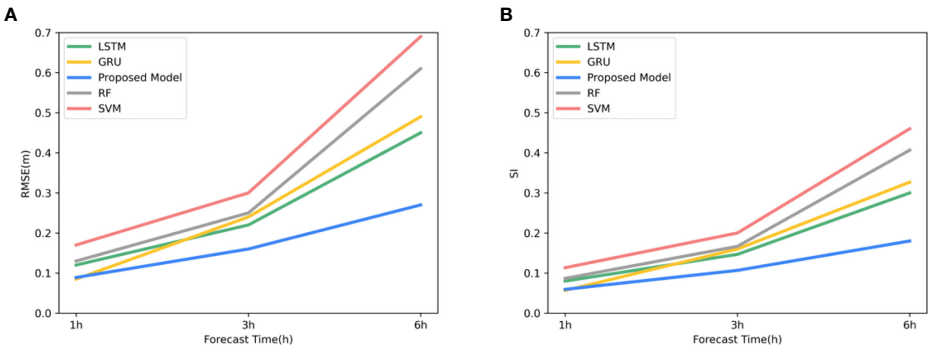


FIGURE 7
The difference in prediction effect of different algorithms: **(A)** RMSE of different algorithms varies with the forecast time, and **(B)** SI of different algorithms varies with the forecast time.

sea area accurately. It may be the reason for the poor performance of the LSTM and GRU algorithms in this study.

4.2 Ablation study

In order to verify the advantages of the wind and wind waves dual-channel compared to the wave-only single channel and the effectiveness of introducing the Mask mechanism, we set up four different sets of experiments. First, a single-channel experiment using only significant wave height data for training and no Mask method (Single-channel). Second, single-channel experiments use only significant wave height data for training but with the Mask method (Single-channel & Mask). Third, experiments using significant wave height and wind speed data for training but without using the Mask method (Dual-channel). Finally, experiments using the significant wave height and wind speed for training with the Mask method (Dual-channel & Mask). It is worth mentioning that the results obtained by the dual-channel network during the prediction process are also dual-channel, containing both the prediction results of the significant wave height and the wind speed. This study aim to study the network's prediction of the significant wave height. Therefore, we take only the significant wave height prediction results from the dual-channel network forecast results when comparing different groups of experiments.

To ensure the validity of these four sets of experiments. The hyperparameters of the four sets of experiments will be set to the same set of values. We randomly selected a period of history and made a 6-hour forecast downward using each of the four models. Figure 9 shows the results of our experiments. It can be seen that the error of the dual-channel network is significantly smaller than that of the single-channel network when only the number of channels is considered, and this situation persists with increasing forecast time. It proves the advantage of the dual-channel network. With the introduction of wind speed, the network takes advantage of the physical correlation between the two to improve our forecasting results for the significant wave height. Similarly, we can see that after the Mask method is used. When the channels are the same, the Mask-RMSE in the two groups of experiments using the Mask method is smaller than in the other two groups of experiments without the Mask method. However, we noticed a particular case. Before 3-h, the error of the single-channel network with the Mask method is smaller than that of the dual-channel network without the Mask method. It may be because the Mask mechanism can dominate the error situation brought by the forecast in the short term. However, this slowly disappears as the forecast time goes on, and both sets of dual-channel experiments slowly outperform the two sets of single-channel experiments. The above findings demonstrate the superiority of the wind and wind waves dual-channel network compared to the wave single-channel network and the effectiveness of the Mask method in this experiment.

4.3 Conventional forecast

After completing the training process of the model, we first explore the performance of the proposed model on intermediate time scales. As input data, we use 6 hours of data from 00:00-12:00 on August 28, 2019. To more visually demonstrate the ability of the proposed model to predict the high-resolution significant wave height of the study sea, we plot the results. Results and error statistics are shown in Figure 10 and Table 1.

In Figure 10, the three subplots (a), (d) and (g) on the left side represent the significant wave height forecasts for 1,3,6-h, respectively. The numerical model data at the corresponding time are shown on the right side. Their evaluation indicators are shown in Table 1. We can see from Figures 10A, B that our model accurately captures the distribution of the significant wave height at 1-h with a Mask-RMSE of only 0.08 and accurately predicts the higher wind waves in the northwestern

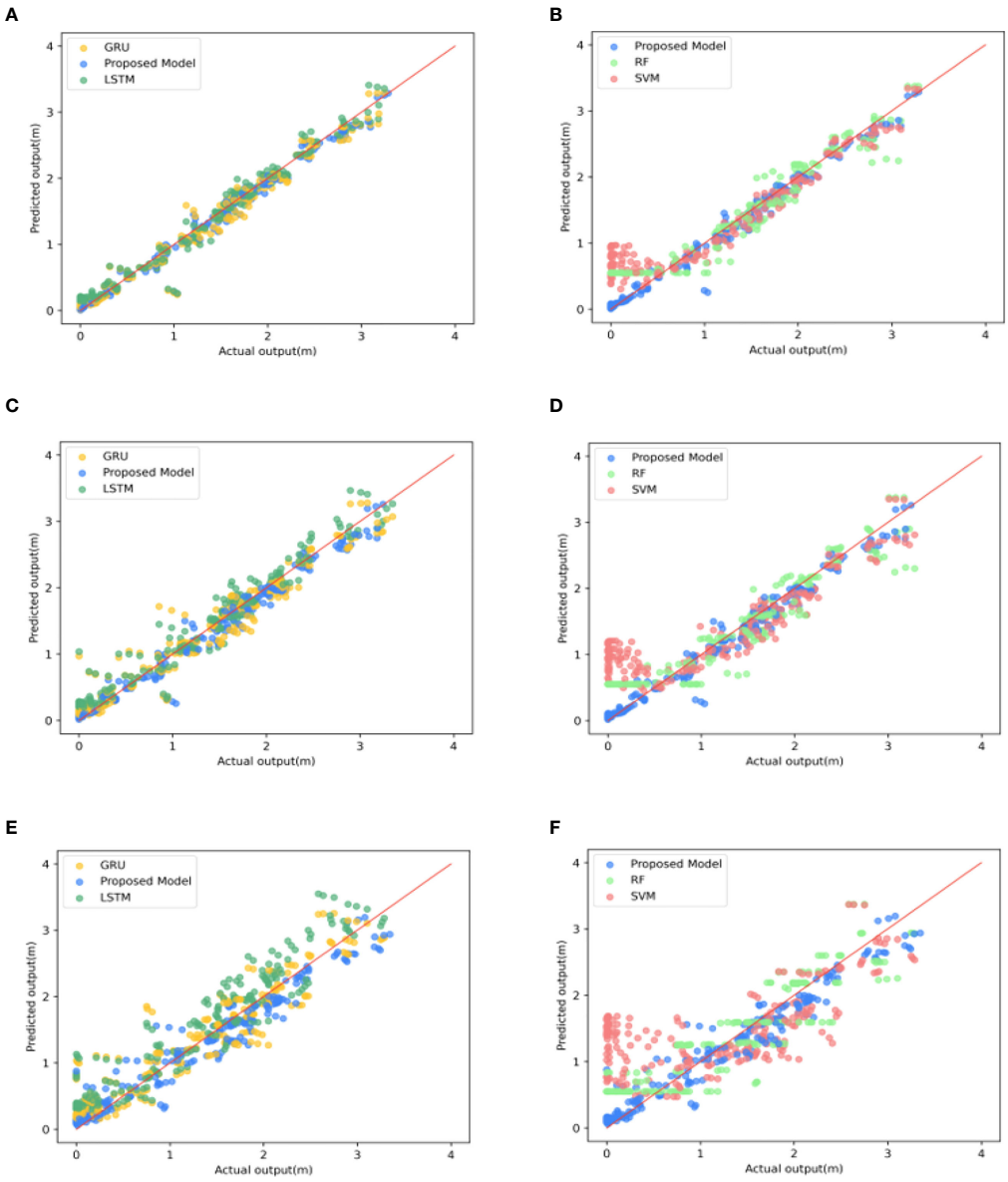


FIGURE 8
Scatter plots of five groups of algorithms in different forecast times: **(A, B)** 1 h, **(C, D)** 3 h, **(E, F)** 6 h.

part of Hainan Island. After increasing the time to 3-h, the proposed model still has high accuracy and low error value, and the R^2 value can be maintained at 0.98. It also has a good prediction ability for distributing significant wave height and wave height size in the region. However, numerically, the predicted values for the western part of Hainan Island in Figure 10D are smaller than the corresponding values. When the time window comes to 6-h, although our model can still capture the significant wave height distribution in the sea, there is a significant difference in the values(Figures 10G, H). Mask-RMSE increases to 0.27. However, this value is perfectly acceptable for such high spatial and temporal resolution data. Overall, the forecasting effect of our model is excellent.

4.4 Analysis of error sources and treatment

The model we use is a dual-channel network. In the prediction process, both SWH and WS channels generate errors. Because the network considers the characteristic correspondence between wind and wind waves, if the error in

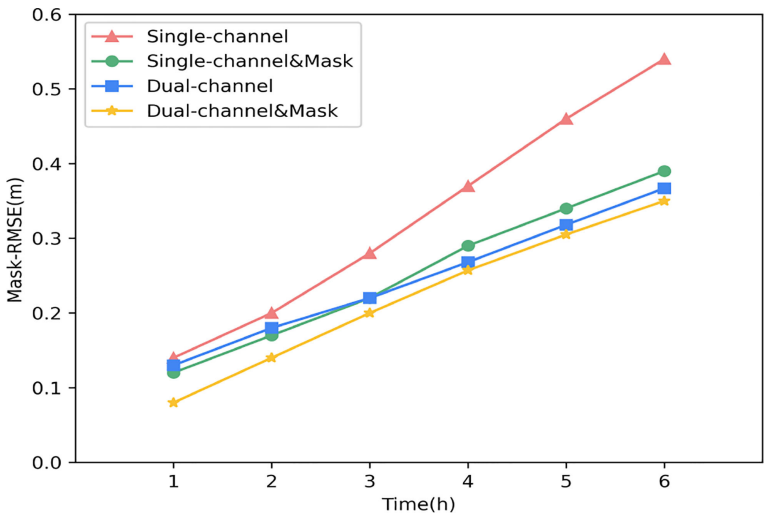


FIGURE 9
Mask-RMSE of the four groups varied with the forecast times.

one channel is too large, it will also decrease the prediction of the other channel. In order to study the variation of the error of the two channels with the forecast time during the forecast, we extended the forecast time to 12-h in Section.4.3 and analyzed the error sources.

For the experimental results in Section.4.3, we analyze the error variance of the two channels in the prediction process separately. The yellow line in Figure 11 shows the trend of the prediction error of WS over time, and the blue line shows the trend of the error of SWH. It can be seen that the Mask-RMSE of both channels increases with the prediction time due to the rolling mechanism. However, on the way up, the blue line shows a steady upward trend over time, while the yellow line shows a strong upward trend, especially after 6 hours. It shows that the predictive validity of the WS channel drops significantly after 6 hours. The correlation between the two channels is considered in the model prediction process. The sharp increase in the error of the WS channel will directly lead to an increase in the error of the SWH channel, increasing the total error of the prediction results. We pioneered a new mechanism called Replace and added the Replace mechanism to the RM process to solve this problem.

The core idea of the Replace mechanism is to replace the predicted wind speed values of the wind speed channel in the network with the wind speed of the numerical model. As shown in Figure 12, the model is run to get the forecast output X_{t+1}, and the RM mechanism uses the obtained X_{t+1} to continue the rolling forecast to get the output results for the next n-1 moments. The black dashed box in Figure 12 explains the specific steps of the Replace mechanism. As we mentioned before, the forecast results of each step of the model include forecast SWH and forecast WS. However, the error of WS increases sharply during the rolling forecast process (Figure 11), leading to an increase in the error of SWH associated with it. Suppose we can solve the problem of a sharp increase of WS in this process. Then the forecast results will be improved. To solve this problem, we perform the following operation for each model and RM mechanism forecast result X_{t+m}: replace the WS obtained from the network forecast with the WS of the numerical model to form a new X_{t+m} consisting of the numerical model WS and the network forecast SWH (red dashed box in Figure 12). This data is used to replace the original X_{t+m} for RM processing.

To verify the effectiveness of the Replace mechanism, we conducted a set of controlled experiments. The error profiles of SWH before and after adopting the Replace mechanism are shown in Figure 13. The yellow line is the error plot of SWH after adopting the Replace mechanism, and the blue line is the forecast error without the Replace mechanism. It can be seen that the adoption of the Replace mechanism significantly reduces the overall forecast error, especially in the medium and long time scales. Compared to the previous one, the Mask-RMSE is even reduced by up to 50%.

4.5 Long time scale forecasting

In previous deep learning wave height forecasting studies. Within the margin of error, the effective forecast duration obtained using hourly data was typically limited to 6-12 hours. If the data resolution is increased, this time will be further shortened. To investigate whether the proposed model can make

10.3389/fmars.2022.983007

predictions on long time series after using the Replace mechanism, we set up two sets of experiments. One group had Replace mechanism, and the other group had no Replace mechanism. We adjusted the timestep to 12 and retrained the model using previous data from each set of experiments. The results of the two experiments are shown in Figure 14, where (a) is the numerical model data, (b) is the prediction result of the model with the Replace mechanism, and (c) is the prediction result of the model without the Replace mechanism.

We predicted the significant wave height for the next 24-hour using the data from 12:00-24:00 on August 2, 2019. The variation of each indicator with time is shown in Table 2. It can be seen that the two sets of experiments still maintain good performance at 1-hour and 3-hour (Figures 14B, C). The significant wave height and the distribution of wind waves in the sea still have a good forecasting effect. When the time reached 8-hour, the effects of the two groups of experiments began to show more apparent differences. The network without the Replace mechanism has a higher prediction of the significant wave height in the eastern part of Hainan Island. In contrast, the network with the Replace mechanism can still predict the change of the significant wave height in the sea more accurately, but the distribution has slightly deviated. When we increase the prediction scale to 12-hour, it can

TABLE 1 The evaluation metrics of the model at different forecast time.

Time (h)	Mask-RMSE	Mask-SI	R^2
1*h*	0.08	0.05	0.99
3*h*	0.16	0.10	0.98
6*h*	0.27	0.17	0.96

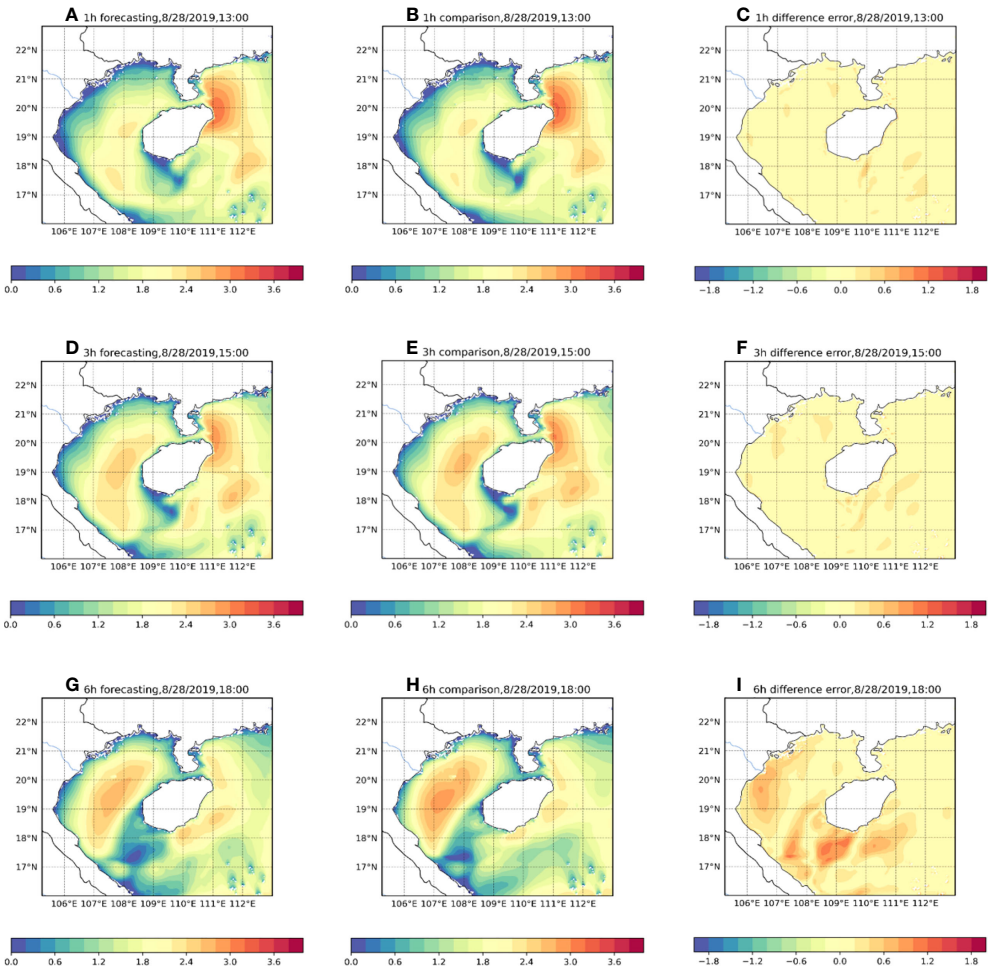


FIGURE 10
Comparison of forecast effect of the model under different forecast times. **(A, D, G)** are the predicted significant wave height effect diagrams for 1h, 3h, and 6h, respectively. **(B, E, H)** is the significant wave height diagram of the numerical model at the corresponding time. **(C, F, I)** is the difference error between forecasting and comparison.

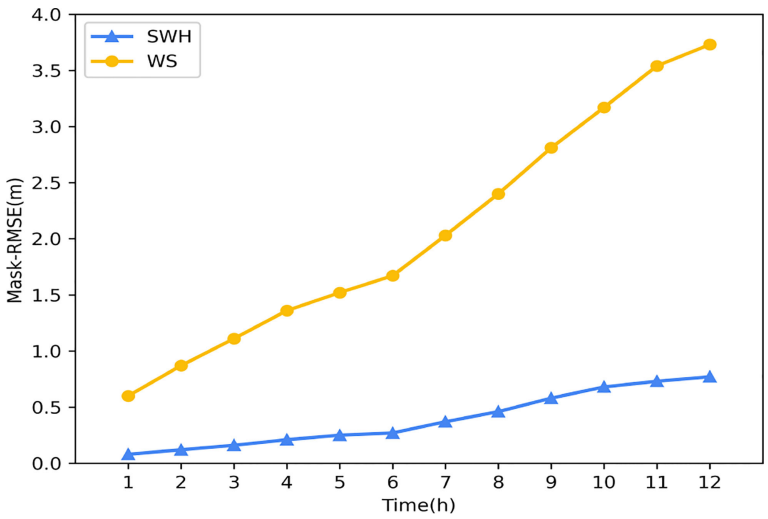


FIGURE 11
The Mask-RMSE of significant wave height and wind speed changes with increasing forecast time.

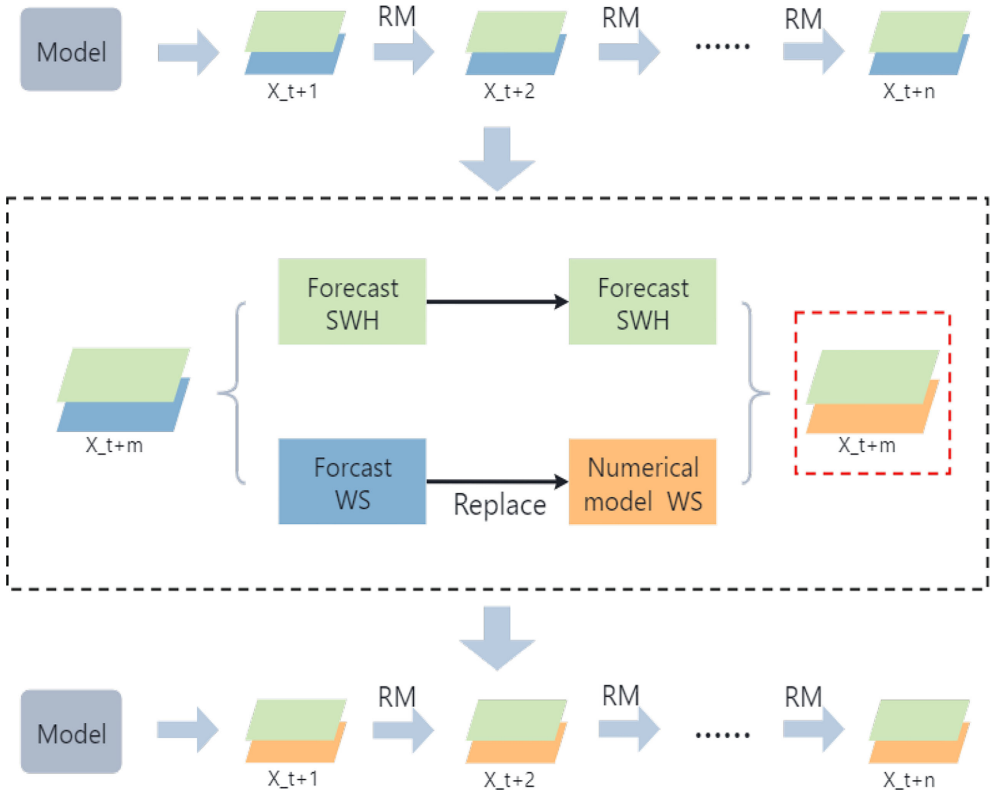


FIGURE 12
The idea of Replace Mechanism. The black dotted box represents the core steps of the RM mechanism, and the red box represents the data obtained through the RM mechanism.

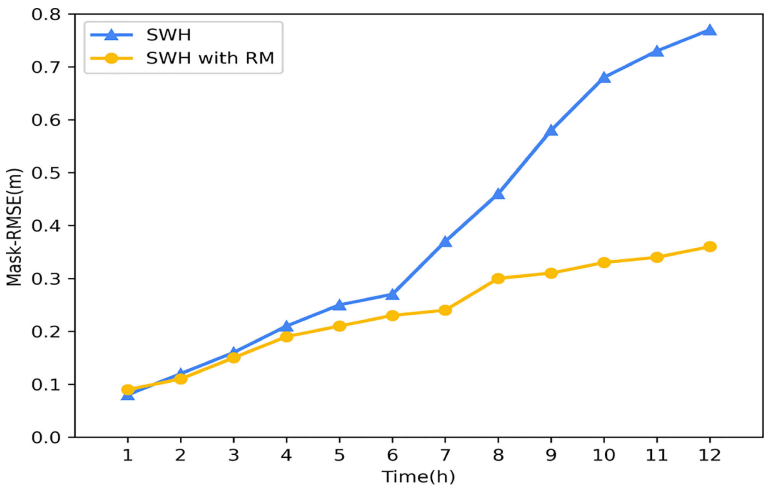


FIGURE 13
Comparison diagram of Mask-RMSE changes of the model before and after using the RM mechanism.

be seen that the model without the Replace mechanism has a significant prediction error for the change of significant wave height in the sea, and R^2 drops to 0.75, basically losing the accuracy of the forecast. However, the model with the replacement mechanism still has good forecasting ability, with a Mask-RMSE of 0.29. Although there are some regions where the significant wave height values are under-predicted, the overall distribution can still be predicted more accurately. When the prediction scale is expanded to 16-hour, the model without the replacement mechanism has lost its predictive power, with R^2 dropping to 0.62. The model with the replacement mechanism also shows a decrease in accuracy. Although the high wind and wave area near Hainan Island can be predicted, the magnitude of the significant wave height obtained from the prediction has been significantly different, with R-squared dropping to 0.78. When the time reaches 24-hour, the model without the replacement

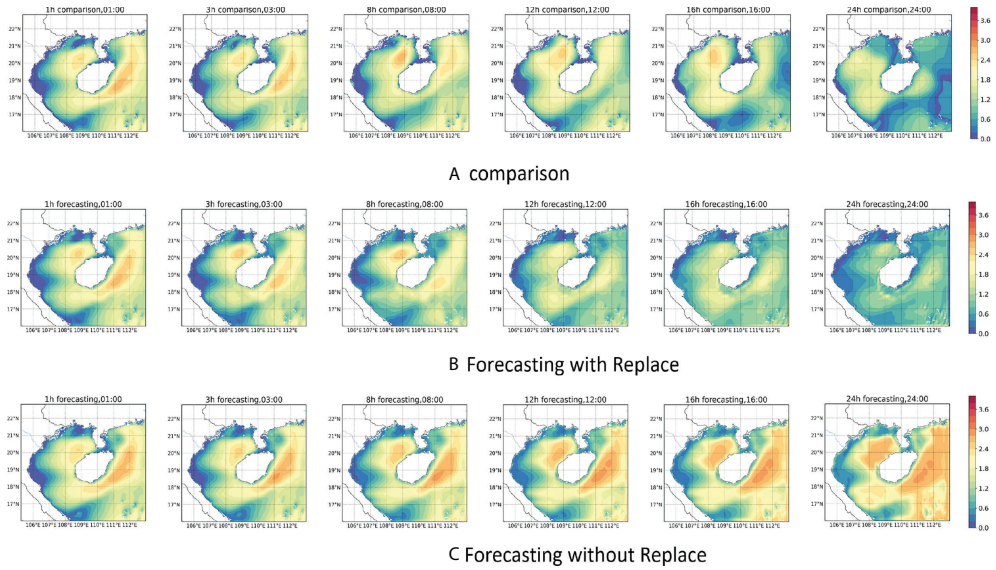


FIGURE 14
24-hour forecast results of significant wave height. **(A)** is the significant wave height diagram of the numerical model. **(B)** is the predicted significant wave height diagram after Replace mechanism is adopted in the model. **(C)** is the predicted significant wave height diagram without Replace mechanism.

TABLE 2 The Evaluation index of the two model at different forecast time.

		Forecast time(h)					
	Evaluation index	1h	3h	8h	12h	16h	24h
Long time forecast with Replace mechanism	Mask-RMSE	0.09	0.15	0.25	0.29	0.40	0.49
	R^2	0.99	0.98	0.96	0.88	0.78	0.69
	Mask-SI	0.06	0.10	0.19	0.23	0.29	0.36
Long time forecast without Replace mechanism	Mask-RMSE	0.10	0.18	0.35	0.45	0.63	0.81
	R^2	0.99	0.97	0.86	0.75	0.62	0.51
	Mask-SI	0.06	0.12	0.27	0.36	0.45	0.59

mechanism already shows confusion in the forecasts. The forecasts obtained from the predictions have little to no relationship with the contrasting values (Figure 14A). Although the model with the replacement mechanism can predict the distribution of significant wave heights in the sea area, the prediction of the magnitude of significant wave heights in the whole sea area also shows a more noticeable difference, and the R-squared drops to 0.69. Through the above experiments, we can see that: the Replace mechanism is very effective in improving the forecasts of the dual-channel network. In particular, it can effectively reduce the forecast error in a more extended time range. With the Replace mechanism, the long-time forecasting capability of the network is substantially improved, and the significant wave height at high spatial and temporal frequencies within 24 hours can be predicted more accurately.

In order to verify the stability of the model's forecasting ability, this study selected the data of SiteA and SiteB for the whole of October 2018, with a total of 744 samples. Moreover, forecast tests with different lengths were conducted using these data, and the results are shown in Figure 15. The red line in Figure 15 represent the comparison values. The yellow line represents the ERA5 reanalysis data. The blue line represents the forecast values of the model. It can be seen that the model has an excellent forecasting effect within 6-h. Both the magnitude and the trend of the values differ very little from the comparison values. When the time comes to 12-h, the model can predict the trend of the significant wave height. However, the values are not stable. When the time comes to 24-h, The model can predict the primary trend, but the values are much different in some cases. It may be because our forecasts are regional, focusing on the trend of the significant wave height within the whole region rather than the significant wave height at a single point. In addition, the insufficient amount of data in this study, which only used two years of data, might be another reason.

5 Conclusion and discussion

The main work of this research is to combine the laws of physical oceanography and use the deep learning method to predict the significant wave height of the entire Beibu Gulf with high spatial and temporal resolution. This study uses a modified ConvLSTM network to explore the Spatio-temporal correlation of historical data and the physical correlation between different marine elements. By comparing with other methods, we show the advantages of the proposed model in dealing with Spatio-temporal data. By introducing a two-dimensional wind speed field, the significant wave height prediction capability of the network is greatly improved. At the same time, this study proposes a Mask method to solve the problem of wave height prediction in sea areas with land areas. This study uses high-resolution wind speed and significant wave height data from 2018 to 2019 to conduct forecast experiments on different forecast time scales. The validity of the model is demonstrated by comparison with the numerical model. The R-squared for the 6-hour forecast reached 0.96. We analyze the error generation during the prediction process and propose an alternative mechanism to alleviate the error diffusion problem of the proposed model and prolong the effective prediction time of the network. The effective forecast time reaches 24 hours. Most importantly, the proposed model has practical application value. We know that numerical models predict waves with the wind as input. The method of this study can also utilize these wind data for wave forecasting and compare the results with short-term numerical models. It will significantly reduce the cost issues associated with numerical models. Typically, the time cost for a numerical model run to complete a prediction is over several hours and requires supercomputer support. In contrast, the model proposed in this study can be run on an ordinary PC, and the time needed for prediction is only a few seconds. It is worth pointing out that the present study was conducted for wind waves. The method proposed in this study may not be effective for a region where the wavesare mainly swell-dominated.

The proposed method also has specific problems. The first is that the RM mechanism leads to the accumulation of errors in the prediction process, and our proposed replacement mechanism can alleviate this problem to some extent. However, it cannot fundamentally solve this problem.

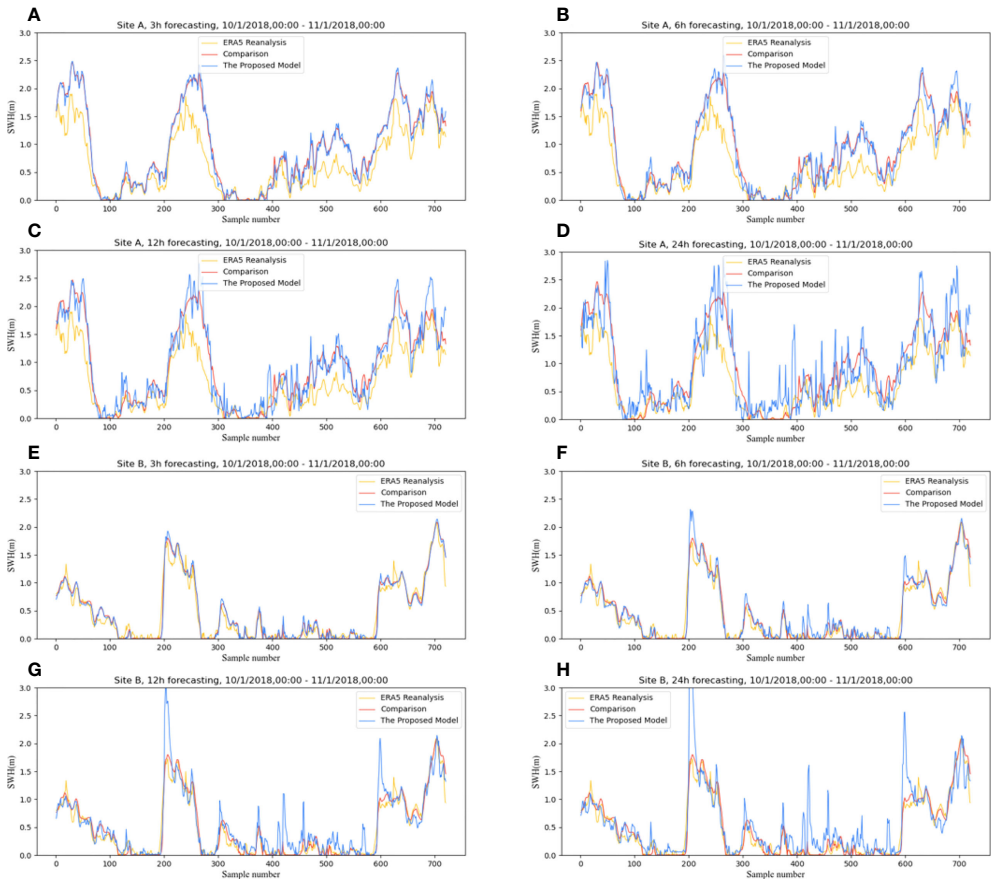


FIGURE 15
24-hour prediction results of significant wave heights for SiteA and SiteB. **(A–D)** respectively represent the 3h, 6h, 12h and 24h forecast results of SiteA, and **(E–H)** respectively represent the 3h, 6h, 12h and 24h forecast results of SiteB.

Secondly, although the proposed model can learn some physical laws of wind speed and significant wave height, it does not really incorporate the dynamic process of the ocean in a sense. How to combine deep learning with the dynamic processes of the ocean is a problem we need to solve in the future. Thirdly, this study is only for the significant wave height, but there are other elements such as wave direction and period. If possible, we will follow up on these elements as well. Finally, we will open up the high spatial and temporal resolution data used in this work. These data will beof great help for subsequent studies.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://doi.org/10.5281/zenodo.6402321.

Author contributions

TS provided overall ideas, writing review and editing, supervision, project administration, and funding acquisition. RH provided concepts, methodology, software, manuscript writing, preparation and research. FM provides formal analysis, data processing, software, and supervision. JW provided data collation and formal analysis. WW provided data management, software. SP provided methodology, method evaluation, data management, project management, and funding acquisition. All authors contributed to the article and approved the submitted version.

Funding

Innovation fund project for graduate students of China University of Petroleum (East China) (No. 22CX04008A).

Project Supported by Key Laboratory of Environmental Change and Natural Disaster of Ministry of Education, Beijing Normal University (Project No. 2022-KF-08).

Acknowledgments

Over the course of my researching and writing this paper, I would like to express my thanks to all those who have helped me. A special acknowledgement should be shown to SP, from whose lectures I benefited greatly, I am particularly indebted to SP who gave me kind encouragement and useful instruction all through my writing. Moreover, I wish to extend my thanks to the library and the electronic reading room for their providing much useful information for my thesis.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Akay, D., and Atak, M. (2007). Grey prediction with rolling mechanism for electricity demand forecasting of turkey. *Energy* 32, 1670–1675. doi: 10.1016/j.energy.2006.11.014

Barnett, T. P. (1968). On the generation, dissipation, and prediction of ocean wind waves. *J. Geophysical Res.* 73, 513–529. doi: 10.1029/jb073i002p00513

Barnett, T., and Kenyon, K. (1975). Recent advances in the study of wind waves. *Rep. Prog. Phys.* 38, 667. doi: 10.1088/0034-4885/38/6/001

Berbić, J., Ocvirk, E., Carević, D., and Lončar, G. (2017). Application of neural networks and support vector machine for significant wave height prediction. *Oceanologia* 59, 331–349. doi: 10.1016/j.oceano.2017.03.007

Bolton, T., and Zanna, L. (2019). Applications of deep learning to ocean data inference and subgrid parameterization. *J. Adv. Modeling Earth Syst.* 11, 376–399. doi: 10.1029/2018ms001472

Booij, N., Ris, R. C., and Holthuijsen, L. H. (1999). A third-generation wave model for coastal regions: 1. model description and validation. *J. geophysical research: Oceans* 104, 7649–7666. doi: 10.1029/98jc02622

Callens, A., Morichon, D., Abadie, S., Delpey, M., and Liquet, B. (2020). Using random forest and gradient boosting trees to improve wave forecast at a specific location. *Appl. Ocean Res.* 104, 102339. doi: 10.1016/j.apor.2020.102339

Cooke, N., Li, T., and Anderson, J. A. (2011). *The tongking gulf through history* (Philadelphia, United States: University of Pennsylvania Press).

Fan, S., Xiao, N., and Dong, S. (2020). A novel model to predict significant wave height based on long short-term memory network. *Ocean Eng.* 205, 107298. doi: 10.1016/j.oceaneng.2020.107298

Fu, Y., Zhou, X., Sun, W., and Tang, Q. (2019). Hybrid model combining empirical mode decomposition, singular spectrum analysis, and least squares for satellite-derived sea-level anomaly prediction. *Int. J. Remote Sens.* 40, 7817–7829. doi: 10.1080/01431161.2019.1606959

Gao, J., Chen, B., and Shi, M. (2015). Summer circulation structure and formation mechanism in the beibu gulf. *Sci. China Earth Sci.* 58, 286–299. doi: 10.1007/s11430-014-4916-2

Gavrikov, A., Krinitsky, M., and Grigorieva, V. (2016). Modification of globwave satellite altimetry database for sea wave field diagnostics. *Oceanology* 56, 301–306. doi: 10.1134/s0001437016020065

Graves, A., and Schmidhuber, J. (2005). Framewise phoneme classification with bidirectional lstm and other neural network architectures. *Neural Networks* 18, 602–610. doi: 10.1016/j.neunet.2005.06.042

Hasselmann, K. (1968). "Weak-interaction theory of ocean waves," in *Basic developments in fluid dynamics* (Cambridge, United States: Academic Press), 117–182. doi: 10.1016/b978-0-12-395520-3.50008-6

Hochreiter, S., Bengio, Y., Frasconi, P., Schmidhuber, J., et al. (2001). *Gradient flow in recurrent nets: the difficulty of learning long-term dependencies*. IEEE 2001, 237–243. doi: 10.1109/9780470544037.ch14. Dataset.

Hochreiter, S., and Schmidhuber, J. (1997). Long short-term memory. *Neural Comput.* 9, 1735–1780. doi: 10.1162/neco.1997.9.8.1735

James, S. C., Zhang, Y., and O'Donncha, F. (2018). A machine learning framework to forecast wave conditions. *Coast. Eng.* 137, 1–10. doi: 10.1016/j.coastaleng.2018.03.004

Jönsson, A., Broman, B., and Rahm, L. (2003). Variations in the baltic sea wave fields. *Ocean Eng.* 30, 107–126. doi: 10.1016/s0029-8018(01)00103-2

Kamranzad, B., Etemad-Shahidi, A., and Kazeminezhad, M. (2011). Wave height forecasting in dayyer, the persian gulf. *Ocean Eng.* 38, 248–255. doi: 10.1016/j.oceaneng.2010.10.004

Kirby, J. T. (1985). "Water wave propagation over uneven bottoms," in *Tech. rep.* (Florida, United States: FLORIDA UNIV GAINESVILLE DEPT OF COASTAL AND OCEANOGRAPHIC ENGINEERING).

Koongolla, J. B., Lin, L., Pan, Y.-F., Yang, C.-P., Sun, D.-R., Liu, S., et al. (2020). Occurrence of microplastics in gastrointestinal tracts and gills of fish from beibu gulf, south china sea. *Environ. pollut.* 258, 113734. doi: 10.1016/j.envpol.2019.113734

Kumar, U., and Jain, V. K. (2010). Time series models (grey-markov, grey model with rolling mechanism and singular spectrum analysis) to forecast energy consumption in india. *Energy* 35, 1709–1716. doi: 10.1016/j.energy.2009.12.021

Li, S., Li, Y., Peng, S., and Qi, Z. (2021). The inter-annual variations of the significant wave height in the western north pacific and south china sea region. *Climate Dynamics* 56, 3065–3080. doi: 10.1007/s00382-021-05636-9

Longuet-Higgins, M. (1963). The generation of capillary waves by steep gravity waves. *J. Fluid Mechanics* 16, 138–159. doi: 10.1017/s0022112063000641

Magdalena Matulka, A., and Redondo, J. M. (2010). Mixing and vorticity structure in stratified oceans. in. *EGU Gen. Assembly Conf. Abstracts.*, 424. Available at: https://ui.adsabs.harvard.edu/abs/2010EGUGA..1215573M

Meng, F., Song, T., Xu, D., Xie, P., and Li, Y. (2021). Forecasting tropical cyclones wave height using bidirectional gated recurrent unit. *Ocean Eng.* 234, 108795. doi: 10.1016/j.oceaneng.2021.108795

Meng, F., Xu, D., and Song, T. (2022). Atdnns: An adaptive time-frequency decomposition neural network-based system for tropical cyclone wave height real-time forecasting. *Future Generation Comput. Syst* 133, 297–306. doi: 10.1016/j.future.2022.03.029

Phillips, O. M. (1957). On the generation of waves by turbulent wind. *J. fluid mechanics* 2, 417–445. doi: 10.1017/s0022112057000233

Shao, W., Sheng, Y., Li, H., Shi, J., Ji, Q., Tan, W., et al. (2018). Analysis of wave distribution simulated by wavewatch-iii model in typhoons passing beibu gulf, china. *Atmosphere* 9, 265. doi: 10.3390/atmos9070265

Shi, X., Chen, Z., Wang, H., Yeung, D.-Y., Wong, W.-K., and Woo, W.-c. (2015). Convolutional lstm network: A machine learning approach for precipitation nowcasting. *Adv. Neural Inf. Process. Syst* 28. doi: 10.48550/arXiv.1506.04214

Song, T., Jiang, J., Li, W., and Xu, D. (2020). A deep learning method with merged lstm neural networks for ssha prediction. *IEEE J. Selected Topics Appl. Earth Observations Remote Sens.* 13, 2853–2860. doi: 10.1109/jstars.2020.2998461

Song, T., Li, Y., Meng, F., Xie, P., and Xu, D. (2022). A novel deep learning model by bigru with attention mechanism for tropical cyclone track prediction in the northwest pacific. *J. Appl. Meteorology Climatology* 61, 3–12. doi: 10.1175/jamc-d-20-0291.1

Sverdrup, H. U., and Munk, W. H. (1947). *Wind, sea and swell: Theory of relations for forecasting* Vol. 601 (Washington, D.C, United States: Hydrographic Office). doi: 10.5962/bhl.title.38751

Tolman, H. L., et al. (2009). User manual and system documentation of wavewatch iii tm version 3.14. *Tech. note MMAB Contribution* 276. Available at: https://polar.ncep.noaa.gov/mmab/papers/tn276/MMAB_276.pdf

Van Aartrijk, M. L., Tagliola, C. P., and Adriaans, P. W. (2002). Ai on the ocean: The robosail project. *In ECAI (Citeseer)* 133, 653–657. Available at: www.robosail.com/research/the_RoboSail_project.pdf

Wei, J., Malanotte-Rizzoli, P., Eltahir, E. A. B., Xue, P., and Xu, D. (2013). Coupling of a regional atmospheric model (regcm3) and a regional oceanic model (fvcom) over the maritime continent. *Climate Dynamics* 43, 1575–1594. doi: 10.1007/s00382-013-1986-3

Zheng, K., Sun, J., Guan, C., and Shao, W. (2016). Analysis of the global swell and wind sea energy distribution using wavewatch iii. *Adv. Meteorology*. 2016 doi: 10.1155/2016/8419580

Zhou, S., Xie, W., Lu, Y., Wang, Y., Zhou, Y., Hui, N., et al. (2021). Convlstm-based wave forecasts in the south and east china seas. *Front. Mar. Sci.* 8. doi: 10.3389/fmars.2021.680079

frontiers | Frontiers in Marine Science

TYPE Original Research
PUBLISHED 06 October 2022
DOI 10.3389/fmars.2022.966621

OPEN ACCESS

EDITED BY
Xi Xiao,
Zhejiang University, China

REVIEWED BY
Patricia Cardoso,
CIIMAR, Portugal
Angeleen Marie Olson,
Hakai Institute, Canada

*CORRESPONDENCE
Yanlong He
yanlhel@163.com

†These authors contributed equally to this work and share first authorship

SPECIALTY SECTION
This article was submitted to Ocean Observation, a section of the journal Frontiers in Marine Science

RECEIVED 11 June 2022
ACCEPTED 16 September 2022
PUBLISHED 06 October 2022

CITATION
He Y, Zhao L, Liu S, Zhao X, Wang Y and Jiang X (2022) Delineation of estuarine ecological corridors using the MaxEnt model to protect marine fishery biodiversity.
Front. Mar. Sci. 9:966621.
doi: 10.3389/fmars.2022.966621

Delineation of estuarine ecological corridors using the MaxEnt model to protect marine fishery biodiversity

Yanlong He[1,2,3*†], Lixia Zhao[1,2,4†], Shouhai Liu[1,2], Xin Zhao[4,5], Yutan Wang[1,2] and Xiaoshan Jiang[1,2]

[1]East China Sea Environmental Monitoring Center, State Oceanic Administration, Shanghai, China, [2]Key Laboratory of Marine ecological monitoring and restoration technologies, The Ministry of Natural Resources, Shanghai, China, [3]Yangtze Delta Estuarine Wetland Ecosystem Observation and Research Station, Ministry of Education & Shanghai Science and Technology Committee, Shanghai, China, [4]Key Laboratory of Ocean Space Resource Management Technology, The Ministry of Natural Resources, Hangzhou, China, [5]Marine Academy of Zhejiang Province, Hangzhou, China

Ecological corridors (ECs) are important management tools to protect biodiversity by linking fragile habitats, especially for highly mobile organisms. ECs in terrestrial landscapes work as passages on land or in water. However, the significance of ECs to migratory species in estuaries has not been well elucidated. Based on annual fishery investigation in the Yangtze estuary and their dominance index rank, three of the top five species, including *Larimochthys polyactis*, *Coilia mystus*, and *Gobiidae*, exhibited absolute dominance in spring during the past 5 years. The temporal and spatial density variance of *C. mystus* supported its short-distance migration pattern. Redundancy analysis and the MaxEnt model predicted optimum habitats for *C. mystus*. *C. mystus* larvae survival was significantly related to salinity, total nitrogen, pH, reactive silicate, dissolved oxygen, surface water temperature, and chlorophyll-a in May and to salinity, surface water temperature, permanganate index, suspended particles, total nitrogen, and total phosphorus in August. The MaxEnt model predicted a broader longitudinal distribution range from offshore to the upstream freshwater area but narrower latitudinal distribution in the southern branch in May than in August. Finally, we delineated migratory corridors connecting optimum habitats for *C. mystus* using the least-cost route method. Optimum habitats close to the coastlines in the south branch might play a significant role in maintaining population or community connectivity in the Yangtze estuary. Our findings provide a perspective and method to quantify and facilitate the harmonious development of socioeconomy and fishery biodiversity conservation.

KEYWORDS

ecological corridors, biodiversity, Yangtze Estuary, *Coilia mystus*, MaxEnt model, fish conservation

Abbreviations: ECs, ecological corridors; YRE, Yangtze River Estuary; SS, suspended particles; DO, dissolved oxygen; N, total nitrogen; P, total phosphorus; Si, reactive silicate; Temp, estuarine surface water temperature; COD_{Mn}, permanganate index; Chla, chlorophyll-a; depth, water column depth; RDA, redundancy analysis.

Introduction

Ecological corridors (ECs) are landscapes linking spatially isolated and dispersed patches in a linear or strip-like distribution, which are constructed for the movement of individuals such as tigers (Seidensticker et al., 2010), fishes (Benson et al., 2007), birds (Klaasen, 1996; Tong et al., 2012), and butterflies (Runge et al., 2015). They also serve as a buffer with certain similarities to habitual kernels benefiting the recovery and adaptation of native species (Lawler et al., 2013; Degteva et al., 2015; McGuire et al., 2016; Keeley et al., 2018). Such landscapes play important roles in sustaining and restoring biodiversity (Hilty et al., 2019) by promoting the flow of energy and materials and genetic exchange in terrestrial environments (Miklos et al., 2019), freshwater systems (Bastian et al., 2015; Hauer et al., 2016), and marine environments (Gillanders et al., 2003; Cowen and Sponaugle, 2009), or a transboundary of the above-mentioned environments. The methods for investigating ECs (Weeks, 2017; Balbar and Metaxas, 2019) have grown steadily during the past several years. Subsequently, more tools are available to project and map ECs in systematic conservation planning, such as the circuit theory (McRae, 2006), graph theory (Urban and Keitt, 2001), least-cost route method (McRae et al., 2016), and individual-based model (Allen et al., 2016). In general, the following aspects are crucial for modeling ECs: the movement behavior of objectives, the physical conditions of their habitats, the degree of disturbances or pressures, and the persistence and adaptation of objectives. To date, global actions have been taken to sustain ecological connectivity by constructing marine protected area networks (Carr et al., 2017). Most of them are based on specific investigating datasets about objectives; for example, specific megafauna or birds migrating across varied jurisdictions, entire passages, or rivers for anadromous or catadromous fish species. However, research on defining the importance of ECs in fishery biodiversity by projecting the dispersion and distribution of model species in estuaries is lacking.

Estuarine ecosystems are located at the ecological intersection between terrestrial and marine ecosystems. They are socioeconomically and ecologically important benefiting from the economic, demographic, natural resource, transportation, and other advantages (Levin et al., 2001; Cowen and Sponaugle, 2009). In recent years, estuarine ecosystems have been considerably altered by global climate change (Gillanders et al., 2011; Scanes et al., 2020) and intense anthropogenic activities (Wang et al., 2022), such as important habitat reclamation, upstream dam construction (Ferguson et al., 2011), overfishing since the 1990s (Wang et al., 2022), spawning and nursing habitat deterioration, and migratory corridor breakdown (Cowen and Sponaugle, 2009). The coupled effects of natural and human activities have altered the temperature, salinity, pH, and nutrients as well as the heterogeneity of seawater characteristics (depending on the mixing process of fresh and salty water) (Pelage et al., 2021). Additionally, it has adversely accelerated the mixing process and promoted the fragmentation of the physical and chemical properties of water columns (Wu and Wu, 2018). Such variance has driven changes in the physiology, morphology, and genetics of marine organisms (Finn, 2007; Johnston and Roberts, 2009; Garcia and Barria de Cao, 2018). For example, ammonia nitrogen and cadmium could be toxic to fish and may decrease their survival rates and biodiversity (Schram et al., 2014). Ocean acidification has changed the preference of catadromous barramundi for tropical estuarine mangroves (Rossi et al., 2018). Spatial heterogeneity in salinity, temperature, and turbidity gradients have increased in estuaries, simultaneously leading to changes in preferred habitats for some marine organisms or different life history stages of the same marine organism. In past decades, some fishes that are highly dependent on estuarine habitats have been decimated; for example, tapertail anchovies (*Coilia nasus*) migrates from the sea to the river (Jiang et al., 2014), the Yangtze River dolphin migrates in the descending river (Zhu et al., 2009), and Chinese sturgeon (Wang et al., 2018) larvae require the brackish water of the estuary to balance body fluid osmolality.

Migration is a characteristic developed by fishes in the process of phylogeny and a long-term adaptation of fish to the environment, which enables the population to obtain more favorable conditions for survival and reproduction (Planes et al., 2009; Christie et al., 2010; Sykes and Shrimpton, 2010). Anadromous fishes live in the ocean but spawn in the middle or upstream of rivers. In contrast, catadromous fishes live in freshwater and spawn in the ocean. They have strict adaptations to the ecological conditions of their habitat (Hanson et al., 2010; Wang et al., 2012; Wang et al., 2018; Chaparro-Pedraza and de Roos, 2019). Eggs and larvae play an important role in energy transfer in the marine ecosystem, and their survival rate and the amount of remaining stock determine the abundance of population replenishment resources (Botsford et al., 2009; Saenz-Agudelo et al., 2010; D'Aloia et al., 2015). Hence, eggs and larvae are indispensable elements in the study of sustainable use of fish resources and important indicators for evaluating fish migration (Almany et al., 2017). In recent years, the eutrophication of estuarine waters and increase in pollutant discharges have led to a significant decrease in the suitability of estuarine fish habitats. The larval survival rate and species replenishment have continued to decrease, especially for the migratory fishes. Hence, it is essential to identify the migratory pathways of key species in estuaries and to protect marine biodiversity by maintaining the connectivity of organisms and their habitats. In this study, we aimed to (1) construct a theoretical framework for identifying key migratory species in estuaries; (2) test the suitability of the MaxEnt model to predict optimal habitats and construct resistance surfaces for migratory fish

species; and (3) delineate or identify ECs as the minimum management space to maintain population connectivity.

Materials and methods

Study site and ichthyoplankton sampling

The Yangtze River Estuary (YRE) is the largest estuary in China, with three-order bifurcations and four outlets for significant volumes of freshwater discharge from the Yangtze River to the East China Sea (Figure 1A). The annual average water temperature is 15°C–16°C. The YRE is a medium-intensity tidal estuary with regular half-day tides outside the mouth and informal half-day shallow tides inside the mouth. From offshore to upstream of the YRE, the duration of flood tide decreases gradually, whereas that of the ebb tide increases and is longer than that of the flood tide. The tidal range of the southern branch of the YRE decreases upstream. The YRE is located in a subtropical monsoon climate zone with oceanic and monsoon characteristics and an annual average precipitation of approximately 1,030 mm, which is mostly concentrated in June and September. A large amount of freshwater carries sediment discharge into the sea, which influences the water quality, hydrologic dynamics, and topography (Wu et al., 2011; Wu et al., 2013). Furthermore, socioeconomic development, including fishing; shipping; and corresponding channel maintenance, pollutant emission, and coastal reclamation, has led to marine ecological degradation and biodiversity extinction through a trophic cascade or habitat destruction.

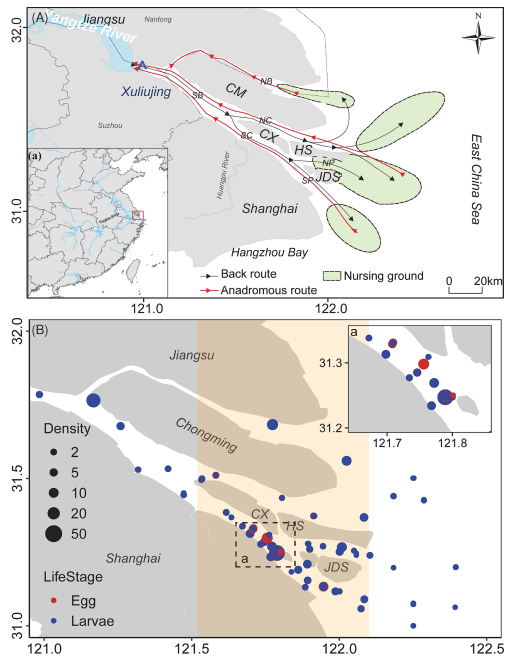


FIGURE 1
Coilia mystus community at Yangtze estuary. **(A)** Migration pattern of *C mystus*; **(B)** spatial density distribution of *C mystus* eggs and larvae. Unit of density: ind./m^3. NB, north branch; SB, south branch; NC, north channel; SC, south channel; NP, north passage; SP, south passage; CM, Chongming Island; CX, Changxing Island; HS, Hengsha Island; JDS, Jiuduansha Island.

Fish species were surveyed once in spring and summer from 2017 to 2021 during the East China Sea monitoring campaign. In our study area, 65 sites were investigated during each campaign (Figure S1). Ichthyoplankton samples were collected using two kinds of beam trawl: I-trawl and a large plankton trawl. I-trawl is the shallow water type, with a 145 cm length, 50 cm horizontal opening, 0.2 m^2 net mouth area, and 0.505 mm screen silk aperture. The I-trawl was trawled once from the bottom to the surface layer at each site. The larvae density of each site was equal to the number of larvae in the sample divided by the volume of the water column, which was defined by the water depth and net mouth area. The large plankton trawl had a net with 280 cm length, 80 cm horizontal opening, 0.5 m^2 mouth area, and 0.505 mm screen silk aperture. Larvae density was equal to the number of larvae in the sample divided by the swept area, which was defined by boat speed, trawling time, and the net mouth area. The collected samples were fixed and preserved with a 37% formaldehyde solution. The species were identified and counted using a stereoscopic microscope (Nikon SMZ25, Japan) based on the external morphology of fish eggs and larvae and the individual morphology, tissues, organs, and phylogenetic characteristics at different development stages. The dominance of the main fish species (Y, Li et al., 2022) was calculated using the following equation: $Y=n_i/N\times f_i$ where n_i represents the number of individuals of species i, f_i is the frequency of species i in each site, and N is the total number of individuals hauled in the campaign. From the five most dominant fish species, only the short-distance migratory species, anchovy (*Coilia mystus*), was selected for the following analysis.

C. mystus is a short-distance anadromous fish that mainly inhabits coastal shallow water or offshore. Its migratory behavior has been inferred from otolith microchemical analysis (Yang et al., 2006a; Yang et al., 2006b; Jiang et al., 2012; Jiang et al., 2014; Yang et al., 2019; Vu et al., 2022). In general, the breeding period of *C. mystus* in the Yangtze estuary is from April to September, including the primary spawning season of May and the minor spawning season of August (Hu et al., 2021). In late April, a small number of sexually mature fishes migrate upstream to the brackish water or freshwater areas of the Yangtze estuary to spawn, and some migrate to Zhenjiang, Jiangsu (Zeng and Dong, 1993). The surviving larvae grow in the deep water around Chongming Island and return to the sea in winter (Zeng and Dong, 1993; Yang et al., 2019). Studies on the feeding habits of *C. mystus* indicated that they mainly fed on crustacean plankton, including Copepods, Mysids, and Decapoda, especially *Acanthomysis longirostris*, *Schmackeria poplesia*, *Tortanus vermiculus*, and *Labidocera euchaeta* (Liu and Xu, 2011). Compared with long-distance migratory species, the growth and reproduction

of *C. mystus* are more dependent on the marine environment of the Yangtze estuary. Hence, we used *C. mystus* as the study object to reference the biodiversity conservation of native resident species and short-distance migratory species.

We monitored the environmental factors influencing the community assemblage or distribution of marine fish species. Environmental factors included the concentration of suspended particles (SS), dissolved oxygen (DO), total nitrogen (N), total phosphorus (P), reactive silicate (Si), salinity, estuarine surface water temperature (Temp), pH, permanganate index (COD_{Mn}), chlorophyll-a (Chla), and water column depth (depth). Salinity, water temperature, and depth were measured *in situ* using a conductivity–temperature–depth (CTD) automatic recording analyzer (SBE25plus, Sea-Bird Electronics, Inc. USA). Seawater samples were collected in 5 L bottles from 0.5 m below the surface and at the bottom of each site to measure the other eight physicochemical factors. DO was measured using the iodometric titration method. The water nutrients, including P, Si, and N, were analyzed using an automatic continuous flow analyzer (QuAAtro, SEAL Analytical GmbH, Germany). pH was monitored by a pH meter (JENCO 6010M, JENCO International, Inc. USA). COD_{Mn} was measured using the alkaline permanganate method, and Chla was measured using fluorescent spectrophotometry (Agilent Cary Eclipse, Agilent Technologies, Inc. USA). SS samples were filtered by preweighted Whatman GF/F microfiber filters (25 mm) (Gao et al., 2019) and weighted by an electronic scale (QUINTIX224-1CN, Sartorius Aktiengesellschaft, Germany); its concentration was determined using the SS weight divided by the sampling volume.

Predicting potential optimal habitats of migratory species

Maximum entropy (MaxEnt) is a program for modeling species niches and spatial distributions from presence-only species records (Elith et al., 2010). From a set of environmental data and known species occurrence, the model expresses a probability distribution where given grids predict the suitability of conditions for the species by comparing the conditions between present locations and the study area. This model has been widely used in predicting the likely distribution of species in terrestrial (Lamb et al., 2008; Yates et al., 2010) and marine ecosystems (Tittensor et al., 2009; Verbruggen et al., 2009). Considering the complex habitat biogeographical features of *C. mystus*, the MaxEnt model may be an effective approach to predict their preferred habitats. MaxEnt analysis was performed using the open-source Java software MaxEnt 3.4.1 (Phillips et al., 2006; Phillips et al., 2017). We selected 75% of fishery survey datasets as the training samples to establish the prediction model and the remaining 25% as the test samples to verify the model. Two performance metrics were chosen: (1) the receiver operating characteristic curve, which evaluates the trade-off between prediction sensitivity and specificity and (2) the associated area under the curve, which presents the values of approximately 0.5 representing a random prediction, and values closer to 1 indicate a better performance of the model (Mantas et al., 2022). The analysis was repeated 10 times and averaged to obtain the final simulation results. The prediction results were imported into ArcGIS 10.2 (ESRI, Redlands, California, USA) and converted to a raster format to generate distribution probability prediction maps. Here, we defined the areas with species distribution probability greater than 90% as optimal habitats.

Optimal habitats located at the Yangtze estuarine mouth or offshore were defined as source patches, and those located at upstream freshwater or brackish water area were defined as sink patches. These definitions refer to the two possible behavior patterns of *C. mystus*. First, *C. mystus* adults migrate from offshore source habitats to upstream sink habitats to spawn, where the source and sink patches correspond to the feeding ground and spawning ground, respectively. Second, *C. mystus* egg and larvae have little motor ability, and they drift back and forth with flood and ebb tides between offshore and freshwater areas. Here, source and sink patches represent the origin and destination areas of larvae drifting with the flood tide, respectively.

Delineation of ecological corridors

Ecological resistance refers to the biotic and abiotic factors in a recipient ecosystem that limits population growth (D'Antonio and Thomsen, 2004). Ecological processes such as animals crossing the landscape carry the property of overcoming spatial resistance, which can be measured in terms of accessibility, cost distance, and minimum cumulative resistance. All these measures of resistance can be considered the extensions of the spatial distance concept. In this study, a graph theory algorithm was used to extract animal migration corridors (Pinto and Keitt, 2009). First, the relationship between the model species and the environmental elements was tested using the canonical correspondence analysis, and the explanatory rate values of the environmental factors in the habitats were analyzed. We selected salinity, DO, and surface water temperature with phosphorate concentrations in May and Chla in August to construct a resistance surface. The weights of selected explanatory environmental factors were calculated based on $y=(a+b)/2$, where a was the predicted contribution value of an environmental factor and b was their explanatory rate in canonical correspondence analysis. Second, kriging interpolation was used to rasterize the point environmental datasets to obtain the resistance surface.

Based on the rasterized resistance surface, the path of cumulative minimum cost distance from the source to the sink patches was generated. The gravity center of source and sink patches were extracted using conventional tools. The cost distance tool was used to calculate the cumulative cost distance from each pixel to the specified source. The cost backlink tool was used to define the direction in which any

pixel extends the minimum cost path to the specified source. The cost path tool was used to generate the path with the smallest cumulative cost distance. Finally, 500-m-wide buffer zones were constructed for the theoretical least-cost route to eliminate the spatial redundancy of corridors delineated by the least-cost route method.

Statistical analysis

The nonparametric Kruskal–Wallis test followed by a *post hoc* comparison was used to analyze the temporal and spatial differences between the average values of life-history characteristics (density, abundance, and maximum and minimum body length) of *C. mystus* larvae and eggs. Redundancy analysis (RDA) was performed to assess the linear relations between multiple dependent and independent variables in a matrix, which was then incorporated into principal component analysis (Xia, 2020). We used the RDA of life-history characteristics on the environmental factors to extract key factors explaining variations in the investigated sites. This statistical analysis was performed using the statistical software package R (R Core Team, 2014).

Results

Dynamics of fish communities in the Yangtze River Estuary

The YRE is an important spawning and nursery ecosystem for some species. According to the fishery survey of the YRE in May during 2017–2021, this area was dominated by estuarine and marine fishes until 2020; however, freshwater fishes of the family Cyprinidae reached the top five in 2021 (Table 1). Among them, marine fish *Larimochthys polyactis*, migratory fish *C. mystus*, and estuarine fish *Gobiidae* exhibited absolute dominance in spring during the past 5 years. *C. mystus* larvae or eggs were detected in the Xuliujing hydrological station located upstream of the YRE. Nonetheless, the larvae density around Changxing Island and Jiuduansha Shoal was 2–10 folds higher than offshore and upstream freshwater areas (Figure 1B). Hence, we speculated that *C. mystus* inhabits coastal or offshore areas and migrates upstream to freshwater or brackish water areas during the breeding season. Their larvae grow around Chongming Island and then migrate back to the sea (Figure 1A). *C. mystus* has two breeding periods: May is the primary reproductive period, and August is secondary. These life-history characteristics were proven by larger larvae density in May (Wilcoxon test, $P < 0.001$) than in August. In particular, the density of *C. mystus* larvae increased significantly from 2020 to 2021 (Figure 2, right panel; Wilcoxon test, $P < 0.001$), which may benefit fishery production.

Linking life-history characteristics of *C. mystus* with environmental variables

RDA was used to correlate the life-history characteristics of *C. mystus* with environmental variables. A total of 11 explanatory variables were considered in this study, including five hydrological variables: Temp, pH, salinity, depth, and DO and six chemical factors: Chla, COD_{Mn}, and the contents of N, P, Si, and SS. Forward selection revealed that the correlation between key factors and the life-history characteristics of *C. mystus* in two breeding seasons varied profoundly (Table S1). In May, 7 of the 11 variables were significant in RDA: salinity, N, pH, Si, DO, surface water temperature, and Chla. However, in August, six variables were significant: salinity, Temp, COD_{Mn}, SS, N, and P.

The first and second axes of the RDA plot demonstrated 39.9% and 7.8% variations in the life-history characteristics of *C. mystus* larvae in May, respectively (Figure 3A; Table S2). The life-history

TABLE 1 Dominant species based on the offshore monitoring voyage in May from 2017 to 2021.

Species	Y-dominance index					Ecotypes
	2017	2018	2019	2020	2021	
Engraulis japonicus	0.0748	0.0119	0.0911	0.0084	/	Marine
Larimochthys polyatis	0.0403	0.0307	0.012	0.0283	0.0023	Marine
Coilia mystus	0.0126	0.0034	0.0109	0.1215	0.0296	Short-distance migration
Gobiidae	0.0006	0.0245	0.0568	0.0137	0.0393	Estuarine
Stolephorus	0.0005	/	/	/	/	Marine
Liza haematocheila	/	0.0238	/	/	/	Estuarine
Callionymidae	/	/	0.0018	/	/	Marine
Cynoglossus sp.	/	/	/	0.0062	/	Estuarine
Cyprinidae	/	/	/	/	0.0122	Freshwater
Cynoglossus semilaevis	/	/	/	/	0.0038	Estuarine

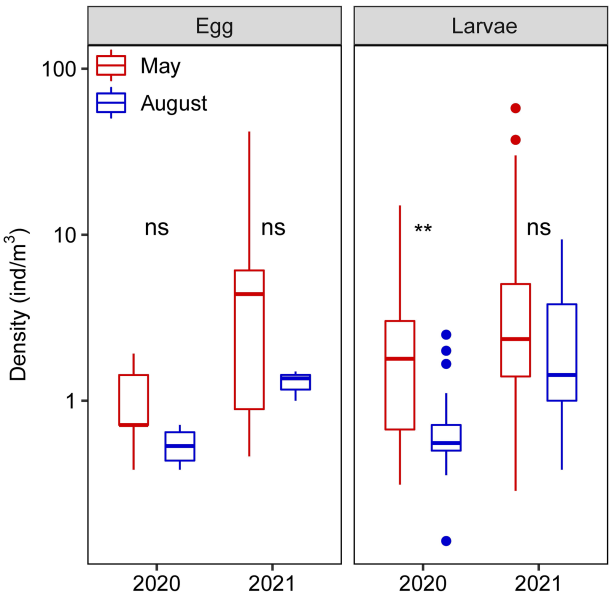


FIGURE 2
Mean density of *C. mystus* eggs and larvae in May and August of 2020 and 2021 at the Yangtze estuary. 'ns' indicates no significance between compared groups; ** indicates $P < 0.01$.

characteristics of *C. mystus* were positively correlated with Chla, N, Si, P, DO, and surface water temperature, whereas pH and salinity negatively contributed to *C. mystus* growth. In August, the first and second axes of the RDA plot demonstrated 17.2% and 9.2% variations in the life-history characteristics of *C. mystus* larvae, respectively (Figure 3B; Table S2). Three chemical factors, P, N, and COD_{Mn}, and two physical factors, Temp and SS, were positively correlated with *C. mystus* growth. Moreover, *C. mystus* growth was negatively correlated with salinity in May.

Optimal habitats of *C. mystus* and its migratory corridors in the Yangtze River Estuary

The average area under the curve values of the model training datasets were 0.983 and 0.971 in May and August, respectively (Figures 4A, B), indicating an accurate prediction of the optimal occurrence of *C. mystus* produced by the MaxEnt model. The predicted *C. mystus* occurrence hotspots with suitability larger than 0.90 were widely spread in the YRE, where the total area of the hotspots reached 89.9 and 271.2 km^2, accounting for 0.41% and 1.53% of the study area in May and August, respectively (Figures 4C, D; Table S3). In May, the predicted occurrence hotspots were mainly concentrated in the south branch (Figure 4C); however, they were almost homogeneously distributed in the south and north branch in August (Figure 4D).

The resistance values were 104.73–165.69 in May (Figure 5A) and 47.49–95.93 in August (Figure 5B), generally decreasing from offshore to upstream of the Yangtze River. ECs delineated in this study showed high spatial variances between May and August. Specifically, almost all ECs were concentrated in the south channel, especially around Jiuduansha Shoal (Figure 5A); however, ECs were almost homogeneously spread in the YRE in August (Figure 5B). Furthermore, most ECs were distributed along the coastlines of Changxing Island, Jiuduansha Shoal, and Hengsha Island. The main nodes among EC interactions were accumulated at western Changxing and Chongming islands and around Jiuduansha Shoal, indicating potential conservation priority for YRE fishery biodiversity protection.

Discussion

Life-history characteristics of *C. mystus* and its fitness to environmental forcing

The life-history characteristics of species are a result of natural evolution and ecological adaptation. The documented spawning ground of *C. mystus* was in the maximum turbidity zone in the YRE (Hu et al., 2021). Our study extended the spawning ground to the Xuliujing hydrologic station (Figure 1B). This extension may indicate two characteristics of short-distance migratory species. First, adults migrating to the upstream freshwater area were an active selection. Runoff is an unneglectable factor controlling larvae

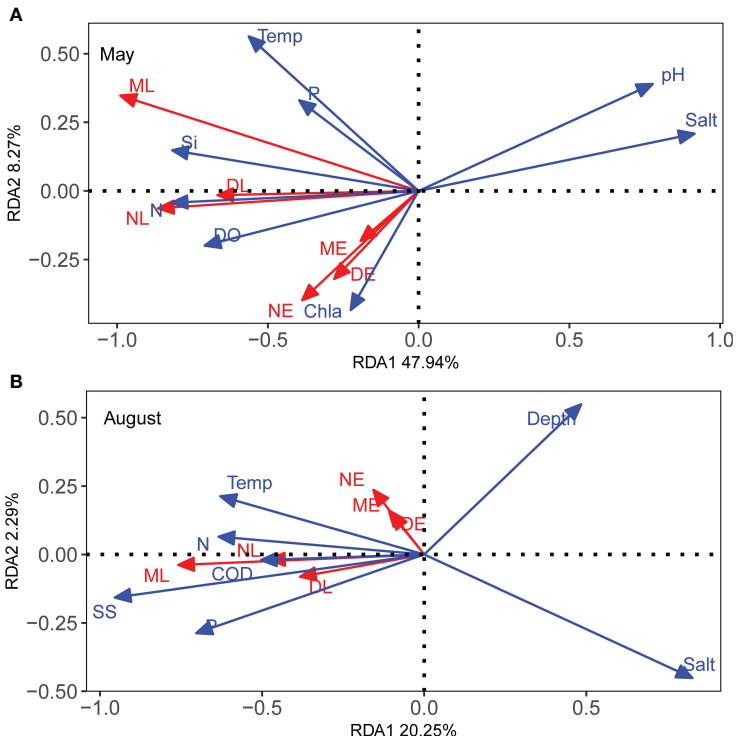


FIGURE 3
Redundancy analysis of the effect of environmental variables (blue arrows) on the life-history characteristics (red arrows) of *C mystus* in May **(A)** and August **(B)**. The response variables include the density, abundance, and mean body length of *C mystus* larvae (*in abbr.* DL, NL, and ML, respectively) and egg (*in abbr.* DE, NE, and ME, respectively). Total variation explained by the redundancy analysis model was 56.21% **(A)** and 22.54% **(B)**. Si, reactive silicate; N, total nitrogen; P, total phosphorus; SS, suspended particles; Salt, salinity; Temp, temperature of surface water.

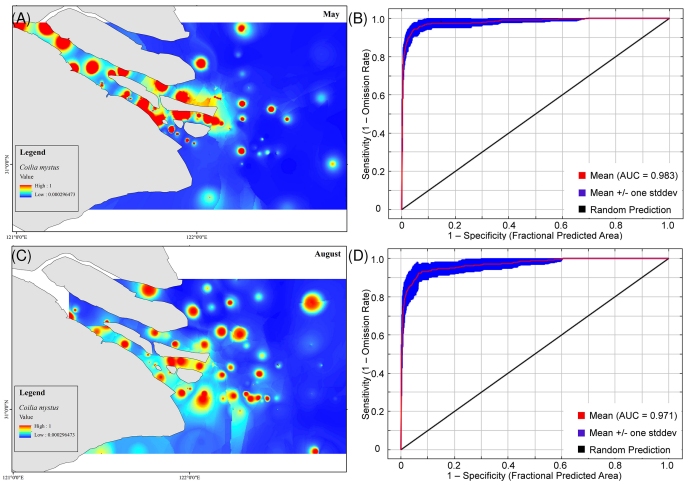


FIGURE 4
Predicted optimal distribution of *C mystus* through the MaxEnt model. **(A, C)** Potential geographical distribution of *C mystus* in Yangtze Estuarine during the May **(A)** and August **(C)** breeding season. **(B, D)** Area under the curve result of MaxEnt modeling (10 runs) for May **(B)** and August **(D)**.

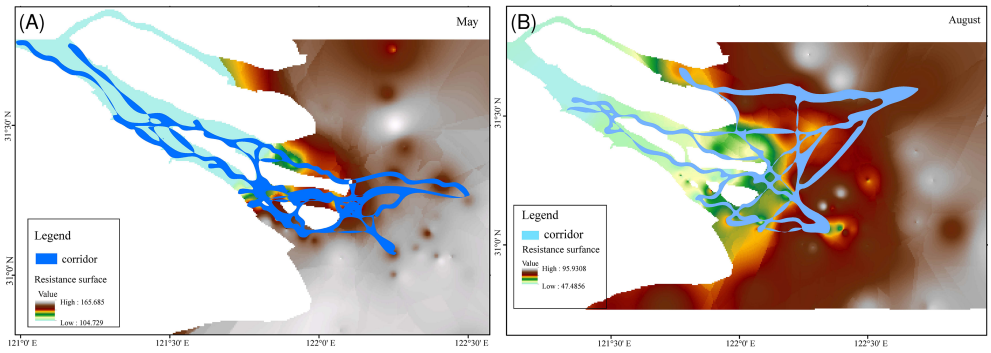


FIGURE 5
Theoretical migratory corridors of *C mystus* based on the least-cost route analysis in May **(A)** and August **(B)**.

drift. If the spawning ground is close to the estuary front, egg and larvae might be carried into the adjacent sea by the tide, where it will be difficult for them to survive because decreased sea water temperature might inhibit the hatching of eggs, and most larvae may become food sources for other fishes. Therefore, some adults would migrate upstream to freshwater, with weak tide and low salinity. Nonetheless, physiological characteristics may be the key factor controlling the migratory distance of *C. mystus*. Compared with long-distance migratory species *C. nasus*, *C. mystus* allocates most of the body lipid for gonadal development and rapidly consumes trunk lipid (Wu et al., 2017). Moreover, when migrating across a large salinity gradient, *C. mystus* needs to evolve the ability to rapidly regulate the equilibration of fluid osmotic pressure, potentially raising the cost of population maintenance. Second, some *C. mystus* adults spawn in the estuary front area, and their larvae drift upstream with tides. In this case, some unknown physiological characteristics may be developed to adapt the ebb and flow tide to sustain their spatial distribution. Some species in the Penaeus family could sense the increase or decrease in water pressure at a depth of 8 m to conduct directional migration. During a flood tide, they enter the estuary along the tide but swim to the bottom to avoid being carried out of the estuary during an ebb tide (Vance and Pendrey, 2008; Wolanski, 2017). Constricted by traditional investigation approaches and the natural characteristics of most marine fish species, providing direct evidence about their migration patterns, life-history characteristics, or adaptative evolution is challenging. Advancements in animal tracking and environmental DNA technology enable greater data collection on migratory behavior and environmental suitability assessment (Xu et al., 2018; Thalinger et al., 2019; Yatsuyanagi and Araki, 2020). Collecting more data using advanced technologies to protect more extinct or dominant species is necessary.

Based on RDA, salinity, COD_{Mn}, SS, N, P, pH, Si, DO, Temp, and Chla may considerably affect the breeding and migratory behavior of *C. mystus*. Among them, physical factors such as salinity, Temp, SS, and DO are correlated with the survival or growth of marine species (Spares et al., 2012; Walsh et al., 2013); however, chemical factors such as Chla, N, P, Si, and COD_{Mn} have not been studied adequately. Higher N, P, and Si promote algal growth, and increased Chla indicates high algal bioactivity. Algae produce large amounts of oxygen through photosynthesis and increase the DO content of the water column. As primary producers, the healthy growth of phytoplankton provides the material basis for fish growth and reproduction. Hence, detecting a positive relationship between Chla, N, P, Si, and *C. mystus* growth is reasonable.

The optimum environmental conditions for *C. mystus* larvae were as follows: salinity, 5‰–12‰; temperature, 20°C–28°C; and DO concentration, 5.2–8.0 mg/L (Hu et al., 2021), indicating that any event causing environmental variables beyond the range of biological adaptation may lead to biodiversity reduction. For example, saltwater intrusion can carry sexually mature *C. mystus* to freshwater through the north branch or to brackish water through the south branch. If the upstream freshwater runoff is low, the tide upwells, and brackish water expands upstream, providing a broad survival space for *C. mystus* larvae; such causality may account for the broader longitudinal distribution of the larvae in May than in August. However, the natural ecosystem is more complex because of the non-linear relationship among all biotic and abiotic factors for the fitness of organisms. Hence, it is necessary to study their fitness to the dynamics of certain or coupled variables.

Marine fishery biodiversity protection through ecological corridors

The YRE is rich in fish biodiversity, supporting 46 freshwater fishes, 53 estuarine sedentary and migratory fishes, and 48 marine fishes (Zhuang et al., 2006). Life-history characteristics revealed that the dominant species varied from season to season, which included *Engraulis japonicus*, *C. mystus*, *Anchoviella commersonii*,

Gobiidae, and *Liza haematocheilus* in spring; *C. nasus*, *C. mystus*, and *Stolephorus chinensis* in summer; *Salanx ariakensis*, and *Benthosema pterotum* in autumn; and *Neosalanx jordani* in winter. Most of these dominant species act as secondary consumers relying on organic detritus or primary consumers (including zooplankton such as copepods, branchiopods, and benthic invertebrates). To date, only a few studies have reported the interspecific competition in the fish community; hence, we speculate that anthropogenic and environmental stressors that directly affect food and habitat quality severely influence the marine fish biodiversity of the YRE.

We investigated the 17 larvae of marine, estuarine, freshwater, and migratory fishes in the YRE during the past few years (Table 1). The spatial distribution of larvae did not match that of adults, which is consistent with the functional orientation of the YRE as migratory passage, spawning, nursing, and feeding grounds. Most larvae species distributed in the maximum turbidity zone developed under the coupled effects of freshwater discharge, tides and currents, the mixture of freshwater and saltwater, and suspended sediment transportation (Shen et al., 1992). Marine fishes use the YRE as a nursing and feeding ground because the convergence of various water masses brings in rich nutrients, and long-distance migratory species need to regulate the equilibration of fluid osmotic pressure here. Some freshwater species, such as those from the Cyprinidae family, can be detected due to complex hydrodynamics here. For example, freshwater discharge from Huangpu River (Figure 1A) southward into the sea along the coastal waters of Shanghai forms a fresh flume. Sometimes, saltwater intrusion might increase salinity in waters near Changxing Island and western Chongming Island. Considering the diversity in the marine species community, their habitat utilization, and the complexity of the marine environment, our study on the short-distance migratory *C. mystus* may provide a breakthrough for researching multispecies habitat utilization and conservation.

C. mystus, a typical short-distance migratory fish in the YRE, mainly lives in the nearshore area. *C. mystus* migration began during the spawning period from the nearshore sea area with 25‰ salinity to the estuarine front area with approximately 4% salinity. Some adults migrated further anadromous to the freshwater. Fish larvae were mainly distributed near Jiuduansha Shoal (Figure 1A), which displays a mixture of brackish and fresh waters with the most intense tidal influence. Affected by tides and runoff, the distribution of *C. mystus* eggs in May was closer to the upstream of the YRE, whereas that in August was closer to the periphery with 4‰–10‰ salinity. In general, migration is an important adaptive evolutionary strategy for fishes to maintain population genetic diversity. Compared with fishes in uniform habitats, migratory fishes need to span different habitats. Highly suitable patches, also called source and sink patches in our study, are important nodes and ecological stepping stones for their migration. The quantity, quality, and connectivity of source and sink patches directly determine whether the species can successfully migrate to their destinations. This study further showed that the estuarine front area was an important ecological aid for anadromous migration, with a larger number of nodes and density of ECs in May and August in this area.

In conclusion, in this study, three priority areas were identified, including the maximum turbidity zone located at the estuarine mouth and two brackish water confluence areas located around Changxing Island and western Chongming Island (Figure 1B). Among these priority areas, *C. mystus* might prefer migrating through the southern branch, especially the south channel in the May breeding season (Figure 5A), but without any bias in August (Figure 5B). Furthermore, most theoretical corridors are distributed along the coastlines except those in the offshore area, which might infer the more important role of islands in fishery production and biodiversity protection. Aside from that, theoretical ECs also ensure the connectivity of some important nature reserves, such as the *C. nasus* germplasm resource protection area at western Chongming Island, YRE Chinese sturgeon Nature Reserve to eastern Chongming Island, and some important wetlands located at Chongming Island, Jiuduansha Shoal, and Nanhui Shoal. ECs also provide an accessible approach to fishery management and biodiversity conservation. On one hand, theoretical ECs provide a least-cost movement route for fish migration and inhabitation, indicating that migratory species can sustain their population connectivity at the least cost. On the other hand, from a social management perspective, imposing a ban on an entire estuary is unnecessary, and differential measures must be taken in corridors and priority areas to facilitate the harmonious development of socioeconomic and biodiversity conservation, especially for the high-traffic YRE.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

YH, LZ, XZ, and XJ designed research; SL and YW contributed field samples collection and species classification; YH and LZ analyzed data and wrote the paper; All authors contributed to discussion of the content, and reviewed the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the Yangtze Delta Estuarine Wetland Ecosystem Observation and Research Station,

Ministry of Education & Shanghai Science and Technology Committee (ECNU-YDEWS-2020), National Natural Science Foundation for Youth Scientists of China (NO.42106171), Open Fund of Key Laboratory of Marine Ecological Monitoring and Restoration Technology, MNR (NO. MEMRT202111) and Open Fund of Key Laboratory of Ocean Space Resource Management Technology, MNR (KF-2021-107).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmars.2022.966621/full#supplementary-material

References

Allen, C. H., Parrott, L., and Kyle, C. (2016). An individual-based modelling approach to estimate landscape connectivity for bighorn sheep (Ovis canadensis). *PeerJ* 4, e2001. doi: 10.7717/peerj.2001

Almany, G. R., Planes, S., Thorrold, S. R., Berumen, M. L., Bode, M., Saenz-Agudelo, P., et al. (2017). Larval fish dispersal in a coral-reef seascape. *Nat. Ecol. Evol.* 1 (6), 148. doi: 10.1038/s41559-017-0148

Balbar, A. C., and Metaxas, A. (2019). The current application of ecological connectivity in the design of marine protected areas. *Glob. Ecol. Conserv.* 17, e00569. doi: 10.1016/j.gecco.2019.e00569

Bastian, O., Grunewald, K., and Khoroshev, A. V. (2015). The significance of geosystem and landscape concepts for the assessment of ecosystem services: exemplified on a case study in Russia. *Landscape Ecol.* 30, 1145–1164. doi: 10.1007/s10980-015-0200-x

Benson, R. L., Turo, S., and McCovey, J. B.W. (2007). Migration and movement patterns of green sturgeon (Acipenser medirostris) in the klamath and trinity rivers, California, USA. *Environ. Biol. Fish.* 79, 269–279. doi: 10.1007/s10641-006-9023-6

Botsford, L. W., Brumbaugh, D. R., Grimes, C. B., Kellner, J. B., Largier, J., Michael, R., et al. (2009). Connectivity, sustainability, and yield: bridging the gap between conventional fisheries management and marine protected areas. *Rev. Fish. Biol. Fish.* 19, 69–95. doi: 10.1007/s11160-008-9092-z

Carr, M., Robinson, S. P., Wahle, C., Davis, G., Kroll, S., Murray, S., et al. (2017). The central importance of ecological spatial connectivity to effective coastal marine protected areas and to meeting the challenges of climate change in the marine environment. *Aquat. Conserv.* 27, 6–29. doi: 10.1002/aqc.2800

Chaparro-Pedraza, P. C., and de Roos, A. M. (2019). Environmental change effects on life-history traits and population dynamics of anadromous fishes. *J. Anim. Ecol.* 88 (8), 1178–1190. doi: 10.1111/1365-2656.13010

Christie, M. R., Tissot, B. N., Albins, M. A., Beets, J. P., Jia, Y., Ortiz, D. M., et al. (2010). Larval connectivity in an effective network of marine protected areas. *PloS One* 5 (12), e15715. doi: 10.1371/journal.pone.0015715

Cowen, R. K., and Sponaugle, S. (2009). Larval dispersal and marine population connectivity. *Annu. Rev. Mar. Sci.* 1, 443–466. doi: 10.1146/annurev.marine.010908.163757

D'Antonio, C. M., and Thomsen, M. (2004). Ecological resistance in theory and practice. *Weed Techno.* 18, 1572–1577. doi: 10.1614/0890-037X(2004)018[1572:ERITAP]2.0.CO;2

D'Aloia, C. C., Bogdanowicz, S. M., Francis, R. K., Majoris, J. E., Harrison, R. G., Buston, P. M., et al. (2015). Patterns, causes, and consequences of marine larval dispersal. *Proc. Natl. Acad. Sci. U.S.A.* 112 (45), 13940–19945. doi: 10.1073/pnas.1513754112

Degteva, S. V., Ponomarev, V. I., Eisenman, S. W., and Dushenkov, V. (2015). Striking the balance: challenges and perspectives for the protected areas network in northeastern European Russia. *Ambio* 44 (6), 473–490. doi: 10.1007/s13280-015-0636-x

Elith, J., Phillips, S. J., Hastie, T., Dudik, M., Chee, Y. E., and Yates, C. J. (2010). A statistical explanation of MaxEnt for ecologists. *Diversity Distrib* 17 (1), 1–15. doi: 10.1111/j.1472-4642.2010.00725.x

Ferguson, J. W., Healey, M., Dugan, P., and Barlow, C. (2011). Potential effects of dams on migratory fish in the Mekong river: Lessons from salmon in the fraser and Columbia rivers. *Environ. Manage.* 47 (1), 141–159. doi: 10.1007/s00267-010-9563-6

Finn, R. N. (2007). The physiology and toxicology of salmonid eggs and larvae in relation to water quality criteria. *Aquat. Toxicol.* 81 (4), 337–354. doi: 10.1016/j.aquatox.2006.12.021

Gao, Y.-C., Yang, Q., Li, H.-J., Wang, X.-C., and Zhan, A.-B. (2019). Anthropogenic pollutant-driven geographical distribution of mesozooplankton communities in estuarine areas of the bohai Sea, China. *Sci. Rep.* 9, 9668. doi: 10.1038/s41598-019-46047-5

Garcia, M. D., and Barria de Cao, S. M. (2018). Anthropogenic pollution along the coast of a temperate estuary: effects on tintinnid assemblages. *Hydrobiologia* 809, 201–219. doi: 10.1007/s10750-017-3465-z

Gillanders, B. M., Able, K. W., Brown, J. A., Eggleston, D. B., and Sheridan, P. F. (2003). Evidence of connectivity between juvenile and adult habitats for mobile marine fauna: An important component of nurseries. *Mar. Ecol. Prog. Ser.* 247, 281–295. doi: 10.3354/meps247281

Gillanders, B. M., Elsdon, T. S., Halliday, I. A., Jenkins, G. P., Robins, J. B., and Valesini, F. J. (2011). Potential effects of climate change on Australian estuaries and fish utilizing estuaries: a review. *Mar. Fresh. Res.* 62, 1115–1131. doi: 10.1071/MF11047

Hanson, N., Fogel, M., Fong, D. W., and MacAvoy, S. E. (2010). Marine nutrient transport: Anandromous fish migration linked to the freshwater amphipod gammarus fasciatus. *Can. J. Zool.* 88 (6), 546–552. doi: 10.1139/Z10-030

Hauer, F. R., Locke, H., Dreitz, V. J., Hebblewhite, M., Lowe, W. H., Muhlfeld, C. C., et al. (2016). Gravel-bed river floodplains are the ecological nexus of glaciated mountain landscapes. *Sci. Adv.* 2 (6), e1600026. doi: 10.1126/sciadv.1600026

Hilty, J. A., Keeley, A. T. H., Lidicker, W. Z.Jr., and Merenlender, A. M. (2019). *Corridor ecology: Linking landscapes for biodiversity conservation and climate adaptation. 2nd ed* (Washington, DC: Island Press).

Hu, L. J., Song, C., Geng, Z., Zhao, F., Jiang, J., Liu, R. H., et al. (2021). Temporal and spatial distribution of *Coilia mystus* larvae and juveniles in the Yangtze estuary during primary breeding season. *J. Fish. Sci. China* 28 (9), 1152–1161. doi: 10.12264/JFSC2020-0343

Jiang, T., Liu, H.-B., Shen, X.-Q., Shimasaki, Y., Oshima, Y., and Yang, J. (2014). Life history variations among different populations of coilia nasus along the Chinese coast inferred from otolith microchemistry. *J. Fac. Agr. Kyushu U.* 59, 383–389. doi: 10.5109/1467650

Jiang, T., Yang, J., Liu, H., and Shen, X. (2012). Life history of coilia nasus from the yellow Sea inferred from otolith Sr:Ca ratios. *Environ. Biol. Fish.* 95 (4), 503–508. doi: 10.1007/s10641-012-0066-6

Johnston, E. L., and Roberts, D. A. (2009). Contaminants reduce the richness and evenness of marine communities: a review and meta-analysis. *Environ. pollut.* 157 (6), 1745–1752. doi: 10.1016/j.envpol.2009.02.017

Keeley, A. T. H., Ackerly, D. D., Cameron, D. R., Heller, N. E., Huber, P. R., Schloss, C. A., et al. (2018). New concepts, models, and assessments of climate-wise connectivity. *Environ. Res. Lett.* 13, 073002. doi: 10.1088/1748-9326/aacb85

Klaasen, M. (1996). Metabolic constraints on long-distance migration in birds. *J. Exp. Biol.* 199 (1), 57–64. doi: 10.1242/jeb.199.1.57

Lamb, J. M., Ralph, T. M. C., Goodman, S. M., Bogdanowicz, W., Fahr, J., Gajewska, M., et al (2008). Phylogeography and predicted distribution of African-Arabian and Malagasy populations of giant mastiff bats, Otomops spp. (Chiroptera: Molossidae). *Acta Chiropterologica* 10 (1), 21–40. doi: 10.3161/150811008X331063

Lawler, J. J., Ruesch, A. S., Olden, J. D., and McRae, B. H. (2013). Projected climate-driven faunal movement routes. *Ecol. Lett.* 16, 1014–1022. doi: 10.1111/ele.12132

Levin, L. A., Boesch, D. F., Covich, A., Dahm, C., Erséus, C., Ewel, K. C., et al. (2001). The function of marine critical transition zones and the importance of sediment biodiversity. *Ecosystems* 4, 430–451. doi: 10.1007/s10021-001-0021-4

Liu, S. H., and Xu, Z. L. (2011). Comparison of zooplankton lists between coilia mystus food contents and collections from the Yangtze river estuary & hangzhou bay. *Acta Ecologica Sinica.* 31 (8), 2263–2271.

Li, Y. X., Yang, S. M., and Yan, Y. T. (2022). Characteristics of phytoplankton assemblages in the central south China Sea. *Mar. pollut. Bull.* 174, 173189. doi: 10.1016/j.marpolbul.2021.113189

Mantas, T. P., CVarotti, C., Roveta, C., Palma, M., Carlo Innocenti, I., Giusti, M., et al. (2022). Mediterranean Sea Shelters for the gold coral *Savalia savaglia*: An assessment of potential distribution of a rare parasitic species. *Mar. Environ. Res.* 179, 1–12. doi: 10.1016/j.marenvres.2022.105686

McGuire, J. L., Lawler, J. J., McRae, B. H., Nuñez, T. A., and Theobald, D. M. (2016). Achieving climate connectivity in a fragmented landscape. *Proc. Natl. Acad. Sci. U.S.A.* 113 (26), 7195–7200. doi: 10.1073/pnas.160281711

McRae, B. H. (2006). Isolation by resistance. *Evolution* 60 (8), 1551–1561. doi: 10.1554/05-321.1

McRae, B., Shah, V. B., and Edelman, A. (2016). Circuitscape: modeling landscape connectivity to promote conservation and human health. *The Nature Conservancy, Fort Collins, CO*, 14. doi: 10.13140/RG.2.1.4265.1126

Miklos, L., Diviakova, A., and Izakovičov., Z. (2019). *Ecological networks and territorial systems of ecological stability* (London: Springer Nature).

Pelage, L., Guazzelli-Gonzalez, J. G., Le Loch, F., Ferreira, V., Munaron, J., Lucena-Fredou, F., et al. (2021). Importance of estuary morphology for ecological connectivity with their adjacent coast: A case study in Brazilian tropical estuaries. *Estuar. Coast. Shelf. S.* 251 (1-2), 1–11. doi: 10.1016/j.ecss.2021.107184

Phillips, S. J., Anderson, R. P., Dudík, M., Schapire, R. E., and Blair, M. E. (2017). Opening the black box: an open-source release of maxent. *Ecography* 40 (7), 887–893. doi: 10.1111/ecog.03049

Phillips, S. J., Anderson, R. P., and Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecol. Model.* 190 (3–4), 231–259. doi: 10.1016/j.ecolmodel.-2005.03.026

Pinto, N., and Keitt, T. H. (2009). Beyond the least-cost path: evaluating corridor redundancy using a graph-theoretic approach. *Landscape Ecol.* 24 (2), 253–266. doi: 10.1007/s10980-008-9303-y

Planes, S., Jones, G. P., and Thorrold, S. R. (2009). Larval dispersal connects fish populations in a network of marine protected areas. *Proc. Natl. Acad. Sci. U.S.A.* 106 (14), 5693–5697. doi: 10.1073/pnas.0808007106

R Core Team (2014). *R: A language and environment for statistical computing* (Vienna: R Foundation for Statistical Computing).

Rossi, T., Pistevos, J. C. A., Connell, S. D., and Nagelkerken, I. (2018). On the wrong track: ocean acidification attracts larval fish to irrelevant environmental cues. *Sci. Rep.* 8, 5840. doi: 10.1038/s41598-018-24026-6

Runge, C. A., Watson, J. E. M., Butchart, S. H., Hanson, J. O., Possingham, H. P., and Fuller, R. A. (2015). Protected areas and global conservation of migratory birds. *Science* 350, 1266–1258. doi: 10.1126/science.aac9180

Saenz-Agudelo, P., Jones, G. P., Thorrold, S. R., and Planes, S. (2010). Connectivity dominates larval replenishment in a coastal coral reef fish metapopulation. *Proc. R. Soc B.* 278, 2954–2961. doi: 10.1098/rspb.2010.2780

Scanes, E., Scanes, P. R., and Ross, P. M. (2020). Climate change rapidly warms and acidifies Australian estuaries. *Nat. Commun.* 11 (1), 1803. doi: 10.1038/s41467-020-15550-z

Schram, E., Roques, J., van Kuijk, T., Abbink, W., Heul, J., Vries, P., et al. (2014). The impact of elevated water ammonia and nitrate concentrations on physiology, growth and feed intake of pikeperch (Sander lucioperca). *Aquaculture* 420, 95–104. doi: 10.1016/j.aquaculture.2013.10.027

Seidensticker, J., Dinerstein, E., Goyal, S. P., Gurung, B., Harihar, A., Johnsingh, A., et al. (2010). "Tiger range collapse and recovery at the base of the Himalayas: a case study," in *Biology and conservation of wild felids* (London: Oxford Unversity Press).

Shen, H., He, S. L., Pan, D. A., and Li, J. F. (1992). A study of turbidity maximum in the changjiang estuary. *Acta Geographica Sinica.* 47 (5), 472–479. doi: 10.11821/xb199205010

Spares, A. D., Stokesbury, M. J. W., O'Dor, R. K., and Dick, T. A. (2012). Temperature, salinity and prey availability shape the marine migration of Arctic char, salvelinus alpinus, in a macrotidal estuary. *Mar. Biol.* 159 (8), 1633–1646. doi: 10.1007/s00227-012-1949-y

Sykes, G. E., and Shrimpton, J. M. (2010). Effect of temperature and current manipulation on smolting in chinook salmon (Oncorhynchus tshawytscha): The relationship between migratory behaviour and physiological development. *Can. J. Fish. Aquat. Sci.* 67 (1), 191–201. doi: 10.1139/F09-166

Thalinger, B., Wolf, E., Traugott, M., and Wanzenböck, J. (2019). Monitoring spawning migrations of potamodromous fish species *via* eDNA. *Sci. Rep.* 9 (1), 15388. doi: 10.1038/s41598-019-51398-0

Tittensor, D. P., Baco, A. R., Brewin, P. E., Clark, M. R., Consalvey, M., Hall-Spencer, J., et al. (2009). Predicting global habitat suitability for stony corals on seamounts. *J. Biogeogr.* 36, 1111–1128. doi: 10.1111/j.1365-2699.2008.02062.x

Tong, M.-X., Zhang, L., Li, J., Zockler, C., and Clark, N. (2012). The critical importance of the rudong mudflats, jiangsu province, China in the annual cycle of the spoon-billed sandpiper calidris pygmeus. *Wader Study Group Bulletin.* 119 (3), 74–77.

Urban, D., and Keitt, T. H. (2001). Landscape connectivity: A graph-theoretic perspective. *Ecology* 82, 1205–1218. doi: 10.1890/0012-9658(2001)082[1205:LCAGTP]2.0.CO;2

Vance, D. J., and Pendrey, R. C.. (2008). Vertical migration of postlarval penaeid prawns in two Australian estuaries: the effect of tide and day/night. *Mar. Freshwater Res.* 59 (8), 671–683. doi: 10.1071/MF07234

Verbruggen, H., Tyberghein, L., Pauly, K., Vlaeminck, C., Van Nieuwenhuyze, K., Kooistra, W., et al. (2009). Macroecology meets macroevolution: evolutionary niche dynamics in the seaweed halimeda. *Global Ecol. Biogeogr.* 18, 393–405. doi: 10.1111/j.1466-8238.2009.00463.x

Vu, A. V., Baumgartner, L. J., Limburg, K. E., Doran, G. S., Mallen-Cooper, M., Gillanders, B. M., et al. (2022). Life history strategies of mekong pangasiid catfishes revealed by otolith microchemistry. *Fish. Res.* 249, 106239. doi: 10.1016/j.fishres.2022.106239

Walsh, C. T., Reinfelds, I. V., Ives, M. C., Gray, C. A., West, R. J., and van der Meulen, D. E. (2013). Environmental influences on the spatial ecology and spawning behaviour of an estuarine-resident fish, macquaria colonorum. *Estuar. Coast. Shelf S.* 118, 60–71. doi: 10.1016/j.ecss.2012.12.009

Wang, H., Wang, P., Xu, C., Sun, Y.-F., Shi, L., Zhou, L., et al. (2022). Can the 'Ten-year fishing ban' rescue biodiversity of the Yangtze river? *Innovation.* 3 (3), 100235. doi: 10.1016/j.xinn.2022.100235

Wang, C.-Y., Wei, Q.-W., Kynard, B., Du, H., and Zhang, H. (2012). Migrations and movements of adult Chinese sturgeon acipenser sinensis in the Yangtze river, China. *J. Fish Biol.* 81 (2), 696–713. doi: 10.1111/j.1095-8649.2012.03365.x

Wang, S.-K., Zhang, T., Yang, G., Wang, Y., Zhao, F., and Zhuang, P. (2018). Migration and feeding habits of juvenile Chinese sturgeon (Acipenser sinensis Gray 1835) in the Yangtze estuary: Implications for conservation. *Aquat. Conservation: Mar. Freshw. Ecosystems.* 28 (6), 1329–1336. doi: 10.1002/aqc.2935

Weeks, R. (2017). Incorporating seascape connectivity in conservation prioritization. *PloS One* 12 (7), e0182396. doi: 10.1371/journal.pone.0182396

Wolanski, E. (2017). Bounded and unbounded boundaries-untangling mechanisms for estuarine-marine ecological connectivity: Scales of m to 10,000 km-a review. *Estuar. Coast. Shelf S.* 198, 378–392. doi: 10.1016/j.ecss.2016.06.022

Wu, H., Deng, B., Yuan, R., Gu, J. H., Shen, F., Zhu, J. R., et al. (2013). Detiding measurement on transport of the changjiang-derived buoyant coastal current. *J. Phys. Oceanogr.* 43 (11), 2388–2399. doi: 10.1175/JPO-D-12-0158.1

Wu, L. H., Tang, W. Q., and Zhang, Y. (2017). Research on the differences of anadromous migratory distance between *Coilia mystus* and *Coilia nasus* based on the transfer process of body lipid. *J. fisheries China* 41 (2), 212–220. doi: 10.11964/jfc.20160310330

Wu, T., and Wu, H. (2018). Tidal mixing sustains a bottom-trapped river plume and buoyant coastal current on an energetic continental shelf. *J. Geophys. Res-Oceans* 123 (11), 8026–8051. doi: 10.1029/2018JC014105

Wu, H., Zhu, J. R., Shen, J., and Wang, H. (2011). Tidal modulation on the changjiang river plume in summer. *J. Geophys. Res.* 116 (C8), 1–21. doi: 10.1029/2011jc007209

Xia, Y. L. (2020). Correlation and association analyses in microbiome study integrating multiomics in health and disease. *Prog. Mol. Biol. Transl.* 171, 309–491. doi: 10.1016/bs.pmbts.2020.04.003

Xu, N., Zhu, B., Shi, F., Shao, K., Que, Y.-F., Li, W. T., et al. (2018). Monitoring seasonal distribution of an endangered anadromous sturgeon in a large river using environmental DNA. *Sci. Nat.* 105, (11–12). doi: 10.1007/s00114-018-1587-4

Yang, J., Arai, T., Liu, H., Miyazaki, N., and Tsukamoto, K. (2006a). Environmental signature in the otolith elemental fingerprint of the tapertail anchovy, *Coilia mystus*, from the changjiang estuary, China. *J. Appl. Ichthyol.* 22 (5), 459–462. doi: 10.1111/j.1439-0426.2006.00776.x

Yang, J., Arai, T., Liu, H., Miyazaki, N., and Tsukamoto, K. (2006b). Reconstructing habitat use of *Coilia mystus* and *Coilia ectenes* of the Yangtze river estuary, and of *Coilia ectenes* of taihu lake, based on otolith strontium and calcium. *J. Fish Biol.* 69 (4), 1120–1135. doi: 10.1111/j.1095-8649.2006.01186.x

Yang, Q., Zhao, F., Song, C., Zhang, T., Zhuang, P., Jiang, T., et al. (2019). Habitat history reconstruction of coilia mystus from the Yangtze river estuary and its adjacent sea area. *J. Fish. Sci. China* 26 (6), 1175–1184.

Yates, C., McNeill, A., Elith, J., and Midgley, G. (2010). Assessing the impacts of climate change and land transformation on banksia in the south West Australian floristic region. *Divers. Distrib.* 16, 187–201. doi: 10.1111/j.1472-4642.2009.00623.x

Yatsuyanagi, T., and Araki, H. (2020). Understanding seasonal migraton of shishamo smelt in coastal regions using environmental DNA. *PloS One* 15, e0239912. doi: 10.1371/journal.pone.0239912

Zeng, Q., and Dong, F.-Y. (1993). Study on the biological characteristics and factor correlation of coilia mystus propagating population. *J. lake Sci.* 5 (2), 164–170.

Zhuang, P., Zhang, T., Li, S.-F., Ni, Y., Wang, Y.-H., Deng, S. M., et al. (2006). *Fishes of the Yangtze estuary* (Beijing: China agriculture press).

Zhu, Y., Liao, W.-G., Feng, S.-X., and Tan, H.-W. (2009). "Analysis of habitat characteristics of the neophocaena phocaenoides asiaeorientalis," in *Advances in hydraulics and hydraoinformatics in China*. Xi'an, China: Xi'an Jiaotong University Press. 140–145.

frontiers | Frontiers in Marine Science

TYPE Original Research
PUBLISHED 19 October 2022
DOI 10.3389/fmars.2022.1008242

OPEN ACCESS

EDITED BY
Junyu He,
Zhejiang University, China

REVIEWED BY
Yuejian Chen,
Tongji University, China
Yuan Wei,
Shanghai University, China

*CORRESPONDENCE
Qing Li
qingli@ahau.edu.cn

SPECIALTY SECTION
This article was submitted to Ocean Observation, a section of the journal Frontiers in Marine Science

RECEIVED 31 July 2022
ACCEPTED 20 September 2022
PUBLISHED 19 October 2022

CITATION
Li Q (2022) Spatio-temporal nonconvex penalty adaptive chirp mode decomposition for signal decomposition of cross-frequency coupled sources in seafloor dynamic engineering. *Front. Mar. Sci.* 9:1008242. doi: 10.3389/fmars.2022.1008242

Spatio-temporal nonconvex penalty adaptive chirp mode decomposition for signal decomposition of cross-frequency coupled sources in seafloor dynamic engineering

Qing Li*

Department of Mechanical Engineering, Anhui Agricultural University, Hefei, China

Electromagnetic field noise and clutter generated from the motion of ocean waves are the main obstacles in the research of magnetotelluric dynamic analysis, and it is difficult to extract the crossed instantaneous frequencies (IFs) of underwater electromagnetic detected (UEMD) data due to the limited resolution of the current time-frequency techniques. To alleviate this bottleneck issue, a new spatio-temporal nonconvex penalty adaptive chirp mode decomposition (STNP-ACMD) is originally proposed for separating each mono-component individually from a complicated multi-component with severely crossed IFs or overlapped components, in this paper. Specifically, the idea of a nonconvex penalty greedy strategy is incorporated into the vanilla ACMD method by using a recursive mode extraction scheme, and the fractional-order characteristic of the observation signal is also considered. Meanwhile, the spatio-temporal matrices were constructed elaborately and then applied to capture coupling characteristics and spatio-temporal relationships among all estimated mono-components. Eventually, a high-resolution adaptive time-frequency spectrum is obtained according to the IFs and instantaneous amplitudes (IAs) of each estimated mono-component. The effectiveness and practicability of the proposed algorithm were verified *via* simulated scenarios and velocity dynamic data of the seafloor from the South China Sea, compared with four state-of-the-art benchmarks.

KEYWORDS

spatio-temporal matrices, nonconvex penalty, adaptive chirp mode decomposition, crossed instantaneous frequency, seafloor dynamic

1 Introduction

Underwater electromagnetic detected (UEMD) data with low frequency have a wide application prospect in deep-sea exploration and deep-sea magnetic field measurements such as oil and mineral exploration, underwater pipe-cable overhaul, fishery, navigation, positioning, ship testing, and anti-submarine demining, due to many merits including high accuracy, stability, and relatively small restrictions by hydrologic weather, and so on. However, the UEMD data with low frequency is susceptible to the electromagnetic noise induced by the ocean current, and the issue of frequency range overlap between UEMD data with low frequency and induced electromagnetic noise is becoming the main obstacle to improving the signal-to-noise ratio (SNR) of UEMD data since the signal recorded *via* the receiver is usually affected by different types of noise, such as electromagnetic field noise, clutter, and marine animals (Park et al., 2019; Nyqvist et al., 2020; Schwalenberg et al., 2020).

The UEMD data, such as radar-echo wave and ocean current sonar, usually exhibit non-stationary, time-varying, and multi-component characteristics, and those data (sound and signals) always contain useful information for engineering applications. Extracting or isolating the constituent modes from their complicated multi-dimensional or multi-component nature is a pivotal task for understanding the dynamic responses in seafloor engineering. Therefore, how to achieve accurate signal decomposition and represent the internal physical characteristics of a system has been a hot research topic.

1.1 Literature review

Over the past few years, multiple methods and studies have been adopted to alleviate this issue. For example, in 2013, MacLennan et al. proposed a flexible and practical denoising technique named equivalent source method (ESM) to remove the additional noise from multicomponent controlled-source electromagnetic method (CSEM) data (MacLennan and Li, 2013). In 2020, Chen et al. utilized the current meter technique (CMT) for the reduction of seawater motion-induced electromagnetic noise from marine magnetotelluric data (Chen et al., 2020). In 2020, the compressive sensing (CS) method with orthogonal matching pursuit (OMP) was presented for reducing noise impacts from marine CSEM data, and the windowed Fourier transform, wavelet transform, and CS method with different training dictionaries were used as benchmarks for comparison (Zhang et al., 2020).

Actually, those non-stationary and complicated time-varying signals usually contain multiple components and are also polluted by additional noise, some of which may overlap (or severely cross) in both time and frequency domains. To tackle this issue, many advanced signal processing methods, algorithms, and techniques have been developed in the past decades. Roughly, those methods can be classified into three main categories: (i) time domain methods, (ii) frequency domain methods, and (iii) time-frequency domain methods, in terms of the processing domain in which the signal decomposition is performed.

In the first class of signal decomposition that is performed in the time domain, the typical representative is pioneered by the empirical mode decomposition (EMD) method (Huang et al., 1998; Tian et al., 2022). The EMD is a fully data-driven and recursive sifting scheme in which a non-stationary multi-component can be decomposed into a set of quasi-orthogonal mono-components called intrinsic mode functions (IMFs) without considering any prior knowledge or statistical information (e.g., temporal or spectral characteristics) of the raw observation (i.e., multi-component data). However, the connatural issues of mode-mixing, ending effect, sensitivity to noise, and lack of mathematical foundations which lead to unexpected and poor results. To alleviate those drawbacks, a series of improved algorithms such as the ensemble EMD (EEMD) (Wu and Huang, 2009), the complete EEMD (CEEMD) (Torres et al., 2011), and the CEEMD with adaptive noise (CEEMDAN) (Lin et al., 2022), time-varying filter-based EMD (TVF-EMD) (Jiang et al., 2020), have been reported. However, these methods are rather restrictive and only overcome parts of the issues mentioned above, and meanwhile, unfortunately, some new obstacles are encountered. For example, the mode-mixing problem can be addressed by the CEEMDAN method, but white noise cannot be eliminated; the problems of mode-mixing and ending effect in the EEMD, and they still cannot achieve the desired performance.

In the second class of signal decomposition that is performed in the frequency domain, the representative models are pioneered by empirical wavelet transform (EWT) (Gilles, 2013) and variational mode decomposition (VMD) methods (Dragomiretskiy and Zosso, 2014). The basic idea of the EWT method is that the wavelet filter bank (WFB) and spectrum support are used for reconstructing each mode, but the spectrum information should be known prior to the support detection algorithm. Moreover, modes with close frequency-band cannot be separated and identified *via* the EWT method. The VMD method is regarded as an adaptive Wiener filter bank in the frequency domain, and all subcomponents are assumed to be narrowband signals that cluster around their respective central frequencies. However, the main shortcoming of the VMD is that parameter α and the number of modes K should be determined before the algorithm executes, and the issues of mode-mixing and bandwidth of the estimated modes will be disturbed if model parameters are not selected properly (e.g., overestimating or underestimating). More importantly, the EWT and VMD methods rely on a restrictive assumption that the quasi-orthogonal narrowband subcomponent should be stationary and that chirp modes whose frequency ranges overlap cannot be decomposed.

In this regard, methods for separating modes with overlapped spectrums have been explored and presented in the past few years. In 2017, Chen et al. proposed a variational

nonlinear chirp mode decomposition (VNCMD) algorithm for analyzing wide-band nonlinear chirp signals (NCSs). This method relies on a new assumption that the wide-band NCS could be transformed into a narrow-band signal *via* signal demodulation analysis, and then three simulated data and two real signals of killer whale whistles are introduced for discussion. However, the fatal flaw of the VNCMD is that, like the VMD method, the number of estimated modes still needs to be pre-determined before the algorithm executes (Chen et al., 2017). In 2019, Chen et al. developed a non-parametric decomposition technique named adaptive chirp mode pursuit (ACMP) for processing multimodal non-stationary wideband chirp signals whose frequency ranges overlap, so as to obtain an optimal initialization of modes frequencies. Eventually, simulated data and systematic application (i.e., rotor fault diagnosis) are presented for performance investigation of the ACMP. However, this method cannot address non-stationary, multimodal signals with crossing IFs, yielding problems of end effect and severe oscillation (Chen et al., 2019). Furthermore, to address the issues of the mode number raised by the VNCMD algorithm and, meanwhile, to extract the fast-fluctuating IFs from non-stationary multimodal signals, the adaptive chirp mode decomposition (ACMD) has been developed, in which a greedy algorithm and a recursive mode extraction scheme are designed elaborately, which is motivated by the VMD method. The adaptive TF spectrum (ATFS) technique is employed for obtaining the IAs and IFs (Chen et al., 2019). In 2020, to tackle the problem of oscillation detection and fault diagnosis, a fast ACMD (FACMD) algorithm was established to process complex multiple oscillations, where the running time of the proposed FACMD algorithm is reduced greatly by setting a reasonable stop threshold compared with the vanilla ACMD (Chen et al., 2020). Currently, in 2022, a two-level ACMD (TL-ACMD) is reported by incorporating the tangent energy and calculating energy eigenvalues of each mono-mode for decomposing the acoustic data of the impeller defects in a centrifugal pump (Vashishtha et al., 2022). Although the decomposition methods listed above have some promising advantages in the simulated domain and industrial areas, the common shortcoming of those decomposition methods is their poor capacity for separating each mono-component individually from a complicated multi-component with severely crossed IFs characteristics or overlapped components.

In the third class of signal decomposition that is performed in the time-frequency domain, the time-frequency representation (TFR) has been developed using the reassignment technique of energy distributions, so as to separate modes with the overlapped spectrum, as well as modes with close frequency-band and the IFs are disjoint. Among various TFR techniques, as a post-processing procedure, the synchrosqueezing transform (SST) has extensive concern since the narrowband subcomponents (or IMFs) are estimated from the enhanced TFR after the synchrosqueezing operation (Daubechies et al., 2011; Li et al., 2022). In 2015, Clausel et al. proposed a new bi-dimensional version algorithm called monogenic synchrosqueezed wavelet transform (MSSWT) for decomposing the demodulated AM–FM images, which is a bivariate extension of the SST, and the concept of intrinsic monogenic mode (IMM) was introduced (Clausel et al., 2015). In 2021, to overcome the issue of mode mixing in decomposing the multicomponent signals, the down-sampled STFT and synchrosqueezing-based transform demodulation were introduced (Meignen et al., 2021). In 2022, Si et al. designed a novel method called multivariate synchrosqueezing transform (MSST) for extracting the high multiple frequency components from deep hole drilling vibration data, which is an extension of the SST to a multivariate version (Si et al., 2022). In 2022, to highlight the time-frequency (TF) resolution of mono-mode signals that are extracted from multicomponent data with an impulse-like waveform, a time-reassigned SST (TSST) has been reported under the framework of STFT, and several numerical experiments were conducted to reveal the performance of the TSST method (Dong et al., 2022). Currently, Li et al. constructed a novel synchrosqueezing polynomial Chirplet transform (SPCT) for mono-mode extraction from multicomponent data in which a multi-kernel operator was designed for rearranging and concentrating the energy of IF in mono-mode. The field seismic data was used for performance analysis of the SPCT algorithm compared with other state-of-the-art TFA techniques such as synchro-extracting transform (SET) and SST (Li et al., 2022). The SST approach implements an energy relocation framework along the frequency direction, which sharpens the time-frequency ridges of modes by focusing their energies on the center of gravity in energy distribution, and thus the estimated mode has higher energy aggregation (Auger et al., 2013). However, it is worth mentioning that signal reconstruction and computationally intensive reassignment techniques cannot be maintained and addressed, which is limited in real applications (Auger et al., 2013). Additionally, how to accurately measure and evaluate the time-frequency energy concentration of a certain component is still a changeling issue that needs to be explored more.

1.2 Motivations and contributions

The aforementioned decomposition methods mainly focus on TF pattern generation of multi-component/signals with a clear range of frequency-band distribution. However, in practical applications, data collected from sensors or receivers is always accompanied by some severely crossed modes in the whole time-span. The ACMD is not suitable for processing high-frequency signals with aliasing and crossover features, yielding the wrong IFs for some intersecting components. Motivated by the ACMD method, a challenge exists in how to apply the greedy strategy to the identification of modes with severely crossed frequencies and separating each overlapped component clearly and accurately, which is the key motivation of this work.

Another crucial issue that needs to be highlighted is that, in most of the scenarios, the aforementioned techniques, such as the VNCMD method, the ACMD method, and the VMD method, mainly focus on channel-wise processing for isolating or extracting the single component from complicated multi-channel data. These techniques, which are employed for the decomposition of complicated time-varying multi-component signals, typically do not fully cater to the coupled nature and spatial–temporal characteristics of each mode during signal estimation. Most of the existing algorithms operate with mode-wise processing, which is not optimal for such modes with spatial–temporal relationships. They may fail to capture the coupled nature of all the components simultaneously, and yield poor decomposability when dealing with multi-dimensional or multi-component data, and also may fail to distinguish the frequency distribution of the components. In addition, in the traditional VNCMD method, the number of estimated components cannot be adaptively determined; i.e., the number of modes should be assumed to be known in advance, which is not feasible in practical applications. Meanwhile, in essence, the mode-wise processing is still conducted in the classic ACMD method, which restricts the application of the ACMD method. By elaborating on the multi-dimensional decomposition techniques, it is therefore natural to investigate the spatio-temporal relationships and coupling nature among estimated components.

In this paper, to address the above drawbacks, a new adaptive signal decomposition algorithm is proposed, and simulation and experimental cases are investigated to verify the availability of the proposed algorithm. The main contributions of this paper can be summarized as follows:

(1) The spatio-temporal nonconvex penalty adaptive chirp mode decomposition (STNP-ACMD) algorithm is proposed for the decomposition of the FM signals whose components with crossing IFs and overlapped components, and the split Bregman algorithm is introduced for solving the constrained optimization problem. The overlapped components or crossed IFs can be separated adaptively by using the recursive mode extraction scheme. The proposed STNP-ACMD algorithm belongs to the second class of signal decomposition algorithms described above. To the best of our knowledge, it is the first time the spatio-temporal nonconvex penalty has been applied to the ACMD framework in this regard, which also retains the reconstruction benefit of the ACMD.

(2) The key contribution is that the spatio-temporal relationship and the coupled nature of the available information (e.g., crossing IFs) among estimated components are exploited by designing and elaborating the temporal matrix and the spatial matrix. In addition, the characteristic of the fractional order of the observation signal is considered, which is rarely reported in the existing literature.

(3) Both simulation and experimental case verification showed the effectiveness of the proposed algorithm for the decomposing signal with crossing IFs and overlapping components. The estimated accuracy of the proposed algorithm has improved greatly compared with four state-of-the-art benchmarks, including the ACMD, the VNCMD, the EEMD, and the VMD methods.

The integrated framework of the rest of the paper is structured as follows: In Section 2, the theoretical preliminaries of variational nonlinear chirp mode decomposition (VNCMD) and adaptive chirp mode decomposition (ACMD) algorithms are reviewed first, and then the proposed spatio-temporal nonconvex penalty adaptive chirp mode decomposition (STNP-ACMD) algorithm and its solution are presented. The numerical simulation is given and discussed in Section 3. In Section 4, the experimental case and its analysis are discussed to verify the performance of the proposed STNP-ACMD algorithm. Conclusions and possible explorations are drawn in Section 5.

2 The STNP-ACMD algorithm

In this section, we first review the classic VNCMD and ACMD algorithms (Chen et al., 2019; Chen et al., 2019), and then the proposed STNP-ACMD algorithm and its solution will be presented later. Meanwhile, the solution strategy of the fractional order of the STNP-ACMD algorithm based on the rescaled range (R/S) method is given.

2.1 Review of the VNCMD and ACMD algorithms

Generally, the fast-aliasing frequency modulation (FM) signal (or called Chirp signal) has the characteristic of time-varying amplitude and frequency contaminated by environmental noise, which can be expressed as,

$$s(t) = \sum_{i=1}^{K} s_i(t) + r(t) = \sum_{i=1}^{K} A_i(t)\cos\left(2\pi\int_0^t f_i(\tau)d\tau + \theta_i\right) + r(t) \quad (1)$$

where $s_i(t)$ represents each Chirp component, K is the number of chirp components. $A_i(t)>0$, $f_i(t)>0$ and θ_i represent the instantaneous amplitude (IA), instantaneous frequency (IF), and initial phase (IP) of the i-th Chirp component $s_i(t)$, respectively. Also, $r(t)$ denotes the zero-mean Gaussian white noise (GWN). According to the demodulation technique, Equation (1) can be rewritten as,

$$s(t) = \sum_{i=1}^{K} a_i(t)\cos\left(2\pi\int_0^t \hat{f}_i(\tau)\,d\tau\right) + b_i(t)\sin\left(2\pi\int_0^t \hat{f}_i(\tau)\,d\tau\right) + r(t) \tag{2}$$

where $\hat{f}_i(\tau)$ stands for frequency function of the $\cos(2\pi\int_0^t \hat{f}_i(\tau)\,d\tau)$ and $\sin(2\pi\int_0^t \hat{f}_i(\tau)\,d\tau)$, the amplitudes $a_i(t)$ and $b_i(t)$ are defined as $a_i(t) = A_i(t)\cos(2\pi\int_0^t (f_i(\tau) - \hat{f}_i(\tau)d\tau + \theta_i)$ and $b_i(t) = -A_i(t)\sin(2\pi\int_0^t (f_i(\tau) - \hat{f}_i(\tau)d\tau + \theta_i)$.

To estimate the components $s_i(t)$, i=1,2,⋯,K from observation, the hypothesis of the variational nonlinear chirp mode decomposition (VNCMD) is developed for minimizing the bandwidths of each mono-component, and the objective cost function for the constrained optimization problem is given as (Chen et al., 2019),

$$\begin{aligned} &\min_{a_i(t),b_i(t),\hat{f}_i(t)} \left\{\sum_{i=1}^{K} \| a_i''(t) \|_2^2 + \| b_i''(t) \|_2^2\right\} \\ &\text{s.t. } s(t) = \sum_{i=1}^{K} a_i(t)\cos\left(2\pi\int_0^t \hat{f}_i(\tau)\,d\tau\right) + b_i(t)\sin\left(2\pi\int_0^t \hat{f}_i(\tau)\,d\tau\right) + r(t) \end{aligned} \tag{3}$$

where K is the number of the estimated signals, which should be pre-determined before the algorithm executes.

Furthermore, the idea of a greedy algorithm is adopted by the pure adaptive chirp mode decomposition (ACMD), and the component signals can be estimated adaptively, and the cost function for constrained optimization problem is given as (Chen et al., 2019),

$$\begin{aligned} &\min_{a_i(t),b_i(t),\hat{f}_i(t)} \left\{\| a_i'(t) \|_2^2 + \| b_i'(t) \|_2^2\right\} \\ &\text{with } s_i(t) = a_i(t)\cos\left(2\pi\int_0^t \hat{f}_i(\tau)\,d\tau\right) + b_i(t)\sin\left(2\pi\int_0^t \hat{f}_i(\tau)\,d\tau\right) + r(t) \end{aligned} \tag{4}$$

where $\alpha \left\| s(t) - s_i(t) \right\|_2^2$ is residual energy of residual signal after extracting the estimated signal, and α >0 is a weighting coefficient. For more details, the iterative procedures of the specific algorithm can be found in (Chen et al., 2019; Chen et al., 2019).

2.2 The proposed STNP-ACMD algorithm and its solution

In this section, the proposed STNP-ACMD algorithm is presented for aliasing FM signals decomposition, and the specific procedures for solving the algorithm for the STNP-ACMD optimization problem are presented later.

Considering the intrinsic characteristics of fractional-order and spatio-temporal coupling of the estimated components, for the i-th component signal, the new cost function for the constrained optimization problem is expressed as,

$$\min_{a_i(t),b_i(t),\hat{f}_i(t)} \left\{ \begin{array}{l} \| a_i''(t) \|_2^2 + \| b_i''(t) \|_2^2 + \alpha \| s(t) \\ - s_i(t) \|_2^2 + \lambda_1 \| \boldsymbol{D}_t a_i''(t) \|_p^p + \lambda_2 \| \boldsymbol{D}_s b_i''(t) \|_p^p \end{array} \right\}$$

$$\text{with } s_i(t) = a_i(t)\cos a_i(t)\cos\left(2\pi\int_0^t \hat{f}_i(t)d\tau\right) + b_i(t)\sin\left(2\pi\int_0^t \hat{f}_i(\tau)\,d\tau\right) + r(t) \tag{5}$$

where p is a fractional-order, $\boldsymbol{D}_t$ and $\boldsymbol{D}_s$ are the temporal matrix and spatial matrix, respectively. The size of matrices $\boldsymbol{D}_t$ and $\boldsymbol{D}_s$ is $mn \times mn$, where m and n are the number of channels and the sampling data point of each channel, respectively.

Both temporal matrix $\boldsymbol{D}_t$ and spatial matrix $\boldsymbol{D}_s$ can be created using the reverse banded sparse-diagonal (RBSD) method. The procedures for developing the temporal and spatial matrices are as follows,

Step 1: Extract the nonzero diagonal of simulated matrix $\boldsymbol{A}$, i.e., $\boldsymbol{A}_t$ and $\boldsymbol{A}_s$, to create $\boldsymbol{D}_1$ and $\boldsymbol{D}_2$ using the RBSD method. The random temporal matrix $\boldsymbol{A}_t$ is simulated by MATLAB with $\boldsymbol{A}_t$ = [−ones(n,1) ones(n,1)], and n is the number of sampling point in one channel. The random spatial matrix $\boldsymbol{A}_s$ is simulated by MATLAB with $\boldsymbol{A}_s$ = [−ones(m,1) ones(m,1)], and m is the number of the estimated component. Then, $\boldsymbol{D}_1$ is obtained with ($\boldsymbol{A}_t$, [0 1], n, n), and $\boldsymbol{D}_2$ is obtained by ($\boldsymbol{A}_s$, [0 1], m, m).

Step 2. Remove the last nonzero point from matrices $\boldsymbol{D}_1$ and $\boldsymbol{D}_2$, then two new matrices $\boldsymbol{D}^T{}_1$ and $\boldsymbol{D}^T{}_2$ are obtained;

Step 3. Constituent an identity sparse matrix $\boldsymbol{I}_t$ with size $m \times m$, where m is the number of channels. Constituent an identity sparse matrix $\boldsymbol{I}_s$ with size $n \times n$, where n is the number of sampling points in one channel;

Step 4. The temporal matrix $\boldsymbol{D}_t$ and spatial matrix $\boldsymbol{D}_s$ can be calculated *via* the Kronecker product in terms of identity sparse matrices, i.e., $\boldsymbol{D}_t$ = Kronecker-product< $\boldsymbol{D}^T{}_1$, $\boldsymbol{I}_t$ >; $\boldsymbol{D}_s$ = Kronecker-product< $\boldsymbol{I}_t$, $\boldsymbol{D}^T{}_2$>. That is, matrices $\boldsymbol{D}_t$ and $\boldsymbol{D}_s$ are defined as $\boldsymbol{D}_t$=$Diag$[−1 1, −1 1, ⋯,0 0, 0 1, n, n]$^T{}_{n\times n}$⊗$Diag$[1, 1, ⋯, 1]$^T{}_{m\times m}$, $\boldsymbol{D}_s$=$Diag$[1, 1, ⋯, 1]$^T{}_{n\times n}$⊗$Diag$[−1 1, −1 1, ⋯, 0 0, 0 1, m, m]$^T{}_{m\times m}$, where symbol ⊗ is the Kronecker product.

Following those procedures of developing both temporal matrix $\boldsymbol{D}_t$ and spatial matrix $\boldsymbol{D}_s$, the coupled nature of the available information within the estimated components can be effectively exploited. Here, Equation (5) can be rewritten in constrained matrix notation as follows,

$$\begin{aligned}\mathbf{x}_i &= \begin{bmatrix} a_i(t) \\ b_i(t) \end{bmatrix} = \min_{u,v,w,z,\hat{f}_i(t)} \\ &\left\{ || \boldsymbol{\Theta}\mathbf{x}_i ||_2^2 + \alpha || s - \boldsymbol{G}_i\mathbf{x}_i ||_2^2 + \lambda_1 || \boldsymbol{D}_t\mathbf{x}_i ||_p^p + \lambda_2 || \boldsymbol{D}_s\mathbf{x}_i ||_p^p \right\} \\ &s.t.\ \ \boldsymbol{u}=\boldsymbol{\Theta}\boldsymbol{x}_i,\ \boldsymbol{v} = \boldsymbol{s} - \mathbf{G}_i\mathbf{x}_i,\ \boldsymbol{w} = \boldsymbol{D}_t\mathbf{x}_i,\ \mathbf{z} = \boldsymbol{D}_s\mathbf{x}_i\end{aligned} \tag{6}$$

where the $\boldsymbol{\Theta}$ = diag($\boldsymbol{\Omega}$, $\boldsymbol{\Omega}$), $\boldsymbol{\Omega}$ is a second-order derivative operator, $x_i = [a_i^T, b_i^T]^T$, and $\boldsymbol{G}_i$ = [$\boldsymbol{C}_i$, $\boldsymbol{S}_i$] with C_i=*diag*[*cos*($\varphi_i(t_0)$),*cos*($\varphi_i(t_1)$),⋯,*cos*($\varphi_i(t_{N-1})$)], S_i=*diag*[*sin*($\varphi_i(t_0)$),*sin*($\varphi_i(t_1)$),⋯, *sin*($\varphi_i(t_{N-1})$)] and $\varphi_i(t) = 2\pi \int_0^t f(x)\hat{}\, d\tau$.

The fractional-order p can be calculated using the rescaled range (R/S) method in the long-range dependence (LRD) domain. The expression of the R/S method is given by (Hurst, 1951; Mason, 2016),

$$\frac{R(n)}{S(n)} = \frac{1}{S(n)} \left\{ \max \sum_{i=1}^{n} [X(t_i) - \frac{1}{n}\sum_{i=1}^{n} X(t_i)] - \min \sum_{j=1}^{n} [X(t_j) - \frac{1}{n}\sum_{j=1}^{n} X(t_j)] \right\} \tag{7}$$

where $R(n)$ is the data renormalization range, and $S(n)$ is standard deviation, i.e., $S(n) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(X(t_i) - \frac{1}{n}\sum_{i=1}^{n} X(t_i))^2}$, $X(t_i)$ is the training data at time t_i, i = 1:k (k denotes the number of data channels). For multiple channel data, the Hurst index H is obtained by the slope value of *log* ($R(n)/S(n)$) vs *log*($S(n)$) in a logarithmic plot. Hence, the fractional-order p is calculated *via* the Hurst index H,

$$p = H - 0.5 = \frac{1}{k}\sum_{i=1}^{k} H(i) - 0.5 \tag{8}$$

According to the procedures of the split Bregman iteration (SBI) algorithm (Goldstein and Osher, 2009; Corsaro et al., 2021), Equation (6) can be expressed as,

$$\min_{u,v,w,z} \left\{ \begin{array}{l} \| \boldsymbol{u} \|_2^2 + \alpha \| \boldsymbol{v} \|_2^2 + \lambda_1 \| \boldsymbol{w} \|_p^p + \lambda_2 \| z \|_p^p + \mu_1 \| \boldsymbol{u} - \boldsymbol{\Theta}x_i - b_1^k \|_2^2 + \\ \mu_2 \| \boldsymbol{v} - \boldsymbol{s} + G_i x_i - b_2^k \|_2^2 + \mu_3 \| \boldsymbol{w} - \boldsymbol{D}_t\boldsymbol{x}_i - b_3^k \|_2^2 + \mu_4 \| \boldsymbol{z} - \boldsymbol{D}_s\boldsymbol{x}_i - b_4^k \|_2^2 \end{array} \right\} \tag{9}$$

where b_1, b_2, b_3, and b_4 are Bregman variables which can be updated as follow,

$$\begin{aligned} b_1^{k+1} &= b_1^k - u + \Theta x_i,\ \ b_2^{k+1} = b_2^k - v + s - G_i x_i \\ b_3^{k+1} &= b_3^k - w + D_t x_i,\ b_4^{k+1} = b_4^k - z + D_s x_i \end{aligned} \tag{10}$$

The above optimization problem can be split into the following five sub-problem,

$$P1:\ \arg\min_x \mu_1 \| u - \Theta x_i - b_1^k \|_2^2 + \mu_2 \| v - s + G_i x_i - b_2^k \|_2^2 + \mu_3 \| w - D_t x_i - b_3^k \|_2^2 + \mu_4 \| z - D_s x_i - b_4^k \|_2^2 \tag{11a}$$

$$P2: \arg\min_u \| u \|_2^2 + \mu_1 \| u - \Theta x_i - b_1^k \|_2^2 \tag{11b}$$

$$P3: \arg\min_v \alpha \| v \|_2^2 + \mu_2 \| v - s + G_i x_i - b_2^k \|_2^2 \tag{11c}$$

$$P4: \arg\min_w \lambda_1 \| w \|_p^p + \mu_3 \| w - D_t x_i - b_3^k \|_2^2 \tag{11d}$$

$$P5: \arg\min_z \lambda_2 \| z \|_p^p + \mu_4 \| z - D_s x_i - b_4^k \|_2^2 \tag{11e}$$

The general soft-threshold (GST) algorithm (Majumdar and Ward, 2012) is reported for the Lp-norm minimization problem $\arg\min_x \| y - x \|_2^2 + \lambda \| x \|_p^p$, and the expression of the GST is expressed as x=SoftTh(y,λ,p)=$sign(y)\times max\{0, |y| - \lambda^{p/2}|y|^{p-1}\}$. Therefore, in this case, sub-problem 1 can be obtained as follows,

$$x_i = \frac{\mu_1\Theta(u - b_1^k) - \mu_2 G_i(v - s - b_2^k) + \mu_3 D_t(w - b_3^k) + \mu_4 D_s(z - b_4^k)}{\mu_1\Theta^T\Theta + \mu_2 G_i^T G_i + \mu_3 D_t^T D_t + \mu_4 D_s^T D_s} \tag{12}$$

The specific derivation process for sub-problem 1 can be referred to in Appendix I.

Next, sub-problems (11b), (11c), (11d) and (11e) can be solved with the GST algorithm (Majumdar and Ward, 2012). We have,

$$u^{k+1} = \text{SoftTh}(\Theta x^{(k+1)} + b_1^k, 1/\mu_1, 2) \tag{13a}$$

$$v^{k+1} = \text{SoftTh}(s - Gx^{(k+1)} + b_2^k, \alpha/\mu_2, 2) \tag{13b}$$

$$w^{k+1} = \text{SoftTh}(D_t x^{(k+1)} + b_3^k, \lambda_1/\mu_3, p) \tag{13c}$$

TABLE 1 The proposed STNP-ACMD algorithm.

Algorithm 1. The proposed STNP-ACMD algorithm

Initiation: input signal s, parameters α, λ_1, λ_2, μ_1,μ_2,μ_3,μ_4, input the initial IF $f_i^0(t)\hat{}$ and construct the matrix G_i^0, *MaxIter N*.

For k = 1 to MaxIter N,
$x^{(k+1)} = x^{(k)}$ from Equation (12);
u^{k+1} = SoftTh($\boldsymbol{\Theta}\mathbf{x}^{(k+1)} + b_1^k, 1/\mu_1, 2$);
v^{k+1} = SoftTh($\mathbf{s} - \mathbf{G}\mathbf{x}^{(k+1)} + b_2^k, \alpha/\mu_2, 2$);
w^{k+1} = SoftTh($\mathbf{D}_t\mathbf{x}^{(k+1)} + b_3^k, \lambda_1/\mu_3, p$);
z^{k+1} = SoftTh($\mathbf{D}_s\mathbf{x}^{(k+1)} + b_4^k, \lambda_2/\mu_4, p$);
$b_1^{k+1} = b_1^k - \mathbf{u} + \boldsymbol{\Theta}\mathbf{x}^{(k+1)}$;
$b_2^{k+1} = b_2^k - \mathbf{v} + \mathbf{s} - \mathbf{G}_i\mathbf{x}^{(k+1)}$;
$b_3^{k+1} = b_3^k - \mathbf{w} + \mathbf{D}_t\mathbf{x}^{(k+1)}$;
$b_4^{k+1} = b_4^k - \mathbf{z} + \mathbf{D}_s\mathbf{x}^{(k+1)}$;
End
Update the signal component $s_i = [C_i, S_i]x_i^{(k+1)}$.

$$z^{k+1} = \mathrm{SoftTh}(D_s x^{(k+1)} + b_4^k, \lambda_2/\mu_4, p) \tag{13d}$$

Then, the demodulated signals can be estimated as,

$$s_i = G_i x_i = [C_i, S_i] x_i \tag{14}$$

Subsequently, the detailed algorithm process of the proposed STNP-ACMD is summarized in Table 1.

According to the formulation of the frequency increment of the demodulated signals (Hou and Shi, 2013) calculated in Equation (14), we have,

$$s_i = G_i x_i = [C_i, S_i] x_i \tag{15}$$

It should be noted that Algorithm 1 is only used for estimating the first signal component. To estimate other components gradually, like traditional signal decomposition methods such as empirical mode decomposition (EMD) and variational mode decomposition (VMD), the first component denoted by $s_1(t)^{\wedge}$ is removed from the original observation signal, that is,

$$R_1(t) = s(t) - \hat{s_1}(t) \tag{16}$$

where $R_1(t)$ denotes the residual signal after removing the first estimated signal from the original observation. Then, the $R_1(t)$ is treated as a new original signal, same procedures in Algorithm 1 will be repeated for extracting the second estimated signal $s_2(t)^{\wedge}$, by parity of reasoning, until the residual signal meets the predetermined threshold, such as $\rho = \Delta s^j(t) = \| s_i^j(t) - s_i^{j-1}(t) \|_2^2 / \| s(t) \|_2^2, 0 < i, j \le K$. Hence, the original observation signal $s(t)$ can be expressed as,

$$s(t) = \sum_{i=1}^{K} s_1(t)^{\wedge} + R_K(t) \tag{17}$$

where $R_K(t)$ is the residual component or additional noise. As a result, the instantaneous frequency (IF) is determined by differentiation $f(t) = \frac{1}{2\pi}\frac{d\psi(t)}{dt}$, and the increment of the IF is calculated as,

$$\begin{aligned}\Delta f_i(t)^{\wedge} &= -\frac{1}{2\pi}\frac{d}{dt}\left(\arctan\left[\frac{b_i^{n+1}(t)}{a_i^{n+1}(t)}\right]\right) \\ &= -\frac{1}{2\pi}\frac{a_i^{n+1}(b_i^{n+1})' - b_i^{n+1}(a_i^{n+1})'}{(a_i^{n+1})^2 + (b_i^{n+1})^2}\end{aligned} \tag{18}$$

Motivated by the approach in Ref (Mcneill, 2016)., the final IF can be updated as,

$$f_i^{n+1}(t) = f_i^n(t) + (\boldsymbol{Q}^{(2)}/\zeta + \boldsymbol{I})^{-1}\hat{\Delta f_i^n}(t) \tag{19}$$

where $Q^{(N)}=(\Omega^{(N)})^T\Omega^{(N)}$, $\boldsymbol{I}$ is an identity matrix, $\boldsymbol{\Omega}$ is a second-order derivative operator, ζ is small positive value.

3 Numerical simulation

A numerical simulation case is conducted to investigate the effectiveness of the proposed approach in terms of the accuracy and robustness versus the traditional ACMD (Chen et al., 2019), the VNCMD method (Chen et al., 2017), the EEMD method (Wu and Huang, 2009), and the VMD method (Dragomiretskiy and Zosso, 2014). A simulated signal that is corrupted by ambient noise is constructed as follows,

$$\begin{cases} x(t) = x_1(t) + x_2(t) + x_3(t) + x_4(t) + x_5(t) + x_6(t) \\ x_1(t) = Q_1(1 + AM_1)\exp(-0.05t)\cos\left[2\pi(2 + 200t - 50t^2 - 2t^3 - 5t^4)\right] \\ x_2(t) = 0.2Q_2(1 + AM_2)\exp(-0.08t)\cos\left[2\pi(4 + 80t + 20t^2 - 1.6t^3 + 0.15t^4)\right] \\ x_3(t) = 0.8\exp(-0.08t)\cos\left[2\pi(5 + 140t + 20t^2 - 2t^3 + 0.2t^4)\right] \\ x_4(t) = 0.5\exp(-0.2t)\cos\left[2\pi(3 + 20t + 40t^2 - 1.5t^3 + 0.4t^4)\right] \\ x_5(t) = heavy\ tailed\ noise \\ x_6(t) = sigma \times \mathrm{randn}(1, N) \end{cases} \tag{20}$$

where the varying speed profiles are Q_1=0.1(1+2t)^2+1, Q_2=0.12(1+2t)^2+2 ,the amplitude profiles are $AM_1 = 0.3\sin(\int_0^t 2\pi(1+2t)dt)$, $AM_2 = 0.3\sin(2\int_0^t 2\pi(1+2t)dt)$, $N = 512$ is the number of sampling points, the sampling frequency f_s = 512 Hz, and the sampling time is 1 s.

The simulated synthetic signal consists of five parts, both $x_1(t)$ and $x_2(t)$ are amplitude modulation frequency modulation (AM–FM) signals, $x_3(t)$ is a frequency modulation (FM) signal, $x_4(t)$ is an additive Gaussian white noise (GWN) with standard deviation *sigma* = 1, and $x_5(t)$ is a heavy tailed noise (HTN). Both GWN and HTN noises are independent with each other. Of particular interest to us is the mode or signal with small amplitude and heavy tail distribution, which is difficult to identify. Therefore, the addition of HTN noise is to create difficulty in signal estimation, artificially. The HTN is expressed with a symmetric alpha-stable (SαS) distribution with parameter $\alpha \neq 1$ and $1 < \alpha < 2$, the expression of SαS distribution is defined as (Kuruoglu et al., 1998; Liu et al., 2022),

$$\begin{aligned} x_5(t) &= S_{\alpha,\rho} \\ &\times \left\{\frac{\sin(\alpha(v + B_{\alpha,\rho}))}{(\cos v)^{1/\alpha}}\left(\frac{\cos(v - \alpha(v + B_{\alpha,\rho}))}{W}\right)^{(1-\alpha)/\alpha}\right\} \end{aligned} \tag{21}$$

where W is the exponential random variable, i.e., W~exprnd(N), random variable $v \in (-\pi/2, \pi/2)$, $B_{\alpha,\rho} = \frac{\arctan(\rho\tan\frac{\pi\alpha}{2})}{\alpha}$, $S_{\alpha,\rho} = [1 + \rho^2\tan^2\frac{\pi\alpha}{2}]^{1/(2\alpha)}$. If the $\alpha = 1$, the THN can be given as,

$$x_5(t) = \frac{2}{\pi}\left[\left(\frac{\pi}{2} + \rho v\right)\tan v - \rho\log\left(\frac{W\cos v}{\frac{\pi}{2} + \rho v}\right)\right] \tag{22}$$

The simulated signal is depicted in Figure 1A, where the AM–FM signals $x_1(t)$ and $x_2(t)$, the FM signal $x_3(t)$, the HTN

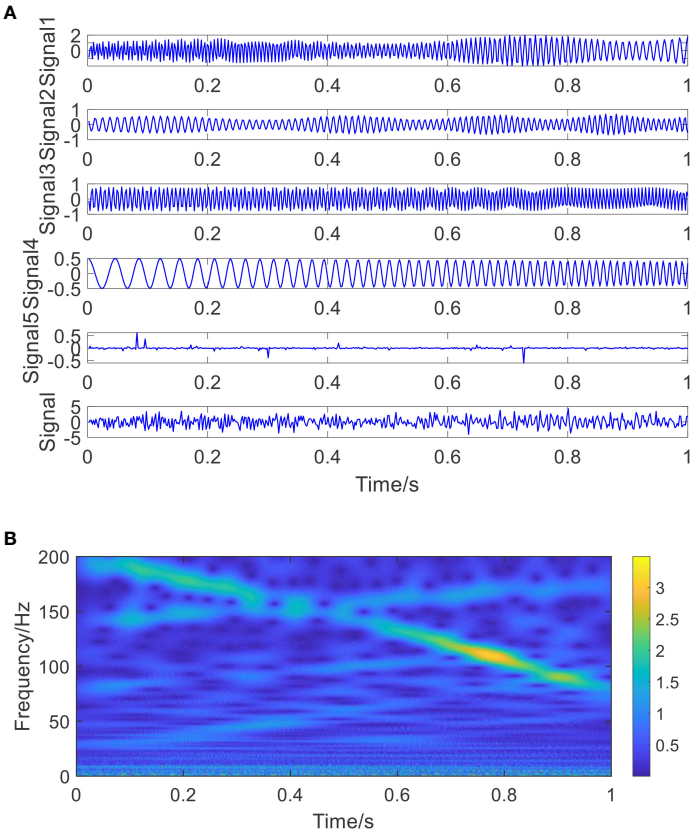


FIGURE 1
The simulated signal and its STFT diagram. **(A)** The simulated signal: from top to bottom are signal $x_1(t)$, $x_2(t)$, $x_3(t)$, $x_4(t)$, heavy tailed noise $x_5(t)$ and synthetic signal $x(t)$; **(B)** the STFT diagram of the simulated synthetic signal.

component with parameters $\alpha = 1.4$, $\rho = 1.2$ and raw synthetic signal are shown in Figure 1A from top to the bottom, respectively. It can be seen that the AM–FM and FM signals are submerged in the additional noises. The short time Fourier transform (STFT) of the simulated synthetic signal is shown in Figure 1B, it can be seen that the modulation frequencies are fuzzy and disorder and the frequency variation trends could not be detected at all in the STFT diagram.

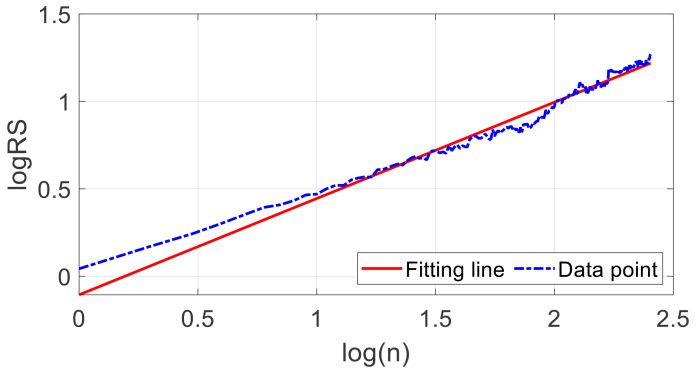


FIGURE 2
The fitted regression lines of the Hurst exponent.

The proposed STNP-ACMD algorithm is utilized to decompose the simulated synthetic signal. Herein, the model parameters are set as follows: $\alpha = 10^{-8}$, $\zeta = 10^{-9}$, $\lambda_1 = \lambda_2 = 0.01\alpha$, $\mu_1 = \mu_2 = \mu_3 = \mu_4 = 0.01$, and the threshold of signal energy is set to be 0.1. The fractional order of the simulated synthetic signal is 0.0508 using the R/S method (Hurst, 1951; Mason, 2016). The fitted regression lines of the Hurst exponent are shown in Figure 2. Note that the Hurst index of the simulated synthetic signal is a random variable and fluctuated around a value of 0.5 due to the interference from the GWN and HTN noises. By iterating, the signal energies of the first five components are 1.0854, 0.2945, 0.1054, 0.1340, and 0.0459. Thus, four estimated signal components are shown in Figure 3, since the signal energy of the last component is 0.0459<0.1. It can be observed that four estimated signal components match the theoretical signals very well. For the most part, the instantaneous frequencies of the estimated components are basically consistent with the original (or true) ones, which shows the availability of the proposed approach with respect to noise perturbation.

For comparison, the ACMD method (Chen et al., 2019), the VNCMD method (Chen et al., 2017), the EEMD method (Wu and Huang, 2009), and the VMD method (Dragomiretskiy and Zosso, 2014) are employed for decomposing the simulated signal. The estimated components generated *via* the ACMD and the VNCMD algorithms are displayed in Figures 4A, 5A, and the IFs of the estimated components generated *via* the ACMD and the VNCMD methods are shown in Figures 4B, 5B. Herein, the number of sub-components is considered and set to be four in the ACMD and VNCMD methods. It can be seen from Figure 4A that the ACMD and the VNCMD methods can slightly extract the components as the proposed does. However, the IFs of the estimated components are mussy and intricate, In the meantime, many wrong IFs are estimated from the ACMD and the VNCMD methods, as shown in Figures 4B, 5B. The main reason behind this shortcoming is that the spatio-temporal coupled information and physical cross-knowledge are not considered by the ACMD and the VNCMD methods, which is also the main limitation of the ACMD and the VNCMD methods in practical applications. It can be concluded that the cross-knowledge or coupled information of the IFs of four pre-estimated components can be demodulated or isolated by the proposed approach, but the ACMD and the VNCMD methods failed.

Additionally, the estimated components generated *via* the EEMD and the VMD methods are displayed in Figures 6A, C, and the instantaneous frequencies of the estimated components generated *via* the EEMD and the VMD methods are shown in Figures 6B, D, respectively. As shown in Figure 6, many spurious IMF components without physical meaning are separated, and the Hilbert Huang Transform (HHT) cannot reveal the actual IF patterns of the estimated components, as shown in Figures 6B, D, which means the EEMD and the VMD methods still suffer from severe issues with the HHT, i.e., models aliasing. The comparison results reveal that the hidden components containing crossover frequencies can be separated and estimated one by one using the proposed STNP-ACMD algorithm. The crossover IFs are much closer approximations to the real ones.

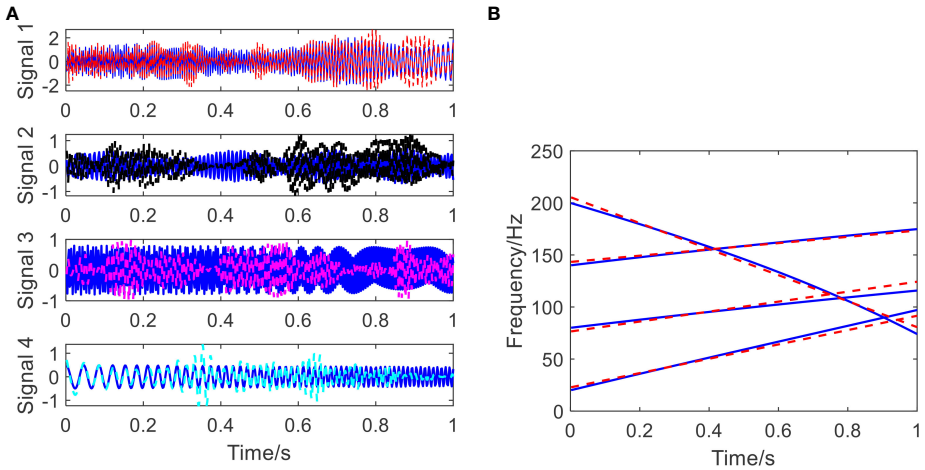


FIGURE 3
The estimated signal components and estimated time frequency plot using the proposed approach. **(A)** The estimated components (blue: true; red, black, carmine, and cyan: estimated ones); **(B)** The estimated time frequency plot (blue: true; red: estimated).

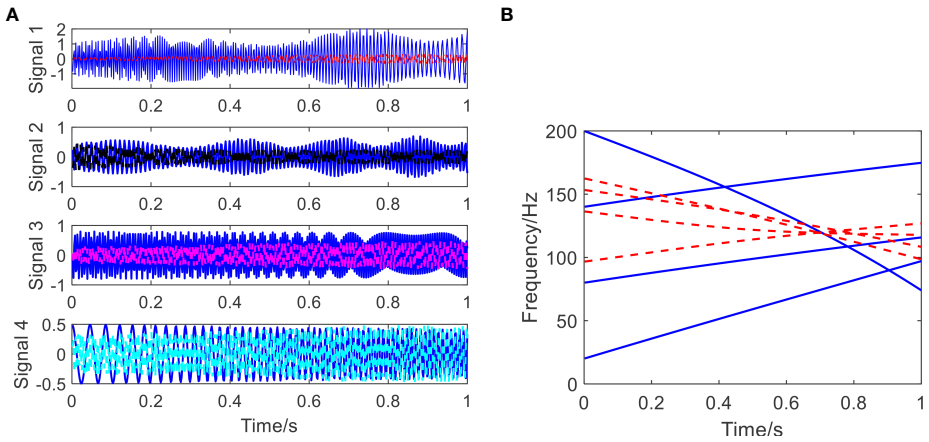


FIGURE 4
The estimated signal components and estimated time frequency plot using the ACMD method. **(A)** The estimated components (blue: true; red, black, carmine and cyan: estimated ones); **(B)** the estimated time frequency plot (blue: true; red: estimated).

4 Experimental case of real-world applications

The simulated case in *Section 3* indicates that the proposed STNP-ACMD algorithm has promising advantages in analyzing simulated signals with crossover instantaneous frequencies. Therefore, in this section, the proposed STNP-ACMD algorithm will be used to extract the signal components and their crossover IFs from seafloor dynamic data and further achieve the exploration of the tidal law and magnetotelluric dynamic analysis in deep seafloor.

4.1 Introduction of experimental datasets

The experimental datasets of seafloor dynamic data are provided by the Institute of Acoustics, Chinese Academy of Sciences (Chang et al., 2019). The diagram of the deep seafloor observation network system in the South China Sea and the observation position of the submarine dynamic platform (represented by the red flag, with 111.0675°E and 17.5811°N) are shown in Figures 7A, B. In this experiment, the acoustic Doppler current profilers (ADCP) of type Teledyne were used for collecting the observation elements including seafloor temperature, seafloor water salinity, seafloor water flow rate #1, seafloor water flow rate

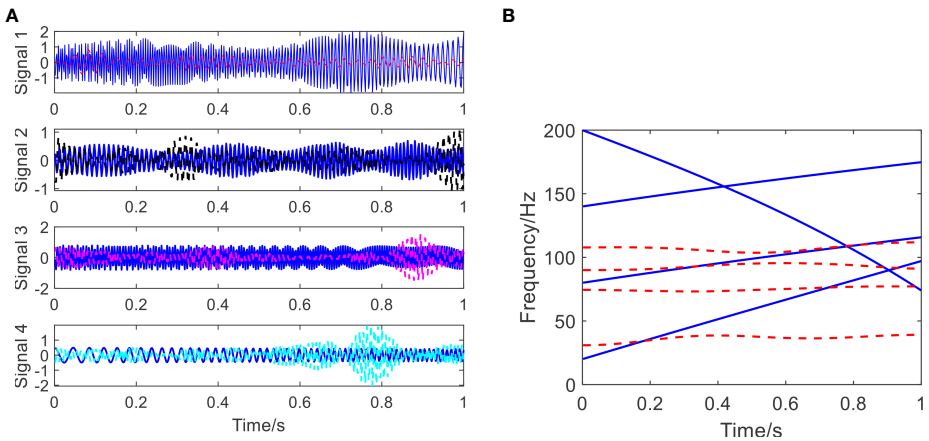


FIGURE 5
The estimated signal components and estimated time frequency plot using the VNCMD method. **(A)** The estimated components (blue: true; red, black, carmine and cyan: estimated ones); **(B)** the estimated time frequency plot (blue: true; red: estimated).

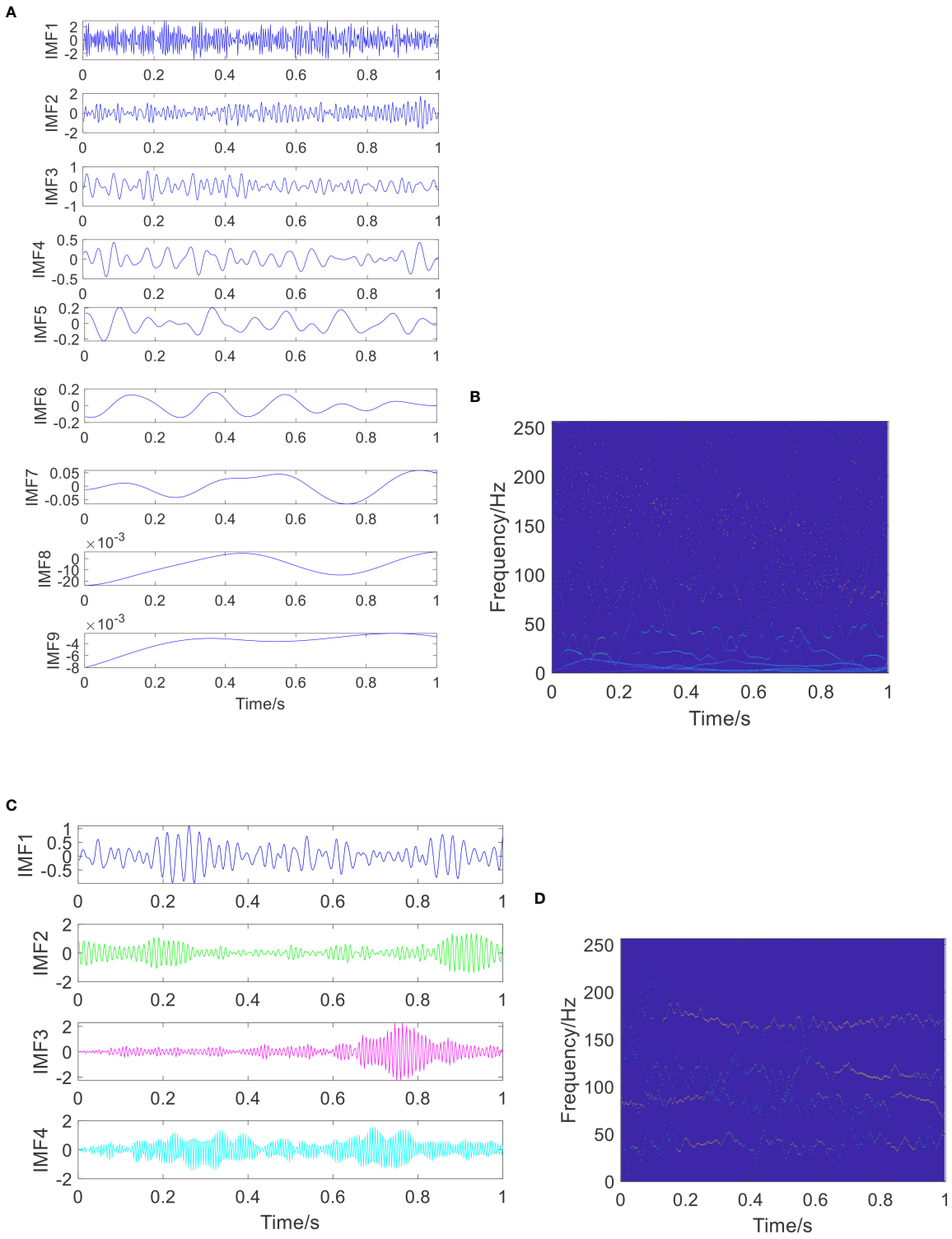


FIGURE 6
The estimated signal components and their estimated time-frequency plots using the EEMD and the VMD methods. **(A)** The estimated components using the EEMD method; **(B)** the estimated time–frequency plot of the estimated components using the EEMD method; **(C)** the estimated components using the VMD method; and **(D)** the estimated time-frequency plot of the estimated components using the VMD method.

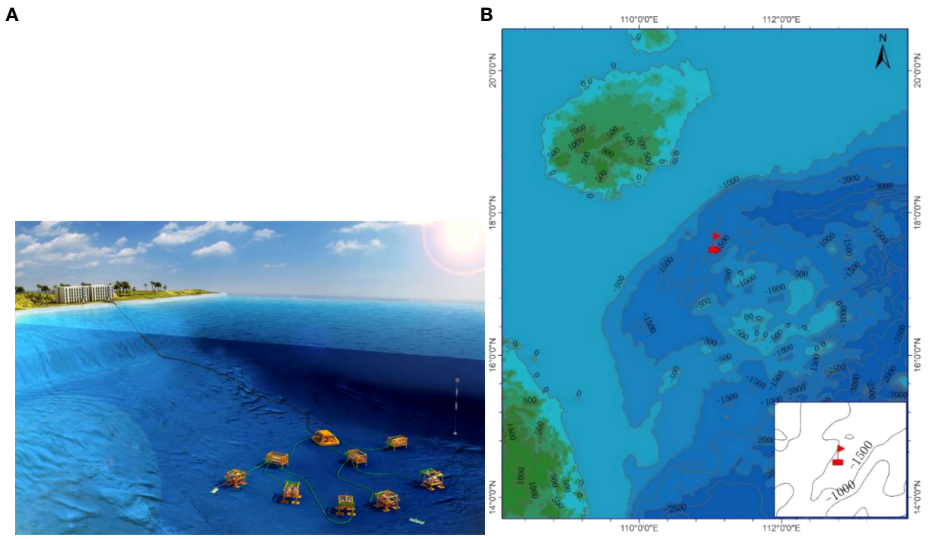


FIGURE 7
The deep seafloor observation network system and observation position (Chang et al., 2019). **(A)** The diagram of the deep seafloor observation network system in the South China Sea; **(B)** the observation position of submarine dynamic platform (represented by red flag).

#2, seafloor water flow rate #3, seafloor water flow rate #4, roll data, bow data, and tilt data. Those datasets were processed with the following operations: autocompletion of default value, verification of data threshold interval, etc.

4.2 Results analysis

In this work, the seafloor water flow rates #1 with 2,048 points are employed. The operating frequency of ADCP was 150

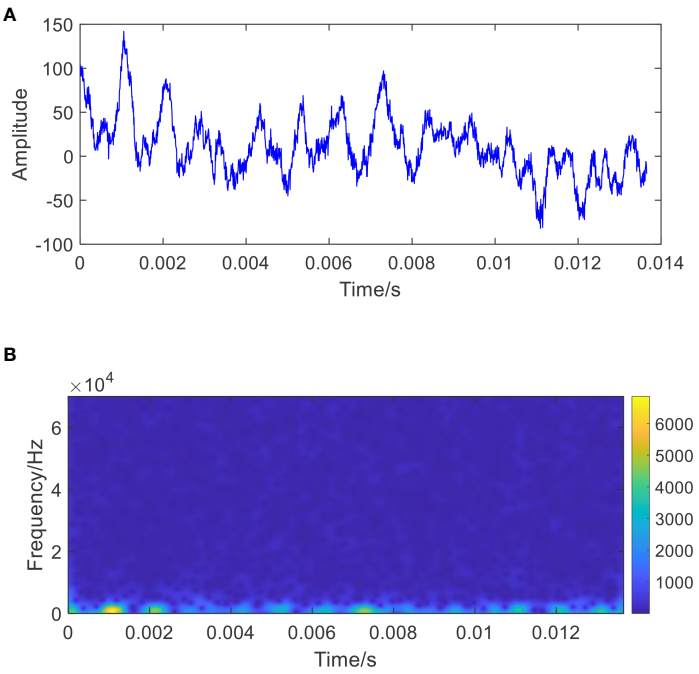


FIGURE 8
Ocean water velocity and its time frequency diagram. **(A)** Ocean water velocity of deep seabed collected by acoustic Doppler current-meter sensor; **(B)** the time frequency diagram of ocean water velocity.

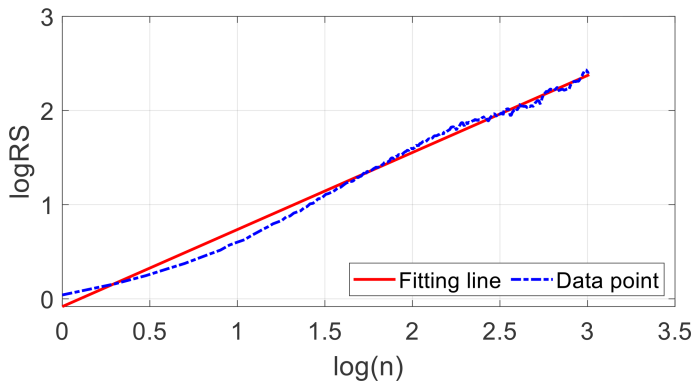


FIGURE 9
The fitted regression lines of the Hurst exponent.

TABLE 2 The signal energy of each estimated component.

Signal components	Signal 1	Signal 2	Signal 3	Signal 4	Signal 5	Signal 6	Residual Signal
Energy	15.2202	0.2643	0.0688	0.0375	0.0334	0.0324	0.0270

kHz. The number of observation layers was 30, each layer was 4 m, and sampling was conducted at an interval of 10 min (Chang et al., 2019). The raw signal waveform and its time frequency diagram of seafloor water flow rates #1 with 2,048 points are shown in Figure 8. It can be seen from Figure 8A that high frequency vibration components are accompanied by a slowly decreasing low frequency trend component, and the frequency information of 2,048 points is mainly concentrated in the low-

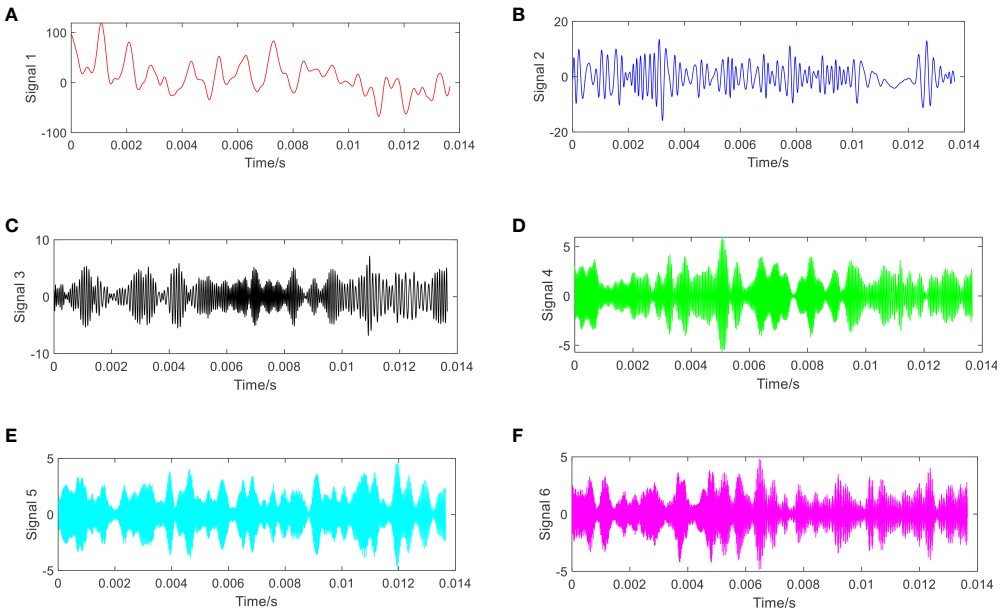


FIGURE 10
The decomposed signal components using the proposed approach. **(A)** Component 1; **(B)** component 2; **(C)** component 3; **(D)** component 4; **(E)** component 5; and **(F)** component 6.

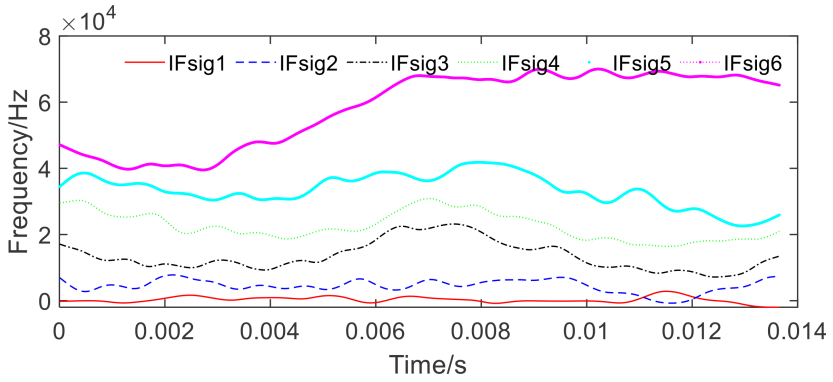


FIGURE 11
The TF diagram of the estimated components using the proposed approach.

frequency band. No obvious instantaneous frequency trends can be found to track the variation of the velocity in the deep seafloor, as shown in Figure 8B.

The proposed STNP-ACMD algorithm is applied to estimate the sub-components and their IFs, where the threshold of signal energy is set to be 0.03. Due to the disturbance of fluctuating trends, detrended and normalization operations were performed on the raw seafloor water flow rates #1 with 2,048 points, and then the fractional order of the simulated synthetic signal is 0.3182 using the R/S method (Hurst, 1951; Mason, 2016). The fitted regression lines of the Hurst exponent are shown in Figure 9, and the signal energies of the first six components and the residual are summarized in Table 2. Therefore, the first six estimated

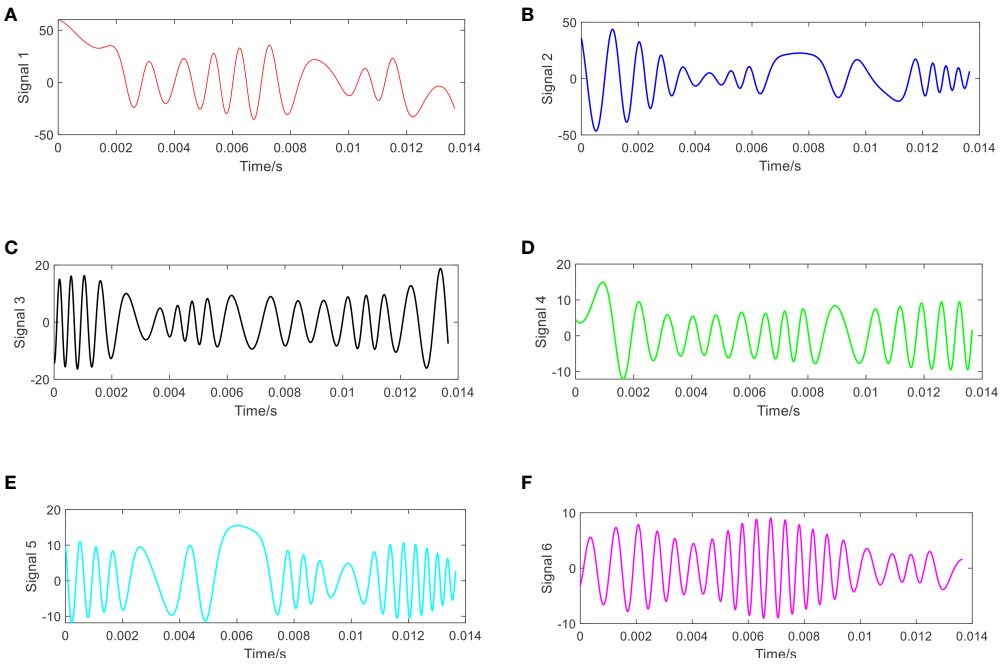


FIGURE 12
The estimated components using the ACMD method. **(A)** Component 1; **(B)** component 2; **(C)** component 3; **(D)** component 4; **(E)** component 5; and **(F)** component 6.

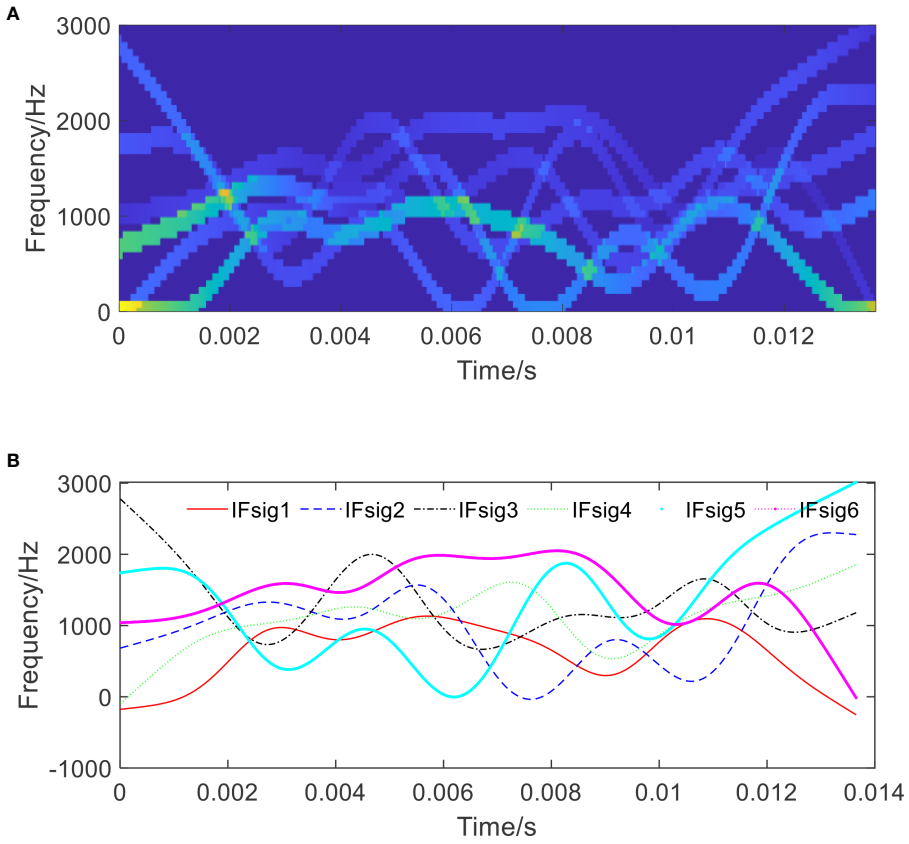


FIGURE 13
The TF diagram and IF plot of the estimated components using the ACMD method. **(A)** The TF diagram of the estimated components; and **(B)** the IF plot of the estimated components.

components are selected and presented in Figure 10. It can be found that the waveform pattern and fluctuating trend of signal 1 are the same as the raw seafloor water flow rates (see Figure 8A). Meanwhile, as demonstrated in Table 2, five high-frequency components, i.e., from signals 2 to 6, with lower energy are also extracted and thus contain more valuable information with respect to interference from other factors such as fauna and plankton. The TF diagram of the estimated components using the proposed approach is shown in Figure 11. It shows that the crossover instantaneous frequencies between signals 1 and 2 can be extracted clearly. Those features above indicate that the patterns of low-frequency trend and high-frequency fluctuation of the raw seafloor water flow rate can be captured and accurately estimated using the proposed approach.

To investigate the superiority of the proposed approach, four benchmarks, including the ACMD method (Chen et al., 2019), the VNCMD method (Chen et al., 2017), the EEMD method (Wu and Huang, 2009), and the VMD method (Dragomiretskiy and Zosso, 2014), are introduced for comparisons. The estimated components generated *via* the ACMD and the VNCMD methods are displayed in Figures 12, 14A, and the IFs of the estimated components generated *via* the ACMD and the VNCMD methods are shown in Figures 13, 14B. Herein, the number of the decomposed components is determined to be 6 in the ACMD and the VNCMD methods, according to the proposed approach. As shown in Figure 12, it shows that the waveform of signal 1, estimated using the ACMD method, is slightly coincident with that of raw seafloor water flow rate, but the other five components still belong to low-frequency components (see Figures 12B–F), and all of the IFs are aliased together, simultaneously. As shown in Figure 14, the whole estimated components that are generated from the VNCMD method are similar to the waveform of the original signal, and the frequency characteristics of the other components cannot be revealed at all. In addition, the same to the results of the ACMD method, all of the instantaneous frequencies are aliased together, simultaneously. That is to say, the high-frequency information of the raw signal is not captured using the ACMD

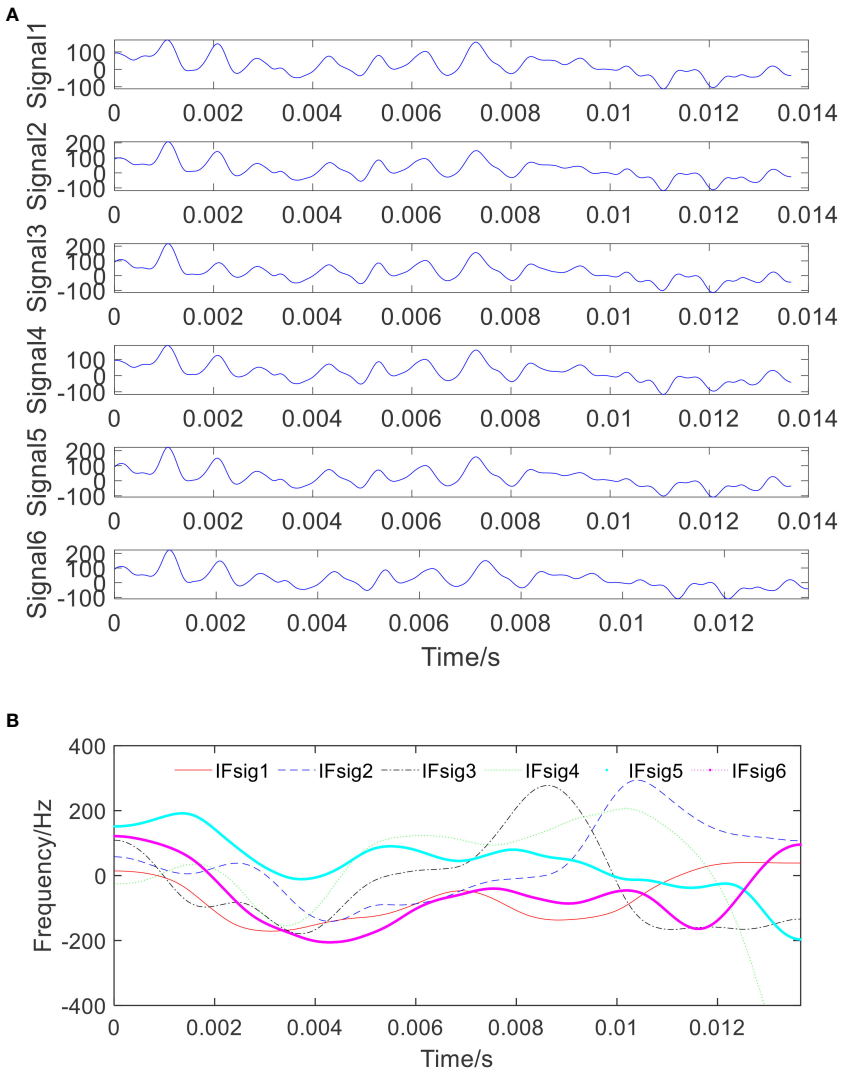


FIGURE 14
The estimated signal components and their estimated time-frequency plots using the VNCMD method. **(A)** The estimated components using the VNCMD method; and **(B)** the estimated time-frequency plot of the estimated components using the VNCMD method.

and the VNCMD methods, which is contrary to the characteristics of the raw signal.

In addition, the estimated components and their IFs generated *via* the EEMD and VMD methods are shown in Figure 15. It shows that the sub-components can be extracted and decomposed through the EEMD and VMD methods from the high frequency to the low frequency. Obviously, the waveform trend of the low frequency cannot also be captured by the EEMD and VMD methods because the wideband modes are still overlapping in the frequency domain. Meanwhile, the modes exhibit poor energy concentration in time–frequency representations and also present a poor correlation with the raw signal of seafloor water flow rates. Therefore, it can be concluded that the EEMD and VMD methods cannot reveal the time-frequency pattern of the signal with crossing regions, and the proposed method can achieve more reliable estimated results and clear fluctuating patterns of the IFs in real-world applications.

5 Conclusions

In this paper, a novel spatio-temporal nonconvex penalty adaptive chirp mode decomposition (STNP-ACMD) algorithm is proposed for recompositing the complicated multi-component whose mono-modes have crossed IFs or overlapped frequency information. The spatio-temporal

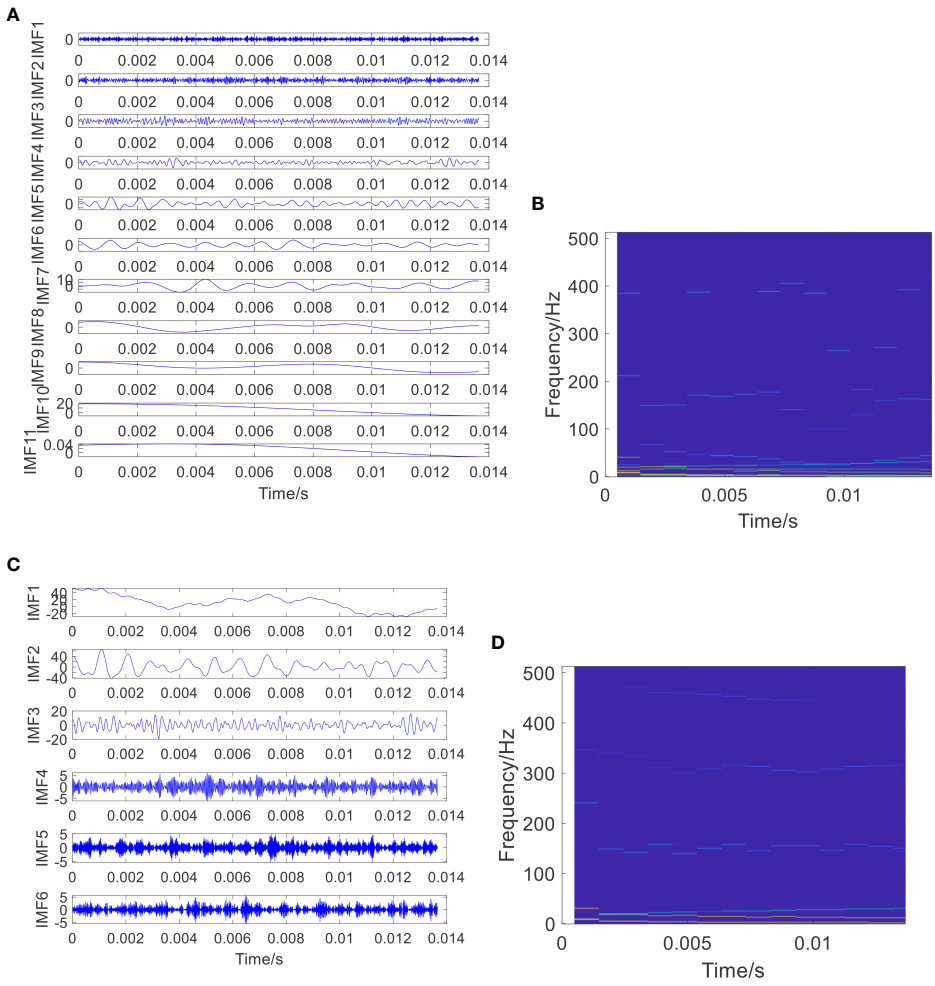


FIGURE 15
The estimated signal components and their estimated time–frequency plots using the EEMD and VMD methods. **(A)** The estimated components using the EEMD method; **(B)** the estimated time–frequency plot of the estimated components using the EEMD method; **(C)** the estimated components using the VMD method; and **(D)** the estimated time–frequency plot of the estimated components using the VMD method.

relationship among estimated components is considered *via* incorporating a nonconvex penalty term into the cost function in terms of a temporal matrix and a spatial matrix. The rationale for exploiting the coupled nature of the available information within the estimated components is given in detail. Meanwhile, the cost function can be effectively solved by the split Bregman algorithm, and the characteristics of the fractional-order of the multi-component are also considered. The results of synthetic simulated cases and specific experimental cases are introduced between the proposed approach and the state-of-the-art benchmarks and indicate that the proposed STNP-ACMD has higher decomposed accuracy and superiority in component estimation and IFs in terms of crossed IFs scenarios.

However, it should be highlighted that due to the fact that the real IF information is unknown, the estimated time–frequency ridge cannot be verified in practical application of seafloor dynamic engineering. The STNP-ACMD method proposed in this work is a deliberate attempt and a promising exploration. In addition, one of the disadvantages in the work focuses on parameter setting. Arbitrary parameter settings are taken in this work, which lacks robustness for time-varying scenarios. Therefore, future investigation will be devoted to enhancing the robustness of the proposed method, including parameter optimization and noise detection, and the multi-dimensional and complicated time-varying scenarios in industrial areas and ocean engineering applications will also be explored.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding author.

Author contributions

The author confirms being the sole contributor of this work and has approved it for publication.

Funding

This work is supported by the National Natural Science Foundation of China (No. 52205079), the Foundation of High-level Talents (No. RC412104), the Natural Science Research Project of Universities in Anhui Province (No. KJ2021A0155), and the State Key Laboratory of Mechanical System and Vibration (No. MSV202015). The insightful comments and profound suggestions from reviewers are greatly appreciated.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Auger, F., Flandrin, P., Lin, Y. T., McLaughlin, S., Meignen, S., Oberlin, T., et al. (2013). Time-frequency reassignment and synchrosqueezing: an overview. *IEEE Signal Process Mag.* 30 (6), 32–41. doi: 10.1109/MSP.2013.2265316

Chang, Y. G., Zhang, F., Guo, Y. G., Song, X. Y., Yang, J., and Liu, R. Y. (2019). The ocean dynamic datasets of seafloor observation network experiment system at the south China sea. *China Sci. Data* 4 (4):48-55. doi: 10.11922/sciencedb.823

Chen, Q. M., Chen, J. H., Lang, X., Xie, L., Lu, S., and Su, H. Y. (2020). Detection and diagnosis of oscillations in process control by fast adaptive chirp mode decomposition. *Control Eng. Pract.* 97, 104307. doi: 10.1016/j.conengprac.2020.104307

Chen, S. Q., Dong, X. J., Peng, Z. K., Zhang, W. M., and Meng, G. (2017). Nonlinear chirp mode decomposition: A variational method. *IEEE T. Signal Process.* 65 (22), 6024–6037. doi: 10.1109/TSP.2017.2731300

Chen, S. Q., Yang, Y., Peng, Z. K., Dong, X. J., Zhang, W. M., and Meng, G. (2019). Adaptive chirp mode pursuit: algorithm and applications. *Mech. Syst. Signal Process.* 116, 566–584. doi: 10.1016/j.ymssp.2018.06.052

Chen, S. Q., Yang, Y., Peng, Z. K., Wang, S. B., Zhang, W. M., and Chen, X. F. (2019). Detection of rub-impact fault for rotor-stator systems: a novel method based on adaptive chirp mode decomposition. *J. Sound Vib.* 440, 83–99. doi: 10.1016/j.jsv.2018.10.010

Chen, K., Zhao, Q. X., Deng, M., Luo, X. H., and Jing, J. E. (2020). Seawater motion-induced electromagnetic noise reduction in marine magnetotelluric data using current meters. *Earth Planets Space* 72, (4). doi: 10.1186/s40623-019-1129-0

Clausel, M., Oberlin, T., and Perrier, V. (2015). The monogenic synchrosqueezed wavelet transform: a tool for the decomposition/demodulation of AM-FM images. *Appl. Comput. Harmon. Anal.* 39, 450–486. doi: 10.1016/j.acha.2014.10.003

Corsaro, S., Simone, V. D., and Marino, Z. (2021). Split bregman iteration for multi-period mean variance portfolio optimization. *Appl. Math. Comput.* 392, 125715. doi: 10.1016/j.amc.2020.125715

Daubechies, I., Lu, J. F., and Wu, H. T. (2011). Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool. *Appl. Comput. Harmon. A* 30 (2), 243–261. doi: 10.1016/j.acha.2010.08.002

Dong, H. R., Yu, G., and Li, Y. Y. (2022). Theoretical analysis and comparison of transient-extracting transform and time-reassigned synchrosqueezing transform. *Mech. Syst. Signal Process.* 178, 109190. doi: 10.1016/j.ymssp.2022.109190

Dragomiretskiy, K., and Zosso, D. (2014). Variational mode decomposition. *IEEE T. Signal Process* 62 (3), 531–544. doi: 10.1109/TSP.2013.2288675

Gilles, J. (2013). Empirical wavelet transform. *IEEE T. Signal Process.* 61 (16), 3999–4010. doi: 10.1109/TSP.2013.2265222

Goldstein, T., and Osher, S. (2009). The split bregman method for L1 regularized problems. *SIAM J. Imag. Sci.* 2, 323–343. doi: 10.1137/080725891

Hou, T. Y., and Shi, Z. (2013). Data-driven time frequency analysis. *Appl. Comput. Harmon. Anal.* 35 (2), 284–308. doi: 10.1016/j.acha.2012.10.001

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., et al. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. R. Soc. A: Mathematical Phys. Eng. Sci.* 454,1971 903–995. doi: 10.1098/rspa.1998.0193

Hurst, H. E. (1951). Long-term storage capacity reservoirs. *Trans. Am. Soc Civil Eng.* 116, 770–808. doi: 10.1061/TACEAT.0006518

Jiang, Y., Liu, S. Y., Zhao, N., Xin, J. Z., and Wu, B. (2020). Short-term wind speed prediction using time varying filter-based empirical mode decomposition and group method of data handling-based hybrid model. *Energ. Convers. Manage.* 220, 113076. doi: 10.1016/j.enconman.2020.113076

Kuruoglu, E. E., Fitzgerald, W. J., and Rayner, P. J. W. (1998). Near optimal detection of signals in impulsive noise modeled with a symmetric /spl alpha/-stable distribution. *IEEE Commun. Lett.* 2 (10), 282–284. doi: 10.1109/4234.725224

Li, R., Chen, H., Fang, Y. X., Hu, Y., Chen, X. P., and Li, J. (2022). Synchrosqueezing polynomial chirplet transform and its application in tight sandstone gas reservoir identification. *IEEE Geosci. Remote S.*, 19:1–5. doi: 10.1109/LGRS.2021.3071318

Lin, Y., Lin, Z. X., Liao, Y., Li, Y. Z., Xu, J. L., and Yan, Y. (2022). Forecasting the realized volatility of stock price index: A hybrid model integrating CEEMDAN and LSTM. *Expert Syst. Appl.* 206, 117736. doi: 10.1016/j.eswa.2022.117736

Liu, X., Du, L., and Xu, S. (2022). GLRT-based coherent detection in sub-Gaussian symmetric alpha-stable clutter. *IEEE Geosci. Remote S. 19* 1-5, 8015405. doi: 10.1109/LGRS.2021.3094847

Li, L., Yu, X. R., Jiang, Q. T., Zang, B., and Jiang, L. (2022). Synchrosqueezing transform meets α-stable distribution: An adaptive fractional lower-order SST for instantaneous frequency estimation and non-stationary signal recovery. *Signal Process.* 201, 108683. doi: 10.1016/j.sigpro.2022.108683

MacLennan, K., and Li, Y. G. (2013). Denoising multicomponent CSEM data with equivalent source processing techniques. *Geophysics* 78 (3), 125–135. doi: 10.1190/geo2012-0226.1

Majumdar, A., and Ward, R. K. (2012). On the choice of compressed sensing priors and sparsifying transforms for MR image reconstruction: An experimental study. *Signal Process. Image Commun.* 27 (9), 1035–1048. doi: 10.1016/j.image.2012.08.002

Mason, D. M. (2016). The Hurst phenomenon and the rescaled range statistic. *Stoch. Proc. Appl.* 126, (12) 3790–3807. doi: 10.1016/j.spa.2016.04.008

Mcneill, S. I. (2016). Decomposing a signal into short-time narrow-banded modes. *J. Sound Vib.* 373, 325e339. doi: 10.1016/j.jsv.2016.03.015

Meignen, S., Pham, D. H., and Colominas, M. A. (2021). On the use of short-time fourier transform and synchrosqueezing-based demodulation for the retrieval of the modes of multicomponent signals. *Signal Process.* 178, 107760. doi: 10.1016/j.sigpro.2020.107760

Nyqvist, D., Durif, C., Johnsen, M. G., Jong, K. D., Forland, T. N., and Sivle, L. D. (2020). Electric and magnetic senses in marine animals, and potential behavioral effects of electromagnetic surveys. *Mar. Environ. Res.* 155, 104888. doi: 10.1016/j.marenvres.2020.104888

Park, D., Kwak, K., Kim, J., Li, J. H., and Chung, W. K. (2019). Underwater localization using received signal strength of electromagnetic wave with obstacle penetration effects. *IFAC-Papers Online* 52 (21), 372–377. doi: 10.1016/j.ifacol.2019.12.335

Schwalenberg, K., Gehrmann, R. A. S., Bialas, J., and Rippe, D. (2020). Analysis of marine controlled source electromagnetic data for the assessment of gas hydrates in the Danube deep-sea fan, black Sea. *Mar. Petrol. Geol.122*, 104650. doi: 10.1016/j.marpetgeo.2020.104650

Si, Y., Kong, L. F., Chin, J. H., Guo, W. C., and Wang, Q. L. (2022). Whirling detection in deep hole drilling process based on multivariate synchrosqueezing transform of orthogonal dual-channel vibration signals. *Mech. Syst. Signal Process.* 167, 108621. doi: 10.1016/j.ymssp.2021.108621

Tian, Y. W., Liu, M. Q., Zhang, S. L., and Zhou, T. (2022). Underwater multi-target passive detection based on transient signals using adaptive empirical mode decomposition. *Appl. Acoust.* 190, 108641. doi: 10.1016/j.apacoust.2022.108641

Torres, M. E., Colominas, M. A., Schlotthauer, G., and Flandrin, P. (2011). "A complete ensemble empirical mode decomposition with adaptive noise," in *IEEE Int. Conf. on Acoust., Speech and Signal Proc. ICASSP, Prague, Czech Republic.* 4144–4147 doi: 10.1109/ICASSP.2011.5947265

Vashishtha, G., Chauhan, S., Yadav, N., Kumar, A., and Kumar, R. (2022). And tangent entropy in estimation of single-valued neutrosophic cross-entropy for detecting impeller defects in centrifugal pump. *Appl. Acoust.* 197, 108905. doi: 10.1016/j.apacoust.2022.108905

Wu, Z., and Huang, N. (2009). Ensemble empirical mode decomposition: a noise-assisted data analysis method. *Adv. Data Sci. Adadp.* 1, 1–41. doi: 10.1142/S1793536909000047

Zhang, P. F., Deng, M., Jing, J. E., and Chen, K. (2020). Marine controlled-source electromagnetic method data de-noising based on compressive sensing. *J. Appl. Geophys.* 177, 104011. doi: 10.1016/j.jappgeo.2020.104011

Appendix

For sub-problem $P1$, the specific derivation process for signal x is given as,

$$
\begin{aligned}
F(x) &= \arg\min_{x} \mu_1 \left\| u - \Theta x_i - b_1^k \right\|_2^2 + \mu_2 \left\| v - s + G_i x_i - b_2^k \right\|_2^2 + \\
&\mu_3 \left\| w - D_t x_i - b_3^k \right\|_2^2 + \mu_4 \left\| z - D_s x_i - b_4^k \right\|_2^2 \\
&= \mu_1 \left[(u - b_1^k)^2 - 2\Theta x_i (u - b_1^k) + \Theta^T \Theta x_i^2 \right] + \\
&\mu_2 \left[(v - s - b_2^k)^2 + 2G_i x_i (v - s - b_2^k) + G_i^T G_i x_i^2 \right] + \\
&\mu_3 \left[(w - b_3^k)^2 - 2D_t x_i (w - b_3^k) + D_t^T D_t x_i^2 \right] + \\
&\mu_4 \left[(z - b_4^k)^2 - 2D_s x_i (z - b_4^k) + D_s^T D_s x_i^2 \right]
\end{aligned}
\tag{A1}
$$

$$
\begin{aligned}
\frac{\partial F(x)}{\partial x} &= -2\mu_1 \Theta (u - b_1^k) + 2\mu_1 \Theta^T \Theta x_i + 2\mu_2 G_i (v - s - b_2^k) + 2\mu_2 G_i^T G_i x_i \\
&-2\mu_3 D_t (w - b_3^k) + 2\mu_3 D_t^T D_t x_i - 2\mu_4 D_s (z - b_4^k) + 2\mu_4 D_s^T D_s x_i
\end{aligned}
\tag{A2}
$$

Let's $\frac{\partial F(x)}{\partial x} = 0$, that is,

$$
\begin{aligned}
&2\mu_1 \Theta^T \Theta x_i + 2\mu_2 G_i^T G_i x_i + 2\mu_3 D_t^T D_t x_i + 2\mu_4 D_s^T D_s x_i \\
&= 2\mu_1 \Theta (u - b_1^k) - 2\mu_2 G_i (v - s - b_2^k) + 2\mu_3 D_t (w - b_3^k) + 2\mu_4 D_s (z - b_4^k)
\end{aligned}
\tag{A3}
$$

$$
x_i = \frac{\mu_1 \Theta (u - b_1^k) - \mu_2 G_i (v - s - b_2^k) + \mu_3 D_t (w - b_3^k) + \mu_4 D_s (z - b_4^k)}{\mu_1 \Theta^T \Theta + \mu_2 G_i^T G_i + \mu_3 D_t^T D_t + \mu_4 D_s^T D_s}
\tag{A4}
$$

frontiers | Frontiers in Marine Science

TYPE Review
PUBLISHED 26 October 2022
DOI 10.3389/fmars.2022.997831

OPEN ACCESS

EDITED BY
Junyu He,
Zhejiang University, China

REVIEWED BY
Kathryn Margaret McMahon,
Edith Cowan University, Australia
Amrit Kumar Mishra,
The University of Hong Kong,
Hong Kong SAR, China

*CORRESPONDENCE
Chiara M. Bertelli
c.m.bertelli@swansea.ac.uk

†These authors have contributed equally to this work and share first authorship

SPECIALTY SECTION
This article was submitted to Ocean Observation, a section of the journal Frontiers in Marine Science

RECEIVED 19 July 2022
ACCEPTED 10 October 2022
PUBLISHED 26 October 2022

CITATION
Bertelli CM, Stokes HJ, Bull JC and Unsworth RKF (2022) The use of habitat suitability modelling for seagrass: A review. *Front. Mar. Sci.* 9:997831. doi: 10.3389/fmars.2022.997831

The use of habitat suitability modelling for seagrass: A review

Chiara M. Bertelli[1*†], Holly J. Stokes[1†], James C. Bull[1] and Richard K. F. Unsworth[1,2]

[1]Faculty of Science and Engineering, Swansea University, Swansea, United Kingdom, [2]Project Seagrass, Bridgend, United Kingdom

Coastal ecosystems, including coral reefs, mangroves, and seagrass, are in global decline. Mitigation approaches include restoration and other managed recovery interventions. To maximise success, these should be guided by an understanding of the environmental niche and geographic limits of foundational species. However, the choices of data, variables, and modelling approaches can be bewildering when embarking on such an exercise, and the biases associated with such choices are often unknown. We reviewed the current available knowledge on methodological approaches and environmental variables used to model and map habitat suitability for coastal ecosystems. While our focus is on seagrass, we draw on information from all marine macrophyte studies for greater coverage of approaches at different scales around the world. We collated 75 publications, of which 35 included seagrasses. Out of all the publications, we found the most commonly used predictor variables were temperature (64%), bathymetry (61%), light availability (49%), and salinity (49%), respectively. The same predictor variables were also commonly used in the 35 seagrass Habitat Suitability Models (HSM) but in the following order: bathymetry (74%), salinity (57%), light availability (51%), and temperature (51%). The most popular method used in marine macrophyte HSMs was an ensemble of models (29%) followed by MaxEnt (17%). Cross-validation was the most commonly used selection procedure (24%), and threshold probability was the favoured model validation (33%). Most studies (87%) did not calculate or report uncertainty measures. The approach used to create an HSM was found to vary by location and scale of the study. Based upon previous studies, it can be suggested that the best approach for seagrass HSM would be to use an ensemble of models, including MaxEnt along with a selection procedure (Cross-validation) and threshold probability to validate the model with the use of uncertainty measures in the model process.

KEYWORDS

habitat suitability modelling (HSM), seagrass, macrophyte, restoration, ensemble, Maxent (maximum entropy)

Introduction

To manage environmental changes and aid conservation of biodiversity, it is vital that we increase our understanding of the relationships between species distribution and their surrounding environment (Bellamy et al., 2013; Matthiopoulos et al., 2020). However, there is often a lack of systematic biological survey data available (Elith et al., 2011) making it impossible to account for all individuals of a species at all times. Recent economic crises have seen major cuts in government spending with environmental protection agencies one of the sectors to suffer, exacerbating this lack of systematic biological data collection (Borja & Elliott, 2013). Habitat suitability models (HSMs) and species distribution models (SDMs) can be used to predict the likelihood of species occurrence based on an understanding of the environmental variables that determine species distribution (Hirzel & Le Lay, 2008; Elith et al., 2011), so that understanding of species-habitat associations can be developed even when biological datasets are limited. HSMs are relied upon for species reintroduction (Adhikari et al., 2012), species management (Vinagre et al., 2006) and restoration projects (Barnes et al., 2007) and can be developed without the need for comprehensive biological datasets. More recently these models have been used to forecast the likely responses to future disturbances and climate change (Jueterbock et al., 2016; Davis et al., 2021), including changes in temperature, ocean acidification and sea level (Convertino et al., 2012; Valle et al., 2014; Wesselmann et al., 2021).

Global decline, resulting from multiple, often synergistic pressures is particularly acute in coastal ecosystems, including coral reefs, mangroves, salt marshes, and seagrass (Waycott et al., 2009; Silliman, 2014; Spivak et al., 2019; Stafford et al., 2021). Typical of these, seagrasses are distributed globally and are among the most ecologically valuable habitats, providing numerous ecosystem services, including carbon sequestration, sediment stabilisation, support for biodiversity, nursery grounds for commercial species, and water purification (Fourqurean et al., 2012; Nordlund et al., 2016; Kerr, 2017; Röhr et al., 2018, Unsworth et al., 2019b) yet these ecosystems are in global decline and face threats from numerous natural and anthropogenic factors (Waycott et al., 2009).

To combat the impacts of coastal ecosystem decline, restoration efforts are increasing in the marine environment, especially of coral reef, mangrove, but also seagrass habitats (Kairo et al., 2009; Guarnieri et al., 2019; Gamble et al., 2021). Seagrass restoration projects have been attempted globally but with varying degrees of success (Marion & Orth 2010; van Katwijk et al., 2016; Unsworth et al., 2019a; Tan et al., 2020), partly due to the dynamic environment seagrasses are often found to grow (McDonald et al., 2016). Large-scale restoration efforts lead to more effective results with a higher proportion of seagrass survival and increases in population growth rate (van Katwijk et al., 2016). Furthermore, ecosystem services have been seen to return following seagrass restoration (Reynolds et al., 2016), in particular where this has been achieved at scale (Orth et al., 2020). However, the use of HSM for informing the restoration of habitats is relatively unexplored, with the majority of marine studies until recently focused on theoretical ecology, movement of non-native species, conservation planning and climate change (Robinson et al., 2017). There is, therefore, the need to understand the environmental niche required for successful seagrass restoration, in order to map that onto suitable habitat across multiple locations which is the motivation of this review of wider marine macrophyte habitat suitability modelling approaches.

The most important variables in predicting habitat suitability can be influenced by the scale of the study, for example, Valle et al. (2013) found that wave exposure and current velocity were the key variables predicting *Zostera marina* distribution when using broadly distributed data, whereas slope and depth were important at predicting species distribution of conservative presence areas. This will have implications for the creation of HSMs for informing restoration and may require the need for different models and variables dependent on the scale, local topography, and location in question.

For the context of seagrass restoration, it should be considered that the environmental requirements for seeds to germinate and seedlings to be able to establish may differ from those of established mature plants that are able to reduce water-flow rates, sediment resuspension and have considerable anchorage within sediments (Maxwell et al., 2016). This difference in environmental requirements may also be the case for transplanted mature shoots. The feedback mechanism that occurs within an established meadow can improve localized conditions for plant growth and for new seedlings to establish. This will be dependent upon meadow characteristics such as shoot density, with sparse and patchy areas having little effect on stabilizing conditions (Maxwell et al., 2016). These factors could have significant effects on successful seagrass restoration. The use of presence data over the range of environmental gradients seagrasses are found is therefore likely to give broader suitability outputs in HSMs. It may also be beneficial to integrate experimental (flume) knowledge into HSM on the physical stability of individual seedlings and plants in varied environments. The availability of fine-scale hydrodynamic data would prove extremely useful in focusing planting efforts.

The aim of this paper was to review the development and use of HSMs in marine macrophyte ecosystems in order to optimize the use of HSM on seagrasses for facilitating restoration. Our specific objectives were to i) interrogate the current body of research using HSM for marine macrophytes, ii) review the selection of appropriate environmental predictor variables, and iii) review the statistical modelling approach used to draw the species-habitat association (including choice of algorithm for model parameterization, model selection (relative goodness-of-fit), and model validation (absolute goodness-of-fit)).

Methods - Literature review of macrophyte HSM studies

We collated data on studies using HSMs to determine the distribution of marine macrophytes. A recent systematic review by Robinson et al. (2017) assembled papers of marine HSM studies published between 1950 and 26 July 2016. Publications from Robinson et al. (2017) that contained marine macrophytes were included in our review and a further search for publications with keywords 'marine macrophyte'+'macroalgae'+' seagrass'+'HSM'+'SDM' was performed using ISI Web of Science (http://apps.webofknowledge.com/) and Google Scholar (https://scholar.google.co.uk/) for papers published between 2016 and 20 Jan 2022. We identified 75 publications, 35 of which were on or included seagrass.

For each of the publications identified, the main factors of interest were extracted into a database, based upon the methods outlined by Robinson et al. (2017). These factors included geographical area and scale of study, taxa, species, approach to modelling, model method (algorithm, selection and validation) and environmental variables used (Supplementary Data, Table S1).

This database was used to cross-examine trends in studies and to elucidate methods currently in use for carrying out HSM.

Results

HSM studies of marine macrophytes dated back to 2007, with seagrass studies beginning in 2009. Although initially there were more macroalgae HSM studies than seagrass there has been a marked increase in seagrass HSM publications since 2016 (77% of published in 2016 and since) so that seagrass studies are more equal (Figure 1). Just over half of macrophyte HSM papers studied multiple species (51%), whereas seagrass studies appear to be more species focused, with 31% covering multiple species. Seagrass species targeted for studies so far, include, *Z. marina* (Canal-Vergés et al., 2016; Beca-Carretero et al., 2020), *Z. noltei* (Valle et al., 2011; Valle et al., 2014), *Z. japonica* (Shafer et al., 2016), *Posidonia oceanica* (Bakirman & Gumusay 2020; Catucci & Scardi 2020), *Cymodocea nodosa* (Chefaoui et al., 2016), *Enhalus acoroides* (Lanuru et al., 2018) and *Halophila stipulacea* (Wesselmann et al., 2021).

The vast majority of HSM studies on macrophytes have been carried out in the temperate northern Atlantic region (48%), the majority of which on macroalgae then seagrasses. This is followed by temperate Australasia (12%) where macroalgae are also dominant above seagrass studies, and central Indo-Pacific (11%) where seagrass is the dominant taxa followed by mangroves (Figure 2).

From the 75 marine macrophyte studies, temperature (usually sea surface, SST) was the most frequently used predictor variable (64%), followed by bathymetry (61%), light availability (4 9%), and salinity (49%). When filtered to only seagrass studies (n=35), bathymetry was most frequently included (74%), followed by salinity (57%), light availability (51%), and temperature (51%). Substrate, wave energy and slope were the next most utilized variables in seagrass HSMs (Figure 3).

In our review, we found a total of 18 algorithms had been used. An ensemble of algorithms was the most dominant method and was used in 22 publications (29%), followed by Maximum entropy (MaxEnt) (n = 15; 20%), and Generalized Additive Models (GAMs), Generalized Linear Models (GLMs) and GIS based algorithms, which were each used in 7-8

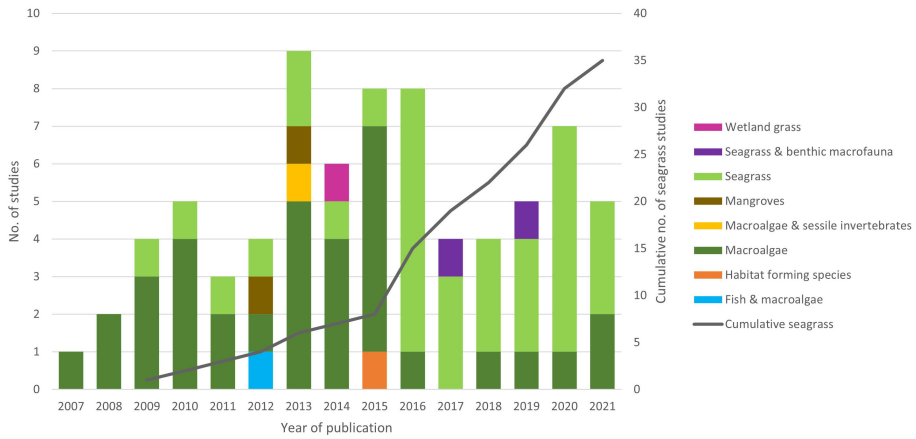


FIGURE 1
Trends in taxa used as focus within marine macrophyte HSM studies with year of publication and cumulative number of seagrass studies on secondary y axis.

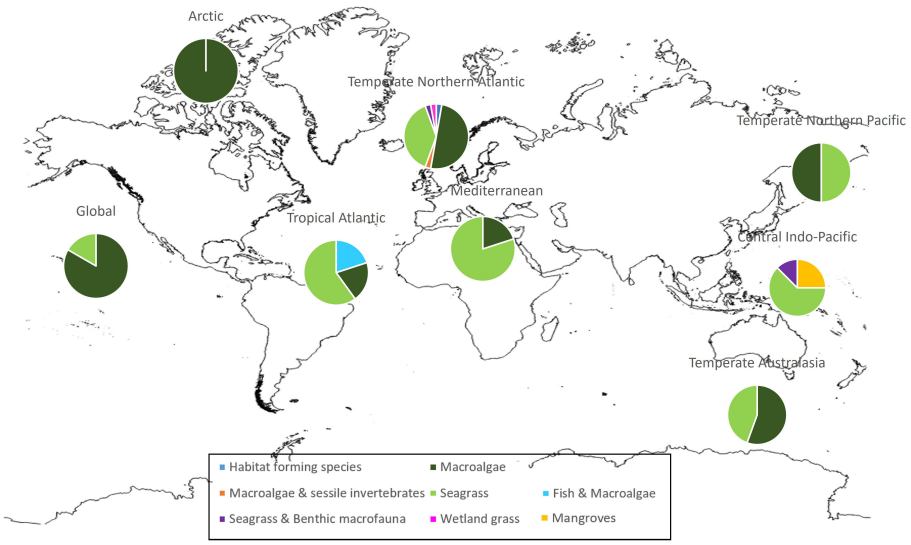


FIGURE 2
Proportion of taxa studied in HSM publications divided by geographical region.

publications (9-11%) (see Supplementary Data; Table S1). Alternatively, 27% of the algorithms were used just one time (e.g. Welch's test, Cellular automata, and Production-environmental suitability model). An ensemble was also the most popular in seagrass HSM publications, with 5 publications using two or more models.

Of the studies found within the review, the majority (40%) were focused on a local scale, covering a relatively small area such as an individual bay or estuary, followed by regional scale studies covering larger areas, such as a whole country or sea (30.7%). 16% of studies were over an intermediate area or scale (>100km coastline), and only 8% at a global scale. Ensemble was

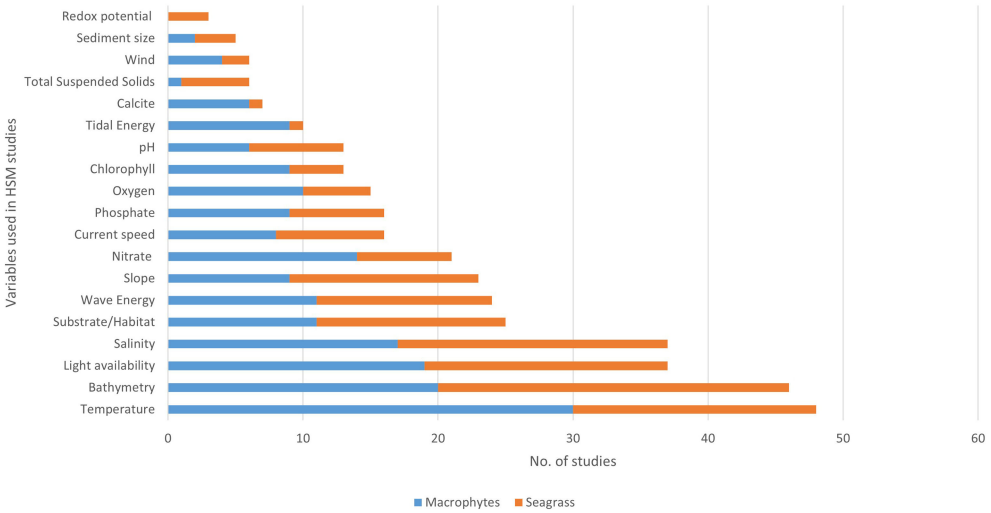


FIGURE 3
Most common environmental variables used in macrophyte HSM based on literature review. Data are divided into number from seagrass HSM (orange) and other macrophyte HSM (blue).

found to be used at a most scales; local, intermediate and regional, but not in any of the global studies. MaxEnt was more popular for larger scale studies including majority of global studies (66.7%), and regional (26%), but only 3 (9.7%) at a local scale. Local studies utilized the same environmental variables overall, with temperature, salinity (both 23%) and bathymetry (22.7%) the most commonly used followed by light (20%), substrate (16%) and wave energy (13%).

There are a total of 11 selection procedures present from the 75 publications. The most popular procedures were cross-validation, variable contribution and stepwise (n = 11, 9, and 8 publications respectively). Cross-validation was also the selection procedure most used in the seagrass HSM publications. When an ensemble of algorithms was used the most popular selection procedure was Cross-validation (29%), yet when the MaxEnt algorithm was used, variable contribution was the most commonly used selection procedure (86%).

Seven model validations were identified in the review of marine macrophyte HSM papers (Chi-squared, Cross-validation, Independent dataset, Markov Chain Monte Carlo, Multimodel inference, Regression methods and Threshold probability); the most popular was Threshold probability (33%). Just 9% of all marine macrophyte HSM publications calculated uncertainty measures, mainly through predictive power.

Discussion

This review shows the increase in trend for using HSM for marine macrophytes, particularly seagrasses in more recent years (since 2016). As a tool HSM is recognized as a viable option for interpolating spatial data for predicting where environmental conditions are conducive to habitat forming species such as seagrasses, macrophytes and mangroves. This information clearly has uses for informing restoration work, although this is still new to the field and as of yet, little evidence can be found to show its actual efficacy. Nonetheless, HSM is a useful tool for focusing efforts of restoration which can be a labor intensive and expensive to carry out.

Predictor (environmental) variables and selection

The majority of studies used over 5 predictor variables, although this ranged from only 1 to over 16 factors (Fabbrizzi et al., 2020). However, if too many environmental parameters are added to an HSM it can lead to over or underfitting a model, therefore, it has been suggested that species presences in a model should be 10 X greater than the number of environmental predictors (Singer et al., 2017). The variables that are used most are not necessarily the most important for determining species presence but could be more commonly used due to data availability. For example, years of climate modelling has refined models used to forecast factors such as sea surface, seabed temperature and salinities at a global scale. The most popular variable used in macrophyte HSM was temperature, but when only seagrass studies were looked at, bathymetry became the most popular. The majority of seagrass studies focused on a single species, with *Z. marina* the most common. This species has wide geographical range and can tolerate a range of temperatures and salinities (Blok et al., 2018). At a regional scale, temperature is not likely to fluctuate greatly and so is arguably not as important as other parameters such as light availability (Abe et al. 2003; Bertelli & Unsworth, 2018). However, localized studies utilized similar environmental variables as the overall average of studies, with few using fine-scale hydrodynamic data for example. The importance of this data is highlighted by Valle et al. (2013), which found current velocity to be the most important variable in determining *Z. marina* distributions within the Ems estuary (Netherlands) in the Wadden Sea, followed by wave exposure and depth. Good environmental data availability is key to providing the data needed for HSMs to run successfully.

It is suggested that including environmental variables that correlate may impede the performance and interpretation of the study (Georgian et al., 2019). Therefore, the removal of correlated variables is a valid method for assisting the selection of model parameters. For MaxEnt, there is a selection criteria called 'MaxEntVariableSelection' which could be used to increase the performance of the model by detecting the most important environmental variables that are not correlated (Jueterbock et al., 2016). The removal of correlated variables is always recommended for variable selection to improve model functionality which can easily be achieved using VIF (Variance Inflation Factor) tests (vifstep and vifcor) in R, (Guisan et al., 2017).

Model algorithm

HSM approaches are varied across studies and models have been designed for different types of data. For example, there are specific models in place for presence-only data such as Ecological Niche Factor Analysis (ENFA) (Valle et al., 2011), MaxEnt (Yesson et al., 2015), and Relative Environmental Suitability (RES) (Davis et al., 2016); and for presence-absence data, including Generalized Linear Models (GLMs) (Adams et al., 2016), and Generalized Additive Models (GAMs) (Nyström Sandman et al., 2013). However, if reliable absence data is not available for a species some studies and methods allow the creation of pseudo-absences or background points to use in their models (Chefaoui et al., 2016).

Various algorithms have been tested to model seagrass (*Z. marina*) distribution at a local scale, including Machine learning methods; Boosted Regression Trees (BRT), MaxEnt, Artificial Neural Networks (ANN), Random Forest (RF), and Regression-based models; GAMs, GLMs and Multivariate Adaptive Regression Spline (MARS) (Valle et al., 2013). The

models used were a mix of presence-absence and presence-only models. Valle et al. (2013) compared these algorithms with presence-only data and pseudo-absences which were created in areas never occupied to avoid overlap with presence data. In this study machine learning methods were found to perform better than regression-based models. However, according to the main findings from comparative studies (listed in chapter 11, Guisan et al., 2017), it is suggested that machine learning methods do not provide better results than regression based models, but can be represented in a more informative format and reveal other properties not always available from other approaches (Guisan et al., 2017). It must be taken into account that when using machine learning approaches, that they are calculated numerous times as outcomes will differ each time, unlike GLM, ENFA or BIOCLIM that will give the same result for the same species presence datasets (Sillero & Barbosa, 2021).

Ensemble models have been shown to perform better than individual models and are a way of avoiding dependence on a single type of model (Georgian et al., 2019). Chefaoui et al., 2016 used pseudo-absence data and an ensemble of six models for presence-absence data. The six presence-absence models were GLM, GAM, GBM (Generalized Boosting Model), RF, MARS, and FDA (Flexible Discriminant Analysis). GAM was most accurate at predicting presences compared to the presence-only MD (Mahalanobis Distance) model used for comparison. MaxEnt was the second most popular algorithm to be used in the reviewed marine macrophyte HSMs (Verbruggen et al., 2009; Tyberghein et al., 2012; Jueterbock et al., 2013; Verbruggen et al., 2013; Gormley et al., 2015; Yesson et al., 2015; Jayathilake & Costello, 2018). MaxEnt, based on maximum entropy, has been found to outperform many other models (Valle et al., 2014). The principle of maximum entropy is that, allowing for constraints, the best probability representation of species distribution that best represents the data is the one with the greatest entropy, i.e. the one which best reproduces the data (Guisan et al., 2017). There is a standalone MaxEnt package available on Java platform which is freely accessible for download with a user-friendly interface (Paquit et al., 2017). This program models habitat suitability from presence-only data (Elith et al., 2011). MaxEnt has also been written into a number of SDM packages in R including 'sdm', enabling it to be run from R.

Many marine HSM papers do not report uncertainty measures but it has been suggested that all HSMs should be accounting for uncertainty levels or estimates of errors as part of the process (Robinson et al., 2017).

Conclusions

There is no set approach used to model habitat suitability of marine macrophytes. Various sources are available to download marine species presence data and it is recommended to use distribution data with at least four years of monitoring, if possible. The number of presences of a species in a model should be 10 x more than the number of environmental predictors. Bathymetry and light availability were popular parameters and are also recommended as essential for seagrass HSM. MaxEnt and ensemble approaches were the most popular methods used in the review. Model choice will be dependent upon presence data available (presence-absence or presence only) but we suggest creating an ensemble approach to combine predictions from several high-performance models to decrease dependence on a single type of model (Table 1). Cross validation was the most used selection procedure and threshold probability the most common validation approach. Best practice measures and a checklist of important model features that need to be reported in marine HSMs are explained in Robinson et al. (2017). Using HSMs to predict the distribution of seagrass depending on environmental variables around the UK will help identify areas that could be suitable for seagrass growth and survival but do not currently show signs of presence. The output of planned HSMs will give a good indication of the most suitable locations with the highest chance of survival and help provide guidance for the future restoration of seagrass habitats around the UK.

TABLE 1 Model approaches from marine macrophyte studies and recommended model approaches.

Algorithm	Model selection	Model validation	Uncertainty	Comment
MaxEnt	Stepwise selection (e.g. 'MaxEntVariableSelection')	Threshold probability	N/A	Presence-only data
Ensemble (RF, GBM, MARS, SVM, GAM)	Cross-validation	Independent data set (Continuous Boyce Index)	Predictive power	Presence-only data used e.g. Folmer et al., 2016
Ensemble (GLM, GBM, GAM, FDA, MARS, RF)	Cross-validation	Threshold probability	Predictive power	Presence-only data but added pseudo-absences. e.g. Chefaoui et al., 2016
Ensemble (BRT, MaxEnt, ANN, RF, SVM, GAM, GLM, MARS)	Multimodel Inference	Threshold probability	N/A	Valle et al., 2013

RF, Random Forests; GLM, Generalized Additive Model; GBM, Generalized Boosting Model; SVM, Support Vector Machines; BRT, Boosted Regression Trees; ANN, Artificial NeuralNetwork; MARS, Multivariate Adaptive Regression Splines; N/A, Not Applicable.

Author contributions

RU and JB contributed to study design. CB and HS completed the literature review and analysis of the metadata. HS wrote the initial draft and CB completed the draft and updated the review. All authors contributed to the article and approved the submitted version.

Funding

This study was part of the ReSOW (Restoring Seagrass for Ocean Wealth) project funded by NERC (Natural Environment Research Council) NE/V01711X/1.

Acknowledgments

This research has been published online as a preprint (Bertelli et al., 2022).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmars.2022.997831/full#supplementary-material

References

Abe, M., Hahimoto, N., Kurashima, A., and Maegwa, M. (2003). Estimation of light requirement for the growth of zostera marina in central Japan. *Fish Sci.* 69, 890–895. doi: 10.1046/j.1444-2906.2003.00704.x

Adams, M. P., Saunders, M. I., Maxwell, P. S., Tuazon, D., Roelfsema, C. M., Callaghan, D. P., et al. (2016). Prioritizing localized management actions for seagrass conservation and restoration using a species distribution model. *Aquat Conserv. Mar. Freshw. Ecosyst.* 26, 639–659. doi: 10.1002/aqc.2573

Adhikari, D., Barik, S. K., and Upadhaya, K. (2012). Habitat distribution modelling for reintroduction of ilex khasiana purk., a critically endangered tree species of northeastern India. *Ecol. Eng.* 40, 37–43. doi: 10.1016/j.ecoleng.2011.12.004

Bakirman, T., and Gumusay, M. U. (2020). A novel GIS-MCDA-based spatial habitat suitability model for Posidonia oceanica in the Mediterranean. *Environ. Monit. Assess* 192. doi: 10.1007/s10661-020-8198-1

Barnes, T. K., Volety, A. K., Chartier, K., Mazzoti, F. J., and Pearlstine, L. (2007). A habitat suitability index model for the Eastern oyster (Crassostrea virginica), a tool for restoration of the caloosahatchee estuary, Florida. *J. Shellfish Res.* 26, 949–959. doi: 10.2983/0730-8000(2007)26[949:AHSIMF]2.0.CO;2

Beca-Carretero, P., Varela, S., and Stengel, D. B. (2020). A novel method combining species distribution models, remote sensing, and field surveys for detecting and mapping subtidal seagrass meadows. *Aquat. Conserv. Mar. Freshw. Ecosyst.* 30, 1098–1110. doi: 10.1002/aqc.3312

Bellamy, C., Scott, C., and Altringham, J. (2013). Multiscale, presence-only habitat suitability models: Fine-resolution maps for eight bat species. *J. Appl. Ecol.* 50, 892–901. doi: 10.1111/1365-2664.12117

Bertelli, C., Stokes, H., Bull, J., and Unsworth, R. K. F. (2022). The use of habitat suitability modelling for seagrass: A review. *Figshare*. doi: 10.6084/m9.figshare.20310429.v2

Bertelli, C. M., and Unsworth, R. K. F. (2018). Light stress responses by the eelgrass, zostera marina (L). *Front. Environ. Sci.* 6, 1–13. doi: 10.3389/fenvs.2018.00039

Blok, S. E., Olesen, B., and Krause-Jensen, D. (2018). Life history events of eelgrass zostera marina l. *populations across gradients latitude temperature. Mar. Ecol. Prog. Ser.* 590, 79–93. doi: 10.3354/meps12479

Borja, Á, and Elliott, M. (2013). Marine monitoring during an economic crisis: The cure is worse than the disease. *Mar. Pollut. Bull.* 68, 1–3. doi: 10.1016/j.marpolbul.2013.01.041

Canal-Vergés, P., Petersen, J. K., Rasmussen, E. K., Erichsen, A., and Flindt, M. R. (2016). Validating GIS tool to assess eelgrass potential recovery in the Limfjorden (Denmark). *Ecol. Modell.* 338, 135–148. doi: 10.1016/j.ecolmodel.2016.04.023

Catucci, E., and Scardi, M. (2020). A Machine Learning approach to the assessment of the vulnerability of Posidonia oceanica meadows. *Ecol. Indic.* 108, 105744. doi: 10.1016/j.ecolind.2019.105744

Chefaoui, R. M., Assis, J., Duarte, C. M., and Serrão, E. A. (2016). Large-Scale prediction of seagrass distribution integrating landscape metrics and environmental factors: The case of cymodocea nodosa (Mediterranean–Atlantic). *Estuaries Coasts* 39, 123–137. doi: 10.1007/s12237-015-9966-y

Convertino, M., Welle, P., Muñoz-Carpena, R., Kiker, G. A., Chu-Agor, M. L., Fischer, R. A., et al. (2012). Epistemic uncertainty in predicting shorebird biogeography affected by sea-level rise. *Ecol. Modell* 240, 1–15. doi: 10.1016/j.ecolmodel.2012.04.012

Davis, T. R., Champion, C., and Coleman, M. A. (2021). Climate refugia for kelp within an ocean warming hotspot revealed by stacked species distribution modelling. *Mar. Environ. Res.* 166, 105267. doi: 10.1016/j.marenvres.2021.105267

Davis, T. R., Harasti, D., Smith, S. D. A., and Kelaher, B. P. (2016). Using modelling to predict impacts of sea level rise and increased turbidity on seagrass distributions in estuarine embayments. *Estuar. Coast. Shelf Sci.* 181, 294–301. doi: 10.1016/j.ecss.2016.09.005

Elith, J., Phillips, S. J., Hastie, T., Dudík, M., Chee, Y. E., and Yates, C. J. (2011). A statistical explanation of MaxEnt for ecologists. *Divers. Distrib* 17, 43–57. doi: 10.1111/j.1472-4642.2010.00725.x

Fabbrizzi, E., Scardi, M., Ballesteros, E., Benedetti-Cecchi, L., Cebrian, E., Ceccherelli, G., et al. (2020). Modeling macroalgal forest distribution at mediterranean scale: Present status, drivers of changes and insights for conservation and management. *Front. Mar. Sci.* 7. doi: 10.3389/fmars.2020.00020

Folmer, E. O., van Beusekom, J. E. E., Dolch, T., Gräwe, U., van Katwijk, M. M., Kolbe, K. P., et al (2016). Consensus forecasting of intertidal seagrass habitat in the Wadden Sea. *J. Appl. Ecol.* 53, 1800–1813. doi: 10.1111/1365-2664.12681

Fourqurean, J. W., Duarte, C. M., Kennedy, H., Marba, N., Holmer, M., Mateo, M. A., et al. (2012). Seagrass ecosystems as a globally significant carbon stock. *Nat. Geosci* 5, 505–509. doi: 10.1038/ngeo1477

Gamble, C., Debney, A., Glover, A., Bertelli, C., Green, B., Hendy, I., et al. (2021). *Seagrass restoration handbook UK & Ireland*. London, UK: Zoological society of London, UK.

Georgian, S. E., Anderson, O. F., and Rowden, A. A. (2019). Ensemble habitat suitability modeling of vulnerable marine ecosystem indicator taxa to inform deep-sea fisheries management in the south pacific ocean. *Fish Res.* 211, 256–274. doi: 10.1016/j.fishres.2018.11.020

Gormley, K. S. G., Hull, A. D., Porter, J. S., Bell, M. C., and Sanderson, W. G. (2015). Adaptive management, international co-operation and planning for marine conservation hotspots in a changing climate. *Mar. Policy* 53, 54–66. doi: 10.1016/j.marpol.2014.11.017

Guarnieri, G., McOwen, C., Papa, L., Papadopoulou, N., Bilan, M., Boström, C., et al. (2019). "Active restoration across marine coastal habitats: A focus on the Mediterranean Sea," in *3rd Mediterranean symposium on the conservation of coralligenous & other calcareous bio-concretions*. Springer, Cham: Coral Reefs of the World, 9

Guisan, A., Thuiller, W., and Zimmermann, N. E. (2017). *Habitat suitability and distribution models with applications in r* (Cambridge: Cambridge University Press). doi: 10.1017/9781139028271

Hirzel, A. H., and Le Lay, G. (2008). Habitat suitability modelling and niche theory. *J. Appl. Ecol.* 45, 1372–1381. doi: 10.1111/j.1365-2664.2008.01524.x

Jayathilake, D. R. M., and Costello, M. J. (2018). A modelled global distribution of the seagrass biome. *Biol. Conserv.* 226, 120–126. doi: 10.1016/j.biocon.2018.07.009

Jueterbock, A., Smolina, I., Coyer, J. A., and Hoarau, G. (2016). The fate of the Arctic seaweed fucus distichus under climate change: An ecological niche modeling approach. *Ecol. Evol.* 6, 1712–1724. doi: 10.1002/ece3.2001

Jueterbock, A., Tyberghein, L., Verbruggen, H., Coyer, J. A., Olsen, J. L., and Hoarau, G. (2013). Climate change impact on seaweed meadow distribution in the north Atlantic rocky intertidal. *Ecol. Evol.* 3, 1356–1373. doi: 10.1002/ece3.541

Kairo, J. G., Bosire, J., Langat, J., Kirui, B., and Koedam, N. (2009). Allometry and biomass distribution in replanted mangrove plantations at gazi bay, Kenya. *Aquat Conserv. Mar. Freshw. Ecosyst.* 19, S63–S69. doi: 10.1002/aqc.1046

Kerr, J. (2017). *Introduction to energy and climate: Developing a sustainable environment, 1st ed* (CRC Press). doi: 10.1201/9781315151885

Lanuru, M., Mashoreng, S., and Amri, K. (2018). Using site-selection model to identify suitable sites for seagrass transplantation in the west coast of South Sulawesi. *J. Phys. Conf. Ser.* 979. doi: 10.1088/1742-6596/979/1/012007

Marion, S. R., and Orth, R. J. (2010). Innovative Techniques for Large-scale Seagrass Restoration Using Zostera marina (eelgrass) Seeds. *Restor. Ecol.* 18, 514–526.

Matthiopoulos, J., Fieberg, J., and Aarts, G. (2020). *Species-habitat associations: Spatial data, predictive models, and ecological insights* (University of Minnesota Digital Conservancy: University of Minnesota Libraries Publishing). doi: 10.24926/2020.081320

Maxwell, P. S., Ekl, J. S., Van Katwijk, M. M., and Brien, K. R. O. (2016). *The fundamental role of ecological feedback mechanisms for the adaptive management of seagrass ecosystems – a review. Biological Reviews* 92 (3), 1521-1538. doi: 10.1111/brv.12294

McDonald, A. M., Prado, P., Heck, K. L., Fourqurean, J. W., Frankovich, T. A., Dunton, K. H., et al. (2016). Seagrass growth, reproductive, and morphological plasticity across environmental gradients over a large spatial scale. *Aquat Bot.* 134, 87–96. doi: 10.1016/j.aquabot.2016.07.007

Nordlund, L. M., Koch, E. W., Barbier, E. B., and Creed, J. C. (2016). Seagrass ecosystem services and their variability across genera and geographical regions. *PloS One* 11, 1–23. doi: 10.1371/journal.pone.0163091

Nyström Sandman, A., Wikström, S. A., Blomqvist, M., Kautsky, H., and Isaeus, M. (2013). Scale-dependent influence of environmental variables on species distribution: A case study on five coastal benthic species in the Baltic Sea. *Ecography (Cop)* 36, 354–363. doi: 10.1111/j.1600-0587.2012.07053.x

Orth, R. J., Lefcheck, J. S., McGlathery, K. S., Aoki, L., Luckenbach, M. W., Moore, K. A., et al. (2020). Restoration of seagrass habitat leads to rapid recovery of coastal ecosystem services. *Sci. Adv.* 6, eabc6434. doi: 10.1126/sciadv.abc6434

Paquit, J. C., Pampolina, N. M., Tiburan, C. L., and Manalo, M. M. Q. (2017). Maxent modeling of the habitat distribution of the critically endangered pterocarpus indicus willd. forma indicus in Mindanao, Philippines. *J. Biodivers Environ. Sci.* 10, 112–122.

Peralta, G., Brun, F. G., Pérez-Lloréns, J. L., and Bouma, T. J. (2006). Direct effects of current velocity on the growth, morphometry and architecture of seagrasses: A case study on zostera noltii. *Mar. Ecol. Prog. Ser.* 327, 135–142. doi: 10.3354/meps327135

Reynolds, L. K., Waycott, M., McGlathery, K. J., and Orth, R. J. (2016). Ecosystem services returned through seagrass restoration. *Restor. Ecol.* 24, 583–588. doi: 10.1111/rec.12360

Robinson, N. M., Nelson, W. A., Costello, M. J., Sutherland, J. E., and Lundquist, C. J. (2017). A systematic review of marine-based species distribution models (SDMs) with recommendations for best practice. *Front. Mar. Sci.* 4, 1–11. doi: 10.3389/fmars.2017.00421

Röhr, M. E., Holmer, M., Baum, J. K., Björk, M., Chin, D., Chalifour, L., et al. (2018). Blue carbon storage capacity of temperate eelgrass (Zostera marina) meadows. *Global Biogeochem Cycles* 32, 1457–1475. doi: 10.1029/2018GB005941

Shafer, D. J., Swannack, T. M., Saltus, C., Kaldy, J. E., and Davis, A. (2016). Development and validation of a habitat suitability model for the non-indigenous seagrass zostera japonica in north america. *Manag Biol. Invasions* 7, 141–155. doi: 10.3391/mbi.2016.7.2.02

Sillero, N., and Barbosa, A. M. (2021). Common mistakes in ecological niche models. *Int. J. Geogr. Inf Sci.* 35, 213–226. doi: 10.1080/13658816.2020.1798968

Silliman, B. R. (2014). Salt marshes. *Curr. Biol.* 24, 348–350. doi: 10.1016/j.cub.2014.03.001

Singer, A., Millat, G., Staneva, J., and Kröncke, I. (2017). Modelling benthic macrofauna and seagrass distribution patterns in a north Sea tidal basin in response to 2050 climatic and environmental scenarios. *Estuar. Coast. Shelf Sci.* 188, 99–108. doi: 10.1016/j.ecss.2017.02.003

Spivak, A. C., Sanderman, J., Bowen, J. L., Canuel, E. A., and Hopkinson, C. S. (2019). Global-change controls on soil-carbon accumulation and loss in coastal vegetated ecosystems. *Nat. Geosci* 12, 685–692. doi: 10.1038/s41561-019-0435-2

Stafford, R., Chamberlain, B., Clavey, L., Gillingham, P. K., McKain, S., Morecroft, M. D., et al. (2021). Nature-based solutions for climate change in the UK: A report by the British ecological society. *Br. Ecol. Soc.* 190.

Tan, Y. M., Dalby, O., Kendrick, G. A., Statton, J., Sinclair, E. A., Fraser, M. W., et al. (2020). Seagrass Restoration Is Possible: Insights and Lessons From Australia and New Zealand. *Front. Mar. Sci.* 7. doi: 10.3389/fmars.2020.00617

Tyberghein, L., Verbruggen, H., Pauly, K., Troupin, C., Mineur, F., and De Clerck, O. (2012). Bio-ORACLE: A global environmental dataset for marine species distribution modelling. *Glob Ecol. Biogeogr.* 21, 272–281. doi: 10.1111/j.1466-8238.2011.00656.x

Unsworth, R. K. F., Bertelli, C. M., Cullen-Unsworth, L. C., Esteban, N., Jones, B. L., Lilley, R., et al. (2019a). Sowing the seeds of seagrass recovery using hessian bags. *Front. Ecol. Evol.* 7, 1–7. doi: 10.3389/fevo.2019.00311

Unsworth, R. K. F., McKenzie, L. J., Collier, C. J., Cullen-Unsworth, L. C., Duarte, C. M., Eklöf, J. S., et al. (2019b). Global challenges for seagrass conservation. *Ambio* 48, 801–815. doi: 10.1007/s13280-018-1115-y

Valle, M., Borja, Á, Chust, G., Galparsoro, I., and Garmendia, J. M. (2011). Modelling suitable estuarine habitats for zostera noltii, using ecological niche factor analysis and bathymetric LiDAR. *Estuar. Coast. Shelf Sci.* 94, 144–154. doi: 10.1016/j.ecss.2011.05.031

Valle, M., Chust, G., del Campo, A., Wisz, M. S., Olsen, S. M., Garmendia, J. M., et al. (2014). Projecting future distribution of the seagrass zostera noltii under global warming and sea level rise. *Biol. Conserv.* 170, 74–85. doi: 10.1016/j.biocon.2013.12.017

Valle, M., van Katwijk, M. M., de Jong, D. J., Bouma, T. J., Schipper, A. M., Chust, G., et al. (2013). Comparing the performance of species distribution models of zostera marina: Implications for conservation. *J. Sea Res.* 83, 56–64. doi: 10.1016/j.seares.2013.03.002

van Katwijk, M. M., Thorhaug, A., Marbà, N., Orth, R. J., Duarte, C. M., Kendrick, G. A., et al. (2016). Global analysis of seagrass restoration: The importance of large-scale planting. *J. Appl. Ecol.* 53, 567–578. doi: 10.1111/1365-2664.12562

Verbruggen, H., Tyberghein, L., Belton, G. S., Mineur, F., Jueterbock, A., Hoarau, G., et al. (2013). Improving transferability of introduced species' distribution models: New tools to forecast the spread of a highly invasive seaweed. *PloS One* 8, 1–13. doi: 10.1371/journal.pone.0068337

Verbruggen, H., Tyberghein, L., Pauly, K., Vlaeminck, C., Van Nieuwenhuyze, K., Kooistra, W. H. C. F., et al. (2009). Macroecology meets macroevolution: Evolutionary niche dynamics in the seaweed halimeda. *Glob Ecol. Biogeogr* 18, 393–405. doi: 10.1111/j.1466-8238.2009.00463.x

Vinagre, C., Fonseca, V., Cabral, H., and Costa, M. J. (2006). Habitat suitability index models for the juvenile soles, solea solea and solea senegalensis, in the tagus estuary: Defining variables for species management. *Fish Res.* 82, 140–149. doi: 10.1016/j.fishres.2006.07.011

Waycott, M., Duarte, C. M., Carruthers, T. J. B., Orth, R. J., Dennison, W. C., Olyarnik, S., et al. (2009). Accelerating loss of seagrasses across the globe threatens coastal ecosystems. *PNAS* 1, 12377–12381. doi: 10.1073/pnas.0905620106

Wesselmann, M., Chefaoui, R. M., Marbà, N., Serrao, E. A., and Duarte, C. M. (2021). Warming threatens to propel the expansion of the exotic seagrass halophila stipulacea. *Front. Mar. Sci.* 8, 1–14. doi: 10.3389/fmars.2021.759676

Yesson, C., Bush, L. E., Davies, A. J., Maggs, C. A., and Brodie, J. (2015). The distribution and environmental requirements of large brown seaweeds in the British isles. *J. Mar. Biol. Assoc. U.K.* 95, 669–680. doi: 10.1017/S0025315414001453

frontiers | Frontiers in Marine Science

TYPE Original Research
PUBLISHED 02 December 2022
DOI 10.3389/fmars.2022.1075855

OPEN ACCESS

EDITED BY
Xi Xiao,
Zhejiang University, China

REVIEWED BY
Xiujuan Shan,
Yellow Sea Fisheries Research Institute (CAFS), China
Matteo Zucchetta,
National Research Council (CNR), Italy

*CORRESPONDENCE
Zuozhi Chen
chenzuozhi@scsfri.ac.cn
Shengwei Ma
mashengwei@scsfri.ac.cn

†These authors have contributed equally to this work and share first authorship

SPECIALTY SECTION
This article was submitted to Ocean Observation, a section of the journal Frontiers in Marine Science

RECEIVED 21 October 2022
ACCEPTED 17 November 2022
PUBLISHED 02 December 2022

CITATION
Qian J, Li J, Zhang K, Qiu Y, Cai Y, Wu Q, Tian H, Ma S and Chen Z (2022) Spatial–temporal distribution of large-size light falling-net fisheries in the South China Sea.
Front. Mar. Sci. 9:1075855.
doi: 10.3389/fmars.2022.1075855

Spatial–temporal distribution of large-size light falling-net fisheries in the South China Sea

Jing Qian[1,2,3†], Jiajun Li[1,2,4†], Kui Zhang[1,2], Yongsong Qiu[1,2], Yancong Cai[1,2], Qiaer Wu[1,2], Han Tian[1,2,3], Shengwei Ma[1,2*] and Zuozhi Chen[1,2*]

[1]South China Sea Fisheries Research Institute, Chinese Academy of Fishery Science, Guangzhou, China, [2]Key Laboratory for Sustainable Utilization of Open-sea Fishery, Ministry of Agriculture and Rural Affairs, Guangzhou, China, [3]College of Marine Sciences, Shanghai Ocean University, Shanghai, China, [4]Key Laboratory of Efficient Utilization and Processing of Marine Fishery Resources of Hainan Province, Sanya Tropical Fisheries Research Institute, Sanya, China

With the rapid modernization of methods for monitoring fishing vessels, different ways of obtaining data on large-size light falling-net vessels have developed. In particular, the vessel monitoring system (VMS) may be used to track fishing activities in the South China Sea (SCS). In this study, we used VMS data to show the characterize and differences in the information transmitted when monitoring large-size falling-net fishing vessels, as well as their advantages and disadvantages. We investigated the spatial distribution of large-size light falling-net fisheries in the SCS using Beidou VMS data from 2017 to 2020. We also compared trajectories of fishing vessels from various provinces. According to these data, vessels operate in the Nansha Islands and the waters of the Pearl River Estuary, Beibu Gulf, Zhongsha, and Xisha Islands, and there is a clear seasonal pattern of fishing in the SCS. There are the differences of vessels in terms of distribution areas and ports of entry and exit. Additionally, we found the operation area of the large-size falling-net fishing vessels divided into four main fishing zones. The analysis compares the number of days of operation in the different fishing areas as well as seasonal variations. Our study shows that the operating area of large-size light falling-net vessels has extended southwards and that the number of days for which fishing vessels operate has increased compared with the last decade. Using the Beidou VMS to study the large-size light falling-net vessels in the SCS can provide more detailed scientific reference data for the SCS fisheries.

KEYWORDS

Beidou VMS, vessel position data, large-size light falling-net fishery, South China Sea, fishing area

Introduction

The South China Sea (SCS) has a diverse natural environment and rich fishery resources. In recent years, marine fishing production in the SCS has been maintained at 3 million tons, accounting for approximately 30% of the country's total. Owing to overfishing and environmental pollution, the main traditional economic fishes in the SCS have gradually declined (Jin et al., 2015; Zhang et al., 2017). However, the outer waters of the SCS still contain rich pelagic fishery resources (Zhang et al., 2018). In the early 1990s, light falling-net fishing vessels stated to appear in the northern part of the SCS. In the spring of 2004, light falling-net fishing vessels were successful in exploring and catching *Symplectoteuthis oualaniensis* in the Zhongsha and Xisha waters. With the number of the light falling-net fishing vessels increasing, the focus gradually shifted from the northern to the central and southern parts of the SCS. Since 2011, the fishery has been expanded to the Nansha waters. The principle of the light falling-net fishing vessels is to use light to trap pelagic phototactic fish and cephalopods such as *Decapterus maruadsi*, *Auxis thazard*, and *Uroteuthis chinensis* (Jiang et al., 2018). Compared with traditional fishing operations (Wei et al., 2019), Large-size light falling-net cause much less environmental damage (Yang et al., 2002), are simpler to operate, and are less technically demanding (Yan et al., 2009). Large-size light falling-net fisheries in the SCS have continued to expand recently. Following technological improvements, light falling-net fishing vessels have become the key to fisheries development in the SCS (Yang et al., 2002; Zheng and Tang, 2016). Exploring the spatiotemporal distribution and changing patterns of large-size light falling-net fishing vessels can provide a scientific basis for the prudent development and protection of fishery resources in the SCS (Yan et al., 2014; Su et al., 2018).

The vessel monitoring system (VMS) is a satellite-based monitoring system consisting of continuous monitoring equipment installed on board fishing vessels. The VMS can obtain and store real-time data for the fishing vessels including, but not limited to, their position, heading, and speed. The information is transmitted to a shore-based monitoring center in real time. Over the past few decades, various countries have developed their own VMS systems. The United States began to research and develop their VMS in 1991 and officially announced a VMS standard in March 1994. New Zealand started to develop a VMS in 1992, and it began using it to monitor fishing vessels in 1994. Australia and its Fisheries Administration have been using a VMS program since 1993, and the European Union began implementing its VMS program in 1994. Additionally, Peru, Malaysia, Russia, and other countries have also promoted the use of the VMS (Cao et al., 2002). The Chinese Beidou vessel position monitoring system began tracking fishing vessels in the SCS in 2000. Although a VMS was developed and popularized in China slightly later, it developed relatively rapidly. Since the end of 2020, all medium and large fishing vessels (These vessels are mostly >12 m in length) from the three provinces of the SCS have been under surveillance (Zheng et al., 2019). The Beidou VMS allows for more scientifically informed management of fishing vessels in the SCS (Zhang, 2020).

Previously, Chinese research on large-size light falling-net fisheries in the SCS focused on using fishing logs and acoustic estimation (Zhang et al., 2016; Zhang et al., 2016; Yan et al., 2021a) to study fishing grounds, fishing season, catch composition (Zhang et al., 2013), variation and distribution (Zou et al., 2014), and fishery resource abundance (Xie et al., 2020). The studies also included the spatial–temporal distribution, biomass of fishery resources (Li et al., 2016), and major economically important species (Zhang et al., 2016) in the SCS. With the development of the Beidou navigation system, the SMS function can be used to collect fisheries data from light falling-net fishing vessels and assess SCS fisheries resources (Feng et al., 2014). VMS is increasingly required for fishing vessels all over the world. Researchers have used VMS data to study the dynamics of purse seine fishing grounds (Rocio et al., 2015), analyze the fishing intensity of trawlers (Lambert et al., 2012), estimate fishing catches, validate log data, and delineate fishing-affected habitats, among various other applications (Mills et al., 2007; Bastardie et al., 2010; Jennings and Lee, 2012). Since 2014, Zhang et al. (Zhang et al., 2014; Zhang et al., 2014; Zhang et al., 2016) have used Beidou VMS data to analyze the fishing effort, fishing traceability, and fishing characteristics of fishing vessels. Some researchers, including Tang, Zhang, and Yu et al. (Zhang et al., 2018; Zhang et al., 2019; Tang et al., 2020; Yu et al., 2021) have used Beidou vessel position data to dynamically monitor and analyze the voyage dynamics, spatial and temporal changes in operations, and operational intensity of different types of fishing vessels. Due to a lack of data, only a few studies have so far described the fishing activities in South China Sea (SCS). Though the VMS data are being used increasingly extensively, there are still few Beidou VMS-based studies on large-size light falling-net fishing vessels in the SCS. VMS was more often used by researchers to analyze the voyages (Zhang et al., 2018; Tang et al., 2020), fishing status (Zhang et al., 2014; Zhang et al., 2016), and operational characteristics of other types of fishing vessels (Zhang et al., 2019; Yu et al., 2021).Here, we investigate large-size light falling-net fishing vessels in the SCS using Beidou VMS data from 2017 to 2020. By analyzing changes in the position of the large-size light falling-net vessels, we were able to monitor the operation of the fishing vessels in the SCS, exploring their spatial and temporal distribution patterns and change patterns. This paper provides an innovative approach to studying the SCS large-size light falling-net fisheries and useful reference for further assessment and management of fishery resources of the SCS.

Materials and methods

Study area

Figure 1 depicts the study area in the South China Sea (16°N - 24°N, 106°E - 121°E). The South China Sea, an important part of the world's large marine ecosystem and one of the world's major fishery production sites, accounts for more than 10% of the total global catch. The excellent natural environment, diverse fishery ecological environment types, and rich fishery resources have resulted in an excellent fishing ground.

VMS data

We obtained the Beidou VMS data collected by the Strategy Research Center for the South China Sea Fisheries. The spatial resolution was approximately 10 m, and the temporal resolution was approximately 3 min. The dataset consisted of records for 121 large-size light falling-net fishing vessels in Guangxi, Guangdong, and Hainan (of which there were 55, 55, and 11 vessels in Guangxi, Hainan, and Guangdong, respectively) in the SCS from 2017 to 2020. The records included the names of the fishing vessels, positioning time, latitude and longitude, speed, heading, and other related information.

Operating days and position of fishing vessels

Preliminary processing of VMS data was required. First, after omitting any duplicate data, we excluded the vessel position data of fishing vessels operating outside the SCS (106°~121°E, 6~24°E). Second, we removed data from fishing vessels that had transmitted less than 1,000 messages of data over 1 month because this provides data points that are too few to be representative of the vessel's behavioral characteristic. The fishing vessels sent many different discrete points during sailing and operation. Therefore, the speed and position of the vessels had to be changed during the month. We trimmed the datasets using the above thresholds. Using the cleaned-up data, we calculated the maximum number of fishing boats and operating days per month from 2017 to 2020.

In this study, we used Matlab 2016a to filter the timestamp, speed, latitude/longitude, and other information for 121 large-size light falling-net fishing vessels obtained from 2017 to 2020. If the data were within the preset range, the vessel data were marked as fishing operation points. Based on the operational characteristics of a large-size light falling-net fishing vessel, its main operation time is from after dark (around 19:00) to dawn (around 05:00 the next day), with no drifting (the speed is kept at 0-2 kn). Thus, a higher velocity of 3-8 kn during that time range ensures the vessel's moving state. The location of each net

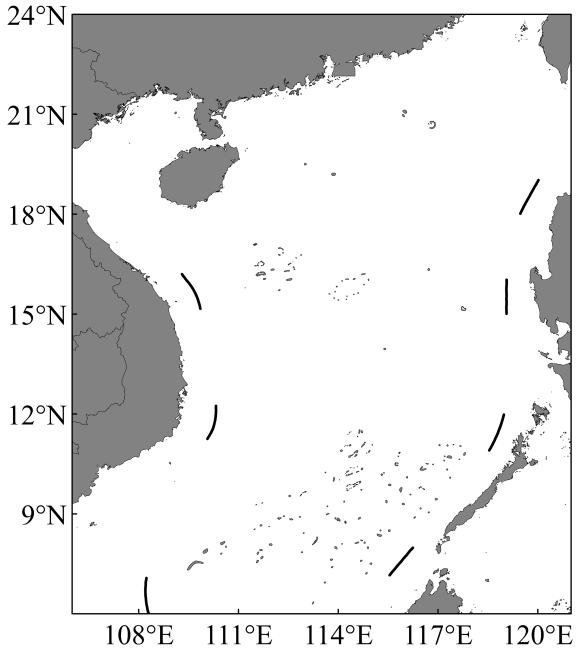


FIGURE 1
South China Sea waters.

operation may change, creating multiple discrete vessel locations (Xie et al., 2020). Excluding the first net, which requires a longer operation time, the net operation time for the other nets was < 30 min. Accordingly, we only chose those vessels with low speed (a vessel velocity of < 3 knots at 1:00 a.m-1:59 am) as fishing candidates and the fishing locations are calculated as the mean location. Besides, fishing vessels which operate for less than 5 days per month are removed to avoid contingencies. We used the mean center algorithm to calculate the fishing locations of the vessels in the SCS. The Equation is as follow (Zhou et al., 2021):

$$x = \frac{\sum_{i=1}^{n} xi}{n} \quad (1)$$

$$y = \frac{\sum_{i=1}^{n} yi}{n} \quad (2)$$

(x,y) represent the coordinates of the vessel's fishing location, (x_i,y_i) represent the coordinates of the vessel's discrete points.

We used Matlab 2016a to create 0.25°-resolution grids and estimated the mean numbers of fishing vessels in each grid cell by month and season.

Determination of fishing vessel port

The Beidou VMS datasets include the vessel name, latitude, longitude, and time. Fishing vessels generally send position information every 3 min. The province of the fishing vessel could be distinguished based on the vessel name. The vessel position points for each fishing vessel were sorted in chronological order. Multiple scattered points were connected to form a trajectory map of the fishing vessel. Finally, the trajectories of the fishing vessels were matched with the ports to determine the main ports from which the fishing vessels from different provinces enter or leave.

Results

Characteristics of VMS data

We mined and analyzed fishing vessel monitoring messages received by the target fishing vessels from 2017 to 2020 and extracted the fishing days of the vessels. Upon comparing the number of large-size light falling-net vessels (Figure 2) and the number of operating days (Figure 3) in the SCS by month from 2017 to 2020, we found that vessels operate in the SCS throughout the year. The period from March to May was the primary production season for large-size light falling-net vessels in the SCS. Position data were received from more than 78% ± 6.7% of the large-size light falling-net vessels, and the average number of fishing days per month exceeded 26.4 (± 2.5) d·vessel^{-1}. From June to August, owing to the fishing moratorium in the SCS, the production of fishing vessels was affected. Position data were received from less than 50% of the vessels. Meanwhile, the mean monthly operation was 25.0 (± 1.3) d·vessel^{-1}. The fishing moratorium ends before September, but the SCS has high winds and harsh operating conditions in the fall. Therefore, approximately 51.6% ± 1.7% of the vessels sent position data. The average monthly operation time was 27.4 (± 0.6) d·vessel^{-1}. From December to February, approximately 53.7% of the large-size light falling-net vessels sent position data, and the average number of monthly fishing days was 23.5 (± 0.9) d·vessel^{-1}.

Frequency of VMS message reception

Given that the theoretical value of Beidou VMS time resolution is 3 min, each vessel can hypothetically transmit 480 position messages per day. Extracting and averaging the operational data of fishing vessels, we found that the actual values differed from the theoretical values (Figure 4). The VMS signal was transmitted at a lower frequency from May to July in 2018 and 2020. Also, the number of messages sent by fishing vessels differed from province to province (Figure 5). In 2017 and 2020, the Guangxi fishing vessels sent fewer VMS messages per day than the Guangdong fishing vessels. In 2018 and 2019, the Guangxi fishing vessels sent more VMS messages than the Guangdong fishing vessels. In contrast, the Hainan fishing vessels consistently sent the lowest number of VMS messages over the 4 years.

Distribution fishing vessels

Based on the vessel position data of fishing vessels from 2017 to 2020, we drew the monthly distribution maps of vessels for different years. Figure 6A shows the spatial pattern of the average monthly operation distribution of large-size light falling-net vessels. There were few fishing vessels in January and February, and the operating location distribution was scattered. The fishing vessels were mainly in the Zhong and Xisha waters. From March to April, the fishing area of large-size light falling-net vessels in the SCS became the largest and was denser. At the same time, the fishing vessels began extending southward, reaching the waters near the Nansha Islands, and the operation area became more concentrated. From May to July, the number of fishing vessels and days decreased significantly. The location was also scattered. Then the number of fishing vessels started to increase in August, and the most of fishing vessels began to shift northward from the outer waters of the SCS. In September, they concentrated more around Wanshan, the waters of the Pearl River Estuary, and the waters of the Dongsha Islands. From October to December, fishing vessels

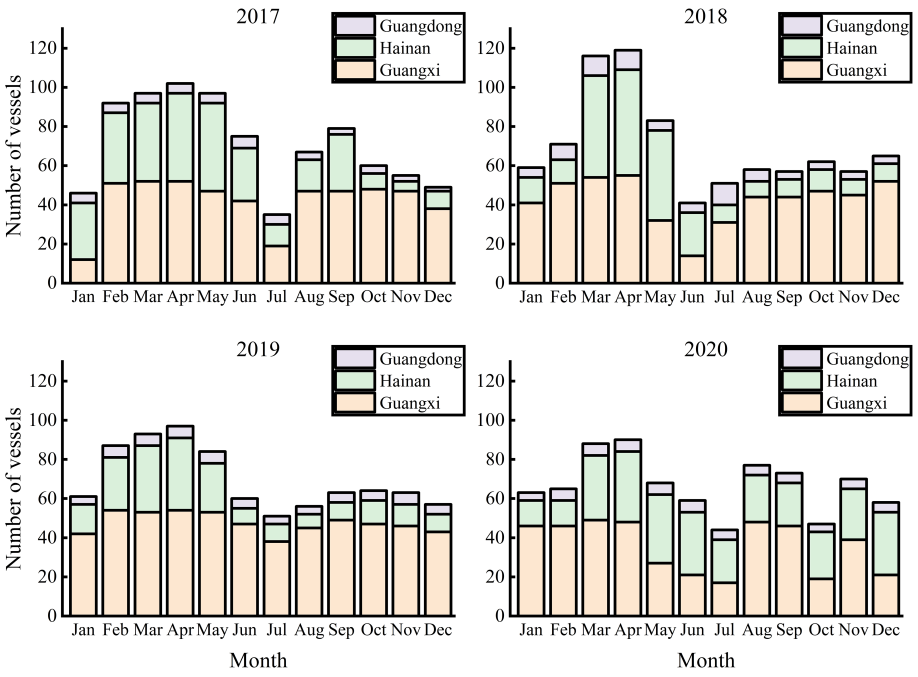


FIGURE 2
Monthly maximum numbers of fishing vessels obtained from vessel monitoring information.

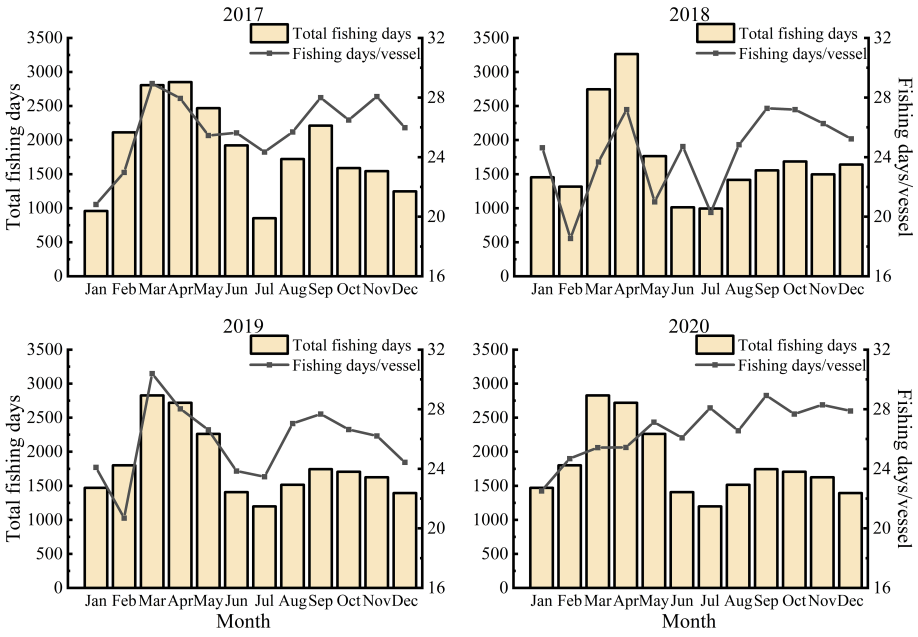


FIGURE 3
Average monthly operating days of fishing vessels from 2017 to 2020.

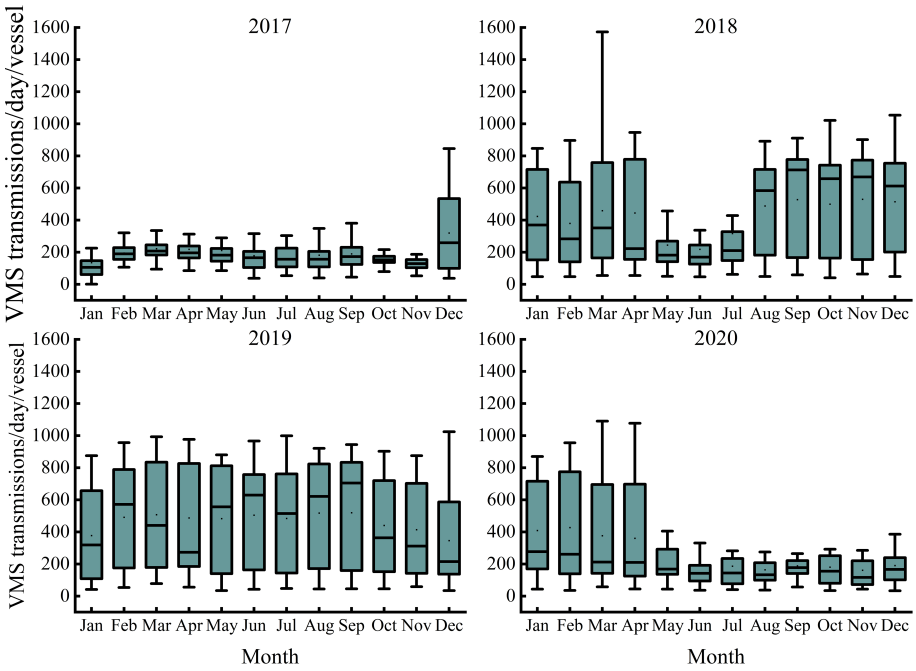


FIGURE 4
Average monthly numbers of Beidou vessel monitoring system transmissions per vessel per day in 2018.

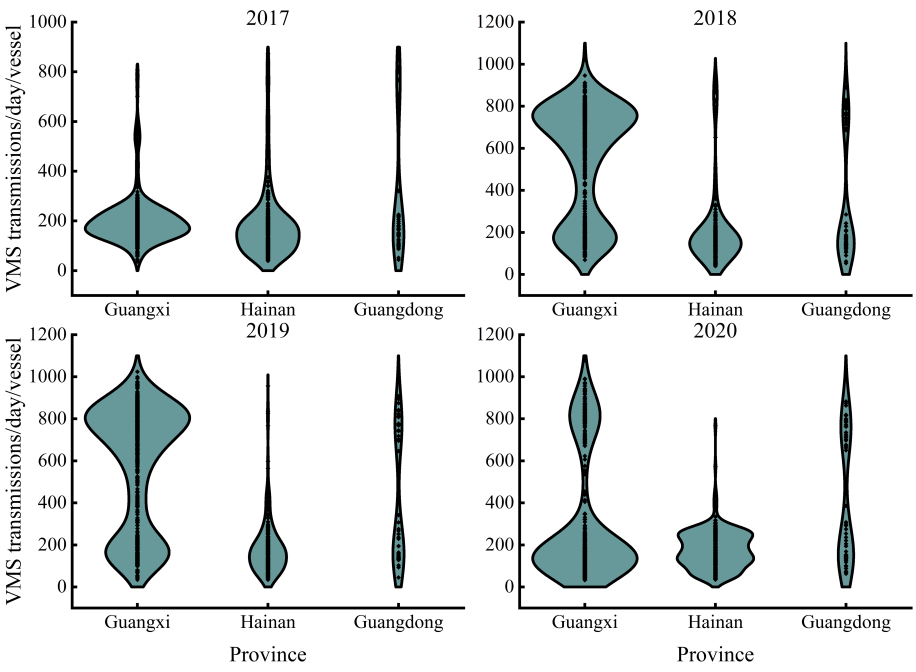


FIGURE 5
Average numbers of vessel monitoring system transmissions per vessel per day for the different provinces.

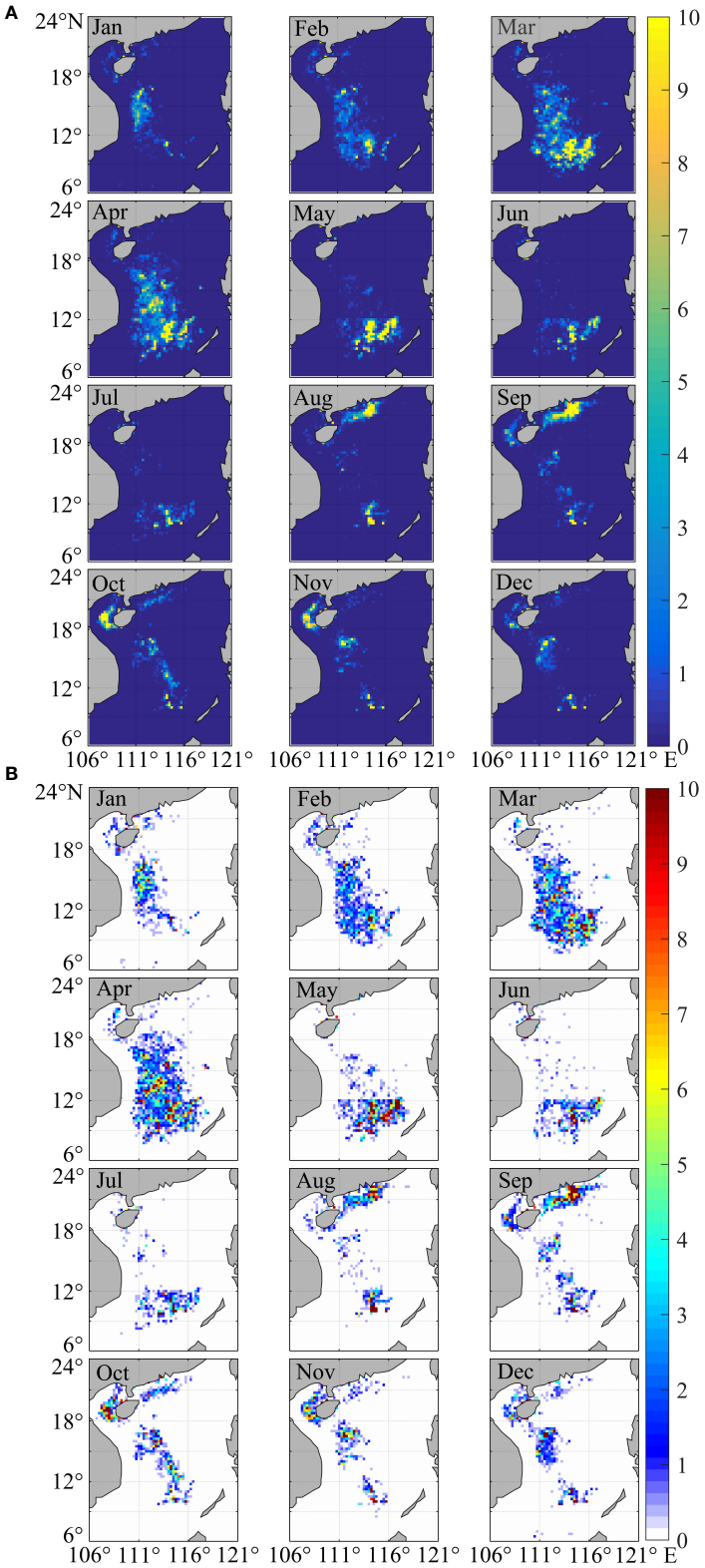


FIGURE 6
Monthly average distribution **(A)** and standard deviation distribution **(B)** of large-size light falling-net vessels in the South China Sea from 2017 to 2020.

were more concentrated in Beibu Gulf, and the waters of the Zhongsha and Xisha Islands. To further explore the characteristics of the position changes of large-size light falling-net vessels during fishing operations, the standard deviation distribution was plotted (Figure 6B). Clearly, the differences in the distribution of fishing vessels among 2017 to 2020 were small. The monthly operations of fishing vessels in the South China Sea exhibited some regularity.

We also analyzed the distribution of fishing vessels based on seasons. Most of the large-size light falling-net fishing vessels were in the Zhongsha, Xisha, and Nansha waters in the spring (Figure 7). Among the four seasons, spring had the maximum number of fishing vessels and the most extensive distribution. The number of fishing vessels operating in the summer decreased abruptly and was distributed mostly in two areas. The vessels were mainly in the waters south of 12°N and the Pearl River mouth. In autumn, the number of fishing vessels rebounded, and the distribution range changed. In addition to the Pearl River Estuary waters, vessels also started to fish in the Beibu Gulf waters. In the winter, the main operating area of fishing vessels changed to the Zhongsha and Xisha Island waters.

We processed the target fishing vessel sailing locations by year and province (district) and then traced the records of these vessels and visualized their trajectory. We found distinctions in the positions of the fishing vessels among the different provinces (districts) (Figure 8). Guangxi fishing vessels were more widely distributed, with several key fishing areas throughout the year. Hainan fishing vessels were less spread out in the Pearl River waters and more spread out in the Zhongsha, Xisha, and Nansha waters. The Guangdong vessels rarely operated in the Beibu Gulf waters throughout the year, staying mainly in the Nansha Islands waters. Additionally, we tracked the trajectory of the fishing vessels throughout the year. The main ports of the fishing vessels are shown in Figure 9. Using the 2017–2020 vessel position data, we calculated the number of fishing vessels entering and leaving the ports. Fishing vessels from different provinces were in several different ports. For example, the fishing vessels of Guangxi were mainly concentrated in the Beihai, Zhapo, Yangpu, Dongfang, and Sanya ports. Of these, the largest proportion is in Beihai every year. Almost all the 55 fishing boats from Guangxi entered the Beihai port. The Zhapo and Yangpu ports also contained more than 45% of the fishing boats entering and leaving. For the fishing vessels of Hainan, the major ports were Sanya (56.3%) and Yazhou (68.0%). The fishing vessels from Guangzhou differed from the other two provinces in that their main ports were Guangzhou and Bohe.

Additionally, we conducted statistics on fishing vessels operating south of 12°N during the fishing moratorium by province (Figure 10). Our results indicate that the fishing vessels' numbers, operating locations, and fishing days varied depending on the province of the fishing vessel. During the fishing ban, the number of fishing vessels belonging to Guangxi province operating south of 12°N was higher in 2017 and 2019. In 2018 and 2020, there were more fishing vessels from Hainan (Table 1).

Distribution of fishing grounds

Comparing previous studies (Zhang et al., 2013), the distribution of the core fishing grounds for these vessels can be divided into four areas (Figure 11). I) 19°30′~23°N, 111° ~116°E, outer continental shelf waters of the Pearl River Estuary; II) 19°~22°N, 106° ~110°E, the Beibu Gulf continental shelf sea area; III) 14°~19°N, 109°~116°E, near the Zhongsha and Xisha Islands waters; and IV: 4°~12°N, 109°30′~118°E, near the Nansha Islands waters. After the statistical screening of fishing days, a comparative graph was made for fishing vessels operating in the four fishing zones in different seasons from 2017 to 2020 (Figure 11). In the spring, vessels mainly operated in fishing areas III and IV, where the number of fishing days in fishing area IV was the maximum number of fishing days among the four fishing areas that year. In the spring of the four years, there were 3,454, 2,815, 4,676, and 3,439 fishing days in fishing area IV, respectively, and they accounted for 19.19%, 19.33%, 29.76%, and 16.82% of the total number of operating days over the entire year. By extension, this means that all of the fishing vessels operate in fishing area IV for over 45% of the entire fishing season. The number of fishing days in fishing area IV in the spring is >40% of the fishing days in fishing area IV throughout the year. The percentage was even up to approximately 60% from 2018 to 2019. In the summer, fishing vessels operate mainly in fishing area IV. Among the vessels, the number of fishing days in area IV in the summer of 2017 was 2,691, accounting for 14.95% of the total operating days. For 2018, 2019, and 2020, the numbers of fishing days in the summer in fishing area IV were 1,144, 1,598, and 1,989 days, respectively, which represent 7.86%, 10.17%, and 14.95% of the total operating days for the whole year. Additionally, the number of fishing vessels in fishing area I increased dramatically in August, creating two distinct fishing zones. In the autumn, there were differences in the distribution of fishing vessels. In addition to the main distribution areas of fishing areas I and II, fishing vessels were more distributed throughout fishing area III from 2018 to 2019 and fishing area IV in 2020. Among the vessels, the number of fishing days in area I in September continued to increase compared with August, accounting for almost 52.88 (± 7.92) % of the total number of fishing days in that month. Moreover, in October and November, there were two prominent fishing zones: fishing areas II and III. Fishing vessels operated in fishing area III for 1,299 and 1,602 days in the winter of 2017 and 2018, respectively. They accounted for 43.59% and 57.21% of the total number of fishing days in the whole season. In the winters of 2019 and 2020, fishing vessels operated in fishing area IV for 1,052 and 1,483 days, accounting for 39.11% and 33.22% of the total number of fishing days for the whole season.

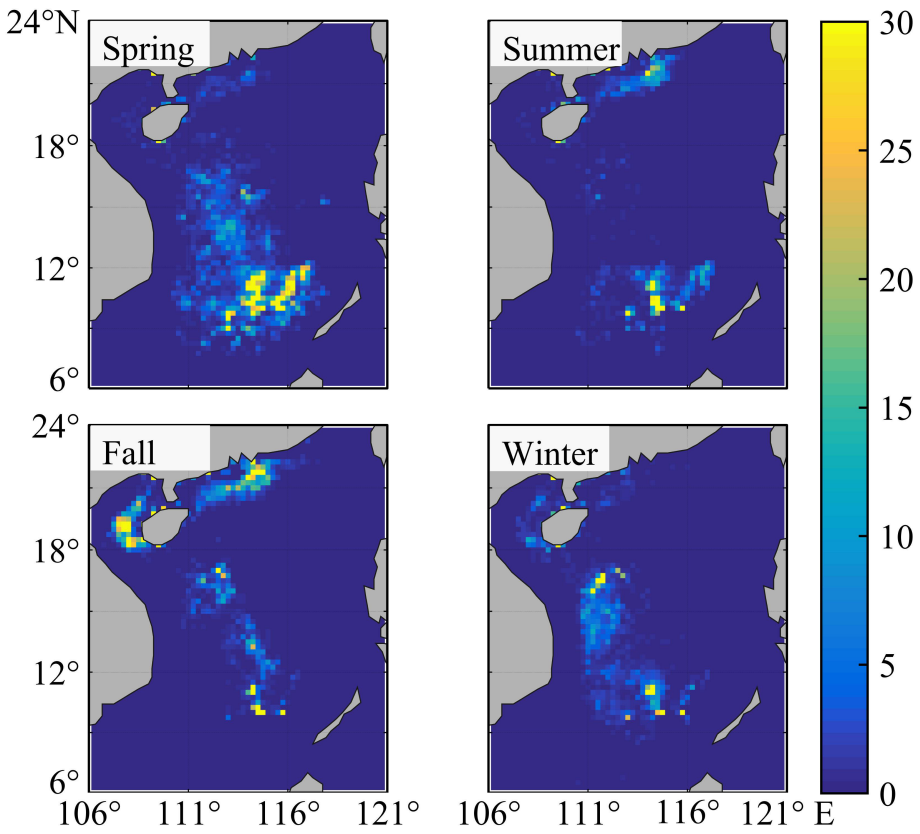


FIGURE 7
Seasonal distribution of large-size light falling-net vessels in the South China Sea.

Discussion

Zhang Peng et al. (Zhang et al., 2013) found that from 2005 to 2010, the operating areas of the light falling-net fishing vessels were mainly in the deep-water areas near the Zhongsha and Xisha Islands, the outer continental shelf of the Pearl River Estuary, and the Beibu Gulf waters. In this study, we found that from 2017 to 2020, the operating area of the light falling-net

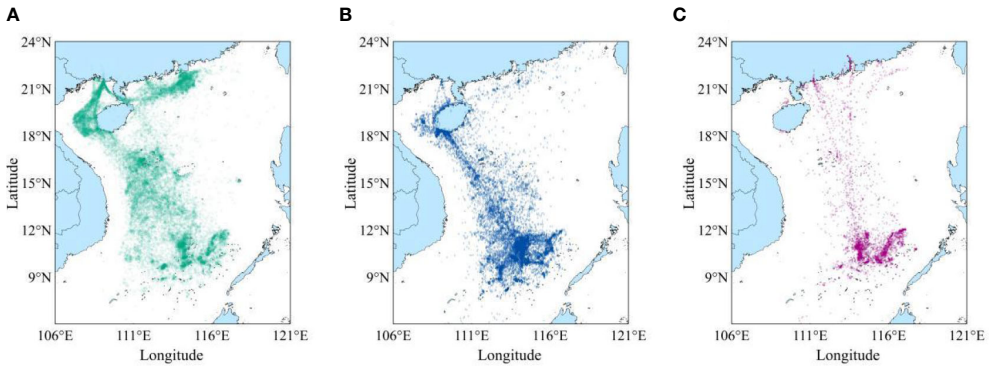


FIGURE 8
Activity of 121 large-size light falling-net vessels according to the vessel monitoring system records. Shown are the distributions of fishing vessels from Guangxi **(A)**, Hainan **(B)**, and Guangdong **(C)**.

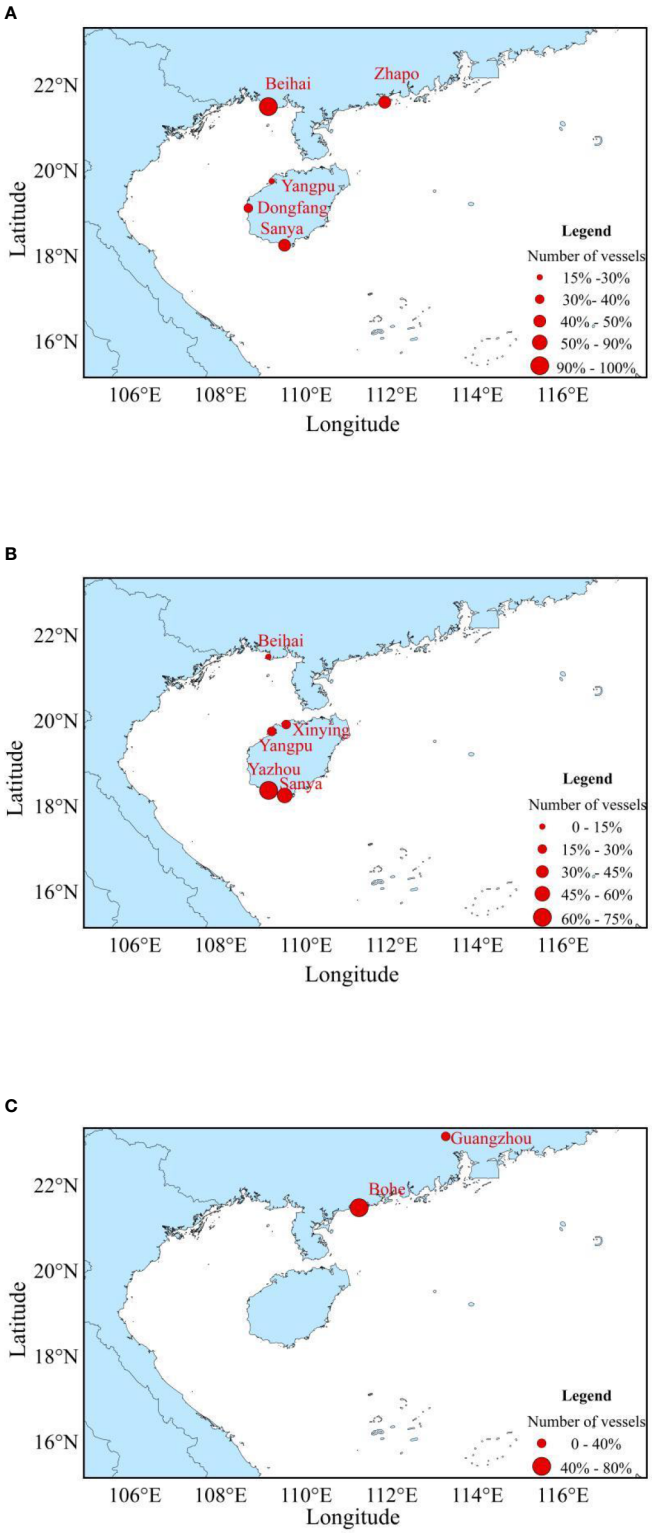


FIGURE 9
Trajectory of the large-size light falling-net vessels from vessel monitoring system records. Shown are the main ports for the fishing vessels of Guangxi **(A)**, Hainan **(B)**, and Guangdong **(C)**.

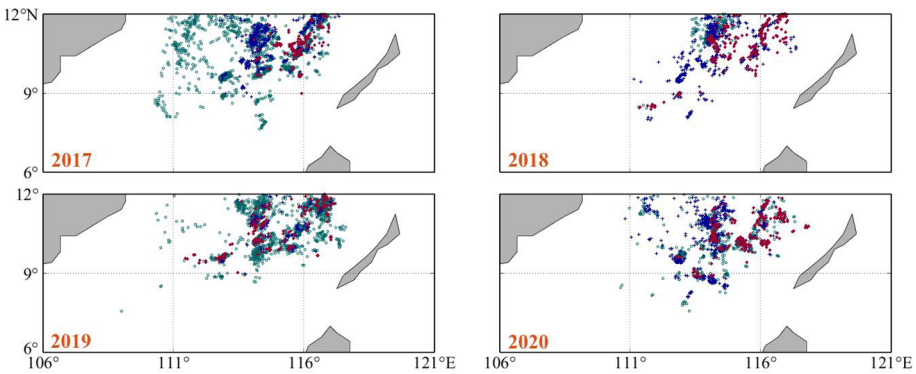


FIGURE 10
Distribution of fishing vessels in the area south of 12°N during the fishing moratorium.

fishing vessels expanded to the waters near the Nansha Islands. Compared with the fishing waters of fishing vessels from 2005 to 2010 (Table 2), the operating areas of the fishing vessels changed greatly throughout the different seasons. Additionally, the main operating season for light falling-net fishing vessels in the Zhong and Xisha waters has changed from spring to winter. The operating time in the sea area off the Pearl River Estuary changed from August to September and December to February (next year) to just August to September. Now the main operation area in spring is the Nansha Sea area. The number of fishing days in different areas throughout the year changed as well. We also found that the light falling-net fishing vessels' material changed from wooden with outdated technical equipment to large steel fishing vessels. As the tonnage, sailing period, and wind resistance of fishing boats have all increased, the operating area of the light falling-net fishing vessel has also extended southward from the sea area closer to the port, and the number of fishing days has also increased. Owing to the increased pressure of offshore fishing and the decline of traditional fishery resources on the continental shelf, the fisherfolk in the northern part of the SCS need to find new outlets. Moreover, the *Sthenoteuthis oualaniensis* caught by large-size light falling-net vessels are in the deep-water areas of the south-central SCS. Thus, the area with the highest number of fishing days per year has changed from the continental shelf waters off the mouth of the Pearl River to the Nansha Islands waters.

Factors influencing fishing vessel distribution

In this study, we compared the distribution areas of fishing vessels in the SCS throughout the year and found that there were differences in the distribution of large-size light falling-net fishing vessels in various provinces (regions) of the SCS. We hypothesized that different fishing vessels have their regular fishing areas. Additionally, large-size light falling-net fishing vessels are affected by various factors including monsoons, typhoons, moratorium system, and traditional festivals such as Chinese New Year. The seasonal distribution pattern of fishing vessels can be identified to some extent by analyzing the spatial variation of large-size light falling-net fishing vessels throughout the year (Zhang, 2017). When comparing the number of operating days of large-size light falling-net fishing vessels from 2017 to 2020, the highest operating intensity of fishing vessels across all seasons was in the spring, followed by autumn, and the lowest was in the winter or summer. In the SCS, large-size falling-net vessels were mainly concentrated in the Zhongsha and Xisha Islands waters and extended southward to the sea areas near the Nansha Islands. Owing to the implementation of the fishing moratorium system (Ministry of Agriculture, 2018), summer is the low period of large-size light falling-net vessel operations, during which almost no fishing vessels operate in the SCS offshore from June to July. However, according to regulations from the Ministry of Agriculture of the PRC, fishing vessels with

TABLE 1 The proportion of fishing vessels in different provinces during the fishing ban.

Year	Guangxi (45.45%)	Hainan (45.45%)	Guangdong (9.1%)
2017	38.84%	33.06%	3.31%
2018	10.74%	30.58%	8.26%
2019	42.98%	18.18%	4.96%
2020	15.70%	28.10%	4.95%

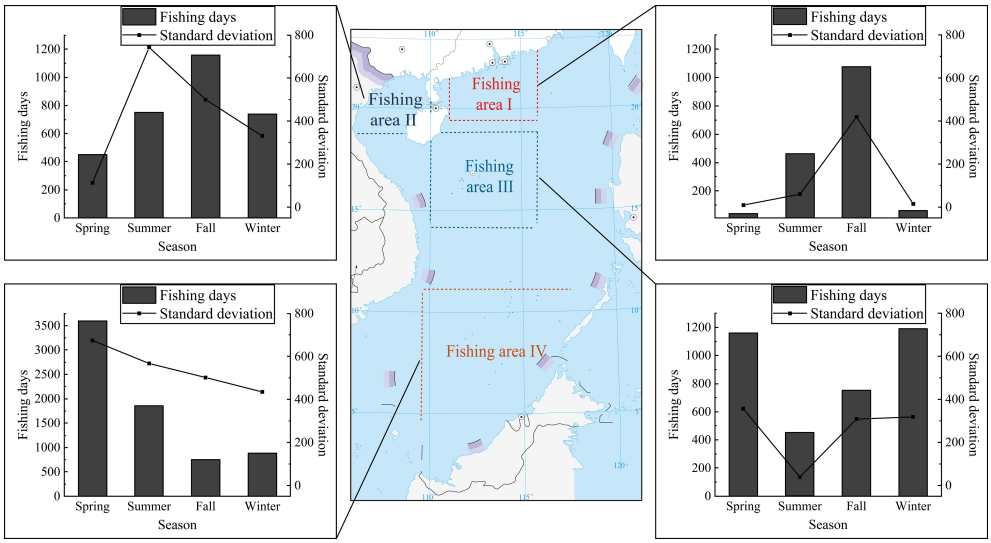


FIGURE 11
Delimitation of fishing zones and seasonal numbers of fishing days in different fishing zones.

a special fishing permit for Nansha are allowed to travel to the Nansha waters south of 12°N to fish during the fishing moratorium. Thus, a small number of fishing vessels remain distributed off the SCS in June and July. At the end of the fishing moratorium, fishing vessels start to operate out of the harbor. Along with the outer SCS, there are also fishing vessels along the coast in August. Because fishing vessels operate in the area south of 12°N during the fishing moratorium, the vessels receiving a special permit need to be determined by lottery. The number of fishing vessels operating in this area varies slightly from year to year in different provinces. In the autumn, large-size light falling-net vessels in the SCS were affected by typhoons. The vessels began moving from the outer sea to the waters of the Pearl River Estuary, the waters near the Dongsha Islands, and the waters of the Beibu Gulf. Besides the summer, winter is the season with the least widespread of fishing vessels. Most fishing vessels operate in offshore waters or even cease production and rest in the harbor owing to the influence of cold air as well as high winds and waves at sea. During this period, the Spring Festival also slows down fishing boat operations. However, because the Chinese Spring Festival does not occur on a fixed date each year, there were annual variations in fishing days and zones between January and February.

In fact, in addition to the above factors, we can link the description of the spatial dynamics to the characteristics of the environment and of the stocks.

The main catches of large-size light falling-net fishing vessels include Sthenoteuthis oualaniensis and Auxis thazard. Relevant studies have shown that SST and SSH are important factors affecting fish growth, reproduction, and migration (Fan et al., 2015; Xie et al., 2020; Yan et al., 2021b). For example, when the sea surface height is higher than the average sea surface, the rich nutrients and bait of the bottom layer are continuously replenished upward, which in turn promotes the growth and reproduction of pelagic fishes.

SCS is affected by the southwest monsoon in summer (June-August), the southwest monsoon retreats from north to south in autumn (September-November), and the northeast monsoon in winter (December-February) (Ministry of Agriculture, 2018).

TABLE 2 A comparison of fishing area for large-size light falling-net vessels in the South China Sea.

Year	Fishing area				
	Zhongsha and Xisha Islands waters	The Pearl River Estuary	Beibu Gulf	Nansha Islands waters	Maximum days of fishing
2005-2010	March-May	August-September December-February (next year)	October-November		The Pearl River Estuary
2010-2020	December-February (next year)	August-September	October-November	March-May	Nansha Islands waters

In addition, SCS is located in subtropical and tropical waters, with a wide variety of plankton and complex composition, suitable for fish growth (Shao and Chen, 2008). The phytoplankton region belongs to the single-peak seasonal variation type in SCS. It gradually rises from March, peaks in June, and then gradually declines, forming a yearly low from October to February. All of these factors can affect the distribution of fishing vessels.

In this study, we specifically compared the behavior of fishing vessels by province. They differed in terms of the sailing area, ports, and fishing. We considered that fishing vessels from different provinces are influenced by fishing policies, culture, and operating experience.

Characteristics of the VMS in the SCS

Currently, there are many modern methods for monitoring fishing vessels in the SCS, such as the Auto identity system (AIS), VMS, electronic monitoring (EM), synthetic aperture radar (SAR), and nightly satellites (Table 3). The main purpose of the AIS is to help prevent the collision of fishing boats and effectively guarantee the navigation safety of ships. It can be divided into Class A and Class B (Li et al., 2021). Presently, the AIS equipment installed by Chinese fishing vessels is Class B. Class B broadcasts at a lower frequency and has a shorter signal transmission distance than Class A. The efficiency of AIS with Class B equipment is extremely low where there are many ships. It is difficult to receive AIS signals in the outer waters of the SCS (Dhar, 2016). Thus, the AIS can barely be used on the near shore of the SCS and is not suitable for the outer sea. In China, the VMS was first applied in the SCS. It has high temporal resolution and wide coverage. From 2000 to the present, almost all fishing vessels in the SCS have been equipped with the VMS. The VMS transmits information through the Beidou satellite with high and stable data transmission quality. The EM is useful because it is easy to record, store, and search for data. But EMS relies on manual records to collect data, and there are cases where the reporting is false, and the data are not standardized. Data collection using the SAR has the characteristics of being all-day, all-weather, wide-range, and multi-parameter. The SAR images can accurately locate vessels through detailed images. Additionally, it is less affected by clouds and rain than the other methods. But it has only two daily imaging opportunities. The SAR processes and stores large amounts of data in ground

TABLE 3 A comparison of selected methods for vessel monitoring in the South China Sea.

Methods	Data Content	Public data	Frequency	Advantages	Disadvantages
AIS	Position, MMSI, ship's length and width, IMO number, name, type speed, heading	Public	Velocity ≤ 2 kn, 3 min Velocity>2 kn, 30 s (The AIS installed on fishing vessels in the South China Sea are Class B)	Real-time, uninterrupted, wide range, high spatial and temporal resolution.	The AIS efficiency is extremely low where there are many ships, especially with Class B equipment. It is difficult to receive AIS signals in the outer waters of the SCS. Only a few analyses of the current behavior and activities of fishing vessels in the northern SCS using AIS data (Guan et al., 2021).
Beidou VMS	Ship ID, longitude, latitude, speed, course, time stamp	Non-public	Generally, 3 min (Varies by region)	VMS was first applied in the SCS and had a wide coverage. Now almost all fishing vessels in the SCS are equipped with VMS. It can monitor the distribution and operation characteristics of fishing vessels in the SCS (Yu et al., 2021; Zhou et al., 2021; Qian et al., 2022).	Only based on the speed and heading in the VMS data to determine the light falling-net cannot determine whether the fishing vessel is operating. The Beidou terminal can only obtain the vessel position data when it is turned on. And it is difficult to share information between different countries (Nicolas et al., 2018).
EMS	Location, various measures for fishing activities, fishing area, fishing time and the catch	Non-public	Almost continuously	Easy recording, easy storing, simple searching and high utilization.	EMS relies on manual records to collect data, and there are cases where the reporting is false and the data is not standardized. Comparing the phishing information of the electronic phishing log with the phishing information obtained from the Beidou location data, there is a deviation (Feng et al., 2019).
SAR	Location	Public	Only two daily imaging opportunities.	Detailed images of ships sailing in the ocean can be provided at different resolutions (up to 3 m resolution). It is less sensitive to clouds and rain.	A large amount of data is usually processed and stored in ground stations. Signal coverage is limited and the main monitoring of fishing vessels in SCS coastal areas. It is also unable to identify fishing vessels by system images alone.
Nightly satellite	Time, fishing vessel lights	Public	Transits once a night	Identify large-size light falling-net fishing vessels according to the lights	Cloud conditions, the zenith angle of satellite observation, and strong lights not from fishing vessels can affect the observation data.

stations. And there are limitations in signal coverage. Thus, the SAR is usually used mainly to monitor the fishing vessels in the SCS coastal areas. The SAR must be integrated with other methods to design and operate a real-time monitoring system for detecting and identifying the ships in the open SCS (Sudhir Kumar, 2019). Nightly satellites monitor large-size light falling-net fishing vessels by their lights (Li et al., 2021; Li et al., 2021). They can determine the vessel position and fishing intensity, as well as the spatial distribution and operation of fishing vessels. However, this method is unlikely to detect other gear types or vessels without bright lights, and it may be affected by factors such as moonlight and the presence of clouds.

Additionally, the VMS in the SCS transmits data through the Beidou satellite, which is characterized by improved confidentiality, higher accuracy, and wider range. Compared with the commonly used fishing log method, the Beidou VMS not only verifies the fishing operation location recorded in the fishing log, but it also achieves higher efficiency and lower manpower and material consumption (McCauley et al., 2016; Feng, 2020).

In addition to studying the spatial-temporal distribution of large-size light falling-net fisheries in the SCS, VMS can effectively obtain information on fishing intensity, analyze the impact of fishing activities on marine biological resources, and study fish stock distribution patterns. We can also use the VMS trajectory characteristics to classify fishing vessel navigation and fishing behavior, as well as quantitatively analyze the evolution of the fishing area of marine fisheries, which has important application value for the scientific development of marine fisheries.

Factors influencing the transmission of information

Considering the factors influencing the information transmitted by the Beidou VMS is important for more accurate monitoring of the distribution of large-size light falling net fishing vessels. The Beidou terminal can only obtain vessel position data when it is turned on. If there are equipment failure, an insufficient power supply, a signal interference, or an artificial shutdown during the monitoring of fishing vessels, it will affect the monitoring results. Here, by comparing the VMS data of fishing vessels in three provinces, we found that there were differences in the frequency of data received in different provinces (districts). Therefore, fishing vessels are likely to be affected by the operating habits of the crew members and management style of the fishing vessels in different provinces (districts) when operating at sea (Zhang et al., 2021). According to Chinese relevant policies, besides the regular domestic oil price subsidies, there will be special oil price subsidies for fishing vessels operating in certain waters during their sailing operations in the waters south of 12°N. Thus, some fishing boats that are not operating in the SCS may open the Beidou VMS and pretend to work to receive subsidies. When fishing vessels are near the shore, they often use other more convenient and economical methods to transmit information such as Beidou short messages, satellite phones, and ultra-short waves (Guo, 2010; Li et al., 2020). The above reasons may make the Beidou VMS more efficient at recording messages in the Outer Sea area than in the nearshore area.

Only through the information of speed and heading of vessel position data cannot judge the operation of the light falling-net fishing vessels. In this study, we set thresholds to analyze fishing vessel operations *via* the operating characteristics of large-size light falling net fishing vessels. We reflected the vessel fishing days by the average number of data points recorded. If there is a problem with the Beidou terminal during operation, it may affect the statistics of the number of fishing vessels going to sea and the analysis of their location at sea. Yan et al. (2015) found that the moon phase significantly affected fishing vessel operations in the SCS with light falling-net vessels. Compared with the dark moon days (from the first to the seventh day of the first month and from the twenty-third to the thirtieth day of the first month), the production of fishing boats on the light moon days (from the eighth to the twenty-second day of the first month) will decrease, especially on the full-moon days (around the fifteenth day of the first month), and a few fishing vessels will not work. In this paper, we did not consider the fact that fishing vessels do not fish during the full moon. Therefore, using the number of days recorded by VMS as the number of fishing vessel operating days would be higher than the actual number of working days for fishing vessels.

Conclusion

In this study, we focused on large-size falling-net fishing vessels in the open SCS. First, using Beidou VMS data from 2017 to 2020, we found a seasonal pattern of change in the distribution of these vessels. The number of operating days for fishing vessels was the highest in the spring, followed by the autumn, and was relatively lower in the winter and summer. We also analyzed the distribution of fishing vessels in the SCS and found that it varies among the different provinces. As the seasons passed, we found that the most concentrated area of fishing vessel distribution shifted and that the main fishing areas and number of fishing days changed. Finally, we compared the VMS with other modern methods for monitoring fishing vessels and found that it is more feasible to study the large, falling-net fisheries in the SCS using the Beidou vessel position data. Using these data will provide more detailed scientific reference information for the SCS fisheries.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding authors.

Author contributions

JQ: Conceptualization, Methodology, Formal analysis, Writing - review & editing, Software. JL: Conceptualization, Methodology, Funding acquisition, Resources, Writing - review & editing. KZ: Writing - review & editing. YQ: Conceptualization, Writing - review & editing, YC: Formal analysis, Writing - review & editing. QW: Funding acquisition, Resources. HT: Data Processing. SM: Funding acquisition, Resources. ZC: Conceptualization, Funding acquisition, Resources. All authors have read and agreed to the published version of the manuscript.

Funding

This work was supported by the Hainan Provincial Joint Project of Sanya Yazhou Bay Science and Technology City (2021JJLH0092), Key Research and Development Project of Guangdong Province (2020B1111030001), and the Central Public-interest Scientific Basal Research Fund, CAFS (2020TD05).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Bastardie, F., Nielsen, J. R., Ulrich, C., Egekvist, J., and Degel, H. (2010). Detailed mapping of fishing effort and landings by coupling fishing logbooks with satellite-recorded vessel geo-location. *J. Fisheries Res.* 106 (1), 41–53. doi: 10.1016/j.fishres.2010.06.016

Cao, S. J., Huang, S. L., and Guo, W. L. (2002). Discussion on adopting the vessel monitoring system in Chinese fishery management. *J. J. Shanghai Fisheries University* 01), 89–93.

Dhar, S. K. (2016). *Addressing challenges with big data for maritime navigation: AIS data within the great lakes system* (History, Computer Science: D. The University of Toledo).

Fan, J. T., Zhang, J., Feng, X., and Chen, Z. Z. (2015). Fishery forecast research of purpleback flying squid in nansha area based on habitat mode. *J. South China Fisheries Sci.* 11 (5), 20–26. doi: 10.3969/j.issn.2095-0780.2015.05.003

Feng, F. (2020). *Current status analysis of marine fishery resources carrying capacity in guangdong province based on electronic fishing logs* (D. Tianjin: Tianjin Agricultural University).

Feng, B., Yan, Y. R., Zhang, Y. M., Yi, M. R., and Lu, H. S. (2014). A new method to assess the population of *Sthenoteuthis oualaniensis* in the south China Sea. *J. Prog. In Fishery Sci.* 35 (04), 1–6. doi: 10.11758/yykxjz.20140401

Feng, F., Zhou, Y. B., Chen, S., Xie, E. G., and Wu, Q. E. (2019). Comparative study on credibility of fishing information of electronic fishing logs of trawlers and fishing information acquired by beidou position data. *J. South China Fisheries Science* 15 (06), 65–74. doi: 10.12131/20190110

Guan, Y. N., Zhang, J., Zhang, X., Li, Z. W., Meng, J. M., Liu, G. W., et al. (2021). Identification of fishing vessel types and analysis of seasonal activities in the northern south China Sea based on AIS data: A case study of 2018. *J. Remote Sensing* 13, 1952. doi: 10.3390/RS13101952

Guo, Y. (2010). Boosting to fishery economy and management by satellite navigation and information-based communication. *J. Chin. Fisheries Economics* 28 (03), 37–42.

Jennings, S., and Lee, J. (2012). Defining fishing grounds with vessel monitoring system data. *J. ICES J. Mar. Science* 69 (1), 51–63. doi: 10.1093/icesjms/fsr173

Jiang, M., Ma, S. W., Ma, Z. H., Zou, J. W., Wu, S. S., Lin, F. H., et al. (2018). Economic benefit analysis of trawlers and light trap and falling net fishing vessels in south China Sea. *J. Insights Aquaculture Biotechnol.* 2 (1), 5.

Jin, X. S., Dou, S. Z., Shan, X. J., Wang, Z. Y., Wan, R. J., and Bian, X. D. (2015). Hot spots of frontiers in the research of sustainable yield of Chinese inshore fishery. *J. Prog. In Fishery Sci.* 36 (01), 124–131. doi: 10.11758/yykxjz.20150119

Lambert, G. I., Jennings, S., Hiddink, J. G., Hintzen, N. T., Hinz, H., Michel, J. K., et al. (2012). Implications of using alternative methods of vessel monitoring system (VMS) data analysis to describe fishing activities and impacts. *J. Ices J. Mar. Science* 69 (4), 682–693. doi: 10.1093/icesjms/fss018

Li, J. J., Cai, Y. C., Zhang, P., Jing, Z. Y., Wu, Q. E., Qiu, Y. S., et al. (2021). Satellite observation of a newly developed light-fishing "hotspot" in the open south China Sea. *J. Remote Sens. Environment* 256 (6), 112312. doi: 10.1016/j.rse.2021.112312

Li, B., Chen, G. B., Guo, Y., Chen, Z. Z., Zhang, J., and Wang, D. X. (2016). Hydroacoustic assessment of spatial-temporal distribution and biomass of fishery resources in the central south China Sea. *J. South China Fisheries Science* 12 (4), 28–37. doi: 10.3969/j.issn.2095-0780.2016.04.004

Li, A., Xu, S., Wang, Y., Wang, Z. Z., and Guo, Y. D. (2020). Research on the application of fishery ultrashort wave shore platform coverage prediction. *J. China Water Transport.* 20 (01), 73–74.

Li, J. J., Zhang, P., Cai, Y. C., Zhang, Q. L., Zhang, K., Jing, Z. Y., et al. (2021). Performance of VMS and nightly satellite in monitoring light fishing vessels in the open south China Sea. *J. Fisheries Res.* 243, 106100. doi: 10.1016/j.fishres.2021.106100

Li, X. E., Zhou, L., Xiao, Y., Wu, W. Z., Su, F. Z., and Shi, W. (2021). Spatial characteristics mining of fishing intensity in the northern south China Sea based on fishing vessels AIS data. *J. J. Geo-Information Science* 23 (05), 850–859. doi: 10.12082/dqxxkx.2021.200328

McCauley, D. J., Woods, P., Sullivan, B., Bergman, B., Jablonicky, C., Roan, A., et al. (2016). Ending hide and seek at sea. *J. Science* 351 (6278), 1148–1150.

Mills, C. M., Townsend, S. E., Jennings, S., Eastwood, P. D., and Houghton, C. A. (2007). Estimating high resolution trawl fishing effort from satellite-based vessel monitoring system data. *J. ICES J. Mar. Science* 64, 248–255. doi: 10.1093/icesjms/fsl026

Ministry of Agriculture (2018). Circular of the ministry of agriculture on adjusting the marine summer fishing moratorium system. J. *Bull. Ministry Agric. Rural Affairs PRC* 03), 54.

Nicolas, L., Guillaume, H., Romy, A., Romain de, J., Béatrice, N., Marzuki, M. I., et al. (2018). Completing fishing monitoring with spaceborne vessel detection system (VDS) and automatic identification system (AIS) to assess illegal fishing in Indonesia. *J. Mar. Pollut. Bull.* 131 (Pt B), 33–39. doi: 10.1016/j.marpolbul.2017.10.016

Qian, J., Li, J. J., Chen, Z. Z., Ma, S. W., Zhang, P., Qiu, Y. S., et al. (2022). An investigation of large-size light falling-net fishing vessels operating in south China Sea based on beidou VMS data. *J. South China Fisheries Sci* 18 (6), 1. doi: 10.12131/20220001

Rocio, J., Omar, S., Mariano, G., Ronan, F., and Sophie, B. (2015). Defining fishing spatial strategies from VMS data: Insights from the world's largest

monospecific fishery. *J. Fisheries Res.* 164, 223–230. doi: 10.1016/j.fishres.2014.12.004

Shao, F., and Chen, X. J. (2008). Relationship between fishing ground of symlectoteuthis oualaniensis and sea surface height in the northwest Indian ocean. *J. Mar. Sci.* 32 (11), 88–92.

Su, L., Chen, Z. Z., Zhang, P., Li, J., Wang, H. H., and Huang, J. X. (2018). Catch composition and spatial-temporal distribution of catch rate of light falling-net fishing in central and southern south China Sea fishing ground in 2017. *J. South China Fisheries Science* 14 (05), 11–20. doi: 10.3969/j.issn.2095-0780.2018.05.002

Sudhir Kumar, C. (2019). Study of synthetic aperture radar and automatic identification system for ship target detection. *J. J. Ocean Eng. Science* 4, 173–182. doi: 10.1016/j.joes.2019.04.002

Tang, X. F., Zhang, S. M., Fan, W., and Pei, K. Y. (2020). Voyage analysis of jiangsu fishing vessels based on beidou position data. *J. Fishery Modernization* 47 (01), 63–71.

Wei, P., Wang, X. H., Ma, S. W., Zhou, Y. B., Huang, Y. B., Su, Y. J., et al. (2019). Analysis of current status of marine fishing in the south China Sea. *J. J. Shanghai Ocean University* 28 (06), 976–982. doi: 10.12024/jsou.201280902393

Xie, E. G., Wu, Q. E., Zhou, Y. B., Zhang, S. M., and Feng, F. (2020). Extraction and verification of operational state characteristics of light shield net vessels based on beidou vessel position data. *J. J. Shanghai Ocean University* 29 (3), 392–400. doi: 10.12024/jsou.20190602714

Xie, E. G., Zhou, Y. B., Feng, F., and Wu, Q. E. (2020). Catch per unit effort (CPUE) standardization of *purpleback flying squid sthenoteuthis oualaniensis* for Chinese large-scale lighting net fishery in the open sea of south China Sea. *J. J. Dalian Fisheries University* 35 (03), 439–446. doi: 10.16535/j.cnki.dlhyxb.2019-122

Yan, Y. R., Feng, B., and Lu, H. S. (2009). Comparative analysis on fishing capability of two light attracting commercial fishing methods around zhongsha and xisha islands sea areas, south China Sea. *J. South China Fisheries Science* 5 (06), 59–64. doi: 10.3969/j.issn.1673-2227.2009.06.011

Yang, L., Lu, H. S., Wu, Z., Zhang, X. F., Yan, Y. R., Liu, G. M., et al. (2002). *Marine fishing gear and fishing method in the south China Sea* (Guangzhou: Guangdong Science & Technology Press).

Yang, L., Zhang, P., Zhang, X. F., and Tan, Y. G. (2002). Main problems for development of marine fishing gears in the south China Sea. *J. Fishery Inf. Strategy* 11), 7–10.

Yan, L., Li, J., Zhang, P., Yang, B. Z., and Wang, T. (2021a). Effects of spatiotemporal and environmental factors on the fishing ground of *Sthenoteuthis oualaniensis* in the south China Sea based on the generalized additive model. *J. Mar. Sci. Bulletin* 40 (2), 217–224. doi: 10.11840/j.issn.1001-6392.2021.02.011

Yan, L., Li, J., Zhang, P., Yang, B. Z., and Wang, T. (2021b). Effects of spatiotemporal and environmental factors on the fishing ground of *Sthenoteuthis oualaniensis* in the south China Sea based on the generalized additive model. *J. Mar. Sci. Bulletin* 40 (02), 217–223. doi: 10.11840/j.issn.1001-6392.2021.02.011

Yan, L., Zhang, P., Yang, L., Yang, B. Z., Chen, S., Li, Y. N., et al. (2015). Effect of moon phase on fishing rate by light falling-net fishing vessels of *Symplectoteuthis oualaniensis* in the south China Sea. *J. South China Fisheries Science* 11 (03), 16–21. doi: 10.3969/j.issn.2095-0780.2015.03.003

Yan, L., Zhang, P., Yang, L., Yang, B. Z., and Tan, Y. G. (2014). Catch composition of light falling-net fishery in the central and southern south China Sea in spring of 2011. *J. South China Fisheries Science* 10 (03), 97–103.

Yu, J., Zhou, Y. B., Tang, Z. C., Chen, G. B., and Zeng, L. (2021). Fishing characteristics of double trawler using vessel monitoring system data. *J. Mar. Sci.* 45 (10), 40–48. doi: 10.11759/hykx20200825001

Zhang, S. Y. (2017). *Research on fishing dynamic changes in the south China Sea using nighttime light data* (D. Nanjing: Nanjing University).

Zhang, R. H. (2020). *A study on fishing behavior of drift gillnet vessels in China based on big data on vessel positions* (D. Shanghai: Shanghai Ocean University).

Zhang, S. M., Fan, W., Zhang, H., Yang, S. L., Tang, F. H., and Zhu, W. B. (2018). Dynamic monitoring and analysis of number of fishing vessel voyages in hainan province based on beidou position data. *J. South China Fisheries Science* 14 (05), 1–10. doi: 10.3969/j.issn.2095-0780.2018.05.001

Zhang, K., Liao, B. C., Xu, Y. W., Zhang, J., Sun, M. S., Qiu, Y. S., et al. (2017). Assessment for allowable catch of fishery resources in the south China Sea based on statistical data. *J. Haiyang Xuebao* 39 (08), 25–33. doi: 10.3969/j.issn.0253-4193.2017.08.003

Zhang, L., Li, Y., Lin, L. S., Yao, Z., Yan, L., and Zhang, P. (2016). Fishery resources acoustic assessment of major economic species in south-central of the south China Sea. *J. Mar. Fisheries* 38 (6), 577–587. doi: 10.13233/j.cnki.mar.fish.2016.06.003

Zhang, S. M., Pei, K. Y., Wu, Z. L., Chen, T. F., Fan, W., and Zhu, W. B. (2021). Quantitative analysis of cruise duration and range of Chinese fishing vessels based on VMS. *J. J. Shanghai Ocean University* 30 (02), 311–319. doi: 10.12024/jsou.20191102856

Zhang, J., Qiu, Y. S., Chen, Z. Z., Zhang, P., Zhang, K., Fan, J. T., et al. (2018). Advances in pelagic fishery resources survey and assessment in open south China Sea. *J. South China Fisheries Science* 14 (006), 118–127. doi: 10.12131/20180037

Zhang, S. M., Tang, F. H., Zhang, H., Fan, W., and Huang, H. W. (2014). Research on trawling tracing based on beidou vessel monitoring system data. *J. South China Fisheries Science* 10 (3), 15–23. doi: 10.3969/j.issn.2095-0780.2014.03.003

Zhang, H., Wu, Z. L., Zhou, W. F., Jin, S. F., Zhang, P., Yan, L., et al. (2016). Species composition, catch rate and occurrence peak time of *Thunnidae* family in the fishing ground of light falling-net fisheries in the nansha islands area of the south China Sea. *J. Mar. Fisheries* 38 (2), 140–148. doi: 10.13233/j.cnki.mar.fish.2016.02.004

Zhang, S. M., Yang, S. L., Dai, Y., Fan, W., and Huang, H. W. (2014). Algorithm of fishing effort extraction in trawling based on beidou vessel monitoring system data. *J. J. Fisheries China* 38 (8), 1190–1199. doi: 10.3724/SP.J.1231.2014.49135

Zhang, P., Zeng, X. G., Yang, L., Peng, C. H., Zhang, X. F., Yang, S., et al. (2013). Analyses on fishing ground and catch composition of large-scale light falling-net fisheries in south China Sea. *J. South China Fisheries Science* 9 (03), 74–79.

Zhang, R. H., Zhang, S. M., Chen, X. Z., Fan, W., and Zhu, W. B. (2019). Research on extract floating gill net hauls and direction based on beidou data. *J. Mar. Fisheries* 41 (2), 169–178. doi: 10.13233/j.cnki.mar.fish.2019.02.005

Zhang, P., Zhang, J., Li, Y., Zhang, R., Lin, L. S., Yan, L., et al. (2016). An exploratory fishing survey of light falling-net fisheries in the central and southern south China Sea in autumn. *J. South China Fisheries Science* 12 (2), 67–74. doi: 10.3969/j.issn.2095-0780.2016.02.010

Zhang, S. M., Zhang, H., Tang, F. H., Fan, W., and Huang, H. W. (2016). Method of extracting trawling effort based on vessel monitoring system. *J. Mar. Sci.* 40 (3), 146–153. doi: 10.11759/hykx20140217002

Zheng, L. B., Ding, T. M., Ai, W. Z., and Lou, R. M. (2019). Application and countermeasures of beidou satellite navigation system in vessels. *J. Shipping Manage.* 41 (10), 31–34. doi: 10.13340/j.jsm.2019.10.011

Zheng, T., and Tang, Y. (2016). Analysis of current status of Chinese marine fishing fleet of south China Sea area. *J. J. Shanghai Ocean Univ.* 25 (04), 620–627. doi: 10.12024/jsou.20160201650

Zhou, Y. B., Xie, E. G., Wu, Q. E., and Feng, F. (2021). The relationship between the resources of *Sthenoteuthis oualaniensis* in the open south China Sea and the marine environment. *J. Acta Oceanologica Sinica* 43 (02), 38–48. doi: 10.12284/hyzb2021008

Zou, J. W., Chen, L. F., Lin, J. J., Lu, Z. F., and Yang, B. H. (2014). Analysis on variation and distribution of central fishing ground for light falling-net in offshore of the south China Sea based on statistics of fishery surveillance to guangxi fishing vessels. *J. South China Fisheries Science* 10 (04), 78–84. doi: 10.3969/j.issn.2095-0780.2014.04.013

frontiers | Frontiers in Marine Science

TYPE Original Research
PUBLISHED 06 January 2023
DOI 10.3389/fmars.2022.1004829

OPEN ACCESS

EDITED BY
Junyu He,
Zhejiang University, China

REVIEWED BY
Matteo Zucchetta,
National Research Council (CNR), Italy
Bradley Thomas Furman,
Florida Fish and Wildlife Research Institute, United States

*CORRESPONDENCE
Chiara M. Bertelli
c.m.bertelli@swansea.ac.uk

SPECIALTY SECTION
This article was submitted to Ocean Observation, a section of the journal Frontiers in Marine Science

RECEIVED 27 July 2022
ACCEPTED 13 December 2022
PUBLISHED 06 January 2023

CITATION
Bertelli CM, Bennett WG, Karunarathna H, Reeve DE, Unsworth RKF and Bull JC (2023) High-resolution wave data for improving marine habitat suitability models.
Front. Mar. Sci. 9:1004829.
doi: 10.3389/fmars.2022.1004829

High-resolution wave data for improving marine habitat suitability models

Chiara M. Bertelli*, William G. Bennett, Harshinie Karunarathna, Dominic E. Reeve, Richard K. F. Unsworth and James C. Bull

Faculty of Science and Engineering, Swansea University, Swansea, Wales, United Kingdom

Habitat suitability modelling (HSM) is a tool that is increasingly being used to help guide decision making for conservation management. It can also be used to focus efforts of restoration in our oceans. To improve on model performance, the best available environmental data along with species distribution data are needed. Marine habitats tend to have ecological niches defined by physical environmental conditions and of particular importance for shallow water species is wave energy. In this study we examined the relative improvements to HSM outputs that could be achieved by producing high-resolution Delft-3D modelled wave height data to see if model predictions at a fine-scale can be improved. Seagrasses were used as an exemplar and comparisons at fine-scale showed considerable differences in the area predicted suitable for seagrass growth and greatly increased the importance of waves as a predictor variable when compared with open-source low resolution wave energy data.

KEYWORDS

seagrass, restoration, habitat suitability modelling, Delft-3D wave modelling, Zostera marina

1 Introduction

There is growing interest in the use of nature-based solutions to help mitigate the worst impacts of climate change. The greatest emphasis around this interest is on the restoration of habitats. With 2021 to 2030 defined as the UN decade of ecosystem restoration the ambition for re-building biodiversity has never been greater. This includes increasing focus on restoration in the ocean (Danovaro et al., 2021) with projects concentrating on restoration of oysters, coral reefs, mangroves, seagrasses and saltmarshes (Waltham et al., 2020).

Conservation research requires long term surveying campaigns over large areas, which needs commitment to a well-funded, strategic programme of measurement. Due to

the practical difficulties this poses, there are limited datasets on habitat quality and extent. A consequence to this is that it is very difficult to direct restoration efforts to places where they will have the best chance of success.

Habitat suitability modelling (HSM) is a tool increasingly used to help guide that decision making. Also known as species distribution models (SDM) and ecological niche models (ENM), they all follow the same premise of capturing the realised niche of a species which can be used for different aims (Naimi and Araújo, 2016; Guisan et al., 2017). HSM can be used to predict distributions of rare and/or vulnerable species for the purposes of conservation management (Thompson et al., 2014; Rowden et al., 2017), assessing changes in distributions of key or invasive species from climate change (Valle et al., 2014) and for informing the most suitable area for habitat conservation and restoration (Adams et al., 2016; Hu et al., 2021).

This type of modelling uses environmental parameters that are known to determine or affect the distribution and presence of the species in question and then projects this to predict where that species should be able to exist if these parameters are met (Guisan et al., 2017). To be able to do this successfully, the availability of good species presence or distribution and environmental parameters is needed to be able to provide models with enough data for testing and training (Guisan et al., 2017; Araújo et al., 2019). Environmental parameters will be made up of the conditions that not only limit, but also influence the distribution and range of the species being studied. Accurate habitat presence data is one of the most important factors for successful HSM as inaccuracies in distribution data will lead to constrained models (Araújo et al., 2019). Presence data need to consist of geographical locations of species and therefore point data (coordinates) are often the most useful and easiest to use.

Seagrass meadows provide an excellent case study for understanding the use of HSMs in the marine environment as they are geographically abundant, have a well-defined environmental range within shallow coastal waters, have a growing interest with regard to restoration initiatives, yet remain poorly mapped (McKenzie et al., 2020; Green et al., 2021). In Europe, the field of seagrass restoration, has grown significantly over the last decade as a result of the increased understanding of the extent of degradation to this marine habitat (Gamble et al., 2021). *Zostera marina* is a meadow forming seagrass species with a wide geographical range across the northern hemisphere, between 27 and 70°N. For *Z. marina*, as a submerged aquatic vegetation, the parameters that affect its distribution include light availability (Dennison and Alberte, 1985; Bertelli and Unsworth, 2018), water depth (affecting light attenuation) (Nielsen et al., 2002), sea water temperature (Marsh et al., 1986; Moore and Jarvis, 2008), salinity (Salo et al., 2014), and physical factors such as exposure to waves and currents (Fonseca and Bell, 1998; van Katwijk and Hermus, 2000; Koch et al., 2001). Seagrasses are generally known to exist in shallow, sheltered coastal areas, in soft sediments (Koch et al., 2006). Around the UK and Ireland, these locations are predominantly east or north-east facing bays where they are sheltered from prevailing wind directions and wave fetch, or within lagoons and estuaries (D'Avack et al., 2019).

A review of previous seagrass HSM studies (Bertelli et al., 2022) found the most commonly used environmental parameters to be seawater temperature, bathymetry, light availability, salinity, wave action, substrate and seabed slope. However, the variables ultimately used will also be somewhat dependent upon what environmental data is readily available and its temporal and spatial resolution. Light availability, specifically Photosynthetically Active Radiation (PAR) which is the spectral range of light that is used for photosynthesis, is arguably one of the most important variables to consider as this will determine where plants can survive within coastal waters (Lee et al., 2007; Short et al., 2007; Kuusemäe et al., 2016). Bathymetry is also key in determining where seagrass can exist, as light decreases with depth as it is attenuated (Duarte, 1991). Within its geographical range, *Z. marina* is usually found within a narrow depth range, typically up to 5-10 m deep depending on water clarity (Davison and Hughes, 1998; Nielsen et al., 2002; Krause-Jensen et al., 2003; Lee et al., 2007; Jackson et al., 2013). *Z. marina* can tolerate a wide temperature range from -1° C in Arctic regions to 30°C in the subtropics. However, temperature will affect respiration rates within plants and will have significant influence on life stages, such as flowering and germination. *Z. marina* is also tolerant to a range of salinities and can be found within estuaries as well as fully oceanic conditions, from 18 PSU to 40 PSU (D'Avack et al., 2019).

Where light limitation is not an issue, such as in uniformly shallow areas unaffected by other water quality issues, other physical parameters will be more important in influencing seagrass presence. Seagrasses are exposed to localised hydrodynamics in the form of waves, tides, wind driven currents and wave driven currents (Koch et al., 2006) and these physical factors have been recognised as important factors in affecting spatial distribution and the minimum depth of colonisation (Stevens and Lacy, 2012). Water movement is important for seagrass growth, but where hydrodynamic energy is too high it can become a limiting factor for seagrass growth (Fonseca and Bell, 1998; Peralta et al., 2006). Morphological changes have been found to be associated with increases in local hydrodynamics. Wave energy (a function of wave height and wave period), also has implications for seed burial and seedling development and is therefore one of the most important environmental variables to consider when choosing restoration sites (Van Katwijk et al., 2009; Infantes et al., 2016; Marion et al., 2021). Restoration needs to break negative feedbacks in the system, and many of these relate to the lower physical limits of seedlings (Temmink et al., 2020). Successful seedling establishment has been found to correspond with lower maximum wave heights and lower orbital velocities (Infantes et al., 2011; Marion et al., 2020) and early patch formation vulnerable to hydrodynamic forces (Furman and Peterson, 2015).

The aims of this study were to investigate the use of habitat suitability modelling for predicting the best locations for potential seagrass restoration in four key locations around Wales, UK, and the benefits of using high-resolution wave data for fine-scale predictions. We defined two key challenges in meeting this aim: 1) identifying up-to-date, high resolution seagrass distribution data and appropriate environmental covariate data; 2) Comparing the benefits of using high resolution wave modelled data to low resolution, freely available hydrodynamic data.

2 Methods

2.1 Site descriptions

The initial consideration was to look at HSM for the entire coasts of UK, Ireland and Channel Isles where *Zostera marina* seagrass is found in many locations and presence data was readily available (Supplementary Data Table S1) and secondly, to focus the HSM at the local potential restoration site scale (Figure 1). These areas were chosen anecdotally as they were perceived by our team of experienced seagrass scientists to contain conditions potentially suitable for seagrass such as shallow waters, easterly facing (away from prevailing winds), calm waters from away from any major swell and soft sediments. The finer focus of the areas was also the result of chosen due to consideration that these sites were of potential for restoration due to existing or historical records of seagrass presence. The south coast of the Llyn Peninsula (Area 1) has small *Z. marina* meadows currently present but largely fragmented into small patches but many anecdotal observations of former distributions. West Anglesey (Area 2) has some small patches, and like the Llyn Peninsula also has anecdotal observations of former distributions. The East of Anglesey (Area 3) has no existing known seagrass but some previous records and descriptions, however, historic mining pollution provides good reasons to assume historic loss. Finally, the area in Pembrokeshire (Area 4) has some old records and anecdotal observations but also has no current known seagrass.

Our study also sought to improve upon the freely available hydrodynamic data by using wave models to create high resolution wave information over the areas of interest and to compare differences in outputs. Finally, our study used these data sources to produce decision tools for informing seagrass restoration sites from model outputs.

2.2 *Zostera marina* presence data

To fulfil the aims of this study we sought to obtain presence data for *Z. marina* around the UK and Ireland from all available sources. This involved checking and refining data points to remove outliers and duplicates. *Z. marina* presence data were obtained from sources in Table S1, uploaded into QGIS (QGIS Development Team, 2019) and georeferenced. This allows visualisation of data layers and identification of erroneous points such as coordinates that placed records on land or in deep (>20-30m) offshore waters where occurrence of *Z. marina* is highly unlikely. These points were removed along with duplicate records from the final dataset. Once data were refined, over 2500 presence points remained around the UK, Ireland and the Channel Islands. These point data (Figure 2) were used for testing suitable HSM methods at a broad scale which would then be refined at a smaller spatial scale for potential restoration sites. Polygon data can also be used, but this would limit the number of data sources and require further analysis to get point coordinates needed for the 'sdm' package and analysis method used for this study.

2.3 Environmental predictor variables for *Zostera marina*

We obtained the highest resolution freely available environmental data of importance to *Z. marina* presence for the purpose of refining restoration site choice at a local scale. A review conducted by the authors (Bertelli et al., 2022) identified the most commonly used environmental variables used for predicting seagrass presence using habitat suitability or species distribution modelling. Highest resolution, freely available sources of environmental data were identified. All data were downloaded between September 2020 and February 2021 and covered temporal ranges from 2001 to the time of download (Table S2) and uploaded into QGIS for visualisation.

Environmental data are available in a range of different resolutions, formats and is created using a variety of methods. Photosynthetically available radiation (PAR) at the seabed (From European Marine Observation and Data Network, EMODNet *via* EUSeaMap) is determined from field and satellite data for light in the water column and then by calculating light attenuation from depth and proximity to coast (EUSeaMap, 2012) at a resolution of ~0.3 km. Wave energy (kinetic energy) at the seabed and kinetic energy from currents at the seabed from EMODNet is based upon modelled outputs from the National Oceanographic Centre (NOC) wave and current models (see EUSeaMap, 2012) to a resolution of ~0.3 km. Data from the Copernicus Marine Environment Monitoring Service (CMEMS, https://www.copernicus.eu/en/services/marine) was used as a source of salinity and temperature as daily averages which are calculated from forecast data which can therefore provide temporal ranges at a resolution of ~1.7 km. Bathymetry was available from EMODNet down to ~70 x 116 m resolution for small high-res downloads, or alternatively from British Oceanographic Data Centre BODC (2020) GEBCO (General Bathymetric Chart of the Oceans) at ~0.2 km. All environmental data layers found to be suitable for habitat suitability modelling are compiled in Supplementary Data (Table S2).

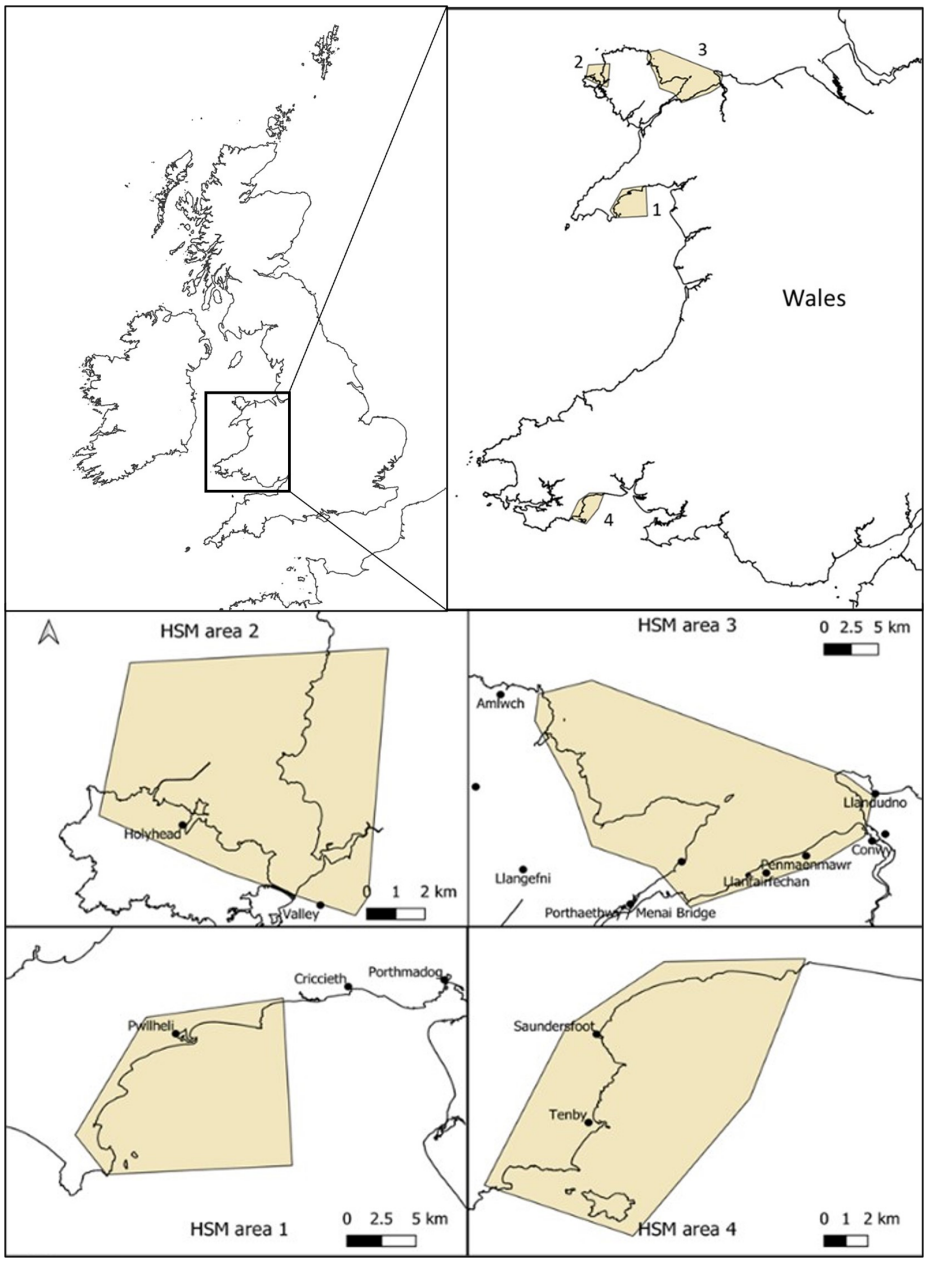


FIGURE 1
Sites for potential *Zostera marina* restoration around Wales (UK) used for high resolution habitat suitability modelling using high resolution wave model data. Areas start at Llyn Peninsula, north Wales (area 1), followed by west Anglesey coast (area 2), east Anglesey (area 3) and finally south Wales, Pembrokeshire (area 4).

2.4 Data modelling

2.4.1 Broad-scale habitat suitability models

Initial habitat suitability models were developed using openly available environmental data layers and *Z. marina* presence data for the whole of the British Isles and Ireland. This was to test the effectiveness of model methods using the available environmental data. Environmental data layers were uploaded into R version 4.0.2 (R Core Team, 2020) as raster files along with *Z. marina* presence point coordinate data as spatial data frames. Five openly available variable datasets were chosen based on suggested predictor variables for seagrass presence identified in Bertelli et al. (2022) - light availability (PAR at seabed), bathymetry, temperature, salinity, wave energy at the seabed and energy from currents at the seabed.

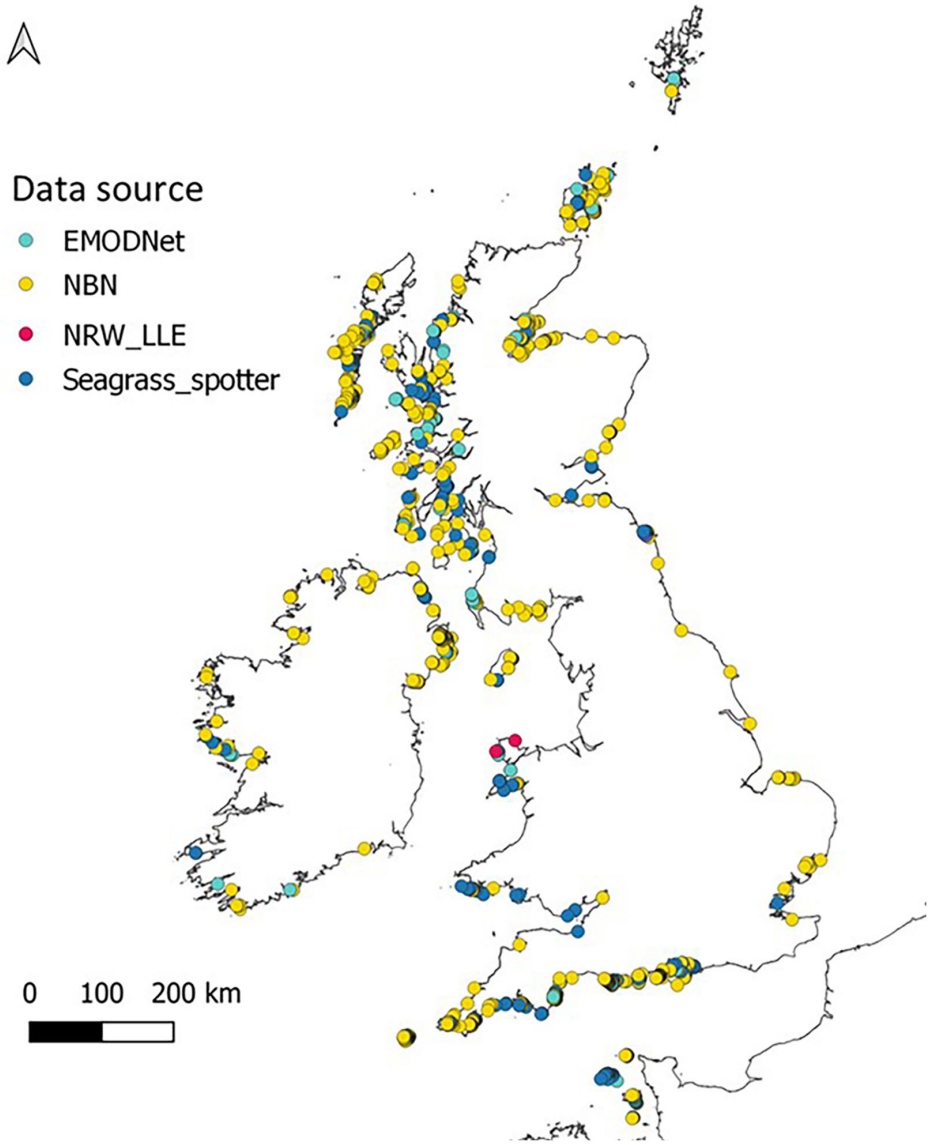


FIGURE 2
Zostera marina presence around the UK, Ireland and the Channel Islands taken from sources listed in Supplementary Data (Table S1) after outliers were removed.

Marine environmental datasets often lack coverage in shallow coastline and pixels covering both land and sea are often excluded (Yesson et al., 2015). To overcome this, a buffering tool was used in QGIS (raster 'fill nodata' tool) to interpolate data from neighbouring pixels to cover coastal areas to a distance of 3 pixels. All environmental data layers were resampled using the 'resample' function and the nearest neighbour method ('ngb') in the 'raster' package in R, so they were all the same resolution as the bathymetry layer (15 arc, sec, ~ 0.4 km), extent and format. All the environmental predictor layers were tested for colinearity and colinear variables below a threshold of 0.9 were removed using VIF (Variance Inflation Factor) tests (*vifstep* and *vifcor*) in R, recommended for variable selection (Naimi et al., 2014; Guisan et al., 2017). Remaining predictor variables were plotted using pairwise correlation plots to visualise and check no collinearity issues remained.

A suite of methods was used for habitat suitability modelling using the 'sdm' package (Naimi and Araújo, 2016) in R. This package allows the use of a wide range of the most common modelling algorithms covering parametric, non-parametric, regression and machine-learning methods to be used all at once. A range of algorithms were chosen to cover the different types of

modelling approaches. These included Generalised Linear Models (GLM), Generalised Additive Models (GAM), Multivariate Adaptive Regression Splines (MARS), Random Forest (RF), Boosted Regression Trees (BRT) and Maximum Entropy (MaxEnt), to include some of the most popular methods in other HSM studies (Valle et al., 2013; Guisan et al., 2017; Bertelli et al., 2022). The models follow a formula of seagrass presence as a function of the predictor variables (a stack of the environmental variable layers) using the different approaches. Working only with presence data, the 'sdm' software takes a range of randomly sampled 'background' points from the study area which are treated as pseudoabsences (Naimi and Araújo, 2016). For the UK and Ireland area, 1050 seagrass presence points were collated with an equal number of pseudoabsences for the HSM. A sample of the presence point data is used to test the model using the bootstrapping method (Naimi and Araújo, 2016). Performance of the models was assessed using area under curve (AUC) of the receiver operating characteristic (ROC) which is produced for each model run (Beca-Carretero et al., 2020). The model-independent variable importance values were also calculated, which are a sensitivity analysis of a model to different predictors (Harisena et al., 2021). This is accomplished by using measures of variable importance that compute a ranking of all the possible species–environment relationships which is one of the most popular measures used (Harisena et al., 2021). Predictions of the probability of presence were then created for each modelling approach. Finally, an ensemble of all the methods (using weighted averages of AUC scores) was used to create prediction outputs in the form of a raster layer indicating probability of suitability for seagrass presence from 0-1 (Naimi and Araujo, 2021).

2.4.2 High resolution wave modelling

To obtain high resolution wave data at each of the four restoration study sites, the openly available computational coastal modelling suite Delft3D was used (Lesser et al., 2004). The Delft3D-WAVE module utilises the third generation spectral SWAN wave model (Booij et al., 1999) to simulate waves in time and space. A Delft3D computational model domain encompassing the Irish Sea with a rectangular 1160 m x 1850 m grid was created to transform offshore waves to the nearshore. High resolution near-shore domains with 50 m x 50 m grid resolution were created for each of the four sites, and nested to the Irish Sea model (Figure 3) to obtain high-resolution wave data required for HSM. Bathymetry for each of the high-resolution domains was created using EMODnet data at 71 m x 116 m resolution. The General Bathymetric Chart of the Oceans (GEBCO) (https://www.gebco.net/data_and_products/gridded_bathymetry_data/) 570 m x 925 m resolution dataset was used to provide additional bathymetry alongside the EMODnet data for the Irish Sea domain. The Irish Sea model was forced by time (hourly) and space-varying (5°) European Centre for Medium-Range Weather Forecasts (ECMWF) ERA5 offshore waves (Hersbach et al., 2018) provided at its offshore boundary. Hourly, 0.5° resolution ERA5 wind data was implemented across the entire model domain to provide atmospheric forcing to the model. Before simulating wave conditions needed for HSM, the Deft3D wave model was validated through a comparison of simulated waves with wave buoy observations at three locations within the larger domain.

A series of wave simulations were carried out using the validated wave model to provide high resolution data at the four sites. The months January, March, August, and October were run for the years 2015 to 2019 providing a total of 20 model runs. Model outputs including significant wave height, mean and peak wave period, and wave direction are provided at one-hour intervals. The time mean significant wave height was calculated for each simulation period, which will be used in HSM. Figure 4 shows an example of this for January.

2.4.3 Fine-scale habitat suitability models

Wave height data from the near-shore wave models, including mean significant wave height, provided higher resolution data for use in the fine-scale HSM. However, the open-source wave energy (EMODNet) data were also clipped to the restoration areas (Figure 1) so that it could also be used within the final models for comparison between the two. As the high-resolution wave data are calculated based on temporal data, four months were chosen based upon seasons that are significant for the life cycle of *Z. marina*. These months were: January as a potential for highest impacts due to wave action and wind speeds, March when seed germination and seedling emergence takes place, August for peak biomass and seed production, and October for seed establishment and as such the suitable planting season for restoration work (Sand-Jensen, 1975; Orth and Moore, 1986; Blok et al., 2018). Salinity and temperature data from Copernicus were obtained for each of these months as an average of the daily means. Other data were not temporally resolved so remained as a single data layer (bathymetry, PAR, slope and currents). All environmental data layers were clipped to the areas of interest (HSM areas 1-4, Figure 2) and resampled to the same resolution as the wave height and bathymetry data using the 'resample' function and the nearest neighbour method ('ngb') in the 'raster' package in R. Seagrass presence records were clipped from the original UK dataset for the areas of interest, which resulted in 32 presence points and a higher number of background 'pseudo-absence' points (n=90) to improve model performance.

Using VIF (Variance Inflation Factor), all the environmental predictor layers were tested for multicolinearity and variables above a standard threshold of 0.9 were removed, used in many studies (Naimi et al., 2014; Perger et al., 2021; Khan et al., 2022). Remaining predictor variables were also plotted using pairwise correlation plots to screen for collinearity (Supplementary Figure S1).

The nine remaining variable data layers (wave height January, mean temp. March, mean temp. August, mean temp. October, mean salinity August, PAR at seabed, bathymetry, currents and slope) were stacked and used as predictor variables for the HSM. The GAM was found to no longer converge, caused by the increase in the number of variables and low number of observations, so another algorithm was used –flexible discriminate analysis (FDA), which is another regression-based model. This process was repeated by replacing the high-resolution wave height data (Delft3D) with the original open source, low-resolution wave energy data (EMODNet) for comparison between models and variable importance.

Areas of suitable habitat were formatted in QGIS to produce a continuous scale of the model output as a 'heat-map' showing probability of suitability (0-1). The layer was then clipped and smoothed to show only the areas with a high probability for both the model results using the low-resolution wave energy data and the high-resolution wave height data. The threshold of 0.6 was chosen as a value that was over 0.5 probability of suitability which returned a reasonable and practical area to work with for future applied restoration. A threshold above 0.6 (0.7 or 0.8) left a considerably reduced area, leaving little scope for practical restoration. A key challenge in the application of HSM is turning the continuous output into a binary output for decision making. We chose 0.6 as an example of this process that practitioners could use but another value could be chosen. Areas of suitable habitat over a threshold of 0.6 were calculated in km^2 using the vector calculator function in QGIS.

3 Results

3.1 Results of broad-scale habitat suitability models

For the seagrass presence points around the UK, Ireland and Channel Islands, the mean values for the predictor variables were calculated to provide an estimated niche for seagrass existence. The mean depth for seagrass presence, calculated from bathymetry, was 6.7 m (± 5.4 Standard Deviation), with mean PAR 12.1 $Mol.phot.m^{-2}.d^{-1}$ (± 8.1 S.D.), wave energy 1671 $N.m^2.s^{-1}$ (± 24577.4 S.D.), energy from currents 119.5 $N.m^2.s^{-1}$ (± 218.5 S.D.), a salinity of 33.8 ppt (± 1.7 S.D.) and temperature of 8.4°C (± 1.1 S.D.).

The predictor variables used for the broad-scale HSM were not found to have any collinearity issues. Variable importance was calculated within the 'sdm' package in R, using methods outlined in Elith et al. (2005) (Naimi and Araújo, 2016). Bathymetry was the most influential variable (63% based on a correlation matrix - COR, and 32.3% based on Area Under Curve - AUC) followed by PAR at the seabed (Table 1). Results from the ensemble models shown in Supplementary Material, Table S4, show that all model algorithms ran successfully with AUC scores ≥ 0.9, Correlation coefficient ≥ 0.76, and True Summary Statistic (TSS) ≥ 0.73.

As bathymetry was found to be the variable of highest importance, large regions of shallow seabed around the UK, Ireland and the Channel Islands were predicted as suitable habitat for seagrass (Figure 5). However, this also included areas of coastline exposed to high wave energy (for example, south-west facing coastlines, see Supplementary Material, Figures S3, S4). Wave energy was only found to have 0.8 to 0.9% importance as a predictor variable, which explains why high energy sites were retained as suitable for seagrass growth (Table 1). This highlights the need for better resolution predictor variables, particularly where wave energy or exposure which is known to have a direct effect on seagrass presence.

3.2 Results of high-resolution wave data

Liverpool Bay, West Pembrokeshire, and Scarweather Centre for Environment, Fisheries and Aquaculture Science (CEFAS) wave buoys provided hourly observations around the Welsh coast. The observations during the period 1st March – 31st March 2016 were utilised for model validation, due to the availability of consistent observations at each of the three buoys. Time series of simulated significant wave heights during this period were directly compared with that of the observed wave heights. The comparisons yielded R^2 values of 0.87 for the West Pembrokeshire and Scarweather buoys, and 0.93 for the Liverpool Bay buoy, indicating that the computational model is able to accurately simulate waves in the Irish Sea.

In Figure 4, the mean January wave height from across the five January model runs (2015-2019) at each of the four study sites is shown, providing insight into the wave characteristics at each location. At all four sites wave heights were greatest offshore, with the greatest for the more exposed Pembrokeshire site (2-2.5 m), and least for the more sheltered East Anglesey site (1.25-1.35 m). The waves quickly reduce in magnitude as the water depth reduces (Figure 3), and where headlands provide natural shelter from the predominant wave approach. This provides large areas nearshore at each site with mean wave heights between 0.5-1 m depending on the site.

3.3 Results of fine-scale habitat suitability models

Many of the mean values for the predictor variables restricted to the seagrass presence in the local restoration areas varied substantially from the broad-scale data. The mean depth for seagrass presence was considerably lower at 0.4 m (± 3.5

TABLE 1 Summary of the variable importance based on Correlation metric and AUC for each of the final models, for the broadscale UK and Ireland model, the fine-scale model using lower resolution wave energy data (open source EMODNet) and fine-scale model using high resolution modelled wave height data (Delft3D).

Summary of relative variable importance	Based on Correlation metric (%)	Based on AUC metric (%)
UK & Ireland model		
Currents	1	0.7
Waves	0.8	0.9
Temperature	4.7	3.6
Salinity	4.1	2
Bathymetry	63	32.3
PAR at seabed	19.3	12.5
Fine-scale model – Low-resolution Wave energy data		
Waves (EMODNet)	8.8	7.9
Mean temp.mar	29.3	16.8
Mean temp.aug	13.6	3.7
Mean temp.oct	19.3	8.3
Mean sal.aug	10.9	4.1
PAR at seabed	31	23.3
Bathymetry	29.6	23
Currents	3	1.9
Slope	14.7	12.9
Fine-scale model – High-resolution modelled wave height data		
Mean wave height.jan	30.1	20.3
Mean temp.mar	5.6	2.3
Mean temp.aug	17.3	11.3
Mean temp.oct	8.1	5.1
Mean sal.aug	3.8	2
PAR at seabed	27.2	16.2
Bathymetry	10.1	7
Currents	2.2	1.2
Slope	15.7	14.2

S.D), wave energy at 160 $N.m^2.s^{-1}$ (± 403.3 S.D.) and energy from currents 51.2 $N.m^2.s^{-1}$ (± 68.8 S.D). Optimal PAR was only slightly higher at 14.1 $Mol.phot.m^{-2}.d^{-1}$ (± 7.3 S.D.) than for the broad-scale seagrass presence. Salinity, temperature, and modelled wave height varied depending on monthly average (see Supplementary Data, Table S3). Mean temperature ranged from 7.7°C in January to 18.7°C in August. Salinity ranged from 30.4 ppt in January to 31.7 ppt in March. Average slope that was calculated from the higher resolution bathymetry data was found to be 89.9°.

Removal of colinear variables left mean wave height for January; mean temperature for March; mean temperature for August; mean temperature for October; mean salinity for August; PAR at seabed; Bathymetry; Energy at seabed due to currents and Slope of seabed. All remaining predictor variables showed a pairwise correlation below 0.7 (Supplementary Material, Figure S1). Variable importance changed significantly in the fine-scale models in comparison to the broadscale models, with wave height becoming much more influential in predicting seagrass presence (Table 1, Supplementary Material, Figure S2). The resulting predictions for suitable habitat for *Z. marina* are shown in Figure 6 for each of the restoration site areas around Wales, with model results shown Supplementary Material, Table S4. All model algorithms resulted in AUC scores of equal to or greater than 0.9 which can be interpreted as good or even excellent predictions based on scales defined in other studies (Araujo et al., 2005). The TSS (True Summary Statistic) is also a useful model validation measure which is independent of prevalence or size of dataset, a limitation for the smaller restoration areas where presence data is low (Allouche et al., 2006). TSS values closer to 1 show higher prediction accuracy, and all models scored above 0.8. Random forest (RF) was found to perform the best (AUC 0.98 ± 0.03 S.D., COR 0.83 ± 0.06 S.D. and TSS 0.95 ± 0.09 S.D.) and flexible discriminate analysis (FDA) the lowest scoring (AUC 0.9 ± 0.03 S.D., COR 0.54 ± 0.17 SD and TSS 0.82 ± 0.1 S.D.). Mean wave height (January) was the most important variable based on the average correlation metric (30.1% ± 24.9 S.D.) and the AUC metric (20.3% ± 20.8 S.D.) for all model runs, followed by PAR at seabed (COR 27.2%, AUC 16.2%, see Table 1).

When the model was repeated with the low-resolution wave energy data in place of the high-resolution wave height data, PAR at the seabed (COR 31% ± 0.3 S.D., AUC 23% ± 0.2 SD), bathymetry (COR 29.6% ± 0.2 SD, AUC 23% ± 0.2 S.D.) and mean temperature in March (COR 29.3% ± 0.3 SD, AUC 16.8% ± 0.3 S.D.) were the most important variables influencing habitat suitability (Table 1). AUC, COR and TSS scores were lower in the model runs when the lower resolution wave energy data was substituted suggesting the model fits were not as successful as when high resolution wave height data was included. Only two methods had an AUC mean score over 0.9 (FDA 0.96 ± 0.05 S.D. and MaxENT 0.92 ± 0.06 S.D.) and none scored a TSS value ≥ 0.8 (Supplementary Material, Table S4).

Figures 6, 7 shows areas with probability ≥ 0.6 of being suitable for *Z. marina* presence for the ensemble model using low-resolution wave energy data and the high-resolution wave height data for the different areas. The total suitable area ≥ 0.6 predicted from the HSM using the high-resolution wave data was more than twice (13.69 km^2) the predicted area from the HSM using the low-resolution wave energy data (6.17C, Figures 6, 7). Within the same areas, the area predicted to be ≥ 0.6 suitable for seagrass from the broad-scale model was 92.3 km^2.

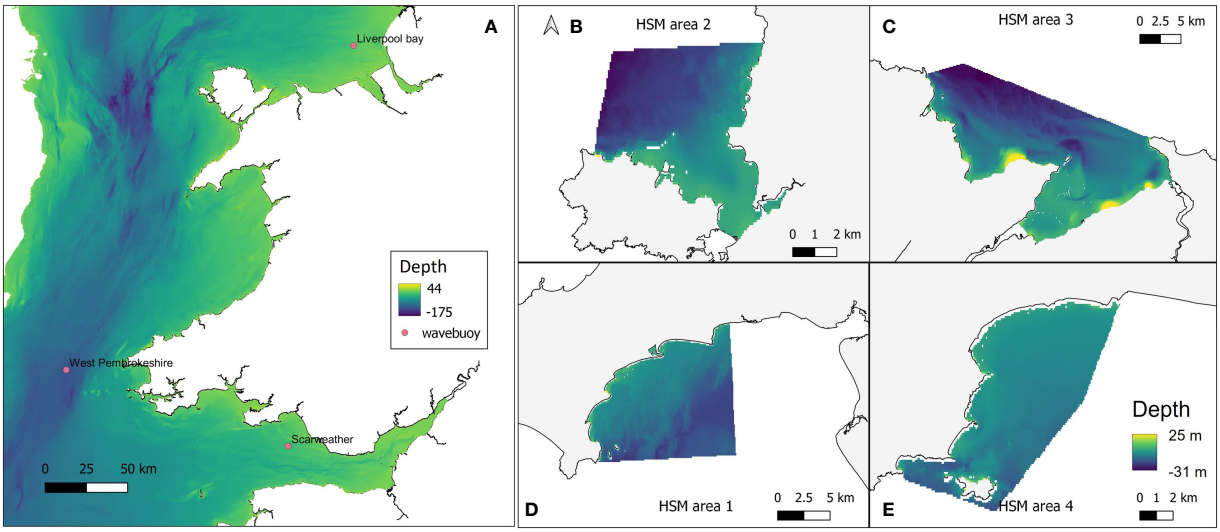


FIGURE 3
(A) Delft3D Large computational domain and bathymetry, encompassing the Irish Sea and extending towards the Atlantic Ocean. Pink markers indicate the positions of the Scarweather, West Pembrokeshire, and Liverpool Bay wave buoys. Grid cells are 1160 m x1850 m resolution. Nested model domains at **(B)** West Anglesey (area 2); **(C)** East Anglesey (area 3); **(D)** Llyn Peninsula (area 1); and **(E)** Pembrokeshire (area 4). Bathymetry data were taken from GEBCO and EMODnet datasets. Some topographical areas were also covered in this data layer which explains the high positive values in depth range, but would not have affected outputs as these areas would not be covered by water so were avoided in modelling.

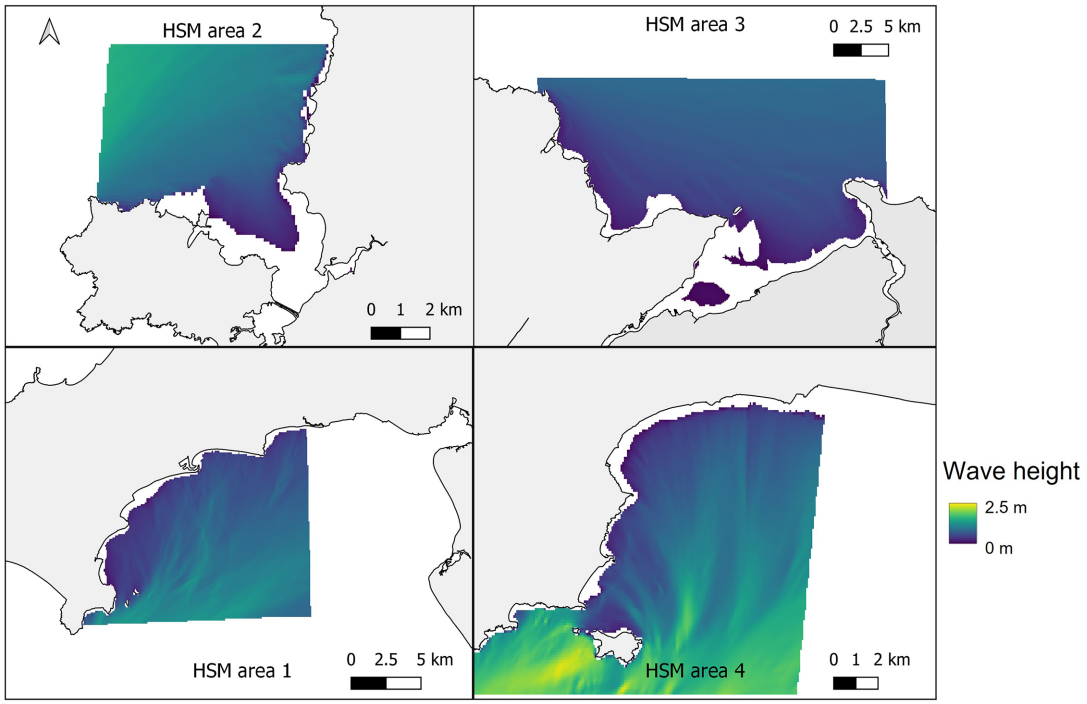


FIGURE 4
Mean significant wave height from across the five January model runs (2015-2019) at the four selected sites. Top left – West Anglesey (area 2), top right – East Anglesey (area 3), bottom left – Llyn Peninsula (area 1), bottom right – Pembrokeshire (area 4).

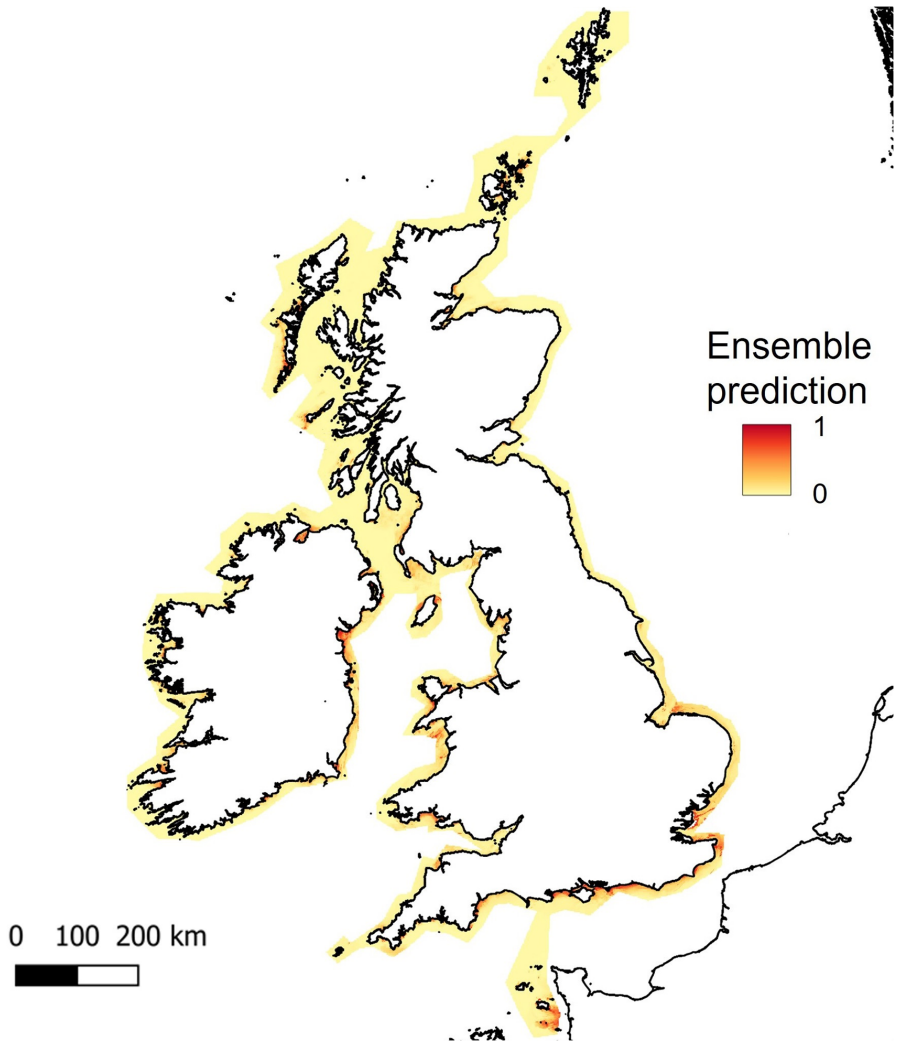


FIGURE 5
Predicted habitat suitability for Zostera marina, based on ensemble model results for UK, Ireland and Channel Isles. The legend shows colour scale for the probability of habitat suitability from yellow (low) to red (high).

4 Discussion

Determining where to invest limited financial resources for restoration or environmental enhancement of species or habitats requires appropriate decision support tools. Habitat suitability models are a useful tool for assisting with this process by identifying areas where species and habitats should be able to exist based upon the environmental data that are available. Here we provide a novel case study that illustrates how the creation of high-resolution environmental datasets, and the use of ensemble modelling techniques can lead to more refined HSMs that better predict these windows of opportunity for restoration, ultimately leading to a reduction in environmental and financial risk attached to major ecological improvement projects.

Our study finds that potential seagrass restoration areas can be defined to a much finer scale and more ecologically representative level by incorporating targeted fine resolution Delft3D wave data in marine habitat suitability models. Model validation results were good for the initial broad-scale models using open access environmental datasets, however the map outputs overpredicted suitability in areas that would not be appropriate for seagrass growth, such as very large swathes of exposed coastlines around Anglesey Island (Natural Resources Wales, 2009) and deep (>15 m) regions such as offshore from the mid Wales coast with wave energy the second to least important variable. The lower resolution of the environmental data available at a broader scale will still identify suitable regions for seagrass growth but the spatial overlap at this scale makes it more difficult for differentiating where conditions are most suitable within smaller areas.

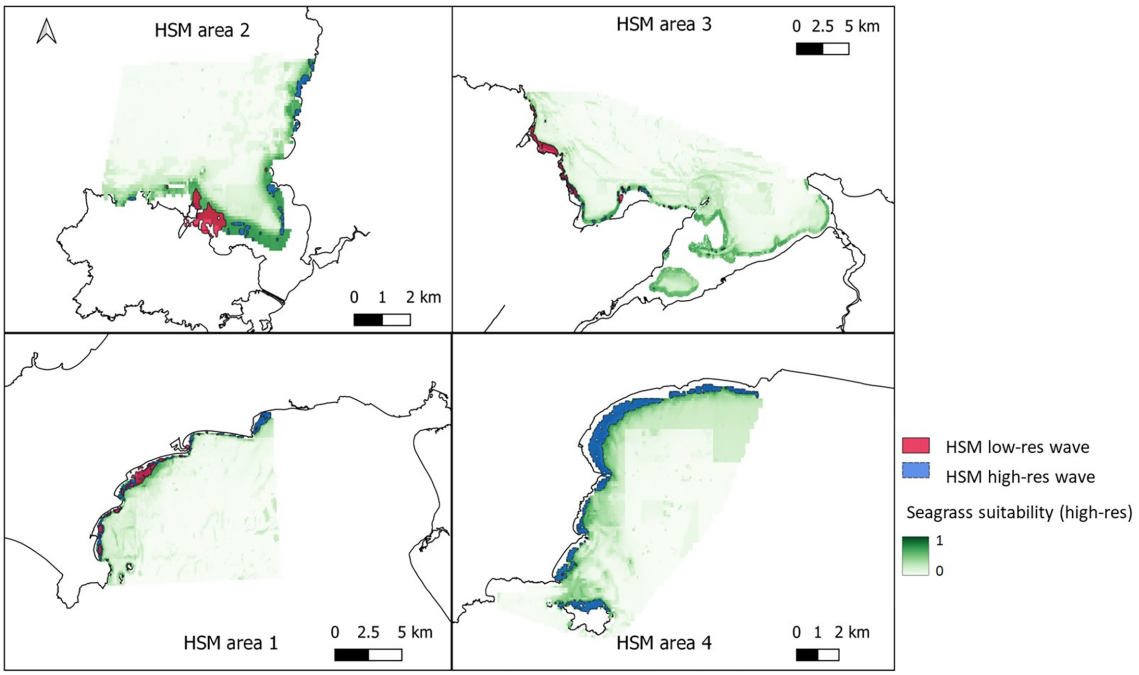


FIGURE 6
Areas of suitable seagrass habitat in green for the four chosen areas around Wales. Top left – West Anglesey (area 2), top right – East Anglesey (area 3), bottom left – Llyn Peninsula (area 1), bottom right – Pembrokeshire (area 4). Maps shows results from ensemble model of 6 methods, using presence only data and all the non-colinear variables including high-resolution wave height data. Overlayed is the vector layer outputs when a threshold of >0.6 probability was applied to the model with low resolution wave data model (pink) and high-resolution wave data model (blue) for comparison.

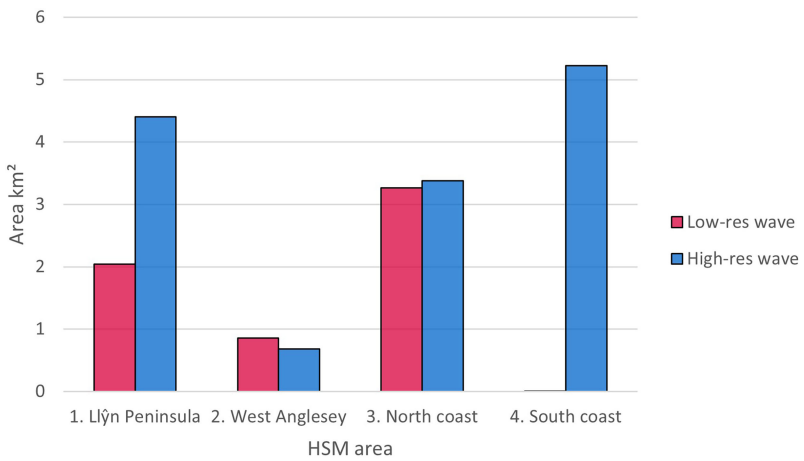


FIGURE 7
Difference in area (km^2) predicted as suitable using ≥0.6 probability threshold for ensemble models using lower resolution wave energy data and high-resolution wave height data.

The results of the ensemble model using Delft3D wave data appear to show appropriate predictions for seagrass growth at a fine scale due to the higher resolution environmental predictors of wave height and bathymetry. Mean wave height for January obtained from the wave model was found to be the most important variable for predicting seagrass suitability which is the month predicted to have the highest wave heights and with suitability increasing as wave height decreases. PAR at the seabed was the second most important variable for predicting seagrass presence in the four areas around Wales.

The open-source wave energy data used in the same ensemble model was less important as a predictor variable, with slope and bathymetry the most important factors. The resulting areas predicted to be highly suitable for seagrass growth was found to be more than 50% lower when using the low-resolution wave data.

The use of ensemble modelling allowed a combination of several model methods to provide predictions which will be more robust to uncertainties that may arise from using a single method (Araújo and New, 2007; Latif et al., 2013). This removes the need for selecting a single 'best' model which limits model selection bias and reduces the need for removing predictor variables which is often the case for step-wise model selection. This method is useful for predictions in new areas where conditions may be different or where data are lacking (Guisan et al., 2017).

The interpretation of HSM outputs can be relatively subjective based upon prior knowledge of sites and the species in question. The AUC, COR and TSS scores allow model comparisons and will indicate those which best fit, however when the outputs are very similar, visual interpretation is needed to assess if the results are realistic. The model results show higher AUC, COR and TSS scores when the high-resolution wave height data were included for the smaller potential restoration areas. More areas were suitable in areas 1 and 4, on the Llyn Peninsula and southwest Wales coast, than the Anglesey and north coast areas (2 and 3). This could be due to areas 2 and 3 having higher areas of intertidal shoreline. Due to large resource requirements, the wave model within Delft3D was not coupled to the tidal processes, and as such wave modelling was carried out for the mean water level. The wave modelling therefore does not provide data in the upper intertidal, limiting prediction of wave heights in shallow extents, especially where land is also obstructing potential wave energy, and therefore this restricts the model areas to subtidal zones. However, this issue was consistent with other open-source environmental data such as PAR at the seabed, which highlights the difficulties in extending marine HSMs into intertidal areas. The wave height data were extrapolated to provide overlap in the shallows, but only by 3 pixels to maintain data integrity as was carried out for other environmental data layers. It can also be argued that the use of high-resolution wave data is only really of use in areas that are more affected by wave action, such as an open bay where there is some gradient of wave exposure such as areas 1 and 4. It is possibly not so effective in areas that are heavily sheltered from prevailing waves e.g. within an archipelago, or in a north-east facing enclosed bay where other hydrodynamics may be having more impact (e.g. tides and currents), which could be the case for the shallow areas around Anglesey - areas 2 and 3. The best way of testing this is to trial restoration in the sites and compare results from the different outputs such as outlined by Thom et al. (2018) which is the next step. Nonetheless, the results do show where restoration should be more successful due to the environmental conditions in all the areas and will provide a good baseline for selecting sites for restoration in the future. This study method provides evidence that the use of high-resolution wave data such as the Delft3D used in this study, can improve model outputs and highlights the importance of wave exposure as a factor in determining seagrass presence.

5 Conclusions

Habitat suitability modelling is a valuable tool for the purposes of conservation management and realising the risks of climate change on key species but also for aiding in decision making for the restoration of our marine systems. This study shows the benefits to obtaining high resolution predictor variable data for HSM at small regional scales. For predicting suitable conditions for seagrass, wave height data at a higher resolution was found to be the most important variable and model outputs were improved and differed considerably in comparison to the use of low-resolution wave energy data in its place. However, it should also be recognised that there is a lack of availability of many environmental variable data for the shallow coastal, including intertidal, areas. This has implications for improving HSM for restoration enabling efforts to be focused on areas where the chance of success will be highest.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding author.

Author contributions

CB: writing – original draft, methodology, investigation, data curation, formal analysis and validation. WB: writing - methodology, data curation, formal analysis and validation. DR, HK, JB, and RU: conceptualization, investigation – review and editing. RU, JB, HK, and DR: funding acquisition, supervision, methodology. All authors contributed to the article and approved the submitted version.

Funding

Initial study was funded by WWF and Sky Seagrass Ocean Rescue project with follow up funding from ReSOW NERC funding NE/V01711X/1 to continue and complete the investigation.

Acknowledgments

This study has been conducted using E.U. Copernicus Marine Service Information; https://doi.org/10.48670/moi-00054 and from EMODNet - Licensed under CC-BY 4.0 from the European Marine Observation and Data Network (EMODnet) Seabed Habitats initiative (www.emodnet-seabedhabitats.eu), funded by the European Commission.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmars.2022.1004829/full#supplementary-material

References

Adams, M. P., Saunders, M. I., Maxwell, P. S., Tuazon, D., Roelfsema, C. M., Callaghan, D. P., et al. (2016). Prioritizing localized management actions for seagrass conservation and restoration using a species distribution model. *Aquat. Conserv. Mar. Freshw. Ecosyst.* 26, 639–659. doi: 10.1002/aqc.2573

Allouche, O., Tsoar, A., and Kadmon, R. (2006). Assessing the accuracy of species distribution models: Prevalence, kappa and the true skill statistic (TSS). *J. Appl. Ecol.* 43, 1223–1232. doi: 10.1111/j.1365-2664.2006.01214.x

Araújo, M. B., Anderson, R. P., Barbosa, A. M., Beale, C. M., Dormann, C. F., Early, R., et al. (2019). Standards for distribution models in biodiversity assessments. *Sci. Adv.* 5, 1–12. doi: 10.1126/sciadv.aat4858

Araújo, M. B., and New, M. (2007). Ensemble forecasting of species distributions. *Trends Ecol. Evol.* 22, 42–47. doi: 10.1016/j.tree.2006.09.010

Araujo, M., Pearson, R., Thuiller, W., and Erhard, M. (2005). Validation of species-climate impact models under climate change. *Glob. Change Biol.* 11, 1504–1513. doi: 10.1111/j.1365-2486.2005.001000.x

Beca-Carretero, P., Varela, S., and Stengel, D. B. (2020). A novel method combining species distribution models, remote sensing, and field surveys for detecting and mapping subtidal seagrass meadows. *Aquat. Conserv. Mar. Freshw. Ecosyst.* 30, 1098–1110. doi: 10.1002/aqc.3312

Bertelli, C. M., Stokes, H. J., Bull, J. C., and Unsworth, R. K. F. (2022). The use of habitat suitability modelling for seagrass: A review. *Front. Mar. Sci.* doi: 10.3389/fmars.2022.997831

Bertelli, C. M., and Unsworth, R. K. F. (2018). Light stress responses by the eelgrass, zostera marina (L). *Front. Environ. Sci.* 6. doi: 10.3389/fenvs.2018.00039

Blok, S. E., Olesen, B., and Krause-Jensen, D. (2018). Life history events of eelgrass zostera marina l. populations across gradients of latitude and temperature. *Mar. Ecol. Prog. Ser.* 590, 79–93. doi: 10.3354/meps12479

Booij, N., Ris, R. C., and Holthuijsen, L. H. (1999). A third-generation wave model for coastal regions 1. model description and validation. *J. Geophys. Res. Ocean.* 104, 7649–7666. doi: 10.1029/98JC02622

Danovaro, R., Aronson, J., Cimino, R., Gambi, C., Snelgrove, P. V. R., and Van Dover, C. (2021). Marine ecosystem restoration in a changing ocean. *Restor. Ecol.* 29, 1–8. doi: 10.1111/rec.13432

D'Avack, E. A. S., Tyler-Walters, H., Wilding, C., and Garrard, S. M. (2019). "Zostera (Zostera) marina beds on lower shore or infralittoral clean or muddy sand," in *Marine life information network: Biology and sensitivity key information reviews.* Eds. H. Tyler-Walters and K. Hiscock Available at: https://www.marlin.ac.uk/habitats/detail/257.

Davison, D. M., and Hughes, D. J. (1998). *Zostera biotopes 1. an overview of dynamic and sensitivity characteristics for conservation management of marine SACs.*

Dennison, W. C., and Alberte, R. S. (1985). Role of daily light period in the depth distribution of zostera marina (eelgrass). *Mar. Ecol. Prog. Ser.* 25, 51–61. doi: 10.3354/meps025051

Duarte, C. M. (1991). Seagrass depth limits. *Aquat. Bot.* 40, 363–377. doi: 10.1016/0304-3770(91)90081-F

Elith, J., Ferrier, S., Huettmann, F., and Leathwick, J. (2005). The evaluation strip: A new and robust method for plotting predicted responses from species distribution models. *Ecol. Model.* 186 (3), 280–289.

EUSeaMap (2012). *EUSeaMap technical appendix no. 1 light data and thresholds.*

Fonseca, M. S., and Bell, S. S. (1998). Influence of physical setting on seagrass landscapes. *Mar. Ecol. Prog. Ser.* 171, 109–121. doi: 10.3354/meps171109

Furman, B. T., and Peterson, B. J. (2015). Sexual recruitment in zostera marina: Progress toward a predictive model. *PloS One* 10, 1–20. doi: 10.1371/journal.pone.0138206

Gamble, C., Debney, A., Glover, A., Bertelli, C., Green, B., Hendy, I., et al. (2021). *Seagrass restoration handbook UK & Ireland.*

GEBCO Bathymetric Compilation Group 2020 (2020). "The GEBCO_2020 grid - a continuous terrain model of the global oceans and land," (NERC, UK: British Oceanographic Data Centre, National Oceanography Centre). doi: 10.5285/a29c5465-b138-234d-e053-6c86abc040b9

Green, A. E., Unsworth, R. K. F., Chadwick, M. A., and Jones, P. J. S. (2021). Historical analysis exposes catastrophic seagrass loss for the united kingdom. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.629962

Guisan, A., Thuiller, W., and Zimmermann, N. E. (2017). *Habitat suitability and distribution models: With applications in R (ecology, biodiversity and conservation)* (Cambridge: Cambridge University Press). doi: 10.1017/9781139028271

Harisena, N. V., Groen, T. A., Toxopeus, A. G., and Naimi, B. (2021). When is variable importance estimation in species distribution modelling affected by spatial correlation? *Ecography (Cop.).* 44, 778–788. doi: 10.1111/ecog.05534

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., et al. (2018). *ERA5 hourly data on single levels from 1979 to present* (Copernicus Climate Change Service (C3S) Climate Data Store (CDS). doi: 10.24381/cds.adbb2d47

Hu, W., Zhang, D., Chen, B., Liu, X., Ye, X., Jiang, Q., et al. (2021). Mapping the seagrass conservation and restoration priorities: Coupling habitat suitability and anthropogenic pressures. *Ecol. Indic.* 129, 107960. doi: 10.1016/j.ecolind.2021.107960

Infantes, E., Eriander, L., and Moksnes, P. (2016). Eelgrass (Zostera marina) restoration on the west coast of Sweden using seeds. *Mar. Ecol. Prog. Ser.* 546, 31–45. doi: 10.3354/meps11615

Infantes, E., Orfila, A., Bouma, T. J., Simarro, G., and Terrados, J. (2011). Posidonia oceanica and cymodocea nodosa seedling tolerance to wave exposure. *Limnol. Oceanogr.* 56, 2223–2232. doi: 10.4319/lo.2011.56.6.2223

Jackson, E. L., Griffiths, C. A., and Durkin, O. (2013). A guide to assessing and managing anthropogenic impact on marine angiosperm habitat - part 1: Literature review. *Natural Engl. Commissioned Rep.*

Khan, Z., Ali, S. A., Parvin, F., Mohsin, M., Shamim, S. K., and Ahmad, A. (2022). Predicting the effects of climate change on prospective banj oak (Quercus

leucotrichophora) dispersal in kumaun region of uttarakhand using machine learning algorithms. *Model. Earth Syst. Environ.* doi: 10.1007/s40808-022-01485-5

Koch, E. W., Sanford, L. P., Chen, S.-N., Shafer, D. J., and Smith, J. M. (2006). System-wide water resources research program and submerged aquatic vegetation restoration research program (Waves in seagrass Systems : Review and technical recommendations). *Eng. Res. Dev. Cent.*

Koch, E. W., Verduin, J. J., and Katwijk, V. (2001). Measurements of physical parameters in seagrass habitats. *Glob. Seagrass Res. Methods*, 325–344. doi: 10.1016/B978-044450891-1/50018-9

Krause-Jensen, D., Pedersen, M. F., and Jensen, C. (2003). Regulation of eelgrass (Zostera marina) cover along depth gradients in Danish coastal waters. *Estuaries* 26, 866–877. doi: 10.1007/BF02803345

Kuusemäe, K., Rasmussen, E. K., Canal-Vergés, P., and Flindt, M. R. (2016). Modelling stressors on the eelgrass recovery process in two Danish estuaries. *Ecol. Modell.* 333, 11–42. doi: 10.1016/j.ecolmodel.2016.04.008

Latif, Q. S., Saab, V. A., Dudley, J. G., and Hollenbeck, J. P. (2013). Ensemble modeling to predict habitat suitability for a large-scale disturbance specialist. *Ecol. Evol.* 3, 4348–4364. doi: 10.1002/ece3.790

Lee, K.-S., Park, S. R., and Kim, Y. K. (2007). Effects of irradiance, temperature, and nutrients on growth dynamics of seagrasses: A review. *J. Exp. Mar. Bio. Ecol.* 350, 144–175. doi: 10.1016/j.jembe.2007.06.016

Lesser, G. R., Roelvink, J. A., van Kester, J. A. T. M., and Stelling, G. S. (2004). Development and validation of a three-dimensional morphological model. *Coast. Eng.* 51, 883–915. doi: 10.1016/j.coastaleng.2004.07.014

Marion, S. R., Orth, R. J., Fonseca, M., and Malhotra, A. (2021). Seed burial alleviates wave energy constraints on zostera marina (Eelgrass) seedling establishment at restoration-relevant scales. *Estuaries Coasts* 44 (2), 352-366. doi: 10.1007/s12237-020-00832-y

Marsh, J., Dennison, W. C., and Alberte, R. S. (1986). Effects of temperature on photosynthesis and respiration in eelgrass (Zostera marina l.). *J. Exp. Mar. Bio. Ecol.* 101, 257–267. doi: 10.1016/0022-0981(86)90267-4

McKenzie, L. J., Nordlund, L. M., Jones, B. L., Cullen-Unsworth, L. C., Roelfsema, C., and Unsworth, R. K. F. (2020). The global distribution of seagrass meadows. *Environ. Res. Lett.* 15, 074041. doi: 10.1088/1748-9326/ab7d06

Moore, K. A., and Jarvis, J. C. (2008). Environmental factors affecting recent summertime eelgrass diebacks in the lower Chesapeake bay: implications for long-term persistence. *J. Coastal Res.* 2008 (10055), 135–147. doi: 10.2112/SI55-014

Naimi, B., and Araújo, M. B. (2016). Sdm: A reproducible and extensible r platform for species distribution modelling. *Ecography (Cop.).* 39, 368–375. doi: 10.1111/ecog.01881

Naimi, B., and Araujo, M. B. (2021). doi: 10.1111/ecog.01881. Package ' sdm .'.

Naimi, B., Hamm, N. A. S., Groen, T. A., Skidmore, A. K., and Toxopeus, A. G. (2014). Where is positional uncertainty a problem for species distribution modelling? *Ecography (Cop.).* 37, 191–203. doi: 10.1111/j.1600-0587.2013.00205.x

Natural Resources Wales (2009). *Marine character areas. MCA 7. HOLYHEAD BAY AND THE SKERRIES.*

Nielsen, S., Sand-Jensen, K., Borum, J., and Geertz-Hansen, O. (2002). Depth colonization of eelgrass (Zostera marina) and macroalgae as determined by water transparency in Danish coastal waters. *Estuaries Coasts* 25, 1025–1032. doi: 10.1007/bf02691349

Orth, R. J., and Moore, K. A. (1986). Seasonal and year-to-year variations in the growth of zostera marina l. (eelgrass) in the lower Chesapeake bay. *Aquat. Bot.* 24, 335–341. doi: 10.1017/CBO9781107415324.004

Peralta, G., Brun, F. G., Pérez-Lloréns, J. L., and Bouma, T. J. (2006). Direct effects of current velocity on the growth, morphometry and architecture of seagrasses: A case study on Zostera noltii. *Mar. Ecol. Prog. Ser.* 327, 135–142.

Perger, R., do Amaral, K. B., and Bianchi, F. M. (2021). Distribution modelling of the rare stink bug ceratozygum horridum (Germar 1839): isolated in small spots across the neotropics or a continuous population? *J. Nat. Hist.* 55, 649–663. doi: 10.1080/00222933.2021.1919328

R Core Team (2020). *R: A language and environment for statistical computing.*

QGIS Development Team (2019). "Open source geospatial foundation project," in *QGIS geographic information system*. Available at: http://qgis.osgeo.org.

Rowden, A. A., Anderson, O. F., Georgian, S. E., Bowden, D. A., Clark, M. R., Pallentin, A., et al. (2017). High-resolution habitat suitability models for the conservation and management of vulnerable marine ecosystems on the Louisville seamount chain, south pacific ocean. *Front. Mar. Sci.* 4. doi: 10.3389/fmars.2017.00335

Salo, T., Pedersen, M. F., and Boström, C. (2014). Population specific salinity tolerance in eelgrass (Zostera marina). *J. Exp. Mar. Bio. Ecol.* 461, 425–429. doi: 10.1016/j.jembe.2014.09.010

Sand-Jensen, K. (1975). Biomass, net production and growth dynamics in an eelgrass (Zostera marina l.) population in vellerup vig, Denmark. *Ophelia* 14, 185–201. doi: 10.1080/00785236.1975.10422501

Short, F., Carruthers, T., Dennison, W., and Waycott, M. (2007). Global seagrass distribution and diversity: A bioregional model. *J. Exp. Mar. Bio. Ecol.* 350, 3–20. doi: 10.1016/j.jembe.2007.06.012

Stevens, A. W., and Lacy, J. R. (2012). The influence of wave energy and sediment transport on seagrass distribution. *Estuar. Coasts* 35 (1), 92–108.

Temmink, R. J. M., Christianen, M. J. A., Fivash, G. S., Angelini, C., Boström, C., Didderen, K., et al. (2020). Mimicry of emergent traits amplifies coastal restoration success. *Nat. Commun.* 11, 1–9. doi: 10.1038/s41467-020-17438-4

Thompson, K., Jackson, E., and Kakkonen, J. (2014). Seagrass (Zostera) beds in Orkney. *Scottish Nat. Herit. Commun.*

Thom, R., Gaeckle, J., Buenau, K., Borde, A., Vavrinec, J., Aston, L., et al. (2018). Eelgrass (Zostera marina l.) restoration in Puget Sound: Development of a site suitability assessment process. *Restor. Ecol.* 26 (6), 1066–1074.

Valle, M., Chust, G., del Campo, A., Wisz, M. S., Olsen, S. M., Garmendia, J. M., et al. (2014). Projecting future distribution of the seagrass zostera noltii under global warming and sea level rise. *Biol. Conserv.* 170, 74–85. doi: 10.1016/j.biocon.2013.12.017

Valle, M., van Katwijk, M. M., de Jong, D. J., Bouma, T. J., Schipper, A. M., Chust, G., et al. (2013). Comparing the performance of species distribution models of zostera marina: Implications for conservation. *J. Sea Res.* 83, 56–64. doi: 10.1016/j.seares.2013.03.002

Van Katwijk, M. M. M., Bos, A. R. R., de Jonge, V. N. N., Hanssen, L. S. A. M., Hermus, D.C.R.C.R., and de Jong, D. J. J. (2009). Guidelines for seagrass restoration: Importance of habitat selection and donor population, spreading of risks, and ecosystem engineering effects. *Mar. pollut. Bull.* 58, 179–188. doi: 10.1016/j.marpolbul.2008.09.028

van Katwijk, M. M., and Hermus, D. C. R. (2000). Effects of water dynamics on zostera marina: Transplantation experiments in the intertidal Dutch wadden Sea. *Mar. Ecol. Prog. Ser.* 208, 107–118. doi: 10.3354/meps208107

Waltham, N. J., Elliott, M., Lee, S. Y., Lovelock, C., Duarte, C. M., Buelow, C., et al. (2020). UN Decade on ecosystem restoration 2021–2030—What chance for success in restoring coastal ecosystems? *Front. Mar. Sci.* 7. doi: 10.3389/fmars.2020.00071

Yesson, C., Bush, L. E., Davies, A. J., Maggs, C. A., and Brodie, J. (2015). The distribution and environmental requirements of large brown seaweeds in the British isles. *J. Mar. Biol. Assoc. United Kingdom* 95, 669–680. doi: 10.1017/S0025315414001453

frontiers | Frontiers in Marine Science

TYPE Original Research
PUBLISHED 11 January 2023
DOI 10.3389/fmars.2022.956813

OPEN ACCESS

EDITED BY
Dongxiao Zhang,
University of Washington and NOAA/Pacific Marine Environmental Laboratory, United States

REVIEWED BY
Di Wang,
Tianjin University, China
Milad Janalipour,
Ministry of Science, Research, and Technology, Iran

*CORRESPONDENCE
Yuanyuan Wang
wangyuanyuanxy@zju.edu.cn

SPECIALTY SECTION
This article was submitted to Ocean Observation, a section of the journal Frontiers in Marine Science

RECEIVED 30 May 2022
ACCEPTED 15 December 2022
PUBLISHED 11 January 2023

CITATION
Wang X, Qin M, Zhang Z, Wang Y, Du Z and Wang N (2023) Typhoon cloud image prediction based on enhanced multi-scale deep neural network.
Front. Mar. Sci. 9:956813.
doi: 10.3389/fmars.2022.956813

Typhoon cloud image prediction based on enhanced multi-scale deep neural network

Xin Wang[1,2], Mengjiao Qin[1,2], Zhe Zhang[1,2], Yuanyuan Wang[2,3]*, Zhenhong Du[1,2] and Nan Wang[4,5]

[1]School of Earth Sciences, Zhejiang University, Hangzhou, China, [2]Zhejiang Provincial Key Laboratory of Geographic Information Science, School of Earth Sciences, Zhejiang University, Hangzhou, China, [3]Ocean Academy, Zhejiang University, Zhoushan, China, [4]College of Science and Technology, Ningbo University, Ningbo, China, [5]Ningbo Bay Area Development Research Base, Ningbo, China

Typhoons threaten individuals' lives and property. The accurate prediction of typhoon activity is crucial for reducing those threats and for risk assessment. Satellite images are widely used in typhoon research because of their wide coverage, timeliness, and relatively convenient acquisition. They are also important data sources for typhoon cloud image prediction. Studies on typhoon cloud image prediction have rarely used multi-scale features, which cause significant information loss and lead to fuzzy predictions with insufficient detail. Therefore, we developed an enhanced multi-scale deep neural network (EMSN) to predict a 3-hour-advance typhoon cloud image, which has two parts: a feature enhancement module and a feature encode-decode module. The inputs of the EMSN were eight consecutive images, and a feature enhancement module was applied to extract features from the historical inputs. To consider that the images of different time steps had different contributions to the output result, we used channel attention in this module to enhance important features. Because of the spatially correlated and spatially heterogeneous information at different scales, the feature encode-decode module used ConvLSTMs to capture spatiotemporal features at different scales. In addition, to reduce information loss during downsampling, skip connections were implemented to maintain more low-level information. To verify the effectiveness and applicability of our proposed EMSN, we compared various algorithms and explored the strengths and limitations of the model. The experimental results demonstrated that the EMSN efficiently and accurately predicted typhoon cloud images with higher quality than in the literature.

KEYWORDS

typhoon, Himawari 8, satellite image prediction, multi-scale feature fusion, deep learning

1 Introduction

Typhoons are synoptic-scale storms that originate in the tropical oceans. They absorb large amounts of water vapor from warm tropical oceans that condense at high altitudes and form a low-pressure center (Anduaga, 2019). As the pressure changes and the earth moves, the incoming air swirls, resulting in a tropical cyclone that can eventually evolve into a typhoon if the sea temperature is sufficiently high. Typhoons are one of the most destructive weather events in coastal areas (Defu et al., 2009) and can cause direct disasters, such as gales, rainstorms, and storm surges, as well as secondary disasters, such as floods, landslides, and mudslides, all of which seriously threaten the safety of individuals and property (Zhang and Chen, 2019; Yu et al., 2020; Jia et al., 2020). During 1970–2010, there were 637 landfalling typhoons worldwide, 195 of which reached major intensity (Weinkle et al., 2012). The average number of typhoons in individual countries is 46 per year, and many typhoons affect multiple countries (Shultz et al., 2005). In addition, the intensity of typhoons is likely to increase (Hoque et al., 2017). Therefore, typhoon research is essential to reducing the damage from typhoon disasters. In addition to developing and establishing typhoon engineering works and reinforcing defenses, accurately predicting typhoon activity is the most fundamental measure.

Typhoons are large-scale meteorological systems that require the observation of vast geographical areas. They also demand real-time responsiveness to minimize damage. Based on these considerations, satellite data are highly suitable for typhoon research owing to their wide coverage, timeliness, and relatively convenient acquisition. Satellite-based techniques have been an effective means of typhoon research because the movement and density of typhoons can be determined based on cloud patterns. Specifically, the outer bands of cumulonimbus clouds indicate the direction of movement, and the cloud pattern surrounding the typhoon eye reflects the location and intensity of the typhoon (Kovordányi and Chandan, 2009). Meteorologists use satellite images to analyze and predict various aspects of typhoons. Zhang et al. (2005) located typhoons using a rotational location method by finding feature points. Jaiswal et al. (Neeru and Kishtawal, 2010) used the helix fitting method to determine typhoon centers based on the special structure of mature typhoons. In addition to traditional morphological methods, deep learning technologies have been used for typhoon detection and location. Wang et al. (Wang et al., 2021) detected typhoon centers using a fast R-CNN. Because typhoons can be located using satellite images, typhoon tracks can also be calculated using imaging sequences. Hong et al. (2017) predicted the typhoon position from a single satellite image and then calculated the typhoon track based on subsequent consecutive images. Thus, future typhoon behavior can be predicted using continuous satellite images.

With the development of artificial intelligence, image prediction technology has significantly improved. For example, convolutional neural networks are a widely used deep learning method for extracting spatial features and are extensively used in image research (Khan et al., 2020). Sequence-learning problems are often solved using recurrent neural networks (RNNs) (Salehinejad et al., 2017). Based on RNNs, Ranzato et al. (2014) presented a recurrent convolutional neural network architecture that connects convolutional layers at the beginning and end of an RNN. Shi et al. (Xingjian et al., 2015) proposed a convolutional LSTM (ConvLSTM) architecture and combined convolutional and recurrent layers to extract dynamic temporal and spatial information. They later proposed a new model, trajectory GRU (TrajGRU), which dynamically learns the location-variant structure of recurrent connections (Shi et al., 2017). Based on ConvLSTM, Wang et al. (2017) proposed a new structure called spatiotemporal LSTM (ST-LSTM), in which spatial and temporal information can be stored simultaneously. Autoencoders (Tschannen et al., 2018) and generative adversarial networks (GANs) (Zhaoqing et al., 2019) are also widely used for image and video prediction. However, because of the absence of high-quality processed satellite datasets and the complex evolution of satellite images (Moskolaï et al., 2021), predicting satellite data remains difficult, and simply applying the image prediction models to this problem cannot yield satisfactory results.

For example, Xu et al. (2019) combined WGAN and LSTM to predict satellite images, but their evaluation method only evaluates the shape of clouds, not their density. Hong et al. suggested an autoencoder model that uses convolutional and ConvLSTM layers to extract features and generate future satellite images. Their study also explored the application of LSTM and ConvLSTM in the model and found that the former was unsuitable for feature extraction from images, which also called into question the efficacy of the model of (Xu et al., 2019) in overcoming the satellite image prediction problem. However, the spatial resolution of the satellite data used in their research was too low to observe clouds. Lee et al. (2019) referred to (Hong et al., 2017b) and proposed an advanced model with better image prediction than those in (Xu et al., 2019) and (Hong et al., 2017b), but it was insufficient. Typhoon research involves large-scale problems that require downsampling operations to mitigate calculation costs. However, repeated downsampling of data results in the loss of detail and the production of poor-quality predicted images. Notably, satellite image prediction models rarely consider this factor.

To solve this problem, we constructed an enhanced multi-scale deep neural network (EMSN) with a feature enhancement method and a multi-scale feature fusion strategy to discover the relationships between historical and future images and predict clearer typhoon cloud images than in the literature. Through the channel attention mechanism, the network extract features from

the input images and focuses on important features. Next, the encode-decode module captures the spatiotemporal relationships of the enhanced features at different scales. Additionally, to further improve the quality of the predicted images, we used skip connections between the encoder and decoder to account for information loss caused by downsampling.

A summary of the contributions of this study is as follows:

(1) Input images from different time steps have different contributions to the model. Hence, we designed a feature enhancement module that uses a channel attention module to extract useful characteristics to enhance important features and suppress useless features.

(2) To consider the variability of typhoon motion and the complexity of feature maps, we designed a symmetrical encode-decode module to capture spatiotemporal rules. A multi-scale feature fusion strategy was implemented to reduce information loss during downsampling and fully use the features of different scales.

(3) Based on the steps, we propose a novel EMSN to predict typhoon cloud images. Extensive experiments based on the Himawari 8 dataset demonstrate the effectiveness of the proposed model.

The remainder of this paper is organized as follows: Section 2 describes the problem and introduces the dataset and the proposed network for typhoon cloud image prediction. Section 3 describes the experiments conducted to demonstrate the effectiveness of the proposed model and discusses the results, and Section 4 presents the conclusions.

2 Methodology

In this section, we describe the data and network structure used in this study. Specifically, we model the typhoon cloud image prediction problem, introduce the study data, area, and our proposed network, and describe the experimental environment.

2.1 Statement of the problem

Typhoon cloud image prediction can be regarded as a video prediction problem for forecasting subsequent frames based on previously observed frames. Our observation area was an $M \times N$ grid, and we used the vector $I \in R^{M \times N}$ to represent the satellite image. We recorded j observed satellite images as $X = \{I_{t-j+1}, \ldots, I_{t-1}\}$ and used $Y = \{I_{t+1}, \ldots, I_{t+k}\}$ to represent the next k frames generated by the prediction model; $\hat{Y} = \{\hat{I}_{t+1}, \ldots, \hat{I}_{t+k}\}$ represents the corresponding ground truth frames.

Therefore, our prediction problem was modeled as a pixel-to-pixel regression problem from X to Y, which aimed to minimize the error between Y and $\hat{Y}$ as follows:

$$\hat{I}_{t+1}, \ldots, \hat{I}_{t+k} = \underset{I_{t+1}, \ldots, I_{t+k}}{argmax}\, p\left(I_{t+1}, \ldots, I_{t+k} | I_{t-j+1}, \ldots, I_{t-1}\right) (1)$$

2.2 Dataset and study area

Our research was based on satellite images from Japan's eighth geostationary meteorological satellite, Himawari 8, which generates images with spatial resolutions of 2 km (6,001 × 6,001) and 5km (2,401 × 2,401) every 10 min. The Advanced Himawari Imager on Himawari 8 has 16 bands: three visible bands, three near-infrared bands, and 10 infrared bands (Bessho et al., 2016).

In general, according to the wavelength range, satellite cloud images can be divided into visible light, water vapor, and infrared images (Figure 1). Visible images have the clearest cloud texture, but visible bands can only be measured during the daytime. Water vapor images only reflect water vapor information in the upper-middle troposphere. Infrared bands most effectively detect clouds of various layers at any time of day

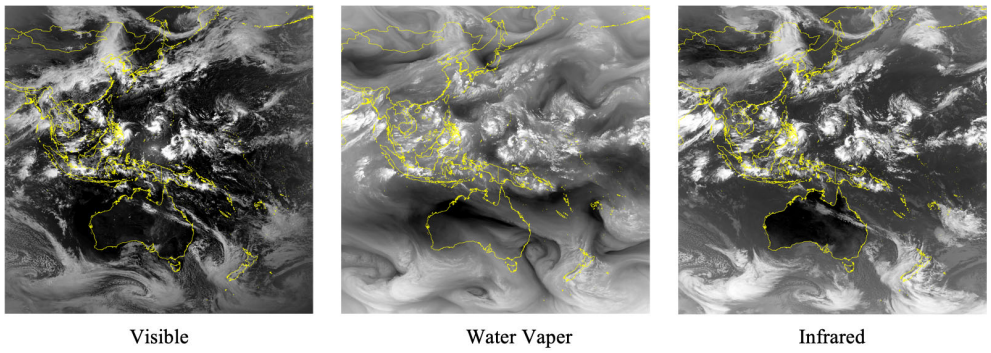


FIGURE 1
Satellite images of H8: left, a visible image; middle, a water vapor image; and right, an infrared image. The country borders in yellow show the spatial scale of the satellite data.

and are suitable for our research. Our work focuses more on building a model for typhoon cloud image prediction than exploring the differences between infrared bands. To conduct calculations efficiently, we used band 13 as the source of experimental data.

Our study aimed to predict future satellite images and observe typhoon activity. Hence, we used only satellite images of typhoon clouds. Data were collected between July 2015 and October 2021, and they comprised 145 typhoons (Japan Meteorological Agency (JMA), 2018). At each location, observations were made at 6-hour intervals. However, because typhoons occurred between these observations, we filtered typhoon-related satellite images at 1-hour intervals. The geographical range of an H8 image is between 60°S–60°N latitude and 80°E–160°W longitude and includes many areas where typhoons have never occurred. Thus, we used part of the Northwest Pacific to collect observations (15°S–50°N, 100°E–165°E) because strong typhoons frequently occur in this area (Figure 2).

2.3 Data processing

Satellite images inevitably have some missing data; thus, we filtered the abnormal data. After sorting by time, the sliding window method was used to create the dataset. A time interval of 3 h (Xu et al., 2019) was used to build sequences. We used this interval because it contained more variation between two adjacent satellite images than the 1-hour interval did (Hong et al., 2017b) and used less training time for the model than the 6-hour interval did (Rüttgers et al., 2019). To determine the hyperparameter of the historical frames (denoted by m), we trained and tested cases with $m = 4$ (Rüttgers et al., 2019), 6 (Hong et al., 2017b), 8 (Xu et al., 2019), and 10. For each case, we used an ($m + 1$)-frame wide sliding window and an ($m/2$) -frame sliding step to create the sequences. Ultimately, $m = 8$ produced the best results. In the cases of $m = 4$ and 6, the movement of typhoons was not sufficiently captured; for $m = 10$, model performance worsened owing to the reduced size of the training dataset. Therefore, we used a 9-frame wide sliding window and a 4-frame sliding step to create the sequences. The resulting dataset contained 5,160 sequences, each comprising eight input frames and one ground truth. These sequences were then divided into training, validation, and test datasets at a ratio of 6:2:2 (i.e., 3,096 for training, 1,032 for validation, and 1,032 for testing).

The original pixel size of H8 satellite images would have caused memory limitations; thus, we resized the images to 481 × 481 pixels and selected patches of 260 × 260 pixels that corresponded to the research area (15°S–50°N, 100°E–165°E). To facilitate multiple downsamplings in the experiments, we cropped the data to 256 × 256 pixels from the upper left corner before normalizing the training, validation, and test datasets by setting $I = \frac{I-\min(I)}{\max(I)-\min(I)}$ to accelerate model convergence.

2.4 Network structure

2.4.1 Overview

We used several lightweight modules to build the model, considering its effectiveness in practical applications. As shown in Figure 3, the proposed EMSN used eight observed satellite images (i.e., $t_1,t_2,\ldots,t_8$) as inputs and predicted the subsequent image (i.e., t_9) as output. EMSN consists of mainly of feature enhancements and encode–decode modules.

The feature enhancement module is a preprocessing module that receives inputs and obtains a set of enhanced feature maps. Subsequently, the encode–decode module captures the spatiotemporal relationships of the feature maps and encodes

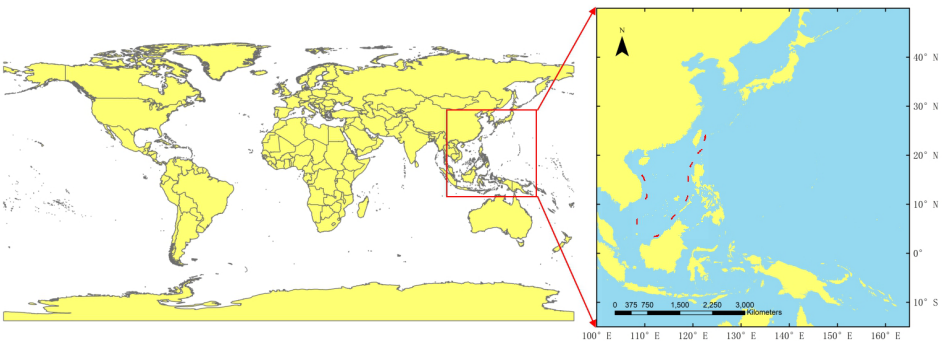


FIGURE 2
Research area map. Part of the Northwest Pacific (15°S–50°N, 100°E–165°E).

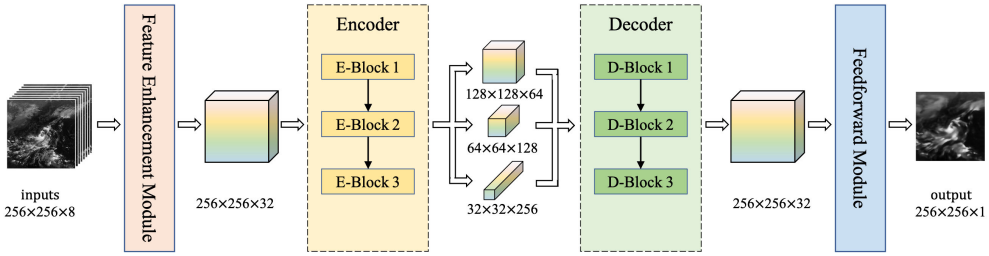


FIGURE 3
Overview of EMSN. EMSN consists of a feature enhancement module, encode–decode module, and feedforward module.

them into hidden states, which are decoded into the output image. The detailed structure of the proposed network is listed in Table 1.

2.4.2 Feature enhancement module

The feature enhancement module is the preprocessing module of the EMSN, which comprises two convolutional layers and a channel attention module. Figure 4 shows the module structure. First, the convolutional layers extract scale-invariant features from the input images. These feature maps have multiple channels, some of which are not useful. A channel attention module was implemented after the convolutional layers to enhance important features and suppress useless features. As shown in Figure 4, the channel attention we used was a squeeze-and-excitation network (Hu et al., 2018) that automatically determined the relevance of each channel and then assigned appropriate weights.

2.4.3 Encode–decode module

The encode–decode module is the core of the EMSN and has two parts: the basic model and the multi-scale feature fusion module. The basic model is a symmetrical autoencoder. A multi-scale feature fusion module was applied between the encoder and decoder to pass the information on (Figure 5).

Both the encoder and decoder of the basic model consisted of three blocks, each of which had three layers: the convolutional layer

TABLE 1 Architecture of EMSN.

Module	Block	Layer	Kernel size	Stride	Output size
Feature Enhancement Module	–	Conv 1	3×3	(1, 1)	256×256×16
	–	Conv 2	3×3	(1, 1)	256×256×32
	–	Channel Attention	3×3	(1, 1)	256×256×32
Encoder	E-Block 1	Conv	3×3	(2, 2)	128×128×64
		ConvLSTM	3×3	(1, 1)	128×128×64
	E-Block 2	Conv	3×3	(2, 2)	64×64×128
		ConvLSTM	3×3	(1, 1)	64×64×128
	E-Block 3	Conv	3×3	(2, 2)	32×32×256
		ConvLSTM	3×3	(1, 1)	32×32×256
Decoder	D-Block 3	ConvLSTM	3×3	(1, 1)	32×32×256
		TransposeConv	4×4	(2, 2)	64×64×128
	D-Block 2	ConvLSTM	3×3	(1, 1)	64×64×128
		TransposeConv	4×4	(2, 2)	128×128×64
	D-Block 1	ConvLSTM	3×3	(1, 1)	128×128×64
		TransposeConv	4×4	(2, 2)	256×256×32
Feedforward Module	–	Conv 1	3×3	(1, 1)	256×256×16
	–	Conv 2	3×3	(1, 1)	256×256×1

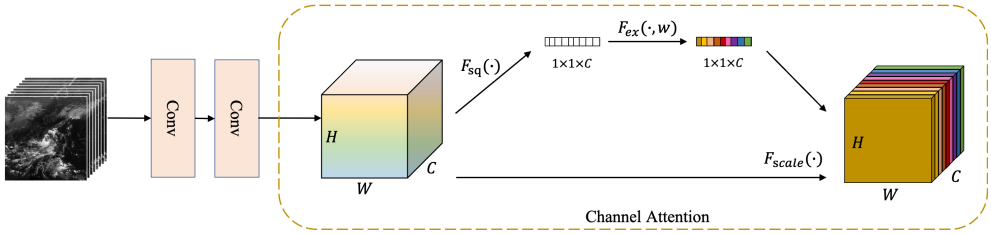


FIGURE 4
Feature enhancement module: two simple convolutional layers and a channel attention module.

(or transposed convolutional layer) for changing feature map sizes, the activation function layer for adding nonlinear factors, and the ConvLSTM layer to capture spatiotemporal relationships.

Owing to the large observation area and complex evolutionary rules of the typhoon cloud image sequences mentioned in Section 1, we used the multi-scale feature fusion module to acquire additional details. The inner structure of the module is shown in Figure 6.

As shown in Figure 6, the multi-scale feature fusion module has two parts: multi-scale spatiotemporal feature capturing and same-stage feature reuse. The encode–decode module uses $\{X_1, X_2,\ldots,X_8\}$ as inputs, where X_i is the enhanced feature of the ith (i=1, 2,…, 8) observed image. Each time the data passed through the convolutional layer; the feature sizes were reduced by half. Thus, the resized features of the three different scales were obtained through the three blocks. These feature maps of different scales contain information at different granularities. As shown in Figure 6, these resized feature maps were passed through the ConvLSTM layers to learn spatiotemporal rules. In Figure 6, h_i^n is the hidden state of the ith (i=1,2,…,7) input frame at the nth (n=1,2,3) ConvLSTM layer, which was used in the encoder to pass spatiotemporal information from the previous frame to the subsequent frame.

Downsampling of the encoder by convolutional layers results in the loss of information on typhoon clouds, and the feature maps obtained by multiple blocks are highly abstract. Therefore, we implemented the same-stage feature reuse strategy to use other low-level features in the decoder. In Figure 6, H_n is the hidden state of the last memory unit in the nth ConvLSTM layer (i.e., h_8^n), which is used as the input of the corresponding ConvLSTM unit in the decoder. Furthermore, F_n is the feature map of the last input frame at the nth convolution layer, and O_n is the output of the ConvLSTM layer in the nth block of the decoder . These values are then summed to acquire the new input of the convolution layers in the nth block of the decoder. Through the encode–decode module, both detailed information and global information were conserved, and the model generated a clear prediction frame.

2.5 Implementation details

We implemented all the models based on Python 3.8 with an NVIDIA GeForce RTX 3090 GPU card. We chose the Adam optimizer for gradient optimization, with a batch size of 16 and L1 loss. The initial learning rate was set to 1e−3, and the learning

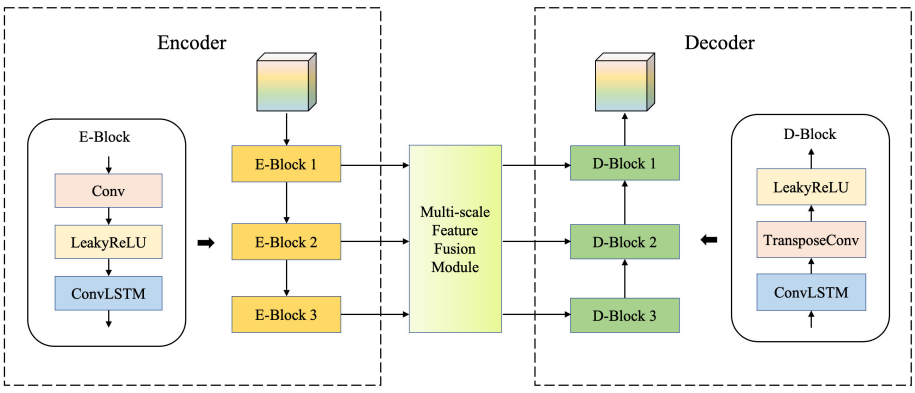


FIGURE 5
Encode–decode module. The symmetric basic model comprises convolutional layers and ConvLSTM layers. A multi-scale feature fusion module is added between the encoder and the decoder.

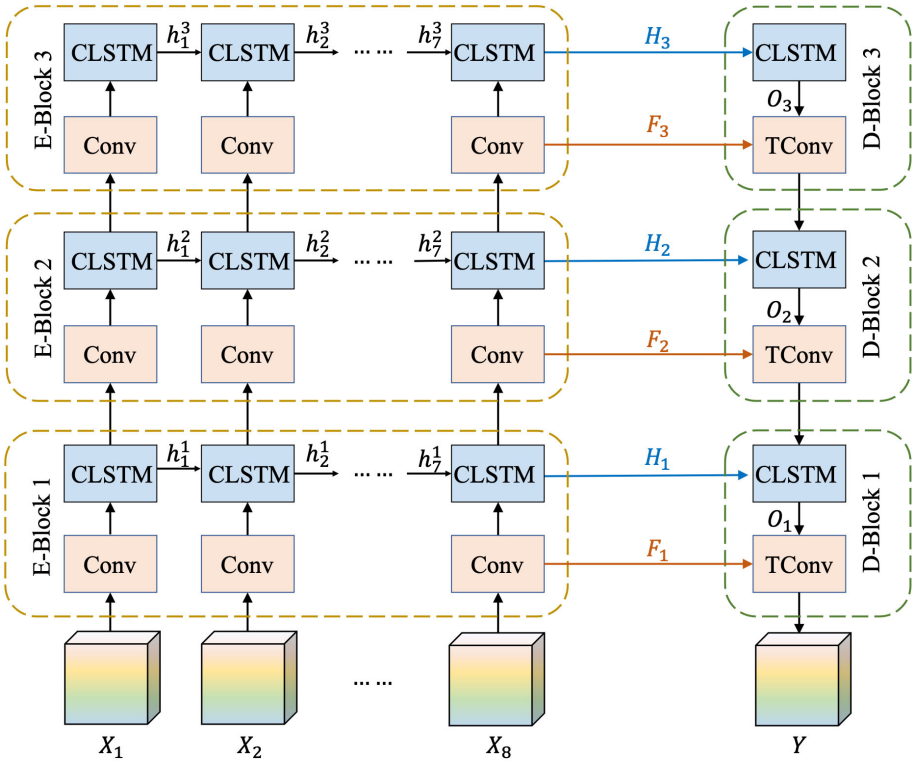


FIGURE 6
Inner structure of feature fusion module, including multi-scale spatiotemporal feature capturing and same-stage feature reusing.

rate was reduced by half when the valid loss did not decrease over 10 epochs. The total training epoch was set to 500, with an early stopping strategy of 75 epochs.

3 Experiments and results

3.1 Evaluation metrics

The mean squared error (MSE), peak signal-to-noise ratio (PSNR), and structural similarity index (SSIM) were used to evaluate the performance of the models. We use and to denote prediction and ground truth, respectively. The criteria are expressed as follows:

$$MSE(p,g) = \frac{1}{WH}\sum_{i=1}^{W}\sum_{j=1}^{H}\left(p_{i,j} - g_{i,j}\right)^2 \quad (2)$$

$$PSNR(p,g) = 10 \times \log_{10}\frac{MAX^2}{MSE(p,g)} \quad (3)$$

$$SSIM(p,g) = \frac{\left(2\mu_p\mu_g + c_1\right)\times\left(2\sigma_{pg} + c_2\right)}{\left(\mu_p^2 + \mu_g^2 + c_1\right)\times\left(\sigma_p^2 + \sigma_g^2 + c_2\right)} \quad (4)$$

where W and H are the width and height of the image, respectively; MAX is the maximum pixel value in the image; μ_p and μ_g are the average pixel values of p and g respectively; σ_p and σ_g are the variances of p and g , respectively; and σ_{pg} is the covariance of p and g . In addition, $c_1=(k_1L)^2$ and $c_2=(k_2L)^2$ both of which are constants used to maintain stability; L is the dynamic range of pixel values; and $k_1=0.01$ and $k_2=0.03$.

MSE and PSNR are the most widely used image quality assessment algorithms and reflect the difference in pixel values. SSIM evaluates the structural similarity between two images. Small values of MSE or large values of PSNR and SSIM indicate that the predicted images are similar to the ground-truth images.

3.2 Experiments on the proposed network

In this section, we introduce the process of finding the best model structure and hyperparameters for our proposed EMSN. All models were trained using the same hyperparameters.

3.2.1 Exploration of the basic model

Because of the state-of-the-art performance achieved by using convolutional and ConvLSTM-based models (Lee et al., 2019), we analyzed various model structures and network hyperparameter settings of the convolutional and ConvLSTM

networks. Table 2 lists the experimental results of the basic model under various settings.

First, we reproduced the model proposed in (Lee et al., 2019), and experiment 1 used the structure of MCSIP's satellite image predictor. Next, we gradually modified the model settings in MCSIP. As mentioned in Section 2.4.3, MCSIP is a feedforward neural network with a structure that results in inevitable information loss. In experiment 2, we applied the U-Net (Siddique et al., 2021) structure to the model. A symmetric three-layer ConvLSTM structure was added before upsampling, and the three hidden states of the encoder ConvLSTM layers were used as inputs to the decoder ConvLSTM layers. The results demonstrated that the hidden states of each ConvLSTM layer benefited the network, and their use in the decoder improved model performance. The MCSIP used nearest neighbor sampling to enlarge the image sizes but, experiment 3 used deconvolution layers. The results showed that deconvolution was superior to nearest neighbor interpolation in our model, possibly because the deconvolution layers dynamically learn the resizing parameters. We then compared the impacts of stacking different numbers of ConvLSTM layers on model performance. The results of experiments 3–6 revealed that the three-layer ConvLSTM module most effectively captured the spatiotemporal relationships among satellite images. As the number of ConvLSTM layers increased, model performance first improved before worsening after this optimum. This initial improvement in performance may have been due to increasing ConvLSTM layers, deepening the network, and—thus—strengthening its ability for nonlinear fitting. The later deterioration of performance may have been due to limited quantities of data and overfitting resulting from too many layers.

Based on our results, we applied the symmetric autoencoder as the basic structure of our model, in which the encoder uses five convolutional layers to downsample and three ConvLSTM layers to capture spatiotemporal relationships. The decoder uses three additional ConvLSTM layers and five deconvolution layers for upsampling.

3.2.2 Ablation experiments on multi-scale components

Typhoon cloud image prediction is a large-scale problem. Features of different scales that contain different information are obtained through convolution layers. Using features computed at multiple scales facilitates detailed predictions. Therefore, to fuse these features, we attempted to add several components to the basic model, the experimental results of which are listed in Table 3.

The models in Section 3.2.1 followed the structure of MCSIP, which downsampled images with several convolution layers before capturing spatiotemporal relationships with ConvLSTM layers. In this structure, models only learn the spatiotemporal rules of the most abstract features because the ConvLSTM layers assess feature maps at the smallest scale. To enable our model to use detailed geometric information, we attempted to capture spatiotemporal relationships at large scales using convolutional layers and ConvLSTM layers in a staggered arrangement. The results of experiments 1 and 2 illustrate that interleaving convolutional layers and ConvLSTM layers helped the model learn more details and improve performance.

In addition, to explore the most suitable feature extraction scale for the typhoon cloud image prediction problem, we conducted experiments on four downsampling scales based on a cross-arranged network. In each iteration, we maintained the total number of convolution layers (i.e., five) but changed the number of convolution layers that halved the feature sizes. According to the results of experiments 2–5, as the scale of the downsampling decreased, the model performance first improved and then worsened after an optimum of 3 as the best downsampling scale.

Furthermore, images inevitably lose information during downsampling, and reusing these large-scale features during the upsampling process can enrich the details of the final output. Therefore, we implemented a skip connection module at the same stage. After considering computational complexity, we chose the add operation (He et al., 2016) rather than the concatenation operation (Huang et al., 2017). Because the input

TABLE 2 Experimental results of basic model with various settings.

No.	ConvLSTM	Symmetric	Deconv	MSE ↓	PSNR ↑	SSIM ↑
1	3	×	×	0.014252	18.515	0.77405
2	3	√	×	0.013397	18.805	0.79977
3	3	√	√	**0.012571**	**19.099**	**0.81418**
4	2	√	√	0.013060	18.937	0.80917
5	4	√	√	0.013051	18.909	0.80050
6	5	√	√	0.013161	18.869	0.80078

No. denotes the experiment number. ConvLSTM denotes the amount of ConvLSTM layers used in the encoder. Symmetric represents the use of a symmetric ConvLSTM structure in the decoder. Deconv signifies the use of a deconvolution layer for upsampling. The best result is marked in bold.

TABLE 3 Experimental results of adding multi-scale components.

No.	Cross	Scale	Skip			MSE ↓	PSNR ↑	SSIM ↑
			Each	Weighted	Last			
1	×	5	×	×	×	0.012571	19.099	0.81418
2	√	5	×	×	×	0.012343	19.355	0.81741
3	√	4	×	×	×	0.012275	19.525	0.81766
4	√	3	×	×	×	0.011575	19.647	0.82954
5	√	2	×	×	×	0.012152	19.432	0.82785
6	√	3	√	×	×	0.011886	19.481	0.82651
7	√	3	×	√	×	0.011565	19.641	0.83096
8	√	3	×	×	√	**0.011394**	**19.672**	**0.83339**

The best result is marked in bold.

and output of the encode–decode module had different dimensions (i.e., the input shape was 8 × 256 × 256 and the output shape was 256 × 256), we attempted three additive methods: only adding the last frame of inputs, separately adding each frame of inputs with the same weights, and separately adding each frame of inputs with learnable weights. Based on the results of experiments 6–8, skipping connections was an effective strategy, and only adding the last frame of inputs was the best additive method.

To illustrate the effectiveness of the feature fusion module more intuitively than in the prior experiment, we considered Typhoon Yutu (October 2018) and visualized the predicted images of the experiments in Table 3 (Figure 7). Our proposed model predicted future typhoon satellite images in the Northwest Pacific and performed well in terms of cloud shape and texture. Figure 7A shows the predicted image of our basic model without the feature fusion module (i.e., experiment 1). Although the basic model effectively captured the overall distribution of clouds, details were missing. Figure 7B shows the predicted image of the model with multi-scale spatiotemporal feature capturing (i.e., experiment 4) and is substantially clearer than that in Figure 7A. Figure 7C shows the predicted image of the model with both multi-scale spatiotemporal feature capturing and same-stage feature reuse (i.e., experiment 8), which performs better than the model in Figure 7B in cloud shape prediction. From Figures 7A-C, the image quality gradually improved, particularly within the red box. Specifically, as shown in Figure 7A, the spiral cloud bands around the center of the typhoon were fragmented rather than coherent and complete. In Figures 7B, C, the spiral characteristics of the typhoon clouds are more obvious than those in Figure 7A. We compared Figures 7B, C, and the spiral cloud bands of the latter are smoother than those of the former and better show the typhoon movement trend than the former does.

Therefore, the multi-scale feature capturing strategy and the same-stage feature reuse method excelled in typhoon cloud image prediction.

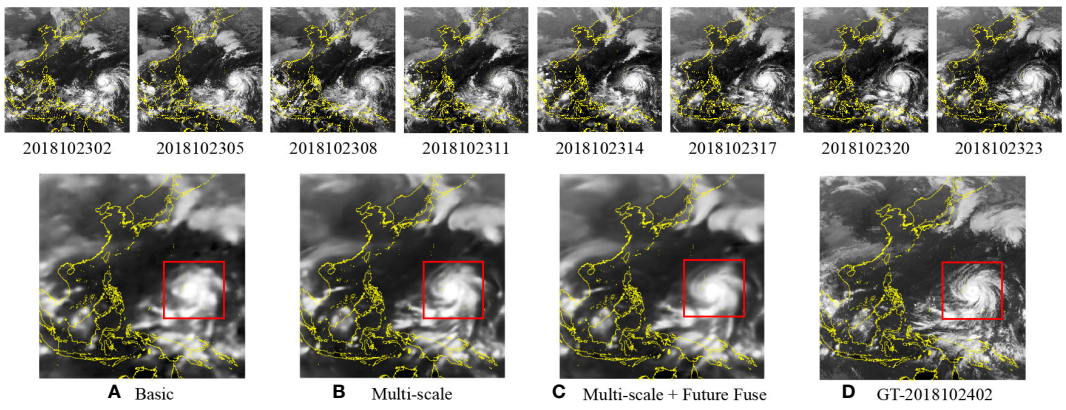


FIGURE 7
Visual results of adding multi-scale components. First-row images are the inputs. Second-row images are outputs of different networks. **(A)** is the basic model; **(B)** is the model with multi-scale spatiotemporal feature capturing; **(C)** is the model with both multi-scale spatiotemporal feature capturing and same-stage feature reuse; and **(D)** is the ground truth.

3.2.3 Ablation experiments on channel attention

The input to our model was a sequence of consecutive satellite images, and the closer an image is to the present time, the greater its impact on the predicted image. After feature extraction by convolution layers, we obtained multiple feature maps with different representations; however, only some of these feature maps benefited our model. Therefore, the dependency of the output on each channel of the input differed; to refine the intermediate features and assign relative importance to each channel, we implemented a channel attention module before the multi-scale feature capturing module. In this section, we review several experiments on the impact of channel attention (Table 4).

SENet is a classic channel attention module with a squeeze-and-excitation structure. The original SENet (Hu et al., 2018) adopted global average pooling to encode the entire spatial feature of a channel as a global feature. It was proposed by (Woo et al., 2018) that average pooling and maximum pooling have different representation effects; thus, we used maximum pooling alone and combined with average pooling in experiments 3 and 4, respectively. Based on these results, the best values of PSNR and SSIM belonged to models with channel attention, demonstrating that it increased model performance. Because the changes in the three indicators were inconsistent, we further evaluated them through visualization and used Typhoon Krosa (August 2019) as an example to increase the robustness of our conclusions (Figure 8).

Figure 8 shows the visualized results of models with different attentions. The cloud texture can be easily distinguished in Figures 8B–D (i.e., models with channel attention), which are clearer than those in Figure 8A (i.e., models without channel attention). Moreover, Figure 8C (i.e., the model with channel attention using maximum pooling) exhibits the highest performance. The red box highlights the differences between the predicted images. In Figure 8C, the details of the typhoon clouds are most evident. The typhoon cloud marked by the red box in Figure 8E (i.e., the ground truth) has a clear eye and ring-shaped eyewall, which is also predicted only in Figure 8C.

Therefore, combining the experimental and visualization results illustrates that channel attention improves model performance and that using maximum pooling to squeeze can result in the best performance.

3.2.4 Comparison of spatiotemporal capture units

As mentioned in Section 1, convolutional GRU (Wang et al., 2018; Tian et al., 2019) (ConvGRU), trajectory GRU (Shi, 2018; Gan and Hui, 2020) (TrajGRU), spatiotemporal LSTM (Feng, 2019; Wang et al., 2021) (ST-LSTM), and convolutional LSTM (Mukherjee et al., 2019; Lin et al., 2020) (ConvLSTM) are widely used in video prediction. To prove that our proposed structure applies to all spatiotemporal capture units and to determine the best one for typhoon cloud image prediction, we compared the variations in the RNN with the EMSN. The experimental results are listed in Table 5.

The model using ConvLSTM had the best performance, with one best result and two second-best results according to the evaluation metrics. The model using ST-LSTM had the worst performance, with the worst results in all three metrics. The predictions were visualized (Figure 9). Each image is similar to that in Figure 9E (i.e., the ground truth), which verifies the effectiveness of the EMSN. However, in the comparison of Figures 9A–D, the latter (i.e., ConvLSTM) shows greater detail and clearer spiral cloud bands than the former. Mature Typhoon Saola (October 2017) is marked with a red box. According to Figure 9E, there is an obvious typhoon eye in the center of the cloud, and the comma-shaped eye area points approximately in the north-south direction. Except for Figure 9D, the images did not predict eye structure well.

Furthermore, we generated different images between the ground-truth satellite images and predicted satellite images (Figure 10). In these images, the brighter the color, the larger the pixel difference, and vice versa. The white box in Figure 10 and the red box in Figure 9 mark the same area (i.e., the typhoon eye area). Figure 10D (i.e., ConvLSTM) is much darker than the other images; thus, ConvLSTM outperforms other RNN variations in the EMSN.

Based on the experimental results, visualization results, and difference images, the proposed EMSN accurately predicts typhoon cloud images, and ConvLSTM is the best spatiotemporal capture unit for the EMSN.

TABLE 4 Experimental results of adding the channel attention module with different pooling methods.

No.	AvgPool	MaxPool	Avg+Max	MSE ↓	PSNR ↑	SSIM ↑
1	×	×	×	**0.011394**	19.672	0.83339
2	√	×	×	0.011834	19.542	0.81879
3	×	√	×	<u>0.011435</u>	**19.733**	0.82982
4	×	×	√	0.011524	<u>19.680</u>	**0.83446**

The best result is marked in bold; the second-best result is underlined.

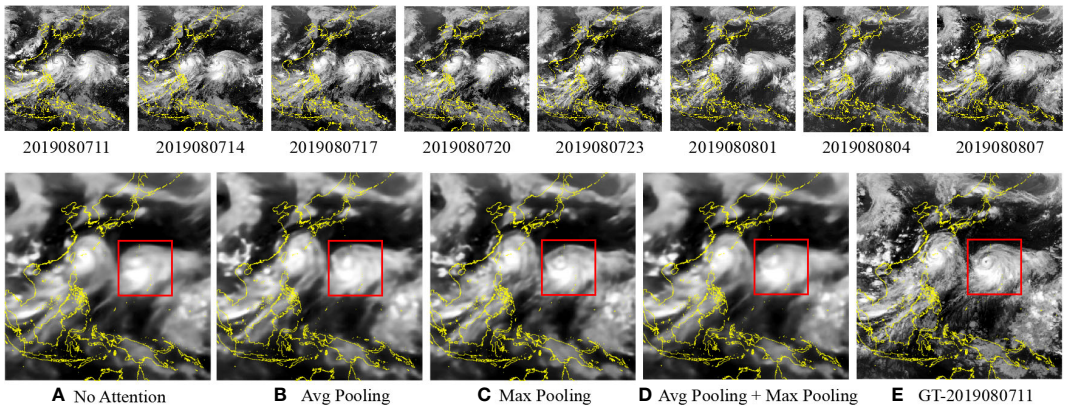


FIGURE 8
Visual results of adding the channel attention module with different pooling methods; the second row displays the outputs. **(A)** without channel attention; **(B)** using average pooling; **(C)** using maximum pooling; **(D)** combining average pooling and maximum pooling; and **(E)** ground truth.

TABLE 5 Results of EMSN with different RNN variations. The best result is marked in bold; the second-best result is underlined.

Model	MSE ↓	PSNR ↑	SSIM ↑
EMSN-ConvGRU	0.011450	19.657	**0.83343**
EMSN-TrajGRU	**0.011366**	<u>19.711</u>	0.82588
EMSN-STLSTM	0.012500	19.198	0.81007
EMSN-ConvLSTM (ours)	<u>0.011435</u>	**19.733**	<u>0.82982</u>

4 Conclusion

In this study, an EMSN was developed for typhoon cloud image prediction. To improve the extraction of useful information from multiple input images, we designed a feature enhancement module to focus on meaningful features and remove worthless features. Additionally, to utilize both high-level semantic features and low-level detail features to generate clearer predicted images than in the literature, we designed a symmetric structure to fuse features at different scales, which used a multi-scale feature capturing strategy and a same stage feature reusing method. Based on a typhoon cloud image dataset collected from Himawari 8, extensive experiments were conducted to demonstrate the effectiveness of EMSN. The

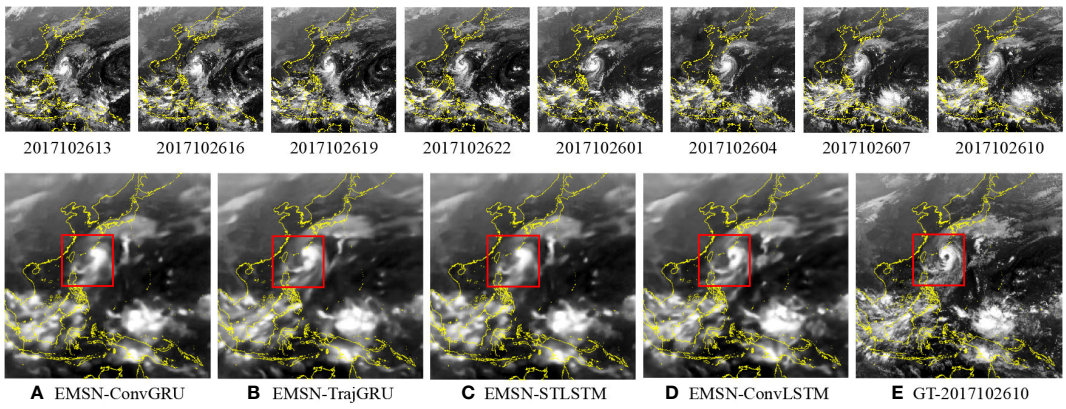


FIGURE 9
Visual results of the EMSN with different RNN structures; second-row images are the outputs. **(A)** EMSN-ConvGRU; **(B)** EMSN-TrajGRU; **(C)** EMSN-STLSTM; **(D)** EMSN-ConvLSTM; and **(E)** ground truth.

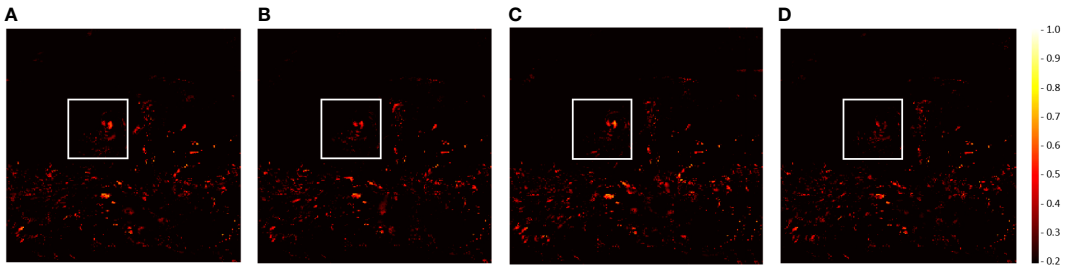


FIGURE 10
Difference images between the ground truth and the predicted results of the EMSN; the ground truth and predicted results are shown in Figure 9. **(A)** EMSN-ConvGRU; **(B)** EMSN-TrajGRU; **(C)** EMSN-STLSTM; and **(D)** EMSN-ConvLSTM.

experimental results showed that EMSN efficiently and accurately predicted typhoon cloud images with higher quality than in the literature.

Although this study succeeded in predicting clear, accurate typhoon cloud images, we did not consider meteorological factors. This limited the model to only learning typhoon movement rules from satellite images. In further research, we intend to add meteorological data such as SST and MSL as model inputs to expand the learnable information and enable the model to learn how the surrounding environment affects typhoons. We also plan to further explore the prediction of typhoon position and intensity based on the predicted cloud images.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: https://www.eorc.jaxa.jp/ptree/.

Author contributions

XW: Conceptualization, investigation, methodology, and writing—original draft. MQ: Methodology, validation, and writing—review and editing. ZZ: Software and data curation. YW: Resources, funding acquisition, and writing—review and editing. ZD: Resources, funding acquisition, and writing—review and editing. NW: Resources and funding acquisition. All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Funding

This work was supported in part by the National Natural Science Foundation of China under Grants 41922043, 41871287, and 42050103, and in part by the Key project of Soft Science Research of Zhejiang Province in the year 2022 under Grant 2022C25021.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Anduaga, A. (2019). "History of typhoon science," in *Oxford Research encyclopedia of climate science*. (Oxford University Press). Available at: https://oxfordre.com/climatescience/display/10.1093/acrefore/9780190228620.001.0001/acrefore-9780190228620-e-721;jsessionid=0C3989FA53A5FCBB227A714BB2E2D2C2.

Bessho, K., Date, K., Hayashi, M., Ikeda, A., Imai, T., Inoue, H., et al. (2016). An introduction to Himawari-8/9–japan's new-generation geostationary meteorological satellites. *J. Meteorological Soc. Japan Ser. II* 94 (2), 151–183. doi: 10.2151/jmsj.2016-009

Defu, L., Pang, L., and Xie, B. (2009). Typhoon disaster in China: prediction, prevention, and mitigation. *Natural Hazards* 49 (3), 421–436. doi: 10.1007/s11069-008-9262-2

Feng, H. (2019). The application of recurrent neural network in nowcasting. *Geophysical Res. Abstracts* 21.

Gan, J., and Hui, Qi (2020). "Research on weather radar nowcasting extrapolation," in *2020 International Conference on Computer Vision, Image and Deep Learning (CVIDL)*. Available at: https://ieeexplore.ieee.org/abstract/document/9270491.

He, K., Zhang, X., Ren, S., and Sun, J. (2016). "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Available at: https://openaccess.thecvf.com/content_cvpr_2016/html/He_Deep_Residual_Learning_CVPR_2016_paper.html.

Hong, S., Kim, S., Joh, M., and Song, S.-K. (2017a). Globenet: Convolutional neural networks for typhoon eye tracking from remote sensing imagery. *arXiv preprint arXiv:1708.03417*.

Hong, S., Kim, S., Joh, M., and Song, S.-K. (2017b). Psique: Next sequence prediction of satellite images using a convolutional sequence-to-sequence network. *arXiv preprint arXiv:1711.10644*.

Hoque, M. A.-A., Phinn, S., Roelfsema, C., and Childs, I. (2017). Tropical cyclone disaster management using remote sensing and spatial analysis: A review. *Int. J. disaster Risk reduction* 22, 345–3545. doi: 10.1016/j.ijdrr.2017.02.008

Huang, G., Liu, Z., van der Maaten, L., and Weinberger, K. Q. (2017). Densely connected convolutional networks. *Proc. IEEE Conf. Comput. Vision Pattern Recognit*. doi: 10.1109/CVPR.2017.243

Hu, J., Shen, Li, and Sun, G. (2018). "Squeeze-and-excitation networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Available at: https://openaccess.thecvf.com/content_cvpr_2018/html/Hu_Squeeze-and-Excitation_Networks_CVPR_2018_paper.html.

Japan Meteorological Agency (JMA) (2018).

Jia, S., Yansong, D., Wenbin, J., Jinmei, L., Xianzhong, L., Wen, N., et al. (2020). Study on evolution process of landslides triggered by typhoon rainstorm. *J. Eng. geol.* 28 (6), 1323–1332. doi: 10.13544/j.cnki.jeg.2019-540

Khan, A., Sohail, A., Zahoora, U., and Qureshi, A. S. (2020). A survey of the recent architectures of deep convolutional neural networks. *Artif. Intell. Rev*. 53 (8), 5455–5516. doi: 10.1007/s10462-020-09825-6

Kovordányi, R., and Chandan, R. (2009). Cyclone track forecasting based on satellite images using artificial neural networks. *ISPRS J. Photogrammetry Remote Sens.* 64 (6), 513–521. doi: 10.1016/j.isprsjprs.2009.03.002

Lee, J.-H., Lee, SS, Kim, HG, Song, S-K, Kim, S, and Ro, YM (2019). Mcsip net: Multichannel satellite image prediction *via* deep neural network. *IEEE Trans. Geosci. Remote Sens.* 58 (3), 2212–2224. doi: 10.1109/TGRS.2019.2955538

Lin, Z., Li, M., Zheng, Z., Cheng, Y., and Yuan, C. (2020). "Self-attention convlstm for spatiotemporal prediction," in *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 34. Available at: https://ojs.aaai.org/index.php/AAAI/article/view/6819.

Moskolaï, W. R., Abdou, W., and Dipanda, A. (2021). Application of deep learning architectures for satellite image time series prediction: A review. *Remote Sens.* 13 (23), 4822. doi: 10.3390/rs13234822

Mukherjee, S., Ghosh, S., Ghosh, S., Kumar, P., and Roy, P. P. (2019). "Predicting video-frames using encoder-convlstm combination," in *2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Available at: https://ieeexplore.ieee.org/abstract/document/8682158.

Neeru, J., and Kishtawal, C. M. (2010). Automatic determination of center of tropical cyclone in satellite-generated IR images. *IEEE Geosci. Remote Sens. Lett.* 8 (3), 460–463. doi: 10.1109/LGRS.2010.2085418

Ranzato, M., Szlam, A., Bruna, J., Mathieu, M., Collobert, R., and Chopra, S. (2014). Video (language) modeling: a baseline for generative models of natural videos. *arXiv preprint arXiv:1412.6604*.

Rüttgers, M., Lee, S., Jeon, S., and You, D. (2019). Prediction of a typhoon track using a generative adversarial network and satellite images. *Sci. Rep.* 9 (1), 1–15. doi: 10.1038/s41598-019-42339-y

Salehinejad, H., Sankar, S., Barfett, J., Colak, E., and Valaee, S. (2017). Recent advances in recurrent neural networks. *arXiv preprint arXiv:1801.01078*.

Shi, X. (2018). Exploring deep learning architectures for spatiotemporal sequence forecasting. *Diss*. doi: 10.14711/thesis-991012671057603412

Shi, X., Gao, Z., Lausen, L., Wang, H., Yeung, D.-Y., Wong, W.-K., et al. (2017). Deep learning for precipitation nowcasting: A benchmark and a new model. *arXiv preprint arXiv:1706.03458*.

Shultz, J. M., Jill, R., and Espinel, Z. (2005). Epidemiology of tropical cyclones: the dynamics of disaster, disease, and development. *Epidemiologic Rev.* 27 (1), 21–35. doi: 10.1093/epirev/mxi011

Siddique, N., Paheding, S., Elkin, C. P., and Devabhaktuni, V. (2021). U-Net and its variants for medical image segmentation: A review of theory and applications. *IEEE Access*. doi: 10.1109/ACCESS.2021.3086020

Tian, L., Li, X., Ye, Y., Xie, P., and Li, Y. (2019). A generative adversarial gated recurrent unit model for precipitation nowcasting. *IEEE Geosci. Remote Sens. Lett.* 17 (4), 601–605. doi: 10.1109/LGRS.2019.2926776

Tschannen, M., Bachem, O., and Lucic, M. (2018). Recent advances in autoencoder-based representation learning. *arXiv preprint arXiv:1812.05069*.

Wang, Y., Long, M., Wang, J., Gao, Z., and Yu, P. S. (2017). "Predrnn: Recurrent neural networks for predictive learning using spatiotemporal lstms," in *Advances in Neural Information Processing Systems 30 (NIPS 2017)*. Available at: https://proceedings.neurips.cc/paper/2017/hash/e5f6ad6ce374177eef023bf5d0c018b6-Abstract.html.

Wang, E. Ke, Wang, F., Kumari, S., Yeh, J.-H., and Chen, C.-M. (2021). Intelligent monitor for typhoon in IoT system of smart city. *J. Supercomputing* 77 (3), 3024–3043. doi: 10.1007/s11227-020-03381-0

Wang, Y., Wu, H., Zhang, J., Gao, Z., Wang, J., Yu, P., et al. (2021). PredRNN: A recurrent neural network for spatiotemporal predictive learning. *arXiv preprint arXiv:2103.09504*. doi: 10.1109/TPAMI.2022.3165153

Wang, X., Weixin, X., and Jiayi, S. (2018). "Learning spatiotemporal features with 3DCNN and ConvGRU for video anomaly detection," in *2018 14th IEEE International Conference on Signal Processing (ICSP)*. Available at: https://ieeexplore.ieee.org/abstract/document/8652354.

Weinkle, J., Maue, R., and Pielke, Jr. R. (2012). Historical global tropical cyclone landfalls. *J. Climate* 25 (13), 4729–4735. doi: 10.1175/JCLI-D-11-00719.1

Woo, S., Park, J., Lee, J.-Y., and Kweon, I. S. (2018). "Cbam: Convolutional block attention module," in *Proceedings of the European Conference on Computer Vision (ECCV)*. Available at: https://openaccess.thecvf.com/content_ECCV_2018/html/Sanghyun_Woo_Convolutional_Block_Attention_ECCV_2018_paper.html.

Xingjian, S. H. I, Chen, Z., Wang, H., Yeung, D.-Y., Wong, W.-K., and Wang-chun, W. O. O. (2015). Convolutional LSTM network: A machine learning approach for precipitation nowcasting. *Adv. Neural Inf. Process. systems*.

Xu, Z., Du, J., Wang, J., Jiang, C., and Ren, Y. (2019). "Satellite image prediction relying on gan and lstm neural networks," in *2019 IEEE International Conference on Communications (ICC)*. Available at: https://ieeexplore.ieee.org/abstract/document/8761462.

Yu, J., Gao, S., Zhang, L., Shen, X., and Guo, L. (2020). Analysis of a remote rainstorm in the Yangtze river delta region caused by typhoon mangkhut (2018). *J. Mar. Sci. Eng.* 8 (5), 345. doi: 10.3390/jmse8050345

Zhang, J., and Chen, Y. (2019). Risk assessment of flood disaster induced by typhoon rainstorms in guangdong province, China. *Sustainability* 11 (10), 2738. doi: 10.3390/su11102738

Zhang, Q. P., Lai, L. L., and Sun, W. C. (2005). "Intelligent location of tropical cyclone center," in *2005 International Conference on Machine Learning and Cybernetics*, Vol. 1. Available at: https://ieeexplore.ieee.org/abstract/document/1526984.

Zhaoqing, P., Weijie, Y, Xiaokai, Y, Asifullah, K, Feng, Y, and Zheng, Y (2019). Recent progress on generative adversarial networks (GANs): A survey.". *IEEE Access* 7, 36322–363335. doi: 10.1109/ACCESS.2019.2905015

frontiers | Frontiers in Marine Science

TYPE Methods
PUBLISHED 12 January 2023
DOI 10.3389/fmars.2022.1056404

OPEN ACCESS

EDITED BY
Zhenhong Du,
Zhejiang University, China

REVIEWED BY
Junhong Bai,
Beijing Normal University, China
Patrick Biber,
University of Southern Mississippi, United States

*CORRESPONDENCE
Grace S. Chiu
gschiu@vims.edu
Molly Mitchell
molly@vims.edu
Julie Herman
herman@vims.edu

SPECIALTY SECTION
This article was submitted to Ocean Observation, a section of the journal Frontiers in Marine Science

RECEIVED 28 September 2022
ACCEPTED 12 December 2022
PUBLISHED 12 January 2023

CITATION
Chiu GS, Mitchell M, Herman J, Longo C and Davis K (2023) Enhancing assessments of blue carbon stocks in marsh soils using Bayesian mixed-effects modeling with spatial autocorrelation — proof of concept using proxy data.
Front. Mar. Sci. 9:1056404.
doi: 10.3389/fmars.2022.1056404

Enhancing assessments of blue carbon stocks in marsh soils using Bayesian mixed-effects modeling with spatial autocorrelation — proof of concept using proxy data

Grace S. Chiu[1*], Molly Mitchell[1*], Julie Herman[1*], Christian Longo[2] and Kate Davis[3]

[1]Virginia Institute of Marine Science, William & Mary, Gloucester Point, VA, United States, [2]Emergency Medicine, RWJBarnabas Health Medical Group, West Orange, NJ, United States, [3]Independent Researcher, Nashotah, WI, United States

Our paper showcases the potential gain in scientific insights about blue carbon stocks (or total organic carbon) when additional rigor, in the form of a spatial autocorrelation component, is formally incorporated into the statistical model for assessing the variability in carbon stocks. Organic carbon stored in marsh soils, or blue carbon (BC), is important for sequestering carbon from the atmosphere. The potential for marshes to store carbon dioxide, mitigating anthropogenic contributions to the atmosphere, makes them a critical conservation target, but efforts have been hampered by the current lack of robust methods for assessing the variability of BC stocks at different geographic scales. Statistical model-based extrapolation of information from soil cores to surrounding tidal marshes, with rigorous uncertainty estimates, would allow robust characterization of spatial variability in many unsampled coastal habitats. In the absence of BC data, we consider a historical dataset (the best available) on soil organic matter (OM)—a close proxy of BC—on 36 tidal (fresh and salt) marshes in the Virginia portion of Chesapeake Bay (CBVA) in the USA. We employ Bayesian linear mixed(-effects) modeling to predict OM by marsh type, soil category, soil depth, and marsh site, whereby site effects are modeled as random. When the random site effects are additionally assumed to exhibit an intrinsic conditional autoregressive (ICAR) spatial dependence structure, this more complex model clearly suggests groupings of marsh sites due to their spatial proximity, even after adjusting for the remaining predictors. Although the actual membership of each group is not a focus of our proof-of-concept analysis, the clear presence of groupings suggests an underlying latent spatial effect at the localized-regional level within CBVA. In contrast, the non-spatially explicit model provides no clear indication of either spatial influence between sites or improvement in predictive power. The polar difference in conclusions between models reveals the potential inadequacy in relying on predictor variables alone to capture the spatial variability of OM across a geographic

domain of this size or larger. We anticipate that spatially explicit models, such as ours, will be important quantitative tools for understanding actual carbon measurements and for assessing BC stocks in general.

KEYWORDS

blue carbon, coastal sediment, spatial regression, conditional autoregressive dependence, Markov random field (MRF), Bayesian modeling and inference

1 Introduction

The storage of "blue carbon" (BC) is the temporary or long-term storage of carbon in aquatic natural systems. Through photosynthesis, shoot and root production and then burial of above- and below-ground biomass, marshes, mangroves, and seagrass beds can store carbon dioxide that is removed from the atmosphere. Depending on the plant community, seasonal, annual, and multiyear storage in biomass and long-term burial of sediments are all considered BC reservoirs (Mcleod et al., 2011), although the contribution of each to the global carbon cycle is different.

Under stable conditions, tidal marsh productivity greatly exceeds the rates of soil carbon metabolism and the carbon stock tends to build over time, resulting in an organic-rich upper layer. Carbon burial in vegetated coastal systems, when extrapolated to global extent, is 2 to 10 times those in shelf/deltaic sediments (Duarte et al., 2005) and exceeds temperate, tropical and boreal forests (when considered separately; Mcleod et al., 2011). The potential for marshes to remove and store carbon dioxide, mitigating anthropogenic contributions to the atmosphere, has made them a critical conservation target (Coverdale et al., 2014). Protecting these critical resources requires an understanding of the existing carbon stocks on a wide geographic scale (McTigue et al., 2019). However, marsh sediment cores tend to be relatively scarce and carbon stocks have been shown to vary with depth and across short geographic scales, leading to broad generalizations in blue carbon stock calculations. In 2011, Mcleod et al. (2011) stated, "There are no definitive studies of spatial variability within mangrove forests, salt marshes, or seagrass meadows other than studies addressing C burial differences." Since then, there have been efforts to estimate spatial variability (e.g., Lavery et al., 2013; Ewers Lewis et al., 2018), but no statistically robust method for assessing the variability in carbon stocks at different scales has been developed. Statistical model-based extrapolation of information from soil cores to surrounding tidal marshes and SAV (submerged aquatic vegetation) beds with rigorous uncertainty estimates would allow robust characterization of many unsampled coastal habitats.

Specifically, conventional methods to assess blue carbon stocks employ rather simplistic regression models that may account for the inherent spatial nature of blue carbon purely through the spatial nature of predictor variables (e.g., marsh type, depth). However, the same predictor variable in one localized spatial region may capture the inherent spatial variability more adequately than it does in another localized region. As such, the additional modeling of the remaining spatial variability — in the form of spatial autocorrelation — that is not already captured by predictor variables is crucial to producing realistic uncertainty estimates for model predictions (Le, 2006).

In this paper, we investigate the potential gain in scientific insights about blue carbon stocks when a spatial autocorrelation component is formally incorporated into the regression model. To investigate methodological approaches, we consider the historical data in Edmonds et al. (1990) on soil organic matter (OM) from 36 tidal fresh, brackish, and salt marshes in the Virginia portion of Chesapeake Bay (CBVA) in the eastern United States. In the absence of BC data for CBVA, our use of OM as a proxy is justified by its close positive relationship with total organic carbon (TOC, which is the measure of blue carbon storage) from marsh sediment samples (Mitchell, 2018). In particular, tidal marsh plants remove carbon dioxide from the atmosphere through photosynthesis. The carbon is used for producing and sustaining plant biomass. Plant and microflora biomass that is not exported from the marsh or respired/decomposed on the marsh surface is stored through burial in the sediments as OM. The potential for export of root biomass is particularly limited, so that almost all of the root biomass is stored in soil until it decomposes (Chmura, 2009). Marsh soils have high organic input, which is typically easily decomposed; however, they are also highly anaerobic, slowing decomposition. Therefore, root biomass decomposes very slowly and carbon does not decline significantly with depth in salt marsh soils (Connor et al., 2001). In addition to plant productivity, tidal marshes have high rates of benthic microflora production (which can equal or exceed macrophyte production; Sullivan and Currin, 2002) that contributes to carbon pools when buried. Rates of biomass decomposition vary with plant type and along a salinity gradient (Craft, 2007) and therefore may result in different OM profiles in different tidal marsh types. OM is composed of multiple components, namely, carbon, nitrogen, phosphorus; TOC only represents the carbon fraction of tidal marsh OM. As such, TOC is typically much lower than organic

matter (e.g., 7% compared to 30%, respectively). That is, while the actual values of OM and TOC will differ, patterns of similarity between marshes are expected to be consistent between OM and TOC. Therefore, we anticipate the insights from the methodology in this paper to be highly relevant to the modeling strategies for assessing blue carbon in general.

The rest of the paper is structured as follows. In Section 2, we describe the data on organic matter and associated predictor variables in Edmonds et al. (1990), and how we reconstruct the sampling sites' geographic coordinates, which are absent in the original data. In Section 3, we discuss the development of our Bayesian linear mixed(-effects) models to predict organic matter by marsh type, soil category, soil depth, and marsh site, whereby marsh site effects are modeled as random. For spatial autocorrelation, we explain our focus on the conditional autoregressive (CAR) structure, also known as a Markov random field (MRF). In particular, we consider an intrinsic CAR (ICAR) structure, which is a degenerate case of CAR. In Section 4, we compare the model results, predictive performance, and scientific implications between two scenarios: (1) assuming spatially autocorrelated marsh site effects based on the reconstructed geographic coordinates, and (2) assuming uncorrelated marsh site effects. Finally, in Section 5, we discuss the broader scientific implications of this statistical approach on estimating blue carbon stocks in marshes.

2 Materials and equipment

2.1 Historical data on organic matter

This paper leverages data published in a historic technical report (Edmonds et al., 1990) on marsh soils throughout the Virginia portion of Chesapeake Bay (CBVA). Chesapeake Bay is the largest estuary in the USA, with more than 18,800 km of shoreline in Maryland and Virginia, and more than 3,800 km^2 of tidal marshes, containing a diverse array of tidal marsh types and ecologies. Within Virginia, CBVA possesses a wide range of salinities from approximately 35 ppt near the mouth of CBVA, to 0 ppt in the upper reaches of the estuarine rivers and in the small tributary creeks found along their edges and is broadly representative of the range of conditions found in mid-Atlantic marshes. Marshes stretch along the entire shoreline, with approximately 25% tidal freshwater marsh, 15% oligohaline marshes, 30% mesohaline marshes and 30% salt marsh (CCRM, 2017). Marsh sediments within CBVA are considered to be predominately mineral with organic matter typically averaging around 30% (or 300 g/kg). A survey of marshes along the York River (a tributary of CBVA), found that organic matter soil contribution ranged from 4–58% (but only exceeded 50% at one site, Mitchell, 2018).

The technical report focused on describing different marsh soil types and did not attempt spatial analysis of organic matter; however, it offers a dataset well suited to the development of a geospatial statistical model because 1) all the cores were taken using the same methods and with standardized core depths, 2) the cores covered broad geospatial and estuarine ranges, sampling everything from tidal fresh to saltwater marshes, and 3) the report included detailed site descriptions including vegetative characteristics. The soil organic matter in the cores, recorded in g/kg of soil, was estimated by the acid-dichromate digestion method (see Supplementary Material).

2.2 Reconstruction of geographic coordinates

Edmonds et al. (1990) did not include precise geographic coordinates (i.e., latitudes and longitudes) for the 38 marshes sampled. However, for each marsh, they described three markers, and the distance and bearing from each marker to the sampling location (e.g., 3.5 miles, 180 ° from the intersection of Route 5 and Route 31). This allowed us to triangulate the location of the sampling site. The uncertainty associated with this triangulation is estimated to be in the tens of meters, thus correctly identifying marshes, but not necessarily the exact sampling point within the marsh. Figure 1 displays the reconstructed locations of the 38 marshes.

2.3 Predictors of organic matter

Depth is a key correlate of organic matter content (OM). Within the marsh sediments, OM tends to decline with depth, more rapidly in the top ∼ 20 cm and less rapidly in deeper sediments as the remaining OM is more refractory (Morris and Bowden, 1986). Edmonds et al. (1990) considered three depth ranges for mineral soils: 0–13 in (∼ 0–33 cm), 13–26 in (∼ 33–66 cm), and 26–40 in (∼ 66–102 cm). (See Supplementary Material for additional information about depth ranges.) Instead of regarding depth as a categorical predictor with ordinal values of *top, mid*, and *bottom*, we take depth as a numerical predictor by taking approximate middle values of the mineral soil ranges in inches, namely, 6, 19, and 33. Employing an explicitly numerical predictor simplifies the regression model structure, in the presence of various other multilevel categorical predictors.

Among the remaining predictor variables, all are categorical, and they also contain overlapping information about the marsh environment:

- Soil category (5 levels, Figure 2A): Edmonds et al. (1990) describe 6 soil categories, namely, Fluvaquent, Hydraquent, Sulfaquent, Sulfihemist, Hydraquentic Humaquept, and Histic Humaquept. Since the Humaquepts are the only categories with modifiers and there is only one of each, we omit those from the

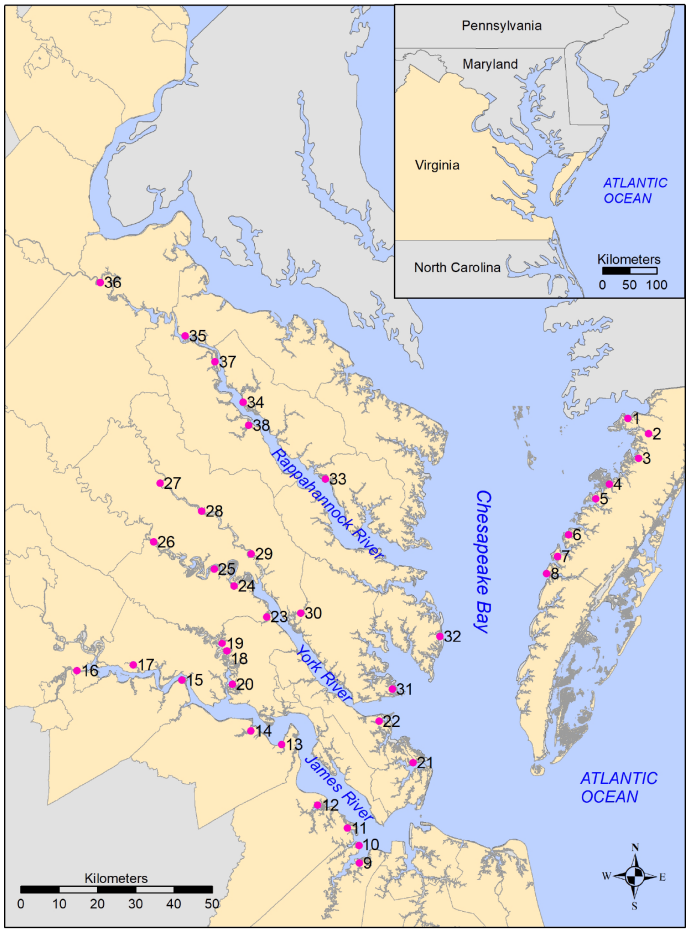


FIGURE 1
Reconstructed locations of marshes (sampling sites) in Edmonds et al. (1990), with inset showing relationship of study area to the Mid-Atlantic coast, USA.

full set of 38 marshes in the sample in order to preserve statistical power, leaving a reduced set of 36 marshes.

- Soil type (12 levels, Figure 2B), as given in Edmonds et al. (1990): In soil surveys, these are formally known as soil components and tend to be locality specific. As such, it is difficult to compare them across the spatial extent of the study.
- Vegetation type (7 levels, Figure 2C), as given in Edmonds et al. (1990): Arrow Arum-PickerelWeed (APW), Big CordGrass (BCG), Black NeedleRush (BNR), Brackish Water Mixed (BWM), Freshwater Mixed Community (FMC), Saltmarsh CordGrass (SCG), and SaltMeadow Combined (SMC).
- Inundation (3 levels, Figure 2D) and Salinity (2 levels, Figure 2E): Tidal marsh vegetation is directly controlled by inundation level and salinity of the environment (Anderson et al., 2022), with particular species or communities in saline, brackish, and freshwater areas, and subsets of those communities at different elevations within the tide range. Therefore, estimated salinity and period of inundation can be inferred from the plant community. In addition, spatially explicit salinity datasets (based on the Chesapeake Bay Program's salinity assignments, as described in Mitchell et al., 2020) were used to help characterize local conditions into 2 levels of salinity (salt, fresh). For inundation, we initially connected the vegetative communities as described in Edmonds et al. (1990) to 3 levels (low, mid, high). Because the distribution of marsh plants are highly regulated by the frequency of inundation, we assigned *low* inundation to reflect a community dominated by high marsh plants, and *high* inundation, a community dominated by low marsh plants. The mid inundation category encompassed sites where the vegetation was evenly mixed between high and low marsh plants. However, preliminary analyses using the

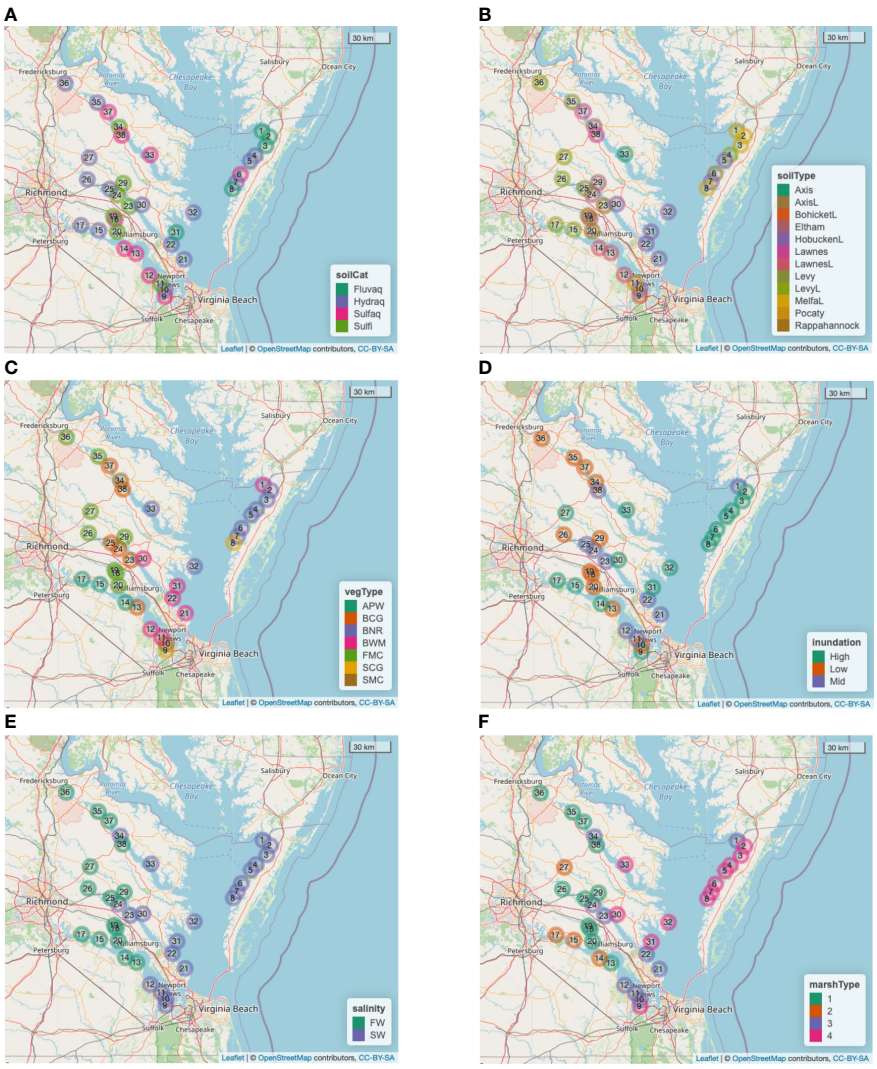


FIGURE 2
Spatial distributions of categorical variables across the 36 marshes considered for modeling (i.e., excluding Sites 16 and 28). **(A)** *Soil category* (Fluvaquents, Hydraquents, Sulfaquents, Sulfihemists). **(B)** *Soil type* ("L" at the end of a category denotes "loam"). **(C)** *Vegetation type*. **(D)** *Inundation*. **(E)** *Salinity* (Freshwater, Saltwater). **(F)** Newly defined *Marsh type* by crossing *Inundation* with *Salinity* ('1'=Low/Mid–FW, '2'=High–FW, '3'=Low/Mid–SW, and '4'=High–SW).

3-level categorization identified the unnecessary breakdown between *mid* and *low* inundation (Longo, 2022). Therefore, in this paper we combine them as a single *not-high (= low/mid)* category.

The large number of categorical predictors relative to the number of marshes poses a challenge to modeling. For a regression model, each categorical predictor contributes a number of unknown model parameters that is no fewer than the number of levels minus one; this base number increases multiplicatively if interactions between predictors are considered. For this reason, we take the following measures to preserve statistical power:

- Omit soil type.
- Omit Sites 16 and 28, the only Humaqueptic sites under soil category.
- Replace the original vegetation type categorization with a new 4-level categorization of marsh type, defined by crossing the 2-level inundation and 2-level salinity. The resulting marsh type categories are '1' = Low/Mid–FW, '2' = High–FW, '3' = Low/Mid–SW, and '4' = High–SW (Figure 2F).

The new definition of marsh type implies collinearity between marsh type and each of inundation and salinity, so that the latter

two cannot enter the regression model as predictors alongside marsh type. Therefore, the final set of regression predictors includes marsh type, depth, soil category, and marsh site.

2.4 Transformation of organic matter measurements

The dependent variable of interest is sediment core OM, recorded in g/kg of soil. For linear mixed modeling, the distribution of the dependent variable should not exhibit substantial skewness. Figure 3 shows that the distribution of these OM data are highly right-skewed, but log-transformation substantially reduces the skewness. Therefore, we model log(OM) as a response of the above predictor variables.

3 Methods

3.1 Bayesian spatial regression modeling

Point-referenced spatial data, such as ours, are often modeled using a classical geostatistical (continuous-space) model, in which the covariance between two spatial locations is described by a proposed functional form (e.g., exponential decay) with unknown parameters (see, e.g., Zimmerman and Stein, 2010). For our data, however, the sampled marshes fall along highly irregular shorelines in the lower Chesapeake Bay. That is, the sampling universe of the study area is far from being a contiguous spatial domain, and it would be rather unreasonable to employ a modeling approach that is meant for continuous space.

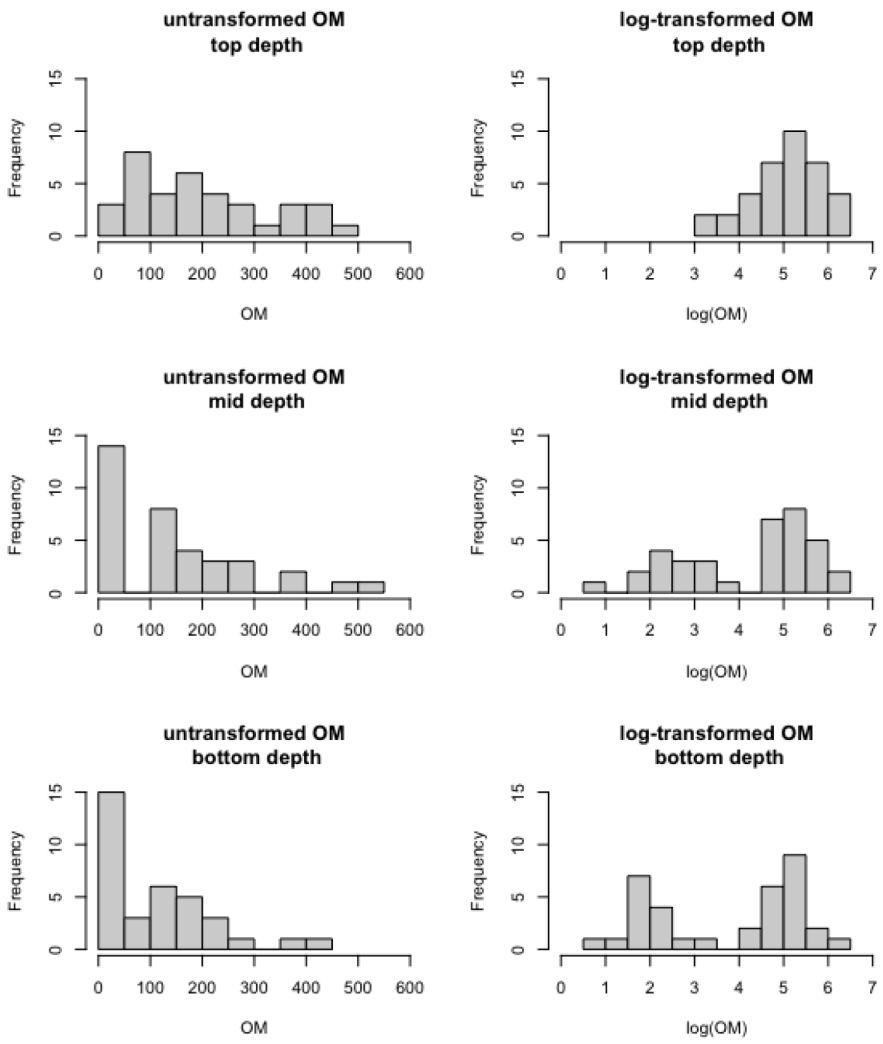


FIGURE 3
Distributions of sediment core organic matter in g/kg soil (left panels) and log(g/kg soil) (right panels) across the 36 marshes observed at the top depth (top panels), middle depth (middle panels), and bottom depth (bottom panels), respectively.

An alternative is to assume a conditional autoregressive (CAR) spatial structure. CAR is a nearest-neighbor dependence structure that is typically employed to model areally aggregated data, in the form of either a lattice (i.e., a regular grid, e.g., Chiu et al., 2013), or a set of irregular polygons (e.g., Hyman et al., 2022). Two areal units are considered immediate neighbors if they share a border. CAR models with first-order dependence only consider nearest neighbors. Higher-order dependence CAR models are possible by further modeling the dependence of a site with its neighbors' immediate neighbors (e.g., Chiu and Lehmann, 2011). For the CAR structure to model point-referenced data, the investigator must decide what constitutes a pair of neighbors. White and Ghosh (2009) developed the stochastic neighborhood CAR, or SNCAR, model for point-referenced data, while avoiding the need for an upfront definition of neighbors. The SNCAR model infers a suitable threshold distance as an unknown model parameter, whereby only those marshes located within each other's threshold radius are considered neighbors. For neighboring marshes, the SNCAR model further assumes an exponential decay covariance.

However, the SNCAR model is highly complex; for a first attempt of spatially explicit modeling of data with potential implications for blue carbon stocks, we choose the simpler first-order intrinsic CAR, or ICAR, dependence structure (e.g., Chiu et al., 2013) with a rather arbitrary threshold. ICAR is a degenerate case of CAR, whereby the conditional spatial correlation between neighbors (given all other marshes) is fixed at 1. For the threshold radius, one extreme would be to take it as the smallest pairwise distance in the dataset, leading to the sparsest possible neighborhood structure, and containing as few as a single pair of neighbors. The other extreme would be to take the largest pairwise distance as the threshold, so that every site is a neighbor of every other site. Mathematically, a certain level of sparsity is required for model estimability; with only 36 marshes, estimating the spatial parameters alongside numerous regression coefficients is rather ambitious. As such, we construct preliminary models using a threshold radius of 16 km, 24 km, and 36 km, respectively, each of which results in a neighborhood structure that is reasonably sparse without having too few pairs of neighbors to inform the statistical inference for the spatial parameters. The preliminary results show little sensitivity to the thresholds (see Supplementary Material), and subsequently we take 24 km as the threshold for formal modeling. The resulting neighborhood structure is depicted as an adjacency matrix in Figure 4.

3.2 Model statements

Our spatial regression model appears as Eqs. (1)–(3) below. It is a linear mixed model, with log-transformed OM regressed on marsh site (categorical), marsh type (categorical), soil category (categorical), depth (numerical, based on ordinal depth ranges), and the interaction between depth and marsh type; marsh site effects are modeled as random, on which the ICAR spatial autocorrelation structure is assumed. The interaction between depth and marsh type allows the slope between log(OM) and depth to differ across marsh types. We do not consider other interaction terms in the model due to either in estimability or no improvement in goodness-of-fit based on preliminary analyses (see Supplementary Material).

For the i -th marsh site (i=1,…,36) at the j -th depth (j=1,2,3 , except for i=11 for which j=1,2 only), let $y_{ijk\ell}$ denote the log(OM) associated with the numerical predictor, depth, denoted by x_{ij} (=6,19,33) , and marsh type k and soil category ℓ . Note that the 107 pairs of (i,j) are nested inside k and ℓ . Further, let w_{ij} denote the standardized values of x_{ij} , i.e., the 107 values of w_{ij} 's have mean 0 and standard deviation 1 (depths are standardized to ensure numerical stability). Then, the regression model is

$$y_{ijk\ell} = \beta_0 + \alpha_{1k} + \gamma_\ell + (\beta_1 + \alpha_{2k})w_{ij} + \phi_i + \epsilon_{ijk\ell} \quad , \tag{1}$$

$$\phi_i \sim \text{ICAR}, \text{Var}(\phi_i|\sigma_\phi^2) = \sigma_\phi^2 \quad , \tag{2}$$

$$\epsilon_{ijk\ell}|\sigma^2 \sim N(0, \sigma^2) \tag{3}$$

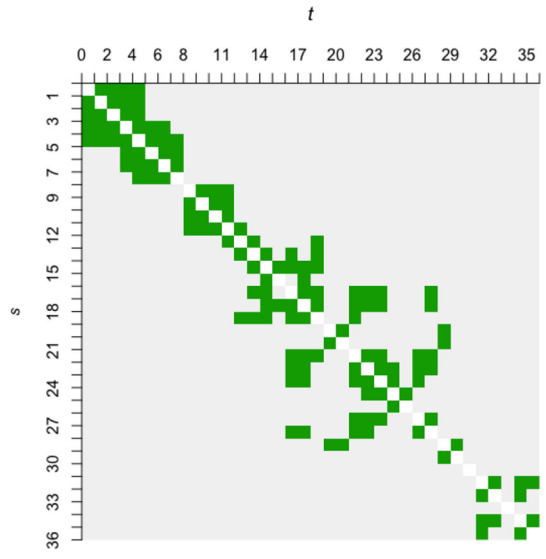


FIGURE 4
The adjacency matrix (symmetric) of 36 × 36 cells based on a 24 km threshold radius. Marshes s and t are considered neighbors if they are located within the threshold radius of each other, and the cell on the s -th row and t -th column of the adjacency matrix appears green. A gray cell denotes non-neighbors. White cells along the diagonal denote *not applicable*, as the notion of neighborhood between a marsh and itself is undefined. Note that the cell labels run from 1 to 36 (not 38) due to the omission of Sites 16 and 28.

where β_0 and β_1 are the overall intercept and slope parameters with respect to depth; α_{1k} and γ_ℓ are the main effects (fixed effects) on $y_{ijk\ell}$ due to marsh type k(=1,…,4) and soil category ℓ(=1,…4) , respectively; α_{2k} is the interaction effect (fixed effect) on $y_{ijk\ell}$ between marsh type and depth; ϕ_i is the spatially autocorrelated random effect due to the i -th marsh site, with variance σ_ϕ^2; and $\epsilon_{ijk\ell}$ is white noise (regression error), with variance σ^2 . Fixed effects are subject to the linear constraint $\alpha_{11}=\alpha_{21}=\gamma_1=0$.

For Bayesian inference, all unknown parameters (appearing only on the right-hand-side of Eqs. (1)–(3)) require prior distributions that reflect the modeler's *a priori* understanding unrelated to the data at hand. Prior distributions can be inspired by historical datasets or mechanistic models. However, when an obvious source of inspiration is absent, priors that are diffuse with respect to exploratory analyses may be employed to reflect the modeler's lack of concrete knowledge about the parameters. We take the latter approach in this paper for Eqs. (4) and (6) using relatively diffuse normal and lognormal distributions, respectively (see Supplementary Material for details):

$$\beta_0, \beta_1, \alpha_{1k}, \alpha_{2k}, \gamma_\ell \sim N\left(0, 10^2\right) \quad \text{for} \quad k, \ell = 2, 3, 4 \quad , \tag{4}$$

$$\sigma_\phi \sim \text{uniform}(0.22, 1) \quad , \tag{5}$$

$$\log(\sigma) \sim N\left(0, 1.5^2\right) \tag{6}$$

We elaborate upon our choice of the uniform prior distribution for σ_ϕ from Eq. (5) in Section 4.

For model comparison, we also fit the simpler mixed model whereby the random marsh site effects ϕ_i are assumed to be uncorrelated, i.e., Eq. 2 is replaced by $\phi_i|\sigma_\phi^2 \sim N(0, \sigma_\phi^2)$; to simplify the code, we use the same log-normal prior distribution from Eq. (6) for both σ and σ_ϕ .

3.3 Model implementation

For exploratory modeling, we employ classical (non-Bayesian) regression that requires minimal computational effort. The specific software packages are the R functions `lme4::lmer` (Bates et al., 2015) and `mgcv::gam` (Wood, 2017) — see Supplementary Material for details. For formal modeling, we implement our Bayesian models in the Stan modeling language interfaced with the `rstan` package (Stan Development Team, 2022). Stan conducts Markov chain Monte Carlo (MCMC) simulations from the posterior distribution, which is the joint probability distribution of all model parameters given the observed data and stipulated model from Section 3.2. In addition to the posterior distribution on which Bayesian inference is based, MCMC simulations also facilitate a wide range of inference summaries and model diagnostics.

The posterior distribution is typically highly complex for spatially explicit models, in which case MCMC would be computationally intensive. Our code is run on a high performance computing cluster for both the ICAR and uncorrelated models. The code and output appear in Supplementary Material.

4 Results

4.1 Technical results

The uncorrelated model converges readily with very good MCMC mixing (statistical behavior of the simulated posterior draws). As anticipated, the ICAR model requires substantially more computational effort. Moreover, if the prior distribution for σ_ϕ in Eq. (5) is replaced with one that extends to the left towards 0, such as a diffuse lognormal or uniform(0.01, 1), then the posterior distribution exhibits multimodality. While a multimodal posterior may not suggest an inherently flawed model, the model as stipulated by Eqs. (1)–(6) results in better MCMC mixing, a unimodal posterior, and some evidence of better cross-validation performance, according to the leave-one-out information criterion (LOOIC) and expected log predictive density (ELPD) *via* the R package `loo`. (Vehtari et al., 2017) (Table 1).

According to Table 1, predictive performance is best for the ICAR model with a uniform(0.22, 1) prior for σ_ϕ , followed by the same model but with a σ_ϕ prior that extends towards 0, then by the uncorrelated model. In Subsection 4.2, we discuss the vastly different conclusions based on the best ICAR model and the uncorrelated model. Before that, first note that the difference in ELPD between the best ICAR model and the uncorrelated model is essentially the size of its SE. Thus, strictly speaking, the best and worst models (and other models in between) are rather statistically consistent with each other with respect to predictive performance. This implies that spatial autocorrelation cannot be clearly ruled out, even if the evidence is mild.

If our key objective here were to make scientific discoveries based on the statistical inference from this dataset, it might be customary to adhere to the principle of parsimony and select the least complex model that is not clearly worse in predictive performance. However, because our least and most complex models yield vastly different scientific insights, more thought must be given to the process of model selection. In fact, our needs do not lie in the statistical inference from this historical proxy dataset. Rather, the objective of our paper is to demonstrate the potential for scientific insights into blue carbon stocks at various spatial scales by employing the methodology of spatial regression modeling; the proxy data are employed here to demonstrate the methodology.

TABLE 1 LOOIC and ELPD of the ICAR and uncorrelated models.

		prior distribution of σ_ϕ	
		uniform(0.22, 1)	extends towards 0
ICAR model	LOOIC	212.2	216.0 to 217.1
	ELPD	**– 106.1**	– 108.0 to – 108.6
uncorr model	LOOIC	—	221.2
	ELPD	—	**– 110.6**
ELPD diff, bold values only	– 4.5 (SE=4.6)		

A lower LOOIC value and higher ELPD value correspond to better leave-one-out predictive performance. The ELPD difference and standard error (SE) are obtained using the `loo_compare()` function in the `loo` package.

4.2 Implications of spatial ICAR vs. uncorrelated random marsh site effects

Spatial regression is a type of model-based spatial smoothing. In this case, the quantities being smoothed over space is the marsh site random effects ϕ_i. Figure 5 shows that the posterior median values of ϕ_i indeed are more spatially similar in size and in sign (positive or negative) for the ICAR model (panels A and C) than for the uncorrelated model (panels B and D). Also note that the posterior medians for the ICAR model are collectively larger with narrower Bayesian confidence intervals (CIs), implying that the marsh site effects on OM are more statistically noticeable. Importantly, there is noticeable spatial grouping of the ϕ_i 's — based on minimal overlapping of CIs in Figure 5C, a possible partition of the marshes into distinct groups is: {1–8} (all Eastern Shore sites), {9–15, 17–20, 23–27, 29, 30, 34–38} (all upstream sites except for sites 9–12), {21, 22, 31, 32} (all sites on the lower western shore), and {33} (the only outstanding upstream site). A finer grouping, although less statistically evident due to overlapping CIs, may be roughly defined as these sets of marshes: {1–8}, {9–12, **17**, *24, 25*}, {13, 14, 15, 18, 19, 20, 23, 29, 30}, {21}, {22, 31, 32}, {26, 27, **36**}, {33}, and {34, 35, 37, 38}. All marshes within each set in the finer grouping — except for the two sets that include site numbers displayed in bold or italic — exhibit spatial proximity in Figure 5A. Among the four "outsider" marshes given in bold and italic, sites 24 and 25 are located next to each other.

Therefore, overall, the ICAR results suggest that when spatial autocorrelation is formally modeled alongside predictor variables that have been documented to be potential drivers of blue carbon stock, a latent spatial grouping emerges, one that is not already accounted for by the predictors. This spatial phenomenon is clearly absent from the results of the uncorrelated model, as demonstrated by the lack of spatial pattern among the posterior medians in Figures 5B, D.

With respect to the predictor variables, their statistical relevance to OM is clearly demonstrated in either model (Figures 5E, F). In particular, soil category matters to OM (at least one γ_ℓ has an 80% CI that is far from 0), and both depth and marsh type are relevant to OM at least through their interaction (the CIs for two of three α_{2k} 's are far from 0), even if the respective main effects may be less statistically evident (0 is inside the CI for β_1 from both models, and none of the CIs for α_{1k} 's from the ICAR model is far from 0).

Finally, in Figures 5E, F, the non-restrictive nature of our prior distributions is evidenced by their flatness, contrasted with the peakedness of the marginal posterior distributions of the individual parameters. In other words, the statistical inference is heavily dominated by the information contributed by the data and minimally by our rather arbitrary choice of the prior distributions.

5 Discussion

The statistical model-based analysis in this paper shows that the inclusion of geospatial considerations when examining differences in tidal marsh OM stocks can result in a markedly different scientific interpretation of variability across marshes. This results in important implications for making inference of blue carbon stocks when it is based only on a few localized samples but extrapolated to unsampled areas. There is an increasing interest in estimating blue carbon stocks both for global climate models and for economic valuation, such as the growing blue carbon market (e.g., van den Bergh and Botzen, 2015). Enhancement of blue carbon stocks is a targeted activity for mitigating climate change (Trumper et al., 2009) but practical implementation of this activity requires a sophisticated understanding of blue carbon stocks and sequestration rates. Macreadie et al. (2019) called out the importance of reducing uncertainties about blue carbon stocks as critical for proper valuation of the resources and the calculation of carbon dioxide emission offsets. The actual number of carbon cores that can be used to infer this information, however, is geospatially limited, and its spatial variability is generally unknown. This has led to a tendency to

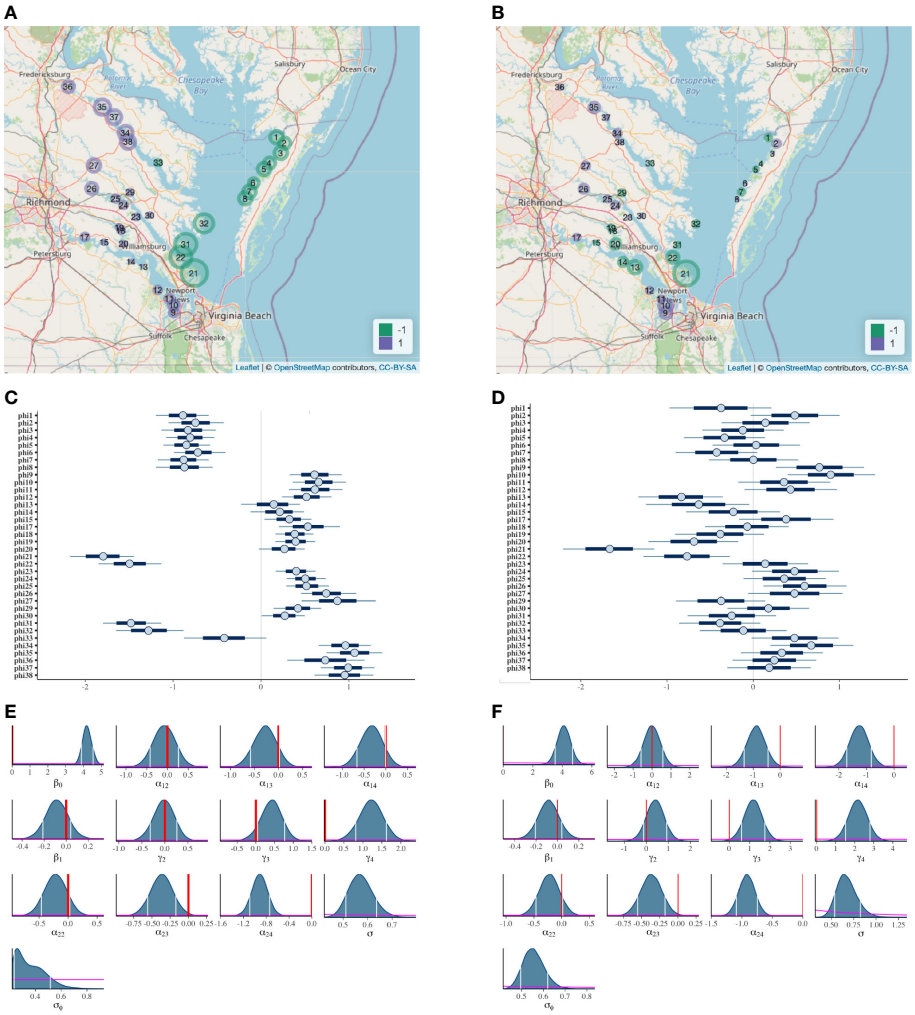


FIGURE 5
Bayesian inference results from the ICAR model of Eqs. (1)–(6) (panels **A**, **C**, and **E**) and uncorrelated model (panels **B**, **D**, and **F**). Panels **(A, B)**: bubble plot of ϕ_i (marsh site effect), where a large purple bubble represents a highly positive value for the estimated marsh site effect (posterior median of ϕ_i), and a large green bubble, a highly negative value. Panels **(C, D)**: Bayesian confidence intervals for ϕ_i at confidence levels of 50% (dark blue line) and 80% (light blue line); circles denote estimated marsh site effects. Panels **(E, F)**: marginal posterior distributions for the regression coefficients and variance parameters, with 80% confidence intervals delimited by white vertical lines, prior distributions shown in magenta, and the value 0 marked in red for the regression coefficients.

use a single value for blue carbon sequestration rates and stocks across all marshes (e.g., Chmura et al., 2003; Duarte et al., 2013; Howard et al., 2014). This approach is recognized as limiting inference of carbon stocks due to the high spatial variability in marsh carbon content, but is a necessary approach where data are limited (Lawrence et al., 2012; Ewers Lewis et al., 2018). Our model-based analysis suggests that, while the use of a single average number may be appropriate in some circumstances, it may be dramatically incorrect in others.

Specifically, in the case of the historical OM data, when geospatial considerations are incorporated (Figure 5C), it is clear that marsh site effects tend to have tighter confidence intervals and their estimates (posterior medians) tend to be much closer to adjacent marshes than marshes further away. The similarity between adjacent marshes is interesting, since (i) it suggests that there are some location-specific marsh characteristics beyond marsh type, depth, or soil category that are affecting OM stocks, and (ii) a single average value of log(OM) would be a grossly inappropriate representation of tidal marsh OM content across the study region, masking the distinction among group averages as revealed by the spatially explicit model. It has been previously shown that autochthonous OM sources such as marsh vegetation are an important control on stocks (since different marsh plants have different levels of productivity, persistence, and recalcitrance; see Ewers Lewis et al., 2020) and our model results support those findings (α_{2k} 's in Figures 5E, F). The geospatial analysis suggests

that OM content of allochthonous sediment contribution maybe equally important to the stock within a given marsh. For example, Figure 5C shows that sites 21, 22, 31, and 32 are less tightly grouped than sites 1–8, although site 32 shares a vegetative community with sites 2–6. Because sites 21, 22, 31, and 32 are facing the mouth of the Chesapeake Bay, they are susceptible to sediment erosion and sand overwash during high energy wave events. This may reduce the retention of autochthonous OM and add allochthonous mineral sediment to the marshes, and could be the reason for site 32 to be distinguishable from 2–6.

In contrast, when geospatial considerations are excluded from the statistical model, groupings between marshes are limited: the values of the estimated marsh site effects are distributed broadly, with confidence intervals that are wider and tend to overlap (Figure 5D). That is, based on the more simplistic uncorrelated model, spatial variability is substantial even after adjusting for the predictor variables marsh type, depth, and soil category. This large variability suggests that a single average value of log(OM) — given a specific combination of marsh type, soil category, and soil depth — might be a justified representation across the entire Chesapeake Bay, if the scientific purpose was to estimate the region's OM stock for use in a climate model, for example. However, based on the widely variable values of the estimated marsh site effect, it is clear that a single number does not represent any given marsh particularly well (Figure 5D). Thus, when the purpose is to estimate the OM content of any particular marsh, the single-value approach could lead to valuable loss of site-specific information about OM stocks. This reasoning also highlights the importance of the distribution of sampling sites for different inference purposes. For large-scale OM estimation, it is clearly important for the sampling sites to be relatively evenly spaced across the spatial region of inference to allow adequate interpolation of OM stocks. For example, according to our geospatial analysis, using data from the Rappahannock River could provide a biased estimate of OM stocks in the upper James River, despite the similarity in salinity distribution and marsh plant vegetation. However, for estimation of OM stocks in a particular marsh — for example, an unsampled marsh on the lower James — samples that are clustered around the marsh of interest may be required.

Note that within a tidal marsh, both plant composition and hydrologic regimes tend to be stable over annual to decadal time frames. Also, although upland soil organic carbon can vary temporally with changes in land use (e.g., Ma et al., 2018), it tends not to when land use is consistent over sampling periods (López-Teloxa et al., 2017). Therefore, our models have not considered temporal variation. In contrast, the tidal marsh OM dataset featured in our paper has allowed us to explore the importance of geospatial consideration in the potential future modeling of blue carbon stocks, because carbon stock datasets with broad geospatial representation and consistent sampling techniques across space are rare to date. We anticipate that our approach and implications discussed above regarding OM will be directly applicable to carbon stock modeling, due to the high correlation between tidal marsh OM and total organic carbon (TOC; Craft et al., 1991). Because TOC values are almost always much lower than OM values (e.g., 4% and 36%, respectively), a rigorous geospatial modeling approach that can discern subtle changes may prove to be especially important. We anticipate applying the statistical methodology of this paper to new TOC data (currently being collated) from an extensive sampling effort of CBVA marshes which was conducted in November 2021 as part of a collaborative grant with the US National Resources Conservation Service. Our methodology should result in improved insights into blue carbon stocks compared to conventional approaches for extrapolating known blue carbon stocks to unsampled marshes.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding authors.

Author contributions

GC, MM, and JH conceived the research and wrote the paper. MM chose and formatted the data, reconstructed the site locations, and constructed the variables *inundation* and *salinity*. Under the guidance of GC: KD conducted graphical data exploration of a subset of the original sampling sites; CL formalized these exploratory results by building and implementing an ICAR model for the same subset. GC extended the model by CL to accommodate the 36 sites of this paper, including model implementation (for all model variants in this paper) and model diagnostics. MM and JH advised on model refinement. All authors contributed to the article and approved the submitted version.

Funding

FO USDA-NRCS-NHQ-SOIL-20-NOFO0001003, 2020 Soil Science Collaborative Research Proposals Award # NR203A750025C010.

Acknowledgments

The authors acknowledge William & Mary Research Computing for providing computational resources and technical support that have contributed to the results reported within this paper (URL: https://www.wm.edu/it/rc). We thank the two reviewers for their valuable input. We also thank C. Friedrichs (VIMS) for suggesting this Special Topic as a publication outlet for our research.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmars.2022.1056404/full#supplementary-material

References

Anderson, S. M., Ury, E. A., Taillie, P. J., Ungberg, E. A., Moorman, C. E., Poulter, B., et al. (2022). Salinity thresholds for understory plants in coastal wetlands. *Plant Ecol.* 223, 323–337. doi: 10.1007/s11258-021-01209-2

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Software* 67, 1–48. doi: 10.18637/jss.v067.i01

CCRM (2017). *Comprehensive coastal inventory program* Gloucester Point, Virginia: Center for Coastal Resources Management (CCRM) Digital Tidal Marsh Inventory Series, Virginia Institute of Marine Science.

Chiu, G. S., and Lehmann, E. A. (2011). "Bayesian Hierarchical modelling: incorporating spatial information in water resources assessment and accounting," in *MODSIM2011, 19th International Congress on Modelling and Simulation*, Perth, Australia. 3349–3355.

Chiu, G. S., Lehmann, E. A., and Bowden, J. C. (2013). A spatial modelling approach for the blending and error characterization of remotely sensed soil moisture products. *J. Environ. Stat* 4(9). Available at: http://www.jenvstat.org/v04/i09.

Chmura, G. L. (2009). "Tidal salt marshes," in *The management of natural coastal carbon sinks*. Eds. D.d'A. Laffoley and G. Grimsditch Gland, Switzerland: International Union for Conservation of Nature and Natural Resources (IUCN).

Chmura, G. L., Anisfeld, S. C., Cahoon, D. R., and Lynch, J. C. (2003). Global carbon sequestration in tidal, saline wetland soils. *Global Biogeochemical Cycles* 17. doi: 10.1029/2002GB001917

Connor, R. F., Chmura, G. L., and Beecher, C. B. (2001). Carbon accumulation in bay of fundy salt marshes: Implications for restoration of reclaimed marshes. *Global Biogeochemical Cycles* 15, 943–954. doi: 10.1029/2000GB001346

Coverdale, T. C., Brisson, C. P., Young, E. W., Yin, S. F., Donnelly, J. P., and Bertness, M. D. (2014). Indirect human impacts reverse centuries of carbon sequestration and salt marsh accretion. *PloS One* 9, e93296. doi: 10.1371/journal.pone.0093296

Craft, C. (2007). Freshwater input structures soil properties, vertical accretion, and nutrient accumulation of Georgia and U.S tidal marshes. *Limnol. Oceanogr.* 52, 1220–1230. doi: 10.4319/lo.2007.52.3.1220

Craft, C. B., Seneca, E. D., and Broome, S. W. (1991). Loss on ignition and kjeldahl digestion for estimating organic carbon and total nitrogen in estuarine marsh soils: Calibration with dry combustion. *Estuaries* 14, 175. doi: 10.2307/1351691

Duarte, C. M., Losada, I. J., Hendriks, I. E., Mazarrasa, I., and Marbà, N. (2013). The role of coastal plant communities for climate change mitigation and adaptation. *Nat. Climate Change* 3, 961–968. doi: 10.1038/nclimate1970

Duarte, C. M., Middelburg, J. J., and Caraco, N. (2005). Major role of marine vegetation on the oceanic carbon cycle. *Biogeosciences* 2, 1–8. doi: 10.5194/bg-2-1-2005

Edmonds, W. J., Hodges, R. L., and Peacock, C. D. (1990). Tidewater Virginia tidal wetland soils: a reconnaissance characterization. *Tech. rep. Virginia Agric. Experiment Station.*

Ewers Lewis, C. J., Carnell, P. E., Sanderman, J., Baldock, J. A., and Macreadie, P. I. (2018). Variability and vulnerability of coastal 'Blue carbon' stocks: A case study from southeast Australia. *Ecosystems* 21, 263–279. doi: 10.1007/s10021-017-0150-z

Ewers Lewis, C. J., Young, M. A., Ierodiaconou, D., Baldock, J. A., Hawke, B., Sanderman, J., et al. (2020). Drivers and modelling of blue carbon stock variability in sediments of southeastern Australia. *Biogeosciences* 17, 2041–2059. doi: 10.5194/bg-17-2041-2020

Howard, J., Hoyt, S., Isensee, K., Pidgeon, E., and Telszewski, M. (2014). *Coastal Blue Carbon: Methods for assessing carbon stocks and emissions factors in mangroves, tidal salt marshes, and seagrass meadows* (Arlington, Virginia, USA: Conservation International, Intergovernmental Oceanographic Commission of UNESCO, International Union for Conservation of Nature (IUCN)).

Hyman, A. C., Chiu, G. S., Fabrizio, M. C., and Lipcius, R. N. (2022). Spatiotemporal modeling of nursery habitat using Bayesian inference: Environmental drivers of juvenile blue crab abundance. *Front. Mar. Sci.* 9. doi: 10.3389/fmars.2022.834990

Lavery, P. S., Mateo, M.-Á., Serrano, O., and Rozaimi, M. (2013). Variability in the carbon storage of seagrass habitats and its implications for global estimates of blue carbon ecosystem service. *PloS One* 8, e73748. doi: 10.1371/journal.pone.0073748

Lawrence, A., Baker, T., and Lovelock, C. (2012). *Optimising and managing coastal carbon: Comparative sequestration and mitigation opportunities across australia's landscapes and land uses: Final report* TierraMar Consulting Pty Ltd., FRDC, Deakin, ACT, Australia: Prepared for the Fisheries Research and Development Corporation (FRDC).

Le, N. D. (2006). "Prediction intervals, spatial," in *Encyclopedia of environmetrics*. Eds. A. H. El-Shaarawi and W. W. Piegorsch (Chichester, UK: Wiley). doi: 10.1002/9780470057339.vap035

Longo, C. (2022). *Bayesian Spatial model development of soil core organic matter as a proxy for blue carbon stocks within the Chesapeake bay* (Williamsburg, Virginia:Undergraduate honors theses, William & Mary). Available at: https://scholarworks.wm.edu/honorstheses/1824.

López-Teloxa, L. C., Cruz-Montalvo, A., Tamariz-Flores, J. V., Pérez-Avilés, R., Torres, E., and Castelán-Vega, R. (2017). Short-temporal variation of soil organic carbon in different land use systems in the ramsar site 2027 'Presa Manuel Ávila camacho' puebla. *J. Earth Syst. Sci.* 126, 95. doi: 10.1007/s12040-017-0881-4

Macreadie, P. I., Anton, A., Raven, J. A., Beaumont, N., Connolly, R. M., Friess, D. A., et al. (2019). The future of blue carbon science. *Nat. Commun.* 10, 3998. doi: 10.1038/s41467-019-11693-w

Ma, R., Shi, J., and Zhang, C. (2018). Spatial and temporal variation of soil organic carbon in the north China plain. *Environ. Monit. Assess.* 190, 357. doi: 10.1007/s10661-018-6734-z

Mcleod, E., Chmura, G. L., Bouillon, S., Salm, R., Björk, M., Duarte, C. M., et al. (2011). A blueprint for blue carbon: toward an improved understanding of the role of vegetated coastal habitats in sequestering CO 2. *Front. Ecol. Environ.* 9, 552–560. doi: 10.1890/110004

McTigue, N., Davis, J., Rodriguez, A. B., McKee, B., Atencio, A., and Currin, C. (2019). Sea Level rise explains changing carbon accumulation rates in a salt marsh over the past two millennia. *J. Geophys. Res.: Biogeosci.* 124, 2945–2957. doi: 10.1029/2019JG005207

Mitchell, M. (2018). *Impacts of Sea level rise on tidal wetland extent and distribution* Gloucester Point, Virginia: Dissertations, theses, and masters projects, William & Mary. doi: 10.25773/v5-5s6z-v827

Mitchell, M., Herman, J., and Hershner, C. (2020). Evolution of tidal marsh distribution under accelerating Sea level rise. *Wetlands* 40, 1789–1800. doi: 10.1007/s13157-020-01387-1

Morris, J. T., and Bowden, W. B. (1986). A mechanistic, numerical model of sedimentation, mineralization, and decomposition for marsh sediments. *Soil Sci. Soc. America J.* 50, 96–105. doi: 10.2136/sssaj1986.03615995005000010019x

Stan Development Team (2022) *RStan: the R interface to Stan. R package version 2.21.5.* Available at: https://mc-stan.org/rstan.

Sullivan, M. J., and Currin, C. A. (2002). "Community structure and functional dynamics of benthic microalgae in salt marshes," in *Concepts and controversies in tidal marsh ecology* (Dordrecht: Kluwer Academic Publishers), 81–106. doi: 10.1007/0-306-47534-06Sullivan2002

Trumper, K., Bertzky, M., Dickson, B., van der Heijden, G., Jenkins, M., and Manning, P. (2009). *The natural fix? the role of ecosystems in climate mitigation: A UNEP rapid response assessment* (Cambridge, UK:United Nations Environment Programme).

van den Bergh, J., and Botzen, W. (2015). Monetary valuation of the social cost of CO2 emissions: A critical survey. *Ecol. Economics* 114, 33–46. doi: 10.1016/j.ecolecon.2015.03.015

Vehtari, A., Gelman, A., and Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Stat Computing* 27, 1413–1432. doi: 10.1007/s11222-016-9696-4

White, G., and Ghosh, S. K. (2009). A stochastic neighborhood conditional autoregressive model for spatial data. *Comput. Stat Data Anal.* 53, 3033–3046. doi: 10.1016/j.csda.2008.08.010

Wood, S. (2017). *Generalized additive models: An introduction with R. 2 edn* (New York, New York:Chapman and Hall/CRC).

Zimmerman, D. L., and Stein, M. (2010). "Classical geostatistical methods," in *Handbook of spatial statistics.* Eds. A. E. Gelfand, P. Diggle, P. Guttorp and M. Fuentes (Boca Raton, Florida:CRC Press). doi: 10.1201/9781420072884-7

frontiers | Frontiers in Marine Science

TYPE Review
PUBLISHED 19 January 2023
DOI 10.3389/fmars.2022.1034188

OPEN ACCESS

EDITED BY
Xi Xiao,
Zhejiang University, China

REVIEWED BY
Xi Zhang,
Ministry of Natural Resources, China
Armin Moghimi,
Leibniz University Hannover, Germany
Muwei Jian,
Shandong University of Finance and Economics, China

*CORRESPONDENCE
Jie Yang
yangjie2016@ouc.edu.cn

SPECIALTY SECTION
This article was submitted to Ocean Observation, a section of the journal Frontiers in Marine Science

RECEIVED 01 September 2022
ACCEPTED 19 December 2022
PUBLISHED 19 January 2023

CITATION
Chen G, Huang B, Yang J, Radenkovic M, Ge L, Cao C, Chen X, Xia L, Han G and Ma Y (2023) Deep blue artificial intelligence for knowledge discovery of the intermediate ocean.
Front. Mar. Sci. 9:1034188.
doi: 10.3389/fmars.2022.1034188

Deep blue artificial intelligence for knowledge discovery of the intermediate ocean

Ge Chen[1,2], Baoxiang Huang[2,3], Jie Yang[1,2]*, Milena Radenkovic[4], Linyao Ge[1], Chuanchuan Cao[1], Xiaoyan Chen[1], Linghui Xia[1], Guiyan Han[1] and Ying Ma[1]

[1]Frontiers Science Center for Deep Ocean Multispheres and Earth System, School of Marine Technology, Ocean University of China, Qingdao, China, [2]Laboratory for Regional Oceanography and Numerical Modeling, Laoshan Laboratory, Qingdao, China, [3]Department of Computer Science and Technology , Qingdao University, Qingdao, China, [4]School of Computer Science and Information Technology, The University of Nottingham, Nottingham, United Kingdom

Oceans at a depth ranging from ~100 to ~1000-m (defined as the intermediate water here), though poorly understood compared to the sea surface, is a critical layer of the Earth system where many important oceanographic processes take place. Advances in ocean observation and computer technology have allowed ocean science to enter the era of big data (to be precise, big data for the surface layer, small data for the bottom layer, and the intermediate layer sits in between) and greatly promoted our understanding of near-surface ocean phenomena. During the past few decades, however, the intermediate ocean is also undergoing profound changes because of global warming, the research and prediction of which are of intensive concern. Due to the lack of three-dimensional ocean theories and field observations, how to remotely sense the intermediate ocean from space becomes a very attractive but challenging scientific issue. With the rapid development of the next generation of information technology, artificial intelligence (AI) has built a new bridge from data science to marine science (called Deep Blue AI, DBAI), which acts as a powerful weapon to extend the paradigm of modern oceanography in the era of the metaverse. This review first introduces the basic prior knowledge of water movement in the ~100 m ocean and vertical stratification within the ~1000-m depths as well as the data resources provided by satellite remote sensing, field observation, and model reanalysis for DBAI. Then, three universal DBAI methodologies, namely, associative statistical, physically informed, and mathematically driven neural networks, are elucidated in the context of intermediate ocean remote sensing. Finally, the unique advantages and potentials of DBAI in data mining and knowledge discovery are demonstrated in a top-down way of "surface-to-interior" *via* several typical examples in physical and biological oceanography.

KEYWORDS

deep blue artificial intelligence, intermediate ocean, ocean remote sensing, associative statistical neural network, physically informed neural network, mathematically driven neural network

1 Introduction

Since the advent of ocean remote sensing technology, the understanding of surface or near-surface ocean phenomena has been improving. Nevertheless, the phenomena mined from satellite remote sensing data are only the tip of the iceberg because that ocean is a huge body of water that is thousands of deep, the maximum depth of which can reach 10,000 meters. Consequently, the comprehensive recognition of the deep blue ocean from the outside to the inside has always been a challenging scientific issue.

The hectometer-scale ocean contains abundant dynamic and ecological processes from turbulence to circulation, from waves to tides, and from oxygen minimum zones to subsurface chlorophyll maximum (SCM), whereas the kilometer-scale ocean is relatively "calm" and significant stratification acts as a barrier to the exchange of matter and energy between the upper and lower waters. However, slight changes in the thermohaline and circulation structures of the kilometer-scale ocean are reminders of climate change on the decadal scale. Before the turn of this century, increasingly mature oceanographic theories and the continuous accumulation of field and remote sensing data have been building the foundation of modern ocean science with the cognition of near-surface ocean phenomena and laws.

Compared with the near-surface layer, the intermediate ocean at the depth of ~100 to ~1000 meters is facing difficulties, such as the lack of corresponding theories and insufficient three-dimensional observation, which are the main factors restricting the further development of marine science (Meng and Yan, 2022). In the 21st century, the *in situ* observation and model reanalysis technologies represented by the Array for Real-time Geostrophic Oceanography (Argo) have developed rapidly, providing unprecedented high-quality data sources for intermediate ocean research.

Meanwhile, prevalent AI technology gradually developed a new branch, termed DBAI, with a unique data-driven advantage of knowledge discovery. DBAI provides the opportunity to accelerate the recognition process of the intermediate ocean, making up for the deficiency of the existing oceanographic theoretical system. In recent years, DBAI technology, in collaboration with basic ocean theory and ocean big data (Li et al., 2020), has made preliminary achievements in internal wave (IW) inversion, stratification spatio-temporal variation and prediction, eddy identification and trajectory prediction, El Nino and Southern Oscillation prediction and other ocean scientific issues Ham et al. (2019); Zhang et al. (2022). Nevertheless, it is undeniable that AI-aided remote sensing of the intermediate ocean is still in its infancy, and the challenges lie in the unclear physical mechanism, insufficient profile data, and the low generalization DBAI. The gap between remote sensing data and knowledge can be filled with the construction of DBAIwith strong knowledge discovery and physical interpretability, thus promoting major discoveries and theoretical innovations in intermediate marine science.

1.1 Water motion in ~100-m ocean

As a large-scale geophysical fluid system, the ocean is forced by celestial bodies, atmosphere, earth rotation, and the system itself, so that the upper water movement covers a very broad three-dimensional space-time scale. Generally speaking, the larger the spatial scale of water movement is, the longer the period is, and the deeper the ocean water layer is affected. As illustrated in Figure 1A.

Ocean motions can be broadly divided into turbulence, wave, and flow. On the small scale (millimeter magnitude) and high-frequency motion spectrum, the velocity (magnitude and direction) of any particle in seawater varies in a disordered manner, which is collectively referred to as turbulence. Turbulence mainly occurs at the surface of the ocean and plays an important role in the multiscale energy cascade of fluids. Its maintenance depends on the external energy supply (wave breaking is the most typical source) and follows fluid dynamics equations (such as Navier–Stokes equations).

The oceanic wave phenomena mainly include waves, tides, IWs, large-scale Rossby waves, and Kelvin waves. The effective wave height and tidal water level change are usually within 10 meters, whereas the surge can reach tens of meters in extreme weather, and the offshore tidal power can even reach the bottom. In addition, the tides have the horizontal scale of a kilometer and are potentially coupled with large-scale Rossby and Kelvin waves (sea surface height anomalies of 10 cm and vertical scales of thousands of meters). Ocean current is defined as a large-scale, relatively stable flow of seawater driven by wind stress and density differences. The core depth of ocean circulation is usually several hundred meters, and the deep compensation undercurrent can reach the depth of kilometers. It is crucial in the composition of the global ocean environment and climate regulation.

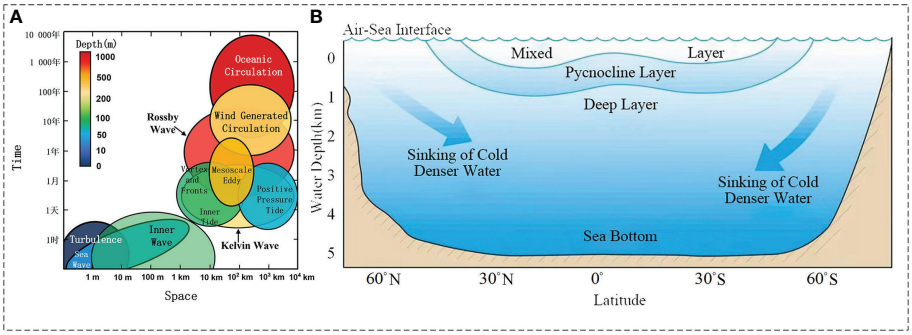


FIGURE 1
Water motion and vertical stratification of the intermediate ocean. **(A)** The time scale, spatial scale and depth of various types of motions. **(B)** Vertical stratification.

1.2 Vertical stratification of the ~1000-m ocean

In addition to the flow and fluctuation in the horizontal direction, the ocean also has typical stratification characteristics in the vertical direction. According to the range of air–sea interaction and the vertical gradient of thermohaline parameters, the ocean water can be divided into the air–sea interface, mixed layer, thermocline, and deep layer from top to bottom (Sprintall and Cronin, 2009) as illustrated in Figure 1B.

Among them, the air–sea interface is related to the momentum, heat, and gas exchange rate between the ocean and the atmosphere, which has become the key problem of climate change and its prediction. Turbulent mixing caused by dynamic effects (wind, wave, current, etc.) and convective mixing caused by thermal effects (evaporation, cooling, densification, etc.) at the air–sea interface makes the upper ocean form a water layer with nearly uniform temperature, salinity, and density field, that is, the mixed layer, which contributes to the heat balance, water cycle, and carbon cycle of the earth system. Meanwhile, climate and large-scale atmospheric circulation changes have direct and important effects. At the lower boundary of the mixing layer, the water layer with rapid changes in temperature, salinity, and density is called the marine thermocline, which is divided into the seasonal and main thermocline. The existence of the thermocline hinders the exchange of heat, oxygen, carbon, and nutrients between the upper and lower water bodies, and its interannual and interdecadal changes are reflected in the climate system and the Marine ecosystem. Below the thermocline is uniform deep water with basically stable temperature, salinity, and density characteristics.

From the perspective of time scale, the air–sea interaction at the interface takes place at any time, the mixed layer and the seasonal thermocline have synoptic and seasonal scales, and the main thermocline and deep water change slowly on decadal and climatic scales. From the perspective of spatial scale, the mixed layer is usually fewer than 100 meters in the low latitude sea area and can be as deep as the main thermocline in the middle latitude. The main thermocline is about 300 m in the low latitude and 700~800 m in the middle latitude. At the high latitude, the mixed layer and the main thermocline gradually rise until the stratification of subpolar water disappears. In general, the air–sea interface and mixing layer are the most active areas of ocean mixing, which are directly affected by the weather system. The stratification structure and thermohaline distribution of the thermohaline circulation are controlled by the seawater subsidence and have climatic-scale characteristics. Therefore, the distribution of temperature, salinity, and density fields in the vertical direction in the low-latitude sea area reflects the radial spatial distribution characteristics of the ocean surface to a certain extent (Trujillo and Thurman, 2011).

1.3 DBAI-aided remote sensing of the intermediate ocean

As an interdisciplinary science based on observation and experiment, marine science, which is characterized by data-intensive and technology integration and linked by water bodies, has experienced a development process from theoretical traction to technology-driven and then to data-driven. Specifically, the model-driven method is to achieve theoretical analytical solutions by modeling problems, just as Western medicine treats patients based on the diagnosis. Data-driven methods search for the approximate optimal solutions by fitting the model to the data, just like "look, hear, question, and feel the pulse" in traditional Chinese medicine, which is guided by traditional Chinese philosophy to carry out "characteristic engineering" with dialectical unity. Western medicine takes Western philosophy as a "mathematical model" to solve the main contradiction quickly and efficiently. "AI for science" is listed as an important trend, indicating that "AI can become a new production tool for scientists and promote the new paradigm change in scientific research" (Appenzeller, 2021). As a cutting-edge technology in the integration of marine and data science, DBAI has the function of "model + data" driving and complementing each other in the way of "integration of traditional Chinese and Western medicine."

This review paper puts forward the viewpoint of philosophy and scientific conception of DBAI for the first time as shown in Figure 2. Although satellite ocean remote sensing is the main approach to observing surface and subsurface ocean phenomena from the "external" perspective, it cannot directly perceive the "internal" process of the intermediate ocean. Benefiting from the *in situ* observation represented by Argo buoys, the vertical profiles compensate for the existing deficiencies of satellite ocean remote sensing with high accuracy. However, it cannot achieve extensive coverage at the same time. Therefore, it is crucial to combine the remote sensing data with the *in situ* data; DBAI technology is highly promising for playing such a role.

The remainder of this review paper proceeds as follows. Section 2 describes the ocean observations regarding the satellite remote sensing in the ~100-m ocean and *in situ* observations in the ~1000-m ocean. In section 3, the DBAI methodology is introduced in detail

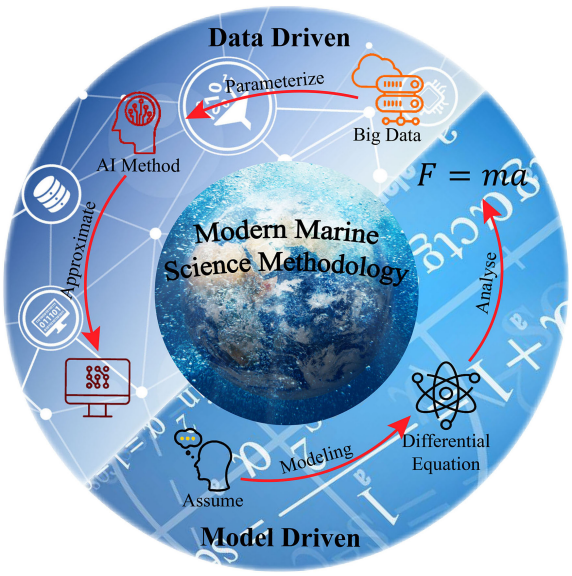


FIGURE 2
Philosophy and scientific conception of DBAI, model-driven and data-driven methods supplement each other, just as Western and Chinese medicine complement each other.

and contains associative statistical, physically informed, and mathematically driven neural networks. Section 4 demonstrates the knowledge discovery with DBAI. Finally, some remarks conclude in Section 5.

2 Observations

Observations are highly significant in oceanographic progress, and DBAI is increasingly being considered as a tool to exchange ocean knowledge and observational data for the intermediate ocean as well as bring together model- and data-driven methods, facilitating conventional tasks better and faster. The observations of the intermediate ocean can be divided into satellite remote sensing in the ~100-m ocean and *in situ* observations in the ~1000-m ocean.

2.1 Satellite remote sensing in the ~100-m ocean

Oceanic remote sensing, as one of the key technologies to promote the development of marine science, has acquired ocean color, dynamics, and environmental parameters with its comprehensive advantages of all-weather, quasi-real-time, large range, high precision, and long-time sequence over the past 50 years. It has fundamentally enhanced our profound understanding of near-surface ocean phenomena and processes such as ocean primary productivity, sea level change, air–sea flux, and ocean circulation and improved the traditional paradigm of global ocean observation and scientific research.

However, current satellite ocean remote sensing is essentially a two-dimensional sea surface remote sensing. As an active optical remote sensing method, light detection and ranging (LiDAR) has the combined technical advantages of all-day range resolution and large penetration depth (more than three times that of passive remote sensing). Moreover, LiDAR can achieve remote sensing at night, in polar regions, and even in cloudy conditions. At present, it is the only known detection method that is expected to realize underwater three-dimensional remote sensing and is also the international frontier in the field of ocean optics and ocean water color remote sensing (Hostetler et al., 2018). Furthermore, LiDAR is widely applied in marine bio-optics, shallow sea topography, fish detection, polar ecology, and upper ocean dynamic processes (IWs, mixed layers, turbulence, foam, etc.). In other words, a new interdisciplinary technology field that is marine LiDAR detection technology has been gradually constructed (Churnside, 2013).

Although the spaceborne marine LiDAR is still in the blank stage, the successfully launched spaceborne atmospheric LiDAR (CALIOP) and terrestrial LiDAR (ICESat-2) have shown good advantages and potential in the preliminary application of ecological oceanography (Behrenfeld et al., 2013; Behrenfeld et al., 2017; Behrenfeld et al., 2019; Lu et al., 2020). With the successive planning and implementation of the NASA PACE mission (Werdell et al., 2019), ESA MESCAL mission (Chepfer et al., 2018), and Guanlan mission (Chen et al., 2019) of China Ocean National Laboratory Qingdao, Spaceborne marine LiDAR with multiband (blue-green wavelength), meter-level resolution and multisystem (hyperspectral, fluorescence, polarization, etc.) observation capability will become a reality. The concept design of observation at ~100 m depth can be seen in Figure 3. Further combined with passive ocean color remote sensing and *in situ* biogeochemical buoys (Biogeochemical Argo, BGC-ARGO) and other observation means, it is expected to achieve the three-dimensional detection and high-precision inversion of bio-optical and physical parameters in four-dimensional space time within the ~100-m depth of the global ocean for the first time (Chen et al., 2021b). For a more intuitive understanding of the abovementioned spaceborne sensor, the characteristics of these satellite sensors are collected and summarized in Table 1 (Amani et al., 2021; Amani et al., 2022a; Amani et al., 2022b; Amani et al., 2022c).

2.2 *In situ* observations in the ~1000-m ocean

The ~1000-m ocean is usually studied with the help of *in situ* observations and reanalysis data. Before the 21st century, *in situ* profile observation typically comprised ship-based and mooring buoy array but was challenged in terms of the low coverage rate of spatiotemporal data, large deviation, and high observation cost. In the 1990s, with the intensification of global climate change and the prominent role of the ocean, the systematic absence of global ocean profile observation data posed great challenges to climate change research (Johnson et al., 2022).

In this context, the Argo program, which aims at the real-time acquisition of global ocean thermohaline data at in the upper 2000 m, was implemented. Up to now, it has provided more than 95% of global thermohaline profile data (Riser et al., 2016), bringing profound changes to intermediate marine scientific research. In the

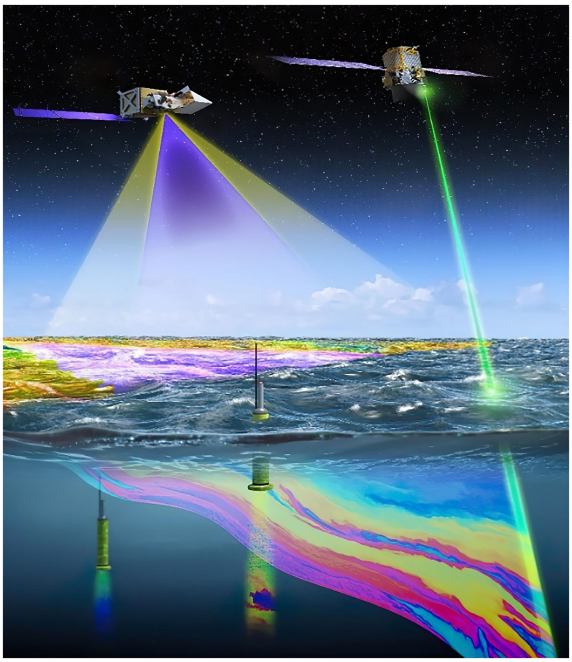


FIGURE 3
Concept design of the integration of active and passive remote sensing and field joint observation at ~100-m depth.

TABLE 1 The characteristics of described spaceborne satellite and mission.

Satellite/Sensor	Temporal Resolution	Spatial Resolution	Wavelength(nm)	Time Period
CALIOP	98.5 minutes	333 m	532, 1064	2006-ongoing
ICESat-2	91 days	0.7 m	1064	2018-ongoing
PACE	1-2 days (OCI)	1 km	340-890 (5steps); 940, 1038, 1250, 1378, 1615, 2130, 2260	Plan
	2 days (HARP-2)	3 km	440, 550, 670, 870	
	-30 days (SPEXone)	2.5 km	385-770(2-4 steps)	
Guanlan	8.45 days	500 m	486, 532	Plan
	5.61 days			

early 21st century, mobile observation and sensor technologies were further developed (Morel and Berthon, 1989). Physical, biological, and chemical sensors with miniature and low power consumption were successfully developed, and BGC-Argo buoys proceeded, providing a broader data basis for exploring the coupling of ocean physics and biogeochemistry. Around 2012, the United States, France, and other countries successively joined the Deep-sea Argo (Deep-Argo) program, planning to build a deep-sea Argo observation network measuring depths up to 6000 m to promote the Argo program to enter the era of all-sea, all-depth, and multidisciplinary ocean observation. Figure 4 illustrates the Labrador Sea temperature profile based on Argo measurement.

Additionally, with the rapid development of ocean observation technology and the significant improvement of computer performance, the development and application of high spatiotemporal resolution, full-sea depth, numerical simulation products based on the earth fluid dynamics theory and the assimilation of multisource observation data have significantly improved the accuracy of model prediction results. Ocean model and climate state observation data products, such as the hybrid coordinate ocean model (HYCOM), global ocean circulation model data set, and world ocean atlas (WOA) global Ocean grid climate state data set, have emerged. Including temperature, salinity, flow field, chlorophyll, inorganic salt, and other multilayer parameter information related to marine thermal, marine dynamics and biogeochemistry, and other fields, provides another foundational data source for intermediate and even deep ocean scientific research. In recent years, *in situ* imagery acquired by underwater imagers has been major ancillary data in marine research. However, different from the remote sensing images, the complicated situations of the marine environment highlight the uneven distribution of data quality in *in situ* images (Jian et al., 2021). *In situ* images enhanced, denoised, or saliently extracted by specific algorithms provide a reliable data source for full ocean-depth scientific research (Jian et al., 2018; Jian et al., 2019).

3 DBAI methodology

DBAI methodology mines the essence of phenomena and reveals the laws behind them, starting from the data dimension and taking scientific calculation as the core. As the key component, the neural network (Cozman, 2021) develops from shallow to deep (Lecun et al., 1998; Ronneberger et al., 2015) and has undergone three stages: associative statistical, physically informed, and mathematically driven. The associative statistical neural networks can extract the information of data space, mine the explicit laws contained in the data, and realize basic knowledge discovery. Physically informed

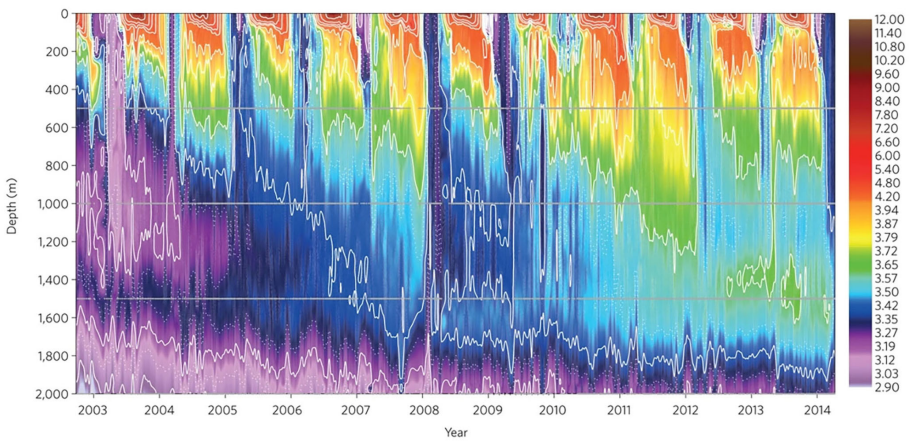


FIGURE 4
Labrador Sea temperature profile obtained from *in situ* Argo observations.

neural networks (PINNs) (Wright et al., 2022) embed disciplined prior knowledge or theory into the network, extract potential knowledge rules from data, improve operation efficiency, and prevent "low-level fallacy." By building a bridge between the traceable mathematical model and scientific knowledge, the mathematically driven neural networks give full play to the excellent performance of the deep neural network in chaotic or extremely complex mass data and achieve valuable knowledge, accelerating scientific discovery and enabling interpretable results. The overall architecture of DBAI is illustrated in Figure 5. The data obtained by satellite remote sensing and *in situ* observations allow for surface-to-interior profiling of ocean phenomena. AI technology bridges the semantic gap between the two data sources, efficiently coupling remote sensing data, which has the external advantage, with *in situ* data, which has an internal character. This coupled paradigm makes DBAI technology, which incorporates mathematically driven, physically informed, and associative statistical, a new engine for knowledge discovery in marine science.

3.1 Associative statistical neural network

The basic method or function of the DBAI methodology is to perform correlation statistics in the data ocean, achieving initial and explicit information mining and knowledge discovery. Data is the fuel of the AI method; the construction methods of the neural network vary with different data features. For example, the image data and rasterized products in remote sensing are general data types, which can be obtained with special equipment. With the development of computer vision (Simonyan and Zisserman, 2015; Huang et al., 2017; Chen et al., 2018), neural networks can mine the local features of objects in these data. Thus, most of the explorative DBAI research is based on Euclidean data to construct convolutional neural networks (He et al., 2017) and then achieves the correlation statistics of data information. In terms of DBAI architecture, the basic building blocks are convolution neural network, recurrent neural network and its extended network, such as the LSTM unit, as shown in Figure 6. The convolution neural network, which is depicted in Figure 6A, provides a new mathematics tool for researching grid data (Goodfellow et al., 2014; He et al., 2016). Furthermore, with continuous iterative updating, the recurrent neural networks, the structures of which are shown in Figure 6B, C, can explore the timeseries features and context information from long-time series data.

As for sensory data, the sea surface temperature, sea surface height, and sea level anomaly (SLA) have time series features that can extract the parameters of the feature for change of temperature and height with time lapse adopting the recurrent neural networks. Then, the model summarizes the overall pattern of the data to predict the future trend. Therefore, associative statistical neural networks are becoming one of the basic methodologies for researchers to detect and predict oceanic phenomena and grasp their patterns. Table 2 displays some research for constructing the neural network based on Euclidean data in marine science from 2018 to now. The application direction and publication time illustrate that DBAI is the environment of mass innovation that can continue to develop.

As previously identified, the data is the beginning of the associative statistical neural networks, which are based on the classification of the data and characteristics of the task to design the network structure. However, in scientific research, it is key to incorporate physical theory into neural networks to enable even higher accuracy.

3.2 Physically informed neural network

To exploit the efficiency of the neural network, the PINN is proposed by Raissi et al. (2019). PINNs follow the specific objective laws of physics described by nonlinear differential equations, which is

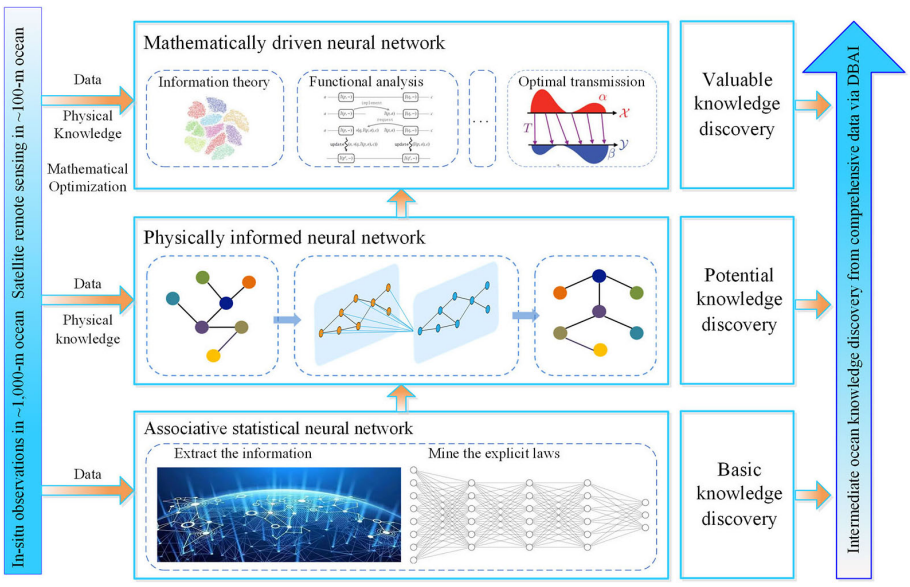


FIGURE 5
Overall architecture of DBAI. The core of DBAI is the associative statistical, physically informed, mathematically driven neural networks.

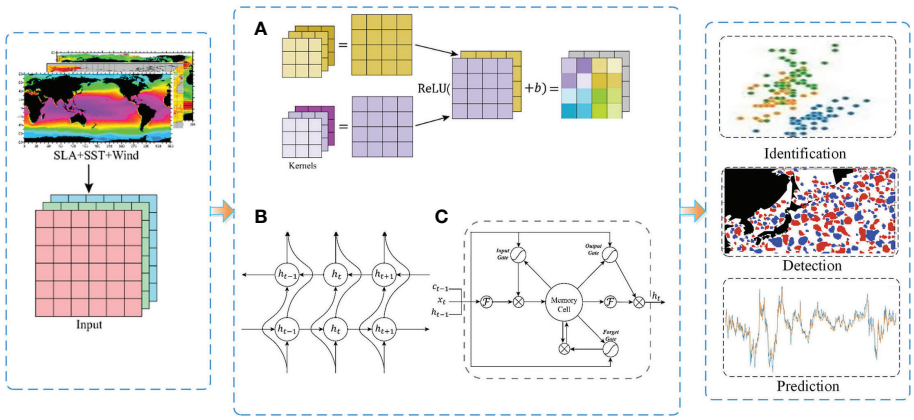


FIGURE 6
Associative statistical neural network schematic; the basic building blocks include **(A)** convolutional neural network, **(B)** recurrent neural network feedforward process, and **(C)** LSTM structure.

a deep learning method that can solve scientific problems (Nakajima et al., 2021). PINNs are a pioneering technology that provides PDE with a new numerical solver. According to the guide method, it can be divided into observation, induction, and learning bias (Karniadakis et al., 2021), which is shown in Figure 7.

The observation bias is designed for the data input stage because the data can reflect the underlying physics. For example, with the universal gravitation $F = \frac{GMm}{r^2}$, Burges function $\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = v\frac{\partial^2 u}{\partial x^2}$, and Hamiton systems $\frac{\partial H}{\partial p} = \frac{\partial p}{\partial t}, \frac{\partial H}{\partial q} = -\frac{\partial q}{\partial t}$, neural networks can learn the physical structure, functions, vector fields, etc., from these physical principles, achieving being physically informed.

Furthermore, in the model-learning process, we expect that some specific features can be prioritized. Thus, the induction bias is introduced in the neural network, which utilizes the universal rules in the physical world to constrain the model. Normally, in the construction process of the network, the induction bias is designed for a neural operator that makes the computational process follow the physical law, for example, the translational and spatial invariance in CNNs and graph embedding computation in graph neural networks.

The learning bias is different from the above guided method; it focuses on the backward neural network. In other words, a dedicated loss function is designed for the model to calculate the error. Suppose that a problem satisfies a constraint, and the output of the network can transform into a mapping of solutions. For instance, the hamiltonian system (Greydanus et al., 2019) description formula transforms into a form of residuals, such as (1).

$$\frac{\partial H}{\partial p} - \frac{\partial p}{\partial t} = 0, \frac{\partial H}{\partial q} + \frac{\partial q}{\partial t} = 0 \tag{1}$$

In this function, p denotes coordinate and q is momentum. According to the residual function, the loss function can be designed as Equation (2).

$$L = \left|\frac{\partial H}{\partial p} - \frac{\partial p}{\partial t}\right|_2 + \left|\frac{\partial H}{\partial q} + \frac{\partial q}{\partial t}\right|_2 \tag{2}$$

The learning bias prompts the trained neural network to conform to the theory that is expressed by differential or partial differential equations. That improves the confidence and prediction accuracy of neural networks significantly. Table 3 enumerates part of the research that PINNs applies in marine science and lists the bias method.

Although some researchers have attempted to predict the real ocean current field based on the isotropic properties of Navier–Stokes equations, physical information is still limited to observational traction using model data or inductive traction using existing network structures. On the contrary, neural networks are

TABLE 2 Partially associative statistical neural networks for mid-level ocean remote sensing.

Network	Network Structure	Correlation Research
	CNNs	ENSO prediction (Ham et al., 2019),
		Eddy heat flux prediction (George et al., 2021),
Associative statistical neural networks		Arctic sea ice seasonality prediction (Andersson et al., 2021),
		Ocean IW amplitude search Zhang et al. (2022),
		Global mesoscale eddy identification (Chen et al., 2021c)
	RNNs	Sea surface temperature prediction (Xie et al., 2020),
		Sea level forecast (Accarino et al., 2021),
		Subsurface temperature field prediction (Su et al., 2021)

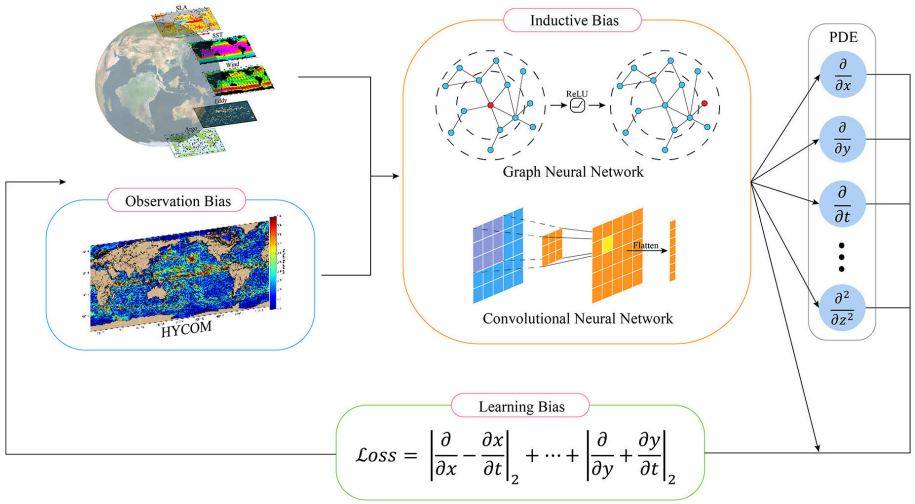


FIGURE 7
Prior knowledge-embedding mechanism for PINNs.

significantly integrated with domain knowledge in basic science, which adopted learning bias to improve the applicability of the model to physical data (Cranmer et al., 2020; Xiong et al., 2021).

The observation, induction, and learning bias can be embedded as physically informed methods for different stages of neural networks, which can organically integrate domain knowledge and neural networks. Compared with the traditional numerical differential equations, the PINNs have excellent robustness in handling noise entrainment within the data. In addition to this, theoretical foundations, such as quasigeostrophic, Rossby wave, Ekman drift, and eddy theory, transform into the learning bias that is an important guarantee for advancing scientific knowledge discovery. In summary, the potential of neural networks in marine science has not yet been fully exploited. In the future, the structure of the neural network should be based on the basic theory of marine science and further advance toward the "AI for ocean science."

3.3 Mathematically driven neural network

As a data-driven method, neural networks can be regarded as a "black box," lacking mathematical interpretability, which is the foundation of building bridges from data to knowledge. It is difficult to fully transform the neural network from a black box to a "white one." However, neural networks can be improved based on the mathematical inference that makes partial network interpretability that is named mathematically driven neural networks. For example, the input and output of a neuron is a linear process that formulates as Equation (3).

$$L = \sum_{i_0=0}^{d+1} (\boldsymbol{a}_{i_1 i_0}^0 \boldsymbol{x}_{i_0}) \tag{3}$$

where $\boldsymbol{a}$, $\boldsymbol{x}$, d denote weight, input, and network width, respectively. Theoretically, the depth and width of a neural network can reach infinite dimensions and can be mathematically represented as an infinitely wide function space, constituting a mathematical "Barak space" (Wojtowytsch, 2020) as in Equation (4).

$$F_\infty : = \left\{ \sum_{i_L=1}^{\infty} a_{i_L}^L \sigma \left(\sum_{i_{L-1}}^{\infty} a_{i_L i_{L-1}}^L \sigma \left(\sum_{i_{L-2}} L \sigma \left(\left(\sum_{i_1=1}^{\infty} a_{i_2 i_1}^1 \sigma \left(\sum_{i_0=0}^{d+1} \left(a_{i_1 i_0}^0 x_{i_0} \right) \right) \right) \right) \right) \right) \right\} \tag{4}$$

where $a_{ij}^l = 0$ for all but finitely many i, j, l. L is the layer number of a neural network, i_L is the neuron of every layer, and σ denotes the activation function. Furthermore, according to the geometrical point, high-dimensional data of the same category in nature is concentrated near some low-dimensional manifold. Therefore, the input of neural networks to the output can be understood as a mapping of differential geometric manifolds (Lei, 2020).

Based on the PINNs, the mathematically driven neural networks provide neural networks with a mathematical explanation. However, its theory, architecture, and application research are in the

TABLE 3 Partial PINNs for marine science.

Network	Bias Method	Correlation Research
Physically informed		Ocean subsurface temperature prediction (Meng et al., 2021),
neural networks	Observation bias	Ocean turbulence prediction (Wang et al., 2021)
		Seasonal Arctic sea ice projections (Andersson et al., 2021),
	Induction bias	Identification of global abnormal mesoscale eddies (Liu et al., 2021),
		Global mesoscale eddy identification (Huang et al., 2022)

preliminary exploration stage. The mathematical theory not only bestowed the interpretability of networks, but also optimized computational efficiency and memory usage. Meanwhile, the properties of neural networks are exploited and compensate for the poor generalization ability of theoretical models to real data. In marine science, hydromechanics is the basic theory of how to introduce it into neural networks, constructing a novel marine physics or mathematical neural network structure as a frontier direction of future development. Moreover, it is able to exploit the excellent performance of deep neural networks in highly complex systems, accelerating knowledge discovery in marine science in massive data.

4 From DBAI to knowledge discovery

The surface data of the ocean obtained by satellite remote sensing is abundant, whereas the profile data of the intermediate ocean is relatively sparse. At present, the critical task of DBAI is to establish the intrinsic connection between surface phenomena and intermediate processes, especially the physical and ecological processes below the surface. With impressive results, the data-driven DBAI methodology has recently made a preliminary exploration of "surface to interior" in ocean science. However, ocean knowledge discovery methods mainly focus on associative statistical neural networks and begin to explode to the PINNs, the research of mathematically driven neural networks is still in exploration. In this review paper, highly challenging typical applications are selected from four aspects of marine morphology, kinematics, dynamics, and ecology to demonstrate intermediate ocean research supported by DBAI as shown in Figure 8.

4.1 3-D identification and trajectory prediction of oceanic eddy

Oceanic mesoscale eddies, which follow the quasi-geostrophic potential vorticity conservation equation, are rotating movements of seawater on scales smaller than Rossby waves. Oceanic mesoscale eddies are the "weather" of the ocean with horizontal scales ranging from several to hundreds of kilometers, vertical scales ranging from tens to thousands of meters, and temporal scales from days to years. Ocean eddies, with their massive quantity, broad distribution, strong entrainment, and high energy, are becoming the ideal proxy for the substance cycling, energy cascade, and multi-sphere coupling in ocean research.

According to their rotation direction, oceanic eddies can be divided into cyclonic eddies (CE) and anticyclonic eddies (AE). The dynamic process of the water body inside the eddy is mapped to the sea surface, resulting in an undulating change in sea surface height of about 1~100 cm, which causes the topological relationship of the sea surface around the eddy to present a relatively stable closed structure. This phenomenon makes the formation of the local maximum (or minimum) value of sea surface height in the center of the AE (CE), which is the fundamental criteria of the application to detect ocean eddies with remote sensing technology (Pegliasco et al., 2022) as shown in Figure 9.

By combining the eddy signal acquired by satellite altimeter with *in situ* observations data such as Argo, the "surface-to-interior" vertical structure and internal dynamic process of the oceanic eddy are gradually being understood (Chaigneau et al., 2011; Zhang et al., 2014). However, the horizontal observation resolution and vertical profile detection capability of existing satellite altimeters are still inadequate. Hence, realizing high-resolution 3-D remote sensing observation of eddies is one of marine science's urgent requirements for the new ocean satellite technology.

4.1.1 Independent eddy identification with Argo profiles

Eddy identification is critical in advancing theoretical knowledge and scientific research on ocean eddies. The current mainstream method of eddy identification is the closed contours method, which is limited by the sampling capacity of the satellite altimeter, resulting in approximately 90% of oceanic eddies being missed (Amores et al., 2018). In addition, this method has limitations for submesoscale eddies with the characteristics of having smaller scale, weaker intensity, and deeper below the sea surface.

To solve the above problem, an idea of ocean eddy identification based on 3-D structure is proposed by using profiling Argo (Chen et al., 2021a). The surface features of oceanic eddies are correlated and modeled with their vertical structures to construct an Argo floats-based algorithm for independent eddy identification. The altimeter-identified eddies are further aligned with Argo profiles to build a multisource eddy data set. The associative statistical neural network incorporating the observation and induction bias is designed to extract abstract features of the eddy vertical structure, thus realizing high-precision eddy identification from a 3-D perspective (Chen et al.,

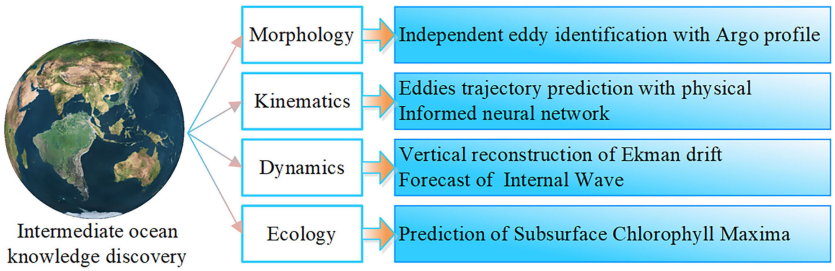


FIGURE 8
Typical applications in physical and biological oceanography with a top-down way of "surface to interior."

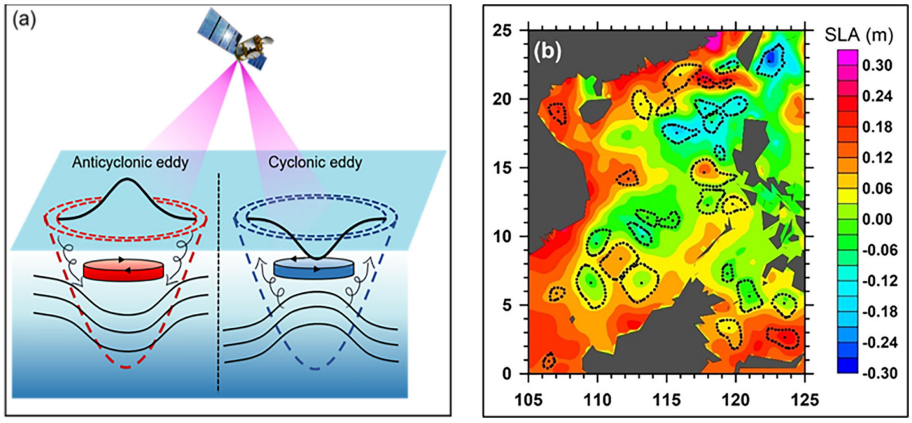


FIGURE 9
Oceanic detection and identification from the altimeter. **(A)** The fundamental criteria of the application to detect an ocean eddy, and **(B)** detection of eddies and identification results.

2021c; Huang et al., 2022) as illustrated in Figure 10A Compared with the traditional mathematical statistics methods, the AI eddy identification algorithm based on Argo 3-D structure not only improves the computational efficiency by more than 10 times, but also achieves 98% accuracy of eddy identification.

The eddy vertical identification algorithm can identify about 36% of the missed eddies. This method can prompt additional consideration of the eddy identification accuracy from satellite altimeter. Figure 10B shows the comparison results of eddy identification using altimeter SLA and the Argo vertical profiles, respectively. The experimental results show that more than 50% of the eddies missed by the altimeter are distributed in the equatorial region with a low satellite sampling rate, weak geostrophic, and strong non-geostrophic. In contrast, the Argo vertical eddy identification algorithm has a unique advantage in non-geostrophic dominated submesoscale eddy identification, which is highly complementary to the altimeter eddy identification method. In addition, the Argo vertical signal of eddies captured by the AI eddy identification algorithm is comparable to or even stronger than the altimeter method. The AI technology demonstrates the feasibility and credibility of eddy identification with 3-D ocean profiles while innovating identification methods and improving identification efficiency. This is essential for promoting the development of eddy identification methods and studying ocean eddies' kinematics and dynamics.

4.1.2 Eddies trajectory prediction with PINN

Eddy motion is primarily prevalent in complex ocean environments (current, wind, and topographic) and is controlled by various physical mechanisms, including potential vorticity and baroclinic instability (Chen et al., 2022a). Naturally, the accuracy of eddy trajectory prediction can be fundamentally improved with the physical constraint of prior knowledge. The eddy trajectory prediction network compliant with physical constraint is proposed as shown in Figure 11A The theoretical approximate phase speed of the first baroclinic mode Rossby waves (C_p) is embedded as the theoretical velocity of the vortex into the machine learning model for trajectory prediction, thus enabling accurate prediction of the eddy trajectory for the next 10 days.

$$C_p = -\beta \times R_d^2 \tag{5}$$

where $\beta = 2\Omega\cos\theta R^{-1}$ is the Rossby parameter, $\omega = 7.29\times10^5$ is an Earth rotation rate, and $R = 6371.39$ km is the radius of Earth. R_d is the Rossby radius of deformation, which is calculated (Chelton et al., 1998) by Equation (6):

$$R_d = \frac{1}{|f(\theta)|\pi}\int_{-H}^{0} N(z)dz(\theta > 5^{\circ}) \tag{6}$$

where $f = 2\Omega\theta$ is the Coriolis parameter, $N(z)=(-(g/\rho)\times\partial\rho(z)/\partial z)^{1/2}$ is the buoyancy frequency, ρ is the potential density, and g is the acceleration due to gravity. The midterm prediction results for more than two million single-track eddies worldwide show that the accuracy of eddy trajectory prediction is improved by about 24% with the embedding of the physical mechanism, and the prediction error is significantly lower than other prediction algorithms (Li et al., 2019; Wang et al., 2022). Figure 11B illustrates the true and predicted trajectories of two typical long-lived eddies. The results visually demonstrate the high consistency between the predicted trajectories and the true trajectories, reflecting the significant advantages of the PINN for eddy trajectory prediction.

4.2 Vertical reconstruction of Ekman drift

The wind is the dominant driving mechanism of ocean circulation. When the wind with constant speed abidingly acts on the vast sea expanse, a steady seawater movement is generated, called drift. Ekman has constructed a precise theoretical drift solution, namely, the Ekman drift theory. Specifically, the drift is the result of the balance between the frictional force generated by the plumb turbulence and the Coriolis force. The latitudinal velocity component (u) and the meridional velocity component (v) of the drift are affected by the depth of seawater (z), and their quantified expressions are shown in Equations (7) and (8).

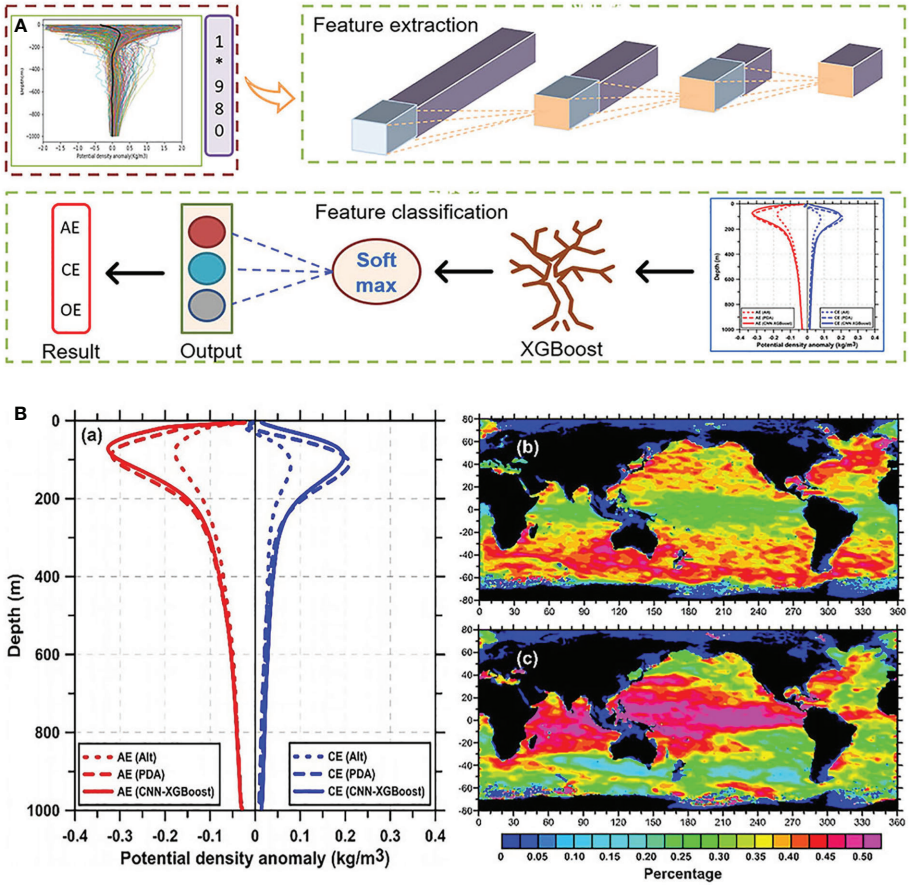


FIGURE 10
(A) The associative statistical neural network for eddy identification with Argo profiles calibrated by the altimeter. **(B)** Comparison of eddy identification results between altimeter and Argo. Independent eddy identification with Argo profiles *via* associative statistical neural network. (10a) The architecture of convolutional neural network with extreme gradient boosting (CNN-XGBoost). (10b) Identification results. (a) Vertical profiles of eddies identified by altimeter (short dashed line) and only Argo (solid and long dashed lines), respectively, (b) Global geographical distribution of identified eddies by altimeter, and (c) global geographical distribution of independently identified eddies by Argo profiles.

$$u = \frac{\tau_y}{\sqrt{2}aA_z}\exp\ (az)\cos\ 45^\circ + az, \tag{7}$$

$$v = \frac{\tau_y}{\sqrt{2}aA_z}\exp\ (az)\sin\ 45^\circ + az \tag{8}$$

Ekman found that the surface current direction is approximately 45° to the right (left) of the wind direction in the Northern (Southern) hemisphere with ideal assumptions. Due to the viscosity of seawater, the flow vector deflects with depth and forms an Ekman spiral as shown in Figure 12A. The simple assessment of global wind-driven drift can be calculated using satellite scatterometer data on sea surface wind fields as illustrated in Figure 12B. However, in the natural ocean, the turbulent viscosity coefficient A_z varies with depth, and the latitudinal wind stress τ_x effect exists simultaneously, coupled with the fact that the wind-driven drift decays faster with depth than the theoretical prediction. These reasons cause the actual profile of wind-generated drift to be different from the theoretical Ekman spiral curve, and its shape is closely related to the sea state of the local ocean. In addition, due to the lack of observational information on the seawater turbulent viscosity coefficient, the two-parameter regression model is used as the mainstream method to reconstruct the wind current field (Lagerloef et al., 1999). The core algorithm of this model is to evaluate wind current using the observations of the *in situ* drifting buoy in the field without the ground-transfer current component. However, the limitation is that the model accuracy is affected by empiricism or climatological parameters. AI technology, as a data-driven method, can effectively avoid the empiricism of the model and have some stability to the climatological parameters in the reconstruction process. Therefore, with the introduction of the DBAI methodology, the accurate quantification of wind-driven current fields is expected to expand novel ideas and methods.

4.3 Forecast of IWs

IWs are fluctuations that oscillate at the surface of two different media within the ocean. Two necessary conditions for generating IWs are seawater stratification and a disturbance source. The wavelength of IWs generally ranges from hundreds of meters to tens of kilometers, period ranges from minutes to hours, and amplitude ranges from several meters to tens of meters. IWs mainly occur in the pycnocline and are most active at 50–800 m of the ocean. In addition, the undulating propagation of IWs and its "dead water effect" affect

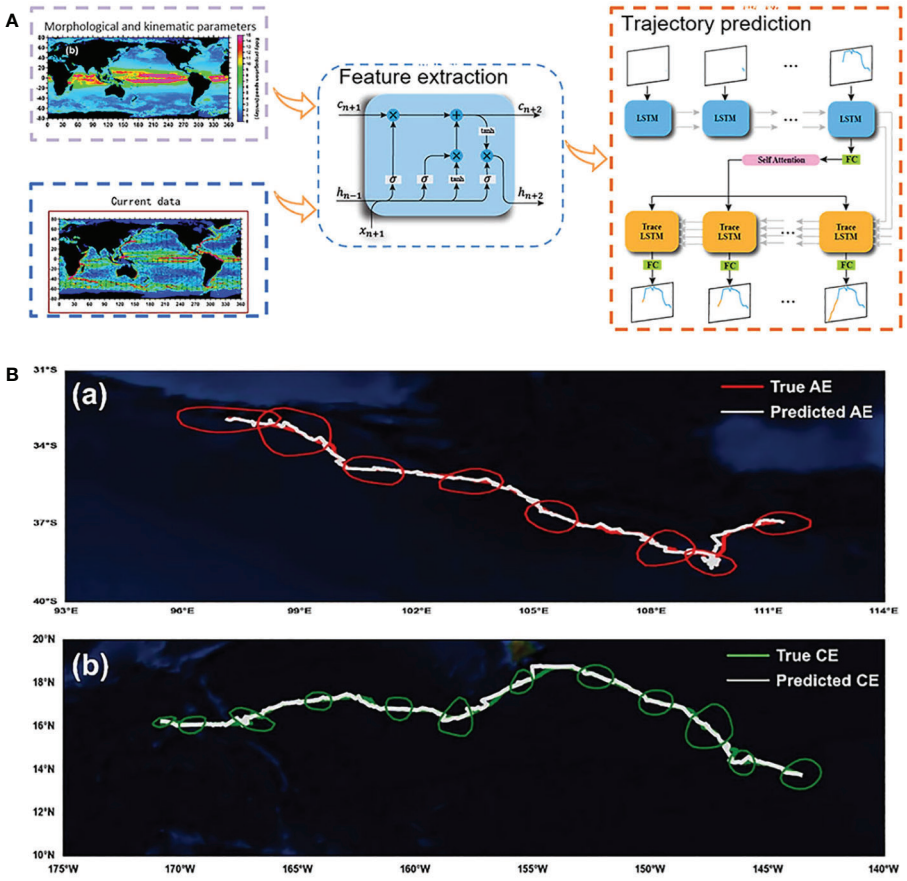


FIGURE 11
(A) The physically informed neural network for eddy trajectory prediction. **(B)** Long-lived eddies trajectory prediction.Eddy trajectory prediction via associative neural network. **(A)** The architecture of the eddy trajectory prediction network compliant with physical constraints. **(B)** Results of long-lived eddies trajectory prediction. (a) AE trajectory prediction, (b) CE trajectory prediction.

local ocean productivity and ocean sound field and even cause great security risks to submarine navigation and ocean engineering.

The synthetic aperture radar (SAR) has become one of the essential remote-sensing tools for monitoring ocean IWs with its observation advantages of large range, all day, all weather, and high resolution. The quantitative expression of the modulation of SAR images by IWs can be derived based on the Korteweg–de Vries (KdV) Equation (9) (Zheng et al., 2001).

$$\delta_{0|IW} = Q\left(\frac{C\eta_0 cos^2\varphi}{h_{md}\Delta\varpi}\right) tgh_{wd}\left[\frac{(x-Vt)}{\Delta}\right] sec\left(h_{wd}^2\right)\left[\frac{(x-Vt)}{\Delta}\right] \quad (9)$$

where $\delta_{0|IW}$ is the backscattering coefficient of the IWs on the SAR image; Q can be treated approximately as a constant, usually determined by the SAR sensor frequency, the angle of incidence, and the dielectric constant of the IWs. C is the linear phase speed; φ is the wave direction; h_{md} is the thickness of the mixing layer; Δ is the characteristic half width of the IWs; ϖ is the mean value of angular frequency of ocean surface waves; g, h_{wd}, x, V, t represent the acceleration due to gravity, seawater depth, spatial position, propagation speed, and propagation time of the IWs, respectively. $sec(g)$is secant function. Equation (9) is also known as the analytical expression of the oceanic IWs presented on the SAR image. Therefore, the spatial position, wavelength, wave direction, and other horizontal parameter information of the IWs can be directly obtained from SAR images. The transient wave speed, depth, amplitude, and other vertical parameters of IWs can then be further inverted, together with the simultaneously measured CTD and historical data to depict the spatial and temporal characteristics of IWs properly as shown in Figure 13.

In recent years, DBAI has been applied successfully to the inversion and prediction of oceanic IWs from satellite images. An associative statistical neural network structure is proposed to address the randomness issue with IWs propagation (Zhang et al., 2021). This network mines the correlation between multiple information of IWs and determines the propagation position of IWs by predicting the propagation speed and direction. Comparisons with the conventional physical model (KdV equation) results show that the neural network can achieve better performance and robustness as shown in Figure 14. This research also further reveals the effect of complicated seafloor topography on IW propagation. Another scholar used the IW amplitudes simulated by physical oceanographic equipment as auxiliary data. The information between the auxiliary data and the spectral characteristics of remote sensing satellite images (MODIS) is correlated and migrated by observational neural networks to invert the IWs amplitudes (Zhang et al., 2022).

The application of DBAI technology in oceanic IWs is still in the exploration stage. These methods mainly learn the deep features of IWs through a large amount of remote sensing data rather than

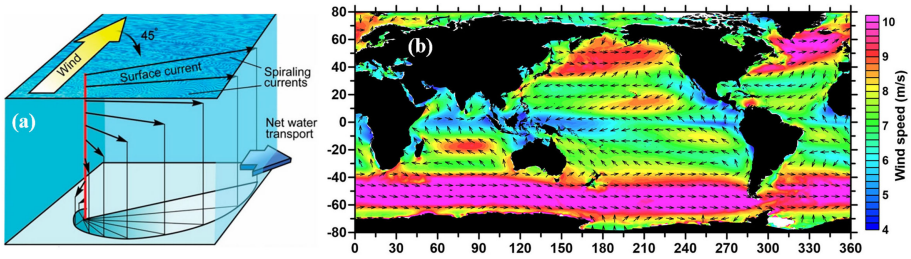


FIGURE 12
Ekman spiral. **(A)** Simulation model of Ekman spiral, **(B)** Global sea surface wind speed distribution.

constructing an exclusive structure of PINNs on the mechanism of IWs generation. Nevertheless, the feasibility of applying DBAI technology to oceanic IWs is confirmed, providing a novel idea for studying oceanic IWs.

4.4 Prediction of subsurface chlorophyll maxima

Biomass distribution is as important in the subsurface ocean as stratification structural features, whereas most of the water color remote sensing satellites can only directly acquire the chlorophyll concentration $[Chla]_{sur}$ from the sea surface. To invert chlorophyll vertical distribution from the sea surface chlorophyll concentration $[Chla]_{sur}$ based on a large number of *in situ* observations, the models of chlorophyll concentration content and its vertical distribution were constructed for different stratified water bodies and trophic levels, integrated over the euphotic layer (Uitz et al., 2006). The detailed calculations of the different models are shown in Equations (10) and (11).

$$Chla_{Z_{eu}} = A \cdot [Chla]_{sur}^{B} \tag{10}$$

$$C(\zeta) = c_0 - c_1 \cdot \zeta + c_2 \cdot exp - [(\zeta - c_3)/c_4]^2 \tag{11}$$

where $Chla_{Z_{eu}}$ represents the chlorophyll *a* content integrated over the euphotic zone, ς is the normalized value of the actual depth relative to the depth of the euphotic layer Z_{eu}, i.e., $\varsigma = z/Z_{eu}$, $C(\varsigma)$ denotes the normalized value of chlorophyll *a* concentration $[Chla(\varsigma)]$ at depth ς relative to the average chlorophyll concentration $\overline{[Chla]_{Zeu}}$ within the euphotic layer, i.e., $C(\varsigma) = [Chla(\varsigma)]/\overline{[Chla]_{Zeu}}$, among them, $\overline{[Chla]_{Zeu}} = Z_{eu}^{-1} \cdot Chla_{Zeu}$. Furthermore, the parameters A, B and c_0~c_4 in (10) and (11) can be obtained by fitting the field data of different stratified water bodies and trophic levels (Uitz et al., 2006). It can be seen that chlorophyll *a* vertical distribution of well-mixed waters ($\frac{Z_m}{Z_{eu}} > 1$, Z_m is the depth of mixed layer) is relatively uniform. In contrast, for stratified water bodies ($\frac{Z_m}{Z_{eu}} < 1$), SCM is generally present in the vertical profile. On the one hand, SCM plays a crucial role in ocean nutrient cycling, energy flow, and biogeochemical cycling, and on the other hand, the chlorophyll vertical distribution, especially the inversion of SCM, is important for accurate estimation of ocean primary productivity. However, the SCM features have been explored only at the regional scale due to the lack of 3-D observation data. With the implementation of the BGC-Argo program, it has become possible to analyze SCM on a global scale while revealing that the seasonal dynamics of SCM have an evident regional character (Cornec et al., 2021).

With the accumulation of biogeochemical 3-D observation data and the development of computer technology, there is an increasing amount of research related to applying neural networks to chlorophyll vertical structure inversion. The associative statistical neural network was developed as shown in Figure 15A and successfully inverted the vertical structure of SCM in the North Pacific region using sea surface parameters observed by satellite remote sensing (Chen et al., 2022b). The application of AI technology to the inversion and prediction of marine SCM can take full advantage of long-term and high-coverage satellite data to obtain the global oceanic SCM data set indirectly, the results are illustrated in Figure 15B. It is crucial for advancing marine ecology and related multidisciplinary research because it provides data for exploring the vertical ecological structure of the ocean.

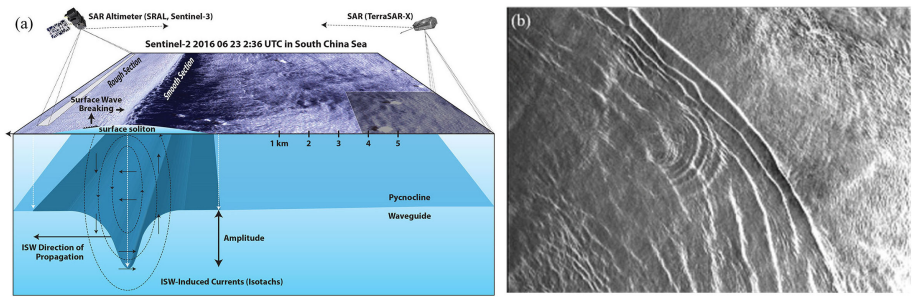


FIGURE 13
Oceanic IWs. **(A)** IWs mechanism for modulated SAR images, **(B)** IWs on SAR images (Magalhães et al., 2021).

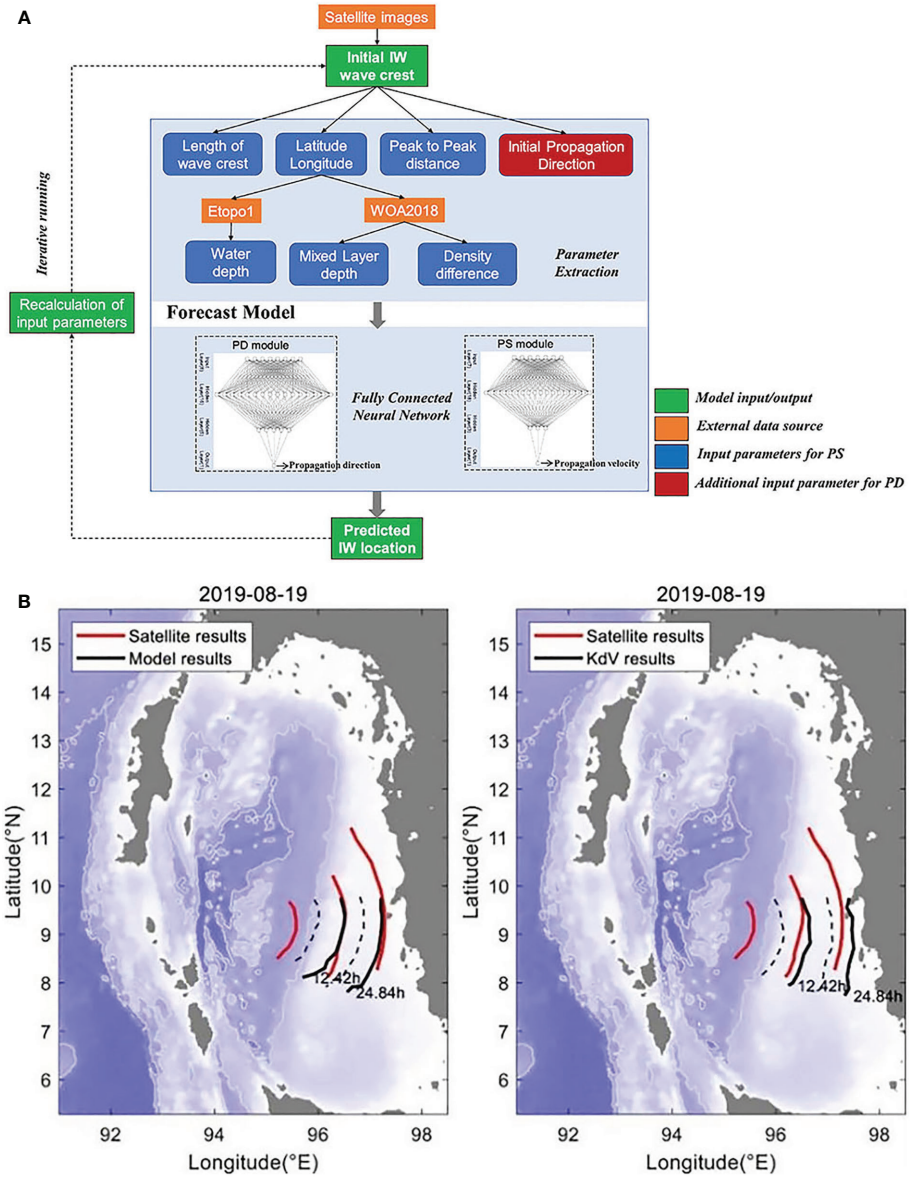


FIGURE 14
(A) The associative statistical neural network for IWs forecasting. **(B)** Forecast results of the IWs. Interval wave forecasting from satellite images. **(A)** The architecture of IWs forecasting neural network. **(B)** Forecast of IW propagations in the Andaman Sea after two semidiurnal tidal cycles. **(A)** shows the results of the associative statistical neural network, and **(B)** shows the results of the KdV equation (Zhang et al., 2021).

5 Conclusions and discussion

From the theoretical development of marine science, the progress of basic theory, observation technology, and computing ability have been drawing and responding to each other, pushing forward the leaps of this interdisciplinary science one after another. Before the middle of the last century, the development of marine science mainly benefited from advances in theory. Based on the limited field observation data, Ekman drift theory, Rossby planetary wave theory, westward strengthening theory of ocean circulation, wind-induced ocean circulation theory, and so on were developed successively, laying the theoretical foundation of modern physical oceanography.

From the perspective of marine science development in the 1970s, the emergence of satellite ocean remote sensing for scientific research provides unprecedented high spatial and temporal coverage data. With remote sensing data, it is the first time to realize the clear appearance of ocean circulation, qualify the global average sea level rise, estimate the global marine primary productivity, etc. Subsequently, the rapid development of computer technology has brought about a significant improvement in computing power, and model systems with high resolution have begun to shine in the cutting-edge applications of operational weather forecasting and climate change prediction. At the end of the 20th century, to jointly cope with the severe challenges brought by global changes, oceanographers devoted themselves to studying the processes occurring in the intermediate and deep sea, filling in the gaps of the existing theoretical system and no longer satisfied with the understanding of the ocean surface and phenomena at large space–time scale.

In this context, the *in situ* observation technology provides profile data to implement research from surface ocean phenomena to

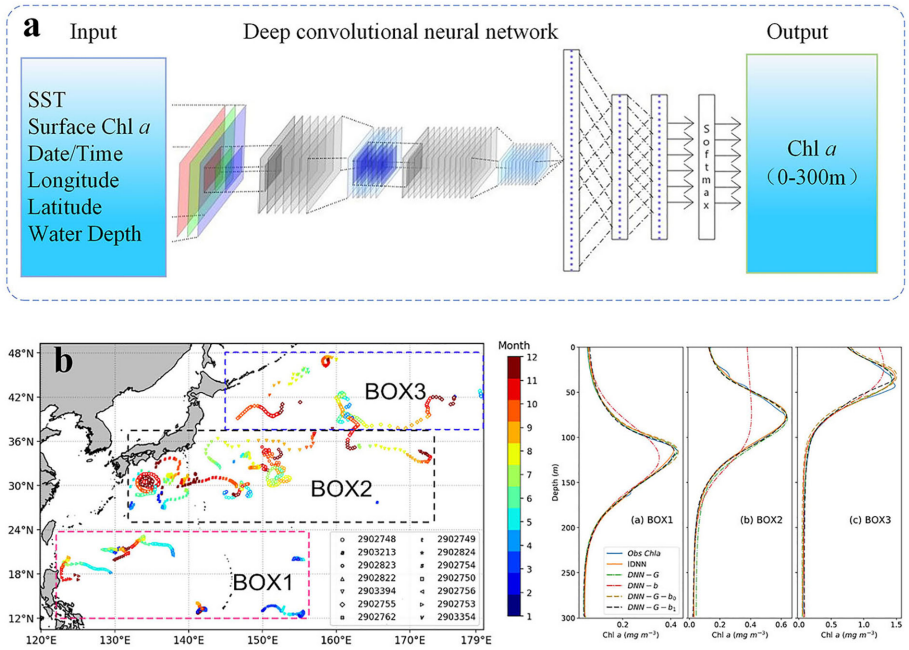


FIGURE 15
(A) The SCM perception network. **(B)** SCM predictions results. Prediction of SCM. **(A)** The architecture of SCM prediction neural network. **(B)** Comparisons of the mean profiles of observed values(blue line), associative statistical neural network predictions (orange line), and multi layer perceptron predictions (green dash line) inferred from remote sensing data in three boxes.

intermediate ocean dynamic and ecological processes. Naturally, how to reveal the intermediate ocean law from the mass fusion data has become a major challenge. "AI for science" has spawned a new paradigm of scientific research, and AI has also become a strategic technology for new scientific research. The development of AI technology requires not only a breakthrough in computational theory but also a full integration of interdisciplinary knowledge. The ultimate goal of AI is to train machines to be brain-like, bridging the knowledge between the human brain and the machine. Moreover, the natural combination of AI technology and oceanography (DBAI method) provides an excellent way to solve the current development dilemma of oceanography. The two have built a new bridge between data science and knowledge discovery, and initially showed unique advantages and great potential in intermediate ocean remote sensing applications from the outside to the inside.

DBAI is poised to achieve "surface-to-interior" knowledge of the middle ocean remote sensing and its processes, covering for the inadequacy of the existing theoretical system. However, the intermediate ocean remote sensing technology supported by AI is still facing two major challenges, one is the ocean data bottleneck the other is the generalization of DBAI methodology. Therefore, we believe that future trends in DBAI will focus on the following two points.

(1) Promote the construction of ocean big data. To fill the current gaps in some ocean parameters and underwater 3-D remote sensing, the development of a new generation of ocean science satellites (e.g., SWOT, Guanlan) is expected to enhance the comprehensive capabilities of satellites in horizontal space, vertical profile, spatial and temporal resolution, and integrated remote sensing. The cornerstone of DBAI is marine big data. With access to massive remote sensing data, the biggest obstacle to the development of DBAI is how to provide reliable ground truth to express information about marine objects. If we want to break through this obstacle, relying on one person or team is unrealistic. Therefore, we should actively participate in ocean big data construction plans, implement extensive international cooperation, and promote the construction of standard marine big data sets.

(2) Strengthening DBAI technology innovation. Most DBAI technologies for ocean science come from the computer-vision community. These techniques aim to enable computers to obtain meaningful pattern information from images, videos, and other visual inputs, often focusing on spatial patterns. These models can be used to construct oceanographic AI models with physical interpretability by combining oceanographic mechanism constraints with remote sensing data-driven to achieve theoretical and technological innovations in DBAI.

With the gradual enrichment of ocean big data sets and the continuous improvement of DBAI technology, DBAI will be widely applied in various marine subdivisions to realize automatic or semiautomatic scientific discovery. We envisage that the new scientific research paradigm of "AI for ocean science" will gradually uncover the mystery of intermediate and even deep oceans, ushering in a new phase in the future of ocean science.

Author contributions

GC conceived the study, co-designed and contributed to data collection, undertook analysis, and led the drafting of the manuscript.

BH, JY, and MR co-designed and contributed to data collection, undertook analysis, and contributed to the drafting of the manuscript. LG, CC, XC, LX, GH, and YM contributed to data collection, data analysis, and wrote sections of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

Funding

This research was jointly supported by the National Natural Science Foundation of China (NO. 42030406 and No. 42276203) and the Natural Science Foundation of Shandong Province (No. ZR2021MD001).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Accarino, G., Chiarelli, M., Fiore, S., Federico, I., Causio, S., Coppini, G., et al. (2021). A multi-model architecture based on long short-term memory neural networks for multi-step sea level forecasting. *Future Generation Comput. Syst.* 124, 1–9. doi: 10.1016/j.future.2021.05.008

Amani, M., Ghorbanian, A., Asgarimehr, M., Yekkehkhany, B., Moghimi, A., Jin, S., et al. (2021). Remote sensing systems for ocean: A review (part 1: Passive systems). *IEEE J. Selected Topics Appl. Earth Observations Remote Sens.* 15, 210–234. doi: 10.1109/JSTARS.2021.3130789

Amani, M., Mehravar, S., Asiyabi, R. M., Moghimi, A., Ghorbanian, A., Ahmadi, S. A., et al. (2022a). Ocean remote sensing techniques and applications: A review (part ii). *Water* 14, 3401. doi: 10.3390/w14213401

Amani, M., Moghimi, A., Mirmazloumi, S. M., Ranjgar, B., Ghorbanian, A., Ojaghi, S., et al. (2022b). Ocean remote sensing techniques and applications: A review (part i). *Water* 14, 3400. doi: 10.3390/w14213400

Amani, M., Mohseni, F., Layegh, N. F., Nazari, M. E., Fatolazadeh, F., Salehitangrizi, A., et al. (2022c). Remote sensing systems for ocean: A review (part 2: Active systems). *IEEE J. Selected Topics Appl. Earth Observations Remote Sens* 15, 1421–1453. doi: 10.1109/JSTARS.2022.3141980

Amores, A., Jordà, G., Arsouze, T., and Le Sommer, J. (2018). Up to what extent can we characterize ocean eddies using present-day gridded altimetric products? *J. Geophysical Research: Oceans* 123, 7220–7236. doi: 10.1029/2018JC014140

Andersson, T. R., Hosking, J. S., Pérez-Ortiz, M., Paige, B., Elliott, A., Russell, C., et al. (2021). Seasonal arctic sea ice forecasting with probabilistic deep learning. *Nat. Commun.* 12, 5124. doi: 10.1038/s41467-021-25257-4

Appenzeller, T. (2021). *The AI revolution in science. report 16.* (Washington, United States: American Association for the Advancement of Science).

Behrenfeld, M. J., Gaube, P., Della Penna, A., O'Malley, R. T., Burt, W. J., Hu, Y., et al. (2019). Global satellite-observed daily vertical migrations of ocean animals. *Nature* 576, 257–261. doi: 10.1038/s41586-019-1796-9

Behrenfeld, M. J., Hu, Y., Hostetler, C. A., Dall'Olmo, G., Rodier, S. D., Hair, J. W., et al. (2013). Space-based lidar measurements of global ocean carbon stocks: Space lidar plankton measurements. *Geophysical Res. Lett.* 40, 4355–4360. doi: 10.1002/grl.50816

Behrenfeld, M. J., Hu, Y., O'Malley, R. T., Boss, E. S., Hostetler, C. A., Siegel, D. A., et al. (2017). Annual boom–bust cycles of polar phytoplankton biomass revealed by space-based lidar. *Nat. Geosci.* 10, 118–122. doi: 10.1038/ngeo2861

Chaigneau, A., Le Texier, M., Eldin, G., Grados, C., and Pizarro, O. (2011). Vertical structure of mesoscale eddies in the eastern south pacific ocean: A composite analysis from altimetry and argo profiling floats. *J. Geophysical Research: Oceans* 116, 1–16. doi: 10.1029/2011JC007134

Chelton, D. B., deSzoeke, R. A., Schlax, M. G., El Naggar, K., and Siwertz, N. (1998). Geographical variability of the first baroclinic rossby radius of deformation. *J. Phys. Oceanography* 28, 433–460. doi: 10.1175/1520-0485(1998)028<0433:GVOTFB>2.0.CO;2

Chen, G., Chen, X., and Cao, C. (2022a). Divergence and dispersion of global eddy propagation from satellite altimetry. *J. Phys. Oceanography* 52, 705–722. doi: 10.1175/JPO-D-21-0122.1

Chen, X., Chen, G., Ge, L., Huang, B., and Cao, C. (2021c). Global oceanic eddy identification: A deep learning method from argo profiles and altimetry data. *Front. Mar. Sci.* 8, 21–35. doi: 10.3389/fmars.2021.646926

Chen, G., Chen, X., and Huang, B. (2021a). Independent eddy identification with profiling argo as calibrated by altimetry. *J. Geophysical Research: Oceans* 126, 1–22. doi: 10.1029/2020JC016729

Chen, J., Gong, X., Guo, X., Xing, X., Lu, K., Gao, H., et al. (2022b). Improved perceptron of subsurface chlorophyll maxima by a deep neural network: A case study with bgc-argo float data in the northwestern pacific ocean. *Remote Sens.* 14, 632. doi: 10.3390/rs14030632

Chen, P., Jamet, C., Zhang, Z., He, Y., Mao, Z., Pan, D., et al. (2021b). Vertical distribution of subsurface phytoplankton layer in south china sea using airborne lidar. *Remote Sens. Environ.* 263, 112567. doi: 10.1016/j.rse.2021.112567

Chen, L.-C., Papandreou, G., Kokkinos, I., Murphy, K., and Yuille, A. L. (2018). Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE Trans. Pattern Anal. Mach. Intell.* 40, 834–848. doi: 10.1109/TPAMI.2017.2699184

Chen, G., Tang, J., Zhao, C., Wu, S., Yu, F., Ma, C., et al. (2019). Concept design of the "guanlan" science mission: China's novel contribution to space oceanography. *Front. Mar. Sci.* 6, 194–208. doi: 10.3389/fmars.2019.00194

Chepfer, H., Noel, V., Chiriaco, M., Wielicki, B., Winker, D., Loeb, N., et al. (2018). The potential of a multidecade spaceborne lidar record to constrain cloud feedback. *J. Geophysical Research: Atmospheres* 123, 5433–5454. doi: 10.1002/2017JD027742

Churnside, J. H. (2013). Review of profiling oceanographic lidar. *Optical Eng.* 53, 51405. doi: 10.1117/1.OE.53.5.051405

Cornec, M., Laxenaire, R., Speich, S., and Claustre, H. (2021). Impact of mesoscale eddies on deep chlorophyll maxima. *Geophysical Res. Lett.* 48, 1–10. doi: 10.1029/2021GL093470

Cozman, F. (2021). How ai is helping the natural sciences. *Artif. Intell.* 598, S5–S7. doi: 10.1038/d41586-021-02762-6

Cranmer, M., Greydanus, S., Hoyer, S., Battaglia, P., Spergel, D., and Ho, S. (2020). "Lagrangian Neural networks," in *ICLR 2020 Workshop on Integration of Deep Neural Models and Differential Equations*. 1–9 (Addis Ababa, Ethiopia: https://OpenReview.net).

George, T. M., Manucharyan, G. E., and Thompson, A. F. (2021). Deep learning to infer eddy heat fluxes from sea surface height patterns of mesoscale turbulence. *Nat. Commun.* 12, 800. doi: 10.1038/s41467-020-20779-9

Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., et al. (2014). "Generative adversarial nets," in *The international conference on neural information processing systems* (Cambridge, MA, USA: MIT Press), 2672–2680. NIPS'14.

Greydanus, S., Dzamba, M., and Yosinski, J. (2019). "Hamiltonian Neural networks," in *The international conference on neural information processing systems*(Vancouver, Canada: Curran Associates, Inc) 1–11.

Ham, Y.-G., Kim, J.-H., and Luo, J.-J. (2019). Deep learning for multi-year enso forecasts. *Nature* 573, 568–572. doi: 10.1038/s41586-019-1559-7

He, K., Gkioxari, G., Dollar, P., and Girshick, R. (2017). "Mask r-cnn," in *IEEE International conference on computer vision* (Venice: IEEE), 2980–2988. doi: 10.1109/ICCV.2017.322

He, K., Zhang, X., Ren, S., and Sun, J. (2016). "Deep residual learning for image recognition," in *2016 IEEE conference on computer vision and pattern recognition* (Las Vegas, NV, USA: IEEE), 770–778. doi: 10.1109/CVPR.2016.90

Hostetler, C. A., Behrenfeld, M. J., Hu, Y., Hair, J. W., and Schulien, J. A. (2018). Spaceborne lidar in the study of marine systems. *Annu. Rev. Mar. Sci.* 10, 121–147. doi: 10.1146/annurev-marine-121916-063335

Huang, B., Ge, L., Chen, X., and Chen, G. (2022). Vertical structure-based classification of oceanic eddy using 3-d convolutional neural network. *IEEE Trans. Geosci. Remote Sens.* 60, 1–14. doi: 10.1109/TGRS.2021.3103251

Huang, G., Liu, Z., van der Maaten, L., and Weinberger, K. Q. (2017). "Densely connected convolutional networks," in *2017 IEEE conference on computer vision and pattern recognition (CVPR)* (Honolulu, HI: IEEE), 2261–2269. doi: 10.1109/CVPR.2017.243

Jian, M., Liu, X., Luo, H., Lu, X., Yu, H., and Dong, J. (2021). Underwater image processing and analysis: A review. *Signal Processing: Image Communication* 91, 116088. doi: 10.1016/j.image.2020.116088

Jian, M., Qi, Q., Dong, J., Yin, Y., and Lam, K.-M. (2018). Integrating qdwd with pattern distinctness and local contrast for underwater saliency detection. *J. Visual communication image representation* 53, 31–41. doi: 10.1016/j.jvcir.2018.03.008

Jian, M., Qi, Q., Yu, H., Dong, J., Cui, C., Nie, X., et al. (2019). The extended marine underwater environment database and baseline evaluations. *Appl. Soft Computing* 80, 425–437. doi: 10.1016/j.asoc.2019.04.025

Johnson, G. C., Hosoda, S., Jayne, S. R., Oke, P. R., Riser, S. C., Roemmich, D., et al. (2022). Argo–two decades: Global oceanography, revolutionized. *Annu. Rev. Mar. Sci.* 14, 379–403. doi: 10.1146/annurev-marine-022521-102008

Karniadakis, G. E., Kevrekidis, I. G., Lu, L., Perdikaris, P., Wang, S., and Yang, L. (2021). Physics-informed machine learning. *Nat. Rev. Phys.* 3, 422–440. doi: 10.1038/s42254-021-00314-5

Lagerloef, G. S. E., Mitchum, G. T., Lukas, R. B., and Niiler, P. P. (1999). Tropical pacific near-surface currents estimated from altimeter, wind, and drifter data. *J. Geophysical Research: Oceans* 104, 23313–23326. doi: 10.1029/1999JC900197

Lecun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proc. IEEE* 86, 2278–2324. doi: 10.1109/5.726791

Lei, N. (2020). A geometric understanding of deep learning. *Engineering* 6, 361–374. doi: 10.1016/j.eng.2019.09.010

Li, G., Cheng, L., Zhu, J., Trenberth, K. E., Mann, M. E., and Abraham, J. P. (2020). Increasing ocean stratification over the past half-century. *Nat. Climate Change* 10, 1116–1123. doi: 10.1038/s41558-020-00918-2

Liu, Y., Zheng, Q., and Li, X. (2021). Characteristics of global ocean abnormal mesoscale eddies derived from the fusion of sea surface height and temperature data by deep learning. *Geophysical Res. Lett.* 48, 1–11. doi: 10.1029/2021GL094772

Li, J., Wang, G., Xue, H., and Wang, H. (2019). A simple predictive model for the eddy propagation trajectory in the northern south china sea. *Ocean Sci.* 15, 401–412. doi: 10.5194/os-15-401-2019

Lu, X., Hu, Y., Yang, Y., Bontempi, P., Omar, A., and Baize, R. (2020). Antarctic Spring ice-edge blooms observed from space by icesat-2. *Remote Sens. Environ.* 245, 111827. doi: 10.1016/j.rse.2020.111827

Magalhães, J. M., Alpers, W., Santos-Ferreira, A. M., and Da Silva, J. C. (2021). Surface wave breaking caused by internal solitary waves. *Oceanography* 34, 166–176. doi: 10.5670/oceanog.2021.203

Meng, Y., Rigall, E., Chen, X., Gao, F., Dong, J., and Chen, S. (2021). "Physics-guided generative adversarial networks for sea subsurface temperature prediction," in *IEEE Transactions on neural networks and learning systems*. (Piscataway, United States: IEEE), 1–14. doi: 10.1109/TNNLS.2021.3123968

Meng, L., and Yan, X.-H. (2022). "Remote sensing for subsurface and deeper oceans: An overview and a future outlook," in *IEEE Geoscience and remote sensing magazine*. Piscataway, United States, IEEE, 2–22. doi: 10.1109/MGRS.2022.3184951

Morel, A., and Berthon, J.-F. (1989). Surface pigments, algal biomass profiles, and potential production of the euphotic layer: Relationships reinvestigated in view of remote-sensing applications: Production computed from space-acquired data. *Limnology Oceanography* 34, 1545–1562. doi: 10.4319/lo.1989.34.8.1545

Nakajima, M., Tanaka, K., and Hashimoto, T. (2021). "Neural schrödinger equation: Physical law as deep neural network," in *IEEE Transactions on neural networks and learning systems*. (Piscataway, United States: IEEE), 1–15. doi: 10.1109/TNNLS.2021.3120472

Pegliasco, C., Delepoulle, A., Mason, E., Morrow, R., Faugère, Y., and Dibarboure, G. (2022). Meta3.1exp: A new global mesoscale eddy trajectory atlas derived from altimetry. *Earth System Sci. Data* 14, 1087–1107. doi: 10.5194/essd-14-1087-2022

Raissi, M., Perdikaris, P., and Karniadakis, G. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comput. Phys.* 378, 686–707. doi: 10.1016/j.jcp.2018.10.045

Riser, S. C., Freeland, H. J., Roemmich, D., Wijffels, S., Troisi, A., Belbéoch, M., et al. (2016). Fifteen years of ocean observations with the global argo array. *Nat. Climate Change* 6, 145–153. doi: 10.1038/nclimate2872

Ronneberger, O., Fischer, P., and Brox, T. (2015). *U-Net: Convolutional networks for biomedical image segmentation* Vol. 9351 (Cham: Springer International Publishing), 234–241.

Simonyan, K., and Zisserman, A. (2015). "Very deep convolutional networks for large-scale image recognition," in *The international conference on learning representations*(San Diego, CA, USA: https://OpenReview.net), 1–14.

Sprintall, J., and Cronin, M. F. (2009). *Upper ocean vertical structure* (Oxford: Academic Press), 217–224. doi: 10.1016/B978-012374473-9.00627-5

Su, H., Wang, A., Zhang, T., Qin, T., Du, X., and Yan, X.-H. (2021). Super-resolution of subsurface temperature field from remote sensing observations based on machine learning. *Int. J. Appl. Earth Observation Geoinformation* 102, 102440. doi: 10.1016/j.jag.2021.102440

Trujillo, A. P., and Thurman, H. V. (2011). *Essentials of oceanography*. (New York, United States: Pearson Education).

Uitz, J., Claustre, H., Morel, A., and Hooker, S. B. (2006). Vertical distribution of phytoplankton communities in open oocean: An assessment based on surface chlorophyll. *J. Geophysical Res.* 111, C08005. doi: 10.1029/2005JC003207

Wang, R., Walters, R., and Yu, R. (2021). Incorporating symmetry into deep dynamics models for improved generalization. *Int. Conf. Learn. Representations.* 1, 1–20.

Wang, X., Wang, X., Yu, M., Li, C., Song, D., Ren, P., et al. (2022). Mesogru: Deep learning framework for mesoscale eddy trajectory prediction. *IEEE Geosci. Remote Sens. Lett.* 19, 1–5. doi: 10.1109/LGRS.2021.3087835

Werdell, P. J., Behrenfeld, M. J., Bontempi, P. S., Boss, E., Cairns, B., Davis, G. T., et al. (2019). The plankton, aerosol, cloud, ocean ecosystem mission: Status, science, advances. *Bull. Am. Meteorological Soc.* 100, 1775–1794. doi: 10.1175/BAMS-D-18-0056.1

Wojtowytsch, W. E. S. (2020). On the banach spaces associated with multi-layer relu networks: Function representation, approximation theory and gradient descent dynamics. *CSIAM Trans. Appl. Mathematics* 1, 387–440. doi: 10.4208/csiam-am.20-211

Wright, L. G., Onodera, T., Stein, M. M., Wang, T., Schachter, D. T., Hu, Z., et al. (2022). Deep physical neural networks trained with backpropagation. *Nature* 601, 549–555. doi: 10.1038/s41586-021-04223-6

Xie, J., Zhang, J., Yu, J., and Xu, L. (2020). An adaptive scale sea surface temperature predicting method based on deep learning with attention mechanism. *IEEE Geosci. Remote Sens. Lett.* 17, 740–744. doi: 10.1109/LGRS.2019.2931728

Xiong, S., Tong, Y., He, X., Yang, S., Yang, C., and Zhu, B. (2021). Nonseparable symplectic neural networks. *Int. Conf. Learn. Representations*. 1, 1–19

Zhang, X., Li, X., and Zheng, Q. (2021). A machine-learning model for forecasting internal wave propagation in the andaman sea. *IEEE J. Selected Topics Appl. Earth Observations Remote Sens.* 14, 3095–3106. doi: 10.1109/JSTARS.2021.3063529

Zhang, Z., Wang, W., and Qiu, B. (2014). Oceanic mass transport by mesoscale eddies. *Science* 345, 322–324. doi: 10.1126/science.1252418

Zhang, X., Wang, H., Wang, S., Liu, Y., Yu, W., Wang, J., et al. (2022). Oceanic internal wave amplitude retrieval from satellite images based on a data-driven transfer learning model. *Remote Sens. Environ.* 272, 112940. doi: 10.1016/j.rse.2022.112940

Zheng, Q., Yuan, Y., Klemas, V., and Yan, X.-H. (2001). Theoretical expression for an ocean internal soliton synthetic aperture radar image and determination of the soliton characteristic half width. *J. Geophysical Research: Oceans* 106, 31415–31423. doi: 10.1029/2000JC000726

frontiers | Frontiers in Marine Science

TYPE Original Research
PUBLISHED 24 January 2023
DOI 10.3389/fmars.2022.1006452

OPEN ACCESS

EDITED BY
Junyu He,
Zhejiang University, China

REVIEWED BY
Xian-Hua Han,
Yamaguchi University, Japan
Qi Wu,
University of Adelaide, Australia

*CORRESPONDENCE
Peng Chen
capricorncp@163.com

SPECIALTY SECTION
This article was submitted to Ocean Observation, a section of the journal Frontiers in Marine Science

RECEIVED 29 July 2022
ACCEPTED 29 December 2022
PUBLISHED 24 January 2023

CITATION
Hu S, Hou R, Ming L, Meifang S and Chen P (2023) A hyperspectral image reconstruction algorithm based on RGB image using multi-scale atrous residual convolution network.
Front. Mar. Sci. 9:1006452.
doi: 10.3389/fmars.2022.1006452

A hyperspectral image reconstruction algorithm based on RGB image using multi-scale atrous residual convolution network

Shaoxiang Hu[1], Rong Hou[2], Luo Ming[1], Su Meifang[1] and Peng Chen[2]*

[1]School of Electronic Science and Engineering, University of Electronic Science and Technology of China, Chengdu, China, [2]Chengdu Research Base of Giant Panda Breeding, Sichuan Key Laboratory of Conservation Biology for Endangered Wildlife, Chengdu, China

Hyperspectral images are a valuable tool for remotely sensing important characteristics of a variety of landscapes, including water quality and the status of marine disasters. However, hyperspectral data are rare or expensive to obtain, which has spurred interest in low-cost, fast methods for reconstructing hyperspectral data from much more common RGB images. We designed a novel algorithm to achieve this goal using multi-scale atrous convolution residual network (MACRN). The algorithm includes three parts: low-level feature extraction, high-level feature extraction, and feature transformation. The high-level feature extraction module is composed of cascading multi-scale atrous convolution residual blocks (ACRB). It stacks multiple modules to form a depth network for extracting high-level features from the RGB image used as an input. The algorithm uses jump connection for residual learning, and the final high-level feature combines the output of the low-level feature extraction module and the output of the cascaded atrous convolution residual block element by element, so as to prevent gradient dispersion and gradient explosion in the deep network. Without adding too many parameters, the model can extract multi-scale features under different receptive fields, make better use of the spatial information in RGB images, and enrich the contextual information. As a proof of concept, we ran an experiment using the algorithm to reconstruct hyperspectral Sentinel-2 satellite data from the northern coast of Australia. The algorithm achieves hyperspectral spectral reconstruction in 443nm-2190nm band with less computational cost, and the results are stable. On the Realworld dataset, the reconstruction error MARE index is less than 0.0645, and the reconstruction time is less than 9.24S. Therefore, in the near infrared band, MACRN reconstruction accuracy is significantly better than other spectral reconstruction algorithms. MACRN hyperspectral reconstruction algorithm has the characteristics of low reconstruction cost and high reconstruction accuracy, and its advantages in ocean spectral reconstruction are more obvious.

KEYWORDS

hyperspectral image reconstruction, deep learning, atrous convolution, marine science, multi-scale

1 Introduction

RGB image is a color image composed of red, green and blue channels. Hyperspectral image is an image composed of dozens or hundreds of channels, which is divided in detail in the dimension of spectrum, as shown in Figure 1A. Hyperspectral images not only contain image information, but also can be expanded in the spectral dimension. For a certain substance, different spectral wavelengths may correspond to different values, and the spectral curve of the substance can be obtained, as shown in Figure 1B. Therefore, hyperspectral images provide richer spectral information than RGB images. Targets with different components and attributes may have similar appearances, but their spectral curves will be quite distinct. This means hyperspectral images have unique advantages in computer vision tasks such as target recognition and image analysis, leading to the gradual application of hyperspectral image analysis in various research fields such as remote sensing, agriculture, medicine and ocean sciences. For the latter, hyperspectral data have been used in the monitoring of marine water quality (such as total suspended solids, chlorophyll, dissolved organic matter, etc.) (Veronez et al., 2018; Peterson et al., 2020; Ying et al., 2021) and marine disasters (such as red tide, oil spills, etc.) (Guga, 2020; Qizhong et al., 2021).

However, hyperspectral images also bring significant capture complexity, high costs (Geelen et al., 2014; Beletkaia and Pozo, 2020), or high light field requirements (Descour and Dereniak, 1995). As a result, reconstructing hyperspectral images from ubiquitous RGB images at a low cost has become a hot research topic. There is a complex correlation between the pixel values of RGB images and hyperspectral images. However, compared with hyperspectral images, RGB images provide less spectral information, which makes it very difficult to reconstruct hyperspectral images from RGB images. Therefore, reconstructing hyperspectral images from RGB images is a very challenging task.

The existing methods of reconstructing hyperspectral images from RGB images can be roughly divided into two categories: the first is to design specific systems based on ordinary RGB cameras. Goel (Goel et al., 2015) designed a hyperspectral camera using time-division multiplexing illumination source to realize reconstruction, which is a universal hyperspectral imaging system for visible light and NIR wavelength. The camera has a frame rate of 150 FPS and a maximum resolution of 1280 × 1024. Wug (Oh et al., 2016) used multiple color cameras for reconstruction, and used different spectral sensitivities of different camera sensors to reconstruct the hyperspectral images, with a resolution of 200x300 and an imaging time of 2 minutes. However, these methods rely on strict environmental conditions or additional equipment. The second method is to directly model the mapping relationship between RGB images and hyperspectral images of ordinary cameras by using the correlation and a large volume of training data. Because this mapping is highly nonlinear, machine learning methods are usually employed.

Some early work (Farnaz et al., 2008; Arad and Benshahar, 2016) expressed this problem as a weighted combination of basis functions, using principal component analysis to extract the basis function from hyperspectral data sets. Other studies use sparse coding to reconstruct hyperspectral data from RGB images (Antonio, 2015; Jia et al., 2017; Aeschbacher et al., 2017). Nguyen et al. (2014) proposed to use radial basis function networks to learn the mapping from RGB image to hyperspectral image, reduce the impact of illumination on network performance, and do white balance processing on the input image. Arad et al. (Antonio, 2015) constructed a new hyperspectral dataset of natural scenes and used it to build a sparse spectral dictionary and its corresponding RGB projection to reconstruct hyperspectral images from RGB images.

Recently, a number of studies have explored end-to-end mapping of the relationship between RGB images and hyperspectral images through neural networks (Shoeiby et al., 2018a; Shoeiby et al., 2018b; Kaya et al., 2019). Galliani, Alvarez Gila, Liu et al. (Galliani et al., 2017; Alvarezgila et al., 2017; Pengfei et al., 2020) applied convolutional neural network and GAN to hyperspectral image reconstruction. By using multi-scale feature pyramid module in GAN, the correlation between local and global features is established. Good reconstruction results have been achieved. Xiong et al. (2017) proposed an HSCNN network for reconstructing hyperspectral images from RGB images and compressed measurements. In order to simplify the sampling process. – proposed an HSCNN-D network, which achieved good reconstruction results. proposed a novel adaptive weighted attention network, which is mainly composed of multiple double residual attention blocks with long and short jump connections. Information on the spatial context is captured through second-order nonlocal operations, and it has more accurate reconstruction effect than HSCNN-D model on noiseless RGB images. proposed a 4-layer hierarchical regression network (HRnet)

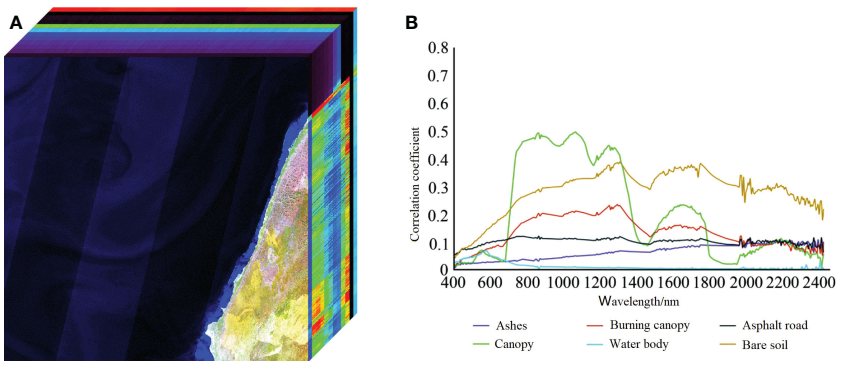


FIGURE 1
Hyperspectral 3D data from the Sentinel-2 satellite along the northern coast of Australia, **(A)** UTM zone 49S, **(B)** Spectral curves of different substances.

with Pixelshuffle layer as the interaction between layers. It uses residual dense blocks to extract features, and uses residual global blocks to build attention, which achieves better reconstruction results on real RGB images. Tao Huang et al. (2021) proposed a method to reconstruct a 3-channel HIS spectral image from a single channel 2D compressed image. This method is based on the maximum a posteriori (MAP) estimation framework using a learned Gaussian Scale Mixture (GSM) priori. It has good reconstruction results on both synthetic and real datasets. Yuanhao Cai et al. (2022) proposed a new Multi stage Spectral wise Transformer (MST++) efficient spectral reconstruction method based on Transformer. The Spectral Wise Multi head Self identification (S-MSA) with spatial sparsity and spectral self similarity is used to form the Spectral Wise Attention Block (SAB). Then, SAB establishes a single-stage spectral converter (SST), which is cascaded by several SSTs, and gradually improves the reconstruction quality from coarse to fine.

Although these studies have found effective methods to convert RGB images into hyperspectral data, they all have shortcomings such as low reconstruction accuracy, complex network structure and high computational cost. Hyperspectral reconstruction in marine environment generally covers visible, near infrared, mid infrared wave bands. However, the existing methods are aimed at the visible or near infrared wave band. Therefore, we need to design a new hyperspectral reconstruction algorithm suitable for the marine environment, make full use of the spatial characteristics of the marine environment, and achieve more accurate hyperspectral reconstruction.

Reconstructing hyperspectral images from RGB images can yield hyperspectral data at a low cost, which is conducive to expanding access to hyperspectral data, and giving full play to the unique advantages of hyperspectral data in computer vision analysis tasks, including in the marine sciences. Although the reconstruction methods based on deep learning have achieved good results, they make insufficient utilization of RGB image spatial features. To solve this problem, we propose a novel multi-scale atrous convolution residual network (MACRN) to reconstruct hyperspectral images from RGB images. The model uses convolution layers with different atrous rates to extract the features of the image at multiple scales and obtain the feature representation of different granularity. We posit that the feature representation of the fused feature map will be more precise and accurate, enhancing the extraction of spatial features from the image. On the basis of ensuring the integrity of local details, more abundant context information will be added, reducing discrepancies between the reconstructed hyperspectral image and the real data.

2 Methods and materials

2.1 Data collection

We used hyperspectral images from Sentinel-2 remote sensing satellite to test our reconstruction algorithm. The original data include remote sensing images of 12 spectral bands in the range of 443-2190nm, namely 443nm, 490nm, 560nm, 665nm, 705nm, 740nm, 783nm, 842nm, 945nm, 1375nm, 1610nm, and 2190nm. The image size was 10800 × 10800 pixels. We generated RGB images for training and testing from hyperspectral images of Sentinel-2, including RGB images without noise (Clean data) and RGB images with noise (Realworld data).

We directly obtained Clean data from hyperspectral data and the CIE spectral response function.

We derived Realworld data by adding noise (Gaussian noise) to the RGB image synthesized by the unknown spectral response function to better simulate the real world scenarios.

We generated 440 training, 10 verification, and 10 test images for each dataset. Figure 2A shows the two types of RGB images in the dataset, which show some differences in color representation. Figure 2B shows the hyperspectral data for the 443 nm, 490 nm, 783 nm and 1375 nm bands.

During MACRN model training, the size of hyperspectral data and RGB image sample pair is 64 × 64 pixels. Overlap area size is 32 pixels when clipping. We performed data enhancement operations such as rotation and mirroring them. The enhancement processing is to rotate within the range of 0° - 180°, the rotation angle is 45°, and make a horizontal flip at each angle at the same time. After enhancement, 1 group 482 × 512 sample pairs generated 3120 groups 64 × 64 sample pairs.

2.2 Problem description

Electromagnetic waves within a certain wavelength range irradiate object surfaces. Because different objects have varying reflectivity to electromagnetic waves of different wavelengths, imaging devices capture the reflection spectra of those surfaces and render images through the conversion of the imaging device. We can express the formation process of the image with Eq. (2.1).

$$I_k = \int_\Omega E(\lambda)S(\lambda)\Phi_k(\lambda)d\lambda \tag{2.1}$$

Where, I_k is the radiant energy recorded by the image sensor k, $E(\lambda)$ is the spectral energy distribution of the illumination source, $S(\lambda)$ is the spectral reflectance of the object surface, $\Phi_k(\lambda)$ is the spectral response function of sensor k at the incident wavelength, λ is the wavelength. If the value of λ is an electromagnetic wavelength in the visible light range of 400nm to 700nm, then k corresponds to the image sensors in the R, G and B bands, and we get an RGB image.

The spectral radiation value $R(\lambda)$ of the target can be obtained by multiplying $E(\lambda)$ and $S(\lambda)$; Thus, Eq. (2.1) can be transformed into Eq. (2.2).

$$I_k = \int_\Omega R(\lambda)\Phi_k(\lambda)d\lambda \tag{2.2}$$

Discretizing Eq. (2.2) yields:

$$I_k = \sum_n R(\lambda_n)\Phi_k(\lambda_n) \tag{2.3}$$

Where, λ_n represents the sampled bands. In our study, n has a value of 12 and corresponds to the wavelength range of 443nm to 2190nm.

Therefore, when I_k is known, find $R(\lambda_n)$. It is a highly uncertain problem to reconstruct the hyperspectral radiation values of 12 bands from RGB images.

We can see that there are huge challenges in how to reconstruct high-precision hyperspectral images from RGB images. Particularly tricky are the problems of insufficient utilization of spatial features of the RGB images, and the unknown spectral response function of the sample data.

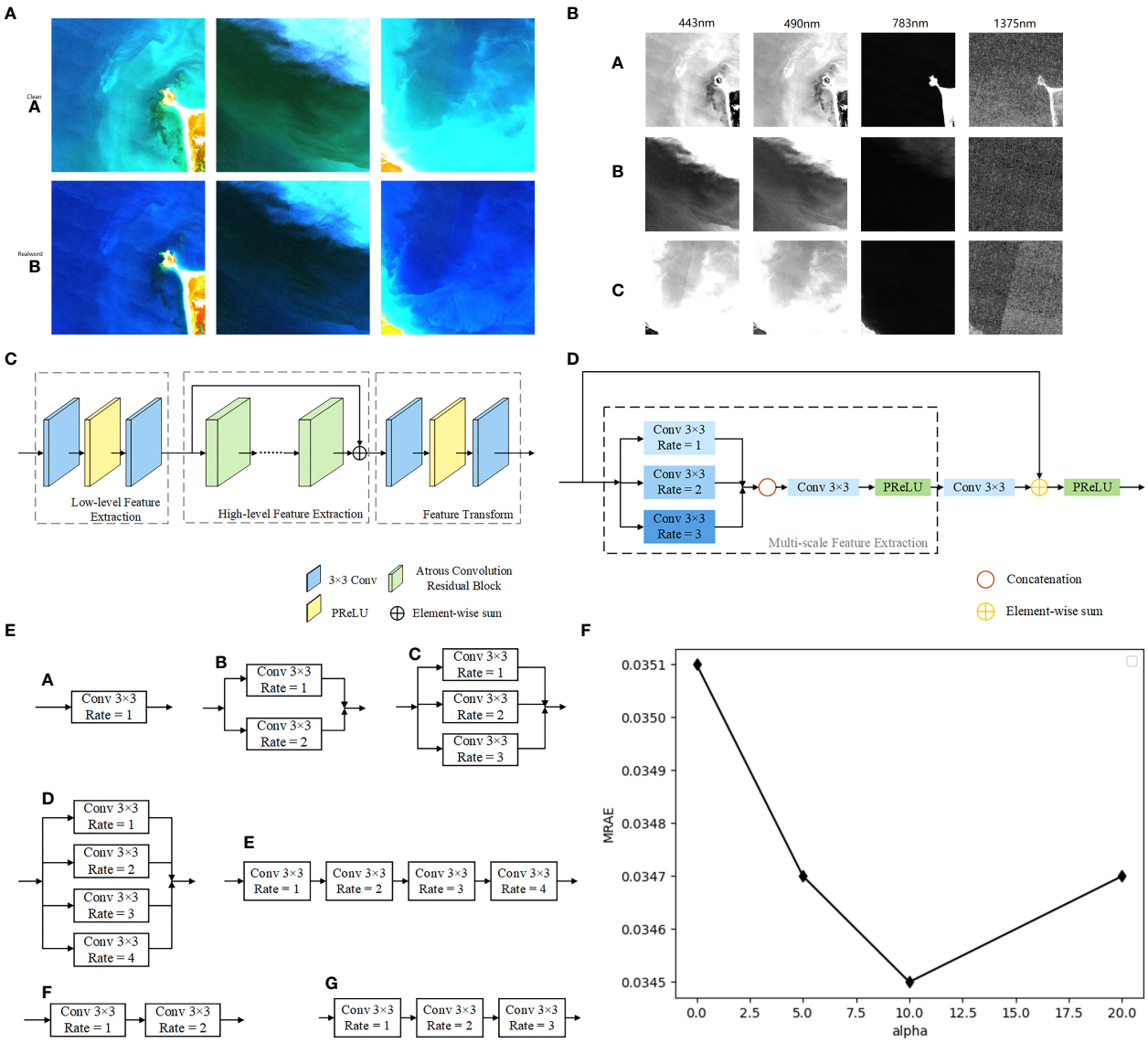


FIGURE 2
(A) A comparison of the two datasets used in this study ((**a**) 'Clean' and **(b)** 'Realworld') at three locations ((8600, 9300), (10000, 4500), (10300, 900) in UTM zone 52S). **(B)** Single band hyperspectral data for the 443nm, 490nm,783nm,1375nmbands at **(a)** coordinates (8600, 9300) **(b)** (10000, 4500and **(c)** (10300, 900), all in UTM zone 52S. **(C)** Overall Structure of MACRN. **(D)** Structure of the Multi-scale ACRB. **(E)** Seven kinds of ACRB, **(a)** C1, **(b)** C1C2, **(c)** C1C2C3, **(d)** C1C2C3C4, **(e)** C1C2C3C4_S, **(f)** C1C2_S, **(g)** C1C2C3_S. **(F)** Effect of different α values on model reconstruction error.

We propose a multi-scale atrous convolution residual network (MACRN) model as a solution for obtaining more features of different granularity without introducing too many parameters, with the aim of reconstructing hyperspectral images with greater accuracy.

2.3 Deep learning network

Multiscale network can extract features at different scales, but it also has a high computing cost, and its requirements for computing devices continue to improve. In order to solve this problem, we replace the ordinary convolution kernel with the atrous convolution, which balances the computational cost while extracting multi-scale features.

The overall structure of our multi-scale atrous convolution residual network (MACRN) is shown in Figure 2C. MACRN includes three main parts: low-level feature extraction, high-level feature extraction, and feature transformation.

2.3.1 Low-level feature extraction

The module is composed of a 3 × 3 convolution layer, a PReLU activation function and a 3 x 3 convolution layer. We use it to extract low-level features from the input RGB images, and its expression is shown in Eq. (2.4).

$$y = W_2 PReLU(W_1 x) \tag{2.4}$$

where x is the input, y is the output of the low-level feature extraction, W_1,W_2 is the weight matrix of the two convolution layers in the module, and $PReLU$ is the PReLU activation function.

2.3.2 High-level feature extraction

The module is composed of multiple cascading atrous convolution residual blocks (ACRB). The network is formed through the stacking of multiple modules to extract the high-level features of the input RGB image. The specific structure of the ACRB will be described in detail in Section 2.3. The output of the low-level feature extraction is connected by the global residual and added element by element to the output of the cascaded ACRB to form the final high-level feature set, which can prevent the phenomenon of gradient dispersion and gradient explosion in the deep network architecture. This process is expressed in Eq. (2.5).

$$z = y + f_{ACRB}^{N}\left(\ldots\left(f_{ACRB}^{1}(y)\right)\right) \tag{2.5}$$

where, z is the output of the depth feature extraction module, $f_{ACRB}(\cdot)$ is the atrous convolution residual block, and N is the number of atrous convolution residual blocks.

2.3.3 Feature transformation

The composition of the feature transformation module is the same as that of low-level feature extraction, which is also composed of a 3 x 3 convolution layer, PReLU activation function and a 3 x 3 convolution layer. This module realizes feature transformation and channel integration. The final reconstructed hyperspectral image is formed, and its expression is shown in Eq. (2.6).

$$output = W_4 PReLU(W_3 z) \tag{2.6}$$

where, *output* is the final output of the whole network, i.e., the hyperspectral image reconstruction result. W_3,W_4 represents two convolution layers respectively, and *PReLU* represents the PReLU activation function.

2.4 Multi-scale atrous convolution residual block

Currently, most methods for hyperspectral image reconstruction using image blocks have a significant dependence on spatial features [33]. When the spatial arrangement of the original image is broken, the reconstruction error often increases greatly, which shows that spatial features play a very important role in the process of hyperspectral image reconstruction. Therefore, the hyperspectral reconstruction error can be further reduced by improving the utilization of spatial features in RGB images.

An important parameter in convolution layers is the size of the convolution kernel, also known as the receptive field, which determines the size of the local area that can be sampled during each iteration. A too small convolution kernel can only extract small local features, while a too large convolution kernel can greatly increase the amount of required computation. Therefore, in order to better extract the spatial features of the image without increasing the computational cost of the network, we use the multi-scale ACRB, which has multiple atrous convolution layers with different atrous rates to control the receptive field. The module can extract the multi-scale spatial features of the image, and then splice the feature maps at different scales to form a multi-scale feature layer. Its structure is shown in Figure 2D.

First, the input characteristic map passes through three 3 x 3 convolution layers with different atrous rates (e.g. the atrous rates for the layers are set to 1, 2 and 3 respectively) to obtain the characteristic f_1, f_2, f_3 at different scales. Then, f_1, f_2, f_3 is spliced together in the channel dimension and integrated through a 3 x 3 convolution layer. Then, the fused multi-scale features are obtained by feature transformation through a PReLU activation function and a 3 x 3 convolution layer. The expression is shown in Eq. (2.7)

$$u = W_6(PReLU(W_5(concat[f_1, f_2, f_3]))) \tag{2.7}$$

where, u is the output of the multi-scale feature extraction module in the atrous convolution residual block, W_5,W_6 represents two convolution layers, *PReLU* represents the PReLU activation function, and *concat* [,·], represents the channel dimension splicing operation.

To prevent gradient dispersion and gradient explosion, we used a jump connection in the atrous convolution residual block to add the module input and the extracted multi-scale features element by element. The model uses the network fitting residual instead of directly learning the identity mapping function, and its expression is shown in Eq. (2.8).

$$v^{(n)} = PReLU\left(u + v^{(n-1)}\right) \tag{2.8}$$

where, $v^{(n)}$ is the output of the nth ACRB.

To choose the appropriate multiscale module construction strategy and verify the effectiveness of the results, we designed seven ACRBs with different structures for experiments. We named them C1, C1C2, C1C2C3, C1C2C3C4, C1C2C3_S, C1C2_S, C1C2C3C4_S (Figure 2E).

Through experimental analysis, C1C2C3 proved to be the most effective in our experimental analysis and was used as the basic structure for the multi-scale feature extraction module.

2.5 Loss function

We divided RGB image samples into two sets according to their spectral response functions. The first set was made up of RGB image samples generated using the CIE color function as the spectral response function. We called these images 'Clean' data because no noise was added. The second type of data set is called Realworld data, which uses other camera sensitivity functions as spectral response functions and adds noise. The noise type is Gaussian noise, SNR is 30-40dB.

2.5.1 Realworld data

Since the spectral response function in the Realworld data is unknown, the loss function only needs to consider the difference between the reconstructed hyperspectral image and the real hyperspectral image.

RMSE is generally used as the loss function when calculating the error between the predicted image and the real image in the image processing task, given that the image brightness levels of different bands of hyperspectral images are usually quite varied. In the process of calculating the root mean square error, the weight difference between the high brightness region and the low brightness region

of the image may be too large. To avoid this situation and balance the deviation between different bands, we use Eq. (2.9) as the loss function of our hyperspectral image reconstruction algorithm:

$$l_H = \frac{1}{N}\sum_{i=1}^{n}\left(\left|I_R^{(i)} - I_G^{(i)}\right|/I_G^{(i)}\right) \tag{2.9}$$

where, N represents the total number of pixels, I_R represents the reconstructed spectral radiation value, and I_G represents the spectral radiation value of the actual hyperspectral image.

2.5.2 Clean data

Since the spectral response function Φ for the Clean data is the CIE color function, the loss function needs to take into account the difference between the reconstructed hyperspectral image and the real hyperspectral image, as well as the difference between the RGB images generated by mapping the two types of images to RGB space.

If the spectral response function Φ is known and the RGB image does not contain noise, we can use the function Φ as *a priori* information to project the reconstructed hyperspectral image into RGB space according to the function Φ. We use Eq. (2.10) to calculate the average absolute error between it and the input RGB image as the loss function.

$$l_R = \frac{1}{N}\sum_{i=1}^{n}\left(\left|I_R^{(i)} - I_G^{(i)}\right|\right) \tag{2.10}$$

where, $I_R^{(i)}$ represents the RGB projection of the reconstructed hyperspectral data by matrix multiplication with Φ and $I_G^{(i)}$ represents the RGB projection of real hyperspectral data by matrix multiplication with Φ.

2.5.3 Overall loss function

The overall loss function of our hyperspectral reconstruction model includes l_H and l_R, and its expression is shown in Eq. (2.11)

$$l = l_H + \alpha l_R \tag{2.11}$$

where α is a variable parameter.

We analyzed the influence of different values of α in the loss function Eq. (2.11) of MACRN on the final reconstruction results. Set the value of α to 0, 1, 5, 10 and 20 for experiments, and the results are shown in Figure 2F. It can be seen from the figure that when α is 10, the reconstruction error is the minimum. Therefore, we set the value of α to 10.

3 Results

3.1 Model parameters

Both the low-level feature extraction module and the feature conversion module of MACRN had 256 convolution cores in each convolution layer, while the atrous convolution residual block had 128 in each convolution layer. The images output by the four atrous convolution layers were spliced together to obtain the channel number of the total characteristic image with 512 channels. Then, the number of channels was condensed down to 256 through a convolution layer with 256 convolution cores and a total of six stacked ACRBs.

The batch size for the model training was 16. An Adam optimizer was used for training, with a first-order attenuation index of $\beta_1 = 0.9$, a second-order attenuation index of $\beta_2 = 0.9$, and a fuzzy factor of $\varepsilon = 10^{-8}$. The learning rate was initialized to 0.0001, and the polynomial function attenuation strategy was used to gradually reduce the learning rate during the training process.

3.2 Quantitative analysis of ACRB

To determine how the number of atrous convolution residual blocks (ACRB) affects the hyperspectral image reconstruction model, we constructed five hyperspectral reconstruction networks with different depths with the number of ACRB ranging from four to eight, which were chosen considering the poor feature extraction ability of the low-level network. The test results are shown in Figure 3A.

In Figure 3A, MRAE (Mean Relative Absolute Error) is a commonly used index to evaluate spectral reconstruction effect, and RMSE (Root Mean Square Error) is used as an auxiliary evaluation standard.

The reconstructed MRAE error reached the minimum at seven ACRB (Fig. 3.1). However, there was very little difference between six and seven ACRBs, so in order to minimize the number of parameters and reduce the complexity of the model, the number of stacks for the ACRB is set to 6.

3.3 Clean data reconstruction

In order to evaluate how the MACRN model against to other reconstruction algorithms, we compared it to HSCNN-R (Shi et al., 2018), HSCNN-D (Shi et al., 2018), HRNet (Zhao et al., 2020), AWAN (Li et al., 2020), MST++ (Cai et al., 2022). To ensure the fairness of the experiment, all models were run under the same experimental conditions, using the same training and test sets.

The MACRN model obtained the lowest error value on the Clean data (Table 1). The HSCNN-R model had the shortest running time, but a large error rate. The AWAN model had the second fastest running time, but MACRN was almost as quick. In conclusion, the MACRN performs well in spectral reconstruction of Clean data.

Figure 3B shows the a heat map (error image) of the hyperspectral images reconstructed from the Clean data for the six models and the real hyperspectral image (Ground truth) in the bands of 443 nm, 490 nm, 783nm and 1375nm.

The error of the reconstructed hyperspectral image in the bands of 443 nm, 490 nm and 783nm was small, while the error in the 1375 nm band was large (Figure 3B). Overall, MACRN produced a high quality reconstruction.

Figure 3C shows the spectral response curves drawn at two spatial positions from the hyperspectral images reconstructed by Clean data of five models. The reconstructed spectral curves of MACRN closely matched the real spectral curves, demonstrating its accuracy in spectral reconstruction relative to the other four models (Figure 3C).

10.3389/fmars.2022.1006452

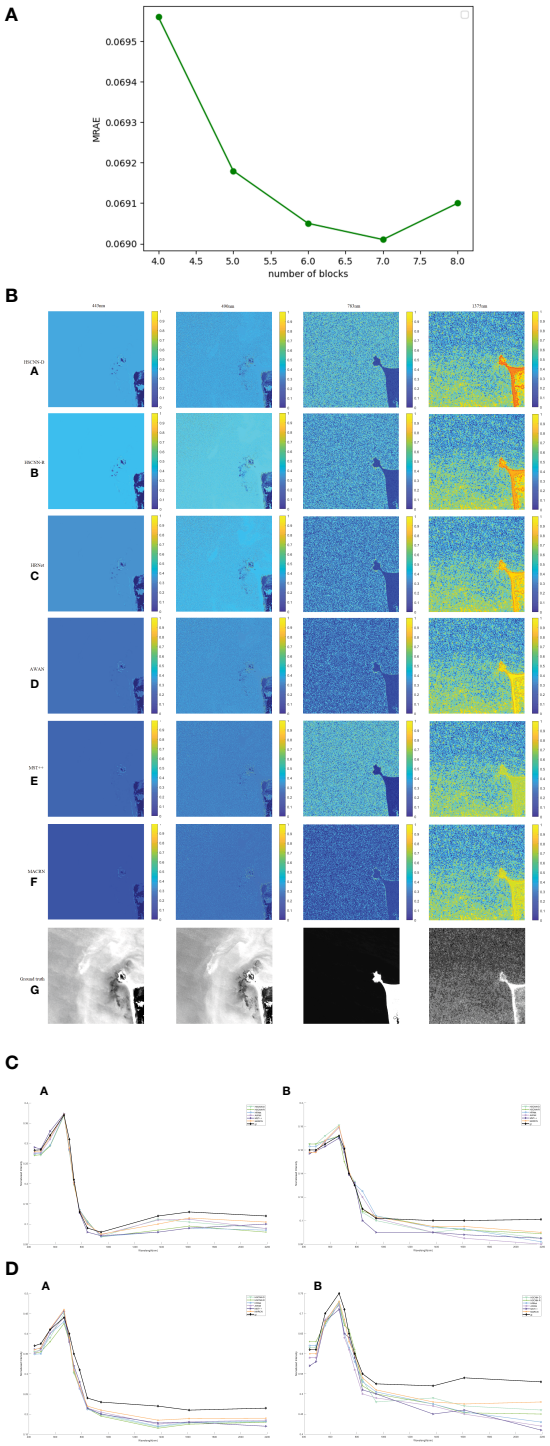


FIGURE 3
(A) Reconstruction error of different ARCB numbers. **(B)** Error image (heat map) of hyperspectral image reconstructions and real hyperspectral images of (from left to right) the 443 nm, 490 nm, 783 nm and 1375 nm bands by the **(a)** HSCNN-D, **(b)** HSCNN-R, **(c)** HRNet, **(d)** AWAN, **(e)** MACRN models run on images from the Clean dataset, compared to **(f)** the true hyperspectral data of the same locations. **(C)** Spectral response curve from the HSCNN-D, HSCNN-R, HRNet, AWAN, MACRN models run on Clean data and true observed data from Sentinel-2 (gt for ground truth) for two points from the images shown in Figure 3.2, **(a)** coordinates (312, 315) and **(b)** (764, 494). **(D)** Spectral response curves from HSCNN-D, HSCNN-R, HRNet, AWAN, MACRN run on Realworld data, and observed data (gt) from two points from Figure 3.2, **(a)** coordinates (312,315) and **(b)** (764,494).

3.4 Realworld data reconstruction

From the Realworld data, MACRN exhibited the smallest errors of the six models (Table 2). HSCNN-R again had the shortest running time, but its reconstruction error was large. The running time of MACRN was close to that of HRNet.

Figure 3D shows the spectral response curves drawn at two spatial positions from the hyperspectral images reconstructed from Realworld data of five models.

All models had higher error rates when run with Realworld data than with Clean data (Fig. 3.4). Compared with other methods, MACRN had the closest fit to the true values (Figure 3D).

The MACRN realizes hyperspectral spectral reconstruction in 443nm - 2190nm band, with less computation and high reconstruction accuracy. On the Realworld dataset, the reconstruction error MARE index is less than 0.0645, the reconstruction error RMSE index is less than 0.0164, and the reconstruction time is less than 9.24S. On the Clean dataset, the reconstruction error MARE index is less than 0.0328, the reconstruction error RMSE index is less than 0.0114, and the reconstruction time is less than 8.58S.

To sum up, for both Clean data and Realworld data sets, the performance of MACRN hyperspectral reconstruction algorithm was relatively stable, and offered good hyperspectral reconstruction relative to other models.

4 Discussion

Tables 1 and 2 show that the calculation cost of MACRN reconstruction algorithm is small. Figures 3C, D show that the error between the reconstructed spectral curve of MACRN and the actual spectral curve is small. In the visible light band, the MACRN reconstruction error has no difference with MST++ (Cai et al., 2022) and AWAN (Li et al., 2020), and is close to the true value. However, in the near infrared band, the MACRN reconstruction accuracy is significantly better than other spectral reconstruction algorithms. Although MST++ (Cai et al., 2022) based on spatial sparsity and spectral self similarity won the first in the RGB spectral reconstruction challenge in 2022. However, MST++ ignored some spatial details and did not consider the near-infrared and mid infrared band features. Therefore, the hyperspectral reconstruction result of MST++ is obviously inferior to those of MACRN in the near infrared and mid infrared bands. This fully shows that ACRB of MACRN captures more spatial details, and the loss function of MACRN is more consistent with the characteristics of the near infrared and mid infrared spectrum.

In the detection application of marine water quality (such as chlorophyll-a, dissolved organic matter, solid suspended solids, etc.), it is easier to find target differences in the near-infrared band, and the reconstruction algorithm that can accurately reconstruct the target's near-infrared band spectrum has a greater application prospect. For example, in marine oil pollution detection (Guga, 2020), because hyperspectral images can obtain a large amount of nearly continuous narrowband spectral information in the visible, near infrared, mid infrared and thermal infrared bands. The hyperspectral images can not only effectively distinguish oil film and water, but also infer the type and time of oil leakage from the different spectral absorption characteristics of offshore oil film. We use the MACRN algorithm to construct hyperspectral images, which can easily and quickly detect and identify objects based on the spectral characteristics. Therefore, the MACRN hyperspectral reconstruction algorithm has the characteristics of low reconstruction cost and high reconstruction accuracy, and its advantages in marine spectral reconstruction are more obvious.

TABLE 1 Comparison of reconstruction results of various models run on the Clean dataset.

Model	MRAE	RMSE	Time (s)
HSCNN-R	0.0385	0.0152	**3.16**
HSCNN-D	0.0375	0.0136	14.39
HRNet	0.0358	0.0140	9.35
AWAN	0.0345	0.0121	7.51
MST++	0.0347	0.0123	11.46
MACRN	**0.0328**	**0.0114**	8.58

The bold numbers represent the best value.

TABLE 2 Comparison of reconstruction results of various models run with the Realworld dataset.

Model	MRAE	RMSE	Time (s)
HSCNN-R	0.0713	0.0188	**3.08**
HSCNN-D	0.0691	0.0185	14.70
HRNet	0.0661	0.0176	9.16
AWAN	0.0666	0.0180	13.94
MST++	0.0665	0.0182	12.73
MACRN	**0.0645**	**0.0164**	9.24

The bold numbers represent the best value.

5 Conclusion

The reconstruction of high-precision hyperspectral images from RGB images is important for expanding the use of hyperspectral analysis in the marine sciences. Currently, most algorithms of reconstructing hyperspectral images from RGB images use convolution neural networks, which is problematic because the receptive field of the convolution kernel is at a small, fixed scale, and thus can only learn the features at a single scale, offers low utilization of information from a global context, and lacks global vision.

We designed a hyperspectral image reconstruction algorithm based on a multi-scale atrous convolution residual network (MACRN), which can better solve one of the problems. The algorithm extracts multi-scale features under different receptive fields, makes full use of the rich colors of RGB images and the spatial structure associated with the colors, and adds richer contextual information, ensuring the integrity of local details. Our hyperspectral reconstruction experiment using Sentinel-2 satellite data from the northern coast of Australia shows that the algorithm can effectively reduce hyperspectral reconstruction errors and deliver hyperspectral image reconstruction of high-precision RGB images with low computational costs.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding author.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Acknowledgments

This research is supported by Sichuan Science and Technology Program (2022NSFSC0020), Chengdu Science and Technology Program (2022-YF09-00019-SN), and the Chengdu Research Base of Giant Panda Breeding (NO. 2020CPB-C09, NO. 2021CPB-B06). We would like to thank Dr. Annalise Elliot at the University of Kansas for her assistance with English language and grammatical editing of the manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Aeschbacher, J., Wu, J., and Timofte, R. (2017). "In defense of shallow learned spectral reconstruction from rgb images," in *International Conference on Computer Vision Workshops*. (Venice, Italy: ICCVW) 471–479.

Agahian, F., Arnirshahi, S. A., and Amirshahi, S. H. (2008). Reconstruction of reflectance spectra using weighted principal component analysis. *Color Res. Appl.* 33 (5), 360–371. doi: 10.1002/col.20431

Alvarezgila, A., Weijer, J., and Garrote, E. (2017). "Adversarial networks for spatial context-aware spectral image reconstruction from rgb," in *2017 IEEE International Conference on Computer Vision Workshops (ICCVW)*. (Venice, Italy: ICCVW). 480–490.

Antonio, R. (2015). "Single image spectral reconstruction for multimedia applications," in *Proceedings of the 23rd ACM international conference on Multimedia*. (Brisbane: ASSOC Computing Machinery, Australia) 251–260.

Arad, B., and Benshahar, O. (2016). "Sparse recovery of hyperspectral signal from natural RGB images," in *European Conference on Computer Vision*. (Amsterdam, Netherlands: Springer-Verlag Berlin) 19–34.

Beletkaia, E., and Pozo, J. (2020). More than meets the eye: Applications enabled by the non-stop development of hyperspectral imaging technology. *PhotonicsViews* 17 (1), 24–26. doi: 10.1002/phvs.202070107

Cai, Y., Lin, J., and Lin, Z. (2022). MST++: Multi-stage spectral-wise transformer for efficient spectral reconstruction. *CVPR*. doi: 10.1109/CVPRW56347.2022.00090

Descour, M., and Dereniak, E. (1995). Computed-tomography imaging spectrometer: experimental calibration and reconstruction results. *Appl. optics* 34 (22), 4817–4826. doi: 10.1364/AO.34.004817

Galliani, S., Lanaras, C., Marmanis, D., Baltsavias, E., and Schindler, K. (2017). Learned spectral super-resolution[EB/OL]. *ArXiv* 03–28. doi: 10.48550/arXiv.1703.09470

Geelen, B., Tack, N., and Lambrechts, A. (2014). Acompact snapshot multispectral imager with a monolithically integrated per-pixel filter mosaic. *Advanced Fabrication Technol. Micro/Nano Optics&Photonics VII Int. Soc. Optics Photonics*, (Osaka, Japan) 661–674. doi: 10.1117/12.2037607

Goel, M., Whitmire, E., Mariakakis, A, Saponas, T. S., Joshi, N., Morris, D., et al. (2015). "Hypercam: hyperspectral imaging for ubiquitous computing applications," in *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (Osaka, Japan) 145–156.

Guga, S. (2020). Extraction of marine oil spill information by hyperspectral mixed pixel decomposition. *Inner Mongolia Sci. Technol. and Economy* 02, 84-89.

Huang, T., Dong, W., and Yuan, X. (2021)Deep Gaussian scale mixture prior for spectral compressive imaging. *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.

Jia, Y., Zheng, Y. Q., Gu, L., Subpa-Asa, A., Lam, A., Sato, Y., Sato, I., et al. (2017). "From rgb to spectrum for natural scenes *via* manifold-based mapping," in *Proceedings of the IEEE International Conference on Computer Vision*. (Venice, Italy: ICCV) 4705–4713.

Kaya, B., Can, Y., and Timofte, R. (2019). "Towards spectral estimation from a single rgb image in the wild," in *2019 IEEE/CVF International Conference on Computer Vision Workshop*. (Seoul, South Korea: ICCVW) 3546–3555.

Li, J. J., Wu, C. X., Song, R, Li, Y. S., and Liu, F. (2020). "Adaptive weighted attention network with camera spectral sensitivity prior for spectral reconstruction from RGB images," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. (ELECTR Network) 1894–1903.

Nguyen, R., Prasad, D., and Brown, M. (2014). "Training-based spectral reconstruction from a single RGB image," in *European Conference on Computer Vision*. (Zurich, Switzerland) 186–201.

Oh, S. W., Brown, M. S., Pollefeys, M., and Kim, S. J. (2016). "Do it yourself hyperspectral imaging with everyday digital cameras," in *Computer Vision and Pattern Recognition*, (Las Vegas Seattle, WA). 1063–1069.

Pengfei, L., Huaici, Z., and Peixuan, Li (2020). Hyperspectral images reconstruction using adversarial networks from single RGB image. *Infrared Laser Eng.* 49 (S1), 143–150. doi: 10.3788/IRLA20200093

Peterson, K. T., Sagan, V., and Sloan, J. J. (2020) Deep learning-based water quality estimation and anomaly detection using landsat-8/Sentinel-2 virtual constellation and cloud computing. *GISCIENCE Remote Sens.* 202, 510–525. doi: 10.1080/15481603.2020.1738061

Qizhong, Z., Endi, Z., Yejian, W., and Farong, G. (2021). Recognition of ocean floor manganese nodules by deep kernel fuzzy c-means clustering of hyperspectral images. *J. Image Graphics* 26 (08).

Shi, Z., Chen, C., Xiong, Z. W., Liu, D., and Wu, F. (2018). "Hscnn+: Advanced cnn-based hyperspectral recovery from rgb images," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*. (Salt Lake City, UT: CVPRW) 1052–1058.

Shoeiby, M., Robles-Kelly, A., Timofte, R., Zhou, R. F., Lahoud, F., Susstrunk, S., et al. (2018a). "Pirm2018 challenge on spectral image super-resolution: Methods and results," in *The European Conference on Computer Vision (ECCV) Workshops*. (Munich, Germany) 56–371.

Shoeiby, M., Robles-Kelly, A., Wei, R., and Timofte, R. (2018b). "Pirm2018 challenge on spectral image super-resolution: Dataset and study," in *The European Conference on Computer Vision (ECCV) Workshops*. (Munich, Germany) 276–287.

Veronez, M. R., Kupssinsku, L. S., and Guimaraes, T. T. (2018). Proposal of a method to determine the correlation between total suspended solids and dissolved organic matter in water bodies from spectral imaging and artificial neural networks. *SENSORS*. doi: 10.3390/s18010159

Xiong, Z. W., Shi, Z., Li, H. Q., Wang, L. Z., Liu, D., Wu, F., et al. (2017). "HSCNN: CNN-based hyperspectral image recovery from spectrally undersampled projections," in *2017 IEEE International Conference on Computer Vision Workshops (ICCVW)*, Venice, Italy. 518–525.

Ying, H. T., Xia, K., and Huang, X. X. (2021). Evaluation of water quality based on UAV images and the IMP-MPP algorithm. *Ecol. Inf.* 61, 101239. doi: 10.1016/j.ecoinf.2021.101239

Zhao, Y., Po, L., and Yan, Q. (2020). "Hierarchical regression network for spectral reconstruction from RGB images," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. (ELECTR Network) 1695–1704.

frontiers | Frontiers in Marine Science

TYPE Original Research
PUBLISHED 02 June 2023
DOI 10.3389/fmars.2023.1059622

OPEN ACCESS

EDITED BY
Junyu He,
Zhejiang University, China

REVIEWED BY
Jianwei Yang,
Nanjing University of Information Science and Technology, China
Lin Xu,
Chengdu University of Traditional Chinese Medicine, China
Dongming Liu,
North University of China, China

*CORRESPONDENCE
Peng Chen
capricorncp@163.com
Zhiwu Liao
liaozhiwu@163.com

RECEIVED 01 October 2022
ACCEPTED 05 May 2023
PUBLISHED 02 June 2023

CITATION
Hu S, Hou R, Liao Z and Chen P (2023) Recognition and location of marine animal sounds using two-stream ConvNet with attention.
Front. Mar. Sci. 10:1059622.
doi: 10.3389/fmars.2023.1059622

Recognition and location of marine animal sounds using two-stream ConvNet with attention

Shaoxiang Hu[1], Rong Hou[2], Zhiwu Liao[3]* and Peng Chen[2]*

[1]School of Automation Engineering, University of Electronic Science and Technology of China, Chengdu, China, [2]Chengdu Research Base of Giant Panda Breeding, Sichuan Key Laboratory of Conservation Biology for Endangered Wildlife, Chengdu, China, [3]Academy of Global Governance and Area Studies, Sichuan Normal University, Chengdu, China

There are abundant resources and many endangered marine animals in the ocean. Using sound to effectively identify and locate them, and estimate their distribution area, has a very important role in the study of the complex diversity of marine animals (Hanny et al., 2013). We design a Two-Stream ConvNet with Attention (TSCA) model, which is a two-stream model combined with attention, in which one branch processes the temporal signal and the other branch processes the frequency domain signal; It makes good use of the characteristics of high time resolution of time domain signal and high recognition rate of frequency domain signal features of sound, and it realizes rapid localization and recognition of sound of marine species. The basic network architecture of the model is YOLO (You Only Look Once) (Joseph et al., 2016). A new loss function focal loss is constructed to strengthen the impact on the tail class of the sample, overcome the problem of data imbalance and avoid over fitting. At the same time, the attention module is constructed to focus on more detailed sound features, so as to improve the noise resistance of the model and achieve high-precision marine species identification and location. In The Watkins Marine Mammal Sound Database, the recognition rate of the algorithm reached 92.04% and the positioning accuracy reached 78.4%.The experimental results show that the algorithm has good robustness, high recognition accuracy and positioning accuracy.

KEYWORDS

voice recognition, location, two-stream ConvNet, YOLO, attention, CMFCC

1 Introduction

In recent years, the development of marine resources has been paid more and more attention by countries all over the world. The sea area is vast and the resource reserve is huge. Marine animals are complex and diverse, and many marine creatures exchange information with sound, such as whales can rely on sound to socialize and locate their prey;

There are many endangered mammals in the sea, and their identification and positioning through sound is an effective means of protection for them. For example, maintaining the number of large whales is crucial to marine ecology, and restoring the number of baleen whales and sperm whales can strengthen the health of the global marine ecological ecosystem. Therefore, using passive acoustic classification and species positioning technology to effectively identify them and estimate their distribution areas has a very important role and research significance in the study of the complex diversity of marine animals. Through this technology, we can reveal the behavior and species density of marine animals. However, due to unknown statistical characteristics and low signal to noise ratio (SNR) conditions, marine mammal voice recognition and localization may be the most challenging task in the field of animal bioacoustics. Therefore, it is a hot spot problem in this field to effectively identify and then estimate the distribution area.

At present, there are two main methods for marine mammal voice recognition: one is classification based on spectral characteristics. (Nanaware et al., 2014) manually classifies the sounds of six marine mammals through the extracted energy and spectral cross-correlation coefficients. (André et al., 2011) identified whether there was a whale voice in the audio by detecting the unique frequency bandwidth of the whale voice, with an accuracy rate of 90%. These methods have good classification effect when there are few marine animals, but it is difficult to distinguish some species with similar sound spectrum. Because marine mammals can emit a variety of different sounds, it is also difficult to identify. Another method is to use machine learning method to classify. This kind of method can identify some sounds that cannot be classified by the spectrum map. (Ibrahim et al., 2016) extracted Mel Frequency Cepstral Coefficients (MFCC) and Discrete Wavelet Transform (DWT) coefficients of whale sounds in the North Atlantic Ocean, and used Support Vector Machine (SVM) to classify their sounds. Experiments show that the method is better than using spectral coefficients to classify their sounds. (González-Hernández et al., 2017) uses 1/6 octave and feedforward neural network to identify eleven species, and each species emits a variety of sounds, such as whistling, shouting and squeaking. The model shows good performance and achieves a classification rate of 90% with low computational cost. (Brown and Smaragdis, 2009) uses MFCC as a feature to classify the sounds of killer whales through hidden Markov model and Gaussian mixture model, achieving 90% recognition rate; (Lu et al., 2021) applies feature fusion method, and uses MFCC, Linear Frequency Cepstral Coefficient (LFCC) and time-domain feature fusion as feature parameters for voice recognition But this method is not effective in low SNR environment; (Mingtuo and Wenyu, 2019) used AlexNet and transfer learning methods to automatically detect and classify killer whales, long fin pilot whales and harp seals with extensive overlapping living areas, and achieved good results, with certain limitations, and the calculation process was cumbersome.

There are three main methods of sound source location: TDE or TDOA, depth learning, and sound energy First of all, the method based on TDE or TDOA is to calculate the time delay or phase difference of sound arriving at different acoustic sensors, and estimate the arrival azimuth of the sound source under the condition that the geometry of the array is known. A typical algorithm is phase transformation (GCC-PHAT) (Pérez-Rubio, 2021; Yoshizawa, 2021), This kind of method is limited by time synchronization, and has poor practicability and high cost.

The method based on depth learning mainly focuses on the sound source localization of supervised learning (Yang et al., 2018; Yangzhou et al., 2019; Jin et al., 2020). The deep neural network is trained by extracting acoustic data features, such as amplitude and phase, collected from different acoustic sensors The cost of this kind of location method system is also very large, which seriously restricts the universality of the location system.

The method based on sound energy is direct and effective (Sheng and Hu, 2004; Dranka and Coelho, 2015; Bo, 2022). However, its effectiveness depends on the propagation attenuation model of sound energy. It is often assumed that the sound energy has a linear attenuation relationship with the propagation distance, which is not true in most real scenes.

It can be seen from the above analysis that these studies are all aimed at the classification of one or several marine mammal audio signals, which cannot fully describe the complex and variable characteristics of sound, and cannot be used to identify more marine mammal species. It is difficult to obtain high recognition accuracy in the case of complex environmental noise. At the same time, these identification algorithms do not provide location information, and cannot estimate the distribution area of marine organisms, nor can they further reveal the behavior and population density of marine animals.

Therefore, we have designed a Two Stream ConvNet with Attention (TSCA) for fast sound recognition and location of marine animal sound features. This model decomposes sound signals into time domain signals and frequency domain signals. One branch of the dual stream network processes time signals and the other branch processes frequency domain signals; The basic network architecture of the two independent streams is YOLO. The model suppresses environmental noise through attention, and realizes information fusion. Finally, the model completes the rapid and accurate identification and location of the sound of marine species.

2 Materials and methods

2.1 Data collection

1) Marine animal sound data set

The voice recognition data set is The Watkins Marine Mammal Sound Database, which contains about 1654 records and 1654 spectrograms, recording more than 31 kinds of marine mammals, and the duration of each sample is 3~20s. 24 filters are selected to extract CMFCC features of sound signals. In order to improve the robustness of identifying sound sources in complex environments, Gaussian noise with different signal-to-noise ratios (SNR) is added to the recording samples to expand the data. SNR is respectively -30db, - 15db, 0db, 15db and 30db.

2) Locate Dataset

Because it is difficult to obtain marine species positioning data sets, we use different human voices to replace the voice of marine

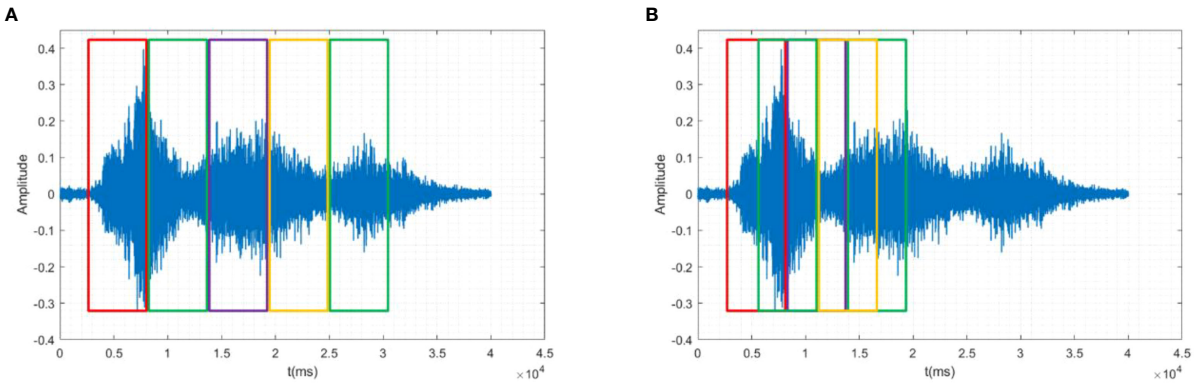


FIGURE 1
Amplitude frequency and phase frequency characteristic diagram of pre emphasis filter, **(A)** Amplitude frequency characteristic diagram, **(B)** Phase frequency characteristic diagram.

species. The voice of people is recorded at a distance of 1km from the three recording sensors. The number of men and women with different timbre is 50, and the sensor spacing is 1m. The number of different recording samples for each person is 30, and the voice in the recording samples is stable, and the voice duration is 5~10s. Considering the existence of various noises in the marine environment, Gaussian noises with different SNR are added to the above recording samples to expand the data, so as to generalize the performance of the system and improve the robustness of identifying sound sources in complex environments. The SNR is respectively - 30db, - 15db, 0db, 15db and 30db. At this time, the number of samples is 3000 × 5 = 15000. Add 5000 groups of soundless source sounding test sets, a total of 20000 test samples.

2.2 Sound processing of marine animals

The sound of marine animals includes static, non static and quasistatic noises (Suleman and Ura, 2007). In order to reduce the influence of background noise, it is necessary to preemphasis the sound signal from the background noise. From the overall perspective, the acoustic signal will change every other period of time, and it is not a steady process. But in a certain period of time, it has a certain degree of stability. This part is usually called a frame. For the global call signal, the final analysis is the time series parameters obtained from each frame.

2.2.1 Preemphasis

Affected by the radiation of the animal's mouth, nose and lip, the call signal collected by the sensor will lose energy, and the high-frequency part (above 800Hz) will fall according to the octave. The high-frequency resolution is low, which affects the recognition results (Yi, 2000). Preemphasis can increase the amplitude of high frequency and reduce the influence of background noise by utilizing the characteristic difference between noise and signal.

In order to compensate for the loss of high-frequency components, the audio signal is pre weighted using Finite Impulse Response (FIR) filter. Formula of the FIR filter $H$is shown in Formula (2.1).

$$H(z) = 1 - az^{-1} \tag{2.1}$$

Where, Z is the transfer function parameter, the value of $a$is close to and slightly less than 1.

When the value of a is 0.98, the amplitude frequency and phase frequency characteristics of the pre emphasis filter are shown in Figure 1. Figure 2 is the Fin whale sound signal graph before and after the pre emphasis processing.

It can be seen from Figure 2 that before pre weighting, the time domain waveform of fin whale calls is relatively discrete, and after pre weighting, it is relatively concentrated and stable. In addition, the low frequency part of the acoustic signal is restrained to some extent, and the middle and high frequency part is effectively improved. The compensation effect of high-frequency loss is good, which is conducive to subsequent feature extraction and recognition.

2.2.2 Framing and windowing

Animal signal is a non-stationary time-varying sound signal, which is easily interfered and affected by vocal production, surrounding environment, vocal tract characteristics and system factors. But in a very short time range (10~30ms), it is considered to be stable. Through a specific window function, the voice signal is windowed and divided into frames for 10~30ms. Each segment is an analysis frame, referred to as a frame for short. Continuous and overlapping are two common framing methods, as shown in Figure 3. To ensure smooth transition between successive frames, we choose overlapping segmentation method. In this algorithm, the sampling frequency of the signal is 22kHz, the frame length is 440 data points, and the corresponding time is 20ms; To ensure the positioning accuracy, 200 points are moved from the frame, and the corresponding time is 10ms.

After framing, in order to make the speech signal globally continuous, and make each frame show some features of periodic function to facilitate subsequent feature extraction, windowing is required. The window function $w(n)$ selected in this paper is Hamming window function, which can be expressed as Eq. (2.2):

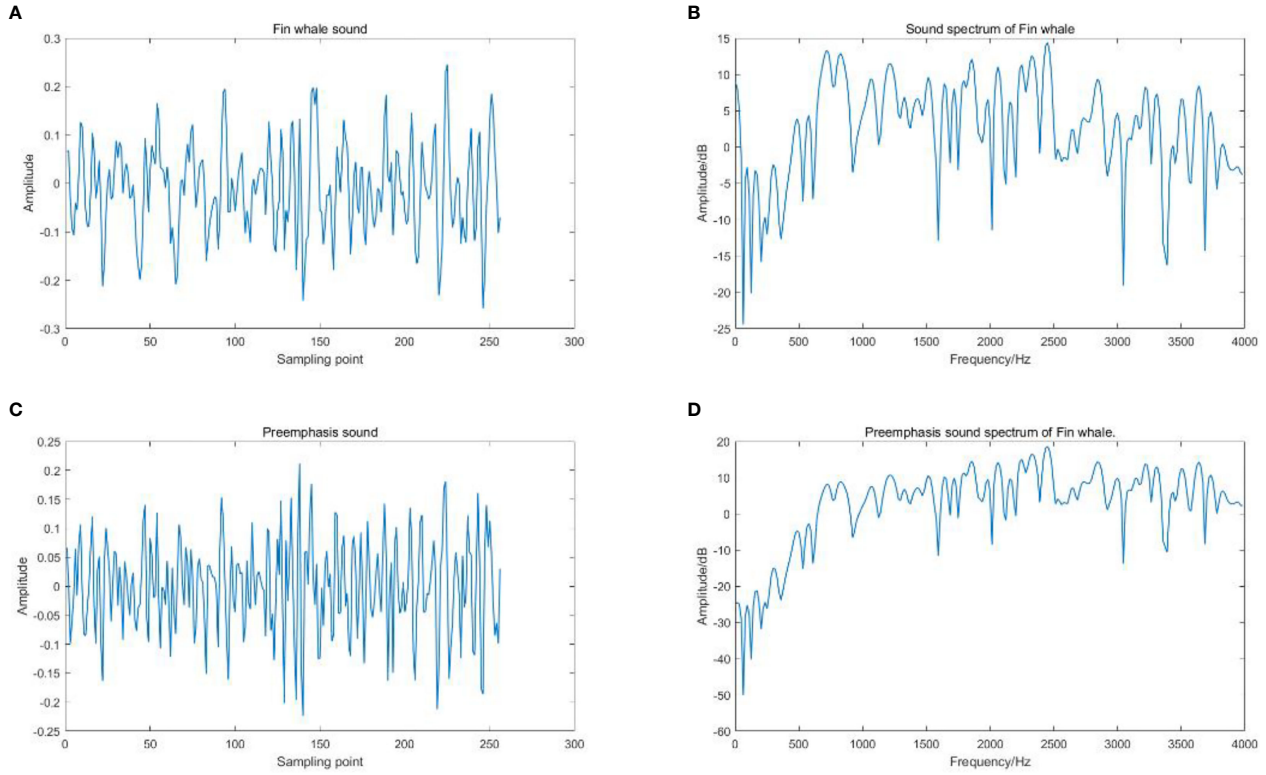


FIGURE 2
The Fin whale sound signal graph before and after the pre emphasis processing, **(A)** Fin whale sound, **(B)** Sound spectrum of Fin whale, **(C)** Preemphasis sound of Fin whale, **(D)** Preemphasis sound spectrum of Fin whale.

$$w(n) = \begin{cases} 0.54 - 0.46\cos\left(\frac{2\pi n}{T-1}\right) & , \quad 0 \ll n \ll T-1 \\ 0 & , \quad other \end{cases} \tag{2.2}$$

Where, T is the length of a frame signal.

The signal before windowing is expressed as $a(n)$, and the signal after windowing is $s(n)$, as shown in Formula 2.3:

$$s(n) = w(n)a(n) \tag{2.3}$$

2.2.3 CMFCC(compound mel-frequency cepstrum coefficient)

CMFCC is Frequency domain characteristics of marine animal sounds. It has been proved that the frequency domain feature CMFCC is the feature with the best recognition rate in voice recognition applications. The extraction method of CMFCC is shown in (Hu et al., 2022). First, the preprocessed signal s (n) is transformed into FFT (Fast Fourier Transformation), its logarithmic energy spectrum is convolved with the filter bank and inversion filter in Meyer frequency domain respectively, and then the output vector is transformed into discrete cosine to obtain CMFCC characteristics. The calculation formula is Eq.2.4. Figure 4 shows the CMFCC characteristics of the orcinus orca's voice.

$$CMFCC(i,j) = \sqrt{\frac{2}{M}}\sum_{m=0}^{m-1}\log[S(i,m)]\cos\left(\frac{\pi j(2m+1)}{2M}\right) \tag{2.4}$$

Where, $CMFCC(i,j)$ is the CMFCC of the i-th frame signal, M is the number of filter banks and inversion filters in the Mel frequency domain, j is the number of CMFCCs, and $S(i,m)$ is the power spectrum in the Mel frequency domain.

2.3 TDOA

The estimation location method based on Time Difference of Arrival (TDOA) has the characteristics of small computation, good real-time, and strong practicability (Benesty et al. 2008). The TDOA method is divided into two steps. First, calculate the time difference (time delay estimation) of the sound source signal arriving at the microphone array, and then establish the sound source location model through the geometry of the microphone array and solve it to obtain the location information (location estimation). As shown in Figure 5, the coordinate axes x, y and z in the figure represent the space distance.

We assume that there is a sound source in the space (denoted as $s(t)$, indicates the position of sound source in space at time t), two microphones (denoted as m1 and m2, their positions in the space are M1 and M2 respectively, and the received signals are $x_1(t)$ and $x_2(t)$. Then the signals received by microphones m1 and m2 are shown in Eq.2.5.

$$\begin{cases} x_1(t) = s(t-\tau_1) + n_1(t) \\ x_2(t) = s(t-\tau_2) + n_2(t) \end{cases} \tag{2.5}$$

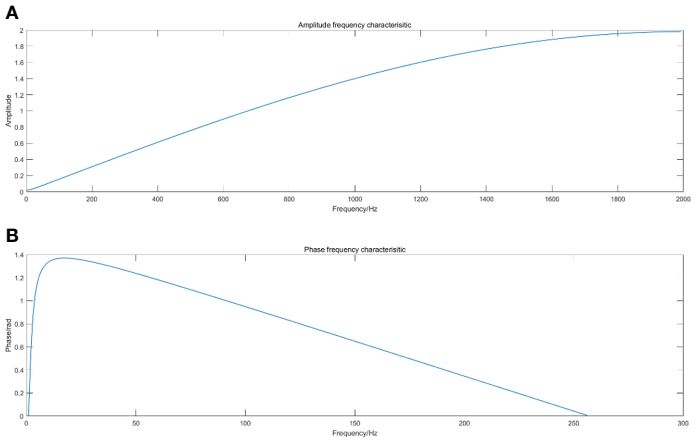


FIGURE 3
Schematic diagram of framing and windowing, **(A)** Continuous framing, **(B)** Overlapping framing.

Where, τ_1 and τ_2 is the delay time for the sound source to reach the two microphones, respectively. $n_1(t)$ and $n_2(t)$ are additive noises. Then the time delay of sound source signal arriving at two microphones is τ,as shown in Eq.2.6.

$$\tau = \tau_1 - \tau_2 \tag{2.6}$$

Here, τ_1 and τ_2 is obtained through the positioning coordinates of TSCA.

We set three microphones to receive the sound signal, and determine the position coordinates of the sound source in the two-dimensional plane. Since the sound of marine animals we detected belongs to the far-field model, the waveform of the sound source arriving at the microphone array is regarded as a plane wave. Therefore, we can calculate the Direction of Arrival (DOA) through the sound signals collected by microphones at two positions. As shown in Figure 6.

According to the geometric relationship of the microphone array, we can determine the angle of the sound source relative to the microphone array θ, as shown in Eq.2.7.

$$\theta = arccos\frac{C\hat{\tau}}{d} \tag{2.7}$$

Where, $\hat{\tau}$ Is the estimated time delay, d is the distance between two sensors, and c is the speed of sound.

Two azimuth angles θ can be obtained by measuring the values of three microphones. The Chan algorithm (Chan and Ho, 1994) solves the position (x, y) of marine animals using Eq.2.8.

$$\begin{bmatrix} x \\ y \end{bmatrix} = -\begin{bmatrix} x_{2,1} & y_{2,1} \\ x_{3,1} & y_{3,1} \end{bmatrix} * \left\{ \begin{bmatrix} R_{2,1} \\ R_{3,1} \end{bmatrix} R_1 + \frac{1}{2} \begin{bmatrix} R_{2,1}^2 - K_2 + K_1 \\ R_{3,1}^2 - K_3 + k_1 \end{bmatrix} \right\} \tag{2.8}$$

Where, $K_1 = x_1^2 + y_1^2$, $K_2 = x_2^2 + y_2^2$, $K_3 = x_3^2 + y_3^2$, $R_1 = \sqrt{(x_1 - x)^2 + (y_1 - y)^2}$, $R_{2,1} = dcos(\theta_1)$, $R_{3,1} = dcos(\theta_2)$, θ_1 is the azimuth angle of the sound source relative to the microphone M1 and M2, θ_2 is the azimuth angle of the sound source relative to the microphone M1 and M3.

2.4 Two-stream ConvNet with attention

Two stream convolutional network appeared in 2014, and it has made considerable progress in the research of action recognition, temporal and spatial behavior detection (Simonyan and Zisserman, 2014).

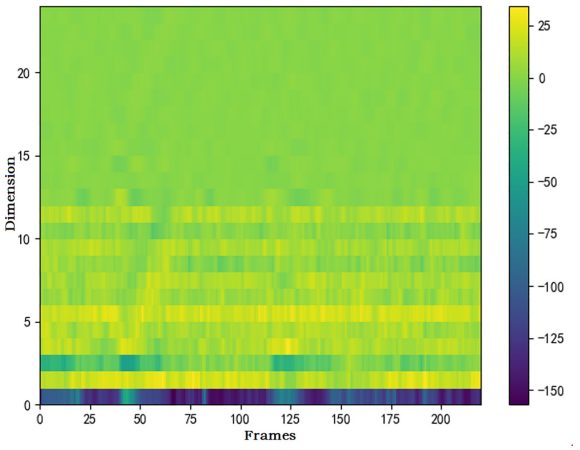


FIGURE 4
The CMFCC characteristics of the orcinus orca's voice.

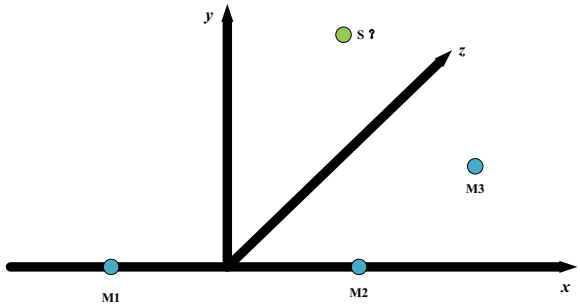


FIGURE 5
The sound source and Microphone array M1, M2, M3.

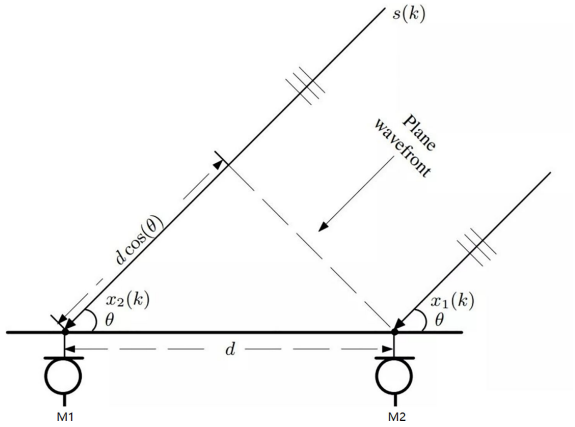


FIGURE 6
Far-field positioning model.

We adopted the idea of Two Stream ConvNet, and designed a Two Stream ConvNet with Attention (TSCA) for fast voice recognition and location of marine animal voice features. The model structure is shown in Figure 7. TSCA uses YOLO net as the basic framework. By embedding time, channel, and space convolution module (TCS), the multi-dimensional network expansion is realized, which improves the model's anti noise ability, and also improves the detection and positioning accuracy of the algorithm.

YOLO net is a fast and high recognition rate target detection and recognition algorithm and network model, which realizes end to end target detection. The anchor box is used to combine the regression problem of classification and target location. It achieves high efficiency, flexibility and good generalization performance. Our goal is to quickly determine the position of marine animals while realizing sound recognition, so we choose YOLO5 as the basic network structure of TSCA.

YOLO net of frequency domain features realizes accurate recognition of marine animal calls through CMTCC feature sequences; YOLO net in time domain uses the high time resolution of call sequence signal to achieve accurate sound location. Since the data frames of the two streams are the same, we fuse the time information of CMFCC channel to YOLO net in the time domain through Time Attention (TAM) after the BackBone of YOLO net to improve the positioning accuracy of YOLO net in the time domain; At the same time, channel and spatial attention are added after the feature layer of the frequency domain feature YOLO net output to improve the anti noise ability of the model.

For the loss function, the two classification cross entropy loss function is fused. Improve the focus loss function to reduce the weight of head class data in the loss function and increase the weight of tail class in the loss function to solve the problem of low accuracy caused by tail data. At the same time, the model combines TDOA module to achieve voice positioning.

2.5 Space time attention module

Attention mechanism has better automatic regulation effect on noise (Ma et al., 2021; Senwei et al., 2021; Zhao et al., 2021). We use attention modules TAM, CAM, SAM (TCS) in the model, as shown in Figure 7. In TCS, channel, time and space information is focused by maximizing and averaging pooling of channels, time and space.

After the BackBone of YOLO net, the time information of CMFCC channel is fused to YOLO net in the time domain through the TAM module. The attention TAM is used to change the weight of each frame time dimension of the time domain YOLO net to achieve dual stream network information fusion, so as to improve the positioning accuracy of YOLO net in the time domain.

In order to better solve the problem of "what" and "where", we draw on the ideas of Convolutional Block Attention Module (CBAM). We use Channel Attention Module (CAM) to focus on "what" in the network; Use the Spatial Attention Module (SAM) to focus on "where" in the network, as shown in Figure 8.

Assume that the input characteristic graph is $F \in R^{C\times H\times W}$, the channel attention diagram is M_c. The spatial attention map is M_s. The formulas of channel attention and spatial attention are as follows Eq.2.9 and Eq.2.10:

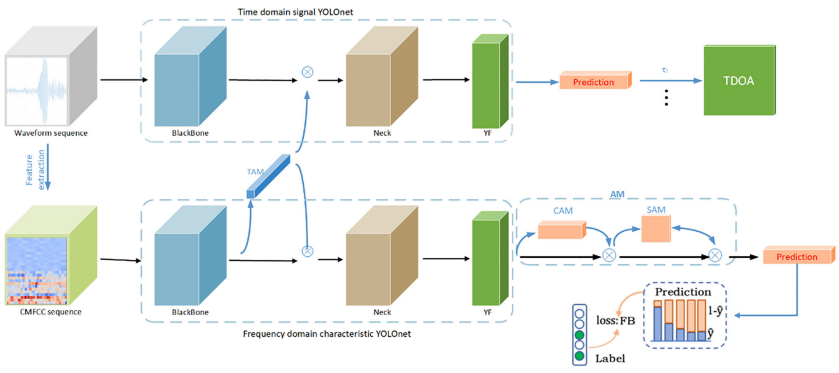


FIGURE 7
TSCA model structure.

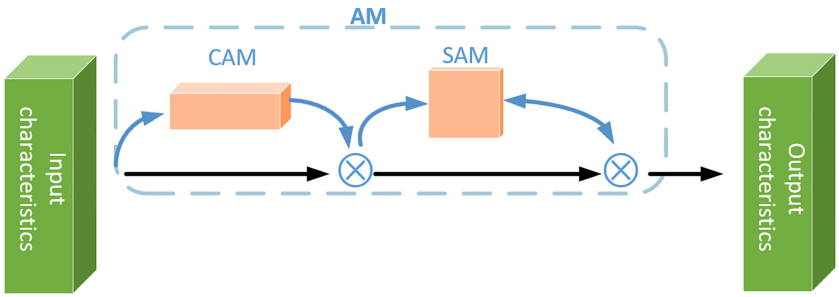


FIGURE 8
Attention module (Channel, Space convolutional block attention module).

$$F' = M_c(F) \otimes F \tag{2.9}$$

$$F'' = M_s(F') \otimes F' \tag{2.10}$$

Where, $M_c \in R^{C\times1\times1}$, $M_s \in R^{1\times H\times W}$, $\otimes$ represents element by element multiplication. In the process of element by element multiplication, attention value will also be propagated to the next level.

2.6 Focus loss function

In the experiment, we found that there was a serious data imbalance problem in the time-space oriented data set of marine animal calls. The distribution of species category tagging data is shown in Figure 9. In the training process of the model, the prediction of the model will be biased towards the head class, making the error rate of the tail class prediction increase.

In classification tasks, the binary cross entropy loss function (BCE Loss) is usually used as the loss function of multi label classification. However, BCE Loss did not take into account the difference between the contributions of the head and tail classes in the long tail data, resulting in low accuracy of model training.

We get a new loss function Focal BCE Loss (FB) by combining the focus loss function (Lin et al., 2017) (Focal Loss) and BCE Loss, as shown in Eq.2.13, to reduce the weight of the head loss function in multi label data and increase the weight of the tail loss function in multi label data.

$$A = -\alpha L(1-\hat{y})^{\gamma} L(\hat{y}) y \tag{2.11}$$

$$B = (1-\alpha)\hat{y}^{\gamma} L(1-\hat{y})(1-y) \tag{2.12}$$

$$FB = A - B \tag{2.13}$$

Where, A represents the Focal Loss, B represents the BCE, L represents the natural logarithmic function, α represents the weighting factor, $\alpha \in [0,1]$, positive class is α, Negative class is 1- α, Y represents the correct label. $\hat{y}$ represents the probability of y=1, $\hat{y} \in (0,1)$. γ is the focus parameter, $(1-\hat{y})^{\gamma}$ is the modulation factor.

When $\hat{y} \to 0$, modulation factor $(1-\hat{y})^{\gamma}$ will be close to 1, so the weight of correct classification will increase.

When $\hat{y} \to 1$, modulation factor $(1-\hat{y})^{\gamma}$ will be close to 0, and the weight for correct classification will decrease.

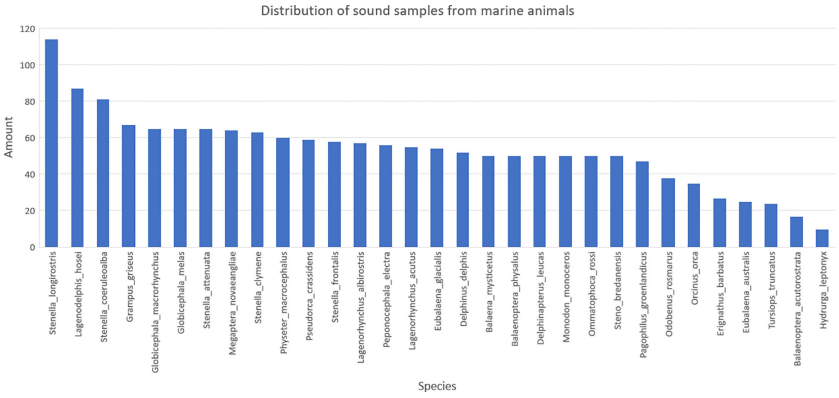


FIGURE 9
The distribution of species category tagging data.

By adjusting the focus parameters γ to reduce the weight of samples that are easy to classify.

When γ= 0, FB is equivalent to BCE Loss. Along with γ Increase of, modulation factor $(1-\hat{y})^{\gamma}$. The influence of, γ= 5. The effect is the best.

With the increase of γ, the influence of modulation factor $(1-\hat{y})^{\gamma}$ will also increase. The experiment found that, γ= 5, the model has the best effect.

3 Results

3.1 Model parameters

Stochastic Gradient Descent (SGD) optimizer is used for model training. Kinetic energy is set to 0.9, weight attenuation is set to 0.00001, initial learning rate is set to 0.075, and epoch is set to 40.

3.2 Loss function analysis

The γ value of the loss function FB is related to the imbalance degree of the data set. In order to analyze the influence of focusing parameter γ in FB on the model detection results, we set different focusing parameter γ, from 0.1 to 10. Here, we use Mean Average Precision (mAP) as the evaluation index to measure the advantages and disadvantages of the algorithm. The experimental result are shown in Figure 10. The solid line represents the reference line, that is, the detection result of the model using BCE loss function, mAP is 83.61%. The dot represents the test result of using FB in the model, and the dotted line is the sixth degree polynomial fitting curve of the dot. The horizontal axis data represents different values of focusing parameter γ from 0.1 to 10. It can be seen that when γ=5, the model detection result reaches 92.04% of the optimal result of γ; $\gamma \in [0,2]$, the model test result mAP increases the fastest. When γ>6, the model test result mAP starts to decline. We can see the effectiveness of the loss function FB in solving long tail data.

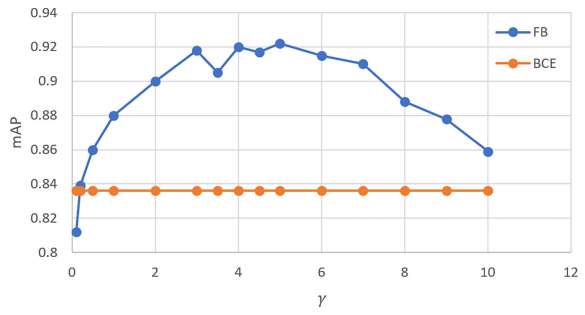


FIGURE 10
mAP of the model when parameter γ of FB takes different values.

3.3 Attention module analysis

In order to verify the effect of spatiotemporal attention module TCS on sound recognition of marine animals, TAM, CAM, SAM, and 3D TCS were added to YOLO net respectively in the experiment. The experimental results of a noise dataset are shown in Table 1. TAM, CAM, SAM, and TCS are represented by T, C, S, and TCS respectively. The line without FB represents the detection results of the above four attention modules added to TSCA respectively. It can be found that after adding TCS, the mAP of the model is the highest. Compared with the benchmark test result of 83.61%, the mAP is increased by 6.43%, followed by the TAM, which is increased by 4.48%.

3.4 Comparison of call classification

1) In order to compare the recognition accuracy of different models under different classifiers, Support Vector Machine (SVM), K Nearest Neighbors (KNN) and Random Forest (RF) are also selected for classification and the results are compared. In order to obtain relatively objective results, 50% of the samples in each experiment are randomly selected as the training set, and the recognition rate is the average of 10 experiments. The identification results of a noise dataset are shown in Table 2.

It can be seen from Table 2 that when using the same feature parameters, the recognition rate of TSCA model is better than the other two classifiers, indicating that TSCA has better recognition performance.

2) Due to the complexity of the marine environment, in order to further verify the robustness of the model in this paper, data with different SNR are tested. The recognition rate is the average of 10 experiments. The recognition results are shown in Figure 11.

It can be seen from Figure 11 that in the process of gradual reduction of the signal-to-noise ratio, the TSCA model has the most gentle decline, which indicates that the TSCA model has good anti noise robustness in the marine environment.

3.5 Positioning

At present, the commonly used technologies for sound source localization products are TDOA, MUSIC and BP neural network sound source localization algorithms. The test samples are self built sample sets. Table 3 summarizes the positioning accuracy of TSCA model and the above algorithms in different environments. Since

TABLE 1 Model detection results of various attention modules (mAP).

	T	C	S	TCS
no FB	85.53%	84.86%	85.18%	86.12%
FB γ =1	88.78%	89.70%	87.86	91.04%
FB γ =5	90.01%	90.24%	89.30%	92.04%

TABLE 2 Classification results of different classifiers.

Data	Classifier	Recognition rate
The Watkins Marine Mammal Sound Database	SVM	89.72%
	RF	84.30%
	KNN	83.24%
	TSCA	92.04%

TABLE 3 The accuracy comparison of algorithms in different environments.

Method	SNR(dB)		
	0	15	30
TDOA	50.3	68.5	79.7
MUSIC	51.3	70.6	80.3
BP	47.8	53.7	70.5
TSCA	52.2	70.8	78.4

the positioning accuracy is less than 20% when the SNR is negative, it is not indicated in the table. The test set used for verification is 5000 in total, including 2000 groups of test samples of silent source sound.

4 Discussion

From Table 2 and Figure 11, we can see that the TSCA model introduces the spatiotemporal attention module TCS and the FB loss function to solve the problem of data heterogeneity (long tail phenomenon). The recognition rate of TSCA model is better than other algorithms in the case of no noise and noise. It shows that the TSCA model can achieve good recognition results in the recognition of the call of marine species.

According to the data in Table 3, although the maximum positioning accuracy of MUSIC and TDOA algorithms is 80.3% and 79.7% respectively. However, the minimum positioning accuracy of both is less than 52%, which is a very significant gap compared with the optimal effect. The reason is that it is difficult to set a threshold parameter for both algorithms, so the peak value of the estimator will shift slightly at the correct position under different conditions. Therefore, the algorithm may not have the best localization accuracy for sound sources with different SNR.

Our TSCA model has higher positioning accuracy than the common BP algorithm under any same conditions. When TSCA is at 0db or 15db, the positioning accuracy is significantly higher than other algorithms. It shows that the TSCA model has good anti-noise ability and can achieve high positioning accuracy in marine species location.

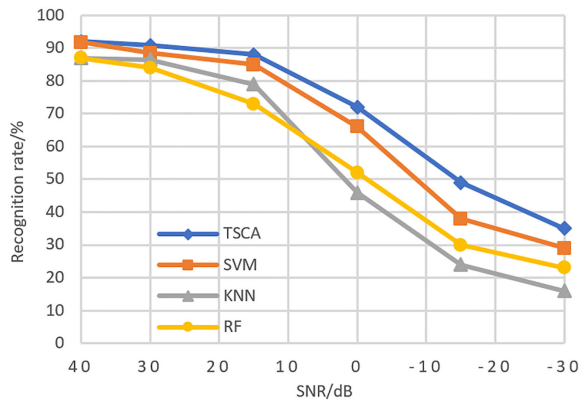


FIGURE 11
Recognition rate of different models under different SNR.

5 Conclusion

The recognition and research of marine organisms' sound is a very important and meaningful work. By recognizing and tracking marine animal targets through sound, it is possible to accurately grasp the distribution, growth status, and behavioral characteristics of marine animals. There are many endangered marine animals in the sea, which can be effectively identified by sound and their distribution areas can be estimated, which plays a very important role in the study of the complex diversity of marine animals. At present, the recognition and classification technology of marine animal sounds largely depends on acoustic characteristics, such as MFCC, LPC, STFT, etc., and classification models, such as GMM, SVM, etc. These recognition technologies can not fully describe the complex and variable characteristics of sound, and often perform poorly. Most of these recognition technologies are used to classify one or several marine animal audio signals. It is impossible to recognize more marine mammal species, and it is impossible to judge the position and distance of marine organisms based on sound.

Therefore, we design a TSCA model to deal with such problems. This model is a dual stream model algorithm based on YOLO net and attention. The algorithm realizes fast localization and recognition of marine species sound from the time high resolution of time domain signal features and the high classification accuracy of frequency domain features. The model uses the loss function FB to strengthen the influence on the tail class of the sample, improve the data imbalance, avoid the over fitting problem, and achieve good results; At the same time, it embeds TAM to achieve dual stream integration; Through TAM, SAM and CAM attention modules, the model can pay attention to more detailed sound features, improve the noise resistance of the model, and achieve high-precision sound recognition and positioning of marine species. The experimental results show that the algorithm has good performance and great practical application potential.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding authors.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Funding

This research is supported by Chengdu Science and Technology Program (2022-YF09-00019-SN), Sichuan Science and Technology Program (2022NSFSC0020, 2023NSFSC1926), and the Chengdu Research Base of Giant Panda Breeding (NO. 2020CPB-C09, NO. 2021CPB-B06).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

André, M., van der Schaar, M., Zaugg, S., Houegnigan, L., Sanchez, A. M., Castell, J. V., et al. (2011). Listening to the deep: live monitoring of ocean noise and cetacean acoustic signals. *Mar. Pollut. Bull.* 63 (1-4), 18–26. doi: 10.1016/j.marpolbul.2011.04.038

Benesty, J, Chen, J, and Huang, Y (2008). *Micrphone array processing.* Berlin, Heidelberg

Bo, X. (2022). *Single sound source localization and tracking technology based on sound energy* (Guangzhou University) Guangzhou, China.

Brown, J. C., and Smaragdis, P. (2009). Hidden Markov and Gaussian mixture models for automatic call classification. *J. Acoustical Soc. America* 125 (6), EL221. doi: 10.1121/1.3124659

Chan, Y. T., and Ho, K. C. A. (1994)Simple and efficient estimator for hyperbolic location. signal processing (Accessed IEEE Transactions).

Dranka, E., and Coelho, R. F. (2015). Robust maximum likelihood acoustic energy based source localization in correlated noisy sensing environments. *IEEE J. Selected Topics Signal Process.* 9 (2), 259–267. doi: 10.1109/JSTSP.2014.2385657

González-Hernández, F. Rubén, Sánchez-Fernández, L. P., Suárez-Guerra, S., and Sánchez-Pérez, L. A. (2017). Marine mammal sound classification based on a parallel recognition model and octave analysis. *Appl. Acoustics.* 119, 17–28. doi: 10.1016/j.apacoust.2016.11.016

Hanny, D. E., Delarue, J., Mouy, X., Martin, B. S., Leary, D., Oswald, J. N., et al. (2013). Marine mammal acoustic detections in the northeastern chukchi Sea, September 2007-July 2011. *Continental Shelf Res.* 67, 127–146. doi: 10.1016/j.csr.2013.07.009

Hu, S. X., Liao, Z. W., Hou, R., and Chen, P. (2022) Characteristic sequence analysis of giant panda voiceprint. *Front. Phys.* 10, 839699. doi: 10.3389/fphy.2022.839699

Ibrahim, A. K., Zhuang, H., Erdol, N., et al. (2016)A new approach for north atlantic right whale upcall detection[C] (Accessed Xi'an: International Symposium on Computer, Consumer and Control).

Jin, L., Yan, J., Du, X., Xiao, X. C., and Fu, D. Y. (2020). RNN for solving time-variant generalized Sylvester equation with applications to robots and acoustic source localization. *IEEE Trans. Ind. Inf.* 16 (10), 6359–6369. doi: 10.1109/TII.2020.2964817

Joseph, R., Santosh, D., Ross, G., and Ali, F. (2016). *You only look once: unified, real-time object detection* (Las Vegas, NV, USA: CVPR), 779-788. doi: 10.48550/arXiv.1506.02640

Lin, T. Y., Goyal, P., Girshick, R., et al. (2017)Focal loss for dense object detection (Accessed Proceedings of the IEEE international conference on computer vision).

Lu, T., Han, B., and Yu, F. (2021). Detection and classification of marine mammal sounds using AlexNet with transfer learning. *Ecol. Inf.* 62. doi: 10.1016/j.ecoinf.2021.101277

Ma, H. W., Liu, G. C., and Yuan, Y. (2021)Enhanced non-local cascading network with attention mechanism for hyperspectral image denoising (Accessed International Conference on Acoustics Speech and Signal Processing ICASSP).

Mingtuo, ZHONG, and Wenyu, C. A. I. (2019). "Marine mammal sound recognition based on feature fusion," in *Electronic sci. & tech.* (Electronic Science and Technology: Xi An, China) 30

Nanaware, S., Shastri, R., Joshi, Y., et al. (2014)Passive acoustic detection and classification of marine mammal vocalizations[C], lucknow (Accessed International Conference on Communication and Signal Processing).

Pérez-Rubio, M. C. (2021). Dynamic adjustment of weighted GCC-PHAT for position estimation in an ultrasonic local positioning system. *Sensors* 21 (21), 7051. doi: 10.3390/s21217051

Senwei, L., ZhongZhen, H., Mingfu, L., and Haizhao, Y. (2021)Instance enhancement batch normalization: an adaptive regulator batch noise (Accessed AAAI Conference on Artificial Intelligence).

Sheng, X., and Hu, Y. H. (2004). Maximum likelihood multiple-source localization using acoustic energy measurements with wireless sensor networks. *IEEE Trans. Signal Process.* 53 (1), 44–53. doi: 10.1109/TSP.2004.838930

Simonyan, K., and Zisserman, A. (2014). "Two-stream convolutional networks for action recognition in videos," in *Advances in neural information processing systems.* NIPS 2014: Montreal, CANADA, 27 27.

Springer (2008). *Micrphone array processing.*

Suleman, M., and Ura, T. (2007)Vocalization based individual classification of humpback whales using support vector machine (Accessed Oceans).

Yang, X., Li, Y., Sun, Y., et al. (2018)Fast and robust RBF neural network based on global K- means clustering with adaptive selection radius for sound source angle estimation (Accessed IEEE Transactions on Antennas and Propagation).

Yangzhou, J., Ma, Z., and Huang, X. (2019). A deep neural network approach to acoustic source localization in a shallow water tank experiment. *J. Acoustical Soc. America* 146 (6), 4802–4811. doi: 10.1121/1.5138596

Yi, K. (2000). *Speech signal processing* (Beijing, China: National Defense Industry Press), 10–70.

Yoshizawa, S. (2021). Underwater acoustic localization based on IR-GCC-PHAT in reverberant environments. *Int. J. Circuits* 15, 164–171. doi: 10.46300/9106.2021.15.18

Zhao, Y., Zhai, D., Jiang, J., and Liu, X. (2021). *ADRN: attention-based deep residual network for hyperspectral image denoising* (Barcelona, Spain: International Conference on Acoustics Speech and Signal Processing ICASSP), 2668–2672.

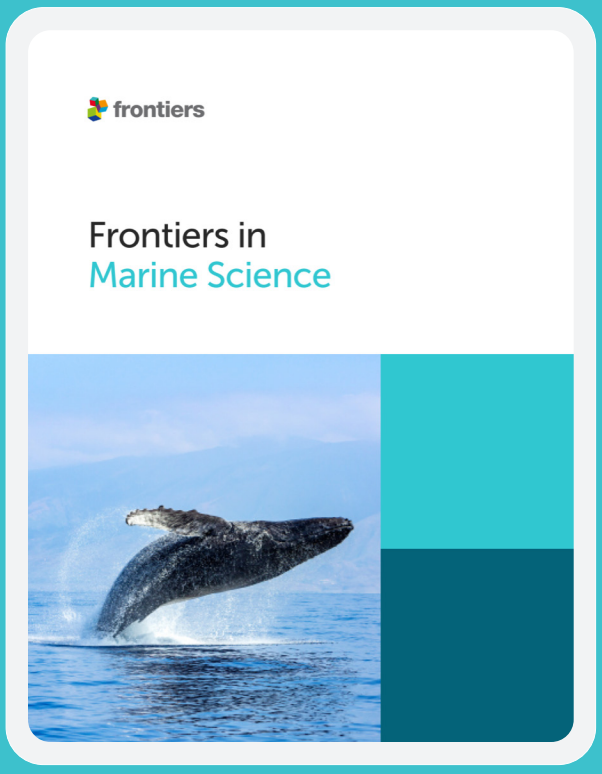
Frontiers in Marine Science

Explores ocean-based solutions for emerging global challenges

The third most-cited marine and freshwater biology journal, advancing our understanding of marine systems and addressing global challenges including overfishing, pollution, and climate change.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact