BIG EARTH DATA INTELLIGENCE FOR ENVIRONMENTAL MODELING

EDITED BY: Peng Liu, Zhe Zhang, Mohd Anul Haq and Yizhe Zhan
PUBLISHED IN: Frontiers in Environmental Science and
Frontiers in Ecology and Evolution

frontiers Research Topics

ISSN 1664-8714
ISBN 978-2-88976-292-7
DOI 10.3389/978-2-88976-292-7

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

BIG EARTH DATA INTELLIGENCE FOR ENVIRONMENTAL MODELING

Topic Editors:
Peng Liu, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences (CAS), China
Zhe Zhang, Texas A&M University, United States
Mohd Anul Haq, Majmaah University, Saudi Arabia
Yizhe Zhan, MetService Wellington, New Zealand

Citation: Liu, P., Zhang, Z., Haq, M. A., Zhan, Y., eds. (2022). Big Earth Data Intelligence for Environmental Modeling. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-88976-292-7

Table of Contents

frontiers | Frontiers in Environmental Science

EDITORIAL
published: 10 May 2022
doi: 10.3389/fenvs.2022.917590

Editorial: Big Earth Data Intelligence for Environmental Modeling

Peng Liu[1]*, *Mohd Anul Haq*[2] *and Yizhe Zhan*[3]

[1]Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, China, [2]Department of Computer Science College of Computer and Information Sciences, Majmaah University, Al Majma'ah, Saudi Arabia, [3]MetService, Wellington, New Zealand

Keywords: Big Earth Data, artificial intelligence, environmental modeling, earth system, remote sensing of environment

Editorial on the Research Topic

Big Earth Data Intelligence for Environmental Modeling

Environmental modeling involves the processes of analyzing the interactions between geophysical, biological, economic, and social systems (Ma et al., 2018; Xu et al., 2020; Zhang et al., 2020; Li et al., 2021). Environmental modeling can be used to understand environmental systems, or for performing interdisciplinary analysis that can support decisions and policy. Model-driven methods have remained dominant, especially for traditional and classical problems that deal with Big Earth Data that come from various data sources with different formats and scales. With recent breakthroughs in data acquisition technology, Big Earth Data can provide globally established, multi-source, multi-scale, high-dimensional, heterogeneous, high-resolution, and highly dynamic datasets. Big Earth Data, with its emerging technologies and approaches, may be an evolution in earth observation, and it also may allow us to use and apply them to various domains in new ways (Sudmanns et al., 2020; Zhang L et al., 2021). At the same time, artificial intelligence is rapidly progressing thanks to advanced algorithm architectures (He et al., 2016), powerful computing devices, and large available datasets (Deng et al., 2009; Liu et al., 2022). The convergence between Big Earth Data and artificial intelligence could open a new era for the advance of environmental modeling. Big Earth Data Intelligence provides new opportunities to understand the environmental modeling of earth systems which helps to resolve problems, such as spatiotemporal complement, data assimilation and fusion (Armstrong et al., 2019; Liu et al., 2021; Zhang L et al., 2021; Zhang Z et al., 2021), uncertainty (Zhang et al., 2014; Zhang et al., 2018), and model calibration. It also promotes development and presents new challenges for remote sensing of the environment (Yuan et al., 2020).

With this topic on Big Earth Data Intelligence in environmental modeling, we try to introduce the latest theory and methods of applying Big Earth Data to environmental science. It contains 14 papers that demonstrate the latest research to advance the science in research areas such as forest, ocean, ecosystem, coast, water, and weather information.

Forest mapping and modeling is an important research area for understanding environmental modeling. The forest cover change in Africa was analyzed by a data-driven method by Xiao et al. and the driving effects of the population, economy, and cultivated land expansion, were analyzed using geographical spatial heterogeneity. Urban plantation tree detection is another very challenging problem because of the irregularity in the shape of a tree canopy. The authors make an attempt to use single shot multibox (Zheng et al.) and apply YOLOv4 (Zheng et al.) for tree detection with high-resolution remote sensing imagery. To satisfy the temporal requirement in forest active fires, Hong et al. use a novel convolutional neural network to learn and search for fire spot features in Himawari-8 satellite images.

The ocean attracts much attention from researchers for its unique role in global change. As an example of the breadth of subjects covered, manuscripts included the attempt by Xiao et al. to generate sea surface temperature products by fusion of microwave and infrared data. By procuring global concentrations of

Edited and reviewed by:
Alexander Kokhanovsky, Telespazio Belgium, Germany

***Correspondence:**
Peng Liu
liupeng@radi.ac.cn

Specialty section:
This article was submitted to Environmental Informatics and Remote Sensing, a section of the journal Frontiers in Environmental Science

Received: *11 April 2022*
Accepted: *20 April 2022*
Published: *10 May 2022*

Citation:
Liu P, Haq MA and Zhan Y (2022) Editorial: Big Earth Data Intelligence for Environmental Modeling. Front. Environ. Sci. 10:917590. doi: 10.3389/fenvs.2022.917590

particulate carbon and nitrogen in different depths, and by analyzing particulate organic carbon-to-nitrogen ratio variations, Xiao et al. pointed out that ocean currents will have a more extensive and profound impact on the ocean carbon and nitrogen cycle. By investigating the physiological responses to different light and nitrogen levels of sargassum muticum, Yan et al. showed that higher light and/or nitrogen levels generally promoted the photosynthesis and growth of the algae. In the study on the coast, Yang et al. use the new deep architecture of vision transformers to extract the waterline for an artificial coast, and Xiao et al. construct a comprehensive platform based on remote sensing images to monitor coastal tidal mudflat ecological development. Different from the ocean, groundwater plays a role in the ecosystem in its own way. Liang et al. carry out research on water table depth prediction in Baoding city, North China Plain by combining wavelet transform with a long short-term memory neural network.

Big Earth Data are guiding us to take an interest in environmental modeling in a new way. The grassland ecosystem was modeled by an optimality-based spatial explicit ecohydrological model at watershed scale by Chen et al., and urban ecology was evaluated by high-resolution remote-sensing data (Huang et al.). Furthermore, remote sensing data such as temporal-spatial data (Ma et al.) and weather radar data (Lu et al.) are good application scenarios for using new artificial intelligence methods and give more support for the study of environmental modeling.

With these issues in mind, we try to present the current state-of-the-art theoretical, methodological, and application research on environmental modeling using Big Earth Data Intelligence. Most of the manuscripts are related to the topic of environmental modeling and geospatial big data analytics. We hope that these high-quality contributions will inspire the readers in their research.

AUTHOR CONTRIBUTIONS

PL wrote the first draft. All authors revised the article and approved the submitted version.

FUNDING

This work was supported by NSFC (Nos. 61731022, 41971397).

ACKNOWLEDGMENTS

We are very grateful to all our colleagues who submitted, reviewed, and edited manuscripts for this Research Topic.

REFERENCES

Armstrong, M. P., Wang, S., and Zhang, Z. (2019). The Internet of Things and Fast Data Streams: Prospects for Geospatial Data Science in Emerging Information Ecosystems. *Cartogr. Geogr. Inf. Sci.* 46 (1), 39–56. doi:10.1080/15230406.2018.1503973

Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., and Fei-Fei, L. (2009). "June. ImageNet: A Large-Scale Hierarchical Image Database," in 2009 IEEE conference on computer vision and pattern recognition (IEEE), 248–255.

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, 770–778. doi:10.1109/cvpr.2016.90

Li, Y., Gao, H., Allen, G. H., and Zhang, Z. (2021). Constructing Reservoir Area-Volume-Elevation Curve from TanDEM-X DEM Data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 14, 2249–2257. doi:10.1109/jstars.2021.3051103

Liu, P., Li, J., Wang, L., and He, G. (2022). "A Review on Remote Sensing Data Fusion with Generative Adversarial Networks (GAN)," in IEEE Geoscience and Remote Sensing Magazine, Just Accepted (March 2022) (IEEE).

Liu, P., Wang, L., Ranjan, R., He, G., and Zhao, L. (2021). A Survey on Active Deep Learning: from Model-Driven to Data-Driven. *ACM Comput. Surv.* 1, 1. doi:10.1145/3510414

Ma, C., Yang, J., Chen, F., Ma, Y., Liu, J., Li, X., et al. (2018). Assessing Heavy Industrial Heat Source Distribution in China Using Real-Time VIIRS Active Fire/hotspot Data. *Sustainability* 10 (12), 4419. doi:10.3390/su10124419

Sudmanns, M., Tiede, D., Lang, S., Bergstedt, H., Trost, G., Augustin, H., et al. (2020). Big Earth Data: Disruptive Changes in Earth Observation Data Management and Analysis? *Int. J. Digital Earth* 13 (7), 832–850. doi:10.1080/17538947.2019.1585976

Xu, B., Li, Y., Han, F., Zheng, Y., Ding, W., Zhang, C., et al. (2020). The Transborder Flux of Phosphorus in the Lancang-Mekong River Basin: Magnitude, Patterns and Impacts from the Cascade Hydropower Dams in China. *J. Hydrology* 590, 125201. doi:10.1016/j.jhydrol.2020.125201

Yuan, Q., Shen, H., Li, T., Li, Z., Li, S., Jiang, Y., et al. (2020). Deep Learning in Environmental Remote Sensing: Achievements and Challenges. *Remote Sens. Environ.* 241, 111716. doi:10.1016/j.rse.2020.111716

Zhang, L., Liu, P., Zhao, L., Wang, G., Zhang, W., and Liu, J. (2021). Air Quality Predictions with a Semi-supervised Bidirectional LSTM Neural Network. *Atmos. Pollut. Res.* 12 (1), 328–339. doi:10.1016/j.apr.2020.09.003

Zhang, Z., Demšar, U., Rantala, J., and Virrantaus, K. (2014). A Fuzzy Multiple-Attribute Decision-Making Modelling for Vulnerability Analysis on the Basis of Population Information for Disaster Management. *Int. J. Geogr. Inf. Sci.* 28 (9), 1922–1939. doi:10.1080/13658816.2014.908472

Zhang, Z., Demšar, U., Wang, S., and Virrantaus, K. (2018). A Spatial Fuzzy Influence Diagram for Modelling Spatial Objects' Dependencies: a Case Study on Tree-Related Electric Outages. *Int. J. Geogr. Inf. Sci.* 32 (2), 349–366. doi:10.1080/13658816.2017.1385789

Zhang, Z., Laakso, T., Wang, Z., Pulkkinen, S., Ahopelto, S., Virrantaus, K., et al. (2020). Comparative Study of AI-Based Methods-Application of Analyzing Inflow and Infiltration in Sanitary Sewer Subcatchments. *Sustainability* 12 (15), 6254. doi:10.3390/su12156254

Zhang, Z., Wang, Z., Li, A., Ye, X., Usery, E. L., and Li, D. (2021). "November). An AI-Based Spatial Knowledge Graph for Enhancing Spatial Data and Knowledge Search and Discovery," in Proceedings of the 1st ACM SIGSPATIAL International Workshop on Searching and Mining Large Collections of Geospatial Data, 13–17.

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.

frontiers in Environmental Science

ORIGINAL RESEARCH
published: 04 August 2021
doi: 10.3389/fenvs.2021.714067

Lightning Strike Location Identification Based on 3D Weather Radar Data

Mingyue Lu[1], Yadong Zhang[1], Zaiyang Ma[2,3,4], Manzhu Yu[5], Min Chen[2,3,4], Jianqin Zheng[6] and Menglong Wang[1]*

[1]Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters, Nanjing University of Information Science and Technology, Nanjing, China, [2]Key Laboratory of Virtual Geographic Environment (Ministry of Education of PRC), Nanjing Normal University, Nanjing, China, [3]State Key Laboratory Cultivation Base of Geographical Environment Evolution (Jiangsu Province), Nanjing, China, [4]Jiangsu Center for Collaborative Innovation in Geographical Information Resource Development and Application, Nanjing, China, [5]Department of Geography, The Pennsylvania State University, University Park, University Park, IL, United States, [6]Wenzhou Meteorological Bureau, Wenzhou, China

Lightning is an instantaneous, intense, and convective weather phenomenon that can produce great destructive power and easily cause serious economic losses and casualties. It always occurs in convective storms with small spatial scales and short life cycles. Weather radar is one of the best operational instruments that can monitor the detailed 3D structures of convective storms at high spatial and temporal resolutions. Thus, extracting the features related to lightning automatically from 3D weather radar data to identify lightning strike locations would significantly benefit future lightning predictions. This article makes a bold attempt to apply three-dimensional radar data to identify lightning strike locations, thereby laying the foundation for the subsequent accurate and real-time prediction of lightning locations. First, that issue is transformed into a binary classification problem. Then, a suitable dataset for the recognition of lightning strike locations based on 3D radar data is constructed for system training and evaluation purposes. Furthermore, the machine learning methods of a convolutional neural network, logistic regression, a random forest, and k-nearest neighbors are employed to carry out experiments. The results show that the convolutional neural network has the best performance in identifying lightning strike locations. This technique is followed by the random forest and k-nearest neighbors, and the logistic regression produces the worst manifestation.

Keywords: lightning strike location, identification, convolutional neural network, 3D weather radar, machine learning

OPEN ACCESS

Edited by:
Peng Liu,
Institute of Remote Sensing and Digital Earth (CAS), China

Reviewed by:
Jan Szturc,
Institute of Meteorology and Water Management - National Research Institute, Poland
Ivica Petrović,
Josip Juraj Strossmayer University of Osijek, Croatia

****Correspondence:***
Zaiyang Ma
zy_ma327@126.com

Specialty section:
This article was submitted to Environmental Informatics and Remote Sensing, a section of the journal Frontiers in Environmental Science

Received: *24 May 2021*
Accepted: *26 July 2021*
Published: *04 August 2021*

Citation:
Lu M, Zhang Y, Ma Z, Yu M, Chen M, Zheng J and Wang M (2021) Lightning Strike Location Identification Based on 3D Weather Radar Data. Front. Environ. Sci. 9:714067. doi: 10.3389/fenvs.2021.714067

INSTRUCTION

Lightning is a spark of electricity in the atmosphere between clouds, the air, or the ground (Maggio et al., 2009). Its high voltage, high temperature, and other physical effects can produce great destructive power in an instant, which is prone to damaging the personal safety of ground personnel. In particular, lightning can easily cause damage to commercial buildings, electrical equipment, homes. Due to the rapid economic development and the massive increase in the amount of electrical equipment, lightning disasters have become one of the ten most serious natural disasters announced by the International Decade for Natural Disaster Reduction (NOAA National Severe Storms Laboratory, 2021). Moreover, lightning is a small-scale strong convective weather phenomenon, which makes it difficult to predict accurately. Since lightning always occurs in convective storms with

small spatial scales and short life cycles, weather radar is one of the best operational instruments that can monitor the detailed 3D (Three dimensional) structures of such storms at high spatial and temporal resolutions. Thus, extracting the features related to lightning automatically from 3D weather radar data to identify lightning strike locations would significantly benefit future lightning predictions.

The current research on lightning can be roughly divided into two categories: numerical analysis and statistical methods. Numerical analysis is a mesoscale numerical weather prediction system that mainly uses the Weather Research and Forecasting (WRF) Model and other models to simulate strong convection processes and discusses the influence of different physical parameterization schemes on the simulation effect (NCAR Mesoscale & Microscale Meteorology Laboratory, 2021). Barthe et al. (2010) estimated total lightning by the WRF model according to the correlation between the flash speed and the available model parameters (e.g., ice water path, ice mass flux product). Zepka et al. (2014) used the WRF model to study potential lightning locations based on the probability of lightning occurrence within the location of interest. Giannaros et al. (2016) used lightning to improve convective representations by controlling the triggers of model convection parameterization schemes on the basis of WRF-LTNGDA. Gharaylou et al. (2020) used the WRF-ELEC model to predict the impact of initial conditions on lightning activity. However, the coarse temporal and spatial resolutions used in numerical analysis method limit their utility in accurate lightning strike location recognition tasks.

Statistical methods are usually applied to study lightning based on prior knowledge. Combining the density-based spatial clustering of applications with noise (DBSCAN) algorithm and the kernel density algorithm, Chen et al. (2017) eliminated insignificant locations with rare lightning strikes. The National Oceanic and Atmospheric Administration (NOAA) and the Cooperative Institute for Meteorological Satellite Studies (CIMSS) developed the empirical probability of severe (ProbSevere) model in 2018. It extracts information related to thunderstorm development from several data sources automatically to produce timely, short-term, statistical forecasts of thunderstorm intensity (Cintineo et al., 2018). Wang et al. (2019) established a probabilistic warning model for strong convective weather, such as hail and lightning, by multiple logistic linear regression. Zhang et al. (2020) presented a density-based convective storm identification method for weather radar data. North et al. (2020) used the heat equation to define a redistribution kernel, and a simple linear advection scheme was shown to work well in a lightning prediction example. Yücelbaş et al. (2021) used effective meteorological parameters to pre-estimate distance-based lightning. Mostajabi et al. (2019) used machine learning techniques to successfully hindcast nearby and distant lightning hazards by looking at single-site observations of meteorological parameters. Karami et al. (2020) presented a machine learning-based method to locate lightning flashes using calculations of lightning-induced voltages on a transmission line. Zhu et al. (2021) presented a machine-learning approach (support vector machines) to classify cloud-to-ground and intracloud lightning. Nevertheless, these methods use a limited number of data factors to analyze the relationships with lightning strike locations, and the recognition effects are often unsatisfactory.

Weather radar is one of the most effective instruments for monitoring the occurrence of lightning. It can be used to indirectly identify the electrification process within a developing thunderstorm because grapples and hail particles return large reflectivity echoes (Wei and Hsieh, 2020). As highly reliable data in the field of meteorological detection, radar data have been widely considered by meteorologists, and many explorations and practices have been carried out. Lu et al. (2017) presented a spatial lattice model based on sampling particles that was proposed to support both the representation and analysis of meteorological information. A 3D modeling strategy was used for weather radar data analysis (Lu et al., 2018). Based on the data of nine weather radar slices at different elevations, Wang et al. (2018) used a convolutional neural network model to identify the spatial structures of three-dimensional abnormal clouds when hail lands. Jiang et al. (2019) utilized multisource convolutional neural networks to extract the features of various weather-related data obtained from Doppler radar to identify thunderstorms and gales. Ling et al. (2020) proposed a new method based on stacked autoencoders to identify abnormal weather radar echo images. Li et al. (2020a) built a dataset from weather radar echo images using different depth models, such as a simple convolutional neural network (CNN), a recurrent neural network (S-RCNN), and a spatiotemporal recurrent convolutional neural network (ST-RCNN). The recognition performances of the learning models on thunderstorms and gales were compared. Zhou et al. (2020) proposed a new semantic segmentation-based deep learning network for cloud-to-ground lightning nowcasting named LightningNet. This model conducts reliable lightning nowcasting by using multisource data.

As mentioned above, the current related experiments are largely based on two-dimensional projection radar data or regenerated product radar data. In fact, when weather radar equipment is working, it scans in the surrounding three-dimensional space, and the obtained radar data have obvious three-dimensional characteristics. These studies ignore the potential three-dimensional characteristics in radar data. It is difficult to restore the real three-dimensional spatial data scene around a target, which has caused relevant research to have certain deficiencies. Machine learning can extract hidden feature information from multidimensional data, and the recognition of lightning strike locations based on three-dimensional radar detection data has become possible.

In this article, we first transform the problem of identifying lightning strike locations into a classification problem. Then, a sliding window is used to construct a lightning feature dataset based on three-dimensional weather radar data and lightning location data. Furthermore, logistic regression, a random forest, k-nearest neighbors, and a convolutional neural network are employed to identify lightning strike locations. The contribution of this paper is a first attempt to apply deep learning methods to the identification of lightning strike locations based on 3D radar data, thereby laying the

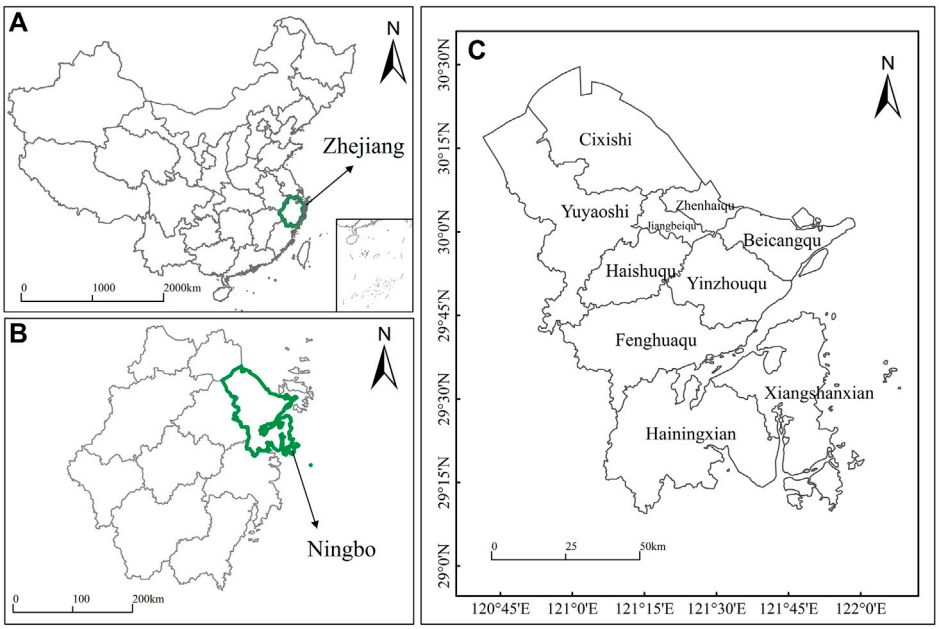

FIGURE 1 | Geographical locations of Ningbo.

foundation for subsequent accurate and real-time lightning location predictions. This work is expected to provide a new method for mitigating and preventing meteorological and lightning disasters.

The rest of this paper is organized as follows. In *Data*, we introduce the utilized data and the study area. In *Methodology*, we provide the details of establishing the dataset based on lightning location data and three-dimensional weather radar data and briefly introduce the classification algorithms used, including logistic regression, a random forest, k-nearest neighbors and a convolutional neural network. In *Experiments*, we present the experimental setup and results. Finally, we conclude and discuss future research directions in. *Conclusions and suggestions for future work*.

DATA

Data and Preprocessing

Lightning location data and three-dimensional weather radar data are provided by the Ningbo Meteorological Bureau. Lightning location data is obtained by the ADTD (Advanced TOA and Direction system) lightning location systems, and ADTD is ground-based advanced time of arrival and direction systems cloud-to-ground lightning detection sensors. At present, the system in the meteorological department has been widely used in China. Its detection efficiency is between 80 and 90%, and the error is generally several hundred meters to several kilometers. The single station detection radius of the lightning positioning system is approximately 300 km (Shi, 2016; Xu and Zhou, 2017). The lightning data contain fields denoting the time, location (latitude and longitude), polarity effect, peak intensity, steepness, and other information of the ground flash return process, providing great help when studying lightning activity. In this article, lightning data with intensities less than 10 KA are removed, and those whose corresponding radar combined reflectance (CR) grid values are less than 10 DBZ are also eliminated to ensure data accuracy. After preprocessing, the lightning location data are used as the ground truth. There is a clear correlation between the occurrence of lightning and radar echoes top heights and echo intensity (Michimoto, 1991; Futyan and Del Genio, 2007). At the same time, weather radar data is considered to be reliable detection data in the meteorological field. Some scholars have used radar data to predict lightning data. Therefore, we used radar echo data to identify lightning, hoping to provide preliminary research for lightning prediction based on radar data.

The weather radar data used in this paper are scanned and generated by the S-band Doppler weather radar system. Doppler weather radar has a high temporal-spatial resolution (1 km/6 min). It provides information about the positions and intensities of precipitation particles and particle motion information. Therefore, it is an effective tool for monitoring microscale and mesoscale convective systems and plays an essential role in detecting severe weather. Weather radars perform a 3D scan of the atmosphere. A radar system scans a full volume every 5–6 min. It scans starting from the lowest elevation angle and then increases the scanning angle gradually. Finally, it provides data at nine elevation angles according to a certain scanning strategy. Each elevation scan forms a cone with an output of 2D raster data of the same size, and all 2D raster data created at different elevation angles constitute 3D raster data with a strict vertical alignment (Han et al., 2019). In this article, 3D

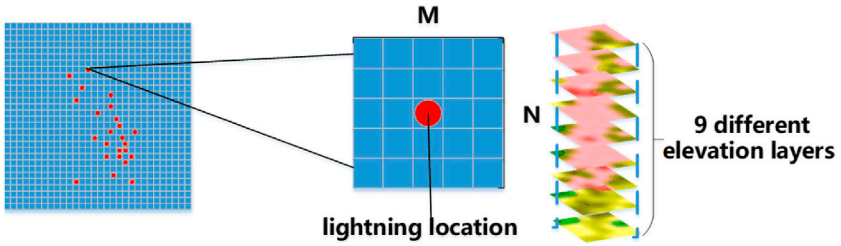

FIGURE 2 | Extraction of feature data with lightning.

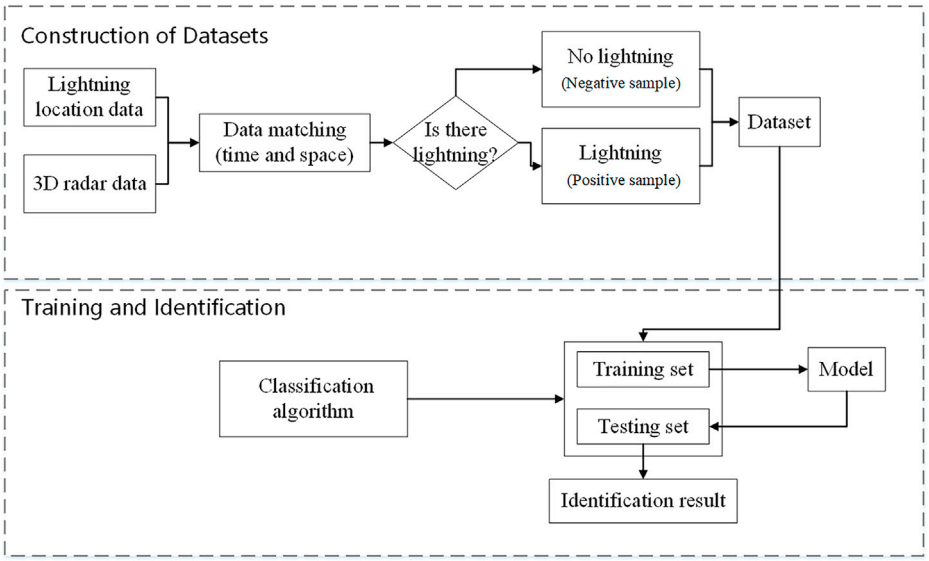

FIGURE 3 | Flowchart of the lightning identification solution used in this article.

weather radar data are used as input data to identify lighting strike locations.

Study Area

Ningbo is a sub-provincial city in northeastern Zhejiang Province, People's Republic of China, as shown in **Figure 1**. The spatial extent is (28° 51' - 30° 33′ N; 120° 55' - 122° 16′ E). It is bounded on the east by the East China Sea and the Zhoushan Archipelago, on the north by Hangzhou Bay (across which it faces Jiaxing and Shanghai), on the west by Shaoxing, and on the south by Taizhou. Thunderstorms frequently occur in Ningbo, and the direct economic losses caused by lightning strikes are as high as millions of yuan each year. Therefore, Ningbo is selected as the study area in this article.

METHODOLOGY

Establishing the Dataset

First, the lightning data and radar data should be matched spatially and temporally. A full radar scan generates one complete set of radar volume data, consuming approximately 5–6 min. Thus, for a specific complete radar volume dataset, the lightning data that occur during the period of the radar data scan are selected, and those beyond the spatial extent of the specific radar data are removed to ensure spatial consistency. Then, the final selected lightning data and the specific radar data form one group in which the lightning data and the radar data are well matched spatially and temporally. Therefore, the dataset in this article is constructed based on groups by the sliding window strategy. First, based on the radar reflectivity raster data, a sliding window with a size of M ×N (M is the size of the row, N is the size of the column) is set up to obtain the feature samples. Each feature sample contains nine layers of radar reflectivity raster data within the sliding window (with a size of M×N). Second, if one or more lightning data points are located on the center grid of the sliding window, the sample in this sliding window is labeled 1 (with lightning). Otherwise, the corresponding sample is labeled 0 (without lightning). **Figure 2** illustrates the extraction of feature data with lightning (samples labeled 1). After the traversal of the sliding window, a dataset with labels of 1 or 0 is established.

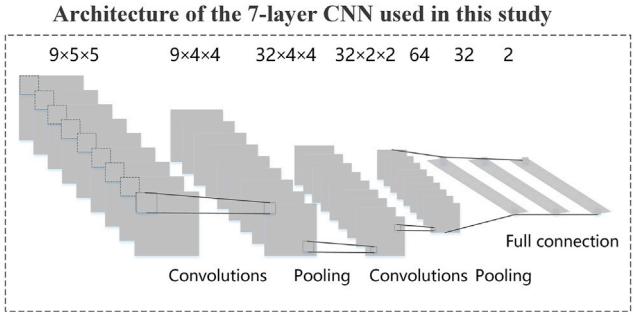

FIGURE 4 | Architecture of the CNN used in this study.

We collect 30,447 samples labeled 1 and 493,557 samples labeled 0 for a total of 524,004 samples from July to September. These data are processed and packaged to build a 3D spatial dataset. We divide into training sets and test sets, and were randomly scrambled during the training. The test subset contained samples for the second half of September 2018 (104,800 samples), whereas the training set included the remaining samples.

During the training period, the extracted feature data (samples labeled 1 or 0) are fed into various classification algorithms to train a two-class classification model. Furthermore, the trained model is applied to identify the new samples. If the output of the classification result is 1, it proves that there is lightning in the sample. If the output is 0, there is no lightning in the sample. **Figure 3** shows an overview of the solution used in this paper. It should be noted that this M×N sliding window can be set according to the specific application. For example, in this article, we set this sliding window size to 5 × 5 (5 columns and five rows).

Classification Algorithms

The identification of lightning strike locations by weather radar data is regarded as a classification problem. We utilize some frequently-used binary classification algorithms, namely, logistic regression (LR) (Wright, 1995; Kleinbaum et al., 2002), K-nearest neighbors (KNN) (Dudani, 1976; Kramer, 2013), a random forest (RF) (Liaw and Wiener, 2002; Pal, 2005) and a convolutional neural network (CNN) (Wang, et al., 2019; Sothe et al., 2020), to conduct our experiments based on the dataset constructed in the previous section. The following is a brief introduction of these approaches.

Logistic regression (LR): LR is essentially a classification method (Cheng et al., 2006). To solve a classification problem, the model is trained according to some known training sets, and then the classes of the new data are predicted. The goal of LR is to find a decision boundary with a sufficient degree of discrimination so that the two categories can be well separated. In this paper, the parameters we used in LR are set as follows, penalty is L2 regularization, Inverse of regularization strength is 1, the maximum number of iterations for the solver to converge is 500, and a binary problem fits for each label.

K-nearest neighbors (KNN): KNN is a classification algorithm (Liu et al., 2019). To determine the category of an unknown sample, KNN uses all the samples of the known categories as references and calculates the distances between the unknown sample and all the known samples. The K known samples that are closest to the unknown sample are selected. According to the majority-voting rule, the unknown sample and the K-nearest samples are classified into one category. The parameters of KNN are we set to: the number of neighbors is 9, leaf size is 5, the number of parallel jobs to run for neighbors search is 1.

TABLE 1 | Confusion matrix.

		Identified class	
		Strike	**Nonstrike**
Actual Class	Strike	TP	FN
	Nonstrike	FP	TN

Random forest (RF): RFs are commonly used in regression and classification, as they improve the prediction accuracy of the resulting model without significantly increasing the amount of required calculations (Gao et al., 2019; Li et al., 2020b). An RF is not sensitive to multivariate common linearity, the results are relatively robust to missing data and unbalanced data, and it can effectively predict the effects of up to thousands of explanatory variables. The parameters of the RF in this article are we set as: the number of trees in the forest is 100, the maximum depth of the tree is 5, random_state is the seed used by the random number generator set to 2.

Convolutional neural network (CNN): A CNN is a type of feedforward neural network that includes convolution calculations and has a deep structure (Lei et al., 2019; Wan et al., 2019). In deep learning, CNNs have achieved great success in image classification. CNNs possess the ability to characterize learning; they can classify input information according to its hierarchical structure and identify similar features at different locations in space. In this paper, the CNN structure contains three parts: an input layer, a hidden layer, and an output layer. The input layer can handle multidimensional data. We put the training dataset into the input layer. The hidden layer is the core of the CNN, including the convolutional layer, pooling layer and fully connected layer. For classification problems, the output layer returns the probability that the input image belongs to a certain category. For us, the output of the output layer is the probability of a lightning strike location.

Figure 4 shows the overall CNN architecture used in this article. The CNN has seven layers, two convolutional layers and two pooling layers that appear alternately. There are three fully connected (FC) layers connected to the last feature map. Finally, the fully connected layers output the final classification results. The size of the dataset input into the CNN is 9 × 5×5, and the filter size is set to 9 × 1×1 (here, nine refers to the weather radar data at nine different heights). 9 × 5×5 in the first layer means that this layer has nine different elevation radars and that each level has a size of 5 × 5.9 × 4×4 in the second layer means that the feature map generated after sample pooling has a size of 9 × 4×4. Other layers are similar ultimately fully connected layer (FC) output results. We use the cross-entropy loss as the objective

TABLE 2 | The results are shown in different models.

Classifier	Precision	FPR	Recall	Accuracy	F-measure	ROC AUC	P-R
LR	0.749	0.251	0.391	0.958	0.513	0.691	0.328
KNN	0.763	0.237	0.558	0.965	0.644	0.774	0.448
RF	0.835	0.165	0.538	**0.967**	0.654	0.765	0.475
CNN	**0.842**	**0.158**	**0.604**	**0.967**	**0.703**	**0.798**	**0.534**

The bold value is the best value in each column.

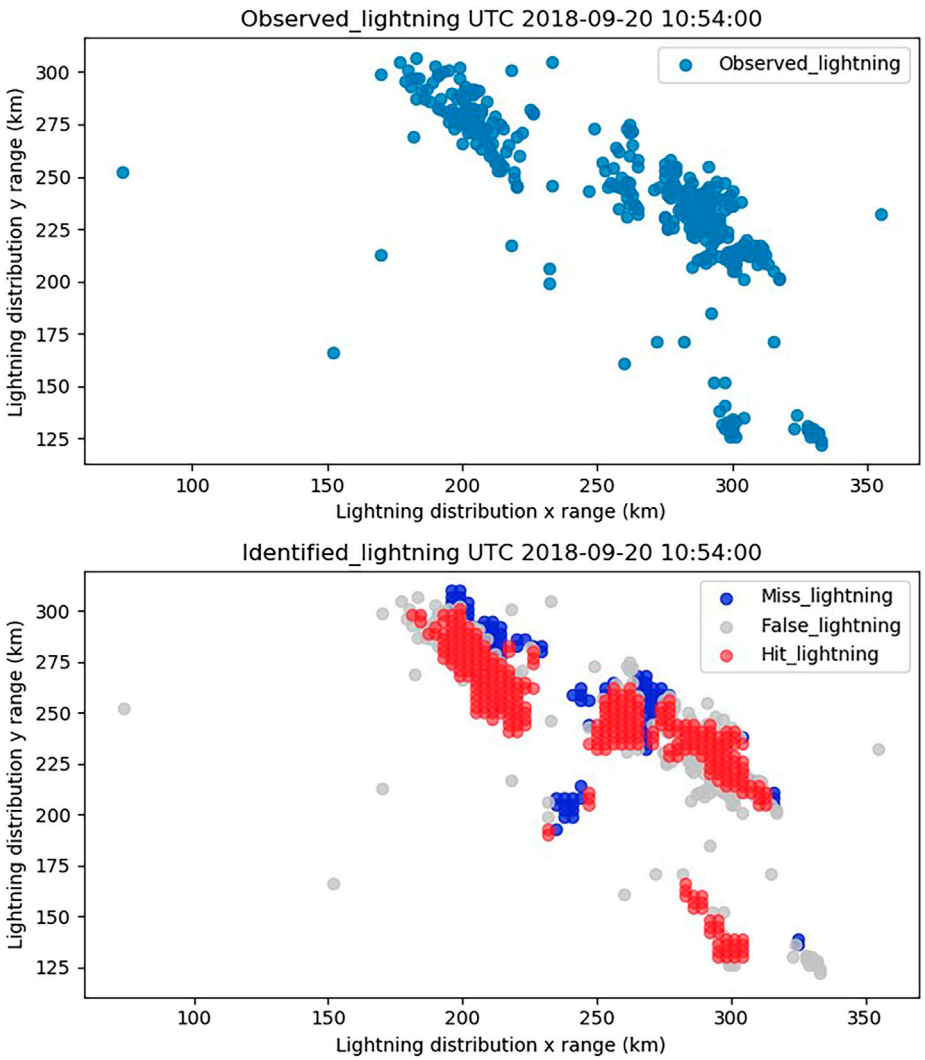

FIGURE 5 | Lightning locations observed and identified by the CNN model on September 20, 2018 (In upper, the blue dot represents the observed lightning. In lower, the dark blue dot (Miss_lightning) is miss identified lightning, the gray dot (False_lightning) represents the falsely reported lightning, and the red dot represents the correctly identified lightning.).

function for training the CNN. For optimization, the Adam optimizer is leveraged to train the network, the batch size is 64, and the number of epochs is set to 100. The learning rate is 0.001. Note that different from conventional machine learning approaches, CNNs can be directly applied to radar images without needing the handcrafted features above.

EXPERIMENTS

Experimental Setup

In this paper, we conduct comparative experiments with the dataset generated in *Methodology*. For these classification methods, the training set is employed for model training, and the test set is

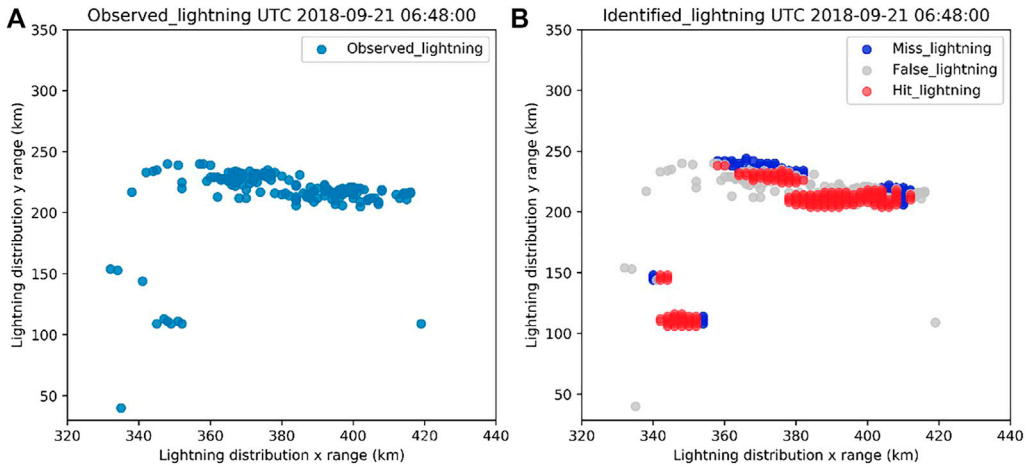

FIGURE 6 | Lightning locations observed and identified by the CNN model on September 21, 2018 (In **Figure 6A**, the blue dot represents the observed lightning. In **Figure 6B**, the dark blue dot (Miss_lightning) is miss identified lightning, the gray dot (False_lightning) represents the falsely reported lightning, and the red dot represents the correctly identified lightning).

used to verify the lightning classification results. In brief, the purpose is to explore the classification outcomes, which are regarded as the identification results of the lightning strike locations.

In addition, all classification algorithms utilized in this article are executed in the python37 environment. These machine learning methods (LR, the RF and KNN) are developed using the scikit-learn (sklearn) library. Sklearn is a simple and efficient tool for predictive data analysis. PyTorch's is used to efficiently implement the CNN approach. The CNN uses a GPU to accelerate the computation process. Moreover, the hardware environment contains a Core i7-9500 (2.6 GHz) CPU and a GeForce GTX 1065 GPU.

Performance Criteria

A confusion matrix can be used to assess the accuracy of binary classification results (Stehman, 1997). The confusion matrix shown in **Table 1** is used. In **Table 1**, Tp and TN represent the numbers of true positive and true negative cases, respectively, while FP and FN denote the numbers of false positive and false negative cases, respectively (Pakdaman et al., 2020). In **Table 1**, "strike" and "nonstrike" indicate the lightning strike and non-lightning-strike events, respectively.

Based on **Table 1**, to evaluate the performance of the proposed algorithms, seven conventional criteria are considered: precision, false positive rate (FPR), recall, accuracy, F-measure, area under the curve (AUC), and precision-recall (P-R) (Pakdaman et al., 2020; Luque et al., 2019; Sofaer et al., 2019). Precision is intuitively the ability of a classifier to not label a negative sample as positive and can be calculated by:

$$\text{Precision} = \frac{TP}{TP + FP} \tag{3}$$

False positive rate (FPR): The proportion of real negative examples predicted to be positive can be calculated by:

$$\text{FPR} = \frac{FP}{FP + TP} \tag{4}$$

Recall is the ability of the classifier to find all the positive samples and can be calculated by:

$$\text{Recall} = \frac{TP}{TP + FN} \tag{5}$$

Accuracy is the proportion of the correct predictions to the total number of predictions:

$$\text{Accuracy} = \frac{TP + TN}{TP + FN + FP + TN} \tag{6}$$

The F-measure can be interpreted as a weighted average of the precision and recall, where the best F-measure is 1 and at the worst is 0. The relative contributions of precision and recall to the F-measure are equal. The formula for the F-measure is:

$$\text{F} - \text{measure} = 2^{*}\frac{precision * recall}{precision + recall} \tag{7}$$

The area under the receiver operating characteristic curve (ROC AUC) is defined as the area enclosed by the coordinate axes under the ROC curve (Luque et al., 2019). The value of this area cannot be greater than 1. The value range of the AUC is between 0.5 and 1. The closer the AUC is to 1.0, the higher the authenticity of the detection method; when it is equal to 0.5, the authenticity is lowest, and the method has no application value.

Precision-recall (P-R) is a useful measure of prediction success when the classes are very imbalanced (Sofaer et al., 2019; Saito and Rehmsmeier, 2015). When the number of positive samples is seriously less than the number of negative samples, the P-R curve can more intuitively express the differences between models than other metrics, which is more appropriate. Since the collected dataset contains unbalanced data, P-R is required. In this paper, we use the area under the P-R curve to express the P-R curve.

Results Analysis

We conduct experiments using the constructed dataset to compare the performance of the four models, namely, LR,

KNN, RF, and a CNN. The precision, false positive rate (FPR), recall, accuracy, F-measure, area under the curve (AUC), and precision recall (P-R) values of the test samples produced by the three models are recorded as the final results. For these indicators, smaller FPRs are better, and the larger the other metrics are, the better they are.

Table 2 shows that the model with the best performance in term of recognizing the lightning strike locations is the CNN, followed by the RF and KNN, and LR has the worst performance. The CNN has the best precision, recall, accuracy, and F-measure, and it has the lowest FPR at 0.158. More importantly, the AUC and P-R performance of the CNN is far superior to that of other models. P-R is an important model identification indicator. The area under the PR curve yielded by the CNN is 0.534, which is 0.059 higher than those of the other best-performing models. This shows that CNN can not only better obtain the hidden data features among the multilayer radar data but also has better model ability and can adapt to complex multidimensional data. Notably, the RF performs best aside from the CNN. Its P-R and other criteria (except Recall and AUC) are the best among the three machine learning methods, indicating that the RF is also an effective algorithm and has a certain effect on identifying lightning strikes. Among all the methods, LR performs worst. Not only are its AUC and P-R the worst, but its FPR is the highest among those of all the models, which shows that LR has difficulty dealing with the identification of lightning strike locations.

In addition, all models have high accuracy, but the PR values of the LR model, DT model, and KNN model do not perform well. This means that when identifying lightning strikes, accuracy alone cannot measure the true accuracy of the given model. This may be due to the imbalance between positive and negative samples in the created dataset. It is undeniable that the CNN model has the best recognition effect, and we use it for a subsequent case analysis.

Case Study 1–September 20, 2018

To verify the CNN model, we apply the model in a real environment. Under the influence of convective cloud clusters, intense lightning activities occurred in Ningbo in northeastern Zhejiang Province of China on September 20, 2018. The China Meteorological Administration (CMA) Public Meteorological Service Centre issued a lightning warning. Therefore, the weather radar data obtained from Ningbo at 10:54 (Universal Time) on that day are selected for use.

Lightning is not particularly stable and normally drifts. The deviation of the thunderstorm center is less than 5%, and the accuracy of the identification location is within the acceptable (Huang et al., 2019). In order to better quantify and evaluate the recognition results of the model, we believe that the identified lightning strike location within 1 km around the lightning strike location is also effective.

Figure 5 shows the observed lightning strike position and the result identified by the CNN model. The upper part of the figure is the distribution of the observed lightning strike position, and the lower part is the result of the CNN model identification. The average probability of correct hit, miss identified, false identified for this case amounted to 0.763, 0.235, 0.237. It can be seen that the CNN can identify the approximate locations of lightning, and the model has a good recognition effect with respect to concentrated lightning, which proves that it is feasible to use a CNN combined with radar data in a real environment to identify lightning. However, the disadvantage is that the range of lightning strike locations identified by the CNN is larger than the actual landing area. These situations may be due to the fact that the model training samples are not sufficient or that some lightning cannot be detected by lightning positioning equipment. It is also possible that we have overlooked certain parameters that have important impacts on lightning recognition. Furthermore, another shortcoming is that the CNN model is not ideal for the recognition of discrete lightning location data. Some scattered lightning bolts are not recognized.

Case Study 2–September 21, 2018

On August 23, 2018, a few thunderstorms occurred in south ningbo. We chose the radar data at 06:48 UTC that day for the examination. **Figure 6** shows Lightning locations observed and identified by the CNN model. **Figure 6A** shows the observed lightning strike locations. and **Figure 6B** shows the results of lightning strikes identified by the CNN. The average probability of correct hit, miss identified, false identified for this case amounted to 0.725, 0.271, 0.275. In this case, the CNN model also can identify the approximate location of the lightning strike locations, but the identification results of the discretely distributed lightning points are poor and cannot be prepared for identification. Nevertheless, it is possible to use the CNN model to identify lightning strike locations.

CONCLUSIONS AND SUGGESTIONS FOR FUTURE WORK

In this article, we convert the problem of identifying lightning strike locations into a binary classification problem, and a sliding window strategy is utilized to construct a dataset suitable for the identification of lightning strike locations based on 3D weather radar data. Then, based on the constructed dataset, four common classification algorithms (LR, an RF, KNN and a CNN) are applied to explore the identification of lightning strike locations. The results show that the CNN has the best performance in terms of the identification of lightning strike locations. Its precision is 0.842, the recall is 0.604, the accuracy is 0.967, the FPR is 0.158, and the area under the P-R curve is also outstanding at 0.534. The CNN is followed by the RF and KNN, LR has the worst performance. This proves that a deep learning method can conduct autonomous learning of spatial feature data with the support of a large amount of multidimensional data and can obtain more hidden data information relationships than other approaches. Lightning strike location recognition based on three-dimensional radar detection data is efficient and can be used to a certain extent. However, because the occurrence of lightning is a relatively low-probability event, it is difficult to identify with weather radar data. The essence of this situation is an imbalanced classification problem. When constructing the dataset, we lack the consideration of a balanced sampling design, which to a certain

extent causes the problem of too large a lightning strike location to being identified by the CNN model. In the future, the impact of unbalanced samples on the identification of lightning strike locations will be considered. In addition, with the rapid development of deep learning, an increasing number of methods and technologies can be applied to the recognition of lightning strikes. Other modeling methods may be able to obtain better results with respect to the recognition of lightning strikes. We need to evaluate these numerous methods in the future. We hope to obtain a more accurate method for identifying lightning strike locations and reduce the false positive rate of lightning detection to provide certain decision-making support for disaster prevention and mitigation.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: https://github.com/zyd530/Data.

AUTHOR CONTRIBUTIONS

Conceptualization and design of study, ML, YZ, and MC; Data Collection, ML and JZ; Data analysis and interpretation, ML, YZ, MY, ZM, and MW; Writing and preparation of original draft, ML, YZ, MY, and ZM; Funding acquisition, ML and MC. All authors contributed to article revision, read and approved the submitted version.

FUNDING

The NSCF Project (41871285).

ACKNOWLEDGMENTS

Many thanks to reviewers for their valuable comments. This paper was supported by the NSCF Project (41871285).

REFERENCES

Barthe, C., Deierling, W., and Barth, M. C. (2010). Estimation of Total Lightning from Various Storm Parameters: A Cloud-Resolving Model Study. *J. Geophys. Res. Atmos.* 115 (D24). doi:10.1029/2010jd014405

Chen, J. P., Tan, B., and Lian, S. Z. (2017). Data Mining for Correlation Rules of Lightning in Hubei Province. *J. Appl. Sci.* 35 (1), 42–50. doi:10.3969/j.issn.0255-8297.2017.01.005

Cheng, Q., Varshney, P. K., and Arora, M. K. (2006). Logistic Regression for Feature Selection and Soft Classification of Remote Sensing Data. *IEEE Geosci. Remote Sensing Lett.* 3 (4), 491–494. doi:10.1109/lgrs.2006.877949

Cintineo, J. L., Pavolonis, M. J., Sieglaff, J. M., Lindsey, D. T., Cronce, L., Gerth, J., et al. (2018). The NOAA/CIMSS ProbSevere Model: Incorporation of Total Lightning and Validation. *Weather Forecast.* 33 (1), 331–345. doi:10.1175/waf-d-17-0099.1

Dudani, S. A. (1976). The Distance-Weighted K-Nearest-Neighbor Rule. *IEEE Trans. Syst. Man. Cybern.* SMC-6 (4), 325–327. doi:10.1109/tsmc.1976.5408784

Futyan, J. M., and Del Genio, A. D. (2007). Relationships between Lightning and Properties of Convective Cloud Clusters. *Geophys. Res. Lett.* 34 (15), L15705. doi:10.1029/2007gl030227

Gao, S., Li, M., Liang, Y., Marks, J., Kang, Y., and Li, M. (2019). Predicting the Spatiotemporal Legality of On-Street Parking Using Open Data and Machine Learning. *Ann. GIS* 25 (4), 299–312. doi:10.1080/19475683.2019.1679882

Gharaylou, M., Farahani, M. M., Mahmoudian, A., and Hosseini, M. (2020). Prediction of Lightning Activity Using WRF-ELEC Model: Impact of Initial and Boundary Conditions. *J. Atmos. Solar-Terrestrial Phys.* 210, 105438. doi:10.1016/j.jastp.2020.105438

Giannaros, T. M., Kotroni, V., and Lagouvardos, K. (2016). WRF-LTNGDA: A Lightning Data Assimilation Technique Implemented in the WRF Model for Improving Precipitation Forecasts. *Environ. Model. Softw.* 76, 54–68. doi:10.1016/j.envsoft.2015.11.017

Han, L., Sun, J., and Zhang, W. (2019). Convolutional Neural Network for Convective Storm Nowcasting Using 3-D Doppler Weather Radar Data. *IEEE Trans. Geosci. Remote Sensing* 58 (2), 1487–1495. doi:10.1109/TGRS.2019.2948070

Huang, L., Sheng, S. U., Yang, X., Jun, H. U., and Liu, Z. (2019). Forecasting of Lightning Trending Improved Density-Based Lls. *Insulators and Surge Arresters* (01), 76–83. doi:10.16188/j.isa.1003-8337.2019.01.013

Jiang, Y., Yao, J., and Qian, Z. (2019). A Method of Forecasting Thunderstorms and Gale Weather Based on Multisource Convolution Neural Network. *IEEE Access* 7, 107695–107698. doi:10.1109/access.2019.2932027

Karami, H., Mostajabi, A., Azadifar, M., Rubinstein, M., Zhuang, C., and Rachidi, F. (2020). Machine Learning-Based Lightning Localization Algorithm Using Lightning-Induced Voltages on Transmission Lines. *IEEE Trans. Electromagn. Compat.* 62 (6), 2512–2519. doi:10.1109/temc.2020.2978429

Kleinbaum, D. G., Dietz, K., Gail, M., Klein, M., and Klein, M. (2002). *Logistic Regression*. New York: Springer-Verlag.

Kramer, O. (2013). "K-nearest Neighbors". in *Dimensionality Reduction with Unsupervised Nearest Neighbors*. Berlin: Heidelberg Springer, 13–23. doi:10.1007/978-3-642-38652-7_2

Lei, B., Yang, L., and Xu, Z. (2019). "Using Convolutional Neural Network to Classify Convective Cloud on Radar Echoes." in ICMO 2019: International Conference on Meteorology Observations; December 2019. Chengdu: IEEE, 1–3.

Li, X., Yang, C., Huang, W., Tang, J., Tian, Y., and Zhang, Q. (2020a). Identification of Cotton Root Rot by Multifeature Selection from Sentinel 2 Images Using Random Forest. *Remote Sensing* 12 (21), 3504. doi:10.3390/rs12213504

Li, Y., Li, H., Li, X., Li, X., and Xie, P. (2020b). On Deep Learning Models for Detection of Thunderstorm Gale. *J. Internet Tech.* 21 (4), 909–917. doi:10.3966/160792642020072104001

Liaw, A., and Wiener, M. (2002). Classification and Regression by random Forest. *R. News* 2 (3), 18–22.

Liu, P., Du, J. Z., Lyu, W. G., and Dou, M. W. (2019). A Modified KNN Classifier for Unbalanced Dataset (In Chinese). *J. Northeast. Univ. (Natural Science)* 40 (7), 932. doi:10.12068/j.issn.1005-3026.2019.07.004

Lu, M., Chen, M., Wang, X., Min, J., and Liu, A. (2017). A Spatial Lattice Model Applied for Meteorological Visualization and Analysis. *Ijgi* 6 (3), 77. doi:10.3390/ijgi6030077

Lu, M., Chen, M., Wang, X., Yu, M., Jiang, Y., and Yang, C. (2018). 3D Modelling Strategy for Weather Radar Data Analysis. *Environ. Earth Sci.* 77 (24), 1–10. doi:10.1007/s12665-018-7985-2

Luque, A., Carrasco, A., Martín, A., and de las Heras, A. (2019). The Impact of Class Imbalance in Classification Performance Metrics Based on the Binary Confusion Matrix. *Pattern Recognition* 91, 216–231. doi:10.1016/j.patcog.2019.02.023

Maggio, C. R., Marshall, T. C., and Stolzenburg, M. (2009). Estimations of Charge Transferred and Energy Released by Lightning Flashes. *J. Geophys. Res. Atmos.* 114 (D14). doi:10.1029/2008jd011506

Michimoto, K. (1991). A Study of Radar Echoes and Their Relation to Lightning Discharge of Thunderclouds in the Hokuriku District. *J. Meteorol. Soc. Jpn.* 69 (3), 327–336. doi:10.2151/jmsj1965.69.3_327

Mostajabi, A., Finney, D. L., Rubinstein, M., and Rachidi, F. (2019). Nowcasting Lightning Occurrence from Commonly Available Meteorological Parameters Using Machine Learning Techniques. *Npj Clim. Atmos. Sci.* 2 (1), 1–15. doi:10.1038/s41612-019-0098-0

NCAR Mesoscale and Microscale Meteorology Laboratory (2021). WRF User Support & Contributor Information. Available at: https://www.mmm.ucar.edu/wrf-user-support-contributor-information (Accessed April 5, 2021).

NOAA National Severe Storms Laboratory (2021). SEVERE WEATHER 101 - Lightning Basics. Available at: https://www.nssl.noaa.gov/education/svrwx101/lightning/ (Accessed April 1, 2021).

North, J., Stanley, Z., Kleiber, W., Deierling, W., Gilleland, E., and Steiner, M. (2020). A Statistical Approach to Fast Nowcasting of Lightning Potential fields. *Adv. Stat. Clim. Meteorol. Oceanogr.* 6 (2), 79–90. doi:10.5194/ascmo-6-79-2020

Pakdaman, M., Naghab, S. S., Khazanedari, L., Malbousi, S., and Falamarzi, Y. (2020). Lightning Prediction Using an Ensemble Learning Approach for Northeast of Iran. *J. Atmos. Solar-Terrestrial Phys.* 209, 105417. doi:10.1016/j.jastp.2020.105417

Pal, M. (2005). Random forest Classifier for Remote Sensing Classification. *Int. J. remote sensing* 26 (1), 217–222. doi:10.1080/01431160412331269698

Saito, T., and Rehmsmeier, M. (2015). The Precision-Recall Plot Is More Informative Than the ROC Plot when Evaluating Binary Classifiers on Imbalanced Datasets. *PloS one* 10 (3), e0118432. doi:10.1371/journal.pone.0118432

Shi, X. (2016). Comparative Analysis on Detection Efficiency of Two ADTD Lightning Location Systems (In Chinese). *Meteorol. Hydrological Mar. Instr.* 33 (3), 6–12. doi:10.19441/j.cnki.issn1006-009x.2016.03.002

Sofaer, H. R., Hoeting, J. A., and Jarnevich, C. S. (2019). The Area under the Precision-Recall Curve as a Performance Metric for Rare Binary Events. *Methods Ecol. Evol.* 10 (4), 565–577. doi:10.1111/2041-210x.13140

Sothe, C., De Almeida, C. M., Schimalski, M. B., La Rosa, L. E. C., Castro, J. D. B., Feitosa, R. Q., et al. (2020). Comparative Performance of Convolutional Neural Network, Weighted and Conventional Support Vector Machine and Random forest for Classifying Tree Species Using Hyperspectral and Photogrammetric Data. *GIScience & Remote Sensing* 57 (3), 369–394. doi:10.1080/15481603.2020.1712102

Stehman, S. V. (1997). Selecting and Interpreting Measures of Thematic Classification Accuracy. *Remote sensing Environ.* 62 (1), 77–89. doi:10.1016/s0034-4257(97)00083-7

Wan, L., Zhang, H., Lin, G., and Lin, H. (2019). A Small-Patched Convolutional Neural Network for Mangrove Mapping at Species Level Using High-Resolution Remote-Sensing Image. *Ann. GIS* 25 (1), 45–55. doi:10.1080/19475683.2018.1564791

Wang, L., Wang, H., and Heng, Z. (2019). A Rapid Identification and Warning Method for Severe Weather via Doppler Radar Based on an Improved TITAN Algorithm. *J. Atmos. Solar-Terrestrial Phys.* 193, 105080. doi:10.1016/j.jastp.2019.105080

Wang, P., Lv, W., Wang, C., and Hou, J. (20182018). "Hail Storms Recognition Based on Convolutional Neural Network." in WCICA 2018: 13th World Congress on Intelligent Control and Automation. Changsha: IEEE, 1703–1708.

Wei, C.-C., and Hsieh, P.-Y. (2020). Estimation of Hourly Rainfall during Typhoons Using Radar Mosaic-Based Convolutional Neural Networks. *Remote Sensing* 12 (5), 896. doi:10.3390/rs12050896

Wright, R. E. (1995). "Logistic Regression," In *Reading and Understanding Multivariate Statistics*. Editor L. G. Grimm and P. R. Yarnold (American Psychological Association), 217–244.

Xu, H., and Zhou, M. (2017). Evaluation and Analysis of Detecting Efficiency Based on ADTD Lightning Location Network. *Henan Sci. Tech.* (01), 126–128.

Yang, L., Wang, Y., Wang, Z., Qi, Y., Li, Y., Yang, Z., et al. (2020). A New Method Based on Stacked Auto-Encoders to Identify Abnormal Weather Radar echo Images. *EURASIP J. Wireless Com Netw.* 2020 (1), 1–15. doi:10.1186/s13638-020-01769-3

Yücelbaş, Ş., Erduman, A., Yücelbaş, C., and Yildiz, F. (2021). Pre-estimation of Distance-Based Lightning Using Effective Meteorological Parameters. *Arabian J. Sci. Eng.* 46 (2), 1529–1539. doi:10.1007/s13369-020-05257-0

Zepka, G. S., Pinto, O., Jr., and Saraiva, A. C. V. (2014). Lightning Forecasting in southeastern Brazil Using the WRF Model. *Atmos. Res.* 135-136, 344–362. doi:10.1016/j.atmosres.2013.01.008

Zhang, Z., Yang, J., and He, Z. (2020). "A Density Based Convective Storm Identification Method for Radar Data." in CIYCEE 2020: IEEE 1st China International Youth Conference on Electrical Engineering. Wuhan: IEEE, 1–6.

Zhou, K., Zheng, Y., Dong, W., and Wang, T. (2020). A Deep Learning Network for Cloud-To-Ground Lightning Nowcasting with Multisource Data. *J. Atmos. Oceanic Tech.* 37 (5), 927–942. doi:10.1175/jtech-d-19-0146.1

Zhu, Y., Bitzer, P., Rakov, V., and Ding, Z. (2021). A Machine-Learning Approach to Classify Cloud-To-Ground and Intracloud Lightning. *Geophys. Res. Lett.* 48, 1–8. doi:10.1029/2020GL091148

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.

frontiers in Environmental Science

ORIGINAL RESEARCH
published: 11 October 2021
doi: 10.3389/fenvs.2021.731452

Integration of One-Pair Spatiotemporal Fusion With Moment Decomposition for Better Stability

Yaobin Ma[1,2], *Jingbo Wei*[2]* *and Xiangtao Huang*[3]

[1]School of Resources, Environmental and Chemical Engineering and Key Laboratory of Poyang Lake Environment and Resource Utilization, Ministry of Education, Nanchang University, Nanchang, China, [2]Institute of Space Science and Technology, Nanchang University, Nanchang, China, [3]Jiangxi Center for Data and Application of High Resolution Earth Observation System, Nanchang, China

Spatiotemporal fusion has got enough attention and many algorithms have been proposed, but its practical stability has not been emphasized yet. Observing that the strategies harnessed by different types of algorithms may lead to various tendencies, an integration strategy is introduced to make full use of the complementarity between different types of spatiotemporal fusion algorithms for better fusion stability. In our method, the images fused by two different types of methods are decomposed into components denoting strength, structure, and mean intensity, which are combined separately involving a characteristic analysis. The proposed method is compared with seven algorithms of four types by reconstructing Landsat-8, Landsat-7, and Landsat-5 images to validate the effectiveness of the spatial fusion strategy. The digital evaluation on radiometric, structural, and spectral loss illustrates that the proposed method can reach or approach the optimal performance steadily.

Keywords: spatiotemporal fusion, Landsat, MODIS, multispectral, fusion, FSDAF

Edited by:
Peng Liu,
Institute of Remote Sensing and Digital Earth (CAS), China

Reviewed by:
Costica Nitu,
Politehnica University of Bucharest, Romania
Guang Yang,
South China Normal University, China
Jining Yan,
China University of Geosciences Wuhan, China
Xinghua Li,
Wuhan University, China

***Correspondence:**
Jingbo Wei
wei-jing-bo@163.com

Specialty section:
This article was submitted to Environmental Informatics and Remote Sensing, a section of the journal Frontiers in Environmental Science

Received: *27 June 2021*
Accepted: *01 September 2021*
Published: *11 October 2021*

Citation:
Ma Y, Wei J and Huang X (2021) Integration of One-Pair Spatiotemporal Fusion With Moment Decomposition for Better Stability. Front. Environ. Sci. 9:731452. doi: 10.3389/fenvs.2021.731452

1 INTRODUCTION

Satellite images with dense time series and high spatial resolution are eagerly needed for remote sensing of abrupt changes in Earth, while they are hardly obtained due to physical constraints and adverse weather conditions (Li et al., 2019). Spatiotemporal fusion algorithms were developed to combine images of different temporal and spatial resolutions to obtain a composite image of high spatiotemporal resolution, which have been put to practice to monitor floods (Tan et al., 2019b) or forests (Chen et al., 2020). The spatiotemporal fusion process usually involves two types of remote sensing images. One type has high temporal and low spatial resolution (hereinafter referred to as low-resolution images), such as MODIS images. The other type has high spatial and low temporal resolution (hereinafter referred to as high-resolution images), such as Landsat images. The one-pair fusion is mostly studied for its convenience that only one pair of known images is required. The one-pair spatiotemporal fusion algorithms can be classified into four types, namely, weight-based, unmixing-based, dictionary pair–based, and neural network–based, as will be discussed.

Weight-based methods search similar pixels within a window in the given high-resolution images and predict the values of central pixels with weights linear to the inverse distance. Gao et al. (2006) proposed the spatial and temporal adaptive reflectance data fusion model (STARFM) with the blending weights determined by spectral difference, temporal difference, and location distance, which is the earliest weight-based method. STARFM was subsequently improved for more complex situations, resulting in the spatiotemporal adaptive algorithm for mapping reflectance change

(STAARCH) (Hilker et al., 2009) and enhanced STARFM (ESTARFM) (Zhu et al., 2010). When land cover type change and disturbance exist, the former can improve the performance of STARFM and the latter can improve the accuracy of STARFM in heterogeneous areas. There are other methods in this category, such as modified ESTARFM (mESTARFM) (Fu et al., 2013), the spatiotemporal adaptive data fusion algorithm for temperature mapping (SADFAT) (Weng et al., 2014), the rigorously weighted spatiotemporal fusion model (RWSTFM) (Wang and Huang, 2017), and the bilateral filter method (Huang et al., 2013).

Unmixing-based methods work out the abundance matrix of endmember fractions by clustering on the known high-resolution images. The first unmixing-based spatiotemporal method may be the multisensor multiresolution technique (MMT) proposed by Zhukov et al. (1999). Later, Zurita-Milla et al. (2008) introduced constraints into the linear unmixing process to ensure that the solved reflectance values were positive and within an appropriate range using the spatial information of Landsat/TM data and the spectral and temporal information of medium resolution imaging spectrometer (MERIS) data to generate images. Wu et al. (2012) proposed a spatiotemporal data fusion algorithm (STDFA) that extracts fractional covers and predicts surface reflectance under the rule of least square errors. Xu et al. (2015) proposed an unmixing method that includes the prior class spectra to smoothen the prediction image of STARFM within each class. Zhu et al. (2016) proposed the flexible spatiotemporal data fusion (FSDAF) (Li et al., 2020b) where a thin plate spline interpolator is used. The enhanced spatial and temporal data fusion model (ESTDFM) (Zhang et al., 2013), the spatial and temporal reflectance unmixing model (STRUM) (Gevaert and Javier Garcia-Haro, 2015), and the modified spatial and temporal data fusion approach (MSTDFA) (Wu et al., 2015b) were also proposed along the framework.

Separately, dictionary pair–based methods introduced coupled dictionary learning and nonanalytic optimization to predict missing images in the sparse domain, where the coded coefficients of high- and low-resolution images are very similar, given the over-complete dictionaries being well designed. Based on this theory, Huang and Song (2012) proposed the sparse representation–based spatiotemporal reflectance fusion model (SPSTFM), which may be the first to introduce dictionary pair–learning technology from natural image super-resolution into spatiotemporal data fusion (Zhu et al., 2016). SPSTFM was developed for predicting the surface reflectance of high-resolution images through jointly training two dictionaries generated by high-resolution and low-resolution difference image patches and sparse coding. After SPSTFM, Song and Huang (2013) developed another dictionary pair–based fusion method, which uses only one pair of high-resolution and low-resolution images. The error-bound-regularized semi-coupled dictionary learning (EBSCDL) (Wu et al., 2015a) and the fast iterative shrinkage-thresholding algorithm (FISTA) (Liu et al., 2016) are also proposed based on this theory. We have also investigated this topic and proposed sparse Bayesian learning and compressed sensing for spatiotemporal fusion (Wei et al., 2017a; Wei et al., 2017b).

Recently, dictionary learning has been replaced with convolutional neural networks (CNNs) (Liu et al., 2017) for sparse representation, which are used in the neural network–based methods to model the super-resolution of different sensor sources. Dai et al. (2018) proposed a two-layer fusion strategy, and in each layer, CNNs are employed to exploit the nonlinear mapping between the images. Song et al. (2018) proposed two five-layered CNNs to deal with the problem of complicated correspondence and large spatial resolution gaps between MODIS and Landsat images. In the prediction stage, they design a fusion model consisting of the high-pass modulation and a weighting strategy to make full use of the information in prior images. These models have small numbers of convolutional layers. Li et al. (2020a) proposed a learning method based on CNNs to effectively obtain sensor differences in the bias-driven spatiotemporal fusion model (BiaSTF). Many new methods are subsequently proposed, such as the deep convolutional spatiotemporal fusion network (DCSTFN) (Tan et al., 2018), enhanced DCSTFN (EDCSTFN) (Tan et al., 2019a), the two-stream convolutional neural network (StfNet) (Liu et al., 2019), and the generative adversarial network–based spatiotemporal fusion model (GAN-STFM) (Tan et al., 2021). It is expected that when a sequence of known image pairs are provided, the missed images can be predicted with the bidirectional long short-term memory (LSTM) network (Zhang et al., 2021).

Although spatiotemporal fusion has received wide attention and a lot of spatiotemporal fusion algorithms were developed (Zhu et al., 2018), the stability of algorithms has not been emphasized yet. On the one hand, the selection of base image pairs greatly affects the performance of fusion, as has been addressed in Chen et al. (2020). On the other hand, the performance of an algorithm is constrained by its type. This could be explained with FSDAF (Zhu et al., 2016) and Fit-FC (Wang and Atkinson, 2018), which are among the best algorithms. The linear model of Fit-FC projects the phase change, which can approach good fitness for the homogeneous landscapes. However, the nearest neighbor and linear upsampling methods used to model spatial differences in Fit-FC are too much rough, and the smoothing in the local window accounts for insufficient details. FSDAF focuses on heterogeneous or changing land covers. Different prediction strategies are used to adapt to heterogeneous and homogeneous landscapes. The thin plate spline for upsampling interpolation shows admirable fitness to the spatial structure. However, it is challenging for the abundance matrix to disassemble the homogeneous landscapes due to the long tail data. An unchanged area may be incorrectly classified as a heterogeneous landscape or changed areas may not be discovered, which leads to wrong prediction directions. To sum up, Fit-FC excels well at predicting homogeneous areas, while FSDAF excels at heterogeneous areas.

The combination of different algorithms is a way to improve the performance consistency in different scenarios. For example, Choi et al. (2019) proposed a framework called the consensus neural network to combine multiple weak image denoisers. Liu et al. (2020) proposed a spatial local fusion strategy to decompose images of different denoised images into structural patches and

reconstruct them. The combined results showed overall superiority than any other single algorithm. These strategies can be transplanted to the results of spatiotemporal fusion to improve the stability of practice.

Observing the complementarity of different spatiotemporal fusion algorithms, in this study, we propose a universal approach to improve the stability. Specifically, the results of FSDAF and Fit-FC are merged with the structure-based spatial integration strategy and the advantages of different algorithms are expected to be retained. The CNN-based methods are not integrated because deep learning has limited performance for a single pair of images, and the unclear theory makes it difficult to locate advantages. Extensive experiments demonstrated that the proposed combination strategy outperforms state-of-the-art one-pair spatiotemporal fusion algorithms.

Our method makes the following contributions:

1) The stability issue of spatiotemporal fusion algorithms is investigated for the first time.
2) A fusion framework is proposed to improve the stability.
3) The effectiveness of the method is proved by comparing with different types of algorithms.

The rest of this article is organized as follows. **Section 2** introduces the FSDAF model and the Fit-FC model in detail. **Section 3** summarizes the fusion based on the spatial structure. **Section 4** gives the experimental scheme and results visually and digitally, which is followed by discussion in **Section 5**. **Section 6** gives the conclusion.

2 RELATED WORK

In this section, the FSDAF and Fit-FC algorithms are detailed for further combination.

2.1 FSDAF

The FSDAF algorithm (Zhu et al., 2016) predicts high-resolution images of heterogeneous regions by capturing gradual and abrupt changes in land cover types. FSDAF integrates ideas from unmixing-based methods, spatial interpolation, and STARFM into one framework. FSDAF includes six main steps.

Step 1: The unsupervised classifier ISODATA is used to classify the high-resolution image at time t_1, and the class fractions A_c are calculated as

$$A_c(i) = N_c(i)/M, \tag{1}$$

where $N_c(i)$ is the number of high-resolution pixels belonging to class c within the *ith* low-resolution pixel and M is the number of high-resolution pixels within one low-resolution pixel.

Step 2: For every band of the two low-resolution images C_{t_1} and C_{t_2} captured at time t_1 and t_2, respectively, the reflectance changes ΔC are used to estimate the temporal change of all classes ΔF_c with the following:

$$\Delta C(i) = C_{t_2}(i) - C_{t_1}(i) = \sum_{c=1}^{L} A_c(i) \cdot \Delta F_c, \tag{2}$$

where L denotes the number of classes.

Step 3: The class–level temporal change is used to obtain the temporal prediction image $F_{t_2}^{TP}$ at time t_2 and calculate the residual R with the following:

$$F_{t_2}^{TP}(j_i) = F_{t_1}(j_i) + \Delta F_c, \tag{3}$$

$$R(i) = \Delta C(i) - \frac{1}{M}\left[\sum_{j=1}^{m}\left(F_{t_2}^{TP}(j_i) - F_{t_1}(j_i)\right)\right]. \tag{4}$$

Here, F_{t_1} is the known high-resolution image at time t_1 and j_i is the coordinate of the *jth* high-resolution pixel within the *ith* low-resolution pixel.

Step 4: The thin plate spline (TPS) interpolator is used to interpolate the low-resolution image C_{t_2} to obtain the spatial prediction image $F_{t_2}^{SP}$ at time t_2.

Step 5: Residual errors were distributed based on temporal prediction $F_{t_2}^{TP}$ and spatial prediction $F_{t_2}^{SP}$,

$$CW(j_i) = \left(F_{t_2}^{SP}(j_i) - F_{t_2}^{TP}(j_i) - R(i)\right) \cdot HI(j_i) + R(i), \tag{5}$$

$$W(j_i) = CW(j_i)/\sum_{j=1}^{M} CW(j_i), \tag{6}$$

$$r(j_i) = M \cdot R(i) \cdot W(j_i). \tag{7}$$

Here, HI denotes the homogeneity index, CW denotes the weight coefficient, W denotes the normalized weight coefficient, and r denotes the weighted residual value. The range of HI is set to (0, 1), and a larger value represents a more homogeneous landscape.

The prediction of the total change of a high-resolution pixel between time t_1 and t_2 is predicted as

$$\Delta F(j_i) = r(j_i) + \Delta F_c. \tag{8}$$

Step 6: The final result $\hat{F}_{t_2}$ is obtained with the information in neighborhood as

$$\hat{F}_{t_2}(j_i) = F_{t_1}(j_i) + \sum_{k=1}^{N} W_k \cdot \Delta F(k). \tag{9}$$

Here, W_k is the neighborhood similarity weight for the *kth* similar pixel and N is the number of similar pixels. For a pixel $F_{t_1}(j_i)$, after the N similar pixels are selected, W_k is calculated with the normalized inverse distance as

$$W_k = (1/d_k)/\sum_{k=1}^{N}(1/d_k), \tag{10}$$

where the distance d_k is defined with the spatial locations between $F_{t_1}(j_i)$ and $F_{t_1}(k)$.

A $\mathbf{w} \times \mathbf{w}$ sized window is centered around $F_{t_1}(j_i)$ and searched for the pixels with a similar spectrum to the center pixel. The spectral difference sd_k between $F_{t_1}(j_i)$ and $F_{t_1}(k)$ in its

neighboring window is defined with the ℓ_2 norm where all bands are involved, that is,

$$sd_k = \sqrt{\sum_b \left[F_{t_1}(k,b) - F_{t_1}(j_i,b)\right]^2 / B}, \tag{11}$$

where b denotes the band number and B denotes the number of bands.

After all the spectral differences in a window are obtained, the first N pixels with smallest values (including the center pixel itself) are identified as spectrally similar neighbors. These pixels will be used to update the value of the central pixel with weights according to their distances from the window's center d_k,

$$d_k = 1 + \sqrt{\left\|\mathrm{loc}\left(F_{t_1}(i)\right) - \mathrm{loc}\left(F_{t_1}(k)\right)\right\|^2} / (\mathbf{w}/2), \tag{12}$$

where $loc(\cdot)$ denotes the 2-dimensional coordinate values and $\mathbf{w}$ is the window size.

FSDAF predicts high-resolution images in heterogeneous areas by capturing both gradual and abrupt land cover type changes and retaining more spatial details. However, it cannot capture small type changes in land covers. The smoothness within each class lessens the intra-class variability. The classification accuracy of unsupervised algorithms will also affect the results as very large images cannot be clustered effectively. To conclude, the performance of FSDAF is dominated by the unmixing process of the global linear unmixing model.

2.2 Fit-FC

Wang and Atkinson (2018) proposed the Fit-FC algorithm based on the linear weight models for spatiotemporal fusion. It uses the low-resolution images at time t_1 and t_2 to fit the linear coefficients and then applies the coefficients to the corresponding high-resolution images at time t_1. In order to eliminate the blocky artifacts caused by large differences in resolution, it performs spatial smoothing of fitting values and error values based on neighborhood similar pixels. Fit-FC includes four main steps.

Step 1: Parameters of linear projection are estimated from low-resolution images, and the low-resolution residual image r is calculated. For every band of the two low-resolution images C_{t_1} and C_{t_2} captured at time t_1 and t_2, respectively, a moving window is used to extract blocks $B_{t_1}(i)$ and $B_{t_2}(i)$ for the *ith* location. Given that two groups of pixels $B_{t_1}(i)$ and $B_{t_2}(i)$ in the local window are known, the least square error is minimized to fit the linear model

$$B_{t_2}(i) = a(i)B_{t_1}(i) + b(i), \tag{13}$$

where $a(i)$ and $b(i)$ are the estimated weight and bias for the *ith* location.

After the linear coefficients are obtained, the low-resolution residual image r is calculated pixel-by-pixel with the following equation:

$$r(i) = C_{t_2}(i) - a(i)C_{t_1}(i) - b(i). \tag{14}$$

Step 2: The matrix of two linear coefficients and residuals are upsampled to the ground resolution of the known high-resolution image. The nearest neighboring interpolation is used for linear coefficients, and the bicubic interpolation is used for residuals.

Step 3: The initially predicted high-resolution image $\tilde{F}_{t_2}$ at time t_2 is calculated with the following equation:

$$\tilde{F}_{t_2}(j_i) = a(j_i) \cdot F_{t_1}(j_i) + b(j_i), \tag{15}$$

where j_i is the coordinate of the *jth* high-resolution pixel within the *ith* low-resolution pixel and $a(j_i)$ and $b(j_i)$ are the upsampled linear coefficients at the same location as the known high-resolution pixels $F_{t_1}(j_i)$.

Step 4: Using information in neighborhood to obtain the final result $\hat{F}_{t_2}$,

$$\hat{F}_{t_2}(j_i) = \sum_{k=1}^{n} W_k \left[\tilde{F}_{t_2}(j_i) + r(j_i)\right], \tag{16}$$

where $r(j_i)$ is the upsampled residual values at the same location as the known high-resolution pixels $F_{t_1}(j_i)$. W_k is the neighborhood similarity weight for the *kth* similar pixel, which is calculated in the same way to FSDAF as is shown in **Eq. 10**.

Fit-FC performs well in maintaining spatial and spectral information and is especially suitable for situations where there is a strong time change and the correlation between low-resolution images is small. However, the fused image smoothens spatial details for visual identification.

3 METHODOLOGY: COMPONENT INTEGRATION

In this section, the structure-based spatial integration strategy by Liu et al. (2020) is adopted to combine the images fused by FSDAF and Fit-FC. According to Liu et al. (2020), an image patch can be viewed from its contrast, structure, and luminance, which is valuable to find local complementarity. However, the patch size in the study by Liu et al. (2020) is not suitable for spatiotemporal applications because, under the goal of data fidelity, current fusion algorithms may produce large errors such that the brightness and contrast of small patches are unreliable. Although the local enhancement can improve visual perception, it may lose data fidelity. Therefore, the decomposition is performed in the whole image. The flowchart of the proposed combination method is outlined in **Figure 1**.

An image $\mathbf{x}$ can be decomposed in the form of moments into three components, namely, strength, structure, and mean intensity,

$$\begin{aligned} \mathbf{x} &= \left\|\mathbf{x} - \mu_{\mathbf{x}}\right\|_2 \cdot \frac{\mathbf{x} - \mu_{\mathbf{x}}}{\left\|\mathbf{x} - \mu_{\mathbf{x}}\right\|_2} + \mu_{\mathbf{x}} \\ &= \|\tilde{\mathbf{x}}\|_2 \cdot \frac{\tilde{\mathbf{x}}}{\|\tilde{\mathbf{x}}\|_2} + \mu_{\mathbf{x}} \\ &= c \cdot \mathbf{s} + l, \end{aligned} \tag{17}$$

where $\|\cdot\|_2$ denotes the l_2 norm of a matrix, $\mu_{\mathbf{x}}$ is the mean value, and $\tilde{\mathbf{x}} = \mathbf{x} - \mu_{\mathbf{x}}$ represents a zero-mean image. The scalar $l = \mu_{\mathbf{x}}$, $c = \|\tilde{\mathbf{x}}\|_2$, and the unit-length matrix $\mathbf{s} = \tilde{\mathbf{x}}/\|\tilde{\mathbf{x}}\|$ roughly represent the strength component, structure component, and mean intensity component of $\mathbf{x}$, respectively.

Each fused image can have its own components through decomposition. By integrating the components of multiple fusion results, the new components may outbreak the limitations of different fusion types. The merging strategy will be discussed in detail below.

The visibility of the image structure largely depends on the contrast, which is directly related to the intensity component. Generally, the higher the contrast, the better the visibility. However, too much contrast may lead to unrealistic representation of the image structure. All input images in this study are generated by spatiotemporal fusion algorithms, and their contrasts are usually higher than those of real images. This is reflected in the residual calculation of FSDAF and Fit-FC where stochastic errors are injected as well as details. Consequently, the image with the lowest contrast has the highest fidelity. Therefore, the desired contrast of the composite images is determined by the minimum contrast of all input images, that is, the fusion results of FSDAF and Fit-FC,

$$\hat{c} = \min(c_1, c_2) = \min\left(\|\tilde{\mathbf{x}}_1\|_2, \|\tilde{\mathbf{x}}_2\|_2\right), \tag{18}$$

where $\tilde{\mathbf{x}}_1$ and $\tilde{\mathbf{x}}_2$ represent the zero-mean fusion images of FSDAF and Fit-FC, respectively.

The structure component is defined by the unit matrix $\mathbf{s}$. It is expected that the structure of the fused image can represent the structures of all the input images effectively, which is calculated with the following:

$$\hat{\mathbf{s}} = \sum_i W_i \mathbf{s}_i / \sum_i W_i, \tag{19}$$

where W_i is the weight to determine the contribution of the *ith* image by its structural component $\mathbf{s}_i$.

To increase the contribution of higher-contrast images, a power-weighting function is given by the following:

$$W_i = \|\tilde{\mathbf{x}}_i\|_p, \tag{20}$$

where $p \geq 0$ is a norm limited in 1, 2, or ∞.

The value of p is adaptive to the structure consistency of the input images, which is measured based on the degree of direction consistency R as

$$R = \left\| \sum_i \tilde{\mathbf{x}}_i \right\| / \sum_i \|\tilde{\mathbf{x}}_i\|. \tag{21}$$

The norm p is empirically set to 1 when $R \leq 0.7$, ∞ when $R \geq 0.98$, and 2 otherwise.

The structural strategy is dedicated to the combination of FSDAF and Fit-FC. For the heterogeneous areas, Fit-FC predicts weak details, while the results of FSDAF are rich and relatively accurate. When the above method is used, the structure of FSDAF accounts for a large proportion. For the homogeneous landscapes, Fit-FC predicts fewer details in a more accurate way, while the results of FSDAF are richer but not accurate. In this case, the two images are mixed in a relatively similar ratio to achieve a tradeoff between detail and accuracy.

The intensity component can be estimated with weights as

$$\hat{l} = \sum_i w_i l_i / \sum_i w_i. \tag{22}$$

Here, w_i is the weight normalized with the Gaussian function as given below:

$$w_i = \exp\left(-\frac{(\mu_i - \mu_c)^2}{2\sigma_i^2}\right), \tag{23}$$

where μ_i and σ_i^2 are the mean value and variance of the *ith* image, respectively. μ_c is a constant approaching the mid-intensity value. The typical value of μ_c is 0.5, which is far higher than the mean value of a linearly normalized remote sensing image for visual improvement.

After the combined values $\hat{c}$, $\hat{\mathbf{s}}$, and $\hat{l}$ are calculated, the target image is restored with the following:

$$\hat{\mathbf{x}} = \hat{c} \cdot \hat{\mathbf{s}} + \hat{l}. \tag{24}$$

The integration strategy is performed band by band, which requires the maximum and minimum normalization of all the input images in unified thresholds.

4 EXPERIMENT

4.1 Experimental Scheme

The datasets for validation are the Coleambally irrigation area (CIA) and Lower Gwydir Catchment (LGC) that were used in Emelyanova et al. (2013). CIA has 17 pairs of Landsat-7 ETM + and MODIS images, and LGC has 14 pairs of Landsat-5 TM and MODIS images. Four pairs of Landsat-8 images are also used for the spatiotemporal experiment, which were captured in November 2017 and December 2017. The path number is 121, and the row number is 41 and 43. These images have six bands, of which the blue, green, red, and near-infrared (NIR) bands are reconstructed. All images are cropped to the size of 1200 × 1200 at the center to avoid the outer blank areas. For the CIA and LGC datasets, four pairs of images were used for training and four pairs of images were used to validate the accuracy. For the Landsat-8 dataset, 2 pairs of images were used for training and the other 2 pairs of images were used to validate the accuracy. In each dataset, the two adjacent pairs of images are set as the known image pair and prediction image pair, respectively. The dates of the predicted images are marked in **Tables 1–6**.

To judge the effectiveness of the proposed method, some state-of-the-art algorithms are compared, including STARFM (Gao et al., 2006), SPSTFM (Huang and Song, 2012), EBSCDL (Wu et al., 2015a), FSDAF (Zhu et al., 2016), Fit-FC (Wang and Atkinson, 2018), STFDCNN (Song et al., 2018), and BiaSTF (Li et al., 2020a). STARFM and Fit-FC use linear weights. FSDAF is an unmixing-based method. SPSTFM and EBSCDL

TABLE 1 | RMSE evaluation of radiometric error for the CIA dataset.

Image	Band	Mean	Stdev	STARFM	SPSTFM	EBSCDL	FSDAF	Fit-FC	STFDCNN	BiaSTF	Proposed
	Red	0.0903	0.0381	0.0177	0.0186	0.0181	0.0165	0.0160	0.0189	0.0186	**0.0156**
1	Green	0.0685	0.0285	0.0120	0.0120	0.0119	0.0110	0.0108	0.0118	0.0123	**0.0102**
2001	Blue	0.0406	0.0236	0.0109	0.0109	0.0110	0.0104	**0.0097**	0.0112	0.0111	0.0098
1109	NIR	0.2166	0.0476	0.0350	0.0324	0.0349	0.0313	0.0312	0.0415	0.0378	**0.0292**
	All	0.1040	0.0809	0.0212	0.0203	0.0213	0.0192	0.0190	0.0242	0.0226	**0.0180**
	Red	0.1413	0.0225	0.0275	0.0275	0.0260	0.0251	0.0250	0.0272	0.0263	**0.0248**
2	Green	0.1029	0.0145	0.0180	0.0185	0.0171	**0.0164**	0.0168	0.0181	0.0170	**0.0164**
2001	Blue	0.0677	0.0105	0.0147	0.0148	0.0142	**0.0137**	0.0137	0.0154	0.0140	0.0138
1204	NIR	0.2539	0.0313	0.0386	0.0380	0.0373	0.0357	0.0355	0.0452	0.0387	**0.0351**
	All	0.1414	0.0853	0.0264	0.0263	0.0253	0.0243	0.0243	0.0289	0.0259	**0.0240**
	Red	0.1002	0.0378	0.0224	0.0239	0.0223	0.0203	0.0204	0.0251	0.0233	**0.0199**
3	Green	0.0825	0.0327	0.0139	0.0152	0.0143	0.0127	0.0124	0.0151	0.0150	**0.0122**
2002	Blue	0.0517	0.0225	0.0114	0.0116	0.0113	0.0105	**0.0102**	0.0114	0.0117	0.0103
0222	NIR	0.2724	0.0606	0.0351	0.0341	0.0332	**0.0324**	0.0330	0.0394	0.0353	**0.0324**
	All	0.1267	0.0998	0.0227	0.0229	0.0220	0.0208	0.0210	0.0252	0.0232	**0.0206**
	Red	0.1070	0.0302	0.0186	0.0178	0.0184	0.0169	0.0166	0.0200	0.0190	**0.0164**
4	Green	0.0817	0.0210	0.0130	0.0121	0.0121	0.0114	0.0117	0.0124	0.0122	**0.0112**
2002	Blue	0.0461	0.0167	0.0121	0.0117	0.0119	0.0115	0.0115	0.0123	0.0121	**0.0113**
0317	NIR	0.2524	0.0727	0.0341	0.0304	0.0331	0.0306	0.0304	0.0377	0.0358	**0.0297**
	All	0.1218	0.0922	0.0214	0.0195	0.0207	0.0193	0.0192	0.0231	0.0220	**0.0188**

TABLE 2 | RMSE evaluation of radiometric error for the LGC dataset.

Image	Band	Mean	Stdev	STARFM	SPSTFM	EBSCDL	FSDAF	Fit-FC	STFDCNN	BiaSTF	Proposed
	Red	0.1149	0.0381	0.0173	0.0236	0.0179	0.0155	0.0180	0.0166	0.0183	**0.0150**
1	Green	0.0937	0.0285	0.0141	0.0196	0.0145	0.0126	0.0144	0.0131	0.0147	**0.0120**
2004	Blue	0.0631	0.0236	0.0121	0.0158	0.0124	0.0111	0.0106	0.0102	0.0119	**0.0101**
0502	NIR	0.2131	0.0476	0.0242	0.0318	0.0258	0.0224	0.0221	0.0239	0.0259	**0.0214**
	All	0.1212	0.0665	0.0175	0.0235	0.0184	0.0160	0.0168	0.0167	0.0184	**0.0152**
	Red	0.1224	0.0225	0.0238	0.0470	0.0292	0.0210	0.0196	0.0586	0.0291	**0.0175**
2	Green	0.0951	0.0145	0.0149	0.0225	0.0166	0.0138	0.0142	0.0223	0.0161	**0.0127**
2004	Blue	0.0701	0.0105	0.0120	0.0159	0.0115	0.0094	0.0144	0.0277	0.0112	**0.0085**
1025	NIR	0.2154	0.0313	0.0483	0.1086	0.0739	0.0335	**0.0193**	0.0429	0.0620	0.0209
	All	0.1257	0.0589	0.0286	0.0607	0.0410	0.0215	0.0171	0.0404	0.0356	**0.0156**
	Red	0.0846	0.0378	0.0300	0.0398	0.0301	0.0297	0.0290	0.0292	0.0309	**0.0288**
3	Green	0.0742	0.0327	0.0254	0.0341	0.0256	0.0253	0.0245	0.0252	0.0260	**0.0245**
2004	Blue	0.0513	0.0225	0.0184	0.0239	0.0187	0.0183	**0.0179**	0.0182	0.0189	**0.0173**
1212	NIR	0.1253	0.0606	0.0402	0.0540	0.0408	0.0408	0.0401	0.0395	0.0412	**0.0392**
	All	0.0839	0.0489	0.0296	0.0395	0.0299	0.0297	0.0290	0.0291	0.0304	**0.0286**
	Red	0.0968	0.0302	0.0141	0.0173	0.0145	0.0134	0.0132	0.0181	0.0149	**0.0129**
4	Green	0.0882	0.0210	0.0114	0.0137	0.0113	0.0103	0.0107	0.0134	0.0111	**0.0098**
2005	Blue	0.0642	0.0167	0.0114	0.0123	0.0106	0.0100	0.0119	0.0118	0.0102	**0.0096**
0113	NIR	0.2120	0.0727	0.0299	0.0406	0.0313	0.0301	**0.0272**	0.0384	0.0317	0.0275
	All	0.1153	0.0706	0.0184	0.0239	0.0189	0.0180	0.0171	0.0230	0.0191	**0.0167**

TABLE 3 | RMSE evaluation of radiometric error for the Landsat-8 dataset.

Data	Band	Mean	Stdev	STARFM	SPSTFM	EBSCDL	FSDAF	Fit-FC	STFDCNN	BiaSTF	Proposed
	Red	0.0374	0.0202	0.0093	0.0082	0.0087	0.0081	0.0079	0.0089	0.0094	**0.0078**
1–41	Green	0.0416	0.0158	0.0074	0.0064	0.0070	0.0067	0.0067	**0.0060**	0.0073	0.0066
2017	Blue	0.0281	0.0127	0.0075	**0.0062**	0.0072	0.0070	0.0072	0.0070	0.0069	0.0069
1219	NIR	0.1784	0.0584	0.0227	0.0220	0.0211	0.0210	0.0209	0.0544	0.0244	**0.0204**
	All	0.0714	0.0700	0.0133	0.0126	0.0125	0.0123	0.0122	0.0279	0.0140	**0.0119**
	Red	0.0435	0.0259	0.0090	0.0091	0.0086	0.0083	0.0084	0.0108	0.0095	**0.0080**
2–43	Green	0.0505	0.0211	0.0078	0.0079	0.0068	0.0068	0.0073	0.0086	0.0075	**0.0067**
2017	Blue	0.0302	0.0150	0.0061	**0.0051**	0.0056	0.0056	0.0059	0.0056	0.0058	0.0055
1219	NIR	0.2326	0.0727	0.0265	0.0230	**0.0228**	0.0229	0.0252	0.0474	0.0265	0.0236
	All	0.0892	0.0925	0.0148	0.0132	**0.0130**	**0.0130**	0.0141	0.0248	0.0148	0.0132

TABLE 4 | SSIM evaluation of structural discrepancy for the CIA dataset.

Image	Band	STARFM	SPSTFM	EBSCDL	FSDAF	Fit-FC	STFDCNN	BiaSTF	Proposed
	Red	0.8861	0.8917	0.8873	0.9062	0.8953	0.8805	0.8808	**0.9064**
1	Green	0.8911	0.9063	0.8989	0.9134	0.9036	0.8989	0.8912	**0.9154**
2001	Blue	0.8860	0.9092	0.9009	0.9143	0.9046	0.8958	0.8951	**0.9150**
1109	NIR	0.9849	0.9874	0.9843	0.9882	0.9872	0.9791	0.9807	**0.9894**
	All	0.9125	0.9240	0.9183	0.9309	0.9232	0.9141	0.9124	**0.9319**
	Red	0.8325	0.8544	0.8539	0.8673	0.8694	0.8507	0.8465	**0.8704**
2	Green	0.8586	0.8685	0.8730	**0.8885**	0.8734	0.8719	0.8676	0.8841
2001	Blue	0.8701	0.8865	0.8865	**0.9010**	0.8837	0.8799	0.8830	0.8948
1204	NIR	0.8243	**0.8558**	0.8457	**0.8558**	0.8470	0.8073	0.8337	0.8549
	All	0.8482	0.8684	0.9667	**0.8800**	0.8705	0.8546	0.8597	0.8780
	Red	0.9570	0.9542	0.9557	0.9654	0.9631	0.9486	0.9506	**0.9658**
3	Green	0.8840	0.8910	0.8868	0.9080	0.9004	0.8782	0.8770	**0.9096**
2002	blue	0.8821	0.9007	0.8968	0.9153	0.9067	0.8962	0.8884	**0.9154**
0222	NIR	0.8750	**0.9021**	0.8960	0.9016	0.8932	0.8724	0.8844	0.8985
	All	0.9010	0.9132	0.9103	0.9236	0.9172	0.9005	0.9018	**0.9238**
	Red	0.9000	0.9229	0.9098	0.9230	0.9212	0.8979	0.9023	**0.9242**
4	Green	0.8945	0.9248	0.9187	**0.9263**	0.9186	0.9159	0.9137	0.9253
2002	Blue	0.8930	**0.9318**	0.9212	0.9299	0.9176	0.9146	0.9156	0.9270
0317	NIR	0.9112	**0.9397**	0.9230	0.9338	0.9286	0.9041	0.9097	0.9338
	All	0.8999	**0.9299**	0.9185	0.9285	0.9216	0.9084	0.9107	0.9277

TABLE 5 | SSIM evaluation of radiometric error for the LGC dataset.

Image	Band	STARFM	SPSTFM	EBSCDL	FSDAF	Fit-FC	STFDCNN	BiaSTF	Proposed
	Red	0.8809	0.8726	0.8760	0.9059	0.8915	0.8949	0.8763	**0.9070**
1	Green	0.8823	0.8711	0.8777	0.9086	0.8991	0.9002	0.8785	**0.9123**
2004	Blue	0.8882	0.8856	0.8833	0.9126	0.9129	0.9064	0.8855	**0.9194**
0502	NIR	0.8555	0.8436	0.8460	0.8788	0.8784	0.8606	0.8475	**0.8851**
	All	0.8772	0.8674	0.8713	0.9017	0.8957	0.8913	0.8725	**0.9061**
	Red	0.9203	0.8207	0.8809	0.9467	0.9597	0.6999	0.8833	**0.9614**
2	Green	0.8573	0.7822	0.8139	0.8840	0.9007	0.7155	0.8166	**0.9018**
2004	Blue	0.9240	0.8987	0.9193	0.9573	0.9469	0.7557	0.9209	**0.9667**
1025	NIR	0.6395	0.4610	0.5004	0.7611	**0.8629**	0.6366	0.5304	0.8607
	All	0.8377	0.7428	0.7821	0.8891	0.9190	0.7057	0.7910	**0.9243**
	Red	0.6315	0.5527	0.6220	0.6128	0.6290	**0.6461**	0.6217	0.6317
3	Green	0.6316	0.5504	0.6207	0.6109	0.6295	**0.6384**	0.6234	0.6300
2004	Blue	0.6139	0.5499	0.6062	0.6038	0.6156	**0.6361**	0.6080	0.6261
1212	NIR	0.6249	0.5497	0.6219	0.6219	0.6213	0.6174	0.6232	**0.6371**
	All	0.6252	0.5498	0.6176	0.6122	0.6240	**0.6347**	0.6188	0.6312
	Red	0.8946	0.8868	0.8927	0.9078	0.9074	0.8602	0.8929	**0.9114**
4	Green	0.8916	0.8851	0.8899	0.9103	0.9060	0.8562	0.8917	**0.9135**
2005	Blue	0.8748	0.8756	0.8793	0.8985	0.8915	0.8433	0.8813	**0.9042**
0113	NIR	0.8572	0.8458	0.8528	0.8667	**0.8843**	0.8154	0.8532	0.8833
	All	0.8797	0.8724	0.8790	0.8961	0.8973	0.8438	0.8800	**0.9032**

TABLE 6 | SSIM evaluation of radiometric error for the Landsat-8 dataset.

Image	Band	STARFM	SPSTFM	EBSCDL	FSDAF	Fit-FC	STFDCNN	BiaSTF	Proposed
	Red	0.9752	0.9811	0.9776	0.9809	0.9814	0.9780	0.9721	**0.9817**
1–41	Green	0.9251	0.9524	0.9435	0.9471	0.9415	**0.9541**	0.9267	0.9461
2017	Blue	0.9791	**0.9865**	0.9788	0.9810	0.9791	0.9853	0.9820	0.9810
1219	NIR	0.9056	0.9090	0.9160	0.9190	0.9095	0.8933	0.8810	**0.9191**
	All	0.9467	0.9572	0.9543	0.9573	0.9534	0.9529	0.9411	**0.9575**
	Red	0.9800	0.9821	0.9801	0.9835	0.9818	0.9811	0.9735	**0.9837**
2–43	Green	0.9842	0.9871	0.9871	**0.9890**	0.9863	0.9854	0.9839	**0.9890**
2017	Blue	0.9778	**0.9866**	0.9803	0.9824	0.9780	0.9864	0.9787	0.9810
1219	NIR	0.9126	**0.9377**	0.9361	0.9369	0.9118	0.9290	0.9165	0.9272
	All	0.9641	**0.9738**	0.9712	0.9734	0.9650	0.9707	0.9636	0.9705

TABLE 7 | Hardware and software for experiment.

Hardware	RAM	CPU	GPU
	62.6G	2 × Intel Xeon E5-2620 v4	2 × Tesla V100
Software	PYTHON	CUDA	PyTorch
	3.6.2	9.0	1.2.0
	MATLAB	RAM	CPU
	R2018b	16.0 GB	Intel(R) Core(TM) i7-6700 CPU at 3.40 GHz

TABLE 8 | SAM evaluation of spectral inconsistency.

Dataset	**CIA**				**LGC**				**Landsat-8**	
Image	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**
STARFM	0.0891	0.0728	0.0723	0.0674	0.0664	0.1215	0.1443	0.0742	0.0646	0.0443
SPSTFM	0.0938	0.0760	0.0638	**0.0567**	0.0681	0.3511	0.1931	0.0802	**0.0577**	**0.0346**
EBSCDL	0.0934	0.0685	0.0665	0.0657	0.0631	0.1675	0.1400	0.0676	0.0637	0.0412
FSDAF	0.0789	0.0644	**0.0620**	0.0595	0.0539	0.0964	0.1513	0.0674	0.0593	0.0403
Fit-FC	0.0674	0.0619	0.0656	0.0587	0.0552	0.0694	0.1419	0.0729	0.0589	0.0426
STFDCNN	0.0853	0.0744	0.0714	0.0686	0.0543	0.1810	**0.1275**	0.0662	0.0888	0.0447
BiaSTF	0.1019	0.0687	0.0725	0.0713	0.0639	0.1614	0.1400	0.0667	0.0639	0.0495
Proposed	**0.0661**	**0.0617**	**0.0620**	0.0569	**0.0516**	**0.0660**	0.1370	**0.0645**	**0.0577**	0.0391

TABLE 9 | ERGAS evaluation of spectral inconsistency.

Dataset	**CIA**				**LGC**				**Landsat-8**	
Image	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**
STARFM	0.2040	0.1863	0.1897	0.1892	0.1541	0.1886	0.3443	0.1493	0.2117	0.1732
SPSTFM	0.2048	0.1873	0.1983	0.1787	0.2069	0.3567	0.4569	0.1798	**0.1848**	0.1631
EBSCDL	0.2057	0.1784	0.1886	0.1840	0.1592	0.2408	0.3481	0.1482	0.2009	0.1597
FSDAF	0.1906	0.1715	0.1731	0.1740	0.1396	0.1523	0.3437	0.1390	0.1936	0.1565
Fit-FC	0.1835	0.1719	0.1712	0.1738	0.1476	0.1567	0.3359	0.1452	0.1933	0.1643
STFDCNN	0.2162	0.1947	0.2038	0.1948	0.1407	0.3462	0.3390	0.1766	0.2410	0.2038
BiaSTF	0.2121	0.1789	0.1968	0.1888	0.1583	0.2203	0.3536	0.1477	0.2080	0.1730
Proposed	**0.1795**	**0.1709**	**0.1696**	**0.1707**	**0.1312**	**0.1249**	**0.3303**	**0.1320**	0.1894	**0.1543**

TABLE 10 | Q4 evaluation of spectral inconsistency (R/G/B).

Dataset	**CIA**				**LGC**				**Landsat-8**	
Image	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**
STARFM	0.8636	0.8543	0.8939	0.8963	0.8947	0.6207	0.6804	0.8811	0.8791	0.9289
SPSTFM	0.8636	0.8360	0.8918	0.9142	0.8386	0.2050	0.2242	0.8200	**0.9132**	0.9338
EBSCDL	0.8688	0.8684	0.9012	0.9099	0.8945	0.5289	0.6806	0.8820	0.8968	0.9422
FSDAF	0.8832	0.8767	**0.9116**	0.9182	0.9107	0.6757	0.6671	0.8924	0.9010	0.9414
Fit-FC	0.8740	0.8749	0.9025	0.9138	0.8955	0.7101	0.6794	0.8881	0.8981	0.9342
STFDCNN	0.8617	0.8695	0.8812	0.8991	0.9096	0.2713	**0.6888**	0.8433	0.9056	0.9178
BiaSTF	0.8614	0.8649	0.8921	0.9047	0.8944	0.5537	0.6836	0.8811	0.8892	0.9331
Proposed	**0.8846**	**0.8774**	0.9104	**0.9186**	**0.9109**	**0.7342**	0.6817	**0.8958**	**0.9132**	**0.9423**

are based on the coupled dictionary learning. STFDCNN and BiaSTF were recently proposed that use the CNNs and deep learning.

The default parameter settings were kept for all competing algorithms. For STFDCNN, the SGD optimizer was used in the training, the batch size was set as 64, the training iterated 300 epochs with the learning rate of the first two layers set to 1×10^{-4} and the last layer to 1×10^{-5}, and the training images were cropped into patches with a size of 64×64 for learning purposes. For BiaSTF, the Adam optimizer was used in the training by setting $\beta_1 = 0.9$, $\beta_2 = 0.999$, and $\epsilon = 10^{-8}$; the batch size was set as 64, the training iterated 300 epochs with the

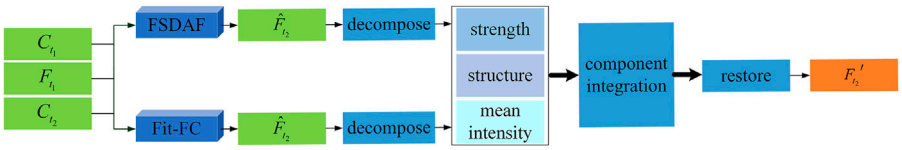

FIGURE 1 | Flowchart of the proposed combination method.

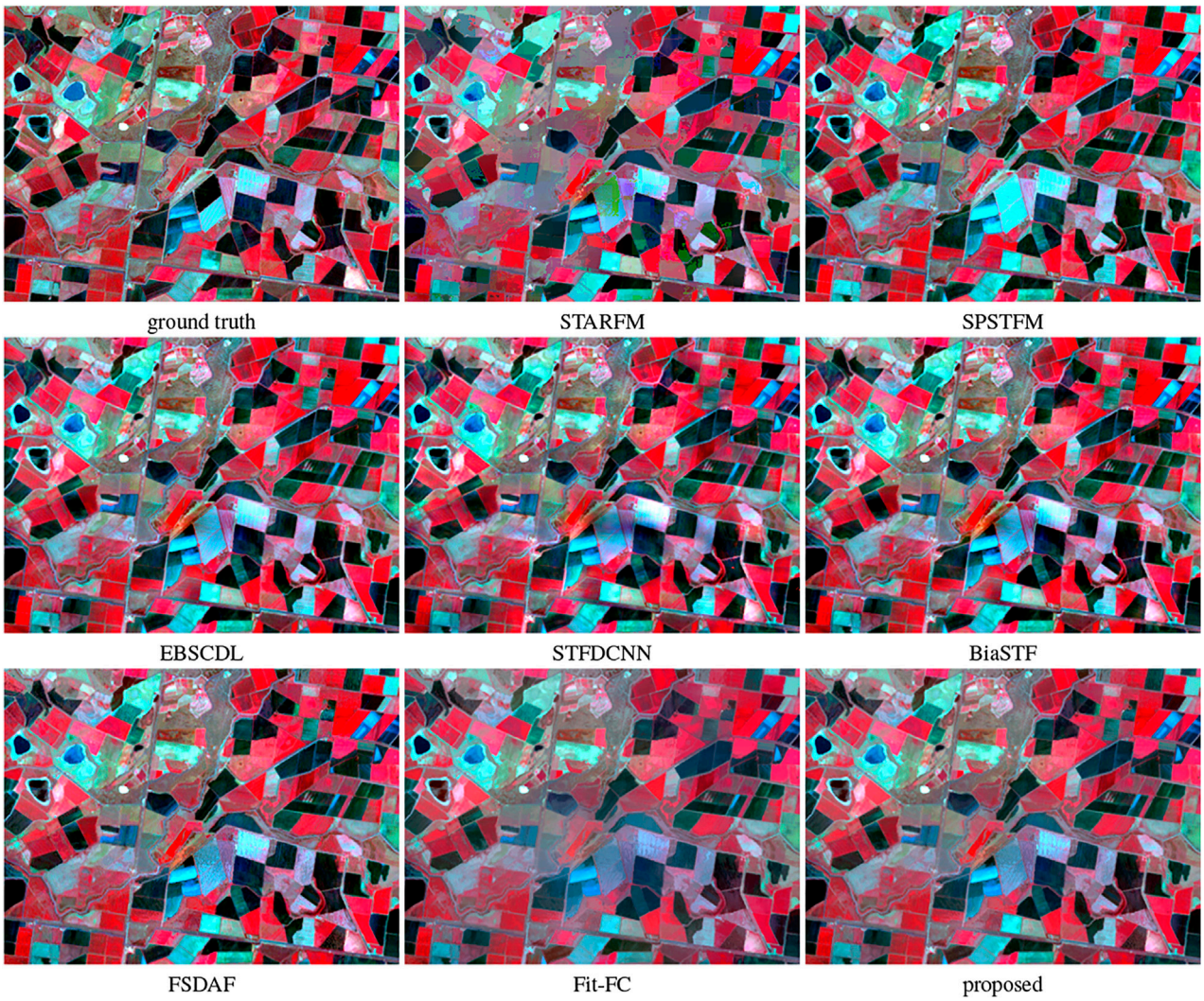

FIGURE 2 | Manifestation of the small region of the NIR, red, and green bands of CIA image 1 for detail observation.

learning rate set as 1×10^{-4}, and the training images were cropped into patches with a size of 128 × 128 for learning purposes. The experimental environment is listed in **Table 7**.

Metrics are used to evaluate the loss of radiation, the structure, and the spectrum. Root-mean-square-error (RMSE) measures the radiometric error. Structural similarity (SSIM) measures the similarity of contours and shapes. The Spectral Angle Mapper (SAM), Erreur Relative Globale Adimensionnelle de Synthese (ERGAS) (Du et al., 2007), and a Quaternion theory-based quality index (Q4) (Alparone et al., 2004) measure the spectral consistency. RMSE and SSIM are calculated band by band, while ERGAS and Q4 are calculated with the NIR, red, green, and blue bands as a whole. The ideal values are 1 for SSIM and Q4 while 0 for RMSE, SAM, and ERGAS.

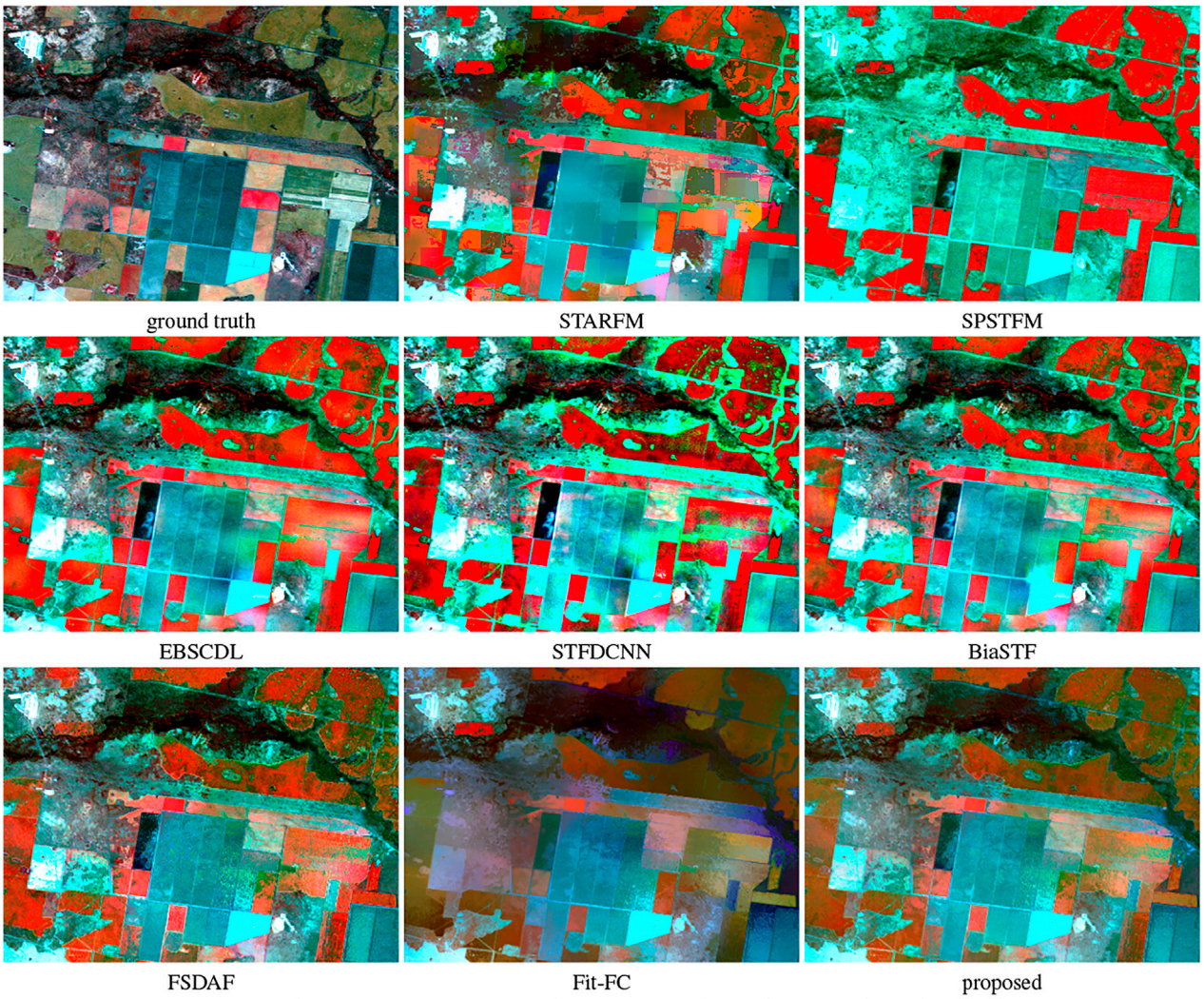

FIGURE 3 | Manifestation of the small region of the NIR, red, and green bands of LGC image 2 for detail observation.

4.2 Radiometric and Structural Assessment

RMSE and SSIM are calculated band by band. To save space, four fusion results are listed for each dataset, which are evaluated with RMSE in **Tables 1–3**, SSIM in **Tables 4–6**, SAM in **Table 8**, ERGAS in **Table 9**, and Q4 in **Table 10**. The best scores are marked in bold, and the better ones between scores of FSDAF and Fit-FC are underlined.

Table 1 shows the radiometric error of Landsat-7 reconstruction. It is clear that FSDAF and Fit-FC can produce more competitive results than dictionary learning– and deep learning–based methods. Compared with FSDAF, Fit-FC works better for image 1 but shows equal advantages for images 2, 3, and 4. The proposed method produces the least radiometric loss in majority cases.

The radiometric error of Landsat-5 is assessed in **Table 2**. It is observed that the performance of FSDAF, Fit-FC, and STFDCNN is accompanied with large fluctuation in image 3 due to the quick change caused by floods. Fit-FC ranks higher than FSDAF for the NIR band. STARFM, EBSCDL, and BiaSTF show better performance than SPSTFM. Again, the proposed method produces the least radiometric loss in most cases.

The radiometric error of Landsat-8 is assessed in **Table 3**. The two dictionary-learning methods, SPSTFM and EBSCDL, perform well in the blue and NIR bands. Fit-FC performs poorly on image 43, making the proposed method slightly worse than FSDAF. It can also be seen that the method proposed in this study is suitable for the fusion of two results with little difference to produce a better result. When the two results differ greatly, the combination shows high stability.

The structural similarity is measured in **Tables 4–6**. The digital differences between algorithms are small. For Landsat-7 (**Table 4**), FSDAF shows strong superiority than Fit-FC, while the advantage is weak for image 2 of Landsat-5 (**Table 5**). STFDCNN and dictionary learning–based methods show good structural reconstruction for

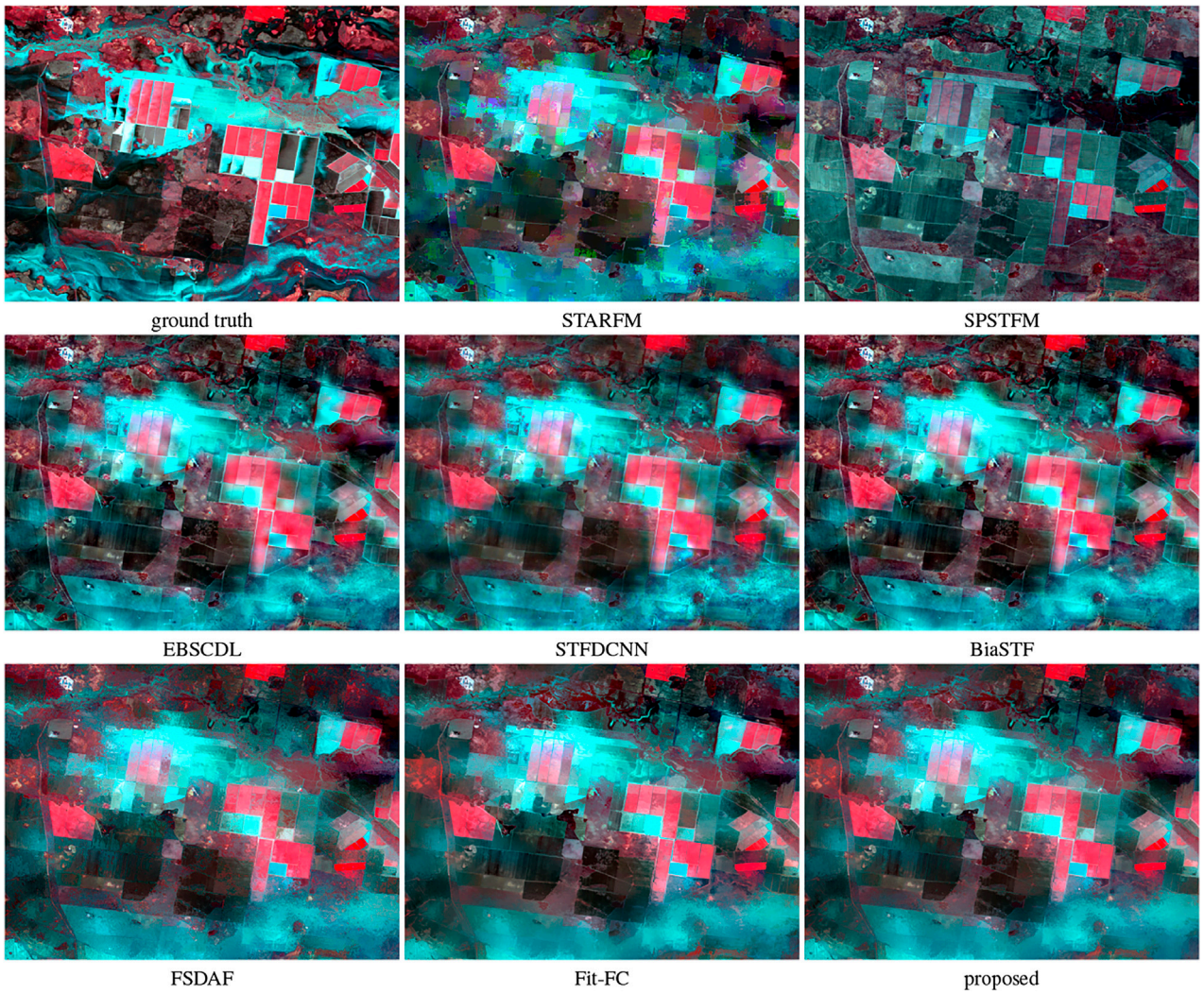

FIGURE 4 | Manifestation of the large region of the NIR, red, and green bands of LGC image 3 for flood observation.

Landsat-7 and Landsat-8. For Landsat-5, STFDCNN works well for images 1 and 3 but poorly for image 2. The proposed method works steadily well in preserving good structures.

4.3 Spectral Assessment

SAM is assessed in **Table 8** with the NIR, red, green, and blue bands as a whole. SPSTFM works well for Landsat-8 but poor for Landsat-5. FSDAF and Fit-FC can produce better results for various datasets. The proposed method gives the best scores for the majority of images.

ERGAS and Q4 for spectral assessment are calculated with the NIR, red, green, and blue bands as a whole. ERGAS is assessed in **Table 9**. The majority of the algorithms work well except for SPSTFM. FSDAF shows better performance than Fit-FC for Landsat-7 but poorer for Landsat-5. The proposed method gives the best scores for all images.

Q4 is listed in **Table 10** for spectral observation with the red, green, and blue bands as a whole. Images 2 and 3 of Landsat-5 are challenging due to the quick change of ground content, where dictionary-based and CNN-based methods produce much poor results. FSDAF and Fit-FC work well for most images. The proposed method shows competitive performance as it gives the best scores for the majority of images.

4.4 Visual Comparison

Four groups of images are demonstrated in **Figures 2–5** for visual identification of the NIR, red, and green bands. All images are linearly stretched with the thresholds by which the brightest and darkest 2% pixels of the ground truth images are reassigned band by band. In this way, the color distortion can be read from the visually enhanced images directly. The manifested images in **Figures 2**, **3**, **5** illustrate that FSDAF produces more details while Fit-FC fuses more consistent colors. Our method adopts both the advantages effectively to approach the true image. The flood area in **Figure 4** shows that none of the algorithms can reconstruct the quick change

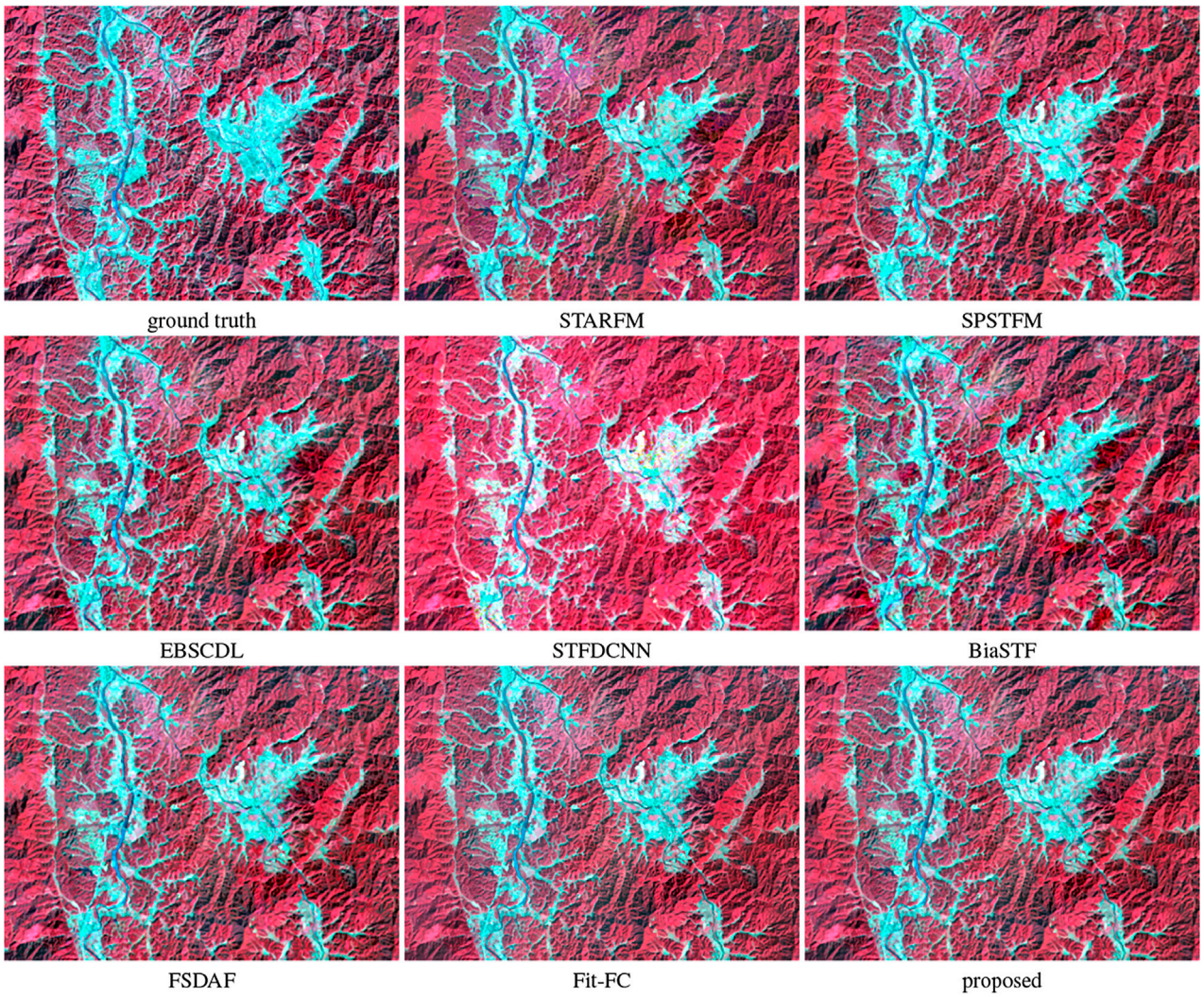

FIGURE 5 | Manifestation of the small region of the NIR, red, and green bands of Landsat-8 image 1 for detail observation.

TABLE 11 | Computational cost.

Algorithm	Code language	Running time (seconds)
STARFM	Python	30
SPSTFM	MATLAB	615
EBSCDL	MATLAB	5150
FSDAF	IDL	660
Fit-FC	MATLAB	1300
STFDCNN	Python	1430
BiaSTF	Python	2200
Proposed	MATLAB	1300 + 660 + 6

in a large region yet despite the effort of FSDAF on changed landscapes.

4.5 Computational Cost

The consumed time in a single prediction is recorded in **Table 11**, in which all the Python code used GPUs (nVidia 2080Ti) for acceleration. It is not fair to compare the time directly because the codes use various programming languages. For our method, the integration process takes only 6 s to combine the fusion results of FSDAF and Fit-FC. Since the fusion algorithms can work in a parallel way, the consumed time for the proposed method is recorded as the longest time plus the combination strategy.

5 DISCUSSION

The stability of our method is worthy of noting. On the one hand, derived from the excellent original methods, our synthetic method hits the highest score in most cases. By comparing the digital evaluation, it is concluded that the proposed method is usually better than the results of FSDAF and Fit-FC, which proves the complementarity indirectly. On the other hand, when our method fails to produce the best results, its score is close to the highest score.

The experiment shows that the proposed method may be improved. The RMSE comparison shows that Fit-FC is weakly better than FSDAF, but the SSIM comparison gives a contrary conclusion. Even though our proposed method is much effective, it does not make full use of the conclusion. To design a more feasible integration strategy, more tests are required to identify the unique advantages of spatiotemporal fusion algorithms, which are prevented in this study by the limited space.

For spatiotemporal fusion, there is no similar method focusing on integrating the fusion results for better performance. The only analogous method was proposed by Chen et al. (2020), who discussed the issue of data selection for performance improvement. Different kinds of algorithms have different advantages. Then, a good algorithm can design complex processes that incorporate multiple kinds for higher quality, or it can integrate the results through post-processing as the method in this article did. Intuitively, the idea in this article can be used for more remote sensing issues, such as pansharpening, denoising, inpainting, and so on.

The main disadvantage of the method is the increased time. As can be seen from **Table 11**, the post-processing time is very short so we have to run two or more different algorithms that extend the total time. This can be partly solved by launching algorithms in a parallel way. Then, the total time is constrained by the slowest algorithm.

The proposed method is usually not sensitive to the data quality of the input images. Some of the fusion results may be poor for specific images, while the proposed method tends to choose the best image block from multiple inputs. For them, the targeted selection of the fusion result, that is, the merger strategy, is the key. By performing this operation block by block, the quality of the whole image is improved.

6 CONCLUSION

Aiming at the insufficient stability of spatiotemporal fusion algorithms, this study proposes to make use of the complementarity of spatiotemporal fusion algorithms for better fusion results. An integration strategy is proposed for the images fused by FSDAF and Fit-FC. Their fusion results are decomposed into a strength component, a structure component, and a mean intensity component, which are packed to form a new fusion image.

The proposed method is tested on Landsat-5, Landsat-7, and Landat-8 images and compared with seven algorithms of four different types. The experimental results confirm the effectiveness of the spatial fusion strategy. The quantitative evaluation on radiometric, structural, and spectral loss shows that images produced by our method can reach or approach the optimal performance.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: https://data.csiro.au/collections/#collection/CIcsiro:5846v1 and https://data.csiro.au/collections/#collection/CIcsiro:5847v1.

AUTHOR CONTRIBUTIONS

JW proposed the idea and wrote the paper. YM made the program and experiment. XH provided suggestions for data processing.

FUNDING

This paper was supported by the National Natural Science Foundation of China (No. 61860130) and the 03 Special and 5G Project of the Jiangxi Province (No. 20204ABC03A40).

REFERENCES

Alparone, L., Baronti, S., Garzelli, A., and Nencini, F. (2004). A Global Quality Measurement of Pan-Sharpened Multispectral Imagery. *IEEE Geosci. Remote Sens. Lett.* 1, 313–317. doi:10.1109/lgrs.2004.836784

Chen, Y., Cao, R., Chen, J., Zhu, X., Zhou, J., Wang, G., et al. (2020). A New Cross-Fusion Method to Automatically Determine the Optimal Input Image Pairs for Ndvi Spatiotemporal Data Fusion. *IEEE Trans. Geosci. Remote Sens.* 58, 5179–5194. doi:10.1109/tgrs.2020.2973762

Choi, J. H., Elgendy, O. A., and Chan, S. H. (2019). Optimal Combination of Image Denoisers. *IEEE Trans. Image Process.* 28, 4016–4031. doi:10.1109/tip.2019.2903321

Dai, P., Zhang, H., Zhang, L., and Shen, H. (2018). "A Remote Sensing Spatiotemporal Fusion Model of Landsat and Modis Data via Deep Learning," in IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium, 7030–7033. doi:10.1109/igarss.2018.8518758

Du, Q., Younan, N. H., King, R., and Shah, V. P. (2007). On the Performance Evaluation of Pan-Sharpening Techniques. *IEEE Geosci. Remote Sens. Lett.* 4, 518–522. doi:10.1109/lgrs.2007.896328

Emelyanova, I. V., McVicar, T. R., Van Niel, T. G., Li, L. T., and van Dijk, A. I. J. M. (2013). Assessing the Accuracy of Blending Landsat-Modis Surface Reflectances in Two Landscapes with Contrasting Spatial and Temporal Dynamics: A Framework for Algorithm Selection. *Remote Sens. Environ.* 133, 193–209. doi:10.1016/j.rse.2013.02.007

Fu, D., Chen, B., Wang, J., Zhu, X., and Hilker, T. (2013). An Improved Image Fusion Approach Based on Enhanced Spatial and Temporal the Adaptive Reflectance Fusion Model. *Remote Sens.* 5, 6346–6360. doi:10.3390/rs5126346

Feng Gao, F., Masek, J., Schwaller, M., and Hall, F. (2006). On the Blending of the Landsat and Modis Surface Reflectance: Predicting Daily Landsat Surface Reflectance. *IEEE Trans. Geosci. Remote Sens.* 44, 2207–2218. doi:10.1109/tgrs.2006.872081

Gevaert, C. M., and García-Haro, F. J. (2015). A Comparison of Starfm and an Unmixing-Based Algorithm for Landsat and Modis Data Fusion. *Remote Sens. Environ.* 156, 34–44. doi:10.1016/j.rse.2014.09.012

Hilker, T., Wulder, M. A., Coops, N. C., Linke, J., McDermid, G., Masek, J. G., et al. (2009). A New Data Fusion Model for High Spatial- and Temporal-Resolution Mapping of forest Disturbance Based on Landsat and Modis. *Remote Sens. Environ.* 113, 1613–1627. doi:10.1016/j.rse.2009.03.007

Huang, B., and Song, H. (2012). Spatiotemporal Reflectance Fusion via Sparse Representation. *IEEE Trans. Geosci. Remote Sens.* 50, 3707–3716. doi:10.1109/tgrs.2012.2186638

Bo Huang, B., Juan Wang, J., Huihui Song, H., Dongjie Fu, D., and KwanKit Wong, K. (2013). Generating High Spatiotemporal Resolution Land Surface

Temperature for Urban Heat Island Monitoring. *IEEE Geosci. Remote Sens. Lett.* 10, 1011–1015. doi:10.1109/lgrs.2012.2227930

Li, X., Wang, L., Cheng, Q., Wu, P., Gan, W., and Fang, L. (2019). Cloud Removal in Remote Sensing Images Using Nonnegative Matrix Factorization and Error Correction. *Isprs J. Photogramm. Remote Sens.* 148, 103–113. doi:10.1016/j.isprsjprs.2018.12.013

Li, Y., Li, J., He, L., Chen, J., and Plaza, A. (2020a). A New Sensor Bias-Driven Spatio-Temporal Fusion Model Based on Convolutional Neural Networks. *Sci. China-Inform. Sci.* 63. doi:10.1007/s11432-019-2805-y

Li, Y., Wu, H., Li, Z.-L., Duan, S., and Ni, L. (2020b). "Evaluation of Spatiotemporal Fusion Models in Land Surface Temperature Using Polar-Orbiting and Geostationary Satellite Data," in IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium, 236–239. doi:10.1109/igarss39084.2020.9323319

Liu, X., Deng, C., and Zhao, B. (2016). "Spatiotemporal Reflectance Fusion Based on Location Regularized Sparse Representation," in 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 2562–2565. doi:10.1109/igarss.2016.7729662

Liu, P., Zhang, H., and Eom, K. B. (2017). Active Deep Learning for Classification of Hyperspectral Images. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* 10, 712–724. doi:10.1109/jstars.2016.2598859

Liu, X., Deng, C., Chanussot, J., Hong, D., and Zhao, B. (2019). Stfnet: A Two-Stream Convolutional Neural Network for Spatiotemporal Image Fusion. *IEEE Trans. Geosci. Remote Sens.* 57, 6552–6564. doi:10.1109/tgrs.2019.2907310

Liu, Y., Xu, S., and Lin, Z. (2020). An Improved Combination of Image Denoisers Using Spatial Local Fusion Strategy. *IEEE Access* 8, 150407–150421. doi:10.1109/access.2020.3016766

Song, H., and Huang, B. (2013). Spatiotemporal Satellite Image Fusion through One-Pair Image Learning. *IEEE Trans. Geosci. Remote Sens.* 51, 1883–1896. doi:10.1109/tgrs.2012.2213095

Song, H., Liu, Q., Wang, G., Hang, R., and Huang, B. (2018). Spatiotemporal Satellite Image Fusion Using Deep Convolutional Neural Networks. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* 11, 821–829. doi:10.1109/jstars.2018.2797894

Tan, Z., Yue, P., Di, L., and Tang, J. (2018). Deriving High Spatiotemporal Remote Sensing Images Using Deep Convolutional Network. *Remote Sens.* 10, 1066. doi:10.3390/rs10071066

Tan, Z., Di, L., Zhang, M., Guo, L., and Gao, M. (2019a). An Enhanced Deep Convolutional Model for Spatiotemporal Image Fusion. *Remote Sens.* 11, 2898. doi:10.3390/rs11242898

Tan, Z. Q., Wang, X. L., Chen, B., Liu, X. G., and Zhang, Q. (2019b). Surface Water Connectivity of Seasonal Isolated Lakes in a Dynamic lake-floodplain System. *J. Hydrol.* 579, 13. doi:10.1016/j.jhydrol.2019.124154

Tan, Z., Gao, M., Li, X., and Jiang, L. (2021). A Flexible Reference-Insensitive Spatiotemporal Fusion Model for Remote Sensing Images Using Conditional Generative Adversarial Network. *IEEE Trans. Geosci. Remote Sens.*, 1–13. doi:10.1109/tgrs.2021.3050551

Wang, Q., and Atkinson, P. M. (2018). Spatio-temporal Fusion for Daily Sentinel-2 Images. *Remote Sens. Environ.* 204, 31–42. doi:10.1016/j.rse.2017.10.046

Wang, J., and Huang, B. (2017). A Rigorously-Weighted Spatiotemporal Fusion Model with Uncertainty Analysis. *Remote Sens.* 9, 990. doi:10.3390/rs9100990

Wei, J., Wang, L., Liu, P., Chen, X., Li, W., and Zomaya, A. Y. (2017a). Spatiotemporal Fusion of Modis and Landsat-7 Reflectance Images via Compressed Sensing. *IEEE Trans. Geosci. Remote Sens.* 55, 7126–7139. doi:10.1109/tgrs.2017.2742529

Wei, J., Wang, L., Liu, P., and Song, W. (2017b). Spatiotemporal Fusion of Remote Sensing Images with Structural Sparsity and Semi-coupled Dictionary Learning. *Remote Sens.* 9, 21. doi:10.3390/rs9010021

Weng, Q., Fu, P., and Gao, F. (2014). Generating Daily Land Surface Temperature at Landsat Resolution by Fusing Landsat and Modis Data. *Remote Sens. Environ.* 145, 55–67. doi:10.1016/j.rse.2014.02.003

Wu, M., Niu, Z., Wang, C., Wu, C., and Wang, L. (2012). Use of Modis and Landsat Time Series Data to Generate High-Resolution Temporal Synthetic Landsat Data Using a Spatial and Temporal Reflectance Fusion Model. *J. Appl. Remote Sens.* 6, 063507. doi:10.1117/1.jrs.6.063507

Wu, B., Huang, B., and Zhang, L. (2015a). An Error-Bound-Regularized Sparse Coding for Spatiotemporal Reflectance Fusion. *IEEE Trans. Geosci. Remote Sens.* 53, 6791–6803. doi:10.1109/tgrs.2015.2448100

Wu, M., Huang, W., Niu, Z., and Wang, C. (2015b). Generating Daily Synthetic Landsat Imagery by Combining Landsat and Modis Data. *Sensors* 15, 24002–24025. doi:10.3390/s150924002

Yong Xu, Y., Bo Huang, B., Yuyue Xu, Y., Kai Cao, K., Chunlan Guo, C., and Deyu Meng, D. (2015). Spatial and Temporal Image Fusion via Regularized Spatial Unmixing. *IEEE Geosci. Remote Sens. Lett.* 12, 1362–1366. doi:10.1109/lgrs.2015.2402644

Zhang, W., Li, A., Jin, H., Bian, J., Zhang, Z., Lei, G., et al. (2013). An Enhanced Spatial and Temporal Data Fusion Model for Fusing Landsat and Modis Surface Reflectance to Generate High Temporal Landsat-like Data. *Remote Sens.* 5, 5346–5368. doi:10.3390/rs5105346

Zhang, L., Liu, P., Zhao, L., Wang, G., Zhang, W., and Liu, J. (2021). Air Quality Predictions with a Semi-supervised Bidirectional Lstm Neural Network. *Atmos. Pollut. Res.* 12, 328–339. doi:10.1016/j.apr.2020.09.003

Zhu, X., Chen, J., Gao, F., Chen, X., and Masek, J. G. (2010). An Enhanced Spatial and Temporal Adaptive Reflectance Fusion Model for Complex Heterogeneous Regions. *Remote Sens. Environ.* 114, 2610–2623. doi:10.1016/j.rse.2010.05.032

Zhu, X., Helmer, E. H., Gao, F., Liu, D., Chen, J., and Lefsky, M. A. (2016). A Flexible Spatiotemporal Method for Fusing Satellite Images with Different Resolutions. *Remote Sens. Environ.* 172, 165–177. doi:10.1016/j.rse.2015.11.016

Zhu, X., Cai, F., Tian, J., and Williams, T. K.-A. (2018). Spatiotemporal Fusion of Multisource Remote Sensing Data: Literature Survey, Taxonomy, Principles, Applications, and Future Directions. *Remote Sens.* 10, 527. doi:10.3390/rs10040527

Zhukov, B., Oertel, D., Lanzl, F., and Reinhackel, G. (1999). Unmixing-based Multisensor Multiresolution Image Fusion. *IEEE Trans. Geosci. Remote Sens.* 37, 1212–1226. doi:10.1109/36.763276

Zurita-Milla, R., Clevers, J., and Schaepman, M. E. (2008). Unmixing-based Landsat Tm and Meris Fr Data Fusion. *IEEE Geosci. Remote Sens. Lett.* 5, 453–457. doi:10.1109/lgrs.2008.919685

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.

frontiers in Environmental Science

ORIGINAL RESEARCH
published: 13 October 2021
doi: 10.3389/fenvs.2021.757471

The Particulate Organic Carbon-to-Nitrogen Ratio Varies With Ocean Currents

Shengjun Xiao[1,2], *Linlin Zhang*[1,2], *Yuhao Teng*[1,2], *Tao Huang*[1,2] *and Wen Luo*[1,2]*

[1]*Key Laboratory of Virtual Geographic Environment, School of Geography, Ministry of Education, Nanjing Normal University, Nanjing, China,* [2]*Jiangsu Center for Collaborative Innovation in Geographical Information Resource Development and Application, Nanjing, China*

Ocean currents could adjust ocean carbon and nitrogen composition which are an important part of the global carbon and nitrogen cycle. We procured global concentrations of particulate carbon and nitrogen in different depths, classified them according to ocean currents (upper 300 m), and analyzed POC-to-PON ratio (particulate organic carbon-to-nitrogen ratio) variations. We found that the regions with currents have a higher ratio than those without currents in the northern hemisphere, except in 50°–60°N (median ratio without currents is 8.38). Warm currents (median ratio ranges from 5.96 to 8.44) have a higher ratio than cold currents (6.19–8.89), except for the East Greenland Current (reach to 8.44) and Labrador Current (reach to 8.89). Meanwhile, we also analyzed the effects of ocean currents' flowing and found that the distributions of the POC-to-PON ratio vary in different current types (e.g., cause of formation and distance from the shore). Generally speaking, the POC-to-PON ratio of the eolian currents and near-ocean currents change fiercer than that of compensation currents and near-coast currents. Ocean currents also have a buffering effect in the variation between surface and deep water, which prevents the severe change of the POC-to-PON ratio. The high-value anomaly of POC-to-PON caused by the confluence of warm and cold currents was also analyzed. It can be deduced that the high ratio in the high-latitude region was mainly caused by the terrigenous organic matter (especially carbon) and low nitrogen.

Keywords: POC, PON, ocean currents, spatial variation, ratio variations

OPEN ACCESS

Edited by:
Peng Liu,
Institute of Remote Sensing and Digital Earth (CAS), China

Reviewed by:
Jingwei Shen,
Southwest University, China
Jianfeng Li,
Chinese Academy of Geological Sciences, China

***Correspondence:**
Wen Luo
luowen@njnu.edu.cn

Specialty section:
This article was submitted to Environmental Informatics and Remote Sensing, a section of the journal Frontiers in Environmental Science

Received: *12 August 2021*
Accepted: *06 September 2021*
Published: *13 October 2021*

Citation:
Xiao S, Zhang L, Teng Y, Huang T and Luo W (2021) The Particulate Organic Carbon-to-Nitrogen Ratio Varies With Ocean Currents. Front. Environ. Sci. 9:757471. doi: 10.3389/fenvs.2021.757471

INTRODUCTION

As an important part of the global biogeochemistry cycle, the ocean carbon and nitrogen cycle has always been concerned. Alfred Redfield presented the deep ocean particulate matter ratio, which later we call "Redfield ratio" (C:N:P = 106:16:1, Redfield, 1934), and it became a central tenet for ocean biogeochemistry and nutrition research (Deutsch and Weber, 2012; Mills and Arrigo, 2010). Later, many scholars studied the carbon, nitrogen, and phosphorus biochemical or stoichiometric variations of phytoplankton in deep or surface water (Dickman et al., 2006). The elemental composition and ratio have always been crucial points in ocean or land water research.

The research about oceans have been conducted very early (Allen, 1921; Stommel, 1957), but most of it focused on the production, for example, Parsons thought that the ocean detritus should be considered a source of food and secondary producers (Parsons and Strickland, 1962). In early times, the stress of ocean research was the nutrition and production of plankton. Then numerous studies of

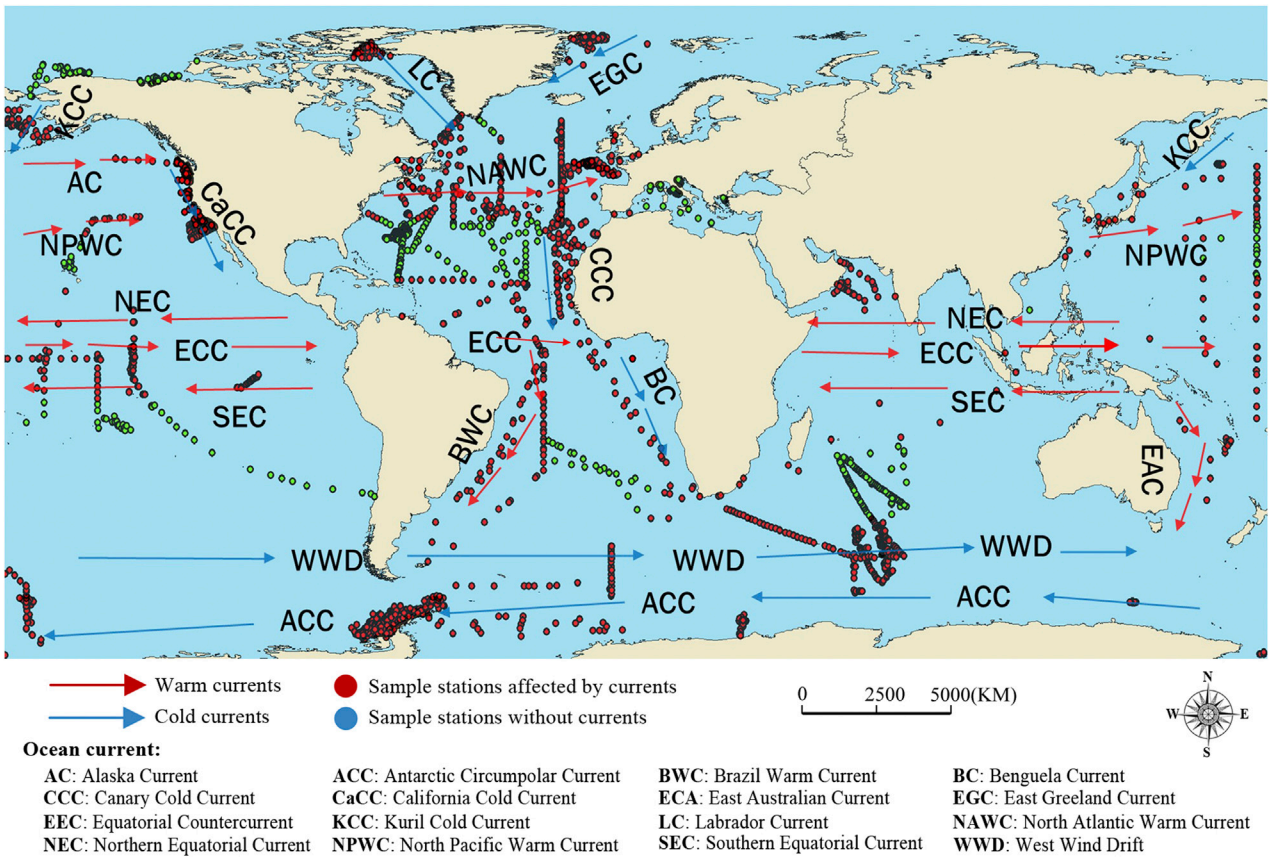

FIGURE 1 | Geographical locations of sample stations used in our study and ocean currents' sketch map.

biology, geography, and chemistry on oceans have been conducted until now. When the Redfield ratio became a hot spot in marine research, Tett connected the Redfield ratio with the phytoplankton growth rate (Tett et al., 1985), and Handa performed chemical studies on the organic matter and carbon cycle in ocean (Handa, 1991), and there had been research into the kinds or variations of the phytoplankton community (Karl et al., 2001; Ho et al., 2010). Singh researched C:N:P stoichiometry in the North Atlantic Ocean (Singh et al., 2015), and Seok Jin Oh conducted research work on spatial–temporal distribution of particulate organic carbon (POC) and nitrogen (PON) in the southwestern area of the East Sea (Seok Jin et al., 2016). Particulate organic matter (POM) reveals carbon or nitrogen sequestration and release in ocean compared with DOM (dissolved organic matter). They are significant processes in the global carbon or nitrogen cycle. Therefore, the distribution and variation of POM in oceans or rivers get more attention.

Recently, many studies have been conducted on variations in particulate carbon, nitrogen, and the phosphorus ratio of oceans, lakes, rivers, and estuaries. Furthermore, the scale from estuaries is regional to global (Kunz et al., 2011; Martiny et al., 2014). Megan Young researched the export of carbon and nitrogen in mangrove-dominated lagoons (Young et al., 2005), and Nicolas Savoye studied particulate organic matter in the Gironde Estuary in France (Savoye et al., 2012); regional research have been conducted in the Strait of Magellan (Fabiano et al., 1999), Southern Ocean at 30°E–80°E (Pasquer et al., 2010), and the northern Beibu Gulf (Kaiser et al., 2014). Globally, Martiny found strong latitudinal patterns in the C-to-P and N-to-P ratios of marine plankton (Martiny et al., 2013; Martiny et al., 2014). Mouw even combined the global ocean particulate organic carbon flux with satellite parameters (Mouw et al., 2016). The researches were not only carried out among the space domain but also the variation of the carbon and nitrogen ratio in time, for example, Rembauville monitored the seasonal dynamic POC export in the Southern Ocean (Rembauville et al., 2015), and Zhao studied the seasonal variation of C:N:P in the western Pacific Ocean (Zhao et al., 2016), and some conducted research both in space and time (Niemi and Michel, 2015). Besides the seasonal variation, there also has been long time scale research about ocean POC or PON (Heinze et al., 2016). There are many scholars trying to explain the distribution and variation of POM in marine environment (Mouw et al., 2016; Barrera et al., 2017). Therefore, variations in the element composition and ratio in space–time have been focused in marine biogeochemistry, and the time variation could be more important with the increasing impact of extreme climatic change and human activities on the environment.

However, few studies suggest the ocean current itself has effects on the variations in the carbon, nitrogen, and phosphorus ratio and spatiotemporal change comprehensively

and systematically. Most usually, researchers study the regional, global, or spatial–temporal variations in carbon or nitrogen in oceans. Even some conducted research on ocean currents' effect on carbon or nitrogen, where they only observed one current C or N variation (Ruiz-Cooley et al., 2014), which are regional research in fact. Ocean currents regulate the global carbon and nitrogen cycle through their most important effect—transmission. They influence the physical, chemical, biological, and geological processes of surface or even deep oceans. The clear majority of studies focusing on POM of phytoplankton in oceans are regional or global, few from the angle of the ocean currents, and some were studies on the element composition and ratio in the Pacific and Atlantic oceans. The salt content and temperature of oceans can be changed gradually corresponding to global warming; afterward, the element composition could be altered. Therefore, the currents' influence in POM of phytoplankton and their ratio have profound meaning to study the carbon and nitrogen cycle as well as its response to global warming.

DATASET AND METHODS

Data

Our data set was archived from Huang (Huang et al., 2018). There are altogether 63,827 sample points in oceans. In order to study ocean currents' influence on the POC-to-PON ratio, all data have been classified according to ocean currents (Orsi et al., 1995; Bischof et al., 2004a; Gyory et al., 2004a; Bonhoure et al., 2004; Rowe et al., 2004; Pidwirny, 2006). Meanwhile, the data which were below 300 m (where currents can hardly affect the POC-to-PON ratio, Monahan et al., 2001; Scott and Anya., 2001) were rejected, except the data of the California Cold Current, Labrador Current, and North Atlantic Warm Current which contained all data to study the deep area variation on the POC-to-PON ratio. The extremely low or high POC-to-PON ratio has been eliminated through all data, and there were still 43,404 examples left after rejection (**Figure 1**). After calculating the POC-to-PON median ratio, we only kept values ranging from 2 to 13 when plotting. We mainly used the median POC-to-PON ratio and included data of temperature, salt content, chlorophyll, pH, and C, N, and P index, which were not complete though.

Spatial Statistical Analysis of Ocean Currents

To identify the global surface ocean variation of the POC-to-PON ratio, we analyzed global distribution from the top 300 m of the water column with and without ocean currents. To reveal the pattern of POC and PON, the currents (except the Benguela Current, Alaskan Current, Canary Cold Current, West Wind Drift, East Greenland Current, Antarctic Circumpolar Current, and the Equatorial Countercurrent in the Atlantic) are divided into upstream and downstream from the middle to understand the variation in the POC-to-PON ratio with respect to the direction of flowing. The direction of the current is also abstract as longitude direction and latitude direction according to the main direction of ocean currents. Because this data set is sampling data, the actual site coordinates are not used. We just classified and summarized the data according to the longitudinal and latitudinal grid. To better study how the material transport and energy transport of ocean currents affect the POC-to-PON ratio, we divide the ocean current into meridional and latitudinal directions and analyze the variation of the POC-to-PON ratio along the meridional and latitudinal directions.

Because the amount of data is huge and the data are not evenly distributed in space and time, the histogram and quantile analysis methods are used to analyze the variation in the POC-to-PON ratio from a statistical point of view. A histogram is used to summarize the value distribution of the data sets. The given POC-to-PON ratio data can be divided into several categories according to the ocean current type or flow direction, and the histogram analysis is performed with every category. For the given category C, it is assumed that the number of observations is N; the histogram analysis is defined as follows:

$$\begin{cases} S(X_i) = \left\{ C:\ X_i - \dfrac{D}{2} < C \le X_i + \dfrac{D}{2} \right\} \\ F(X_i) = \dfrac{\mathrm{count}\left(S(X_i)\right)}{N} \end{cases} \tag{1}$$

where $S(X_i)$ is used to divide the original observation data into different subsets, which are also known as histogram bins; X_i denotes the central point of the *ith* bin, and D denotes the width of the bins. After the division step, $F(X_i)$ can be used to calculate the frequency of the *ith* bin. And operator count$(S(X_i))$ is used to count the observation number belonging to the *ith* bin.

Quantile analysis is a common method used for the quantitative analysis of data set distribution. For the given category C, its m quantile analysis is defined as follows:

$$\{C_m\}_{i=1}^{n} = \begin{cases} C_{\left([np]+1\right)} & , np\ is\ not\ an\ \mathrm{integer} \\ \dfrac{\left[C_{(np)} + C_{(np+1)}\right]}{2} & , np\ is\ an\ \mathrm{integer} \end{cases} \tag{2}$$

In the formula, [] is the rounding operation; there are np samples smaller than the quantile, and np may not be an integer. C_m represents the value of the *mth* position in the sample. Quantile analysis can generally be expressed by boxplot. Boxplot results reflect the upper and lower boundaries, upper quartile, lower quartile, and the median of the data set.

RESULT

We classified the datasets according to ocean currents and explored the distribution of values, as shown in **Figure 2**. As the southern hemisphere has less currents, it has the largest wind current—the West Wind Current. The median C-to-N ratio of the West Wind Current is 7.53 in the Atlantic, 7.76 in India, and 6.57 in the Pacific Ocean, all of them close to the R ratio, especially in the Pacific. Moreover, the Antarctic Cycle Current has a lower median ratio, 6.45. Consequently, in the Antarctic and open ocean zones, both have cold currents. The POC-to-PON ratio is very close to the R ratio. This is consistent with some conclusions,

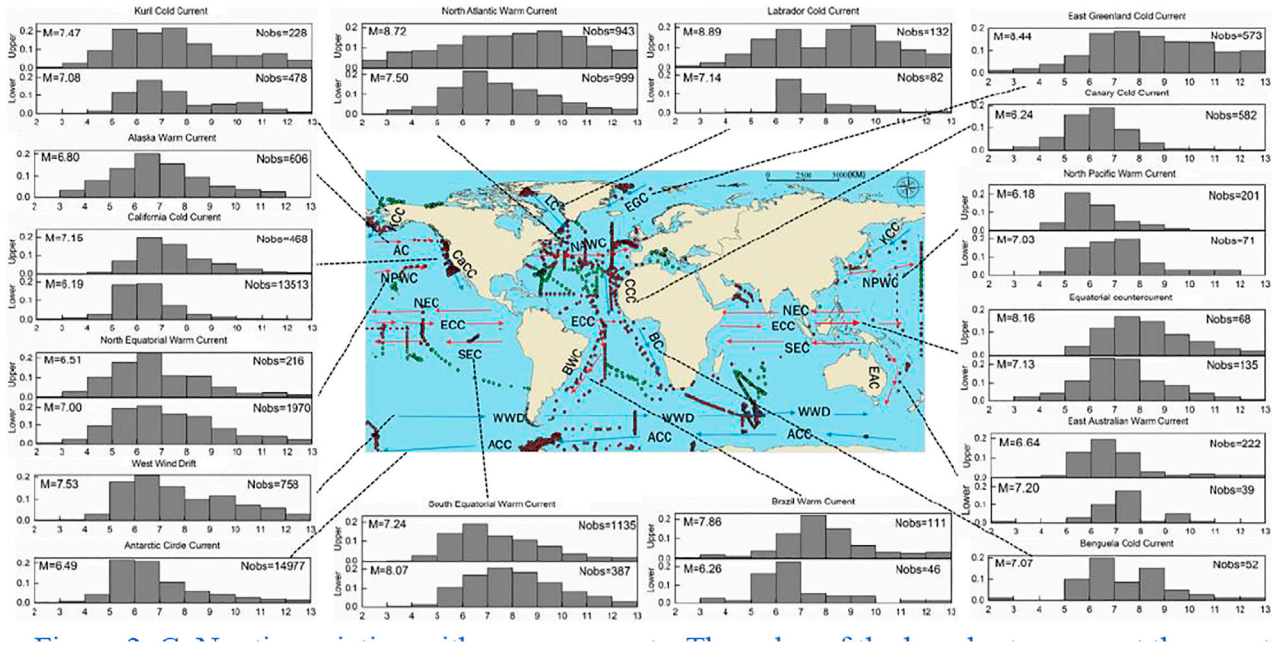

FIGURE 2 | C-to-N ratio variation with ocean currents. The value of the bar chart represents the count percentage of the histogram bins.

that is, open ocean, few or without human activities, has a POC-to-PON ratio close to 6.63.

DISCUSSION

Spatial Variation of the POC-to-PON Ratio With Ocean Currents

Global POC-to-PON Ratio Variation With and Without Ocean Currents

To identify the global surface ocean variation in the POC-to-PON ratio, we analyzed global distribution from the top 300 m of the water column with and without ocean currents. As a result, the median ratio in currents affects the regions' range from 6.14 to 8.63, which is close to the Redfield ratio in general, and it is higher in the northern hemisphere and high-latitude areas (**Figure 3A**). In contrast to the POC-to-PON ratio in regions with and without currents, we found that the ratio is higher with currents in the northern hemisphere, except 50°–60°N (median ratio without currents is 8.38), which probably is because of the few samples, but in the southern hemisphere, the ratio is little different. Therefore, in the northern hemisphere, ocean currents had an increased POC-to-PON ratio but little influence in the southern hemisphere. Actually, the distribution is complex in the northern hemisphere.

We globally analyzed the median POC-to-PON ratio of upper 300 m sea water and classified the data ocean current and the regions without currents. Some researchers have conducted studies about the latitudinal patterns of the C-to-N ratio in oceans (Martiny et al., 2013 Martiny et al., 2014), Although some global median C-to-N ratios differ from the canonical Redfield ratio, most of them are close to 6.6 in general. As other study, the C-to-N ratio is higher in low latitudes and warm ocean regions (Martiny et al., 2013), and the latitude dominates the global element ratio of oceans. However, whether the median C-to-N ratio of regions has ocean currents or not is different. The existence of ocean currents indeed affects the global C-to-N ratio because of the function—transport. The main factors that influence the POC-to-PON ratio of plankton are solar radiation, temperature, nutrients, and the kinds of plankton communities (Barrera et al., 2017). The currents adjust global C and N by delivering heat, nutrients, and so on. The region with currents has a higher median ratio than areas without currents (**Figure 3B**), although the data of regions without ocean currents are limited. Therefore, ocean currents could increase the POC-to-PON ratio. The ability of regulating the surface global ocean carbon and nitrogen ratio is related to the range of currents affecting the temperature of water, cause of information, flowing direction, and other characters of ocean currents.

The variation of the C-to-N ratio is remarkable in the northern hemisphere, which is because of complicated conditions, than the southern hemisphere—more significant currents as well as the terrene. The southern hemisphere has broader and more open oceans; there is little influence by currents which mainly contain the West Wind Drift and Antarctic Circle Current, as well as less human activities. Our results vary from others distinctly in the upper latitude of the northern hemisphere, and the median ratio in the upper latitude surface of oceans is above 6.63, and it even exceeds eight in 80°N–90°N, most possibly because of the terrigenous organic matter (especially carbon) and low nitrogen in the Arctic Circle. There still is another

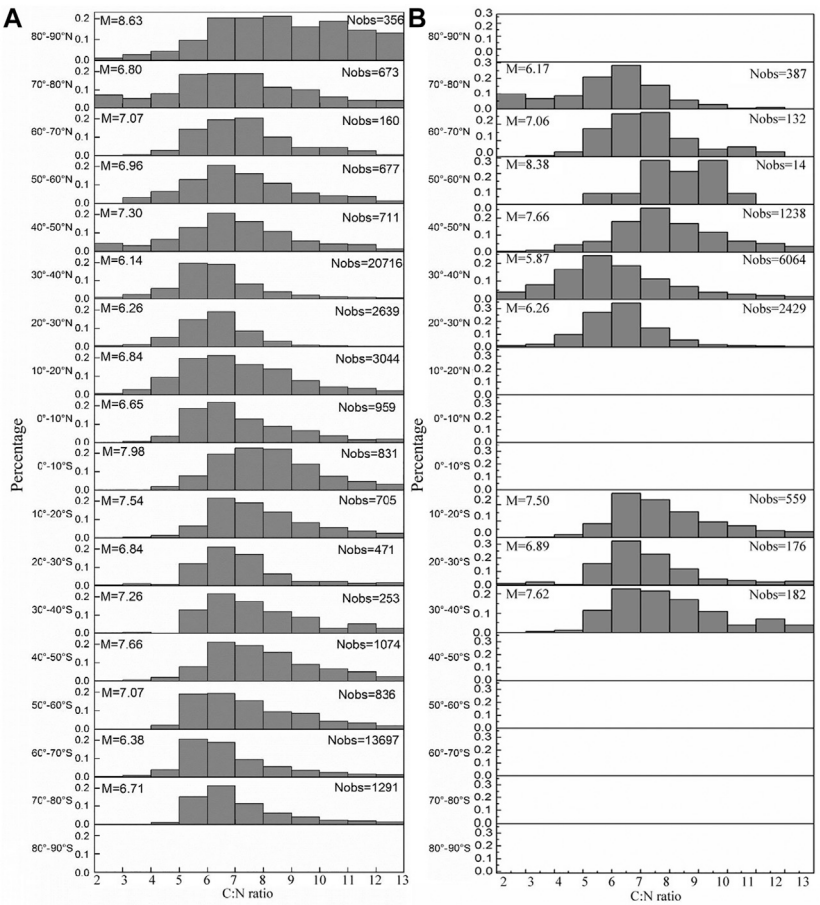

FIGURE 3 | Latitudinal patterns of the global C-to-N ratio (upper 300 m). **(A)** Only including the region affected by ocean currents. **(B)** the region without ocean currents. The blank indicates no data. M is the median ratio.

likelihood—global warming. Because of ice and snow melting in the Arctic Circle, the sea water has been desalinating and the deep ocean circulation probably has been affected to great extent. There have been several research studies about the variation of POM in the Arctic Circle (Emerson et al., 2001; Harada, 2016). The C-to-N ratio in north upper latitudes is likely to increase by currents converging in the future because of global warming, while how and to what extent are unclear.

POC-to-PON Ratio Variation With Ocean Currents

In fact, the ocean current affects the POC-to-PON ratio directly by transporting nutrition, heat, and plankton to change the composition of organic matter, and they also change the ratio indirectly by regulating the environmental conditions. The scale of the current affect could alter the organic matter element ratio as well (Sterner et al., 2008). In our study, ratios are usually higher large scale, and it will be closer to the R ratio in small scale.

Different temperatures of ocean currents have different influences on the POC-to-PON ratio of global surface oceans. In our study, the region with warm currents has higher median C-to-N ratio, except the East Greenland Current and Labrador Current. This is identical with most studies—higher temperature and higher carbon-to-nitrogen ratio (Martiny et al., 2013; Martiny et al., 2014). According to our data, warm currents have higher temperature and concentration of chlorophyll than cold currents, whereas the East Greenland Current has some samples which have high concentration of chlorophyll. Thus, the East Greenland Current has a much higher POC-to-PON ratio, and it is probably a rich terrigenous supply. Low-latitude districts receive more solar radiation. The temperature of surface sea water is higher, so plankton has a favorable environment for growing in higher temperature. Thus, warm currents have a higher POC-to-PON ratio than cold currents, but only in middle and low latitudes.

How about the warm currents at the same latitude? The North Atlantic Warm Current has a much higher ratio than the canonical Redfield value and that of the North Pacific Warm Current. Meanwhile, the North Equatorial Warm Current has a median ratio of 5.96–8.27, from lower to upper reaches, both deviate from the R ratio. The situation at the same latitude is complex and varying as the environment, including land–sea distribution, nitrogen, plankton, and water temperature. The sample points of the North Atlantic Warm Current are more from offshore to central ocean compared with the North Pacific

Warm Current; it could comprise several terrigenous particulate organic matter (POM). The North Atlantic is a bigger carbon sink, and there has been a high carbon value and terrestrial POM that could influence the carbon and nitrogen cycle, which has been proved (Head et al., 1996). Although at some latitude, the ocean current could change the latitude pattern of the POC-to-PON ratio, which is affected by the distance to the shore as well.

In our study, all the cold currents, except the Antarctic flow, are divided into upstream and downstream and have a decrease in the POC-to-PON ratio with flowing. It is likely that cold currents usually flow from high to low latitude; besides, the Brazil Warm Current and North Atlantic Warm Current have the same trade. However, the East Australian Current and North Pacific Warm Current are different; their ratio increases from upstream to downstream. They run from low to high latitude, and their regular patterns are like cold currents. As for the Brazilian Warm Current, there is the Falkland Cold Current joint which causes a rich nutritional environment and pulls down the ratio downstream.

Effects of Ocean Currents in POC-to-PON Ratio Variation

POC-to-PON Ratio Variation Along the Direction of Current Flow

Warm Currents

As shown in **Figure 3**, the range of POC-to-PON median ratios of the warm currents is 5.96–8.44, and most of the warm currents have much higher values than the Redfield ratio. For the meridional current, all the three currents flow from low latitudes to high latitudes (**Figure 4A**). Among the three currents, both the Japan Warm Current and the East Australia Warm Current are compensation currents, whose POC-to-PON ratio structures are relatively stable (falling first and then rising, and the Japan Warm Current rises abnormally high and far more than the R ratio, which may be caused by the confluence of warm and cold currents). The Brazil Warm Current is an eolian current, whose POC-to-PON ratio is evolving with drastic changes (especially in the front and end of the ocean current). This indicates that ocean currents of different causes have different mechanisms of action on the POC-to-PON ratio.

The latitudinal current is more common for warm currents (**Figure 4B**), and the median POC-to-PON ratio of most warm currents decreases with their flow direction. However, there exists a marked difference between the near-coast current and the near-ocean current. For the near-coast Alaska Current and the North Indian Ocean monsoon Circulation, the POC-to-PON ratio is relatively stable (fluctuates in the early stage and stabilizes in the later stage), and the near-ocean North Atlantic Current and the North Pacific Current are found to vary significantly (with an obvious downward trend and fluctuates significantly in the later stage). We can also find that the Mexico Gulf Current rises rapidly in the later stage because of the confluence of warm and cold currents.

Identical with longitudinal currents, almost all latitudinal currents have a decreased POC-to-PON ratio with their flow direction, except the North Atlantic Current and North Pacific Current. There are many rivers that pour into the Northern Atlantic; they carry massive POM from the land which changes the C and N proportion. Different rivers have different effects (Pasquer et al., 2010). Thus, the North Atlantic Current has an increased ratio with its flow direction, reiterating what we think determined by the composition and proportion of injected rivers. As for the North Pacific Current, the increased ratio is probably because of the missing water from the Equatorial Countercurrent which leans toward north and is influenced by the El Nino phenomenon (Chavez et al., 1999).

Cold Currents

As shown in **Figure 5**, the POC-to-PON ratios of the cold currents range from 6.19 to 8.89 (median), which are higher than the R ratio. The higher the latitude, the higher is the C-to-N ratio. Same as most warm currents, the ratio of cold currents decreases with flow direction. Most cold currents are longitudinal, compared with latitudinal currents which mainly transmit nutrition and organisms, and the longitudinal currents adjust heat of sea water in addition.

Compared with latitudinal currents, most cold currents have lower POC-to-PON ratios, except the cold currents in high latitudes in the northern hemisphere (the East Greenland Current and Labrador Current). The reason could be the terrigenous organic matter (especially carbon) or the low nitrogen in the Arctic Circle (Wheeler et al., 2002; Martiny et al., 2013). On the other hand, the confluence of warm and cold currents also made the POC-to-PON ratio unusually high.

POC-to-PON Ratio Variation With Ocean Currents Move Vertically

In addition to the horizontal motion of matter and energy, ocean currents will further strengthen the vertical motion. The element ratio changes with depth (Martiny et al., 2013; Auguè res and Loreau, 2015), and the ratio is low in deep water. The ratio of deep water under some currents (below 300 m) is also shown. Deep water of the Labrador Cold Current has a high C-to-N ratio reaching up to 8.17, whereas the North Atlantic Warm Current does not have too low ratio in deep areas (**Figure 6**). POM of surface ocean water could subside to as deep as the effect of gravity, so deep water could not decrease extremely if the surface water has a high ratio. However, in upwelling areas, such as the deep of the California Cold Current, the ratio of deep water is usually higher than the surface, and it is also above the R ratio. In upwelling zones, surface water is rich in nutrition, yet deep water is short of nutrition. Many studies have confirmed the ratio is low in oligotrophic areas. We can deduce that the C-to-N ratio in deep areas does not reduce drastically if surface water has a high ratio; it will likely continue to have a higher ratio. The ratio tends to be higher in deep water, if the C-to-N ratio of surface water is low. In other words, the currents advance the ratio in deep water, especially in low-ratio areas, particularly in upwelling areas, while this trend is overt in cold currents; the ratio decreases much in deep water of regions with warm currents.

Effects of the Confluence of Warm and Cold Currents

The confluence of warm and cold currents could mix sea water resulting in nutrients upwelling from deep water, phytoplankton booming, and fishing grounds being often found there. We chose

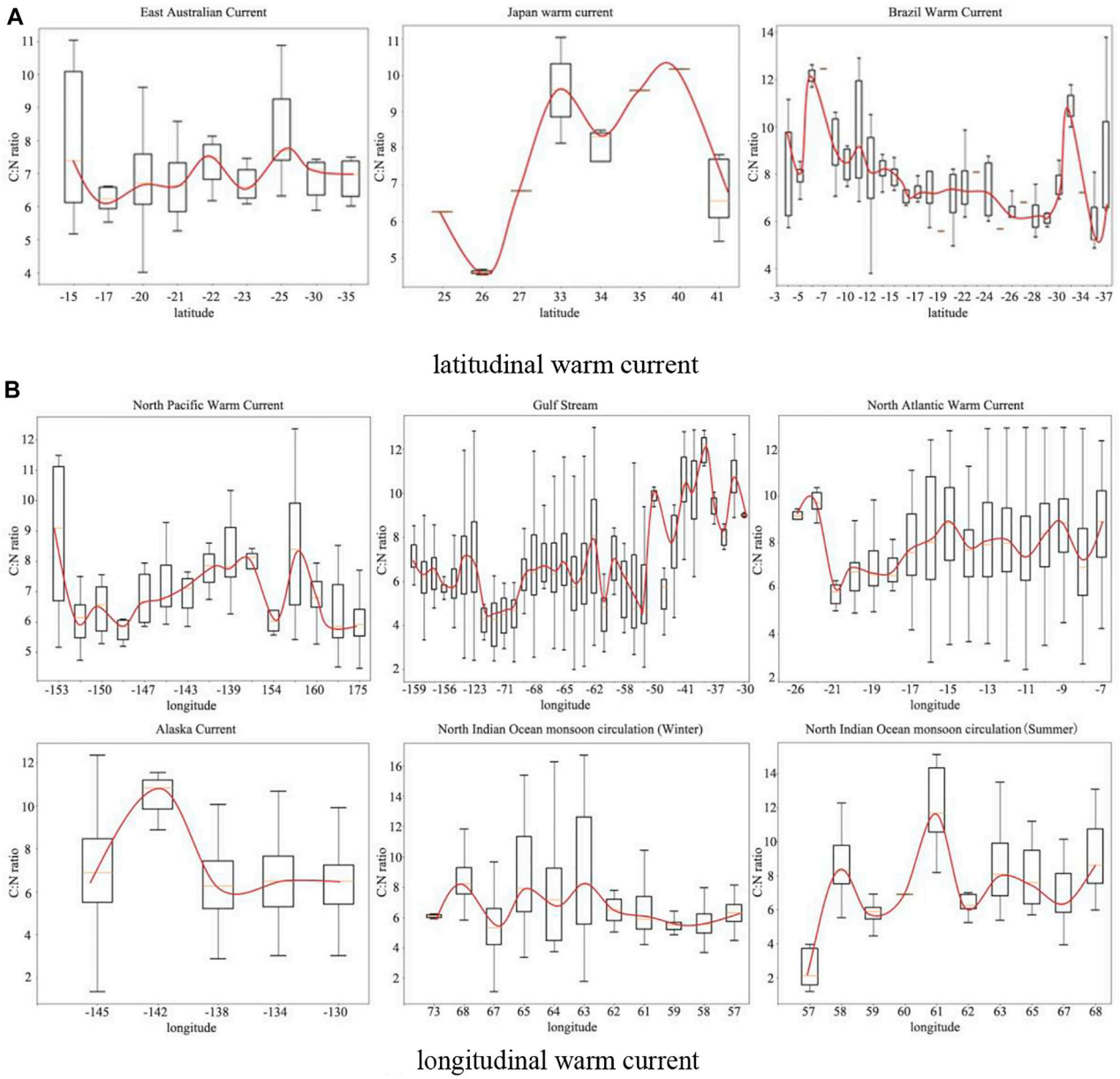

FIGURE 4 | POC-to-PON ratio variation along warm currents. **(A)** Latitudinal warm current. **(B)** Longitudinal warm current.

two typical cold and warm intersections to analyze its impact on the ratio. As shown in **Figure 7**, the C-to-N ratio in the confluence was much higher. However, we found that the ratio would increase at the intersection first and then decrease after mixing, for example, the Mexico Gulf Warm Current intersects with the Labrador Cold Current, and the median ratio changes from 5.80 to 9.80 and then to 7.60 (**Figure 7A**). The same trend appears between the North Atlantic Current and the East Greenland Current (**Figure 7B**). Consequently, the confluence of warm and cold currents could raise the POC-to-PON ratio.

We can also find that from warm to cold currents, the POC-to-PON ratio increases sharply, but from cold to warm currents, the ratio increases slowly. The reason may be that the POC-to-PON ratio of the cold current is much larger, which further proves that the main reason for the high ratio in the high-latitude region is the terrigenous organic matter (especially carbon) and the low nitrogen (Barrera et al., 2017).

CONCLUSION

We procured global particulate organic carbon and nitrogen of phytoplankton in oceans, analyzed the POC-to-PON ratio of the global variations of surface water and the median ratio based on ocean currents (with or without ocean currents). We found that

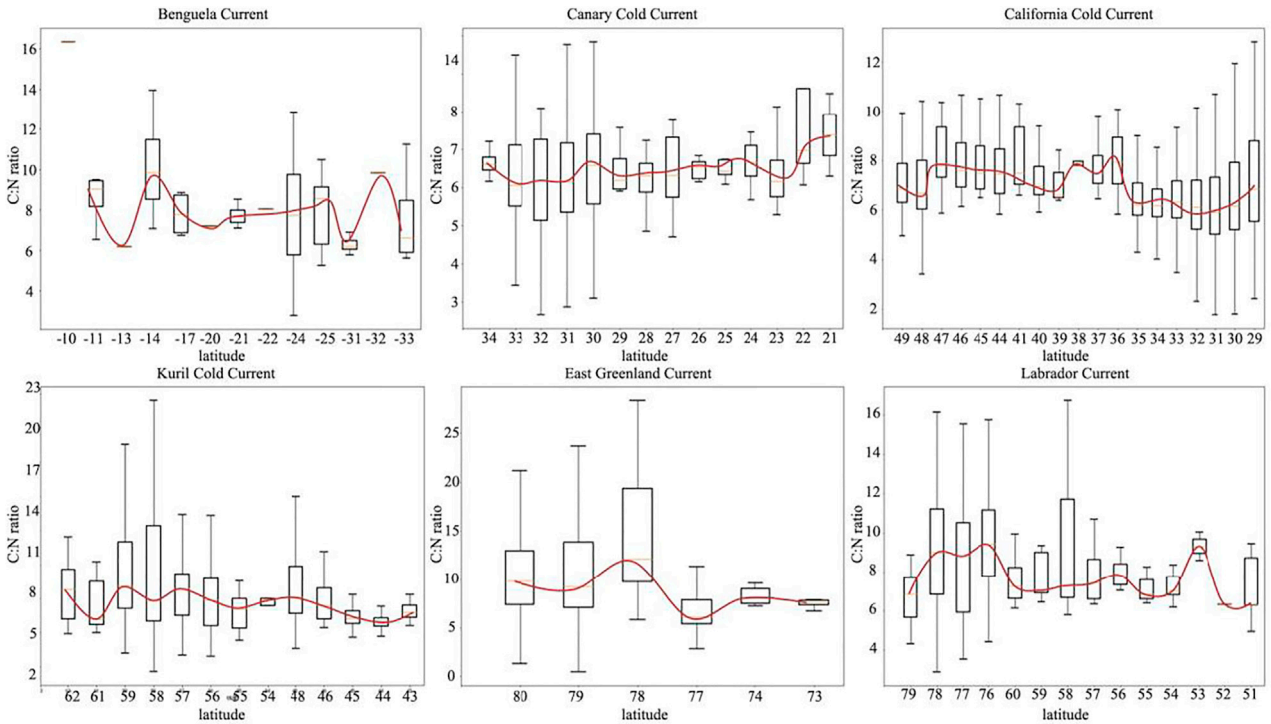

FIGURE 5 | POC-to-PON ratio variation along cold currents.

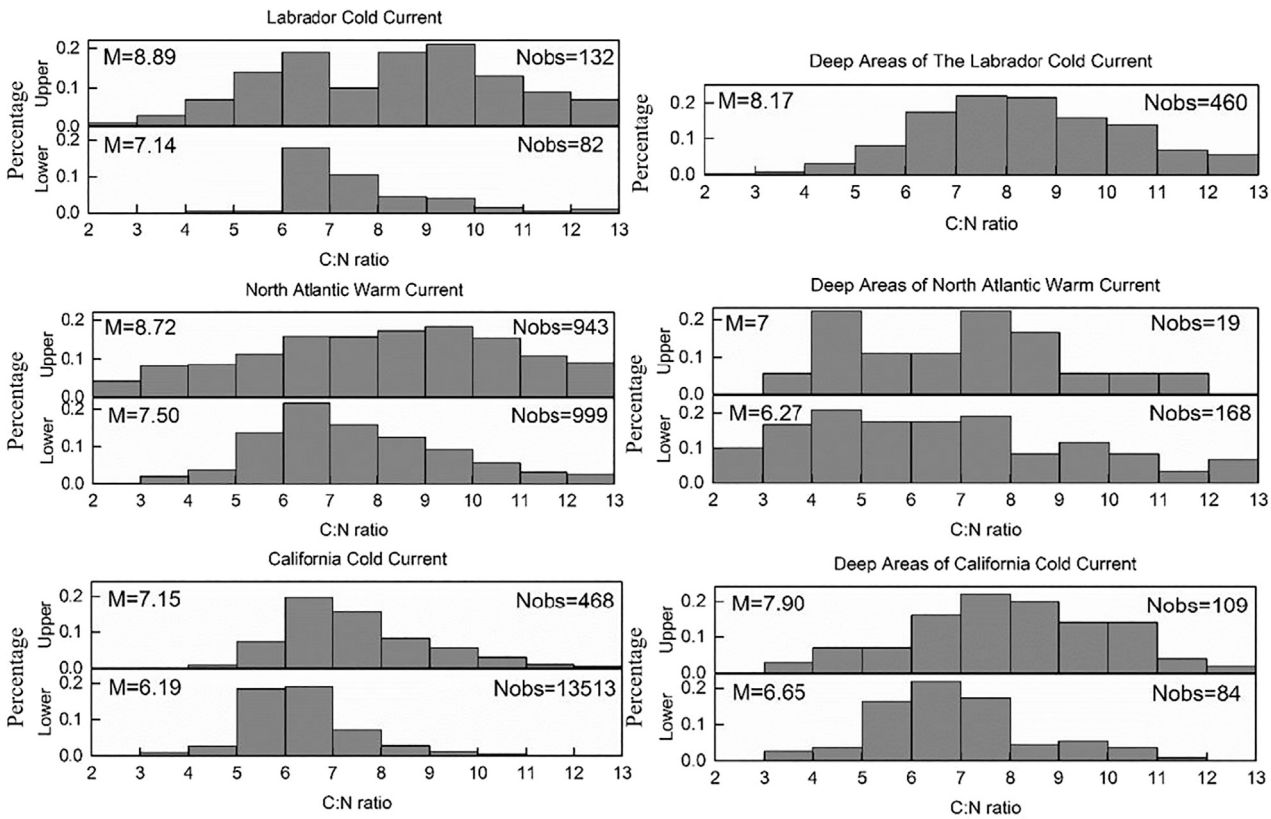

FIGURE 6 | Frequency of C-to-N ratio in deep water. M is the median ratio.

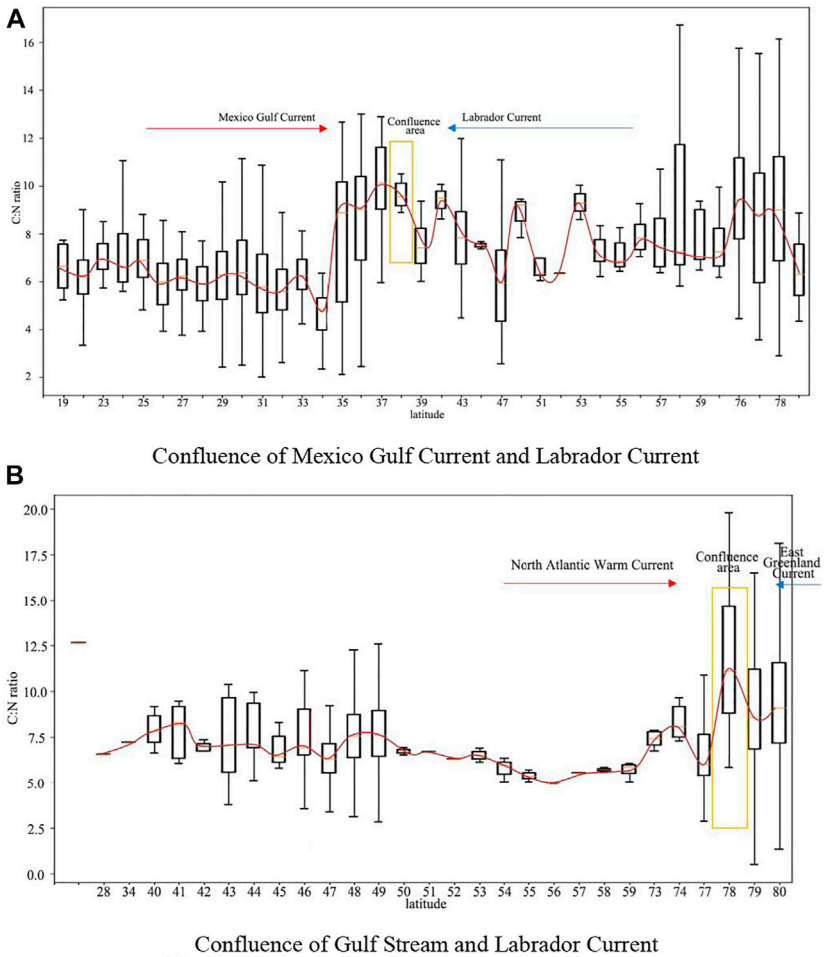

FIGURE 7 | Confluence of warm and cold currents. **(A)** Confluence of the Mexico Gulf Current and Labrador Current. **(B)** Confluence of the Gulf Stream and Labrador Current.

the global ocean median ratio, especially in the southern hemisphere, is close to the Redfield ratio, whereas the ratio is a little higher in the northern hemisphere, it even reaches 8.63 in 80°N–90°N. In our study, ocean currents indeed could affect the POC-to-PON ratio by adjusting heat, nutrition, and delivering plankton.

The higher ratio arises in warm upper latitudinal currents, whereas it is lower in cold lower latitudes. On the other hand, the distributions of the POC-to-PON ratio vary in different ocean current types (e.g., cause of ocean currents and distance from shore). Generally speaking, the POC-to-PON ratio of the eolian current and near-ocean current change fiercer than that of the compensation current and near-coast current. Meanwhile, ocean currents also have a buffering effect in the variation between surface and deep water, which prevents the severe change of the POC-to-PON ratio. The analysis of the high-value anomaly of the POC-to-PON ratio with the confluence of warm and cold currents showed that the changing process of warm currents and cold currents is different, which may promote our understanding of high ratios in high-latitude regions.

Our study shows that ocean currents will have more important, extensive, and profound impact on the ocean carbon and nitrogen cycle. With the process of global warming, the effect of currents should get more attention. And we should focus on the variation in the northern hemisphere, especially the Arctic because of its complex conditions and human activity effects.

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://figshare.com/s/a7cb3953273a6baa5e3e.

AUTHOR CONTRIBUTIONS

SX and WL contributed to the article's ideas and methods. LZ, TH, and YT classified the data and drew the graphs. SX, WL, and LZ wrote the article with considerable input from LZ. YT, and TH revised and checked the language of the article.

FUNDING

This research has been supported by the National Natural Science Foundation of China (grant nos. 41976186 and 41773097) and the Youth Top Talent funded by Nanjing Normal University.

REFERENCES

Allen, W. E. (1921). Some Work on Marine Phytoplankton in 1919. *Trans. Am. Microscopical Soc.* 40 (4), 177–181. doi:10.2307/3221971

Auguères, A.-S., and Loreau, M. (2015). Regulation of Redfield Ratios in the Deep Ocean. *Glob. Biogeochem. Cycles* 29 (2), 254–266. doi:10.1002/2014gb005066

Barrera, F., Lara, R. J., Krock, B., Garzón-Cardona, J. E., Fabro, E., and Koch, B. P. (2017). Factors Influencing the Characteristics and Distribution or Surface Organic Matter in the Pacific-Atlantic Connection. *J. Mar. Syst.* 175, 36–45. doi:10.1016/j.jmarsys.2017.07.004

Bischof, B., Mariano, A. J., and Ryan, E. H. (2004a). The North Equatorial Counter Current. Ocean Surface Currents. Available at: http://oceancurrents.rsmas.miami.edu/atlantic/north-equatorial-cc.html (Accessed 2021).

Bischof, B., Rowe, E., Mariano, A. J., and Ryan, E. H. (2004b). The Brazil Current. Ocean Surface Currents. Available at: http://oceancurrents.rsmas.miami.edu/atlantic/brazil.html (Accessed 2021).

Bonhoure, D., Rowe, E., Mariano, A. J., and Ryan, E. H. (2004). The South Equatorial Sys Current. Ocean Surface Currents. Available at: https://oceancurrents.rsmas.miami.edu/atlantic/south-equatorial.html (Accessed 2004).

Chavez, F. P., Strutton, P. G., Friederich, G. E., Feely, R. A., Feldman, G. C., Foleyand, D. G., et al. (1999). Biological and Chemical Response of the Equatorial Pacific Ocean to the 1997-98 El Niño. *Science* 286 (5447), 2126–2131. doi:10.1126/science.286.5447.2126

Dauby, P., Frankignoulle, M., Gobert, S., and Bouquegneau, J.-M. (1994). Distribution of POC, PON, and Particulate Al, Cd, Cr, Cu, Pb, Ti, Zn and $\delta^{13}C$ in the English Channel and Adjacent Areas. *Oceanologica Acta* 17 (6), 643–657.

Deutsch, C., and Weber, T. (2012). Nutrient Ratios as a Tracer and Driver of Ocean Biogeochemistry. *Ann. Rev. Mar. Sci.* 4 (4), 113–141. doi:10.1146/annurev-marine-120709-142821

Dickman, E. M., Vanni, M. J., and Horgan, M. J. (2006). Interactive Effects of Light and Nutrients on Phytoplankton Stoichiometry. *Oecologia* 149 (4), 676–689. doi:10.1007/s00442-006-0473-5

Emerson, S., Mecking, S., and Abell, J. (2001). The Biological Pump in the Subtropical North Pacific Ocean: Nutrient Sources, Redfield Ratios, and Recent Changes. *Glob. Biogeochem. Cycles* 15 (3), 535–554. doi:10.1029/2000gb001320

Fabiano, M., Povero, P., Danovaro, R., and Misic, C. (1999). Particulate Organic Matter Composition in a Semi-enclosed Periantarctic System: the Straits of Magellan. *Scientia Marina* 63 (4), 89–98. doi:10.3989/scimar.1999.63s189

Gyory, J., Mariano, A. J., and Ryan, E. H. (2004a). The Benguela Current. Ocean Surface Currents. Available at: http://oceancurrents.rsmas.miami.edu/atlantic/benguela.html (Accessed 2021).

Gyory, J., Mariano, A. J., and Ryan, E. H. (2004b). The Canary Current. Ocean Surface Currents. Available at: http://oceancurrents.rsmas.miami.edu/atlantic/canary.html (Accessed 2021).

Gyory, J., Mariano, A. J., and Ryan, E. H. (2004c). The Labrador Current. Ocean Surface Currents. Available at: http://oceancurrents.rsmas.miami.edu/atlantic/labrador.html (Accessed 2021).

Handa, N. (1991). Chemical Studies on Organic Matter and Carbon Cycle in the Ocean. *J. Oceanographical Soc. Jpn.* 47, 49–61. doi:10.1007/bf02301755

Harada, N. (2016). Review: Potential Catastrophic Reduction of Sea Ice in the Western Arctic Ocean: Its Impact on Biogeochemical Cycles and marine Ecosystems. *Glob. Planet. Change* 136, 1–17. doi:10.1016/j.gloplacha.2015.11.005

Head, E. J. H., Harrison, W. G., Irwin, B. I., Horne, E. P. W., and Li, W. K. W. (1996). Plankton Dynamics and Carbon Flux in an Area of Upwelling off the Coast of Morocco. *Deep Sea Res. Oceanographic Res. Pap.* 43 (11), 1713–1738. doi:10.1016/s0967-0637(96)00080-5

Heinze, C., Hoogakker, B., and Winguth, A. (2016). Ocean Carbon Cycling during the Past 130,000 Years – a Pilot Study on Inverse Paleoclimate Record Modelling. *Clim. Past Discuss.* 12 (10), 1–48. doi:10.5194/cp-2016-35

Ho, T., Quigg, A., Finkel, Z. V., Milligan, A. J., Wyman, K., Falkowski, P. G., et al. (2010). The Elemental Composition of Some marine Phytoplankton. *J. Phycology* 39 (6), 1145–1159. doi:10.1111/j.0022-3646.2003.03-090.x

Huang, C., Jiang, Q., Yao, L., Yang, H., Lin, C., Huang, T., et al. (2018). Variation Pattern of Particulate Organic Carbon and Nitrogen in Oceans and Inland Waters. *Biogeosciences* 15, 1827–1841. doi:10.5194/bg-15-1827-2018

Kaiser, D., Unger, D., and Qiu, G. (2014). Particulate Organic Matter Dynamics in Coastal Systems of the Northern Beibu Gulf. *Continental Shelf Res.* 82 (1), 99–118. doi:10.1016/j.csr.2014.04.006

Karl, D. M., Bidigare, R. R., and Letelier, R. M. (2001). Long-Term Changes in Plankton Community Structure and Productivity in the North Pacific Subtropical Gyre: The Domain Shift Hypothesis. *Deep Sea Res. Part Topical Stud. Oceanography* 48 (8–9), 1449–1470. doi:10.1016/s0967-0645(00)00149-1

Kunz, M. J., Anselmetti, F. S., WÜEst, A., Wehrli, B., Vollenweider, A., Thüring, S., et al. (2011). Sediment Accumulation and Carbon, Nitrogen, and Phosphorus Deposition in the Large Tropical Reservoir Lake Kariba (Zambia/Zimbabwe). *J. Geophys. Res.* 116 (G3), 1–13. doi:10.1029/2010jg001538

Martiny, A. C., Vrugt, J. A., and Lomas, M. W. (2014). Concentrations and Ratios of Particulate Organic Carbon, Nitrogen, and Phosphorus in the Global Ocean. *Sci. Data* 1 (4), 140048. doi:10.1038/sdata.2014.48

Martiny, A. C., Pham, C. T. A., Primeau, F. W., Vrugt, J. A., Moore, J. K., Levin, S. A., et al. (2013). Strong Latitudinal Patterns in the Elemental Ratios of marine Plankton and Organic Matter. *Nat. Geosci* 6 (4), 279–283. doi:10.1038/ngeo1757

Martiny, A. C., Vrugt, J. A., Primeau, F. W., and Lomas, M. W. (2013). Regional Variation in the Particulate Organic Carbon to Nitrogen Ratio in the Surface Ocean. *Glob. Biogeochem. Cycles* 27 (3), 723–731. doi:10.1002/gbc.20061

Monahan, E. C., Dam, H. G., and Hans, G. (2001). Bubbles: An Estimate of Their Role in the Global Oceanic Flux of Carbon. *J. Geophys. Res.* 106 (C5), 9377–9383. doi:10.1029/2000jc000295

Mouw, C. B., Barnett, A., Mckinley, G. A., Gloege, L., and Pilcher, D. (2016). Global Ocean Particulate Organic Carbon Flux Merged with Satellite Parameters. *Earth Syst. Sci. Data* 8 (2), 1–24. doi:10.5194/essd-8-531-2016

Mouw, C. B., Barnett, A., Mckinley, G. A., Gloege, L., and Pilcher, D. (2016). Phytoplankton Size Impact on Export Flux in the Global Ocean. *Glob. Biogeochem. Cycles* 30, 1542–1562. doi:10.1002/2015gb005355

Niemi, A., and Michel, C. (2015). Temporal and Spatial Variability in Sea-Ice Carbon: Nitrogen Ratios on Canadian. *Arctic shelves* 3, 000078. doi:10.12952/journal.elementa.000078

Oh, S. J., Jeong, S., and Kim, S.-Y. (2016). Spatio-Temporal Distribution of Particulate Organic Carbon (POC) and Nitrogen (PON) in the Southwestern Area of East Sea. *Jkosmee* 19 (2), 144–150. doi:10.7846/jkosmee.2016.19.2.144

Orsi, A. H., Whitworth, T., III, and Nowlin, W. D., Jr. (1995). On the Meridional Extent and Fronts of the Antarctic Circumpolar Current. *Deep Sea Res. Oceanographic Res. Pap.* 42 (5), 641–673. doi:10.1016/0967-0637(95)00021-w

Parsons, T. R., and Strickland, J. D. H. (1962). Oceanic Detritus. *Science* 136 (3513), 313–314. doi:10.1126/science.136.3513.313

Pasquer, B., Mongin, M., and Johnston, N. (2010). Distribution of Particulate Organic Matter (POM) in the Southern Ocean during BROKE-West (30°E - 80°E). *Deep-Sea Res. Part* 57 (9), 779–793. doi:10.1016/j.dsr2.2008.12.040

Pidwirny, M. (2006). *Surface and Subsurface Ocean Currents*. 2nd Edition. Fundamentals of Physical Geography. Available at: http://www.physicalgeography.net/fundamentals/8q.html (Accessed 2021).

Redfield, A. C. (1934). "On the Proportions of Organic Derivatives in Sea Water and Their Relation to the Composition of Plankton," in *James Johnstone Memorial Volume* (Liverpool, U.K: Liverpool Univ. Press).

Rembauville, M., Salter, I., Leblond, N., Gueneugues, A., and Blain, S. (2015). Export Fluxes in a Naturally Iron-Fertilized Area of the Southern Ocean - Part 1: Seasonal Dynamics of Particulate Organic Carbon export from a Moored Sediment Trap. *Biogeosciences* 12 (11), 3153–3170. doi:10.5194/bg-12-3153-2015

Rowe, E., Mariano, A. J., and Ryan, E. H. (2004). The North Atlantic Current. Ocean Surface Currents. Available at: http://oceancurrents.rsmas.miami.edu/atlantic/north-atlantic.html (Accessed 2021).

Ruiz-Cooley, R. I., Koch, P. L., Fiedler, P. C., and McCarthy, M. D. (2014). Carbon and Nitrogen Isotopes from Top Predator Amino Acids Reveal Rapidly Shifting Ocean Biochemistry in the Outer California Current. *Plos One* 9 (10), e110355. doi:10.1371/journal.pone.0110355

Savoye, N., David, V., Morisseau, F., Etcheber, H., Abril, G., Billy, I., et al. (2012). Origin and Composition of Particulate Organic Matter in a Macrotidal Turbid Estuary: The Gironde Estuary, France. *Estuarine, Coastal Shelf Sci.* 108 (4), 16–28. doi:10.1016/j.ecss.2011.12.005

Scott, D., and Anya, W. (2001). Is Southern Ocean Organic Carbon and Biogenic Silica export Enhanced by Iron-Stimulated Increases in Biological Production? Sediment Trap Results from SOIREE. *Deep Sea Res. Part Topical Stud. Oceanography* 48 (11-12), 2681–2701. doi:10.1016/S0967-0645(01)00014-5

Singh, A., Baer, S. E., Riebesell, U., Martiny, A. C., and Lomas, M. W. (2015). C: N: P Stoichiometry at the Bermuda Atlantic Time-Series Study Station in the North Atlantic Ocean. *Biogeosciences* 12 (12), 9275–9305. doi:10.5194/bg-12-6389-2015

Sterner, R. W., Andersen, T., Elser, J. J., Hessen, D. O., Hood, J. M., McCauley, E., et al. (2008). Scale-Dependent Carbon:nitrogen:phosphorus Seston Stoichiometry in marine and Freshwaters. *Limnol. Oceanogr.* 53 (3), 1169–1180. doi:10.4319/lo.2008.53.3.1169

Stommel, H. (1957). A Survey of Ocean Current Theory. *Deep Sea Res. (1953)* 4 (3), 149–184. doi:10.1016/0146-6313(56)90048-x

Tett, P., Droop, M. R., and Heaney, S. I. (1985). The Redfield Ratio and Phytoplankton Growth Rate. *J. Mar. Biol. Ass.* 65 (2), 487–504. doi:10.1017/s0025315400050566

Young, M., Gonneea, M. E., Herrera-Silveira, J., and Paytan, A. (2005). Export of Dissolved and Particulate Carbon and Nitrogen from a Mangrove-Dominated Lagoon, Yucatan Peninsula, Mexico. *Int. J. Ecol. Environ. Sci.* 31 (3), 189–202. Available at: https://cpb-us-e1.wpmucdn.com/sites.ucsc.edu/dist/5/1050/files/2020/08/IJEES31-189.pdf

Zhao, J., Li, Y., and Wang, F. (2016). Seasonal Variation of the Surface North Equatorial Countercurrent (NECC) in the Western Pacific Ocean. *Chin. J. Ocean. Limnol.* 34 (6), 1332–1346. doi:10.1007/s00343-016-5119-9

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.

frontiers in Environmental Science

ORIGINAL RESEARCH
published: 27 October 2021
doi: 10.3389/fenvs.2021.748913

A Genetic Algorithm–Assisted Deep Neural Network Model for Merging Microwave and Infrared Daily Sea Surface Temperature Products

Changjiang Xiao[1,2], *Chuli Hu*[3,4]*, *Nengcheng Chen*[3,4], *Xiang Zhang*[4], *Zeqiang Chen*[4] and *Xiaohua Tong*[1]

[1]College of Surveying and Geo-Informatics, Tongji University, Shanghai, China, [2]Frontiers Science Center for Intelligent Autonomous Systems, Shanghai, China, [3]School of Geography and Information Engineering, China University of Geosciences (Wuhan), Wuhan, China, [4]National Engineering Research Center of Geographic Information System, China University of Geosciences (Wuhan), Wuhan, China

OPEN ACCESS

Edited by:
Peng Liu,
Institute of Remote Sensing and Digital Earth (CAS), China

Reviewed by:
Lei Guan,
Ocean University of China, China
Zenghong Liu,
Ministry of Natural Resources, China
Manoj Singh,
University of Petroleum and Energy Studies, India

***Correspondence:**
Chuli Hu
huchl@cug.edu.cn

Specialty section:
This article was submitted to Environmental Informatics and Remote Sensing, a section of the journal Frontiers in Environmental Science

Received: *28 July 2021*
Accepted: *09 September 2021*
Published: *27 October 2021*

Citation:
Xiao C, Hu C, Chen N, Zhang X, Chen Z and Tong X (2021) A Genetic Algorithm–Assisted Deep Neural Network Model for Merging Microwave and Infrared Daily Sea Surface Temperature Products. Front. Environ. Sci. 9:748913. doi: 10.3389/fenvs.2021.748913

Sea surface temperature (SST) is an important factor in the global ocean–atmosphere system, being vital in a variety of climate analyses and air–sea interaction research studies. However, estimating daily SST with both high precision and high spatial completeness remains a challenge. This article attempts to solve this problem by merging two complementary daily SST products, that is, the 25 km-resolution Advanced Microwave Scanning Radiometer for EOS (AMSR-E) SST and 4 km-resolution Moderate Resolution Imaging Spectroradiometer (MODIS) SST, using a genetic algorithm–assisted deep neural network model (GA-DNNM). The merged SST with a spatial resolution of 4 km and a temporal resolution of 1 day is achieved. Experiments in the Asia and Indo-Pacific Ocean (AIPO) region in 2005 were conducted to demonstrate the feasibility and advantages of the proposed method. Results showed that the spatial coverages of the original MODIS SST and AMSR-E SST are ranging from 25.0 to 48.1%, and 31.5 to 47.6%, respectively, while the merged SST achieves a spatial coverage ranging from 56.1 to 73.1%, with improvements ranging from 50.2 to 131.7% relative to the original MODIS SST. Comparisons with drifting buoy observations indicate that the merged SST is accurate, with an average bias of 0.006°C and an average RMSE of 0.502°C, in places where the MODIS SST data are missing before being merged in the AIPO area, and with an average bias of –0.082 °C, and an average RMSE of 0.603°C for the merged SST in the whole study area.

Keywords: sea surface temperature (SST), AMSR-E SST, MODIS SST, data fusion, genetic algorithm, deep neural network model

INTRODUCTION

Sea surface temperature (SST) is an important physical parameter of the oceans, playing a fundamentally important role in the exchange of energy, momentum, and moisture between the oceans and atmosphere (Wentz et al., 2000). The SST's changes may alter marine ecosystems, affect global climate significantly, influence the development and evolution of tropical storms and hurricanes, and potentially contribute to droughts and floods in some areas (Wentz et al., 2000;

USEPA, 2021). SST with high spatiotemporal resolution, spatial coverage, and accuracy is of vital importance to forecasting weather and monitoring climate change (Reynolds & Smith, 1995; Reynolds et al., 2002; Guan & Kawamura, 2004; Guo, 2010; Li et al., 2013; Tang et al., 2015; Zhu et al., 2018; Xiao et al., 2019).

Satellite observations, including infrared (IR) and microwave (MW), are the major sources based on which the global SST products are derived. IR SST is the earliest satellite derived one that emerged in the 1970s (Wentz et al., 2000). The IR SST usually has high spatial resolutions but is vulnerable to cloud contaminations (covering about half of the Earth), fog, and aerosols, leading to sparse spatial coverage and large amounts of missing data (Tang et al., 2015). In contrast, microwaves can penetrate clouds with little attenuation, and thus, MW SST can provide a fairly high spatial coverage of the sea under all weather conditions, except for rain (Wentz et al., 2000). However, MW SST has lower resolution than IR SST. Besides, its accuracy near coastlines is low, and it may not even be retrieved near lands (Li et al., 2013). It can be concluded that both IR SST and MW SST have advantages and disadvantages, which means they can only derive SST under certain circumstances alone. However, they are complementary to each other. Therefore, we can utilize these two types of SST complementarily to obtain SST with desirable qualities based on the idea of synergy (Zhang and Chen, 2016).

There have already been research studies on conflating MW SST and IR SST (Chao et al., 2009; Donlon et al., 2012; Guan & Kawamura, 2004; Guo, 2010; Li et al., 2013; Tang et al., 2015; Wang & Xie, 2007; Zhu et al., 2018). The mostly used methods are objective analysis (OA), optimum interpolation (OI), data assimilation, and Bayesian methods. OA, based on the Gauss–Markov theorem, was first introduced into oceanographic applications by Bretherton et al. (1976). However, statistical information about the field to be interpolated should be known or the field should be smooth (McIntosh, 1990). The OI method can increase the spatial completeness. However, it smoothens the fine spatial characteristics, which limits applications near the coastal area (Li et al., 2013; Tang et al., 2015). Besides, prior knowledge of the statistics of errors of input data is also required, which however is hard to know (Bretherton et al., 1976; Tang et al., 2015). There are primarily two data assimilation methods applied to merging SSTs, including the VARiational (VAR) approach and Kalman filter (KF). Using the same mathematical principle with OI, the VAR approach has a disadvantage that the variances of the background error and the covariances of the observational error are usually subjectively specified due to the difficulties in ascertaining them (Li et al., 2013; Tang et al., 2015). The KF needs to transform scales before merging, which may introduce extra uncertainties (Zhu et al., 2018). The Bayesian hierarchal model (BHM) and Bayesian maximum entropy (BME) are two typical Bayesian methods for merging multiple SSTs. The BHM-based methods assume that the value of pre-fusion data satisfies a special distribution. They use the prior knowledge as parameters and conclude the posterior average value to be the fused value (Guo, 2010). Therefore, prior knowledge is still a necessity, and bad or insufficient prior knowledge may lead to inaccurate fusion results. BME has been successfully applied to merging IR SST and MW SST of different spatial resolutions to produce high-resolution and high-accuracy SST (Li et al., 2013; Tang et al., 2015). The BME method can resolve the scale transformation problem of KF, but prior knowledge is still needed.

Unlike the previous methods, the deep neural network model represents a nonlinear computational method for learning knowledge from data and predicting complex trends, no matter what distributions the errors are subjected to, or how complex the relationships hidden in the data are (Yue et al., 2017; Zare Abyaneh et al., 2016). It has been successfully applied to numerous areas such as speech recognition (Dahl et al., 2012), human face recognition (Le, 2011), crop yield prediction (Kaul et al., 2005; Panda et al., 2010), crop type classification (Cai et al., 2018), weather forecasting (Valverde Ramírez et al., 2005), environmental monitoring (Li et al., 2017), and image fusion (Wu et al., 2018). However, neural networks tend to get trapped in local extreme values during training. Therefore, some researchers have tried to solve this problem by combing the neural network approach with optimization methods such as genetic algorithms (GA), and have achieved better performance and improved results consequently (Mahmoudabadi et al., 2009; Tahmasebi & Hezarkhani, 2012).

Therefore, considering the complex patterns and uncertainties in the satellite data, the fact that current methods usually require prior knowledge about the error statistics of input data which however is sometimes hard to ascertain, and the advantages of genetic algorithm-assisted deep neural network model (GA-DNNM) in learning patterns of data and dealing with uncertainties, no matter how complex the patterns are and how the data are distributed, we adapt the GA-DNNM to model the relationship between IR SST and MW SST data, and merge these two data to produce high-quality SST products which can further benefit climate analyses and air–sea interaction studies. Therefore, this research aims to 1) develop a GA-DNNM model to capture complex relationships between IR SST and MW SST, and evaluate the accuracy of such relationships over different time frames; 2) exploit the relationships to produce merged SST using IR SST and MW SST; and 3) evaluate the quality of the merged SST with drifting buoy observations (ground truth).

The study area is targeted at the area joined by the Asia and Indo-Pacific Ocean (AIPO) (Chang-Xiang et al., 2010). The major contributions of this article include 1) a novel GA-DNNM method specifically developed and demonstrated to be feasible and accurate for the task of merging IR SST and MW SST, and 2) the merged SST whose spatial resolution is 4 km, temporal resolution is 1 day, and spatial coverage is much improved.

The reminder of the article is structured as follows. In Section 2, the study area and data are introduced. **Section 3** describes the method, including data preprocessing, deep neural network model design, and genetic algorithm–based deep neural network model parameter optimization. In **Section 4**, the experimental results are given, and the accuracy of the GA-DNNM and quality of the merged SST are comprehensively evaluated. Finally, the conclusions are given in **Section 5** with potential future work.

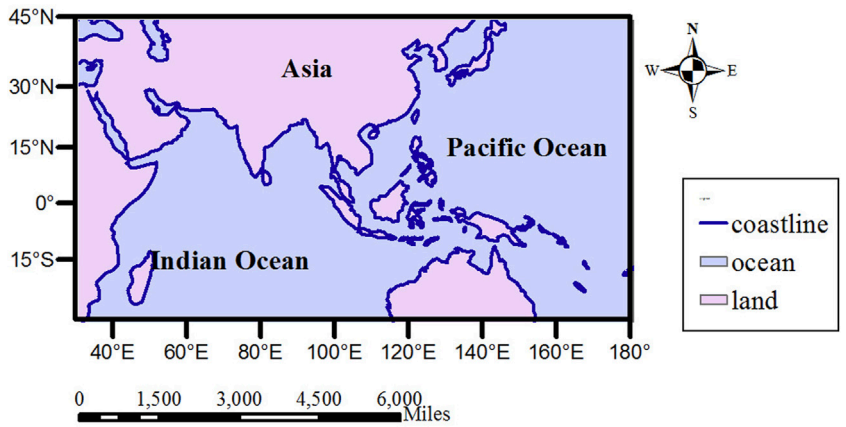

FIGURE 1 | Study area.

TABLE 1 | Specification of data used in this research.

SST data	Spatial range	Time range	Spatial resolution (km)	Temporal resolution (hour)	Depth
MODIS SST	30°S ~ 45°N, 30°E ~ 180°E	2005	4	12	~ um
AMSR-E SST			25	12	~1 mm
Drifting buoy SST			—	6	0.2–0.3 m

STUDY AREA AND DATA

Study Area

The study area AIPO lies between 30°S and 45°N, 30°E, and 180°E, as shown in **Figure 1**. The ocean–atmosphere interaction over AIPO has significant impacts on the short-term climate variations and predictions in China and surrounding areas (Wu et al., 2006; Li et al., 2013). Therefore, it is of vital importance to provide SST with high accuracy, high spatial completeness, and high spatiotemporal resolution in this region.

Data

This research uses two kinds of satellite-derived SSTs, that is, moderate-resolution imaging spectroradiometer (MODIS) SST, that is, IR SST, and advanced microwave scanning radiometer for EOS (AMSR-E) SST, that is, MW SST for merging, and drifting buoy observations as the ground truth for validation purpose, as illustrated in **Table 1** and detailed in the following subsections.

MODIS SST

MODIS SST that is used in this research is the MODIS Aqua Global Level 3 Mapped Thermal SST products derived from the 11 and 12 μm thermal infrared bands, produced and distributed by the Ocean Biology Processing Group (OBPG) at the NASA GSFC (OBPG, 2015; Werdell et al., 2013). Daily, weekly, monthly, and annual MODIS products can be obtained at the spatial resolutions of both 4.63 and 9.26 km, and for both day and night passes. To avoid diurnal warming caused by solar heating of the ocean surface, and to provide high spatiotemporal resolution SST, the daily 4 km-resolution nighttime SST products at 1:30 am local time are chosen. The version of this dataset is v2014.0 released on August 31, 2015. The time span of this dataset is from July 4, 2002 to present, and in this study, the daily MODIS SST data in 2005 are chosen, with 363 images in total (the MODIS SST on November 17, 2005 and that on November 20, 2005 are excluded).

The SST data used are in the format of netCDF with two layers, including a temperature data layer and a data quality control layer. The data quality control layer has the same number of pixels as the temperature data layer, recording a quality label of the corresponding SST on the temperature data layer. The quality flags are as following: 0 represents good, 1 represents questionable, 2 represents clouds, and 255 represents gross clouds, land, and other errors. In this research, the MODIS SST pixels with the quality flag equaling 0 are used for the modeling process. The temperature data represent the temperature at the depth of a few micrometers, with a valid retrieval range of −2°C–32°C (Armstrong, 2007).

AMSR-E SST

AMSR-E SST is derived from the remote sensing data of AMSR-E on NASA's EOS Aqua spacecraft, produced by Remote Sensing Systems (RSS), and sponsored by the NASA AMSR-E Science Team, and the NASA Earth Science MEaSUREs Program (Wentz et al., 2014). The data version is v7 released in October 2011. The daily SST products provided by RSS are orbital data that are mapped to 0.25°C grid, and divided into two maps based on ascending (1:30 pm) and descending (1:30 am) passes. To be

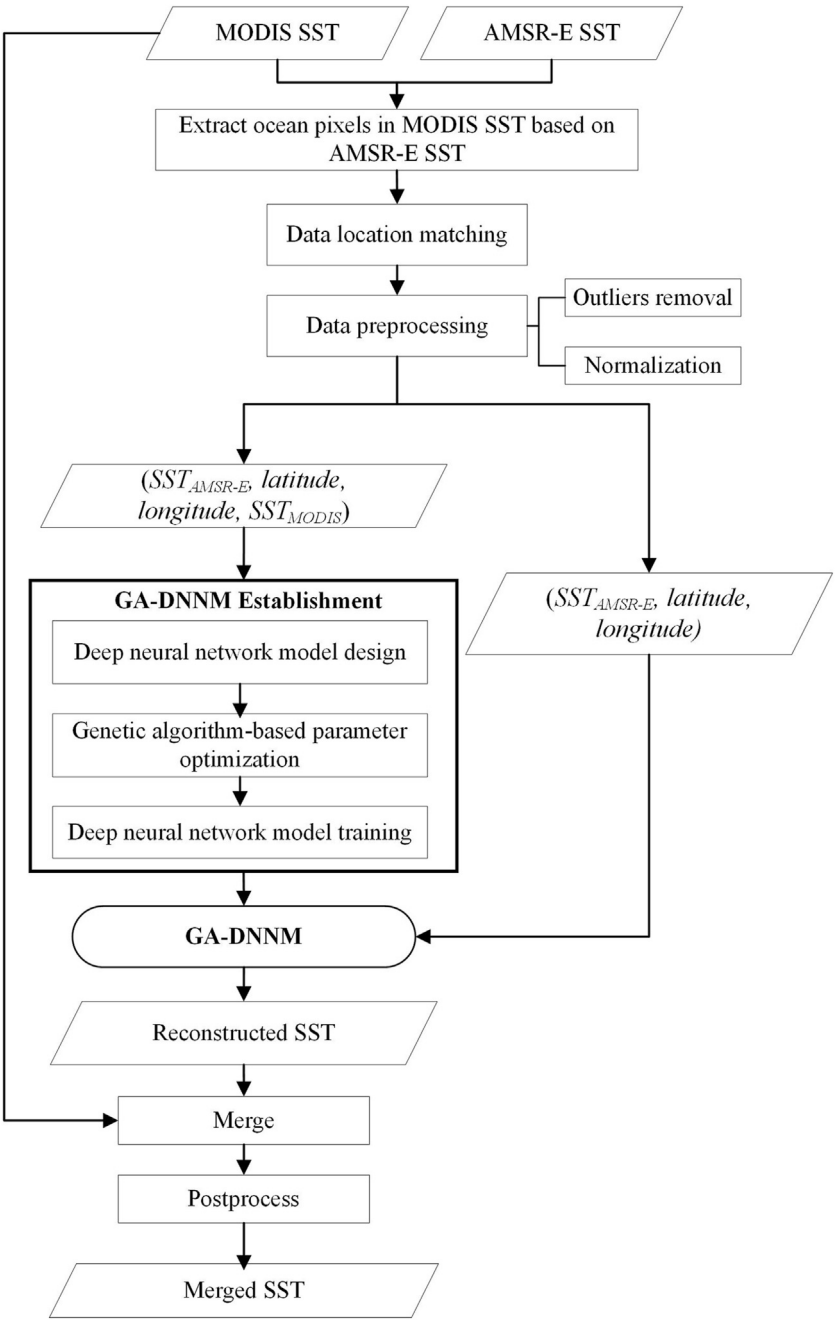

FIGURE 2 | Workflow for merging MW and IR SSTs based on GA-DNNM.

consistent with the MODIS SST data in time, the data measured at 1:30 am are chosen.

The AMSR-E SST data are the temperature of the top layer of water, which is about 1 mm thick. The original data values are in the range of 0–255, with 0–250 indicating valid geophysical data, 251 indicating missing SST, 252 indicating sea ice, 253 indicating bad observations which are not used in composite maps, 254 indicating no observations, and 255 indicating land mass. The original data values have to be scaled to get meaningful SST, which is achieved by multiplying the scale factor (0.15) and adding the offset (−3.0) (RSS, 2021). Therefore, the valid value for AMSR-E SST is -3°C–34.5°C.

Drifting Buoy SST

Drifters are expendable satellite-tracked systems which drift in response to ocean currents and winds. Currently, there are more than 1,000 drifters circulating in the world ocean, measuring SST and other properties (e.g., atmospheric pressure, sea salinity,

wind speed, and wind direction) at unprecedented resolution as ocean currents carry them along. The drifting buoy observations are usually used to correct the satellite measurement of ocean environmental parameters.

The drifting buoy SST data that are used in this research are collected, processed, and quality-controlled by the Atlantic Oceanographic and Meteorological Laboratory (AOML) (Lumpkin & Centurioni, 2019). The measurements are obtained at a depth of 0.2–0.3 m. The raw observations are interpolated into quarter-day intervals at 00:00, 06:00, 12:00, and 18:00 UTC using an optimal interpolation procedure. For the purpose of minimizing the possible diurnal warming of *in situ* SST and avoiding cold bias of the AMSR-E SST and MODIS SST, the minimum value of the observations of a day is chosen as daily ground truth for validation of the merged SST (Li et al., 2013; Tang et al., 2015). Before using the drifting buoy data, we remove the gross errors which are beyond the range of −1.8 ℃–35 ℃ (Høyer et al., 2012). The chosen drifting buoy SST at the same day is then mapped to 4 km × 4 km grids for the MODIS SST and merged SST, and 25 km × 25 km grids for AMSR-E SST by averaging the drifting buoy data belonging to the same grid as the corresponding satellite SST.

METHODS

The workflow of the method is depicted in **Figure 2**. First, ocean pixels are extracted in MODIS SST, based on AMSR-E SST. Then AMSR-E SST and MODIS SST pixels are matched based on locations, and quad-tuples ($SST_{AMSR\text{-}E}$, *latitude*, *longitude*, and SST_{MODIS}) are obtained. After preprocessing, including outlier removal and normalization, the quad-tuples are used by the genetic algorithm to obtain optimal initial parameters for the neural network model. Then the optimized neural network is trained with the quad-tuples to establish a mapping function between ($SST_{AMSR\text{-}E}$, *latitude*, and *longitude*) and SST_{MODIS}. The mapping function is later used to reconstruct the MODIS SST where MODIS SST is missing, but AMSR-E SST exists. The final merged SST is achieved by combining the reconstructed MODIS SST with the original MODIS SST and performing necessary post-processing. The following subsections detail the main steps of the method.

Extracting Ocean Pixels in MODIS SST Based on AMSR-E SST

Extracting ocean pixels from the satellite SST data is the prior step for further evaluating the spatial coverage of satellite SST before and after merging. It can be easily achieved for AMSR-E SST because the land pixels are marked separately in AMSR-E SST with a flag value 255. However, we cannot directly determine land pixels in MODIS SST because the MODIS SST quality control layer uses the same flag 255 to represent land, gross clouds, and other errors. By using the cross-check method proposed in the studies by Li et al. (2013) and Zhu et al. (2018), we can extract the land pixels in MODIS SST and further obtain the ocean pixels with the assistance of AMSR-E SST. The principle of this method is formulated as

$$IsLandPixel(MODIS\ \ SST) = \begin{cases} True & if\ \ flag(MOIDS\ \ SST) = 255\ \ and\ \ flag(AMSR-E\ \ SST) = 255 \\ False & otherwise \end{cases} \tag{1}$$

Namely, for a pixel in MODIS SST flagged with 255 (potential land), if the AMSR-E SST pixel that spatially overlaps the most with the target MODIS pixel is marked as land, then the MODIS SST pixel is identified as land. Otherwise, it is regarded as an ocean pixel with gross clouds and other errors. When the land pixels are identified in MODIS SST, ocean pixels can then be easily extracted.

Data Location Matching

To achieve high-resolution and high-spatial-coverage merged SST, the missing high-resolution MODIS SST pixels should be reconstructed based on the low-resolution cloud-free AMSR-E SST, where the AMSR-E SST has value. Therefore, an important step of our method is establishing a mapping relation between the MODIS SST and AMSR-E SST at the same location. To achieve this, first we must match MODIS SST and AMSR-E SST where the values of both SST exist in the study area. The output of the matching is quad-tuples ($SST_{AMSR\text{-}E}$, *latitude*, *longitude*, and SST_{MODIS}), which will feed into the deep neural network model for model establishment. **Algorithm 1** achieves this goal, where $grid\ \ resolution_{AMSR\text{-}E\ \ SST} = 0.25^{\circ}$, and *ceil(x)* function rounds x to the smallest integer that is bigger than or equal to x.

Algorithm 1 | Algorithm for creating quad-tuples

1. **procedure** CREATING QUAD-TUPLES
2. **for** each MODIS SST pixel $P_{MODIS\ SST}$ that has value SST_{MODIS} with quality flag = 0, denoting its center point coordinate as ($lat_{MODIS,}$, lon_{MODIS}) **do**
3. Calculate the latitude index $Index_{lat}$ and longitude index $Index_{lon}$ of the center point of the MODIS SST pixel in the corresponding AMSRE SST pixel, denoted as $P_{AMSR\text{-}E\ SST}$ and calculated using the following formulas:
$$Index_{lat} = ceil\left(\frac{lat_{MODIS} - lat_{\min}}{grid\ \ resolution_{AMSR-E\ \ SST}}\right)$$
$$Index_{lon} = ceil\left(\frac{lon_{MODIS} - lon_{\min}}{grid\ \ resolution_{AMSR-E\ \ SST}}\right)$$
4. Obtain the corresponding AMSR-E SST value $SST_{AMSR\text{-}E}$ of $P_{AMSR\text{-}E\ SST}$ by $Index_{lat}$ and $Index_{lon}$
5. Create a quad-tuple ($SST_{AMSR\text{-}E}$, *latitude, longitude, and* SST_{MODIS}) where *latitude* = lat_{MODIS} and *longitude* = lon_{MODIS}
6. **end for**
7. **end procedure**

Outlier Removal and Normalization

Before feeding the quad-tuples ($SST_{AMSR\text{-}E}$, *latitude*, *longitude*, and SST_{MODIS}) for model establishment, first we must perform some preprocessing for data quality control, including outlier removal and data normalization. The outlier removal can help avoid the decrease in modeling accuracy caused by outliers in the training data (Khamis et al., 2005). The normalization of the quad-tuples can enhance the neural network's training speed and performance (Puheim & Madarász, 2014).

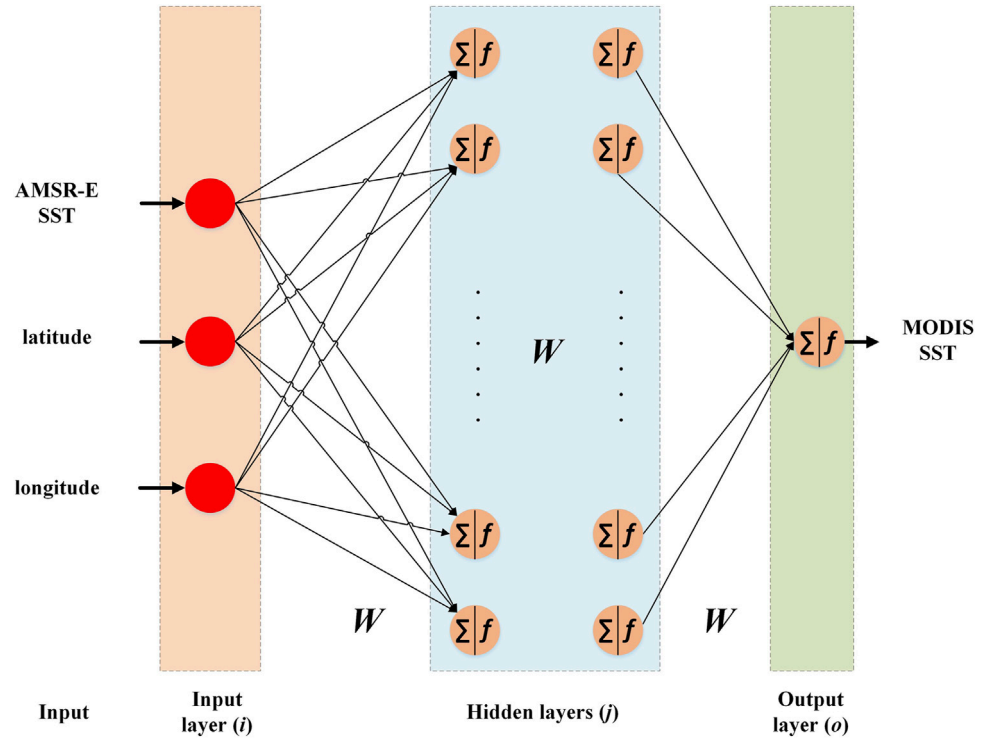

FIGURE 3 | Structure of a feedforward deep neural network model.

For outlier removal, we calculate the difference between AMSR-E SST and MODIS SST of each quad-tuple and obtain a difference value set. The quad-tuple with its difference value falling outside of 3 standard deviations of mean of the difference set are flagged as an outlier and removed. Then, the remaining quad-tuples go to the next step for normalization.

To perform normalization, we first split the N rows (N is the number of quad-tuples obtained after outlier removal) of quad-tuples ($SST_{AMSR\text{-}E}$, *latitude*, *longitude*, and SST_{MODIS}) into N rows of triple-tuples ($SST_{AMSR\text{-}E}$, *latitude*, and *longitude*) and N rows of scalar value SST_{MODIS}. Then, we map values of each row of both the triple-tuples and the scalar values to [-1 1] by using the following equation.

$$y = \frac{(y_{\max} - y_{\min}) * (x - x_{\min})}{(x_{\max} - x_{\min})} + y_{\min} \quad (if \;\; x_{\max} = x_{\min}, \;\; then \;\; y = x), \tag{2}$$

where $y_{max} = 1$, $y_{min} = -1$, x_{max}, x_{min} are the maximum and minimum values of each row, x is the value to be normalized, and y is the normalization result of x. The normalized triple-tuple will be used as input and the normalized scalar value will be used as the desired output of the deep neural network model during both the genetic algorithm–assisted parameter optimization process and the deep neural network model training process.

Genetic Algorithm–Assisted Deep Neural Network Model

Deep neural network models are good at modeling nonlinear and complex relationships among variables. Therefore, in this research, we use a deep neural network model for modeling the relationships between MODIS SST and AMSR-E SST together with locations, namely, obtaining the relationship $SST_{MODIS} = f(SST_{AMSR-E}, latitude, longitude)$. However, the usually adopted gradient-based optimizers for training neural networks usually lead to a local optimum instead of a global optimum. Therefore, in this study, a global search method, that is, the genetic algorithm, is utilized to help prevent the deep neural network model from being trapped in a local optimum, which has been demonstrated in previous studies (Sexton et al., 1998; Tahmasebi & Hezarkhani, 2012). The details of the model establishing process, including deep neural network model design, genetic algorithm–based parameter optimization, deep neural network model training, and performance validations, are explained as follows.

Deep Neural Network Model

The deep neural network model used in this research is a feed-forward deep neural network model, the structure of which is shown in **Figure 3**. It consisted of an input layer, one or more hidden layers, and an output layer. Each layer is consisted of a number of neurons. Neurons between the layers are connected with varying weights (denoted as W in **Figure 3**). The weighted sum (denoted using the operator $\sum$ in **Figure 3**) of all the inputs

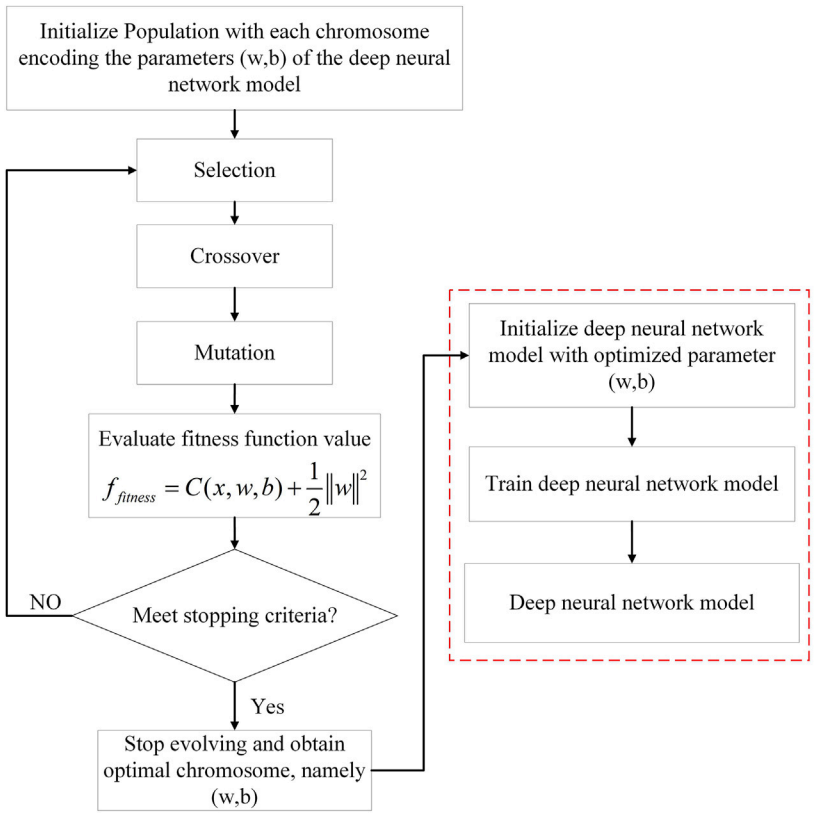

FIGURE 4 | Workflow of using GA to optimize the initial parameters of the deep neural network model.

TABLE 2 | Configuration of genetic algorithm.

Initial population size	Elite count	Crossover fraction	Mutation rate	Generations
200	10	0.8	0.01	100

to a neuron plus a bias is activated by an activation function $f(\cdot)$, producing the output of the neuron.

The deep neural network model is trained by using a backpropagation (BP) algorithm. The interconnecting weights and bias are updated iteratively to minimize the output error, which is usually a mean-square-error (MSE) between the targeted outputs and actual outputs of the neural network over all the training samples (Zare Abyaneh et al., 2016), and is calculated as

$$\mathrm{E} = \frac{1}{N}\sum_{x,w,b}\left\|y(x) - \hat{y}(x)\right\|^2, \tag{3}$$

where $\boldsymbol{b}$ and $\boldsymbol{w}$ denote all the bias and weights in the network, respectively. N denotes the number of inputs, $\boldsymbol{x}$ is the input of the network, and $y(x)$, $\hat{y}(x)$ are the vectors of the activated output of the network and targeted output, respectively.

During the backpropagation, the weights and bias are updated using a gradient descent strategy. In each iteration, the gradient is first calculated using

$$\begin{aligned}\frac{\partial E}{\partial w} &= \left(\frac{\partial E}{\partial w_1}, \frac{\partial E}{\partial w_2}, \ldots, \frac{\partial E}{\partial w_\ell}\right)\\ \frac{\partial E}{\partial b} &= \left(\frac{\partial E}{\partial b_1}, \frac{\partial E}{\partial b_2}, \ldots, \frac{\partial E}{\partial b_k}\right).\end{aligned} \tag{4}$$

Then, each weight and bias are updated using the increment,

$$\Delta w_i = -\eta\frac{\partial E}{\partial w_i} \text{ for i} = 1,\ 2, \text{and } \ell, \tag{5}$$

$$\Delta b_j = -\eta\frac{\partial E}{\partial b_j} \text{ for j} = 1,\ 2, \text{and } k, \tag{6}$$

where η is the learning rate, which is a constant.

The following parameters need to be determined in the deep neural network model: the number of hidden layers, the number of neurons in each hidden layer, the activation function for each layer, and the learning rate. The parameters in this study are set by combining experience and experiments. We choose a 3-layer architecture with three neurons for the input layer which receives

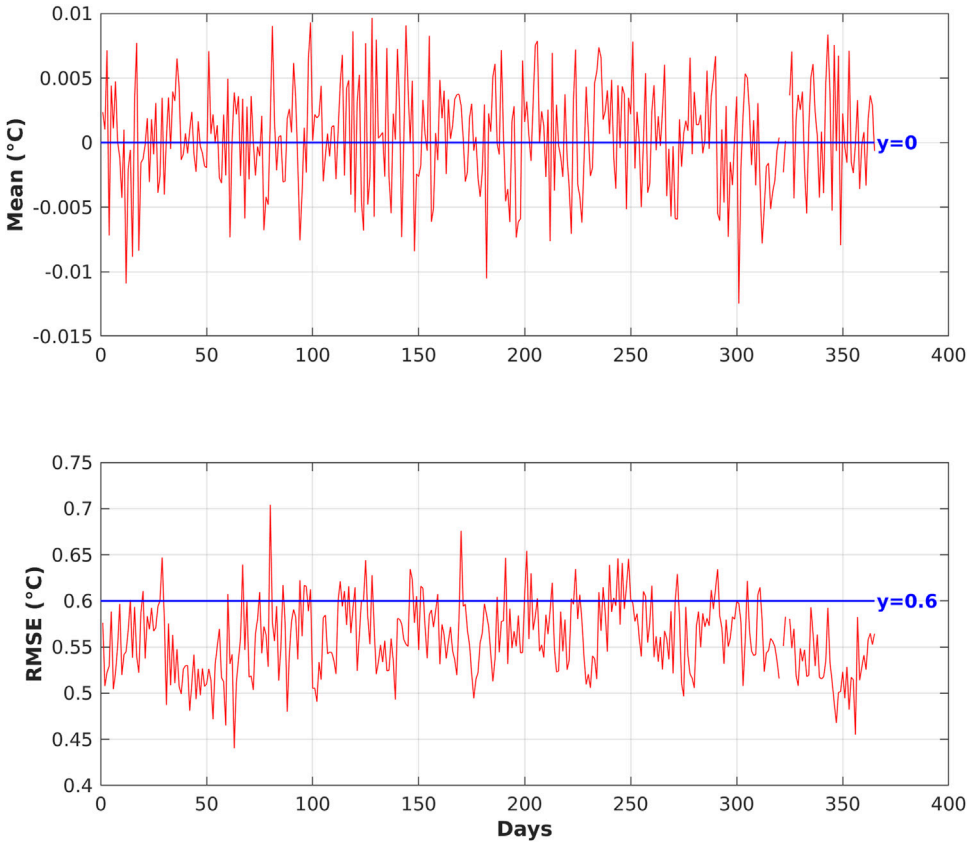

FIGURE 5 | Statistics of GA-DNNM performance on the daily test dataset in 2005.

the input triple-tuples ($SST_{AMSR\text{-}E}$, *latitude*, and *longitude*), seven neurons for the hidden layer, and one neuron for the output layer which outputs the estimated SST_{MODIS} value. The number of hidden layers and number of neurons in each hidden layer are determined by the process that we first chose several configurations of the number of hidden layers and the number of neurons in each hidden layer, then we compare the prediction performance of each configuration, and finally, the one that achieves the best performance is chosen. The sigmoid function is used as the activation function for the hidden layer, and the linear function is used as the activation function for the output layer. The learning rate is set to 0.05.

Genetic Algorithm–Based Deep Neural Network Parameter Optimization

By using the gradient descent method, the cost function is driven to a low value which however is without global convergence guarantee. Besides, the gradient-based training method is sensitive to the values of initial parameters (i.e., weights and bias). Thus, to prevent the deep neural network model from being trapped in a local minimum, the GA approach is adopted.

GA is a meta-heuristic method for solving optimization problems. Some researchers have demonstrated that GA can be used to help the neural network achieve global optimum (Mahmoudabadi et al., 2009; Wang et al., 2016; Yu & Xu, 2014). GA is based on the process of natural selection (Whitley et al., 1990), during which a population of individual solutions is repeatedly modified, and the population finally reaches an optimal solution through successive generations based on the following rules.

- Selection: select individuals as parents in the current generation to reproduce next generation based on their fitness.
- Crossover: combine the genes of parents to produce children as individuals in the next generation.
- Mutation: introduce random changes to a chromosome to produce children for the next generation.

Specifically, for optimizing the deep neural network model in this study, parameters of the neural network, including weights and bias ($\boldsymbol{w}$, $\boldsymbol{b}$), are encoded to a chromosome, and a population of such chromosomes is created and initialized. The fitness of each chromosome is evaluated using

$$f_{\text{fitness}}(chromosome) = C(x, w, b) + \frac{1}{2}\|w\|^2, \tag{7}$$

where $C(x,w,b)$ is the MSE of the deep neural network model whose parameters are specified by the chromosome. The L2 regularization term $\frac{1}{2}\|w\|^2$ is also added to the fitness function to

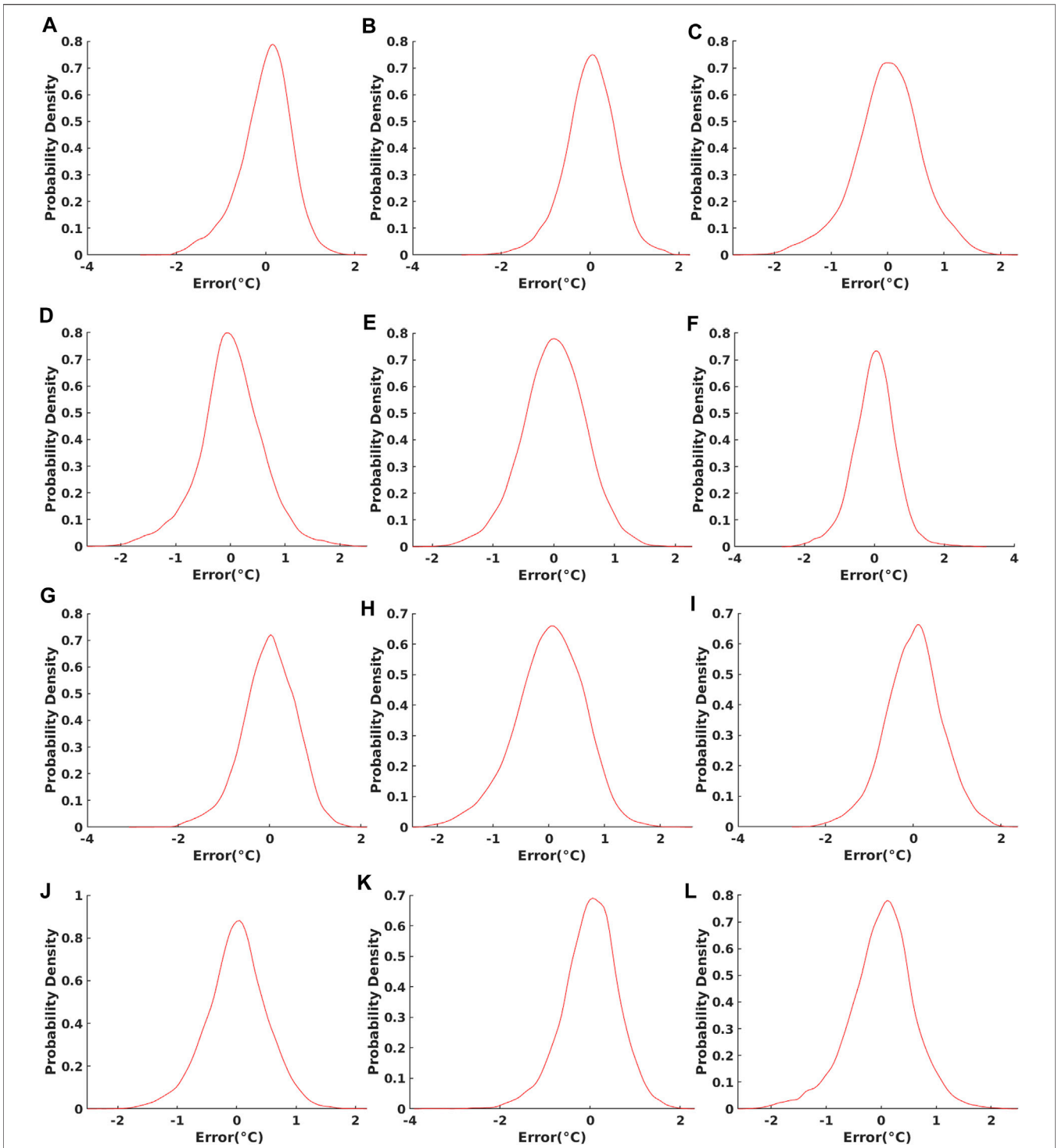

FIGURE 6 | Estimated probability density of residuals of the proposed GA-DNNM for prediction on test data. Subfigures **(A)–(L)** are the estimated probability densities of residuals for the first day of January to December in 2005, respectively.

balance weights. Chromosome with the least fitness function value is considered as the fittest individual, namely, the best solution in the current generation of population. Based on the rules defined above, the population evolves from generation to generation and finally stops at a specific generation according to the stopping criteria. The chromosome in generation G_t with the best fitness is the final optimal solution we wanted to seek, represented as $optimalInitialParam = (w, b)_{G_t}$. This *optimalInitialParam* is then

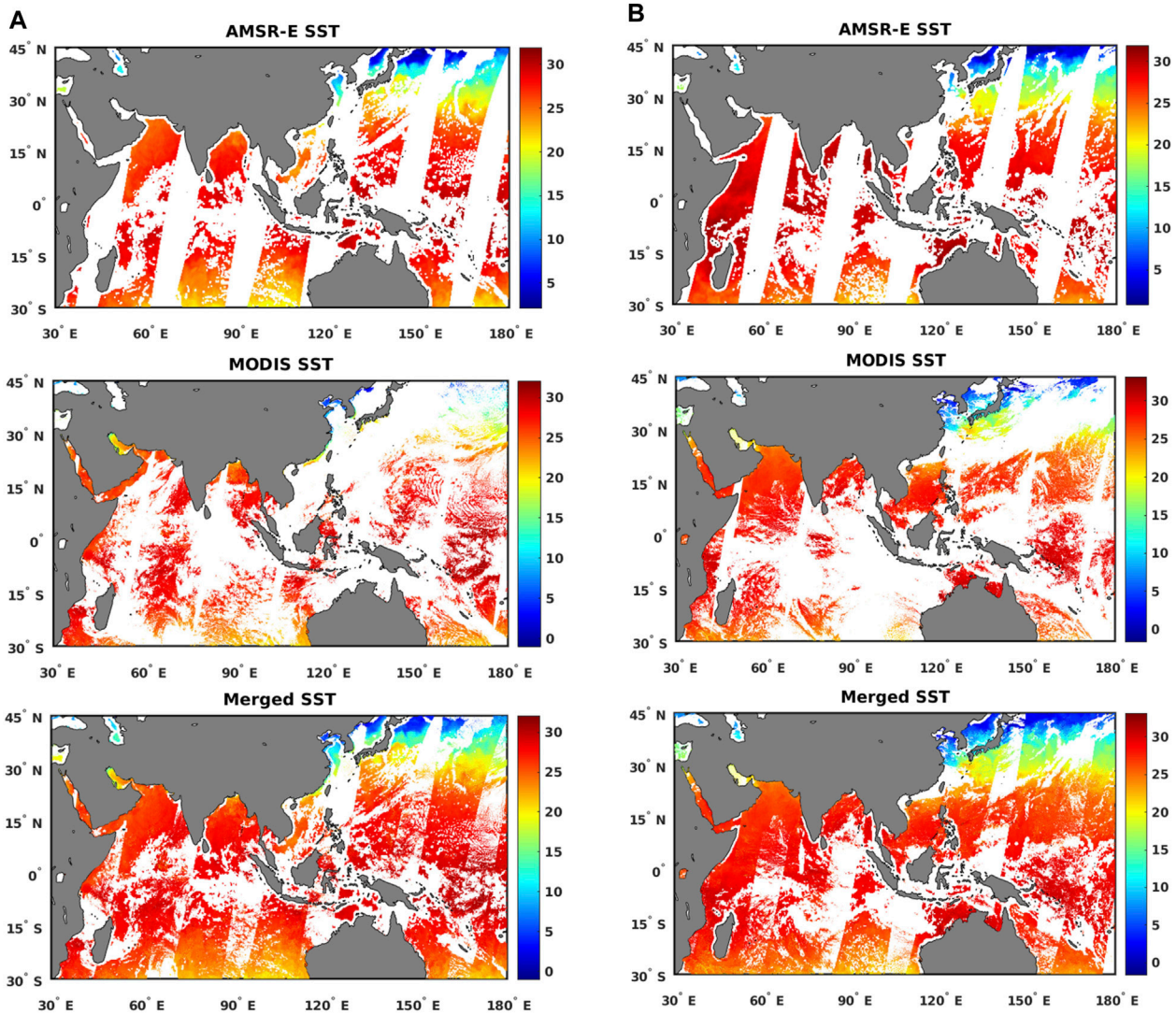

FIGURE 7 | Spatial patterns of AMSR-E SST, MODIS SST, and merged SST on selected days in each season in 2005 with white color representing missing ocean pixels and gray color representing land pixels **(A)** from top to bottom: the three SSTs on January 1, 2005; **(B)** from top to bottom: the three SSTs on April 1, 2005; **(C)** from top to bottom: the three SSTs on July 1, 2005; and **(D)** from top to bottom: the three SSTs on October 1, 2005.

utilized to initialize the deep neural network model for training. With a very high probability, the *optimalInitialParam* can make the neural network converge to a global optimum instead of a local optimum quickly. The whole process is illustrated in **Figure 4**.

There are several parameters in GA that need to be set, including initial population size, number of elite children (individuals with top fitness and directly selected to the next generation of population without any change), crossover fraction, and mutation rate. In this study, we set these parameters empirically, as listed in **Table 2**.

Performance Validation of the Model

Before applying the GA-DNNM to the merging AMSR-E SST and MODIS SST, the performance of the model should be first validated. In this study, we randomly select 90% of the normalized quad-tuples obtained in section ***Outlier Removal and Normalization*** to train the neural network and the remaining 10% to test the generalization performance of the trained network. Two indexes are utilized for performance evaluation: the mean error and root-mean-square-error (RMSE), which is defined as

$$mean\,error = \frac{1}{n}\sum_1^n d_i, RMSE = \sqrt{\frac{\sum_1^n d_i^2}{n}}, \tag{8}$$

where d_i is the error vector calculated by the difference between the desired MODIS SST value and the estimated MODIS SST value of the GA-DNNM, and n is the total number of test samples.

Performance validation results obtained during the experimental period are shown in **Figure 5**. From the testing results, the mean

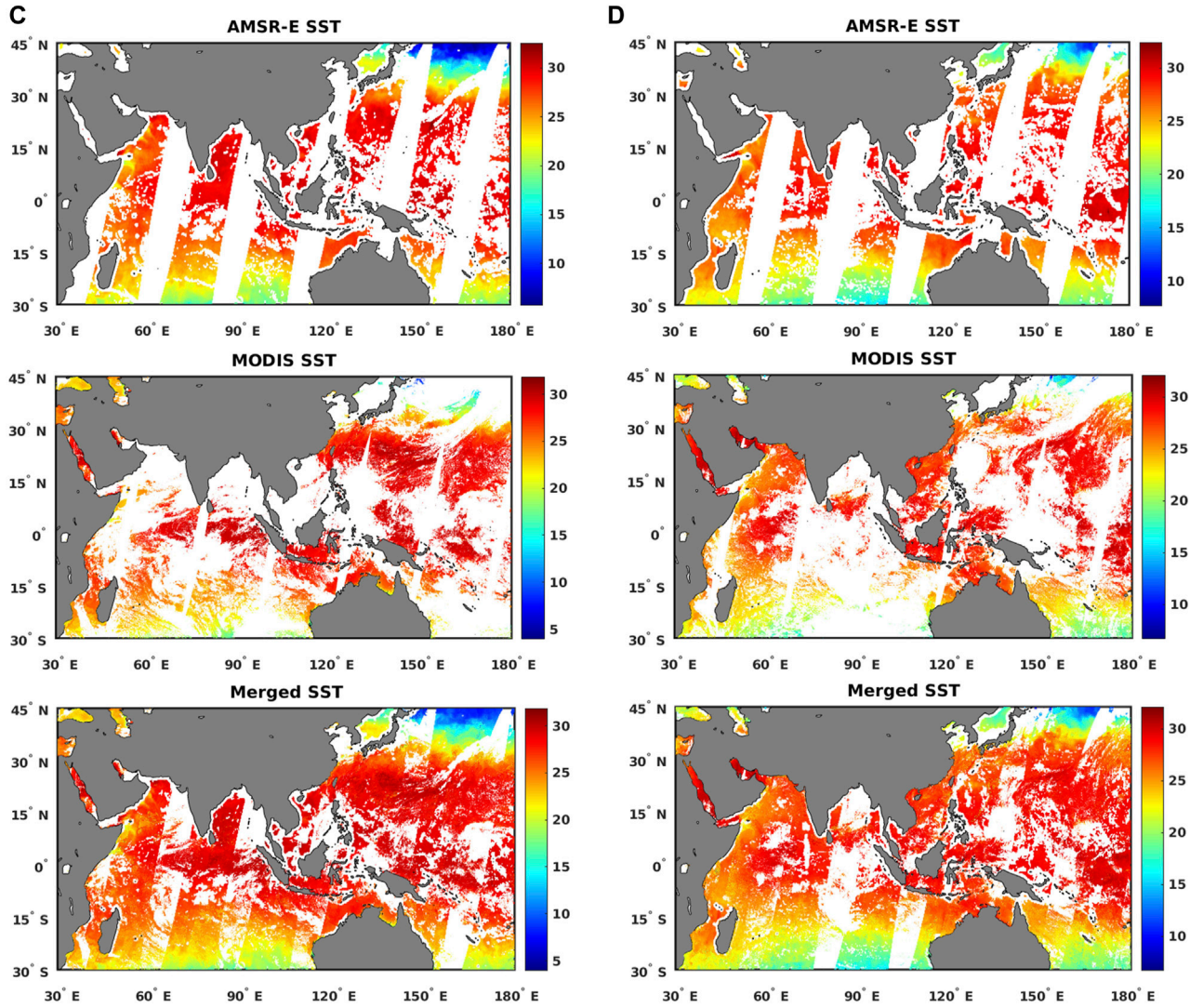

FIGURE 7 | (Continued).

errors are almost equal to 0 °C, and 84.02% of the RMSEs are below 0.6°C. The estimated probability densities of the residuals of prediction on the test dataset of the first day of each month in 2005 are shown in **Figure 6**. It can be seen that the residuals are concentrated around 0. The validation results demonstrate a high generalized prediction accuracy of the GA-DNNM. Therefore, the GA-DNNM is capable of establishing the relationship between AMSR-E SST and MODIS SST through learning from the training dataset and can be further applied to merging these two SSTs.

Post-Processing

When the merged SST has been obtained by GA-DNNM, we post-process it by removing pixels with gross error. The gross error pixels are those whose SST values are beyond the range of −3°C–35°C, which is the union of the valid data range of the MODIS SST and that of the AMSR-E SST.

RESULTS AND DISCUSSION

For evaluating the proposed method, experiments are conducted on each day of 2005, expect for November 17, 2005 when the AMSR-E SST's spatial coverage is 0.0% in the study area, and November 20, 2005 when the AMSR-E SST's spatial coverage is 0.0524% in the study area and has no match with the drifting buoy observations. 4 km daily merged SST products with improved quality are generated in the AIPO area.

Comparison of the Spatial Coverage of MODIS SST, AMSR-E SST, and Merged SST

The spatial coverage is a critical index for measuring the quality of SST. In this section, we evaluate the spatial coverage of MODIS SST, AMSR-E SST, and merged SST both visually and quantitatively.

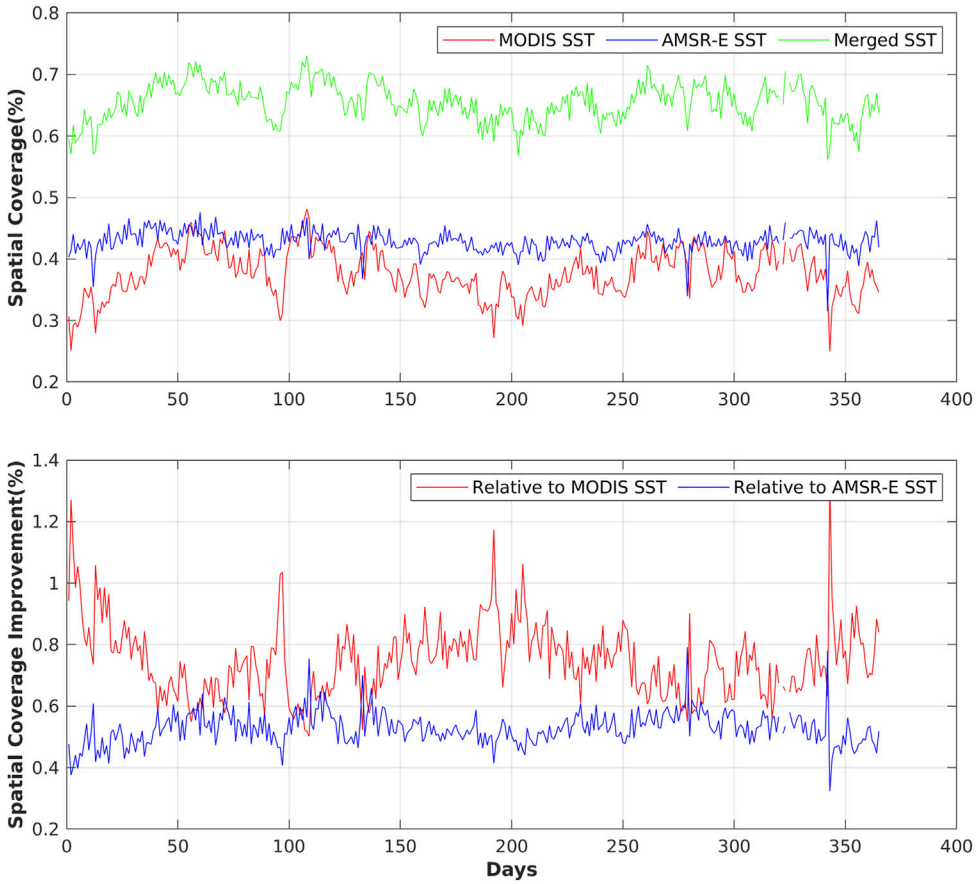

FIGURE 8 | Comparison of spatial coverage of the daily MODIS SST, daily AMSR-E SST, and daily merged SST in 2005 in the AIPO area.

Intuitively, from **Figure 7**, we can see that the spatial coverage and continuity of SST are greatly improved after merging. The atmospheric contaminations and costal effects have been eliminated extensively. Further, in a quantitative way, we examine the spatial coverage of the three SSTs in the year 2005, in the study area using the following formula.

$$spatialCoverage = \frac{N_{valid\ SST}}{N_{ocean}}, \tag{9}$$

where $N_{valid\ SST}$ and N_{ocean} denote the number of valid SST pixels and total ocean pixels, respectively. The number of ocean pixels is obtained using the method introduced in section ***Extracting Ocean Pixels in MODIS SST Based on AMSR-E SST***.

The quantitative results are shown in **Figure 8**. The spatial coverage of the original MODIS SST, original AMSR-E SST, and merged SST are ranging from 25.0 to 48.1%, 31.5 to 47.6%, and 56.1 to 73.1%, respectively. The merged SST has much higher spatial coverage than MODIS SST and AMSR-E SST, with a minimum improvement by 50.2% on April 19, 2005 and maximum improvement by 131.7% on December 9, 2005 compared with MODIS SST. The improvement of the spatial coverage relative to AMSR-E SST ranges from 32.3 to 79.2%. The spatial coverage of AMSR-E SST is quite stable, while there is more fluctuation for MODIS SST due to the vulnerability of the MODIS sensor to various atmospheric contaminations such as cloud cover, thick fogs, and concentrated aerosols. The spatial coverage of merged SST has the same fluctuation characteristics as MODIS SST because of the stability of AMSR-E SST and fluctuation of MODIS SST.

Validation of Reconstructed SST and Merged SST With Drifting Buoy Observations

To validate the reconstructed SST and merged SST (SST in the whole study area), a linear regression of the MODIS SST with the drifting buoy observations, the AMSR-E SST with the drifting buoy observations, the reconstructed SST with the drifting buoy observations, and the merged SST with the drifting buoy observations are performed each for each day in the study period. R-square (R^2), RMSE, mean bias (Bias), and correlation coefficient are used for quantitatively evaluating the accuracy of SST.

To be concise, we select 1 day in each season to illustrate the accuracy of merged results, as shown in **Figure 9A–D**. From **Figure 9**, it can be seen that R^2 and correlation coefficient of reconstructed SST are with little difference with those of AMSR-E SST but are much greater than those of MODIS SST. The RMSE of reconstructed SST in the time frame mostly lies between that of the AMSR-E SST and MODIS SST, that is, greater than AMSR-E

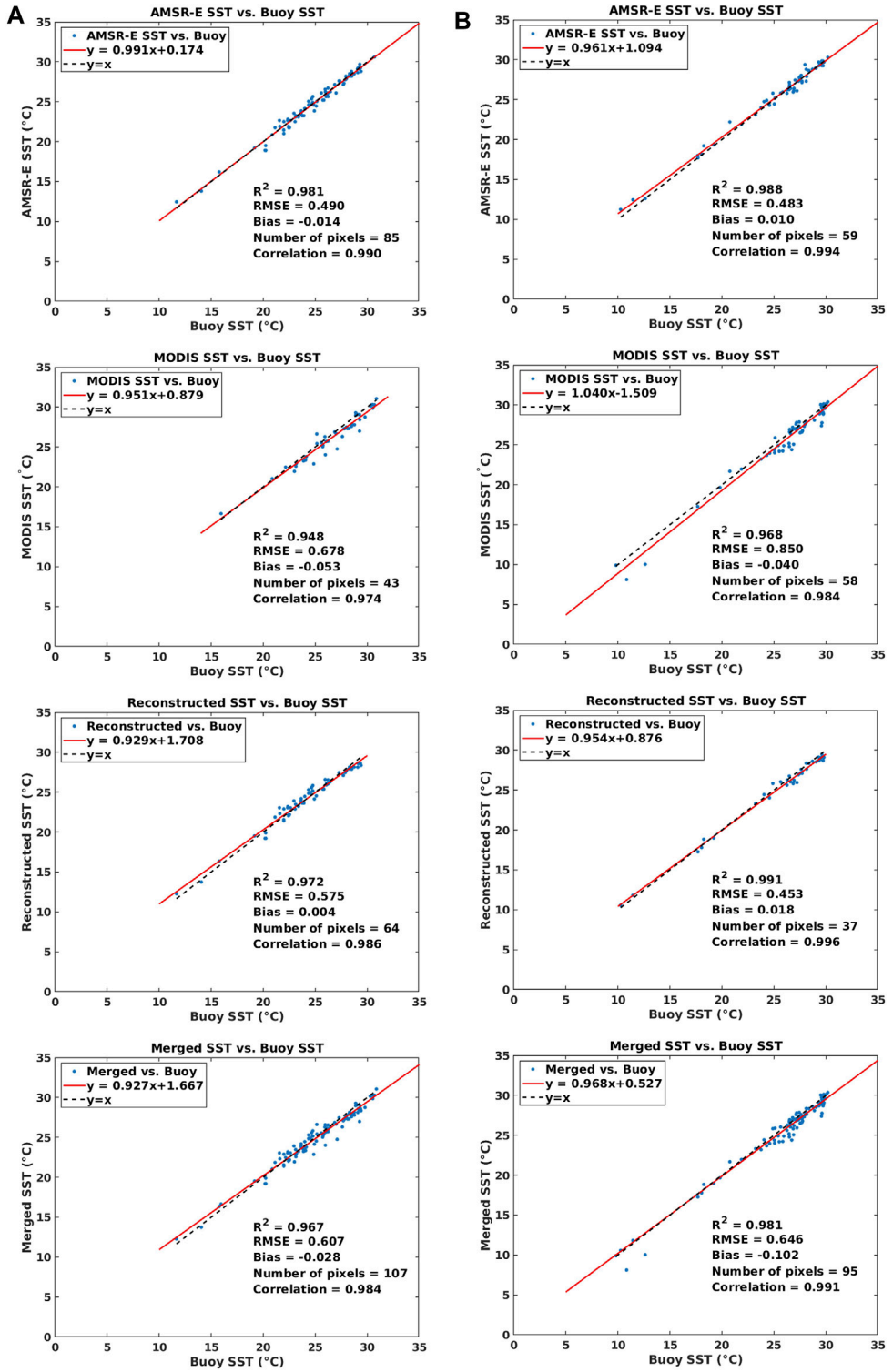

FIGURE 9 | Validation of reconstructed SST and merged SST with drifting buoy observations on selected date in each season in 2005. From top to bottom: **(A)** validation on January 1, 2005, **(B)** validation on April 1, 2005, **(C)** validation on July 1, 2005, and **(D)** validation on October 1, 2005.

SST and smaller than MODIS SST. The bias of the reconstructed SST is also much smaller than that of the MODIS SST. As with the finally merged SST, its R^2 and correlation coefficient are greater than those of MODIS SST, and the RMSE greater than that of AMSR-E SST and reconstructed SST but smaller than that of MODIS SST. The reason why merged SST has bigger RMSE than

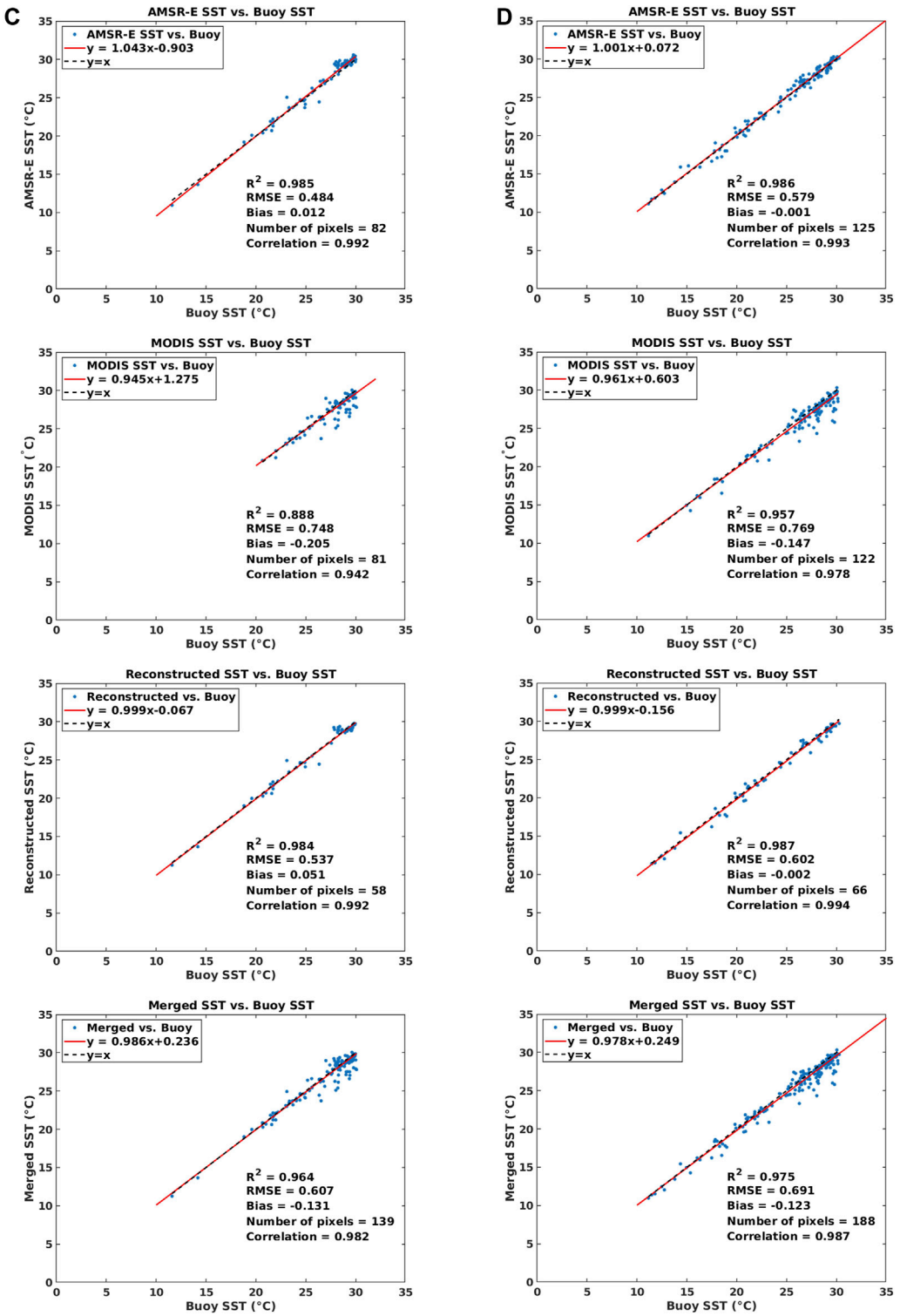

FIGURE 9 | (Continued).

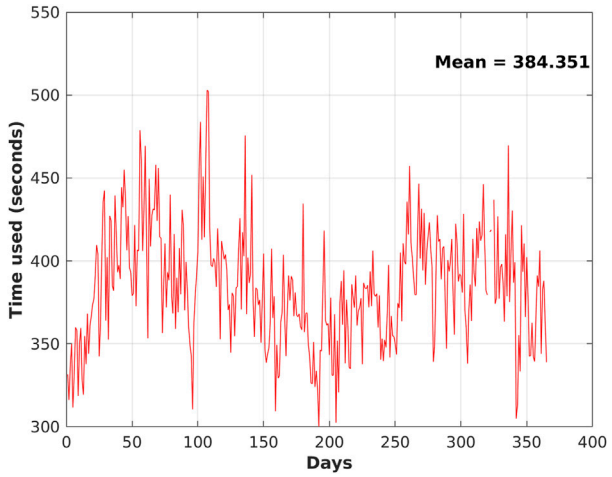

FIGURE 10 | Time consumed for each day's data processing.

reconstructed SST is that during the integration of reconstructed SST and MODIS SST to produce merged SST, the error of MODIS SST may be introduced. The RMSE and bias of the merged SST are acceptable, with higher accuracy than MODIS SST, and meanwhile keeps the same spatial resolution (4 km) and temporal resolution (1 day) as MODIS SST.

The average RMSE and average bias of the reconstructed SST are 0.502°C and 0.006°C, respectively. The average RMSE and average bias of the merged SST in the AIPO area are 0.603°C and −0.082°C, respectively. Errors of the merged SST may come from three aspects: 1) error of AMSR-E SST and MODIS SST: the merged SST is based on the AMSR-E SST and MODIS SST. Therefore, errors existing in two merging source of SST may contribute to errors in merged SST. 2) Errors of GA-DNNM can also be a source of errors for the merged SST; 3) difference of measured depth: AMSR-E SST, MODIS SST, and buoy SST measured at ~ um (skin SST), ~ 1 mm and 0.2–0.3 m (bulk SST), respectively. The merged SST can be seen measuring the same depth as MODIS SST, which however is coupled with the atmosphere–ocean exchange of heat and momentum closely, making the bulk-skin difference a quantity which varies with quite short time and space scales (Emery et al., 2001; Zhu et al., 2018).

Efficiency Analysis of the Proposed Method

The time taken for the whole processing process each day, including data preprocessing, data location matching, GA-DNNM establishment, and SST merging, is shown in **Figure 10**. The time fluctuates for different days, with the longest time being 502.964 s on April 17 and the lowest time being 300.072 s on July 11, and the average time for each day being 384.351 s. It is little bit time-consuming, due to two reasons: 1) In the genetic algorithm, each chromosome carries 36 genes (calculated based on the structure of the neural network designed) that need to be optimized, and the number of inputs used for evaluating fitness of individuals in a population in each generation is firmly large (around 200,000–350,000) which involves lots of computation. 2) The procedure runs on a desktop with one Intel (R) Core (TM) i9-9,920X CPU at 3.5GHz and 48.0 GB RAM, whose computing resources and computing capabilities are limited. The fluctuating characteristics of the time consumed in each day are primarily because of the varying number of inputs for the genetic algorithm, neural network model training, and SST reconstruction. In future research, the configuration of the genetic algorithm may be further optimized, and high-performance computing (HPC) infrastructure and technologies (Wright and Wang, 2011) may be used to improve the efficiency.

CONCLUSIONS

SST is a crucial parameter for oceanic and atmospheric models. It plays an important role for weather forecasting and climate change monitoring. Therefore, getting high-resolution SST both in time and space, as well as high spatial coverage, is of vital importance. Satellite observations are the major sources based on which large-area SST is derived. However, due to the difference in the imaging mechanism, different satellite observations have different limitations. Infrared satellite sensors usually have high spatial resolutions but are vulnerable to various atmospheric contaminations such as cloud cover, thick fogs, and concentrated aerosols, while microwave sensors can penetrate clouds and aerosols but usually with low resolution and cannot obtain data near coasts. Consequently, a single sensor usually cannot achieve desirable SST.

This study therefore merges SST data from both infrared sensor (MODIS SST) and microwave sensor (AMSR-E SST) synergistically to produce daily SST with a spatial resolution of 4 km which has a much higher spatial coverage than the SST of each sensor, much higher spatial resolution than SST of microwave sensor, and higher accuracy than SST of infrared sensor. During this process, a genetic algorithm–assisted deep neural network model is established and evaluated. The validation of the reconstructed SST with drifting buoy observations each day during the year 2005 (363 days of data are analyzed) shows an average RMSE and average bias of 0.502°C and 0.006°C, respectively, and an average RMSE and average bias of 0.603°C and −0.082°C, respectively, for the merged SST in the whole study area. With the high generalized prediction accuracy, the model can be used for extended merging of the MODIS SST and AMSR-E SST in other years.

With the improved SST, extensive climate applications promise to be better supported, and the marine environment including spatiotemporal patterns and variability can be better monitored and understood than using SST from a single sensor alone. Furthermore, the method is applicable to merging SST at a global scale, which can provide improved data for and further benefit global and regional climate research and applications.

The GA-assisted optimization strategy is both computation- and data-intensive, which takes significant time for the GA-DNNM workflow. For future work on larger geographic areas, cyberGIS and high-performance computing approaches may be developed to accelerate and enhance the workflow (Liu & Wang, 2015; Wang & Goodchild, 2019). Besides, the proposed model

currently could only be applied to the locations where AMSR-E SST is available, making it hard to achieve daily merged SST with 100% spatial coverage. How to expand the proposed model to incorporate more kinds of satellite-derived SSTs and drifting buoy observations to produce spatially seamless SST may also be a future direction.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. These data can be found here: The MODIS SST data analyzed for this study can be obtained at PO. DACC website: https://podaac.jpl.nasa.gov/dataset/MODIS_AQUA_L3_SST_THERMAL_DAILY_4KM_NIGHTTIME_V2014.0. The AMSR-E SST analyzed for this study can be obtained at the Remote Sensing Systems website: http://data.remss.com/amsre/bmaps_v07/y2005/). The drifting buoy data analyzed for this study can be obtained at NOAA National Centers for Environmental Information website: https://doi.org/10.25921/7ntx-z961.

AUTHOR CONTRIBUTIONS

CX, CH, and NC conceived and designed the experiments; CX performed the experiments; CX, ZC, and XZ analyzed the data; CX wrote the paper; and XT helped revise the paper.

FUNDING

This research was supported by the National Nature Science Foundation of China (NSFC) Program (Nos. 42001372 and 42071380), the National Key R&D Program (No. 2018YFB2100501), the Project funded by China Postdoctoral Science Foundation (Nos. 2019M661621 and 2021T140513), Shanghai Municipal Science and Technology Major Project (No. 2021SHZDZX0100), and Shanghai Municipal Commission of Science and Technology Project (No.19511132101).

ACKNOWLEDGMENTS

The authors would like to thank the following data, tool, and service providers: OBPG at the NASA Goddard Space Flight Center and NASA Physical Oceanography Distributed Active Archive Center (PO.DACC) for providing MODIS SST data and corresponding data reading software; Remote Sensing Systems (RSS) for providing AMSR-E SST data and corresponding data reading routines (AMSR data are produced by Remote Sensing Systems and were sponsored by the NASA AMSR-E Science Team and the NASA Earth Science MEaSUREs Program. Data are available at www.remss.com); NOAA National Centers for Environmental Information for providing AOML Global Drifter interpolated data. The authors would also like to thank Professor Shaowen Wang at the University of Illinois at Urbana-Champaign for his help in the investigation of this work.

REFERENCES

Armstrong, E. (2007). MODIS Sea Surface Temperature (SST) Products. Available at: ftp://podaac-ftp.jpl.nasa.gov/allData/modis/L3/docs/modis_sst.html (Accessed March 1, 2021).

Bretherton, F. P., Davis, R. E., and Fandry, C. B. (1976). A Technique for Objective Analysis and Design of Oceanographic Experiments Applied to MODE-73. *Deep Sea Res. Oceanographic Abstr.* 23 (7), 559–582. doi:10.1016/0011-7471(76)90001-2

Cai, Y., Guan, K., Peng, J., Wang, S., Seifert, C., Wardlow, B., et al. (2018). A High-Performance and In-Season Classification System of Field-Level Crop Types Using Time-Series Landsat Data and a Machine Learning Approach. *Remote Sensing Environ.* 210, 35–47. doi:10.1016/j.rse.2018.02.045

Chang-Xiang, Y., Jiang, Z., and Ji-Ping, X. (2010). An Ocean Reanalysis System for the Joining Area of Asia and Indian-Pacific Ocean. *Atmos. Oceanic Sci. Lett.* 3 (2), 81–86. doi:10.1080/16742834.2010.11446848

Chao, Y., Li, Z., Farrara, J. D., and Hung, P. (2009). Blending Sea Surface Temperatures from Multiple Satellites and *In Situ* Observations for Coastal Oceans. *J. Atmos. Ocean. Technol.* 26 (7), 1415–1426. doi:10.1175/2009JTECHO592.1

Dahl, G. E., Dong Yu, D., Li Deng, L., and Acero, A. (2012). Context-Dependent Pre-trained Deep Neural Networks for Large-Vocabulary Speech Recognition. *IEEE Trans. Audio Speech Lang. Process.* 20 (1), 30–42. doi:10.1109/TASL.2011.2134090

Donlon, C. J., Martin, M., Stark, J., Roberts-Jones, J., Fiedler, E., and Wimmer, W. (2012). The Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) System. *Remote Sensing Environ.* 116, 140–158. doi:10.1016/j.rse.2010.10.017

Emery, W. J., Castro, S., Wick, G. A., Schluessel, P., and Donlon, C. (2001). Estimating Sea Surface Temperature from Infrared Satellite and *In Situ* Temperature Data. *Bull. Amer. Meteorol. Soc.* 82 (12), 2773–2785. doi:10.1175/1520-0477(2001)082<2773:ESSTFI>2.3.CO;2

Guan, L., and Kawamura, H. (2004). Merging Satellite Infrared and Microwave SSTs: Methodology and Evaluation of the New SST. *J. Oceanogr.* 60 (5), 905–912. doi:10.1007/s10872-005-5782-5

Guo, P. (2010). "Study on Bayesian Hierarchal Model-Based SST Data Fusion Methods," in Proc. Remote Sensing of the Ocean, Sea Ice, and Large Water Regions 2010. Editors R. B. Charles, P. M. Stelios, N. Xavier, and V.-R. Miguel (Washington, United States: International Society for Optics and Photonics), 7825O. doi:10.1117/12.864912

Høyer, J. L., Karagali, I., Dybkjær, G., and Tonboe, R. (2012). Multi Sensor Validation and Error Characteristics of Arctic Satellite Sea Surface Temperature Observations. *Remote Sensing Environ.* 121, 335–346. doi:10.1016/j.rse.2012.01.013

Kaul, M., Hill, R. L., and Walthall, C. (2005). Artificial Neural Networks for Corn and Soybean Yield Prediction. *Agric. Syst.* 85 (1), 1–18. doi:10.1016/j.agsy.2004.07.009

Khamis, A., Ismail, Z. I., Khalid, K. H., and Mohammed, A. T. (2005). The Effects of Outliers Data on Neural Network Performance. *J. Appl. Sci.* 5 (8), 1394–1398. doi:10.3923/jas.2005.1394.1398

Le, T. H. (2011). Applying Artificial Neural Networks for Face Recognition. *Adv. Artif. Neural Syst.* 2011, 1–16. doi:10.1155/2011/673016

Li, A., Bo, Y., Zhu, Y., Guo, P., Bi, J., and He, Y. (2013). Blending Multi-Resolution Satellite Sea Surface Temperature (SST) Products Using Bayesian Maximum Entropy Method. *Remote Sensing Environ.* 135, 52–63. doi:10.1016/j.rse.2013.03.021

Li, T., Shen, H., Yuan, Q., Zhang, X., and Zhang, L. (2017). Estimating Ground-Level PM2.5 by Fusing Satellite and Station Observations: A Geo-Intelligent Deep Learning Approach. *Geophys. Res. Lett.* 44 (23), 11985–11993. doi:10.1002/2017GL075710

Liu, Y. Y., and Wang, S. (2015). A Scalable Parallel Genetic Algorithm for the Generalized Assignment Problem. *Parallel Comput.* 46, 98–119. doi:10.1016/j.parco.2014.04.008

Lumpkin, R., and Centurioni, L. (2019). *Data from: Global Drifter Program Quality-Controlled 6-hour Interpolated Data from Ocean Surface Drifting*

Buoys. NOAA National Centers for Environmental Information. doi:10.25921/7ntx-z961

Mahmoudabadi, H., Izadi, M., and Menhaj, M. B. (2009). A Hybrid Method for Grade Estimation Using Genetic Algorithm and Neural Networks. *Comput. Geosci.* 13 (1), 91–101. doi:10.1007/s10596-008-9107-9

McIntosh, P. C. (1990). Oceanographic Data Interpolation: Objective Analysis and Splines. *J. Geophys. Res.* 95 (C8), 13529–13541. doi:10.1029/JC095iC08p13529

OBPG (2015). *Data from: MODIS Aqua Level 3 SST Thermal IR Daily 4km Nighttime v2014.0. OBPG, v2014.0*. CA, USA: PO.DAAC. doi:10.5067/MODSA-1D4N4

Panda, S. S., Ames, D. P., and Panigrahi, S. (2010). Application of Vegetation Indices for Agricultural Crop Yield Prediction Using Neural Network Techniques. *Remote Sensing* 2 (3), 673–696. doi:10.3390/rs2030673

Puheim, M., and Madarász, L. (2014). "Normalization of Inputs and Outputs of Neural Network Based Robotic Arm Controller in Role of Inverse Kinematic Model," in Article Presented at 12th International Symposium on Applied Machine Intelligence and Informatics (Sami), IEEE, Herl'any, Slovakia (IEEE). doi:10.1109/SAMI.2014.6822439

Reynolds, R. W., Rayner, N. A., Smith, T. M., Stokes, D. C., and Wang, W. (2002). An Improved *In Situ* and Satellite SST Analysis for Climate. *J. Clim.* 15 (13), 1609–1625. doi:10.1175/1520-0442(2002)015<1609:AIISAS>2.0.CO;2

Reynolds, R. W., and Smith, T. M. (1995). A High-Resolution Global Sea Surface Temperature Climatology. *J. Clim.* 8 (6), 1571–1583. doi:10.1175/1520-0442(1995)008<1571:AHRGSS>2.0.CO;2

RSS (2021). AMSR-2/AMSR-E. Available at: http://www.remss.com/missions/amsr/(Accessed March 1, 2021).

Sexton, R. S., Dorsey, R. E., and Johnson, J. D. (1998). Toward Global Optimization of Neural Networks: A Comparison of the Genetic Algorithm and Backpropagation. *Decis. Support Syst.* 22 (2), 171–185. doi:10.1016/S0167-9236(97)00040-7

Tahmasebi, P., and Hezarkhani, A. (2012). A Hybrid Neural Networks-Fuzzy Logic-Genetic Algorithm for Grade Estimation. *Comput. Geosciences* 42, 18–27. doi:10.1016/j.cageo.2012.02.004

Tang, S., Yang, X., Dong, D., and Li, Z. (2015). Merging Daily Sea Surface Temperature Data from Multiple Satellites Using a Bayesian Maximum Entropy Method. *Front. Earth Sci.* 9 (4), 722–731. doi:10.1007/s11707-015-0538-z

USEPA (2021). Climate Change Indicators: Sea Surface Temperature. Available at: https://www.epa.gov/climate-indicators/climate-change-indicators-sea-surface-temperature (Accessed May 28, 2021).

Valverde Ramírez, M. C., de Campos Velho, H. F., and Ferreira, N. J. (2005). Artificial Neural Network Technique for Rainfall Forecasting Applied to the São Paulo Region. *J. Hydrol.* 301 (1-4), 146–162. doi:10.1016/j.jhydrol.2004.06.028

Wang, S., and Goodchild, M. F. (2019). *CyberGIS for Geospatial Innovation and Discovery*. Dordrecht, Netherlands: Springer. doi:10.1007/978-94-024-1531-5

Wang, S., Zhang, N., Wu, L., and Wang, Y. (2016). Wind Speed Forecasting Based on the Hybrid Ensemble Empirical Mode Decomposition and GA-BP Neural Network Method. *Renew. Energ.* 94, 629–636. doi:10.1016/j.renene.2016.03.103

Wang, W., and Xie, P. (2007). A Multiplatform-Merged (MPM) SST Analysis. *J. Clim.* 20 (9), 1662–1679. doi:10.1175/JCLI4097.1

Wright, D., and Wang, S. (2011). The Emergence of Spatial Cyberinfrastructure. *PNAS* 108 (14), 5488–5491. doi:10.1073/pnas.1103051108

Wentz, F. J., Gentemann, C., Smith, D., and Chelton, D. (2000). Satellite Measurements of Sea Surface Temperature through Clouds. *Science* 288 (5467), 847–850. doi:10.1126/science.288.5467.847

Wentz, F. J., Meissner, T., Gentemann, C., and Brewer, M. (2014). *Data from: Remote Sensing Systems AQUA AMSR-E Daily Environmental Suite on 0.25 Deg Grid. R.S. System, V7*. Santa Rosa, CA: Remote Sensing Systems. Available at: http://www.remss.com/missions/amsr/.

Werdell, P. J., Franz, B. A., Bailey, S. W., Feldman, G. C., Boss, E., Brando, V. E., et al. (2013). Generalized Ocean Color Inversion Model for Retrieving marine Inherent Optical Properties. *Appl. Opt.* 52 (10), 2019–2037. doi:10.1364/AO.52.002019

Whitley, D., Starkweather, T., and Bogart, C. (1990). Genetic Algorithms and Neural Networks: Optimizing Connections and Connectivity. *Parallel Comput.* 14 (3), 347–361. doi:10.1016/0167-8191(90)90086-O

Wu, G., Li, J., Zhou, T., Lu, R., Yu, Y., Zhu, J., et al. (2006). The Key Region Affecting the Short-Term Climate Variations in China: the Joining Area of Asia and Indian-Pacific Ocean. *Adv. Earth Sci.* 21 (11), 1109–1118. doi:10.11867/j.issn.1001-8166.2006.11.1109

Wu, W., Qiu, Z., Zhao, M., Huang, Q., and Lei, Y. (2018). Visible and Infrared Image Fusion Using NSST and Deep Boltzmann Machine. *Optik* 157, 334–342. doi:10.1016/j.ijleo.2017.11.087

Xiao, C., Chen, N., Hu, C., Wang, K., Gong, J., and Chen, Z. (2019). Short and Mid-term Sea Surface Temperature Prediction Using Time-Series Satellite Data and LSTM-AdaBoost Combination Approach. *Remote Sensing Environ.* 233, 111358. doi:10.1016/j.rse.2019.111358

Yu, F., and Xu, X. (2014). A Short-Term Load Forecasting Model of Natural Gas Based on Optimized Genetic Algorithm and Improved BP Neural Network. *Appl. Energ.* 134, 102–113. doi:10.1016/j.apenergy.2014.07.104

Yue, L., Shen, H., Zhang, L., Zheng, X., Zhang, F., and Yuan, Q. (2017). High-quality Seamless DEM Generation Blending SRTM-1, ASTER GDEM V2 and ICESat/GLAS Observations. *ISPRS J. Photogrammetry Remote Sensing* 123, 20–34. doi:10.1016/j.isprsjprs.2016.11.002

Zare Abyaneh, H., Bayat Varkeshi, M., Golmohammadi, G., and Mohammadi, K. (2016). Soil Temperature Estimation Using an Artificial Neural Network and Co-active Neuro-Fuzzy Inference System in Two Different Climates. *Arab. J. Geosci.* 9 (5), 377. doi:10.1007/s12517-016-2388-8

Zhang, X., and Chen, N. (2016). Reconstruction of GF-1 Soil Moisture Observation Based on Satellite and *In Situ* Sensor Collaboration under Full Cloud Contamination. *IEEE Trans. Geosci. Remote Sensing* 54 (9), 5185–5202. doi:10.1109/TGRS.2016.2558109

Zhu, Y., Bo, Y., Zhang, J., and Wang, Y. (2018). Fusion of Multisensor SSTs Based on the Spatiotemporal Hierarchical Bayesian Model. *J. Atmos. Ocean. Technol.* 35 (1), 91–109. doi:10.1175/JTECH-D-17-0116.1

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.

frontiers in Ecology and Evolution

ORIGINAL RESEARCH
published: 01 November 2021
doi: 10.3389/fevo.2021.759732

Physiological Responses of *Sargassum muticum*, a Potential Golden Tide Species, to Different Levels of Light and Nitrogen

Fang Yan[1†], Lingxue Li[1†], Daode Yu[2], Cuiju Cui[3], Shasha Zang[1,4], Zhiguang Xu[1,4] and Hongyan Wu[1,4*]*

[1] College of Life Sciences, Ludong University, Yantai, China, [2] Marine Science Research Institute of Shandong Province, Qingdao, China, [3] School of Agriculture, Ludong University, Yantai, China, [4] Key Laboratory of Marine Biotechnology in Universities of Shandong, Ludong University, Yantai, China

OPEN ACCESS

Edited by:
Peng Liu,
Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences (CAS), China

Reviewed by:
Yaping Wu,
Jiangsu Ocean University, China
Wang Xulei,
Institute of Oceanology, Chinese Academy of Sciences (CAS), China
Sylwia Śliwińska-Wilczewska,
University of Gdańsk, Poland
Ping Li,
Shantou University, China
Gang Li,
South China Sea Institute of Oceanology, Chinese Academy of Sciences (CAS), China
Andrew J. Davies,
University of Rhode Island, United States

***Correspondence:**
Zhiguang Xu
bigwide@163.com
Hongyan Wu
sdwuhongyan@126.com

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

Specialty section:
This article was submitted to Environmental Informatics and Remote Sensing, a section of the journal Frontiers in Ecology and Evolution

Received: 17 August 2021
Accepted: 08 October 2021
Published: 01 November 2021

Citation:
Yan F, Li L, Yu D, Cui C, Zang S, Xu Z and Wu H (2021) Physiological Responses of Sargassum muticum, a Potential Golden Tide Species, to Different Levels of Light and Nitrogen. Front. Ecol. Evol. 9:759732. doi: 10.3389/fevo.2021.759732

Sargassum golden tides have bloomed frequently in many sea areas throughout the world, and negatively impacted on the local marine ecology. *Sargassum muticum* commonly inhabits rocky shores. It is now distributed worldwide due to its invasiveness, and recently drifting individuals have been observed on the coasts of Canary Islands. However, as a potential golden tide alga, physiological, and ecological studies of this species have not been frequently explored. To investigate the responses of *S. muticum* to light and nitrogen, two key environmental factors in golden tide formation, we established three light levels (LL, low light, 10 μmol photons m^{-2} s^{-1}; ML, medium light, 60 μmol photons m^{-2} s^{-1}, and HL, high light, 300 μmol photons m^{-2} s^{-1}) and two nitrogen levels (LN, low nitrogen, 25.0 μM of natural seawater; HN, high nitrogen, 125.0 μM), and cultivated the thalli under different conditions for 12 days before measuring the physiological properties of alga. The results showed that higher light and/or nitrogen levels enhanced the relative algal growth rate. The maximum net photosynthesis rate of alga increased with the light, while it remained unaffected by the nitrogen. The HN treatment had no effect on the apparent photosynthetic efficiency of algae in the LL culture, while increased it in the ML and HL cultures. The irradiance saturation point of photosynthesis was approximately 300 μmol photons m^{-2} s^{-1} with no significant difference among the six treatments, except for a slight increase under HLHN in contrast to the LLHN and MLLN treatments. HL treatment decreased the maximum quantum yield of photosynthesis (F_v/F_m) in both nitrogen levels. In the HN culture, ML and HL led to lower values of photoinhibition, indicating higher survivability in the alga. The HN culture led to higher nitrogen uptake but had no effects on F_v/F_m and the contents of pigments and soluble protein, regardless of culture light level. Based on these results, we speculate that drifting individuals of *S. muticum* would be possible to form a golden tide owing to its rapid growth rate at light level of 300 μmol photons m^{-2} s^{-1}, when they encountered the sustained lower light level on the sea surface (≤300 μmol photons m^{-2} s^{-1}). A high nitrogen supply caused by eutrophication of seawater might facilitate this process. Our results provide an important reference for the prediction of golden tides formed by *S. muticum*.

Keywords: floating macroalgae, eutrophication, photosynthesis, photoinhibition, growth

INTRODUCTION

In recent years, there have been massive inundations of the genus *Sargassum* aggregating on the surface of sea, known as "golden tides," in the central Atlantic Ocean, Gulf of Mexico (Smetacek and Zingone, 2013; Schell et al., 2015; Wang et al., 2019; Qi et al., 2020), East China Sea and Yellow Sea (Qi et al., 2017; Xing et al., 2017; Liu et al., 2018). When the golden tide blooms as an invasive genus, mats of floating *Sargassum* can change the ecosystem structure of the original sea area (van Tussenbroek et al., 2017; Baker et al., 2018; Caselle et al., 2018), negatively impact offshore tourism and maritime transport (Williams and Feagin, 2010; Milledge and Harvey, 2016), and lead to the deterioration of seawater quality after they sink and decay (Cruz-Rivera et al., 2015). If an episode was to occur in aquaculture areas, golden tides could cause enormous losses to the local mariculture industry, for example, blooming of *Sargassum horneri* golden tide in the Yellow Sea in December 2016, caused economic loss of 0.5 billion CNY due to the damaged local seaweed aquaculture (Xing et al., 2017).

The occurrence of golden tides has been linked to changes in marine environmental factors, such as nutrients, temperature, and light (Sfriso and Facca, 2013; Smetacek and Zingone, 2013). Satellite remote sensing monitoring data have confirmed that the formation of golden tides often occurs in eutrophic oceanic areas with the availability of upwelled (Brooks et al., 2018; Wang et al., 2019) and river (Sfriso and Facca, 2013; Louime et al., 2017; van Tussenbroek et al., 2017) plume nutrients, which provide sufficient nitrogen and phosphorus for the growth of golden tide algae. The fact that high nitrogen resulted in high growth rate has been found in in several algal species known to cause golden tides, namely, *S. horneri* (Yu et al., 2019), *Sargassum natans* and *Sargassum fluitans* (Wang et al., 2019). Therefore, many scientists hypothesize that eutrophication plays an important role in inducing the formation of *Sargassum* golden tides (Sfriso and Facca, 2013; Smetacek and Zingone, 2013; Louime et al., 2017; van Tussenbroek et al., 2017; Wang et al., 2019).

Alternatively, golden tide algae drifting on the sea surface receive more light, compared with those individuals attached to the seabed, owing to the light attenuated by seawater (Kouassi and Zika, 1992; Tedetti and Sempéré, 2006; Wu et al., 2010). Typically, the majority of macroalgae manifest higher photosynthesis and growth rates at higher light levels (Xu et al., 2014; Erin Cox and Smith, 2015). However, levels of light that are too high could depress macroalgal photosynthesis, which is designated photoinhibition (Yakovleva and Titlyanov, 2001; Murata et al., 2007). However, that floating *Sargassum* receiving more light energy, may grow explosively to form a golden tide when exposed to nutrient enrichment, for example *S. natans* and *S. fluitans* (Wang et al., 2019). Moreover, the upper and lower sides of the floating *Sargassum* layer will also suffer from differing light conditions owing to the thickness of the floating layer. Unfortunately, little is known about the effects of light levels on the photosynthesis related to growth of floating *Sargassum*, particularly when enriched by nutrients.

Currently, three species have been observed to form golden tides. *S. natans* and *S. fluitans* are the predominant species in golden tides that bloom in the central Atlantic Ocean and the Gulf of Mexico (Smetacek and Zingone, 2013; Schell et al., 2015; van Tussenbroek et al., 2017; Wang et al., 2019), while *S. horneri* is the only species that causes golden tides in the East China Sea and Yellow Sea (Qi et al., 2017; Xing et al., 2017; Liu et al., 2018). Although no other *Sargassum* species have been currently observed to form golden tides, whether they will do so in the case of appropriate environmental conditions in the future remains unclear. *Sargassum muticum* is widely distributed on the coast of the Western Pacific North and has recently invaded the Northern European and North American coasts (Engelen et al., 2015), resulting in significant negative impacts on local marine ecosystems (Cacabelos et al., 2013; Salvaterra et al., 2013; Sánchez and Fernández, 2018; Belattmania et al., 2020). Most seriously, many drifting alga individuals were observed on the coasts of Canary Islands (Eastern Atlantic Ocean) in March 2020 (Álvarez-Canali et al., 2021). Therefore, we suggest that more intensive study on this potential golden tide species is merited. The aim of this study was to investigate the physiological responses of *S. muticum* to light and nitrogen levels, two key factors during the formation of golden tides, which expectedly provides references for the future prediction of golden tides caused by this species.

MATERIALS AND METHODS

Materials

Samples of *S. muticum* were collected from an attached population at a depth of 2 m in Lidao Bay, Rongcheng City, Shandong Province, China (37°15′ N, 122°35′ E) on April 5, 2019, where temperature is 20°C, salinity is 30 PSU, NO_3^-, 25.0 μM and dissolved inorganic phosphorus (DIP) is 1.6 μM in the natural seawater. Sporophytes of approximately 60 cm long were selected and brought to the laboratory in an insulated cooler of 4°C in 2 h. After cleaned with autoclaved seawater, segments that were approximately 5 cm long were randomly cut from several branches and pre-cultured in autoclaved natural seawater (salinity, 30 PSU; NO_3^-, 25.0 μM; NH_4^+, 0.8 μM and DIP, 1.6 μM) for 24 h. The temperature was set at 20°C, and the light level was 60 μmol photons m^{-2} s^{-1} with light and dark periods of 12 h:12 h. The media were continuously aerated.

Experimental Design

After preincubation for 24 h, 5–6 segments [∼4.0 g of fresh weight (FW)] of the algae were cultured in a conical flask with 2 L of modified Provasoli's enriched seawater (PESI, Provasoli, 1968). The DIP concentration of 5 μM in PESI medium was obtained by the addition of NaH_2PO_4 to ensure adequate supply of P. In addition, $NaNO_3$ was added separately for the high nitrogen levels. According to the observation that inorganic

nitrogen concentration in golden tide bloom area increased by 3–4 times (Wang et al., 2019), two nitrogen levels were established in the experiment: low nitrogen (LN, the concentration in natural seawater, 25.0 μM) and high nitrogen (HN, 125.0 μM). The effect of NH_4^+ was ignored due to its low content of 0.8 μM in both LN and HN treatments. The media were continuously aerated and renewed every other day. The culture density was maintained at 2.0 g FW L^{-1} during the experiment by removing the excess growth part of the thalli when the media was renewed.

The containers were positioned in growth chambers (MGC-250P, Yiheng Technical Co., Ltd., Shanghai, China) with constant temperature of 20°C and 12 h:12 h light and dark periods. By regulating the light source of growth chambers, light that reached the algae was set at three levels (LL, low light, 10 μmol photons m^{-2} s^{-1}; ML, middle light, 60 μmol photons m^{-2} s^{-1}; HL, high light, 300 μmol photons m^{-2} s^{-1}),based on the daily mean irradiance in spring received by the attached thalli [the lowest and highest values in the distribution depths of 2−6 m, referred to Liu et al. (2021)] and the floating individuals (Rumyantseva et al., 2019), respectively. Light source was provided by LED lamps emitting white light in chambers, and the light levels were measured using a handheld optical quantum meter (QRT1, Hansatech, Norfolk, United Kingdom). The samples were cultured under different nitrogen and light conditions for 12 days, with three replicates of each treatment. At the end of 12-day culture, the growth rate, photosynthesis, contents of pigments and soluble protein, and nitrogen uptake of the algae were determined.

Determination of the Growth Rate

The growth of algae under different nitrogen and light conditions was measured during the final 2 days of culture in order to reflect acclimated physiological performance of thalli, and the relative growth rate (RGR) was calculated using the following formula:

$$RGR = 100 \times (lnN_t - lnN_0)/t$$

Where N_t represents the FW on day t (day 12); N_o represents the initial FW (day 10), and t represents the time interval (2 days). The FW was determined through electronic analytical balance (ME104, Mettler-Toledo, Switzerland) after removing surface water drops of thalli by lightly blotting with tissue paper.

Estimates of Respiration and Photosynthesis

Dark respiration (R_d) and the rates of evolution of photosynthetic O_2 were determined using a Clarktype oxygen electrode (Chlorolab-3, Hansatech, Norfolk, United Kingdom) at 20°C. Approximately 0.15 g FW of algae harvested from the culture flask on day 12 was transferred to the oxygen electrode cuvette that contained 8 ml of autoclaved natural seawater, and the media were stirred during measurement. The dark respiration rate was measured after covering the cuvette with an opaque cloth. The photosynthetic O_2 evolution rates (E_k) were measured during 5 min under six levels of light (20, 60, 120, 200, 400, and 600 μmol photons m^{-2} s^{-1}). White light was provided by a metal halide lamp, and different levels of irradiance were provided by altering the distance between light source and the oxygen electrode cuvette. The light levels were measured through a handheld optical quantum meter (QRT1, Hansatech, Norfolk, United Kingdom). Therefore, the P−E curve (photosynthetic rates at different irradiance levels) was obtained using the model of Eilers and Peeters (1988) and fitting the data by iteration:

$$P = E/(a \times E^2 + b \times E + c)$$

Where P are the net photosynthesis rates (μmol O_2 h^{-1} g^{-1} FW) at different irradiance levels; E is the irradiance (μmol photons m^{-2} s^{-1}), and a, b, and c are the adjustment parameters. The parameter "a" is the photoinhibition term. The apparent photosynthetic efficiency (α), maximum net photosynthetic rate (P_{max}) and irradiance saturation point (E_k) are expressed as a function of the parameters a, b, and c as follows:

$$\alpha = 1/c$$

$$P_{\max} = 1/[b+2(a \times c)^{0.5}]$$

$$E_k = (c/a)^{0.5}$$

Determination of the Maximum Photochemical Yield of Photosystem II

The chlorophyll fluorescence yield data were measured using a plant efficiency analyzer (Handy PEA, Hansatech, Norfolk, United Kingdom) during the 6th hour of the photoperiod at end of 12 days of culture. The thalli were dark adapted for 15 min to allow all the reaction centers to open before the measurements were taken. The minimal fluorescence (F_0) was induced by low irradiation of the measuring light (approximately 0.15 μmol photons m^{-2} s^{-1}), and the maximum fluorescence (F_m) was obtained after exposed to a single saturating pulse irradiance of 5,000 μmol photons m^{-2} s^{-1} for 800 ms. The maximum quantum yield of photosystem II (PSII) was then estimated as:

$$F_v/F_m = (F_m - F_0)/F_m$$

Measurement of Pigments and Soluble Proteins

Approximately 0.1 g FW thalli from each treatment were ground with an aqueous solution of 90% acetone. The solution was brought to 8 ml and extracted for 12 h at 4°C in the dark. After centrifugation (4,000 r min^{-1}, 15 min), the absorption spectrum of 400–700 nm of supernatant was determined using a spectrophotometer (DU 530, Beckman Coulter, Fullerton, CA, United States). The contents (mg g^{-1} FW) of chlorophyll (Chl a, Chl c) and carotenoids (Car) were calculated as described by Jeffrey and Humphrey (1975) and Parsons et al. (1984), respectively.

$$\text{Chl a} = (11.47 \times A_{664} - 0.40 \times A_{630}) \times V/M$$

$$\text{Chl c} = (24.36 \times A_{630} - 3.73 \times A_{664}) \times \text{V/M}$$

$$\text{Car} = [7.6 \times (A_{480} - 1.49 \times A_{510})] \times \text{V/M}$$

Where A was the absorbance value of supernatant at differing wavelengths; V was the volume of solution at a constant volume (8 mL), and M was the FW of algae (g).

To extract the soluble protein, approximately 0.2 g of FW thalli from each treatment were ground in a mortar, and the volume was brought to 5 mL with phosphate buffer (0.1 mol L^{-1}, pH 6.8). After centrifugation at 5,000 *g* for 10 min, the soluble protein was estimated from the supernatant using the Bradford (1976) assay with bovine serum albumin as the standard.

Assessment of Nitrogen Uptake

The rate of nitrogen uptake of the thalli was estimated from the decrease in rate of nitrate in the culture seawater over a 2-day interval when the medium was being renewed. The concentration of nitrate was determined as described by Strickland and Parsons (1972), and the rate of nitrogen uptake was calculated using the following equation:

$$\text{Uptake rate} = (N_0 - N_t) \times \text{V/M/t}$$

Where N_0 and N_t were the concentrations of nitrate in the medium at the initial and end of culture, respectively. V was the volume of the culture medium; M was the initial FW of the thalli (g), and t was the culture time (2 days).

Data Analysis

The results were expressed as the means of replicates ± standard deviation. The data were analyzed using SPSS v. 21 (IBM, Inc., Armonk, NY, United States). The data under every treatment conformed to a normal distribution (Shapiro–Wilk, $P > 0.05$), and the variances can be considered equal (Levene's test, $P > 0.05$). Two-way analysis of variance (ANOVA) was conducted to assess the effects of light and N on RGR, photosynthetic parameters from the P–E curves, dark respiration rate, F_v/F_m, pigments, soluble protein, and nitrate uptake rate. Tukey's honest significance difference (HSD) was conducted for a *post hoc* investigation. A confidence interval of 95% was set for all tests.

RESULTS

Growth Rate

The RGRs of *S. muticum* cultured under different light and nitrogen conditions are shown in **Figure 1**. Light and nitrogen had no interactive effect on RGR of *S. muticum* (ANOVA, $P = 0.088$), but each factor had a main effect (ANOVA, $P < 0.001$ for light; ANOVA, $P < 0.001$ for nitrogen). An increase in the light intensity enhanced the RGR at both LN andHN treatments (Tukey's HSD, $P < 0.05$). Compared with LL, treatment with ML and HL increased the RGR of algae by 55.9 and 190.6% at LN levels, respectively, while under HN condition, the RGR were enhanced by 60.3% under ML and 168.9% under HL treatments. The enrichment of nitrogen also significantly promoted the growth of thalli under all three light conditions (Tukey's HSD, $P < 0.05$). HN increased the RGR by 11.7, 14.8, and 3.4% in LL, ML, and HL treatments, respectively, compared with LN conditions.

Photosynthesis and Respiration

The irradiance saturation point (E_k), photoinhibition term (a), apparent photosynthetic efficiency (α), and maximum net photosynthesis rate (P_{max}) are presented in **Table 1**. The irradiance saturation point (E_k) was approximately 300 photons m^{-2} s^{-1}, with no significant difference between the six treatments (Tukey's HSD, $P > 0.05$), with the exception of a slight increase under HLHN in contrast to the LLHN and MLLN treatments (Tukey's HSD, $P < 0.05$). Light and nitrogen had an interactive effect on P_{max} (ANOVA, $P < 0.001$), α (ANOVA, $P < 0.001$) and the photoinhibition term "a" (ANOVA, $P < 0.05$). In all six treatments of different light and nitrogen levels, the thalli cultured in MLHN and HLHN expressed lower "*a*" values (Tukey's HSD, $P < 0.05$), suggesting that algae that had adapted to HN and HL conditions were more tolerant of strong light inhibition. Regardless of the nitrogen concentration, P_{max} increased with the light level of culture, with the exception of an insignificant difference between MLHN and HLHN (Tukey's HSD, $P > 0.05$). At the ML and HL levels, the nitrogen level insignificantly affected P_{max} (Tukey's HSD, $P > 0.05$), while in LL treatments, HN enhanced the P_{max} of thalli by 28.3%, compared with LN (Tukey's HSD, $P < 0.05$). In the LN culture, the light level had no effect on the α of algae (Tukey's HSD, $P > 0.05$). However, in HN culture, the value of α differed significantly between the three light levels (Tukey's HSD, $P < 0.05$). The highest (0.169 ± 0.005) and lowest (0.105 ± 0.012) values of α occurred at ML and LL, respectively.

In the LL and ML culture, the dark respiration rate (R_d) of thalli remained unaffected interactively by different levels of culture light and nitrogen (ANOVA, $P > 0.05$, **Figure 2**). However, HL culture (300 μmol photons m^{-2} s^{-1}) remarkably increased R_d in both nitrogen concentrations, compared with the LL and ML treatments (Tukey's HSD, $P < 0.05$). In addition, in HL culture, HN enhanced the R_d by 24.6%, in contrast to that in LN culture (Tukey's HSD, $P < 0.05$).

The Maximum Photochemical Yield

The maximum photochemical yields (F_v/F_m) of the algae cultured at different light and nitrogen levels are shown in **Figure 3**. Light and nitrogen had no interactive effect on F_v/F_m of algae (ANOVA, $P > 0.05$). The nitrogen levels had an insignificant effect on F_v/F_m regardless of the culture light levels (ANOVA, $P > 0.05$). In both LN and HN cultures, the values of F_v/F_m expressed insignificant differences between the LL and ML conditions (Tukey's HSD, $P > 0.05$), while it decreased remarkably in the HL treatment

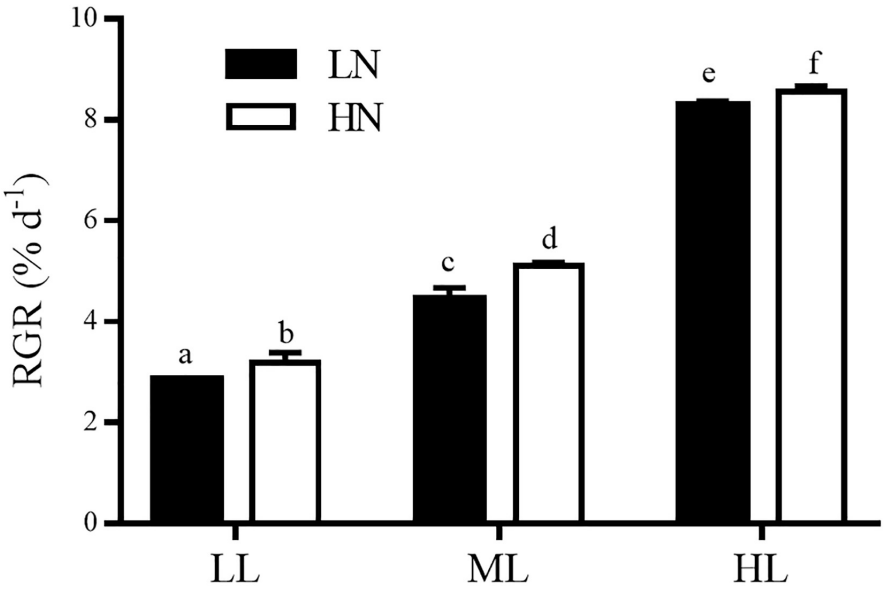

FIGURE 1 | Relative growth rates (RGRs) in *Sargassum muticum* grown at different light and nitrogen levels, which were determined during 2 days at the end of 12 days of incubation. Significant ($P < 0.05$) differences among the treatments are indicated by different lowercase letters. Vertical bars represent ± SD for the means of three samples.

TABLE 1 | Photosynthetic parameters from the *P–E* curves of *Sargassum muticum* cultured under different light and nitrogen conditions.

	a	P_{max}	α	E_k
LLLN	$(1.23 \pm 0.11) \times 10^{-4a}$	84.11 ± 8.44^{a}	0.085 ± 0.010^{a}	310.48 ± 4.28^{ab}
LLHN	$(1.09 \pm 0.15) \times 10^{-4a}$	107.91 ± 3.40^{b}	0.105 ± 0.012^{a}	297.88 ± 5.17^{a}
MLLN	$(1.23 \pm 0.04) \times 10^{-4a}$	123.91 ± 3.89^{c}	0.093 ± 0.003^{a}	295.53 ± 1.00^{a}
MLHN	$(0.65 \pm 0.01) \times 10^{-4b}$	135.43 ± 3.80^{cd}	0.169 ± 0.005^{b}	303.38 ± 4.69^{ab}
HLLN	$(1.17 \pm 0.11) \times 10^{-4ac}$	152.73 ± 3.32^{e}	0.107 ± 0.007^{a}	309.92 ± 10.76^{ab}
HLHN	$(0.80 \pm 0.14) \times 10^{-4bc}$	146.52 ± 2.46^{de}	0.129 ± 0.008^{c}	314.67 ± 2.00^{b}

a, the photoinhibition term; P_{max} (μmol O_2 h^{-1} g^{-1} FW), the maximum net photosynthesis rate; α [(μmol O_2 h^{-1} g^{-1} FW)/(μmol photons m^{-2} s^{-1})], the apparent photosynthetic efficiency; E_k (μmol photons m^{-2} s^{-1}), the irradiance saturation point. Different superscript letters indicate significant differences in one parameter between the treatments ($P < 0.05$).

compared with the LL and ML treatments (Tukey's HSD, $P < 0.05$).

Photosynthetic Pigments

The contents of photosynthetic pigments (Chl a, Chl c, and Car) of the thalli that had adapted to different light and nitrogen conditions are shown in **Figure 4**. Light and nitrogen had no interactive effect on the contents of all three pigments (ANOVA, $P = 0.671$ for Chl a, $P = 0.419$ for Chl c, and $P = 0.512$ for Car). Regardless of the culture light levels, the contents of Chl a, Chl c, and Car were not affected by the nitrogen level in cultures (ANOVA, $P = 0.057$ for Chl a, $P = 0.417$ for Chl c, and $P = 0.148$ for Car). However, light had main effects on three pigments (ANOVA, $P < 0.001$ for Chl a, $P < 0.001$ for Chl c, and $P < 0.01$ for Car). In the LN or HN treatment, HL treatment reduced the contents of all three photosynthetic pigments by approximately 30−45% compared with the LL and ML conditions (Tukey's HSD, $P < 0.05$), while no significant difference was observed between the LL and ML culture (Tukey's HSD, $P > 0.05$).

Soluble Proteins and Nitrogen Uptake

Light and nitrogen had no interactive effect on the content of soluble proteins (ANOVA, $P > 0.05$, **Figure 5**). The culture nitrogen levels had no impact on the content of soluble proteins at LL, ML, or HL treatment (ANOVA, $P > 0.05$). However, light had a main effect on it (ANOVA, $P < 0.001$). The LL culture markedly enhanced the content of soluble protein in thalli grown under LN or HN conditions (Tukey's HSD, $P < 0.05$), despite no significant difference was observed between the ML and HL cultures (Tukey's HSD, $P > 0.05$).

Light and nitrogen had an interactive effect on the nitrogen uptake rates of algae (ANOVA, $P < 0.001$, **Figure 6**), and each factor had a main effect (ANOVA, $P < 0.001$ for light; ANOVA, $P < 0.001$ for nitrogen). Compared with the LN treatment, HN culture increased the rate of uptake of nitrogen by 7.6 times at low light, 1.3 times in ML treatment, and 0.8 times in HL treatment. In LN cultures, ML and HL improved the rate of nitrogen uptake of algae by 3.0 and 8.5 times, respectively, compared with the

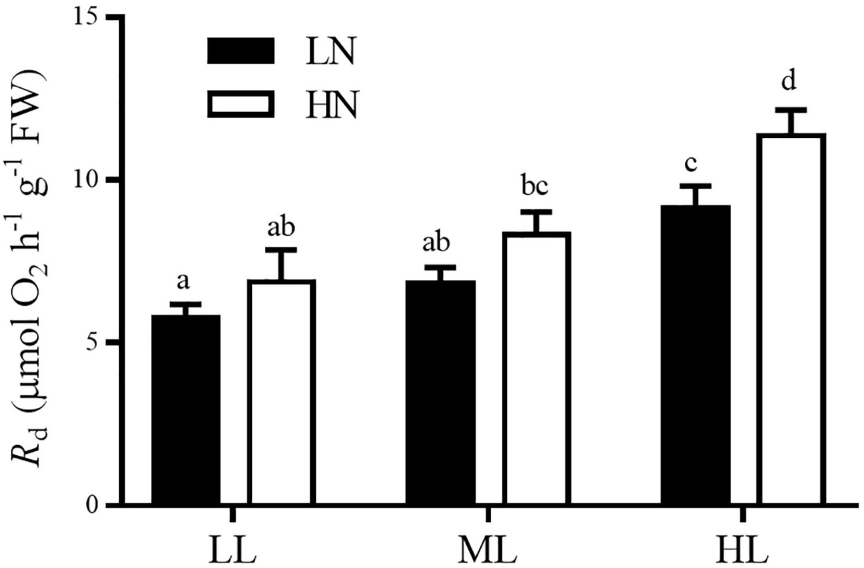

FIGURE 2 | Dark respiration rates of *S. muticum* grown at different light and nitrogen levels for 12 days. Significant ($P < 0.05$) differences among the treatments are indicated by different lowercase letters. Vertical bars represent ± SD for the means of three samples.

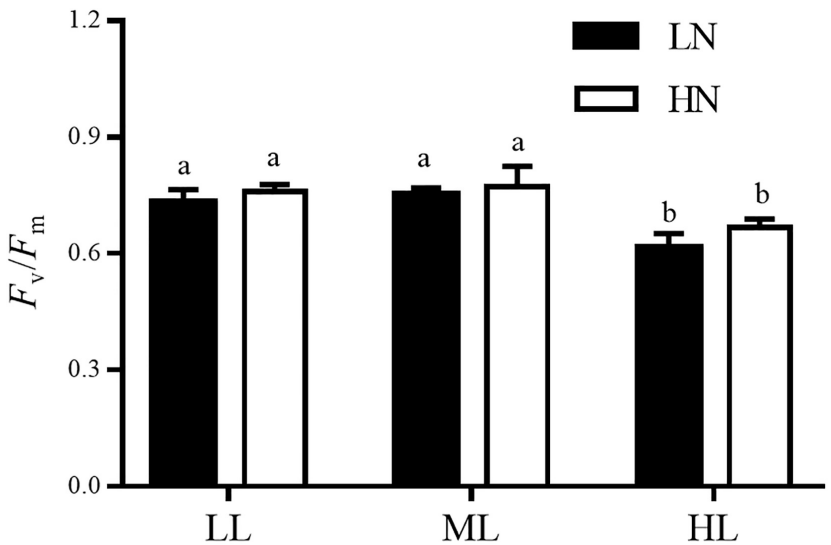

FIGURE 3 | The maximum quantum yield of photosystem II (F_v/F_m) in *S. muticum* grown at different light and nitrogen levels for 12 days. Significant ($P < 0.05$) differences among the treatments are indicated by different lowercase letters. Vertical bars represent ± SD for the means of three samples.

LL culture, while ML and HL improved it by 9.0 and 96.0%, respectively, when cultured in HN treatment.

DISCUSSION

In this study, higher light culture led to a higher RGR in *S. muticum*. The growth of macroalgae is closely related to algal photosynthesis (Lobban et al., 1985; Xu and Gao, 2009). Light provides the initial energy for photosynthesis, and the increase in culture light level can usually promote the photosynthesis of macroalgae by affecting electron transfer, pigment synthesis, and/or the activity of enzymes associated with carbon utilization (Xu et al., 2014; Wu et al., 2018; Dudgeon and Kübler, 2020; Eismann et al., 2020). In this study, a higher culture light level induced a greater maximum net photosynthesis rate, which indicated a higher photosynthetic capacity, and ultimately contributed to the higher growth rate of *S. muticum*.

The apparent photosynthetic efficiency (α) is regarded as the indication of efficiency of the alga to utilize light energy under limited light intensity (Eilers and Peeters, 1988). In this study, α showed the highest value in middle light culture of

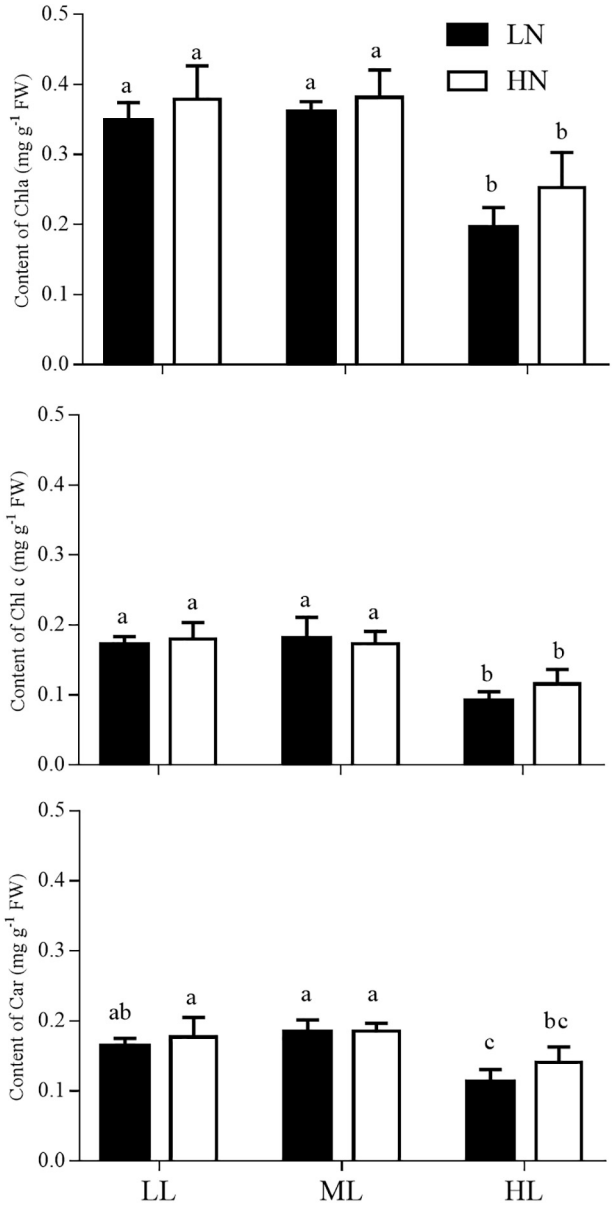

FIGURE 4 | Contents of Chl a, Chl c and Car in *S. muticum* grown at different light and nitrogen levels for 12 days. Significant ($P < 0.05$) differences among the treatments are indicated by different lowercase letters. Vertical bars represent ± SD for the means of three samples.

60 μmol photons m^{-2} s^{-1} and comparatively decreased in high light culture of 300 μmol photons m^{-2} s^{-1}. Moreover, the value of F_v/F_m also decreased in high light culture, suggesting that *S. muticum* thalli suffered from light stress in the high light culture of 300 μmol photons m^{-2} s^{-1}. Such findings may be explained by the reduction of photosynthetic pigments in the high light culture compared with low and middle light culture. Algae can alter the contents of reaction center and light-harvesting pigments to adapt to different culture light intensities (Makarov, 2012). Previous studies have shown that high light intensity may reduce the photosynthetic pigment yields of algae in minutes [reviewed by Eismann et al. (2020)]. Photopigments, particularly light-harvesting pigments, usually play an important role in photoprotection and are damaged first under the high light stress, resulting in the reduction of contents (Xu and Gao, 2012; Quintano et al., 2019). However, in this study, the decrease induced by high light was not found in the growth rate and P_{max} of *S. muticum*. The possible explanation could be that the light stress at 300 μmol photons m^{-2} s^{-1} did not damage the reaction center protein owing to the protection by its pigments, or the improved repair rate was enough to alleviate such photodamage to the reaction center (Murata et al., 2007; Zhang et al., 2020). As indirect evidence, the elevated dark respiration rate in the high light culture at 300 μmol photons m^{-2} s^{-1} is likely to provide sufficient energy to repair the damage of photosynthetic reaction center in *S. muticum* (Xu and Gao, 2009, 2012).

Excessive light energy generally leads to the damage of photosynthesis reaction center of algae, and such a deleterious effect is often counteracted by the protective strategies of algae, such as the accumulation of photoprotective compounds, repair mechanisms, and the removal of reactive oxygen species (Bonomi-Barufi et al., 2020; Zhang et al., 2020). In the case that damage exceeds protection, the algae will reduce their evolution of photosynthetic oxygen, which is designated "photoinhibition" (Murata et al., 2007). In this study, the irradiance saturation point of oxygen E_k was ∼300 μmol photons m^{-2} s^{-1} with no significant difference in all six treatments of different light and nitrogen levels. Beyond this irradiance point, the rate of evolution of photosynthetic oxygen of algae decreased drastically with the increase in light intensity, i.e., photoinhibition was induced in *S. muticum*. In contrast, many species of macroalgae have light saturation points at ∼600 μmol photons m^{-2} s^{-1} (Xu and Gao, 2012; Xu et al., 2014). Especially, *S. horneri*, a golden tide species, did not exhibit photoinhibition even at a light level of 600 μmol photons m^{-2} s^{-1}, and it maintained enhanced photosynthetic and growth rates with the increase in light intensity (Liu et al., 2019). In the north temperate zone, which is inhabited by *S. muticum* and *S. horneri*, the daily mean irradiance of sea surface is approximately 300–800 μmol photons m^{-2} s^{-1} during the spring, summer, and autumn (Rumyantseva et al., 2019), the seasons in which golden tides have been observed. At such light levels, compared with *S. horneri*, drifting *S. muticum* is not prone to form a golden tide owing to the induction of photoinhibition.

Nitrogen is an essential nutrient for the survival and growth of algae, and an increased supply of nitrogen generally results in the promotion of growth rate in algae, owing to enhanced photosynthesis (Gao et al., 2018; Liu et al., 2019; Jiang et al., 2020). Such enhanced growth rates by N were also found in three golden tide species: *S. horneri* (Yu et al., 2019), *S. natans* and *S. fluitans* (Wang et al., 2019). The fact that enhanced photosynthesis induced by high nitrogen is usually attributed to an enhanced electron transfer rate, pigment content, or enzyme activity associated with carbon fixation (e.g., RUBISCO and carbonic anhydrase among others) in the photosynthetic system of algae (Giordano et al., 2005; Xu et al., 2014; Gao et al., 2018; Liu et al., 2019; Rugiu et al., 2020). In this study, high nitrogen enhanced the growth rate of *S. muticum* cultured at

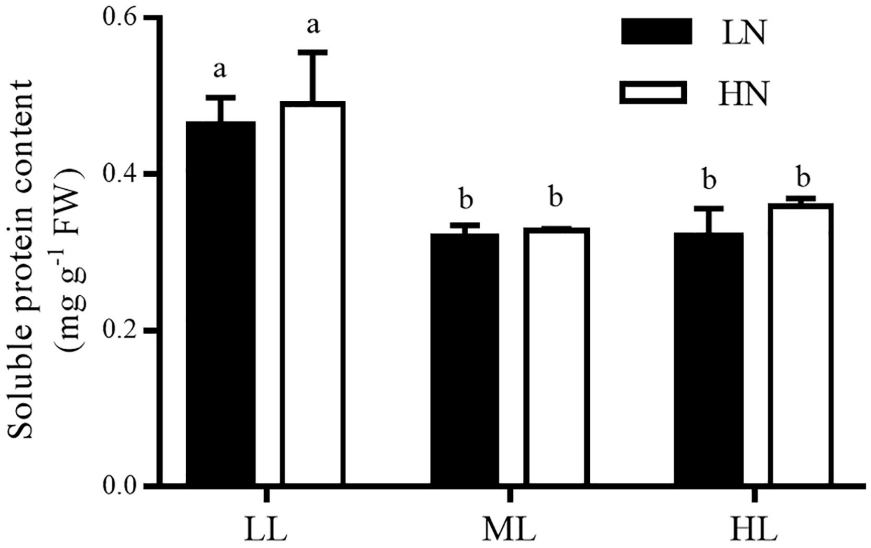

FIGURE 5 | Content of soluble proteins in *S. muticum* grown at different light and nitrogen levels for 12 days. Significant ($P < 0.05$) differences among the treatments are indicated by different lowercase letters. Vertical bars represent ± SD for the means of three samples.

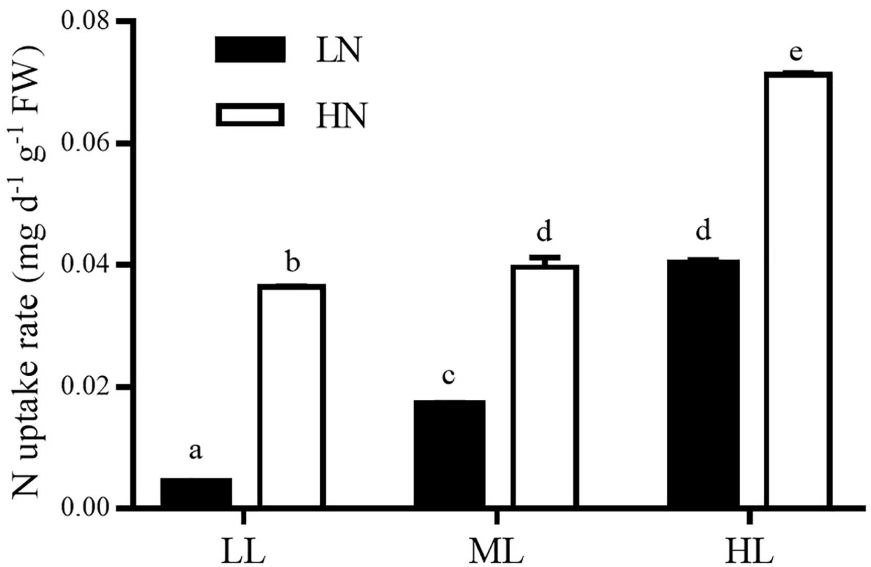

FIGURE 6 | Nitrogen uptake rates of *S. muticum* grown at different light and nitrogen levels, which were determined during 2 days at the end of 12 days of incubation. Significant ($P < 0.05$) differences among the treatments are indicated by different lowercase letters. Vertical bars represent ± SD for the means of three samples.

three light levels. Correspondingly, the photosynthesis of thalli was increased by culture in high nitrogen, reflected by the increased "P_{max}" at low light and increased "α" at middle and high light levels. Such findings were consistent with previous reports in other species of macroalgae, including *Ulva prolifera*, *S. horneri*, *Gracilariopsis lemaneiformis*, *Saccharina latissima*, etc., (Xu et al., 2014; Liu et al., 2019; Yu et al., 2019; Jiang et al., 2020; Rugiu et al., 2020). Moreover, in the case of an adequate supply of nitrogen, the macroalgae can rapidly absorb it to supply the demand for nitrogen for growth and other physiological processes (Xu and Gao, 2012; Jiang et al., 2020), and excessive nitrogen may be stored in cells as nitrogen pools in the forms of pigments, soluble proteins, and/or amino acids (Teichberg et al., 2010; Yu et al., 2019; Jiang et al., 2020). In this study, a high supply of nitrogen induced a high rate of nitrogen uptake in *S. muticum*, while neither the contents of pigments nor soluble proteins were affected by the high nitrogen treatment. Therefore, we hypothesized that the greater amounts of nitrogen absorbed in the high nitrogen culture contributed to the higher rates of photosynthesis and growth in *S. muticum*. In addition, a higher

acquisition of nitrogen may be used to improve the capability of the algae to endure high light stress (Murata et al., 2007; Xu and Gao, 2012), which was reflected by the lower "*a*" value of the photoinhibition term in *S. muticum* cultured under high nitrogen conditions.

CONCLUSION

Sargassum muticum is naturally distributed in the rocky shores of the Western Pacific North and has been colonized to the Northern European and North American coasts as an invasive species (Sánchez and Fernández, 2018; Belattmania et al., 2020). Owing to the concern that this species could form a golden tide (Álvarez-Canali et al., 2021), we investigated its physiological responses to different light and nitrogen levels in this study. Our results showed that higher light and/or nitrogen levels generally promoted the photosynthesis and growth of the algae. However, when the intensity of light exceeded 300 μmol photons m^{-2} s^{-1}, its photosynthesis was inhibited, regardless of nitrogen conditions set in this study. Based on these findings, we hypothesized that the lower value of irradiance saturation point resulted in drifting individuals of *S. muticum* that were unable to tolerate the light intensity of sea surface on sunny days (Rumyantseva et al., 2019). Therefore, for this *Sargassum* species, it would not easily form a golden tide, compared with *S. horneri*. However, when encountered a sustained suitable light level (≤300 μmol photons m^{-2} s^{-1}), it would also be possible for the drifting individuals of *S. muticum* to form a golden tide due to its rapid growth rate (∼8% d^{-1} at light level of 300 μmol photons m^{-2} s^{-1}). And a high supply of nitrogen might facilitate this process.

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

FY and LL conceived and designed the experiment with help from ZX and HW. FY carried out the experiment, with help from LL and DY. FY and LL performed the data analyses, with help from DY, CC, and SZ. FY wrote the manuscript with contributions from LL, DY, CC, SZ, ZX, and HW. All authors reviewed and gave their approval for the final manuscript.

FUNDING

This work was supported by the Shandong Provincial Natural Science Foundation of China to ZX (ZR2020MD092), HW (ZR2019MC015), and SZ (ZR2020QC025), and grants from the open project of Rongcheng Marine Industrial Technology Research Institute, the Ludong University awarded to ZX (KF20180001).

REFERENCES

Álvarez-Canali, D., Sangil, C., and Sansón, M. (2021). Fertile drifting individuals of the invasive alien *Sargassum muticum* (Fucales, Phaeophyceae) reach the coasts of the Canary Islands (eastern Atlantic Ocean). *Aquat. Bot.* 168:103322. doi: 10.1016/j.aquabot.2020.103322

Baker, P., Minzlaff, U., Schoenle, A., Schwabe, E., Hohlfeld, M., Jeuck, A., et al. (2018). Potential contribution of surface-dwelling *Sargassum* algae to deep-sea ecosystems in the southern North Atlantic. *Deep-Sea Res.* 148, 21–34. doi: 10.1016/j.dsr2.2017.10.002

Belattmania, Z., Chaouti, A., Engelen, A. H., Serrao, E. A., and Sabour, B. (2020). Spatiotemporal variation of the epifaunal assemblages associated to *Sargassum muticum* on the NW Atlantic coast of Morocco. *Environ. Sci. Pollut. R.* 27, 35501–35514. doi: 10.1007/s11356-020-09851-5

Bonomi-Barufi, J., Figueroa, F. L., Korbee, N., Momoli, M. M., and Oliveira, M. C. (2020). How macroalgae can deal with radiation variability and photoacclimation capacity: the example of *Gracilaria tenuistipitata* (Rhodophyta) in laboratory. *Algal Res.* 50:102007. doi: 10.1016/j.algal.2020.102007

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72, 248–254. doi: 10.1016/0003-2697(76)90527-3

Brooks, M. T., Coles, V. J., Hood, R. R., and Gower, J. F. R. (2018). Factors controlling the seasonal distribution of pelagic *Sargassum*. *Mar. Ecol. Prog. Ser.* 599, 1–18. doi: 10.3354/meps12646

Cacabelos, E., Olabarria, C., Viejo, R. M., Rubal, M., Veiga, P., Incera, M., et al. (2013). Invasion of *Sargassum muticum* in intertidal rockpools: patterns along the Atlantic Iberian Peninsula. *Mar. Environ. Res.* 90, 18–26. doi: 10.1016/j.marenvres.2013.05.008

Caselle, J. E., Davis, K., and Marks, L. M. (2018). Marine management affects the invasion success of a nonnative species in a temperate reef system in California, USA. *Ecol. Lett.* 21, 43–53. doi: 10.1111/ele.12869

Cruz-Rivera, E., Flores-Díaz, M. C., and Hawkins, A. (2015). A fish kill coincident with dense *Sargassum* accumulation in a tropical bay. *B. Mar. Sci.* 91, 455–456. doi: 10.5343/bms.2015.1048

Dudgeon, S., and Kübler, J. E. (2020). A multistressor model of carbon acquisition regulation for macroalgae in a changing climate. *Limnol. Oceanogr.* 65, 2541–2555. doi: 10.1002/lno.11470

Eilers, P. H. C., and Peeters, J. C. H. (1988). A model for the relationship between light intensity and the rate of photosynthesis in phytoplankton. *Ecol. Model.* 42, 199–215. doi: 10.1016/0304-3800(88)90057-9

Eismann, A. I., Reis, R. P., da Silva, A. F., and Cavalcanti, D. N. (2020). Ulva spp. carotenoids: responses to environmental conditions. *Algal Res.* 48:101916. doi: 10.1016/j.algal.2020.101916

Engelen, A. H., Serebryakova, A., Ang, P., Britton-Simmons, K., Mineur, F., Pedersen, M. F., et al. (2015). Circumglobal invasion by the brown seaweed *Sargassum muticum*. *Oceanogr. Mar. Biol.* 53, 81–126. doi: 10.1201/b18733-4

Erin Cox, T., and Smith, C. M. (2015). Photosynthetic rapid light curves for *Padina sanctaecrucis* vary with irradiance, aerial exposure, and tides in Hawaii's micro-intertidal zones. *Mar. Biol.* 162, 1061–1076. doi: 10.1007/s00227-015-2649-1

Gao, G., Xia, J. R., Yu, J. L., and Zeng, X. P. (2018). Physiological response of a red tide alga (*Skeletonema costatum*) to nitrate enrichment, with special reference to inorganic carbon acquisition. *Mar. Environ. Res.* 133, 15–23. doi: 10.1016/j.marenvres.2017.11.003

Giordano, M., Beardall, J., and Raven, J. A. (2005). CO_2 concentrating mechanisms in algae: mechanisms, environmental modulation, and evolution. *Annu. Rev. Plant Biol.* 56, 99–131. doi: 10.1146/annurev.arplant.56.032604.144052

Jeffrey, S. W., and Humphrey, G. F. (1975). New spectrophotometric equations for determining chlorophylls a, b, c_1 and c_2 in higher plants, algae and natural phytoplankton. *Biochem. Physiol. Pflanzen* 167, 191–194. doi: 10.1016/S0015-3796(17)30778-3

Jiang, H., Liao, X. H., Zou, D. H., Huang, B., and Liu, Z. W. (2020). The regulations of varied carbon-nitrogen supplies to physiology and amino acid contents in *Gracilariopsis lemaneiformis* (Gracilariales, Rhodophyta). *Algal Res.* 47:101818. doi: 10.1016/j.algal.2020.101818

Kouassi, A. M., and Zika, R. G. (1992). Light-induced destruction of the absorbance property of dissolved organic matter in seawater. *Toxicol. Environ. Chem.* 35, 195–211. doi: 10.1080/02772249209357816

Liu, F., Liu, X. F., Wang, Y., Jin, Z., Moejes, F. W., and Sun, S. (2018). Insights on the *Sargassum horneri* golden tides in the Yellow Sea inferred from morphological and molecular data. *Limnol. Oceanogr.* 63, 1762–1773. doi: 10.1002/lno.10806

Liu, T., Ma, Z. L., Li, H., and Xu, Z. G. (2019). Effects of light intensity and nitrate level on growth and photosynthetic characteristics of *Sargassum horneri*. *Chin. J. Ecol.* 38, 762–769. doi: 10.13292/j.1000-4890.201903.026 (in Chinese with English abstract).

Liu, W., Xin, M. L., Zhou, J., Zhan, D. M., Ding, G., and Wu, H. Y. (2021). Habitat hierarchies distribution of *Sargassum muticum* in Lidao bay, Shandong, China based on habitat suitability index model. *Chin. J. Appl. Ecol.* 2, 1061–1068. doi: 10.13287/j.1001-9332.202103.034 (in Chinese with English abstract).

Lobban, C. S., Harrison, P. J., and Duncan, M. J. (1985). *The Physiological Ecology of Seaweeds*. New York, NY: Cambridge University Press, 29–34.

Louime, C., Fortune, J., and Gervais, G. (2017). *Sargassum* invasion of coastal environments: A growing concern. *Amer. J. Environ. Sci.* 13, 58–64. doi: 10.3844/ajessp.2017.58.64

Makarov, M. V. (2012). Adaptation of the light-harvesting complex of the Barents Sea brown seaweed *Fucus vesiculosus* L. to light conditions. *Dokl. Biol. Sci.* 442, 58–61. doi: 10.1134/S0012496612010176

Milledge, J. J., and Harvey, P. J. (2016). Golden tides: problem or golden opportunity? The valorization of *Sargassum* from beach inundations. *J. Mar. Sci. Eng.* 4, 1–19. doi: 10.3390/jmse4030060

Murata, N., Takahashi, S., Nishiyama, Y., and Allakhverdiev, S. I. (2007). Photoinhibition of photosystem II under environmental stress. *BBA-Bioenergetics* 1767, 414–421. doi: 10.1016/j.bbabio.2006.11.019

Parsons, T. R., Maita, Y., and Lalli, C. M. (1984). *A Manual of Chemical & Biological Methods for Seawater Analysis*. New York, NY: Pergamon Press, 173. doi: 10.1016/B978-0-08-030287-4.50034-7

Provasoli, L. (1968). "Media and prospects for cultivation of marine algae," in *Cultures and Collections of Algae*, eds A. Watanabe and A. Hattori (Kyoto: Japanese Society for Plant Physiology), 63–75.

Qi, L., Hu, C. M., Mikelsons, K., Wang, M. H., Lance, V., Sun, S. J., et al. (2020). In search of floating algae and other organisms in global oceans and lakes. *Remote Sen. Environ.* 239:111659. doi: 10.1016/j.rse.2020.111659

Qi, L., Hu, C. M., Wang, M. Q., Shang, S. L., and Wilson, C. (2017). Floating algae blooms in the east China sea. *Geophys. Res. Lett.* 44, 11501–11509. doi: 10.1002/2017GL075525

Quintano, E., Celis-Plá, P. S. M., Martínez, B., Díez, I., Muguerza, N., Figueroa, F. L., et al. (2019). Ecophysiological responses of a threatened red alga to increased irradiance in an *in situ* transplant experiment. *Mar. Environ. Res.* 144, 166–177. doi: 10.1016/j.marenvres.2019.01.008

Rugiu, L., Hargrave, M. S., Enge, S., Sterner, M., Nylund, G. M., and Pavia, H. (2020). Kelp in IMTAs: small variations in inorganic nitrogen concentrations drive different physiological responses of *Saccharina latissimi*. *J. Appl. Phycol.* 33, 1021–1034. doi: 10.1007/s10811-020-02333-8

Rumyantseva, A., Henson, S., Martin, A., Thompson, A. F., Damerell, G. M., Kaiser, J., et al. (2019). Phytoplankton spring bloom initiation: the impact of atmospheric forcing and light in the temperate North Atlantic ocean. *Prog. Oceanogr.* 178:102202. doi: 10.1016/j.pocean.2019.102202

Salvaterra, T., Green, D. S., Crowe, T. P., and O'Gorman, E. J. (2013). Impacts of the invasive alga *Sargassum muticum* on ecosystem functioning and food web structure. *Biol. Invasions* 15, 2563–2576. doi: 10.1007/s10530-013-0473-4

Sánchez, I., and Fernández, C. (2018). Impact of the invasive species *Sargassum muticum* (Phaeophyta) on tide pool macroalgal assemblages of northern Spain. *Estuar. Coast. Shelf S.* 210, 45–54. doi: 10.1016/j.ecss.2018.06.003

Schell, J. M., Goodwin, D. S., and Siuda, A. N. S. (2015). Recent *Sargassum* inundation events in the Caribbean: shipboard observations reveal dominance of a previously rare form. *Oceanography* 28, 8–10. doi: 10.5670/oceanog.2015.70

Sfriso, A., and Facca, C. (2013). Annual growth and environmental relationships of the invasive species *Sargassum muticum* and *Undaria pinnatifida* in the lagoon of Venice. *Estuar. Coast. Shelf S.* 129, 162–172. doi: 10.1016/j.ecss.2013.05.031

Smetacek, V., and Zingone, A. (2013). Green and golden seaweed tides on the rise. *Nature* 504, 84–88. doi: 10.1038/nature12860

Strickland, J. D. H., and Parsons, T. R. (1972). *A Practical Handbook of Seawater Analysis*. Ottawa: The Alger Press Ltd, 293.

Tedetti, M., and Sempéré, R. (2006). Penetration of ultraviolet radiation in the marine environment. A review. *Photochem. Photobiol.* 82, 389–397. doi: 10.1562/2005-11-09-IR-733

Teichberg, M., Fox, S. E., Olsen, Y. S., Valiela, I., Martinetto, P., Iribarne, O., et al. (2010). Eutrophication and macroalgal blooms in temperate and tropical coastal waters: nutrient enrichment experiments with Ulva spp. *Glob. Change Biol.* 16, 2624–2637. doi: 10.1111/j.1365-2486.2009.02108.x

van Tussenbroek, B. I., Arana, H. A. H., Rodríguez-Martínez, R. E., Espinoza-Avalos, J., Canizales-Flores, H. M., González-Godoy, C. E., et al. (2017). Severe impacts of brown tides caused by *Sargassum* spp. on near-shore Caribbean seagrass communities. *Mar. Pollut. Bull.* 122, 272–281. doi: 10.1016/j.marpolbul.2017.06.057

Wang, M. Q., Hu, C. M., Barnes, B. B., Mitchum, G., Lapointe, B., and Montoya, J. P. (2019). The great Atlantic *Sargassum* belt. *Science* 365, 83–87. doi: 10.1126/science.aaw7912

Williams, A., and Feagin, R. (2010). *Sargassum* as a natural solution to enhance dune plant growth. *Environ. Manage.* 46, 738–747. doi: 10.1007/s00267-010-9558-3

Wu, H. L., Gao, G., Zhong, Z. H., Li, X. S., and Xu, J. T. (2018). Physiological acclimation of the green tidal alga *Ulva prolifera* to a fast-changing environment. *Mar. Environ. Res.* 137, 1–7. doi: 10.1016/j.marenvres.2018.02.018

Wu, Y. P., Gao, K. S., Li, G., and Helbling, E. W. (2010). Seasonal impacts of solar UV radiation on the photosynthesis of phytoplankton assemblages in the coastal water of the South China Sea. *Photochem. Photobiol.* 86, 586–592. doi: 10.1111/j.1751-1097.2009.00694.x

Xing, Q. G., Guo, R. H., Wu, L. L., An, D. Y., Cong, M., Qin, S., et al. (2017). High-resolution satellite observations of a new hazard of golden tides caused by floating *Sargassum* in winter in the yellow sea. *IEEE Geosci. Remote. Sens. Lett.* 14, 1815–1819. doi: 10.1109/LGRS.2017.2737079

Xu, Z. G., and Gao, K. S. (2009). Impacts of UV radiation on growth and photosynthetic carbon acquisition in *Gracilaria lemaneiformis* (Rhodophyta) under phosphorus-limited and replete conditions. *Funct. Plant Biol.* 36, 1057–1064. doi: 10.1071/FP09092

Xu, Z. G., and Gao, K. S. (2012). NH_4^+ enrichment and UV radiation interact to affect the photosynthesis and nitrogen uptake of *Gracilaria lemaneiformis* (Rhodophyta). *Mar. Pollut. Bull.* 64, 99–105. doi: 10.1016/j.marpolbul.2011.10.016

Xu, Z. G., Wu, H. Y., Zhan, D. M., Sun, F. X., Sun, J. Z., and Wang, G. C. (2014). Combined effects of light intensity and NH_4^+-enrichment on growth, pigmentation, and photosynthetic performance of *Ulva prolifera* (Chlorophyta). *Chin. J. Oceanol. Limn.* 32, 1016–1023. doi: 10.1007/s00343-014-3332-y

Yakovleva, I. M., and Titlyanov, E. A. (2001). Effect of high visible and UV irradiance on subtidal *Chondrus crispus*: stress, photoinhibition and protective mechanisms. *Aquat. Bot.* 71, 47–61. doi: 10.1016/S0304-3770(01)00167-X

Yu, J., Li, J. Y., Wang, Q. H., Liu, Y., and Gong, Q. L. (2019). Growth and resource accumulation of drifting *Sargassum horneri* (Fucales, Phaeophyta) in response to temperature and nitrogen supply. *J. Ocean U. China* 18, 1216–1226. doi: 10.1007/s11802-019-3835-4

Zhang, D., Beer, S., Li, H., and Gao, K. S. (2020). Photosystems I and II in *Ulva lactuca* are well protected from high incident sunlight. *Algal Res.* 52:102094. doi: 10.1016/j.algal.2020.102094

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.

frontiers in Environmental Science

ORIGINAL RESEARCH
published: 30 November 2021
doi: 10.3389/fenvs.2021.755587

Single Shot MultiBox Detector for Urban Plantation Single Tree Detection and Location With High-Resolution Remote Sensing Imagery

*Yueyuan Zheng and Gang Wu**

College of Information, Beijing Forestry University, Beijing, China

Using high-resolution remote sensing images to automatically identify individual trees is of great significance to forestry ecological environment monitoring. Urban plantation has realistic demands for single tree management such as catkin pollution, maintenance of famous trees, landscape construction, and park management. At present, there are problems of missed detection and error detection in dense plantations and complex background plantations. This paper proposes a single tree detection method based on single shot multibox detector (SSD). Optimal SSD is obtained by adjusting feature layers, optimizing the aspect ratio of a preset box, reducing parameters and so on. The optimal SSD is applied to single tree detection and location in campuses, orchards, and economic plantations. The average accuracy based on SSD is 96.0, 92.9, and 97.6% in campus green trees, lychee plantations, and palm plantations, respectively. It is 11.3 and 37.5% higher than the latest template matching method and chan-vese (CV) model method, and is 43.1 and 54.2% higher than the traditional watershed method and local maximum method. Experimental results show that SSD has a strong potential and application advantage. This research has reference significance for the application of an object detection framework based on deep learning in agriculture and forestry.

Keywords: single shot multibox detector, Urban forest, tree detection, tree location, high-resolution remote sensing image

OPEN ACCESS

Edited by:
Peng Liu,
Institute of Remote Sensing and Digital Earth (CAS), China

Reviewed by:
Costica Nitu,
Politehnica University of Bucharest, Romania
M. Iqbal Saripan,
Putra Malaysia University, Malaysia
Wei Liu,
Institute of Automation (CAS), China

****Correspondence:***
Gang Wu
wugang@bjfu.edu.cn

Specialty section:
This article was submitted to Environmental Informatics and Remote Sensing, a section of the journal Frontiers in Environmental Science

Received: *09 August 2021*
Accepted: *12 November 2021*
Published: *30 November 2021*

Citation:
Zheng Y and Wu G (2021) Single Shot MultiBox Detector for Urban Plantation Single Tree Detection and Location With High-Resolution Remote Sensing Imagery. Front. Environ. Sci. 9:755587. doi: 10.3389/fenvs.2021.755587

1 INTRODUCTION

Single tree detection based on remote sensing images is a crucial technology for establishing a single tree database and monitoring single tree plantation resources, which is of great significance to urban landscape planning and ecological environment monitoring (Congalton et al., 2014; Faridatul and Wu, 2018; Ahl et al., 2019). Single tree detection is a cross-research field of computer vision, measurement, single tree management, and remote sensing (Kupidura et al., 2019; Zhang et al., 2020; Belcore et al., 2021). Researchers began to explore single tree detection methods a long time ago. As early as 1995, Gougeon et al. (Gougeon, 1995) used aerial photos to carry out single tree identification; they searched for the local minimum value at the bottom of a tree for the first time. Larsen et al. (Larsen and Rudemo, 1998) used an improved template matching method to detect crown vertices of a single tree. Poullot et al. (Pollock, 1996) used remote sensing imagery to determine the location of a single tree by selecting a moving window from 15 × 15 to 30 × 30.

Depending on the size of the canopy in the image, they determined the location of a single tree and used the local ray method to depict the young conifer forest crown. Wang et al. (Wang et al., 2004) used the watershed method to depict the canopy boundaries of white cloud fir forests. Zhang Ning et al. (Zhang et al., 2014) improved the application of the peak climbing method to the problem of canopy extraction and experimented in Quickbird images. The accuracy of experimental samples reached more than 85%. Jiang Renrong et al. (Jiang et al., 2016) used hydro analyzation and regional growth fusion methods for lychee single tree detection and canopy depiction. The overall accuracy was 78.69%. Yu et al. (Yu et al., 2018) applied the iterative threshold method to canopy extraction. The matching rate of the iterative threshold method was only 60.15%, due to complicated and discrepant texture and over-splitting phenomenon in a single tree canopy.

In recent years, researchers applied convolutional neural network (CNN) to single tree detection, solving the problems of traditional single tree detection methods (Liu et al., 2017; Zhao et al., 2020; Zhang et al., 2021). For example, traditional single tree detection methods detected seed points or matching templates by pixels, so traditional single tree detection methods required prior knowledge to specify characteristic parameters of different scenes. The traditional single tree detection method had low stability. The above problems could be solved by introducing CNN (Sharma et al., 2016; Chen et al., 2017). CNN could learn features automatically and abstract local low-level features into high-level features such as the color and the contour of trees gradually without specified parameters in advance (Chu et al., 2017; Mokroš et al., 2018). CNN showed good advantages in single tree detection (Roska and Chua, 2008; Wang et al., 2020). Weijia et al. (Weijia et al., 2016) applied the deep learning approach to detect densely planted Malaysian oil palm trees for the first time. Guirado et al. (Guirado et al., 2017) proposed a CNN-based shrub detection method. Dong Tianyang et al. (Dong et al., 2018) proposed a cascaded convolutional neural network of single tree detection in 2018. They applied it to Google Earth images in 2019 (Dong et al., 2019), and found that it was hard to repeat, had small object leakage, was inefficient, and that it was challenging to meet practical requirements.

At present, there is no single tree detection method that can adapt to various stands (Liu et al., 2019). It is a significant research direction to use existing technology to improve the accuracy of single tree detection and simplify the single tree detection process (Deng et al., 2010). Currently, object detection based on deep learning is divided into one-stage object detection and two-stage object detection. One-stage object detection is also known as end-to-end, which only takes one step to obtain results. Single shot multibox detector (SSD) is one of the most widely used in one-stage object detection. SSD uses multi-size convolutional layers to predict, adding data enhancement of small objects, which has the advantages of high accuracy and high efficiency (Li et al., 2019).

In this study, we have improved the SSD backbone feature extraction network and optimized the aspect ratio of a preset box in single tree detection. The SSD model of single tree detection has been reduced in terms of parameters and computation burden. The optimal SSD model of single tree detection was obtained by comparing the experimental results of multiple groups. The optimal SSD model was used to detect quantity identification and location management of urban plantations. The optimal SSD model achieved better accuracy than the traditional watershed method, traditional local maximum method, and the latest template matching method, the CV model method.

2 METHODOLOGIES

This section focuses on the principles, training method, and other details of SSD. The principle of SSD is shown in **Figure 1**. During training, we only need to enter the original image and file that marks the original image's actual box. In multiple feature maps (e.g., 9 × 9 (**Figure 1B**) and 5 × 5 (**Figure 1C**)), SSD uses convolution kernel to calculate the category confidence of the detected box and the offset between the actual box and detected box. During training, we match these preset boxes with the actual box first. For example, if three preset boxes match three trees, they are regarded as positive samples. The remaining preset boxes are treated as negative samples.

2.1 Single Tree Detection Process

The overall process of single tree detection based on SSD is shown in **Figure 2**. The process of single tree detection based on SSD is divided into five parts, including:

1) Collect high resolution remote sensing image data.
2) Separate training data from validation data and label training data.
3) The SSD model is trained to obtain single tree characteristic parameters.
4) Single tree detection is carried out by single tree characteristic parameters.
5) Evaluation of test results.

A small number of factors can affect the quality of the SSD model, such as data richness, feature extraction accuracy, and robustness.

Simplified SSD Object Detection Framework

In this study, SSD is an object detection framework with 300 × 300×3 as the input image. SSD mainly divides into three parts: a central feature extraction network, feature layer processing network, and stacking adjustment parameters. The SSD model is simplified and shown in **Figure 3**. The backbone feature extraction network uses a visual geometry group network (VGG16) with the partial convolution layer removed. The optimal SSD model four extracts four feature layers for object detection, and the sizes of these four feature layers are 38 × 38×256, 19 × 19×512, 10 × 10×512, and 5 × 5×256, respectively. The 38 × 38×256 feature layer can be understood as dividing 300 × 300 image evenly into 38 × 38 parts. The center of each part generates preset boxes of different sizes and proportions for an

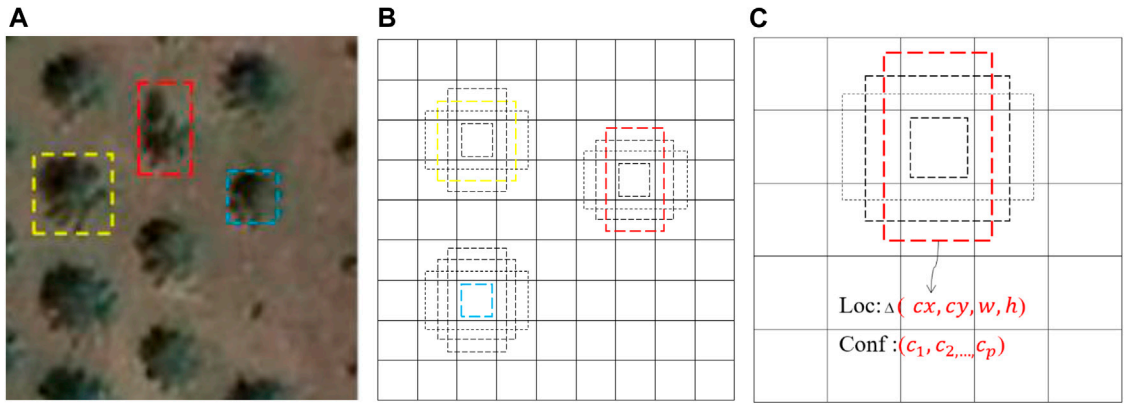

FIGURE 1 | Principles of SSD. **(A)** Actual box. **(B)** 9 × 9 grid. **(C)** 5 × 5 grid.

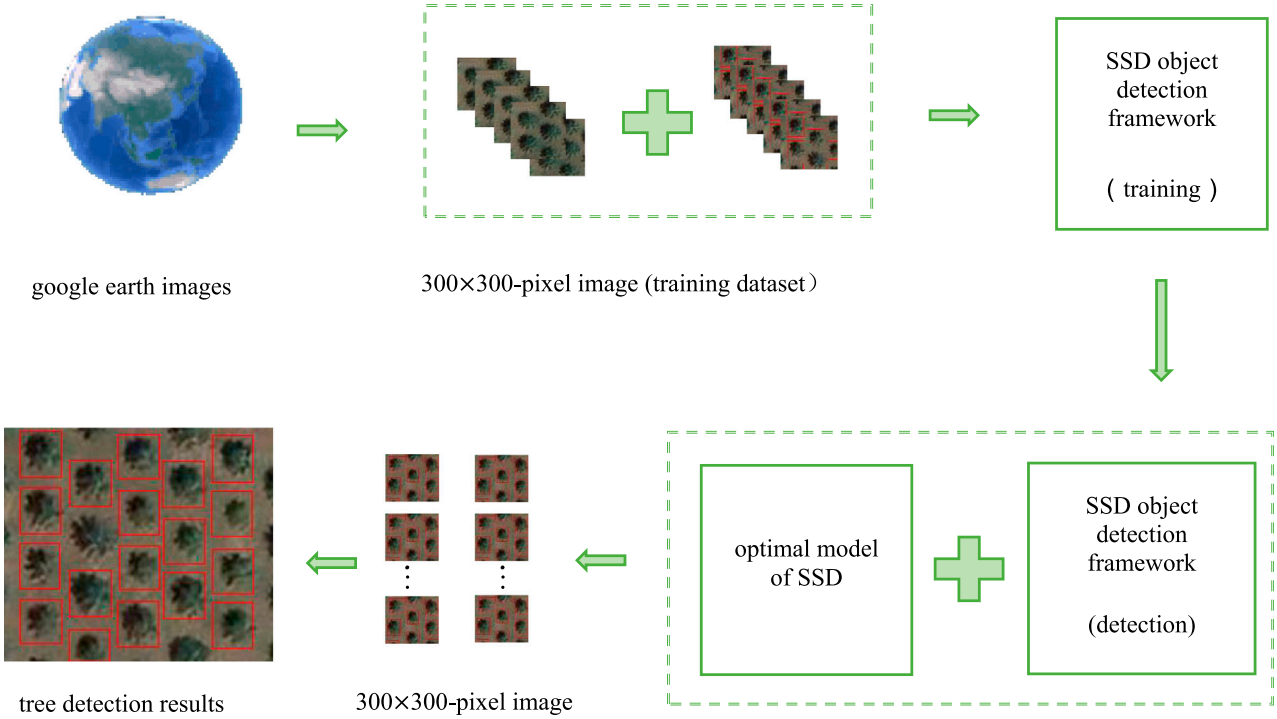

FIGURE 2 | The overall process of single tree detection based on SSD.

anchor point. The preset box is introduced in **Section 2.4**. The size of the feature map decreases gradually. The large feature map predicts small objects, while the small feature map predicts large objects.

For feature layer processing, the extracted four feature layers are convolved twice, one convolution to extract category confidence, another convolution to extract position adjustment parameters. Four parameters are needed to control the position of each preset box, including offset abscissa of point, ordinate of point, height, and width.

2.2 Non-Maximum Suppression

The main idea of non-maximum suppression is to search for local maximum values and suppress non-maximum values. As can be known from **Section 2.4**, SSD produces many preset boxes, overlapping between each preset box, and each preset box has a category confidence score. By introducing a non-maximum suppression method, we find the best location for a single tree by removing excess preset boxes and only retaining the optimal preset boxes. The process of non-maximum suppression is as follows:

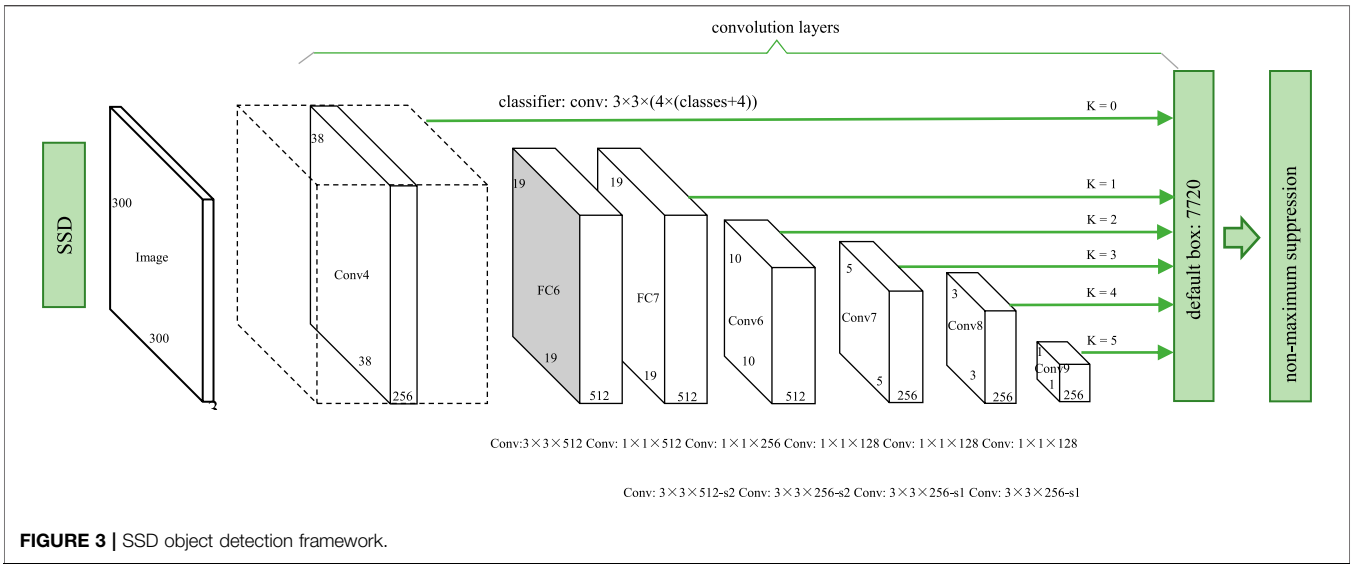

FIGURE 3 | SSD object detection framework.

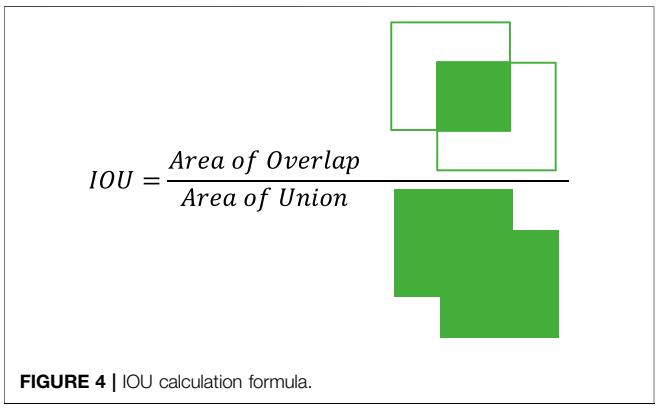

FIGURE 4 | IOU calculation formula.

1) Sort each preset box by category confidence.
2) Select the preset box with the highest confidence in the category as the output box and remove it from the list of preset boxes.
3) Calculate the area of all preset boxes.
4) Calculate the intersection over union (IOU) value of the output box and other preset box. As shown in **Figure 4**, the intersection area of two boxes divides by their union area.
5) Remove a preset box with an IOU greater than threshold from the list of preset boxes.
6) Repeat steps (1) to (5) until the list of preset boxes is empty.

As shown in **Figure 5**, the non-maximum suppression method is used to select what is most likely a tree canopy among detected objects. A single left tree outputs three preset boxes, and a single right tree outputs two preset boxes. The score sequence of preset boxes is {0.92, 0.86, 0.81, 0.49, 0.42}, and the highest score is 0.92. The preset box of 0.92 is taken as a detected single tree prediction box, and then the IOU value of the remaining preset boxes and prediction box is calculated. If the IOU of the preset boxes of 0.81 and 0.49 and the prediction box of 0.92 exceed threshold, the two preset boxes of 0.81 and 0.49 will be deleted from the list. The remaining two IOU areas of 0.86 and 0.42 are less than the set threshold, and they are rearranged as {0.86 and 0.42} according to score. The highest score is 0.82. The preset box of 0.82 is used as the detected single tree prediction box. The final test results are obtained by excluding preset box 0.42 through IOU.

2.4 Preset Box

The scale of the preset box follows a linear increment rule, increasing linearly as the size of the feature map decreases:

$$S_k = S_m + \frac{S_{max} - S_{min}}{4}(k-1),\ k \in [1,\ m] \tag{1}$$

In **Eq. 1**, m refers to the number of feature layers. Four feature layers are extracted, but $m = 3$, because the first feature layer (Conv4) is set separately. S_k represents the ratio of the preset box size to the image, and S_{min} and S_{max} represent minimum and maximum values of ratio. In this study, S_{min} is set to 0.2 and S_{max} to 0.9. For the first feature layer, the minimum ratio of the preset box to the original picture is $\frac{S_{min}}{2} = 0.1$, the size of preset box is $300 \times 0.1 = 30$. According to **Eq. 1** calculation, the S_k of each feature layer is $S_1 = 0.2$, $S_2 = 0.375$, and $S_3 = 0.55$. The scale of each feature layer preset box is 30, 60, 112.5, and 165.

The shape of the crown of a single tree is mainly round and oval, so the default width to height ratio of the frame that comes to mind at first is 1:$\frac{1}{2}$ or 1:1. However, when cutting the picture for detection, the width to height ratio of the half tree frame is more than $\frac{1}{2}$. In this study, we select $a_r \in \{1, 2, 3, \frac{1}{2}, \frac{1}{3}\}$, S_k refers to the actual scale of preset box, and width ($w_k^a = S_k\sqrt{a_r}$) and height ($h_k^a = \frac{S_k}{\sqrt{a_r}}$) of the present box are calculated. Most tree crowns are more circular. And each feature map will have an S_k preset box of $a_r = 1$ and an $S_k^{'}$ scale. Besides, there is a preset box of $S_k^{'} = \sqrt{S_k S_{k+1}}$ and $a_r = 1$, so that each feature map has two width to height ratios of preset boxes with an aspect ratio of 1: 1, they are 1 and 1$^{'}$. The last feature layer needs $S_{m+1} = 300 \times \frac{71}{100} = 213$ to

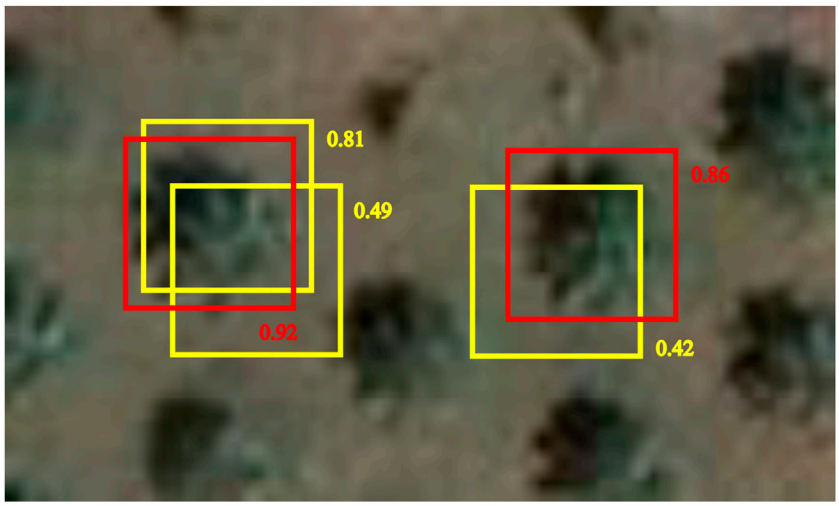

FIGURE 5 | Non-maximum suppression method single tree detection results.

calculate $S_m^{'}$. Therefore, there are six preset boxes ($a_r \in \{1, 2, 3, \frac{1}{2}, \frac{1}{3}, 1^{'}\}$) for each feature map and each anchor point.

The coordinate of anchor points can be obtained by point formula ($\frac{i+0.5}{f_k} \times 300, \frac{j+0.5}{f_k} \times 300$), where $i, j \in (0, |f_k|)$ and $|f_k|$ are the size of the k-th feature map.

As shown from above, four feature layers are extracted, respectively 38 × 38×256, 19 × 19×512, 10 × 10×512, and 5 × 5×256. The number of preset boxes at each feature map is 4, 6, 6, and 6, so there are 8,692 preset boxes in each original image. The single tree position in the original image is retrieved using the thick preset box.

2.5 Loss Function

The loss function is defined on a single sample, and it is the error of a sample.

$$L(x, c, l, g) = \frac{1}{N}\left(L_{conf}(x, c) + \alpha L_{loc}(x, l, g)\right) \tag{2}$$

where L is the loss function of SSD. The loss function of SSD is divided into two parts: location loss and class confidence loss. L_{loc} is location loss, and L_{conf} is class confidence loss. The confidence loss is SoftMax loss over multiple classes confidences, c stands for confidence loss. N is the number of matched preset boxes. If N = 0, loss is set to 0. The positioning loss is smoothing loss (Girshick; R.2015) between the preset box and actual box. l is the predicted box, g is the actual box. α is used to adjust the ratio between class confidence loss and location loss. By default, α = 1.

$$L_{loc}(x, l, g) = \sum_{i \in Pos}^{N} \sum_{m \in \{cx, cy, w, h\}} x_{ij}^{k} smooth_{L1}\left(l_i^m - \hat{g}_j^m\right) \tag{3}$$

$$\hat{g}_j^{cx} = \left(g_j^{cx} - d_i^{cx}\right)/d_i^w,\ \hat{g}_j^{cy} = \left(g_j^{cy} - d_i^{cy}\right)/d_i^h \tag{4}$$

$$\hat{g}_j^w = \log\left(\frac{g_j^w}{d_i^w}\right),\ \hat{g}_j^h = \log\left(\frac{g_j^h}{d_i^h}\right) \tag{5}$$

$$L_{conf}(x, c) = -\sum_{i \in Pos}^{N} x_{ij}^p log\left(\hat{c}_i^p\right) - \sum_{i \in Neg} \log\left(\hat{c}_i^0\right) \tag{6}$$

$$\hat{c}_i^p = \frac{\exp\left(c_i^p\right)}{\sum_p exp\left(c_i^p\right)} \tag{7}$$

where $smooth_{L1}$ is smooth loss, Pos is positive samples, Neg is negative samples, (cx, cy) is the regressed offsets for the center of the preset box. w is the width of the preset box, h is the height of the preset box, d is the preset box, and $\hat{g}$ is the actual box that has been offset. p refers to category, $p = 0$ represents the background. $x_{ij}^p = \{1, 0\}$ is an indicator for matching the i-th preset box to thet j-th actual box of p.

2.6 Increased Accuracy of Small Object

The main process of data enhancement is as follows:

1) Use entire original picture.
2) Take a small piece from the original image, and minimum overlap between this small piece and actual box is 0.1, 0.3, 0.5, 0.7, or 0.9.
3) Take a piece randomly from the original picture.

The ratio of sub-block size to original image size is between 0.1 and 1, and the width to height ratio is between 0.5 and 2. If the center of the actual single tree frame is within the intercepted sub-block, overlap is retained as the actual box of the sub-block. Scale the size of each subblock to 300 × 300, and each sub-block has a probability of 0.5 to flip horizontally before training.

Distorting images is a way to enhance data, including randomly changing contrast, brightness, saturation, and tone of image and randomly disrupting the three RGB channels. After data enhancement, training images have richer features, which can enhance the robustness of the model. In this study, we prove that this strategy can effectively improve the detection accuracy of the SSD model.

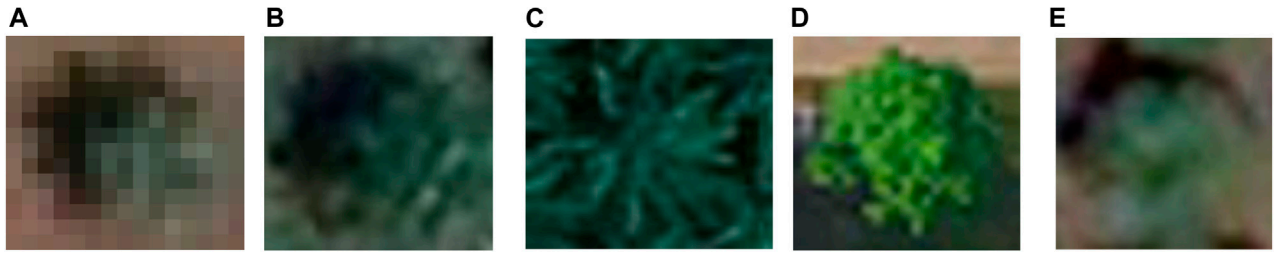

FIGURE 6 | Training data sampling ground images. **(A)** Open palm plantation. **(B)** Complex background palm plantation. **(C)** Dense palm plantation. **(D)** Campus green tree. **(E)** Orchard.

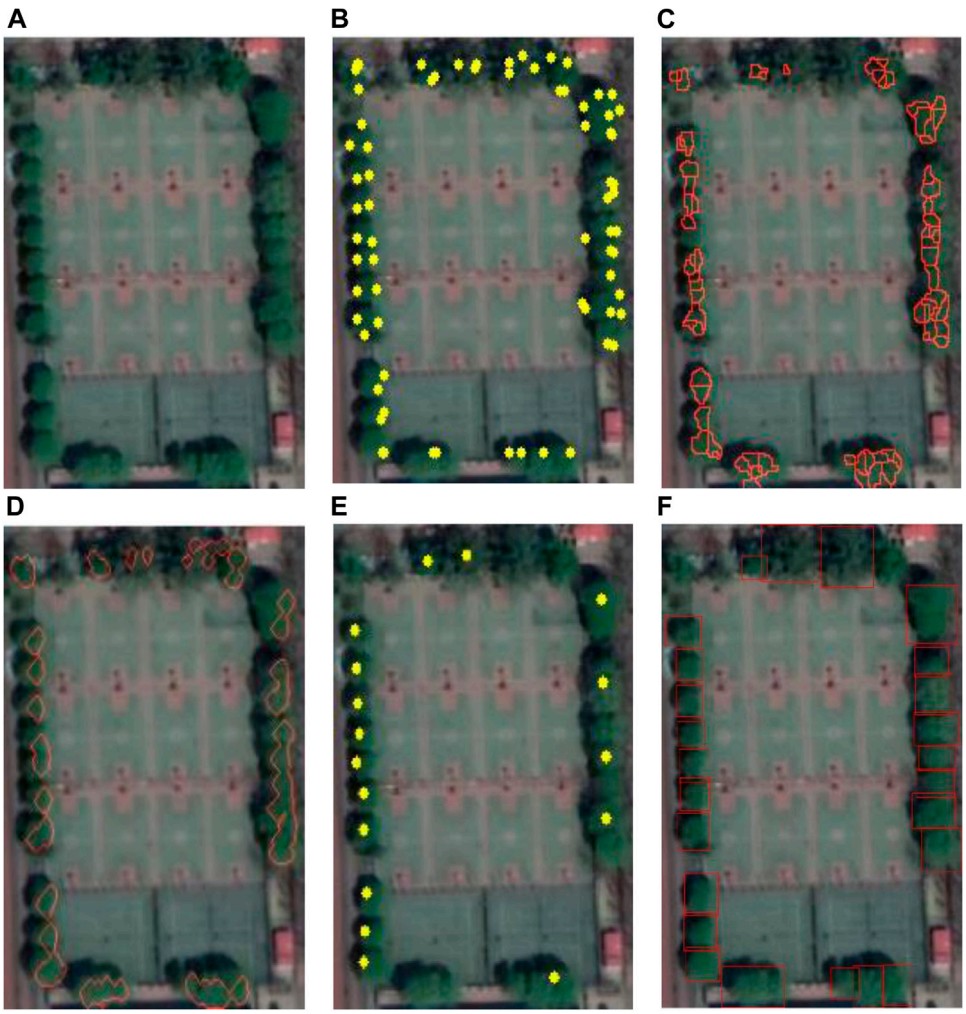

FIGURE 7 | Results of different methods in campus. **(A)** Original image. **(B)** Local max. **(C)** Watershed. **(D)** CV model. **(E)** Template matching. **(F)** SSD.

3 EXPERIMENTS

3.1 Study Area

High-resolution remote sensing images of three groups of representative different city types are used in the study, including campus, orchard, and economic plantation. In June 2019, a campus green tree was located at 116°20′8.76″E, 40°0′6.52″N, at Beijing Forestry University in China. In December 2017, a litchi plantation was located at 113°53′26.34″E, 22°38′41.22″N in Bao'an district, Shenzhen city, China. A palm plantation was collected at

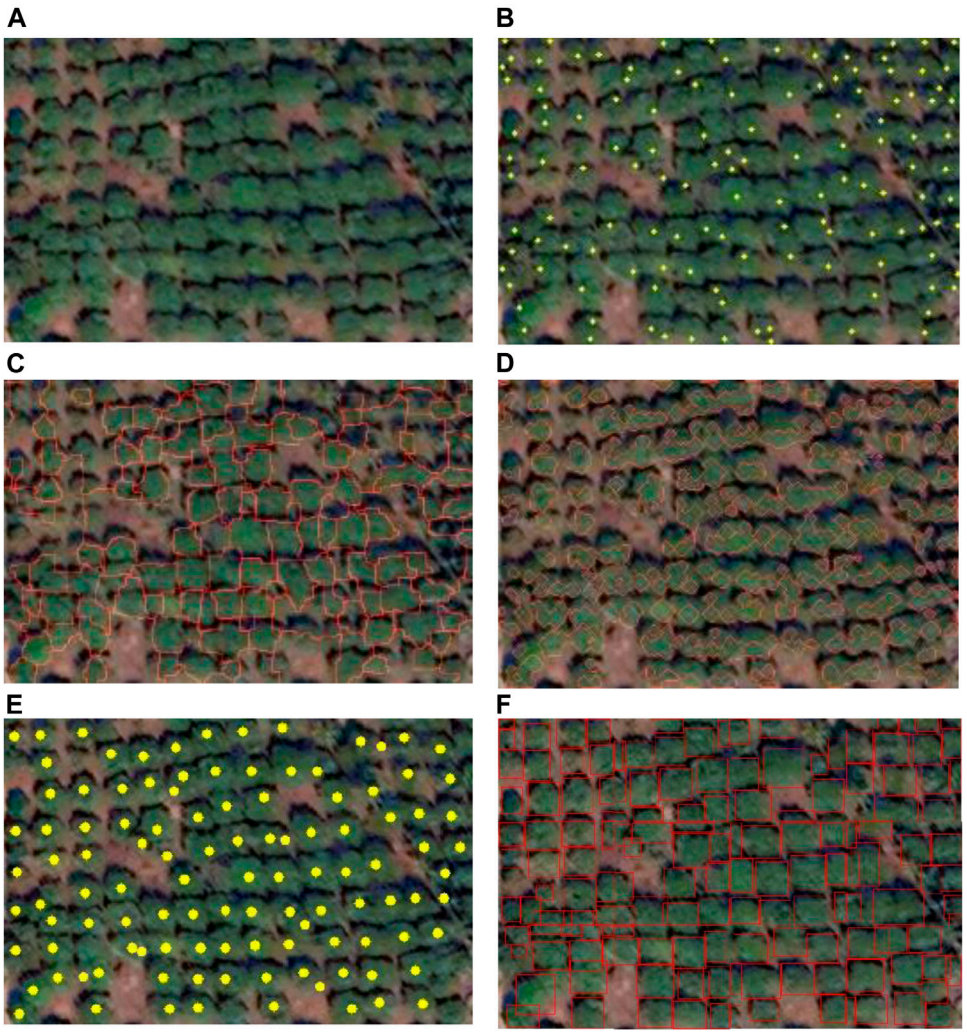

FIGURE 8 | Results of different methods in lychee plantation. **(A)** Original image. **(B)** Local max. **(C)** Watershed. **(D)** CV model. **(E)** Template matching. **(F)** SSD.

98°20′53.22″ E, 8°27′18.45″ N in Phang Nga province, Thailand. High-resolution remote sensing images are used in our experiments, with a spatial resolution of 0.27m, a scale of 800:1, and a visual field height of 1 km.

3.2 Training Dataset and Sample Dataset

The training data were collected from Google Earth images, and the training data of each experimental group are shown in **Figure 6**. Training data of palm trees were collected from around palm trees. Campus green tree training data were collected from universities in Beijing, China. Training data for lychee plantation were collected from around the lychee plantation.

Experimental results of the campus sample plot are shown in **Figure 7A**. The position data of a single reference tree were obtained from field measurements.

The sample plot representing the orchard is shown in **Figure 8A**. The position data of the single reference tree are obtained by visual annotation. The palm plantation area representing economic plantation is divided into three groups according to the characteristics of the palm plantation area. The location data of the single reference palm tree were obtained by visual annotation. Open economic plantation refers to the plantation area with a canopy density between 0.4 and 0.6.

The palm sample plot representing open economic plantation is shown in **Figure 9A**. Dense plantation refers to a plantation where the canopy density of palm plantation is between 0.7 and 1. Sample plots of palm trees representing dense economic plantations are shown in **Figure 9C**. Background detection in plantation areas also has a great impact on single tree detection. Especially, background color is like tree crown color, resulting in the background being wrongly identified as a single tree. In addition, complex ground features and shadow generated by sunlight in the background make the shape of a single tree abnormal, which greatly increases the difficulty of single tree detection. In this study, a plantation area with a complex background is experimented as a type alone. The sample

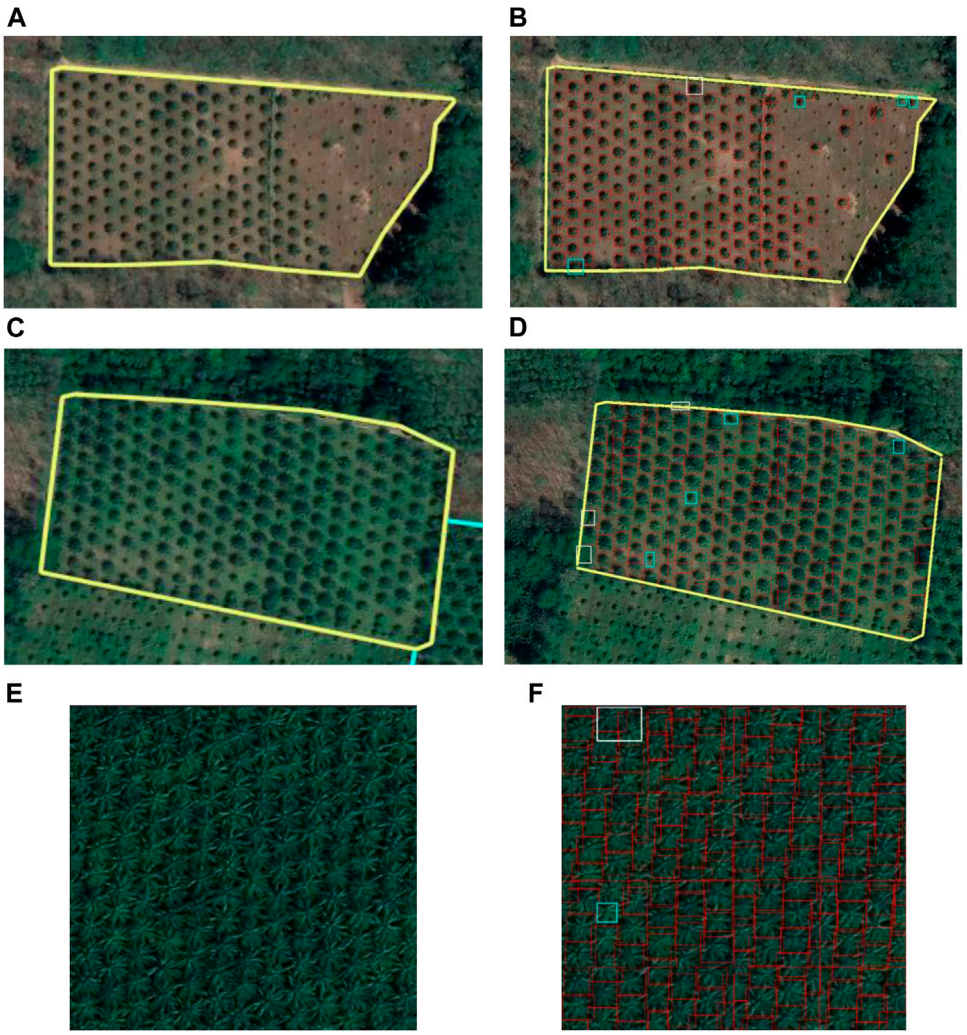

FIGURE 9 | Results of different methods in palm. **(A)** Original image of open palm plantation. **(B)** SSD of open palm plantation. **(C)** Original image of complex background palm plantation. **(D)** SSD of complex background palm plantation. **(E)** Original image of dense palm plantation. **(F)** SSD of dense palm plantation.

TABLE 1 | Training model parameters.

Training model	Optimal SSD	SSD
Epochs	400	400
Batch-size	8	8
Input image size	300×300	300×300
Score	0.5	0.5
Weight file size (MB)	47.5	90.5

plot representing a complex background economic plantation is shown in **Figure 9E**.

3.3 Constant Parameters

Table 1 lists the main parameters used in the experiment. When a complete dataset passes through a neural network once and returns once, the process is called an "Epoch". When data cannot be passed through a neural network at one time, the dataset needs to be divided into several batch-sizes. Each "Batch-size" is equivalent to a new dataset. The "Score" is a confidence score. "Weight file size (MB)" is the size of the model.

3.4 Optimal SSD Model

The open economic plantation is experimentally studied. Palm trees are detected through the SSD object detection framework. Experimental results are shown in **Table 2**. The "trunk feature extractor" refers to the main network structure used in feature extraction. After the main network structure reframes, many feature layers are obtained. And some feature maps are selected in several feature maps to build preset boxes. The base scale of the preset box is related to the feature map. The number of parameters refers to the size of the model.

As shown in **Table 2**, SSD has been improved in many areas for single tree detection, such as omitting some VGG16 feature

TABLE 2 | Experimental results of sample plot 1 under different SSD parameters.

Detection framework	SSD	Simplified SSD			
Model	Model 1	Model 2	Model 3	Model 4	Model 5
Trunk feature extractor	VGG16	Simplified VGG16	Simplified VGG16	Simplified VGG16	Simplified VGG16
The selected feature layer	K0~K5	K0~K5	K0~K4	K0~K3	K0~K3
Preset box base scale	0.1,0.2,0.38,0.55,0.71,0.88	0.1,0.2,0.38,0.55,0.71,0.88	0.1,0.2,0.38,0.55,0.71	0.1,0.2,0.38,0.55	0.1,0.2,0.38,0.55
Preset box aspect ratio	1,2, $\frac{1}{2}$,3, $\frac{1}{3}$,1	1,2, $\frac{1}{2}$,3, $\frac{1}{3}$,1	1,2, $\frac{1}{2}$,3, $\frac{1}{3}$,1	1,2, $\frac{1}{2}$,3, $\frac{1}{3}$,1	1,2, $\frac{1}{2}$,1
Parameter /mb	90.7	60.3	52.5	46.7	46.5
Average accuracy	95.3%	95.9%	95.4%	97.0%	93.9%
Single picture time /s	0.92	0.63	0.57	0.46	0.43

layers, network depth, and unwanted feature layers. The extracted feature layers by model 4 retains only K0-K3, which is approximately 13.6 and 5.8 MB less than model 2 and model 3 parameters, but accuracy does not decrease. In the process, it is found that there are more false positives in model 2 and model 3, and the deletion of feature layer can reduce false positives. Almost, no single tree has a crown height and width of 3 and $\frac{1}{3}$ in the regular top view. SSD model 5 removes $a_r \in \{3, \frac{1}{3}\}$, which makes detection speed increase to 0.43 s, omission rate increase, and accuracy decrease, because sample images are clipped during the experiment. The SSD model can recognize a separated half tree as a tree. After removing $a_r \in \{3, \frac{1}{3}\}$, an incomplete canopy cannot be detected, resulting in some single trees being missed.

Precision-recall is one of the most useful weapons to detect the efficiency of the object detection model. As shown in **Figure 10**, model 5 performs the worst, AP = 89.36%; model 4 has the best performance, AP = 94.37%, in the precision-recall curve of all models. In model 4, the SSD object detection framework only extracts K0~K3 feature layers and sets the aspect ratio to $a_r \in \{1, 2, 3, \frac{1}{2}, \frac{1}{3}, 1^{'}\}$. The accuracy of model 4 is 97%. Under the condition of reducing parameters and time, the accuracy of SSD is improved to the greatest extent.

3.5 Evaluation Criteria of Detection Accuracy

For a variety of single tree detection methods, the evaluation of their detection excellence depends on evaluation standard. At present, there is no unified evaluation standard. The spatial position difference between a ground reference single tree and a detected single tree can be considered as correct detection within a specific range. The geometric center of the actual box is the position of a single tree. The point coordinate of single tree detection and a single reference tree are denoted as M_i and E_j. There are three possibilities for the results of single tree detection: correct detection, error detection, and omission. A set threshold $\varepsilon > 0$, $d\ (M_i, M_i)$ is denoted as the distance between the two points M_i and E_j. M_i is traverse:

1) When $d\ (M_i, E_j) < \varepsilon$, it is considered that the detection of a single tree matches the single reference tree, and it is a correct detection.
2) If there is $d\ (M_i, E_j) > \varepsilon$ for any M_i, there is no reference single tree matching with a detected single tree. The single tree detected is considered as a false detection.
3) If E_j neither conforms to case (1) or case (2), E_j is omission.

Based on the above conditions, N_r is the number of reference single trees, N_a is the number of detected single trees, and N_{match} is the correct number of detected single trees in detected single trees. The calculation formula of all values is shown in **Table 3**, N_{leave} is the number of undetected reference single trees, and is also the difference value between N_r and N_{match}, N_{error} is the difference value between N_{match} and N_a. The recall rate is represented by symbol N_{mat}, N_{om} is commission rate, N_{com} is omission rate, and M is accuracy.

4 SINGLE TREE DETECTION RESULTS

The optimal SSD model is applied to the single tree detection of an urban plantation, and experimental results are compared with the latest single tree detection method. Optimal SSD model 4 is called the SSD model in the following.

4.1 Campus

The available view of single tree detection around the campus is shown in **Figure 7**. The statistical analysis of experimental results is shown in **Table 4**. The accuracy of the five methods differs significantly. From the experimental results, the traditional local maximum method and watershed method are generally effective in single tree detection. The watershed method has the worst result. The accuracy of the watershed method is 32.9%. The latest template matching method and CV model method have a good effect. The accuracy of the SSD model is 96% and has the highest accuracy in the experimental results of five methods. Specifically, the SSD model gains high scores in recall rate and omission rate. The commission rate of the template matching method is zero. The omission rate of template matching is 29.1%.

4.2 Orchard

The visual results of the experiment in the lychee plantation are shown in **Figure 8**. The accuracy of the five methods is significantly different, as shown in **Table 4**. SSD has the highest accuracy of five methods. The accuracy of SSD is 92.9%. The lowest accuracy is 31.8%. The accuracy of the local maximum method is the lowest. The accuracy of SSD is 61.1% more than the accuracy of the local maximum method. The accuracy of watershed is 52.9%. Obviously, SSD's single tree detection effect is far better than the traditional single tree

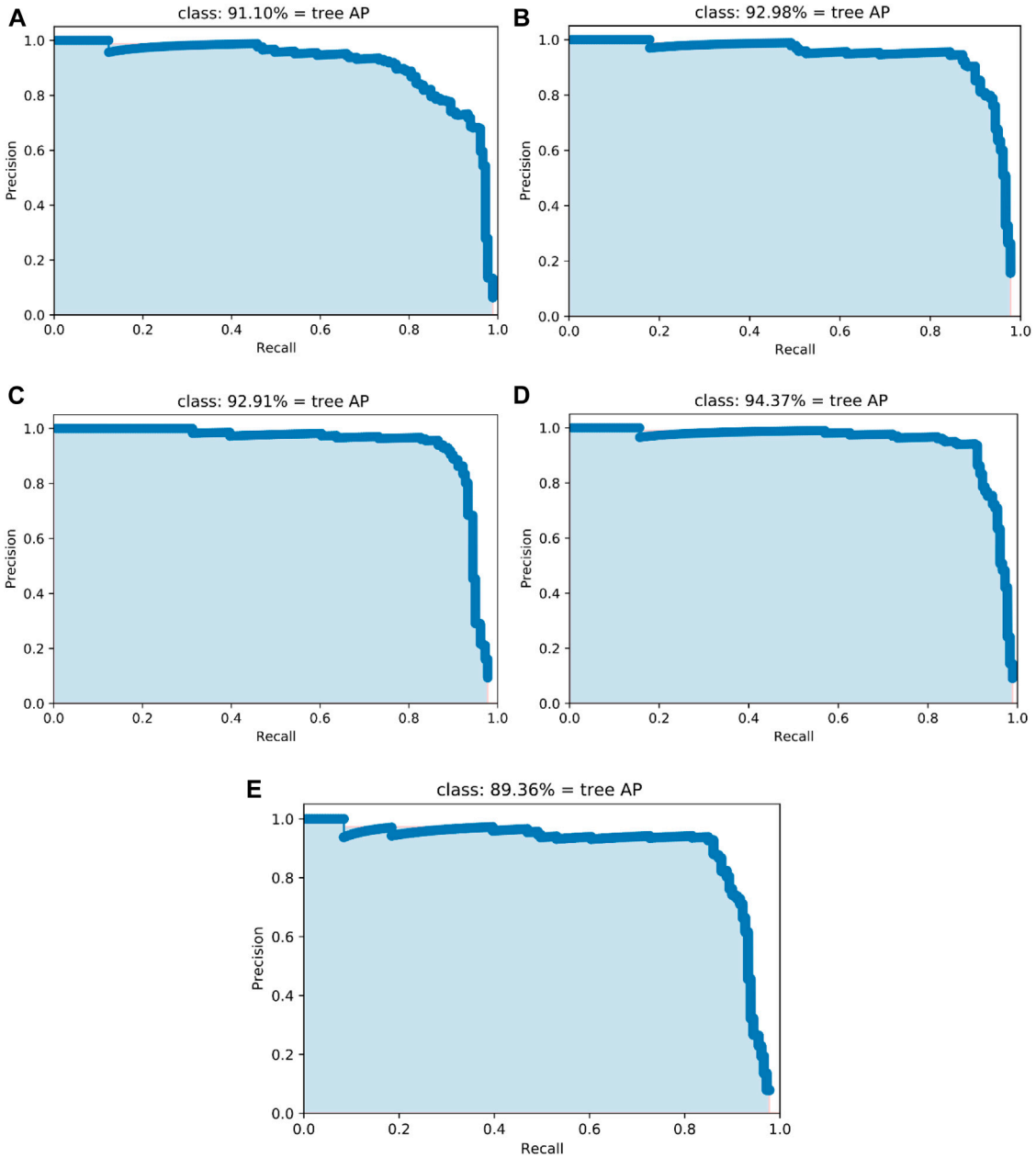

FIGURE 10 | Precision-recall curves for each method. **(A)** Model 1 **(B)** Model 2. **(C)** Model 3. **(D)** Model 4. **(E)** Model 5.

TABLE 3 | Single tree detection evaluation indicators.

Evaluation indicators	Formula
Recall	$N_{mat} = \frac{N_{match}}{N_r}$
Commission rate	$N_{om} = \frac{N_{error}}{N_a}$
Omission rate	$N_{com} = \frac{N_{leave}}{N_r}$
Accuracy	$M = \frac{N_{match}}{N_{match}+N_{leave}+N_{error}} \times 100$

detection method. Compared with the latest CV model and template matching, the accuracy of the SSD model is 31% higher than the CV model and 8.3% higher than the template matching method. The SSD model obtains a higher score in recall rate, omission rate, and commission rate.

4.3 Economic Plantation

Palm trees are not only a treasure, but are also one of the world's most important sources of oil (De Aguiar et al., 2020). Thailand is one of the main planting bases of palm trees, while China's largest area of palm trees is mainly distributed in Red River county, Yunnan province, China. Palm plantations represent economic plantations.

The single palm tree detection experimental effect is shown in **Figure 9**, a detected box is red and a detected box outside the yellow line is cleared. Palm trees that were detected in error or

TABLE 4 | Single tree inspection accuracy evaluation.

The image number	Method	N_r	N_a	N_{match}	N_{leave}	N_{error}	M	N_{mat}	N_{com}	N_{om}
Campus	SSD	24	25	24	0	1	96.0%	100.0%	0.0%	4.0%
	Watershed	24	73	24	0	49	32.9%	100.0%	0.0%	67.1%
	Local maximum	24	71	24	0	47	33.8%	100.0%	0.0%	66.2%
	Template matching	24	17	17	7	0	70.8%	70.8%	29.1%	0.0%
	CV model	24	21	17	7	4	60.7%	70.8%	29.1%	19.0%
Lychee plantation	SSD	111	107	105	6	2	92.9%	94.6%	5.4%	1.9%
	Watershed	111	149	90	21	59	52.9%	81.1%	18.9%	39.6%
	Local maximum	111	204	76	35	128	31.8%	68.5%	31.5%	62.7%
	Template matching	111	116	104	7	12	84.6%	93.7%	6.3%	10.3%
	CV model	111	127	91	20	36	61.9%	82.0%	18.0%	28.3%

TABLE 5 | Single tree inspection accuracy evaluation.

The image number	Method	N_r	N_a	N_{match}	N_{leave}	N_{error}	M	N_{mat}	N_{com}	N_{om}
Open palm plantation	SSD	164	166	162	2	4	96.4%	98.8%	1.2%	2.4%
	Watershed	164	117	108	56	9	62.4%	65.9%	34.1%	7.7%
	Local maximum	164	213	132	32	81	53.9%	80.5%	19.5%	38.4%
	Template matching	164	157	155	9	2	93.4%	94.5%	5.5%	1.3%
	CV model	164	143	119	45	24	63.3%	72.6%	27.4%	16.8%
Complex background palm plantation	SSD	224	223	220	4	3	97.0%	98.2%	1.8%	1.3%
	Watershed	224	123	118	106	5	51.5%	52.7%	47.3%	4.1%
	Local maximum	224	162	93	131	69	31.4%	41.5%	58.5%	42.6%
	Template matching	224	217	203	21	14	85.3%	90.6%	9.4%	6.5%
	CV model	224	202	111	113	91	35.2%	49.6%	50.4%	45.0%
Dense palm plantation	SSD	148	148	147	1	1	98.7%	99.3%	0.7%	0.7%
	Watershed	148	183	136	0	47	74.3%	91.8%	0.0%	31.7%
	Local maximum	148	162	124	24	38	66.7%	83.8%	16.2%	23.5%
	Template matching	148	152	147	1	5	96.1%	99.3%	0.7%	3.3%
	CV model	148	173	130	18	43	68.1%	87.9%	12.1%	24.9%

missed are flagged. A white marker denotes a mistakenly detected single tree, and a blue marker means the missed detected of a single tree.

4.3.1 Open Economic Plantation

On behalf of the open economic plantation, the open palm plantation is sample plot 1. The test results of the open palm plantation are shown in **Figure 9B** and **Table 5**. The SSD model has the highest accuracy, the highest recall rate, and the lowest omission in the experimental results of the five methods. The accuracy of SSD is 96.4%. The recall rate of SSD is 98.8%. The omission rate of SSD is 1.2%. Two palm trees are missed among 164 reference palm trees in SSD-detected results. Mistakenly, the template matching method identifies two background objects as palm trees. The lowest commission rate reaches 1.3%. the SSD model has the second lowest commission rate. The commission rate of the SSD model is 2.4%. The accuracy of the local maximum is 53.9%. Watershed has the lowest recall rate. The recall rate of watershed is 65.9%. Watershed has the highest omission rate. The highest omission rate of watershed is 34.1%. The commission rate of the local maximum is 38.4%.

4.3.2 Complex Background Economic Plantation

In this study, sample plot 2 is a complex background palm plantation, and the experimental results are shown in **Figure 9D** and **Table 5**. The SSD model has the highest accuracy in the experimental results of the five methods. The accuracy of SSD is 97.0%. The SSD model has the highest recall in the experimental results of the five methods. The recall of SSD is 98.2%. The SSD model has the lowest omission rate in the experimental results of the five methods. The lowest omission rate of SSD is 1.8%. Four palm trees are missed among 224 reference palm trees in SSD-detected results. Mistakenly, the SSD model identifies three background objects as palm trees. The lowest commission rate reaches 1.3%. The SSD model has the lowest commission rate. The accuracy of the local maximum is 31.4% and is the lowest accuracy. The recall rate of the local maximum is 41.5% and is the lowest recall rate. The omission rate of the local maximum is 58.5% and is the highest omission rate. The CV model has the highest commission rate. The commission rate of the CV model is 45.0%. Mistakenly, the CV model identifies 91 background objects as palm trees.

4.3.3 Dense Economic Plantation

Dense palm plantations have many problems, such as the fact that dense tree crowns completely block sunlight, obscure single tree crowns, and have high canopy density, which have increased the difficulty of gaining artificial statistics for single trees. Therefore, it is difficult to improve the accuracy of single tree detection and location in dense palm plantations when using the traditional method. In this study, sample plot 3 is a dense palm plantation, and experimental results are shown in **Figure 9F** and **Table 5**.

The SSD model has the highest accuracy in the experimental results of the five methods. The accuracy of SSD is 98.7%. The SSD model has the highest recall rate in the experimental results of the five methods. The recall rate of SSD is 99.3%. The omission rate of watershed is zero. The omission rate of the SSD model and template matching method is tied for second place with one missed detection among 148 reference palm trees. The omission rate of the SSD model and template matching method is 0.7%. The SSD model has the lowest commission rate in the experimental results of the five methods. Mistakenly, the SSD model identifies one background object as a palm tree. The commission rate of the SSD model is 0.7%. The accuracy of the local maximum is the lowest in the experimental results of the five methods. The accuracy of the local maximum is 66.7%. The lowest recall rate is the local maximum in the experimental results of the five methods. The recall rate of the local maximum is 83.8%. The omission rate of the local maximum is the highest in the experimental results of the five methods. As can be seen from **Table 5**, 24 reference trees are missed. The omission rate of the local maximum is 16.2%. Watershed has the highest commission rate in the experimental results of the five methods. The commission rate of watershed is 31.7%. The watershed mistakes 47 background objects as reference palm trees.

5 DISCUSSION

At present, it takes a lot of manpower and material resources to identify and locate tree species over a large area and in a scattered plantation depending on collecting information or identifying pictures with the naked eye. The manager of an artificial plantation divides it into four stages: young plantation, middle plantation, mature plantation, and overmature plantation. Each stage has different characteristics of individual trees, and machine recognition is more stable and produces fewer errors than human eye recognition. Researchers have developed a variety of methods to extract individual tree information from high-resolution remote sensing images instead of the human eye (Liu et al., 2016; Iqbal et al., 2021). However, the existing single tree detection methods still have shortcomings (Gebreslasie et al., 2011; Millikan et al., 2019; De Aguiar et al., 2020; Dersch et al., 2020).

In this study, the comparison of five experiment groups proves that single tree detection based on SSD has a better effect. According to the problems existing in single tree detection, the experimental research is carried out one by one:

1) Experiments with different canopy densities have been completed, this article makes a comparison between an open plantation and dense plantation.
2) The detection effect of the SSD model has been verified in a plantation area with a complex background.
3) The application effect of the SSD model has been verified in urban single tree detection, including a single tree on a campus, a single tree in an orchard, and a single tree in an economic plantation.

The summary of the five experiment groups shows that the lowest accuracy is the local maximum method, with an average detection accuracy of 43.52%. Local maximum extracts the maximum value of an area. If there is no single tree in the local area, the local maximum value will be wrongly judged as a single tree. If there are multiple single trees in the local area, the local maximum can only identify a single tree with the largest value. Due to the above reasons, local maximum is mediocre in single tree detection. The average detection accuracy of the watershed method is 54.8%. If a tree has too many branches, watershed can easily identify it as two trees. This leads to the high commission rate of watershed. If the distance between two trees is very close, watershed will identify it as a tree, which leads to a high omission rate of detection. The average detection accuracy of the latest template matching method is 86.04%. Template matching is not as detailed as the SSD model in extracting single tree crown features. The average detection accuracy of the CV model method is 57.84%. The CV model combines the advantages of local maximum and watershed. However, the CV model also has the problems of over-segmentation and under-segmentation. The SSD model has the highest average accuracy in the experimental results of the five methods with an average accuracy of 96.32%, because the SSD model can capture single tree canopy features from high-resolution remote sensing images well. The average recall rate of the SSD model is 97.94% which is the highest average recall rate. When category confidence exceeds 0.5, the SSD model identifies an object as a single tree. The average commission rate of the SSD model is the lowest average commission rate, which is 1.7%. The SSD model has the lowest average omission rate in the experimental results of the five methods. The average omission rate of SSD is 2.06%. Therefore, the SSD model has the best single tree detection performance.

Weijia Li et al. (Weijia et al., 2016) applied a convolutional neural network method based on deep learning to the detection of general palm trees, with an average recall rate of 96%. Dong Tianyang et al. (Dong et al., 2018) proposed a single tree detection method based on a progressive cascade convolutional neural network, which is applied to an open plantation and green trees with an average recall rate of 90%. In this study, the SSD model is applied to urban single tree detection, the average recall rate of SSD is 97.94%. So, the SSD model is more comprehensive and detailed in extracting single tree features. The SSD model is better than the common convolutional neural network in single tree detection.

6 CONCLUSION

Automatic identification and location of single trees based on high-resolution remote sensing images is of great significance for ecological park planning, plantation management, and large-scale ecological environment quality monitoring. In this study, the SSD model is applied to single tree detection of high-resolution remote sensing images. The study addresses a number of problems with previous approaches, such as too large trunk branches of a single tree, serious adhesion between crowns, and missing and false detection problems in complex backgrounds. The SSD model is applied to single tree detection in urban plantation. Accurately, the SSD model can capture the canopy features of single trees in high-resolution remote sensing images. The SSD model not only segments single tree crowns, but also recognizes single tree crowns in complex backgrounds. In the process of single tree detection, the SSD model has stronger anti-interference ability and is almost unaffected by the branches of a single tree. The SSD model has excellent performance in all aspects and shows good application potential. This study can be used as a reference for other agricultural and forestry products. It is also hoped that other techniques can be used to describe the crown contour of a single tree.

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

Conceptualization: YZ; Methodology: YZ; Software YZ; Validation: GW; Formal analysis: YZ; Investigation: YZ; Resources: YZ; Data Curation: YZ; Writing -Original Draft: YZ; Writing -Review and Editing: YZ; Visualization: YZ; Funding acquisition: GW.

REFERENCES

Ahl, R., Hogland, J., and Brown, S. (2019). A Comparison of Standard Modeling Techniques Using Digital Aerial Imagery with National Elevation Datasets and Airborne LiDAR to Predict Size and Density Forest Metrics in the Sapphire Mountains MT, USA. *Ijgi* 8 (1), 24. doi:10.3390/ijgi8010024

Belcore, E., Angeli, S., Colucci, E., Musci, M. A., and Aicardi, I. (2021). Precision Agriculture Workflow, from Data Collection to Data Management Using FOSS Tools: An Application in Northern Italy Vineyard. *Ijgi* 10 (4), 236. doi:10.3390/ijgi10040236

Chen, C., Wang, Y., Li, Y., Yue, T., and Wang, X. (2017). Robust and Parameter-free Algorithm for Constructing Pit-free Canopy Height Models. *Ijgi* 6 (7), 219. doi:10.3390/ijgi6070219

Chu, H.-J., Huang, M.-L., Tain, Y.-C., Yang, M.-S., and Höfle, B. (2017). Historic Low Wall Detection via Topographic Parameter Images Derived from Fine-Resolution DEM. *Ijgi* 6 (11), 346. doi:10.3390/ijgi6110346

Congalton, R., Gu, J., Yadav, K., Thenkabail, P., and Ozdogan, M. (2014). Global Land Cover Mapping: A Review and Uncertainty Analysis. *Remote Sensing* 6, 12070–12093. doi:10.3390/rs61212070

De Aguiar, A., Dos Santos, F., Santos, L., and Filipe, V. M. (2020). Vineyard Trunk Detection Using Deep Learning - an Experimental Device Benchmark[J]. *Comput. Elect. Agric.* 175, 1. doi:10.1016/j.compag.2020.105535

Deng, G., Li, Z., and Wu, H. (2010). *Tree crown Recognition Algorithm on High Spatial Resolution Remote Sensing Imagery*. IEEE.

Dersch, S. D. S., Heurich, M. H. M., Krueger, N. K. N., and Krzystek, P.(2020). Combining Graph-Cut Clustering with Object-Based Stem Detection for Tree Segmentation in Highly Dense Airborne Lidar point Clouds[J]. *ISPRS J. Photogrammetry Remote Sensing* 172, 207–222. doi:10.1016/j.isprsjprs.2020.11.016

Dong, T. Y., Jian, Z., Gao, S., and Shen, Y. (2018). Single-Tree Detection in High-Resolution Remote-Sensing Images Based on a Cascade Neural Network[J]. *Int. J. Geo-Information* 7 (9), 367. doi:10.3390/ijgi7090367

Dong, T., Shen, Y., Zhang, J., Ye, Y., and Fan, J. (2019). Progressive Cascaded Convolutional Neural Networks for Single Tree Detection with Google Earth Imagery. *Remote Sensing* 11 (15), 1786. doi:10.3390/rs11151786

Faridatul, M. I., and Wu, B. (2018). Automatic Classification of Major Urban Land Covers Based on Novel Spectral Indices. *Ijgi* 7 (12), 453. doi:10.3390/ijgi7120453

Gebreslasie, M. T., Ahmed, F. B., Van Aardt, J. A. N., and Blakeway, F. (2011). Individual Tree Detection Based on Variable and Fixed Window Size Local Maxima Filtering Applied to IKONOS Imagery for Even-Aged Eucalyptus Plantation Forests. *Int. J. Remote Sensing* 32 (15), 4141–4154. doi:10.1080/01431161003777205

Gougeon, F. A. (1995). A Crown-Following Approach to the Automatic Delineation of Individual Tree Crowns in High Spatial Resolution Aerial Images. *Can. J. Remote Sensing* 21, 274–284. doi:10.1080/07038992.1995.10874622

Guirado, E., Tabik, S., Alcaraz-Segura, D., Cabello, J., and Herrera, F. (2017). Deep-learning versus OBIA for Scattered Shrub Detection with Google Earth Imagery: Ziziphus lotus as Case Study. *Remote Sensing* 9 (12), 1220. doi:10.3390/rs9121220

Iqbal, M. S., Ali, H., Tran, S. N., and Iqbal, T. (2021). Coconut Trees Detection and Segmentation in Aerial Imagery Using Mask Region-based Convolution Neural Network. *IET Comput. Vis.* 15 (6), 428–439. doi:10.1049/cvi2.12028

Jiang, R. R., Wang, C. Y., Shen, L. Q., and Wang, P. (2016). A Method for Lichee's Tree-crown Information Extraction Based on High Spatial Resolution Image. *Trans. Chin. Soc. Agric. Machinery* 47 (09), 17–22. doi:10.6041/j.issn.1000-1298.2016.09.003

Kupidura, P., Osińska-Skotak, K., Lesisz, K., and Podkowa, A. (2019). The Efficacy Analysis of Determining the Wooded and Shrubbed Area Based on Archival Aerial Imagery Using Texture Analysis. *Ijgi* 8 (10), 450. doi:10.3390/ijgi8100450

Larsen, M., and Rudemo, M. (1998). Optimizing Templates for Finding Trees in Aerial Photographs. *Pattern Recognition Lett.* 19 (12), 1153–1162. doi:10.1016/s0167-8655(98)00092-0

Li, M. H., Chen, Y. C., Zhou, S. F., and Xiao, S. Z. (2019). Extractionand Recognition of Individual Tree Information on Aerial Image Data Used Watershed Algorithm. *J. Northeast For. Univ.* 47 (9), 58–62. doi:10.13759/j.cnki.dlxb.2019.09.011

Liu, P., Zhang, H., and Eom, K. B. (2017). Active Deep Learning for Classification of Hyperspectral Images. *IEEE J. Sel. Top. Appl. Earth Observations Remote Sensing* 10, 712–724. doi:10.1109/jstars.2016.2598859

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., and Scott, R. (2016). "SSD: Single Shot Multi Box Detector[J]," in European Conference on Computer Vision (Cham: Springer).

Liu, Y. F., Pan, Y., and Li, H. (2019). Study of crown Information Extraction of Picea Schrenkiana Var. Tianschanica Based on High-Resolution Satellite Remote Sensing Data. *Remote Sensing Land Resour.* 31 (04), 112–119. doi:10.6046/gtzyyg.2019.04.15

Millikan, P. H. K., Silva, C. A., Rodriguez, L. C. E., Oliveira, T. M. d., Carvalho, M. P. d. L. C. e., and Carvalho, S. d. P. C. e. (2019). Automated Individual Tree Detection in Amazon Tropical Forest from Airborne Laser Scanning Data. *Cerne* 25 (3), 273–282. doi:10.1590/01047760201925032630

Mokroš, M., Liang, X., Surový, P., Valent, P., Čerňava, J., Chudý, F., et al. (2018). Evaluation of Close-Range Photogrammetry Image Collection Methods for Estimating Tree Diameters. ISPRS Int. *J. Geo-inf* 7 (3), 93. doi:10.3390/ijgi7030093

Pollock, R. (1996). *J. The Automatic Recognition of Individual Trees in Aerial Images of Forests Based on A Synthetic Tree Crown Image Model[D]*. Vancouver: The University of British Colombia, 172.

Roska, T., and Chua, L. O. (2008). The CNN Universal Machine:[J]. *J. Circuits Syst. Comput.* 12 (04), 377–388. doi:10.1142/S0218126603000921

Sharma, R., Xu, Z., Sugumaran, R., and Oliveira, S. (2016). Parallel Landscape Driven Data Reduction & Spatial Interpolation Algorithm for Big LiDAR Data. *Ijgi* 5 (6), 97. doi:10.3390/ijgi5060097

Wang, L., Gong, P., and Biging, G. S. (2004). Individual Tree-Crown Delineation and Treetop Detection in High-Spatial-Resolution Aerial Imagery. *Photogramm Eng. Remote Sensing* 70 (3), 351–357. doi:10.14358/pers.70.3.351

Wang, Y., Jiang, T., Liu, J., Li, X., and Liang, C. (2020). Hierarchical Instance Recognition of Individual Roadside Trees in Environmentally Complex Urban Areas from UAV Laser Scanning Point Clouds. *Ijgi* 9 (10), 595. doi:10.3390/ijgi9100595

Weijia, L., Haohuan, F., Le, Y., and Cracknell, A. (2016). Deep Learning Based Oil Palm Tree Detection and Counting for High-Resolution Remote Sensing Images[J]. *Remote Sensing* 9 (1), 22. doi:10.3390/rs9010022

Yu, X. Z., Wang, R. R., and Cheng, W. J. (2018). Improved Watershed Algorithm Is Used in the Segmentation of Uav Remote Sensing Image Canopy. *J. fujian Agric. For. Univ.* 47 (04), 428–434. doi:10.13323/j.cnki.j.fafu(nat.sci.).2018.04.008

Zhang, L., Liu, P., Zhao, L., Wang, G., Zhang, W., and Liu, J. (2021). Air Quality Predictions with a Semi-supervised Bidirectional LSTM Neural Network. *Atmos. Pollut. Res.* 12 (1), 328–339. doi:10.1016/j.apr.2020.09.003

Zhang, M., Du, H., Mao, F., Zhou, G., Li, X., Dong, L., et al. (2020). Spatiotemporal Evolution of Urban Expansion Using Landsat Time Series Data and Assessment of its Influences on Forests. *Ijgi* 9 (2), 64. doi:10.3390/ijgi9020064

Zhang, N., Zhang, X. L., and Ye, L. (2014). Tree crown Extraction of High Resolution Remote Sensing Image Segmentation Based on Improved Peak Climbing Method. *Trans. Chin. Soc. Agric. Machinery* 45 (12), 294–300. doi:10.6041/j.issn.1000-1298.2014.12.042

Zhao, L., Zeng, Y., Liu, P., and Su, X. (2020). Band Selection with the Explanatory Gradient Saliency Maps of Convolutional Neural Networks. *IEEE Geosci. Remote Sensing Lett.* 17 (12), 2105–2109. doi:10.1109/lgrs.2020.3012140

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.

frontiers in Environmental Science

ORIGINAL RESEARCH
published: 03 December 2021
doi: 10.3389/fenvs.2021.780434

Combined Wavelet Transform With Long Short-Term Memory Neural Network for Water Table Depth Prediction in Baoding City, North China Plain

Zehua Liang[1], *Yaping Liu*[1,2]*, *Hongchang Hu*[3], *Haoqian Li*[1], *Yuqing Ma*[1] and *Mohd Yawar Ali Khan*[4]

[1]College of Resource Environment and Tourism, Capital Normal University, Beijing, China, [2]Beijing Laboratory of Water Resources Security, Capital Normal University, Beijing, China, [3]Department of Hydraulic Engineering, State Key Laboratory of Hydroscience and Engineering, Tsinghua University, Beijing, China, [4]Department of Hydrogeology, Faculty of Earth Sciences, King Abdulaziz University, Jeddah, Saudi Arabia

OPEN ACCESS

Edited by:
Peng Liu,
Institute of Remote Sensing and Digital Earth (CAS), China

Reviewed by:
Hao Zhang,
Henan Agricultural University, China
Qiulan Zhang,
China University of Geosciences, China

***Correspondence:**
Yaping Liu
y.liu@cnu.edu.cn

Specialty section:
This article was submitted to Environmental Informatics and Remote Sensing, a section of the journal Frontiers in Environmental Science

Received: *21 September 2021*
Accepted: *08 November 2021*
Published: *03 December 2021*

Citation:
Liang Z, Liu Y, Hu H, Li H, Ma Y and Khan MYA (2021) Combined Wavelet Transform With Long Short-Term Memory Neural Network for Water Table Depth Prediction in Baoding City, North China Plain. Front. Environ. Sci. 9:780434. doi: 10.3389/fenvs.2021.780434

Accurate estimation of water table depth dynamics is essential for water resource management, especially in areas where groundwater is overexploited. In recent years, as a data-driven model, artificial neural networks (NNs) have been widely used in hydrological modeling. However, due to the non-stationarity of water table depth data, the performance of NNs in areas of over-exploitation is challenging. Therefore, reducing data noise is an essential step before simulating the water table depth. This research proposed a novel method to model the non-stationary time series data of water table depth through combing the advantages of wavelet analysis and Long Short-Term Memory (LSTM) neural network (NN). A typical groundwater over-exploitation area, Baoding, North China Plain (NCP), was selected as a study area. To reflect the impact of anthropogenic activities, the variables harnessed to develop the model includes temperature, precipitation, evaporation, and some socio-economic data. The results show that decomposing the time series of the water table depth into three sub-temporal components by Meyer wavelets can significantly improve the simulation effect of LSTM on the water table depth. The average NSE (Nash-Sutcliffe efficiency coefficient) value of all the sites increased from 0.432 to 0.819. Additionally, a feedforward neural network (FNN) is used to compare forecasts over 12-months. As expected, wavelet-LSTM outperforms wavelet-FNN. As the prediction time increases, the advantages of wavelet-LSTM become more evident. The wavelet-LSTM is satisfactory for forecasting the water table depth at most in 6 months. Furthermore, the importance of input variables of wavelet-LSTM is analysed by the weights of the model. The results indicate that anthropogenic activities influence the water table depth significantly, especially in the sites close to the Baiyangdian Lake, the largest lake in the North China Plain. This study demonstrates that the wavelet-LSTM model provides an option for water table depth simulation and predicting areas of over-exploitation of groundwater.

Keywords: water table depth, long short-term memory neural network, wavelet transform, over-exploitation area, feedforward neural network, North China Plain

1 INTRODUCTION

Groundwater, an important water resource, is being over-exploited due to the rapid population growth and economy, especially in arid and semi-arid areas. Excessive exploitation of aquifers has caused severe land subsidence, increased groundwater recharge area, and led to pollution and salinization of groundwater (Li et al., 2020). The NCP, one of the most heavily influenced regions through anthropogenic activities, has emerged as the largest groundwater depression cone in the world (Tang et al., 2013; Chen et al., 2020). Previous studies have shown that the water table in the NCP exhibited a long-term decline rate of −17.8 ± 0.1 mm/yr from 1971–2015 (Gong et al., 2018).

At present, physical models, such as MODFLOW (Modular Ground-Water Flow Model) (Xu et al., 2012; Lachaal et al., 2012; Xiang et al., 2020), HYDRUS (Huang et al., 2016), GMS (Groundwater Modeling System) (Roy et al., 2015), have been widely used in groundwater resources evaluation and management. For example, Xu et al. (2012) integrated the SWAP (Soil–Water–Atmosphere–Plant) package into MODFLOW to simulate the regional groundwater flow system. Xiang et al. (2020) evaluated the balance between groundwater protection with crop production based on the results of MODFLOW combined with DSSAT (Decision Support System for Agrotechnology Transfer). Maihemuti et al. (2021) employed HYDRUS to evaluate the effects of groundwater on plant distribution. However, these physical models usually require boundary conditions and a large number of hydraulic parameters for calibration. When hydrogeological data is lacking, the data-driven model based on NNs shows advantages.

Over the past decades, many studies have applied NN methods, such as FNN, ANFIS (Adaptive-network-based fuzzy inference system) to predict water table or water table depth (Coppola et al., 2003; Daliakopoulos et al., 2005; Nayak et al., 2006; Altunkaynak, 2007; Chen et al., 2010; Taormina et al., 2012; Nourani and Mousavi, 2016). Compared to physics models, the data required by NNs is easier to collect and quantify (Mohanty et al., 2013). In addition, some studies have shown that the simulation effect of NN is better than that of numerical model in certain scenarios (Altunkaynak, 2007; Mohanty et al., 2013). For example, Zealand et al. (1999) employed FNN to predict short-term streamflow. In their study, the WIFFS model (Winnipeg Flow Forecasting System) was used as a conventional numerical model for a contrastive study. They found that the average RMSE (root mean square error) of about 52.8 m^3/s was obtained *via* FNN, which was better than obtained *via* WIFFS (64.5 m^3/s). Mohanty et al. (2013) evaluated the performance of MODFLOW and FNN in the short-term prediction of water table. Their study show that the NSE value obtained by NN varied in the range of 0.90 – 0.96, up from 0.55 – 0.95 by MODFLOW.

Nevertheless, these traditional NN methods may not deal with time series data effectively because they cannot preserve previous information (Zhang et al., 2018). To deal with time series data in groundwater modelling, some researchers employed Recurrent Neural Network (RNN), as its output can be associated with previous state of the network (Coulibaly et al., 2001; Chang et al., 2014). However, due to the disappearance of the gradient, the performance of RNN in long-term backpropagation is limited. Therefore, a special RNN, LSTM, is widely used to solve long-term sequence prediction problems, including some hydrological domains. For example, Zhang et al. (2018) used the LSTM to predict the water table depth in Hetao Irrigation District, and compared the results with traditional FNN. They found that LSTM's prediction is much more accurate than that of FNN. They also pointed out that the single hidden layer is better than the double hidden layer. Hewage et al. (2021) found that LSTM performs better than numerical models in weather forecasting, but numerical models have obvious advantages in long-term prediction. Kratzert et al. (2018) used the LSTM network to simulate precipitation in multiple watersheds. They found that in the case of insufficient data, previous training parameters can be recorded and used to simulate the precipitation in other watersheds to achieve satisfactory results.

Although NNs have received a lot of attention in hydrological modeling, NN may not adequately handle nonlinear and non-stationary data (Ebrahimi and Rajaee, 2017). Due to the high autocorrelation of the time series data, NNs tend to produce a forecast that is very similar to the last observed data (de Vos and Rientjes, 2005). The prediction results of NNs are always continuations of historical trends and do not accurately reflect high-frequency and irregular changes for multi-step predictions (Zhang et al., 2021). In addition, most of the measured and observed hydrological time series contain noise. Therefore, eliminating data noise to manage non-stationary data better is essential in hydrological modeling (Nourani and Mousavi, 2016).

As an effective data preprocessing method, wavelet analysis provides a time-frequency representation of signals with many different periods in the time domain. It can decompose time series data into approximate and detailed parts to extract potential information from noisy data (Daubechies, 1990). The combination of wavelet transform analysis, and NN has been used in various fields of hydrology, including streamflow prediction (Tiwari and Chatterjee, 2010; Adamowski and Sun, 2010; Nanda et al., 2016), precipitation prediction (Nourani et al., 2009) and drought forecasting (Kim and Valdés, 2003). Furthermore, wavelet transform combined with an NN also has important applications in groundwater modeling. For example, Gorgij et al. (2017) used an NN based on wavelet analysis and a genetic program model to predict the water table in the eastern plain of Iran. Ebrahimi and Rajaee (2017) used NNs, multiple linear regression and support vector regression combined with wavelet analysis to predict the monthly water table of the Qom plain in Iran and have found

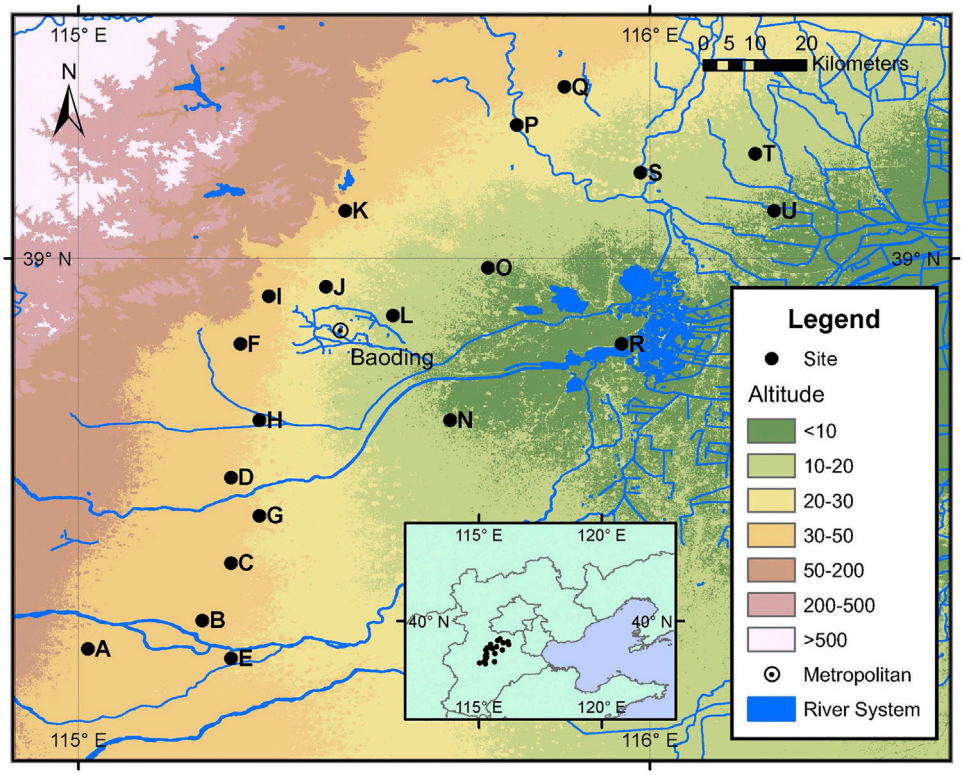

FIGURE 1 | Location of all sites in the study area.

that the wavelet transform analysis improved the prediction effect of these models. Therefore, considering the periodicity and randomness of the water table time series events, the wavelet-based NN model can be used as an efficient method to deal with nonlinear and non-stationary water table time series.

This study focuses on combing wavelet analysis with NNs to establish a novel data-driven model for non-stationary time series data of water tables in areas of over-exploitation. Furthermore, the influence of various factors on water table is discussed through analysing the importance of input variables, which provides a reference for local groundwater resource management. The city of Baoding in the NCP was chosen as the study area. The specific objectives of this study are: 1) evaluating the simulation effect of wavelet-LSTM model, 2) forecasting water table over the 12 months using the wavelet-LSTM model, 3) analysing the contribution of each variable to the changes in water table based on the weight of the NN and the land use distribution.

2 DATA AND METHODOLOGY

2.1 Study Area and Data Sources

The study area is located in Baoding City, Hebei Province, in the middle of the NCP, between 113°40′—116°20′E, 38°10′—40°00′N, This region belongs to a temperate continental monsoon climate zone. The average annual precipitation is about 500 mm, and the annual evaporation is about 1,430 mm. Over the past 40 years, the coldest month (average temperature −2.7°C), and the lowest monthly average precipitation (2.4 mm) occurred in January. The hottest month (average temperature 27.1°C), and the highest monthly average precipitation (155.5 mm) occurred in July. We obtained monthly water table depth data from 20 observation wells from 2000 to 2016 from the local hydrological bureau. The locations are shown in **Figure 1**.

The study area mainly includes alluvial fans and alluvial plains, and the lithology is composed of gravels, sands, silts, and silty clays etc. Due to the scarcity of surface water resources in the study area, groundwater is the leading water resource. Agriculture and industries as a major grain producer and steel producer, respectively, in China, accounts for the most significant proportion of water consumption. Studies have shown that groundwater is almost the only source of irrigation water (Xiao et al., 2017). In addition, Hebei Province has historically been the largest steel-producing province in China, with a steel output of 2.184 billion tons in the past decade, accounting for 23% of the country's total production. As a high water consumption industry, the development of the steel industries has contributed significantly to the depletion of groundwater in the region.

As shown in **Figure 2**, steel prices and API (Agricultural Price Index) negatively affect the depth of the water table. The three peaks appeared in 2005, 2009, and 2011, respectively, corresponding to the three valleys of the water table. Generally, the periods of high prices correspond to the periods of strong demand. In other words, during

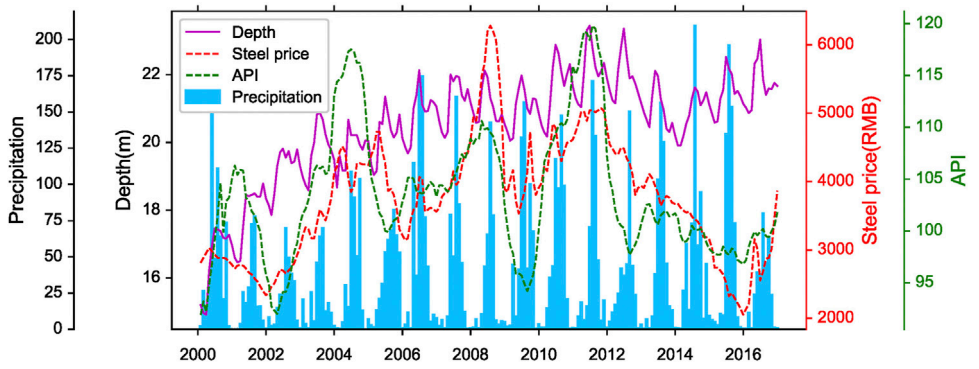

FIGURE 2 | Comparison of the average water table depth and precipitation, steel price and API.

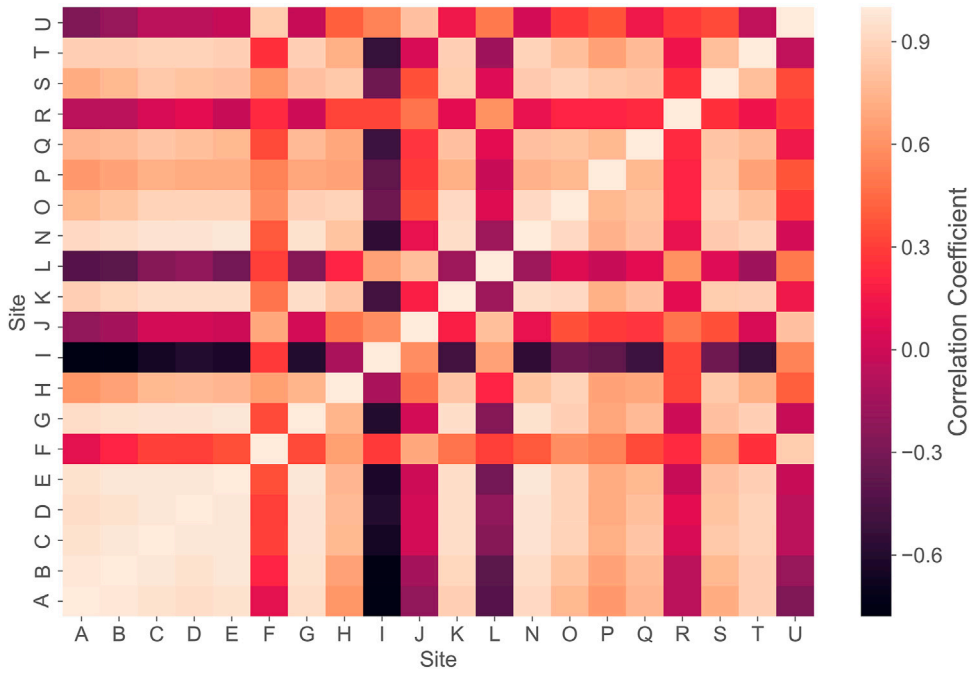

FIGURE 3 | Correlation heatmap of all sites.

high prices, the production activities of steel and agriculture increased significantly, resulting in a large consumption of water, which in turn causes the water table to fall.

2.2 The LSTM Model

NN is a model that simulates the biological brain to achieve the artificial intelligence effect. The basic NN consists of an input layer, an output layer and a hidden layer. Each neuron is connected to the other by weights, and the training process is the process of updating weights. The NN activation function requires a nonlinear function that maps the input to a finite interval that determines whether the neuron is activated.

FNN is a simple NN that is widely used. All layers of the FNN are dense layers, and the parameters are propagated unidirectional from the input layer to the output layer and are updated by the error backpropagation algorithm. The NN parameters are the weights on each connection, and these weights are obtained by learning processes. Backpropagation algorithms based on the gradient descent method are often used to train NNs. In a NN, if we associate the hidden-layer state with each instant, we call it RNN. RNN is generally used for processing time series data because it uses information from the previous moment in each step. In this paper, the activation function we adopted between hidden layers is "tanh." However, the calculations of the gradient of networks weight is essentially a continuous product operation. The gradients tend to zero or infinity exponentially with the length of the sequence increasing. It is the vanishing and exploding gradient problems.

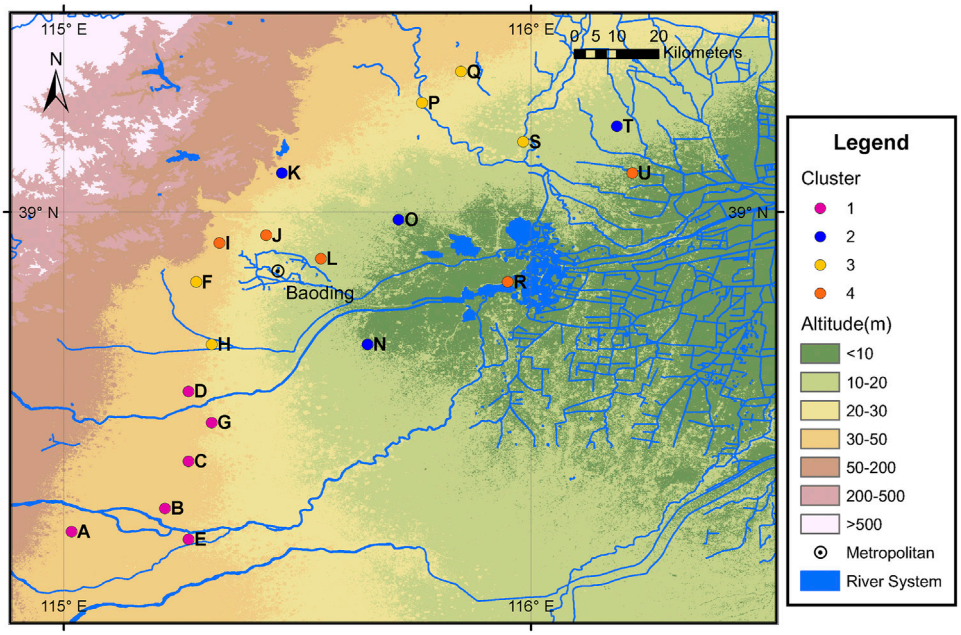

FIGURE 4 | Clustering of all sites.

In this case, the model will ignore the previous state information. To solve this problem, the LSTM NN has been proposed (Hochreiter and Schmidhuber, 1997). A forget gate is added to the LSTM to manage the network's "memory" to remember the model's state for a long time. The following equation describes the computational procedure of LSTM:

$$f_t = \sigma(b_f + U_f s_t + V_f h_{t-1}) \quad (1)$$

$$i_t = \sigma(b_i + U_i s_t + V_i h_{t-1}) \quad (2)$$

$$o_t = \sigma(b_o + U_o s_t + V_o h_{t-1}) \quad (3)$$

$$c_t = tanh(b_c + U_c s_t + V_c h_{t-1}) \quad (4)$$

$$m_t = f_t \otimes m_{t-1} + i_t \otimes c_t \quad (5)$$

where, c_t is calculated by s_t and h_{t-1}, forget gate and input gate are employed to control m_t. In RNN, h_t is the state of its hidden layer, while in LSTM, m_t (memory) is added to remember its long-term state and c_t to represent its cell state of the current input. In this study, "sigmoid" is employed as activation function of its forget gate.

2.3 Discrete Wavelet Transform

The idea of wavelet transform is to decompose the original sequence into different subsequences to provide detailed information about the multi-scale properties of time series. The superior function of wavelet transforms to reflect information on the time, location and frequency of a signal simultaneously (Cohen and Kovacevic, 1996). Wavelet transform is generally divided into continuous wavelet transform (CWT) and discrete wavelet transform (DWT). Due to information redundancy, DWT is usually recommended in hydrological forecasting (Quilty and Adamowski, 2018; Rajaee et al., 2019). Unlike CWT, DWT uses a specific subset of all zoom and translation values. In DWT, the original sequence is decomposed by a scale function for approximating the original sequence, and the wavelet function is used to describe the details of the original sequence. The scale function and wavelet function of the DWT decomposition can be defined as follows:

$$\phi_{j,k}(t) = 2^{j/2}\phi(2^j t - k) \quad (6)$$

$$\psi_{j,k}(t) = 2^{j/2}\psi(2^j t - k) \quad (7)$$

where $\phi(t)$ is scale function, $\psi(t)$ is wavelet function, j and k are dilation factor and translation factor respectively.

Meanwhile, let V_j, W_j is a space spanned by $\phi_{j,k}(t)$, $\psi_{j,k}(t)$ respectively, W_j is a orthogonal complement space of V_j:

$$V_0 = V_1 \subset V_2 \subset V_3 \cdots \subset V_{j-1} \subset V_j \subset \cdots = W_{j-1} \oplus V_{j-1} \quad (8)$$

Thus, each V_j can be decompose to W_{j-1} and V_{j-1}:

$$V_j = W_{j-1} \oplus W_{j-2} \oplus W_{j-3} \oplus \cdots \oplus W_0 \oplus V_0 \quad (9)$$

In this study, DWT is applied to decompose the water table time series. The processed sub-time series are input to the LSTM model with meteorological data, socio-economic data as variables.

2.4 Data Processing

The input format of the LSTM or FNN is a multidimensional tensor. The input data is typically preprocessed in a three-dimensional tensor format like (*samples*, *timesteps*, *features*) for time series data. In this study, air temperature (K), precipitation (mm), evapotranspiration (mm) data, normalized difference vegetation index (NDVI) data,

TABLE 1 | Comparison of performance between LSTM model and wavelet-LSTM model.

Site		NSE		R		RMSE(m)	
		Training	Testing	Training	Testing	Training	Testing
A	LSTM	0.978	0.575	0.989	0.931	0.514	0.647
	wavelet-LSTM	0.995	0.887	0.998	0.949	0.241	0.335
B	LSTM	0.987	0.763	0.994	0.971	0.345	0.323
	wavelet-LSTM	0.996	0.894	0.998	0.952	0.202	0.217
C	LSTM	0.956	0.542	0.986	0.908	0.630	0.590
	wavelet-LSTM	0.989	0.797	0.995	0.932	0.321	0.392
D	LSTM	0.957	0.411	0.982	0.900	0.549	1.079
	wavelet-LSTM	0.989	0.901	0.995	0.953	0.277	0.443
E	LSTM	0.974	0.921	0.992	0.972	0.482	0.201
	wavelet-LSTM	0.997	0.943	0.998	0.973	0.173	0.170
F	LSTM	0.934	0.668	0.976	0.855	1.076	0.919
	wavelet-LSTM	0.983	0.828	0.992	0.919	0.551	0.660
G	LSTM	0.950	0.685	0.977	0.908	0.568	0.503
	wavelet-LSTM	0.980	0.954	0.990	0.978	0.361	0.192
H	LSTM	0.943	0.424	0.972	0.773	0.727	1.122
	wavelet-LSTM	0.960	0.732	0.980	0.906	0.613	0.765
I	LSTM	0.938	0.441	0.970	0.814	1.836	1.638
	wavelet-LSTM	0.975	0.866	0.989	0.952	1.163	0.802
J	LSTM	0.837	–0.440	0.918	0.766	0.622	0.808
	wavelet-LSTM	0.930	0.773	0.965	0.919	0.406	0.321
K	LSTM	0.946	0.168	0.972	0.714	0.678	0.710
	wavelet-LSTM	0.979	0.831	0.990	0.912	0.421	0.319
L	LSTM	0.818	0.593	0.905	0.881	0.518	0.678
	wavelet-LSTM	0.931	0.939	0.965	0.969	0.318	0.263
N	LSTM	0.912	0.430	0.965	0.741	0.700	0.626
	wavelet-LSTM	0.985	0.784	0.993	0.890	0.293	0.385
O	LSTM	0.936	0.583	0.967	0.766	0.699	0.735
	wavelet-LSTM	0.969	0.800	0.985	0.897	0.488	0.509
P	LSTM	0.860	0.554	0.941	0.758	0.987	0.697
	wavelet-LSTM	0.957	0.859	0.979	0.928	0.544	0.392
Q	LSTM	0.870	0.487	0.939	0.881	0.675	0.691
	wavelet-LSTM	0.932	0.747	0.966	0.893	0.490	0.485
R	LSTM	0.440	–0.417	0.664	0.226	1.057	1.981
	wavelet-LSTM	0.711	0.523	0.844	0.730	0.760	1.149
S	LSTM	0.904	0.143	0.958	0.755	0.816	1.464
	wavelet-LSTM	0.961	0.816	0.981	0.906	0.518	0.679
T	LSTM	0.876	0.516	0.940	0.735	1.163	1.831
	wavelet-LSTM	0.962	0.749	0.982	0.869	0.639	1.319
U	LSTM	0.942	0.602	0.977	0.788	0.347	0.518
	wavelet-LSTM	0.978	0.763	0.990	0.889	0.214	0.399

agricultural price index (API) and steel price data are harnessed as input variables of LSTM of FNN models. Air temperature, precipitation and evapotranspiration data in NetCDF (Network Common Data Form) format are resampled to monthly data. Monthly API and steel pricing data were collected from the website of the National Bureau of Statistics of China (http://www.stats.gov.cn/tjsj/), while Moderate Resolution Imaging Spectroradiometer (MODIS) provided NDVI data. Also, because the variables are different in order of magnitude, to make their scales uniform, the data has been normalised through the following equation to be a dimensionless value between 0 and 1:

$$x_{scaled} = \frac{(x - x_{min})(max - min)}{(x_{max} - x_{min})} + min \tag{10}$$

where x_{scaled} is normalized data, x_{min} and x_{max} represent minimum and maximum value of the data respectively.

2.5 Model Evaluation

The NSE, RMSE (the root mean square error) and R (correlation coefficient) are harnessed to evaluate the performance of the model:

$$NSE = 1 - \frac{\sum_{i=1}^{n}(O_i - P_i)^2}{\sum_{i=1}^{n}(O_i - \bar{O})^2} \tag{11}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(O_i - P_i)^2}{n}} \tag{12}$$

$$R = \frac{\sum_{i=1}^{n}(O_i - \bar{O})(P_i - \bar{P})}{\sqrt{\left[\sum_{i=1}^{n}(O_i - \bar{O})^2\right]\left[\sum_{i=1}^{n}(P_i - \bar{P})^2\right]}} \tag{13}$$

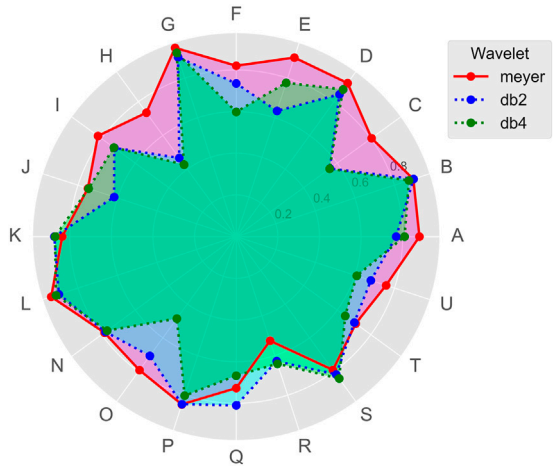

FIGURE 5 | NSE value of three models using db2, db4, and meyer wavelet respectively.

where O_i is observed value at time i, P_i is predicted value at time i, $\bar{O}$ is the mean value of O_i, $\bar{P}$ is the mean value of P_i. The NSE value range from negative infinity to 1 while the correlation coefficient R from −1 to 1. The prediction is ideal if NSE and correlation coefficient are close to 1 or RMSE to 0.

3 RESULTS AND DISCUSSIONS

3.1 The LSTM Model

The correlation between the two sites was examined to reduce the noise influence of the water table data as much as possible. According to the correlated heat map of the water table depth at each site (**Figure 3**), the 20 sites are divided into four clusters (**Figure 4**). Data from the first 14 years is used for training purposes, and the data from the next 3 years is used for testing purposes. As each cluster, model's output represents the water table depth prediction of all sites included in this cluster.

Table 1 shows the NSE, RMSE, and correlation coefficients of all sites during the training and testing periods using the LSTM model and wavelet-LSTM model. It is evident that the NSE of all sites during the training period is greater than 0.8, and the correlation coefficients are greater than 0.9. During the testing period, the NSE at all sites was significantly lower than the NSE during the training period and was even negative at sites J, K, R, and S. The results indicate that quite a significant overfitting phenomenon occurs. From a spatial point of view, the simulation performance of the densely distributed area (cluster 1) is better, while the sparsely distributed area (cluster 4) has poor simulation performance. In addition, the closer to the lake, the weaker the results are. It may be due to the fact that the water table depth near the lake is strongly affected by the lake. However, hydrological data for the lake is lacking.

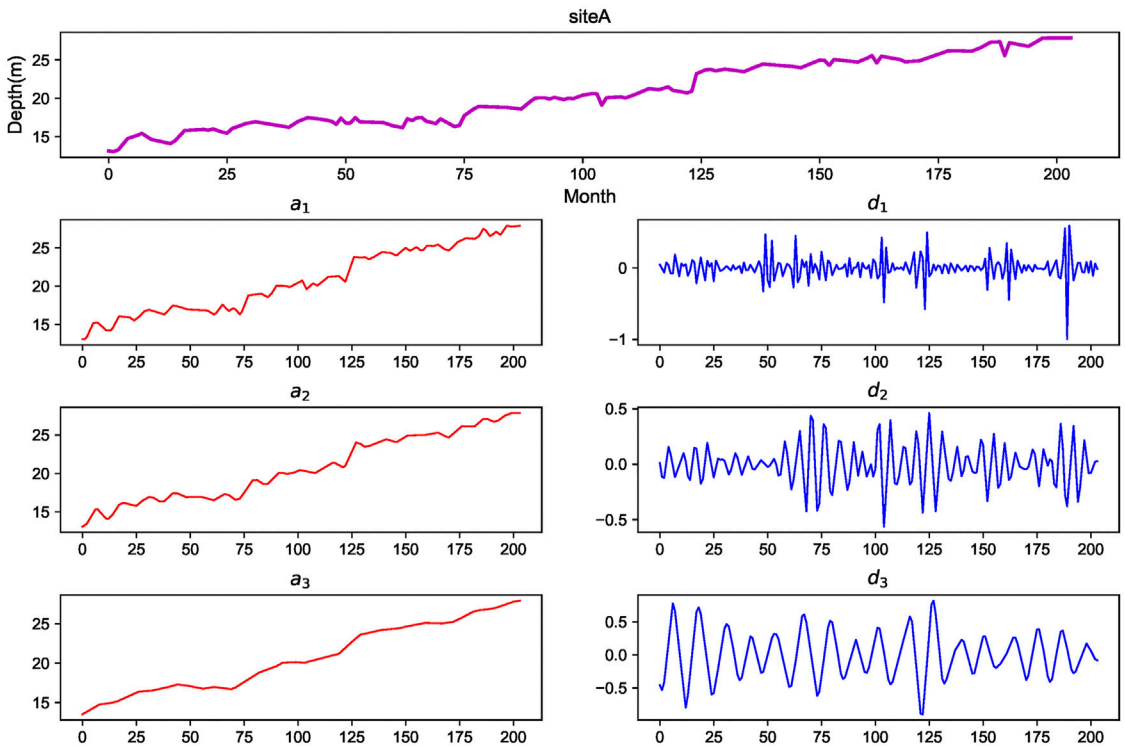

FIGURE 6 | Decomposition of the time series of site A.

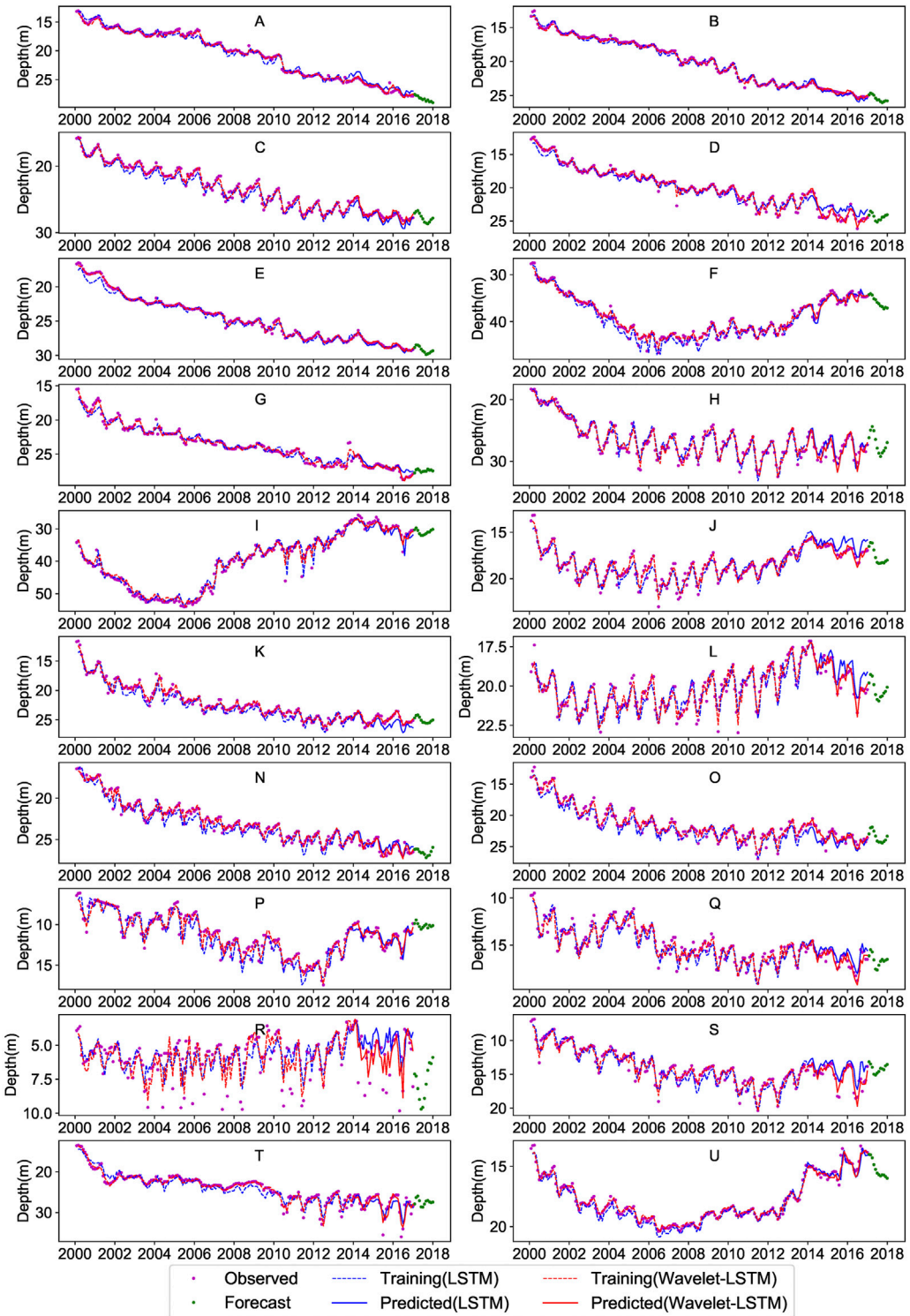

FIGURE 7 | Comparison of observed, simulated and forecast water table depth applied the single LSTM model and the wavelet-LSTM model at each site.

3.2 The Wavelet-LSTM Model

The processed sub-time series decomposed using wavelets is used as the input of the LSTM model to improve the output. On the selection of wavelet function, Daubechies family wavelet (db*N*, *N* refers to the number of vanishing moments) (Khan et al., 2020) and Meyer wavelet (Freire et al., 2019)are commonly used; Haar wavelet are also often used for comparative studies (Liu et al., 2012; Ebrahimi and Rajaee, 2017).

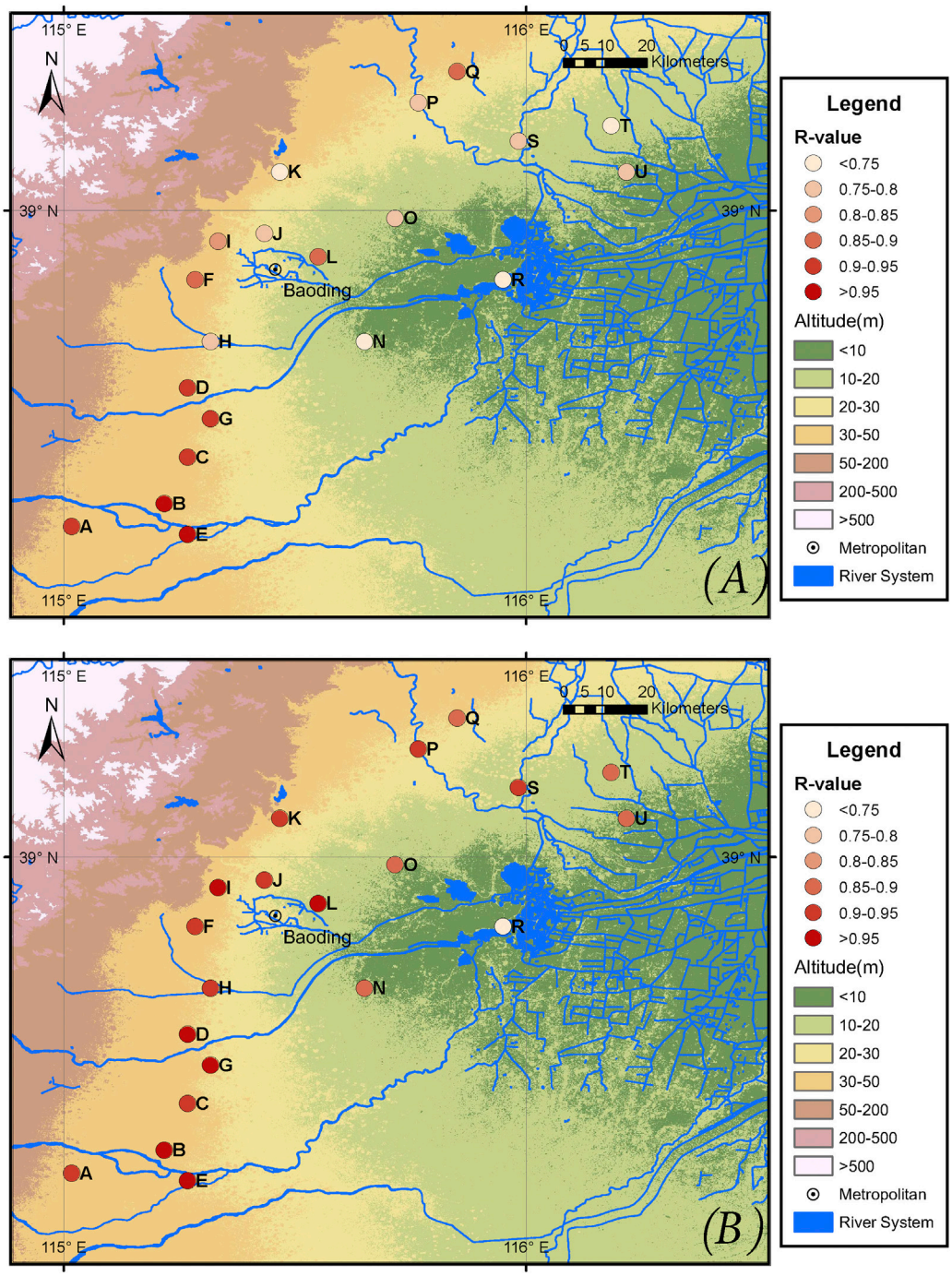

FIGURE 8 | Spatial distribution of *R* values using LSTM model and wavelet-LSTM model during the testing period: **(A)** LSTM **(B)** wavelet-LSTM.

Maheswaran and Khosa (2012) proposed that a wavelet with a compact support is suitable for processing time series with short memory with short-duration transient features while wavelets with wider support for time series with long term features. Nourani et al. (2009) used db4 and Meyer wavelet to decompose the time series with two decomposition levels to simulate monthly precipitation data. Gorgij et al. (2017) used a db4 wavelet to decompose the monthly water table data with two levels. Nanda et al. (2019) used a db2 wavelet to decompose the daily time series with five levels to simulate the daily streamflow data. Therefore, the wavelet function and levels of decomposition should be carefully determined according to different conditions.

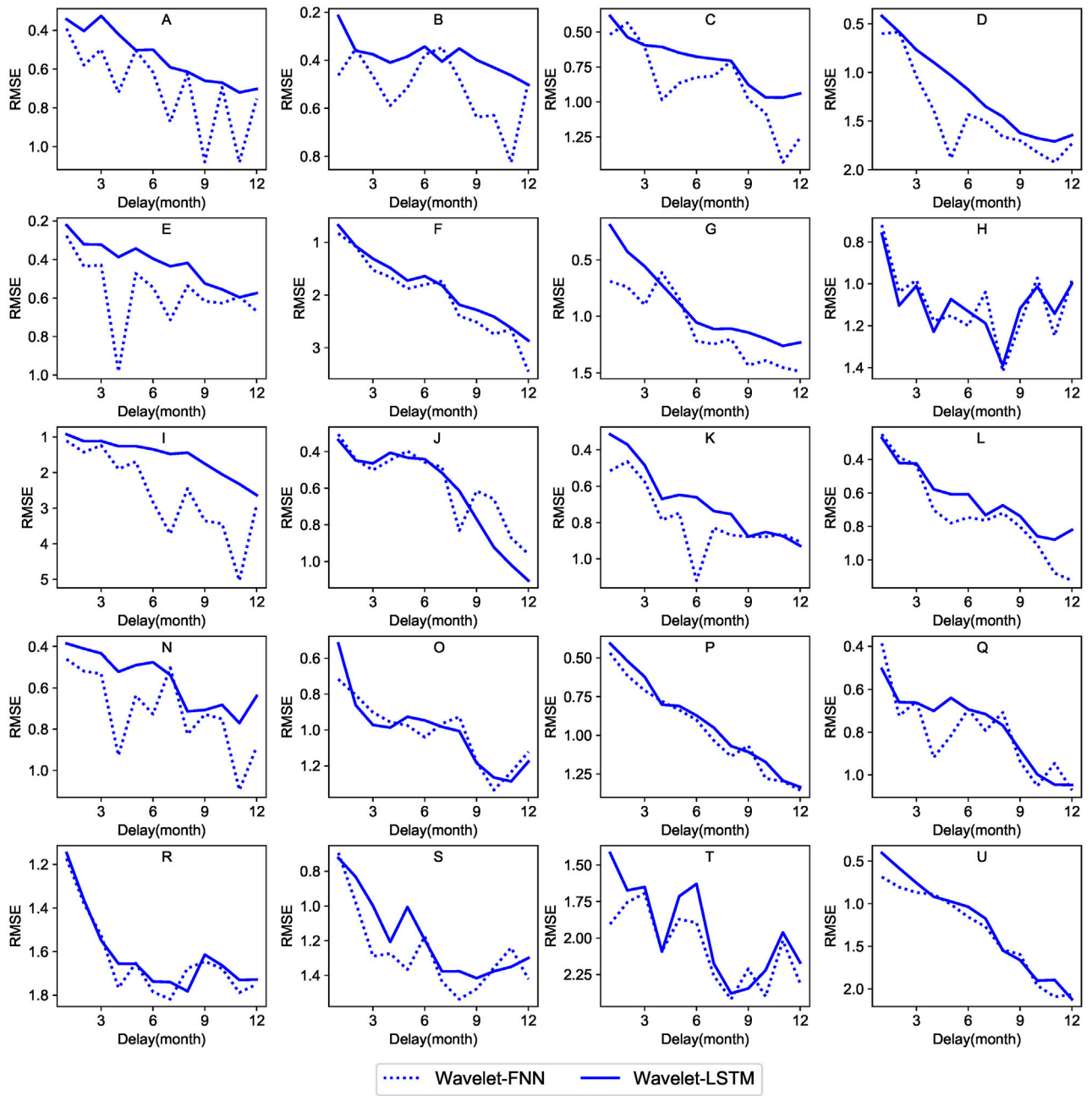

FIGURE 9 | Comparison of RMSE values of water table depth using wavelet-LSTM and wavelet-FNN models with 1–12 months delays.

In this study, db2, db4, and Meyer wavelet are used for comparison. The NSE value of the three wavelets used by the model in the testing phase are shown in **Figure 5**. It can be seen that although db2 and db4 wavelets may be close or even slightly better than Meyer wavelet at some sites, the advantages of Meyer wavelets are evident on most sites. It should be noted that, as wavelet components are input into the model as variables, the decomposition level could not be unduly high. Because it is not practical to apply the network effectively when the number of training samples is limited while the dimension of the feature space is large (Liu et al., 2017). The sub-time series of the data of site A decomposed by Meyer wavelets are shown in **Figure 6**. The component d_3 (three decomposition levels) can be seen to have a significant periodic variation feature. As a result, three levels of decomposition were used.

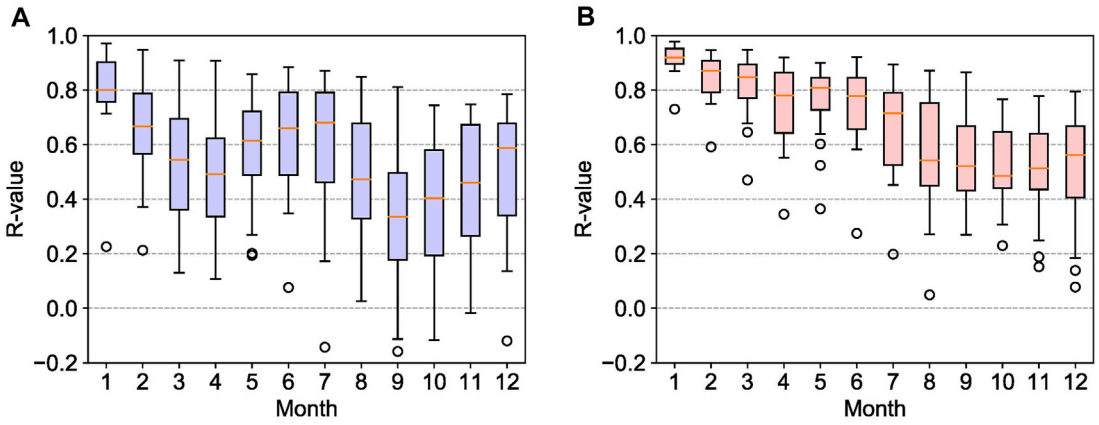

FIGURE 10 | Comparison of *R* values between LSTM model and wavelet-LSTM model during the testing period for 1–12 months forecast: **(A)** LSTM **(B)** wavelet-LSTM.

As shown in **Table 1**, the performance of the wavelet-LSTM model is significantly better than that of the single LSTM model. The simulation results of both models during training and testing periods are shown in **Figure 7**. During the training period, the LSTM model without wavelet transform does not accurately simulate the water table under extreme conditions (peaks and troughs) and is subject to overfitting during the testing process. For example, when a single LSTM model is used, the NSE value for sites J, K, and S are 0.837, 0.946, and 0.904, respectively, while in the testing phase, the NSE value are −0.440, 0.168, and 0.143, respectively. After using the wavelet-LSTM model, the NSE value reached 0.773, 0.831, and 0.816, respectively for sites J, K, and S. It should be noted that site R is close to Baiyangdian Lake, the study area's primary surface water body. Despite the lacking of hydrological data for Baiyangdian Lake, utilizing the wavelet-LSTM model enhanced the simulation effect of the R site from −0.417 to 0.523. The results indicate that the phenomenon of overfitting was significantly improved. From the comparison of **Figures 8A,B**, it can be seen that compared to a single LSTM, the simulation effect of each site has been considerably improved under the LSTM model coupled by wavelets. The delayed response of water table depth data to weather conditions and our inability to obtain socio-economic data with higher spatial resolution makes it impossible to use a single LSTM model to capture the exact characteristics of the water table series accurately. However, the wavelet transform is very suitable for dealing with the non-stationary and stochastic nature of groundwater variability.

As described in **Section 2.2**, unlike LSTM-NNs, FNN has no memory and cannot record the state of individual inputs. Therefore, the wavelet transform is combined with FNN (wavelet-FNN) and compared with wavelet-LSTM on the water table forecasting effect. **Figure 9** shows the RMSE comparison of the combined wavelet transform with FNN and LSTM-NN over the next 1–12 months. As expected, FNN is not as efficient as the LSTM model for time series data. Although the RMSE increases with prediction time increases, wavelet-LSTM still performs better than wavelet-FNN simulation for almost all sites. This phenomenon is more evident as the prediction time increases, reflecting the features of the wavelet-LSTM, which can memorise information for a long time.

It should be noted that the underground funnels are mainly distributed in the southwest of the study area, namely sites A, B, C, D, E, and G. For these sites, the advantages of wavelet-LSTM are particularly evident, and the RMSE of wavelet-LSTM is even less than half of wavelet-FNN in individual sites, illustrating the applicability of wavelet-LSTM in overexploited areas. Therefore, it can be concluded that LSTM-NN is better than FNN in long-term prediction in areas where anthropogenic activities strongly influence groundwater. It further shows that the wavelet-LSTM model can effectively simulate the non-stationary water table variation in the overexploited area.

3.3 Forecast of the Future Water Table Depth

Given that meteorological data, socio-economic data are unknown, we need to use the present value of these parameters at this time to forecast the water table depth for the unknown future. To predict the value of weather data, socio-economic and other variable data for the unknown future, we need to use the present value of these parameters at this time. To ensure as much precision as possible, we respectively use the wavelet-LSTM model with a delay of 1–12 months to predict the water table depth in the next 1–12 months. The green dotted line indicates the results of the future predictions (**Figure 7**). **Figure 10** compares correlation coefficient *R* values between the LSTM model and wavelet-LSTM model during the delayed testing period for 1–12 months. Although the *R* values decrease as the prediction delay increases, the performance of wavelet-LSTM is remarkably better than the single LSTM model in 6 months' prediction, as expected. Wavelet-LSTM also shows higher stability.

Furthermore, the results show that for a single LSTM model, the 6-months forecast is sometimes more reliable

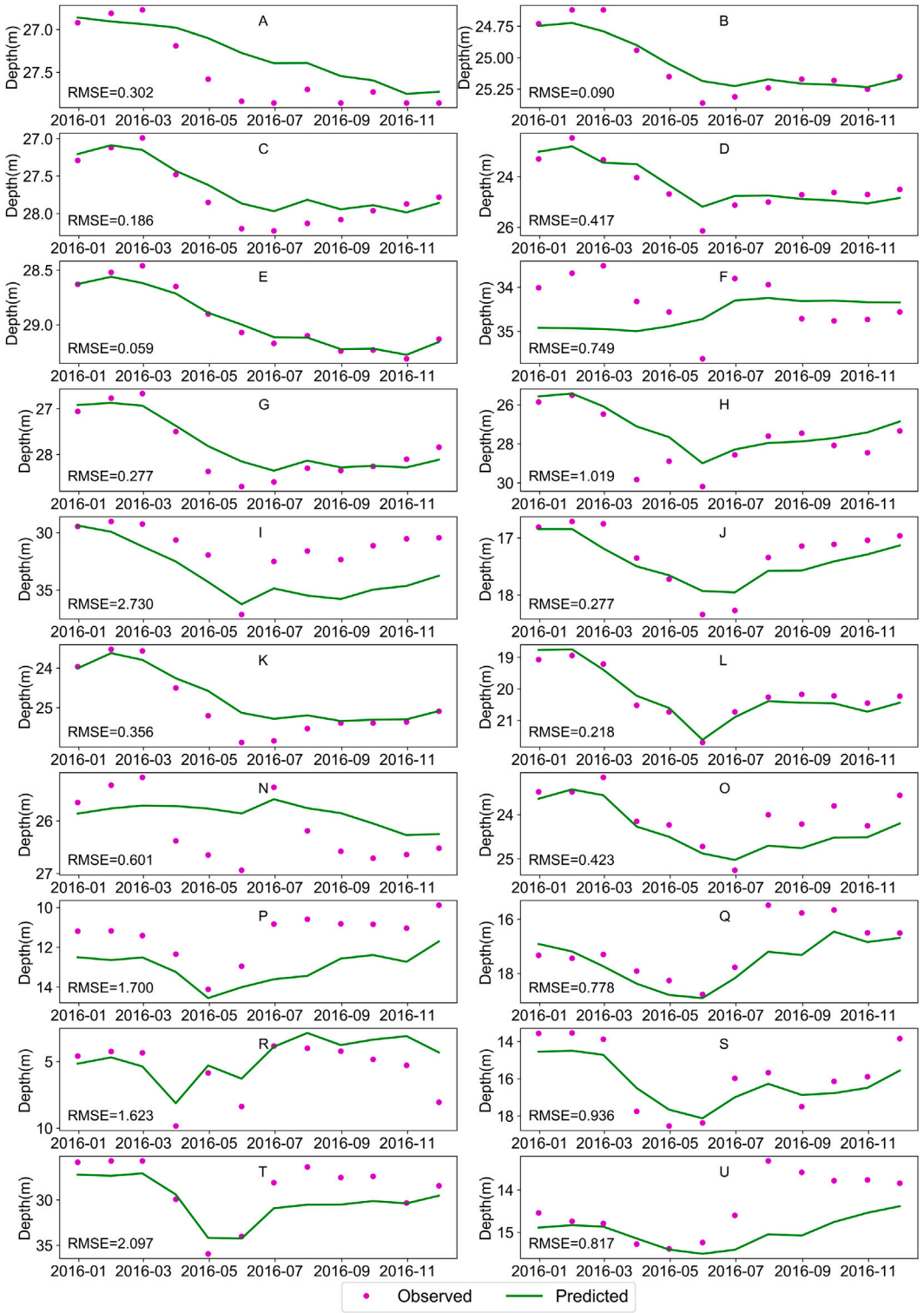

FIGURE 11 | Recursive prediction using wavelet-LSTM model for 1–12 months.

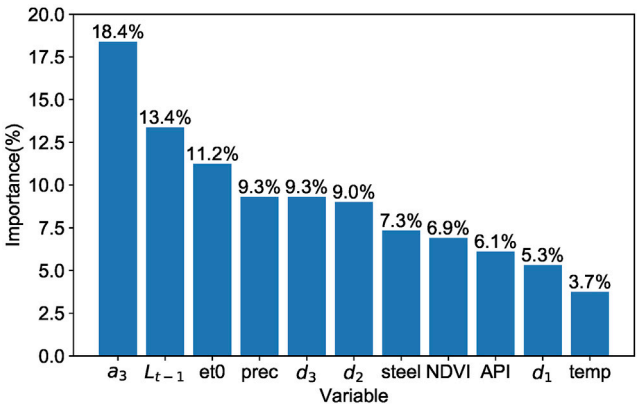

FIGURE 12 | Importance percentage of each variable.

than the 4 or 5-months forecast. However, the former has a longer time frame; the 12-months forecast shows better than 9, 10 or 11 months of superior performance. Nevertheless, for wavelet-LSTM, the advantages in the 6 and 12 months forecasts are not obvious. In other words, wavelet transform increases the model's dependence on the autocorrelation of the data.

In addition, to evaluate the response of groundwater to changes in various variables (such as climate change, economic development, etc.) in the future, future simulated values of these variables are entered into the model. Then the recursive method is used to predict the water table depth in the future gradually. **Figure 11** shows the 12-months recursive forecast using the January 2016 forecast values. In this method, the meteorological and socio-economic data are real values, and the wavelet decomposition data will be

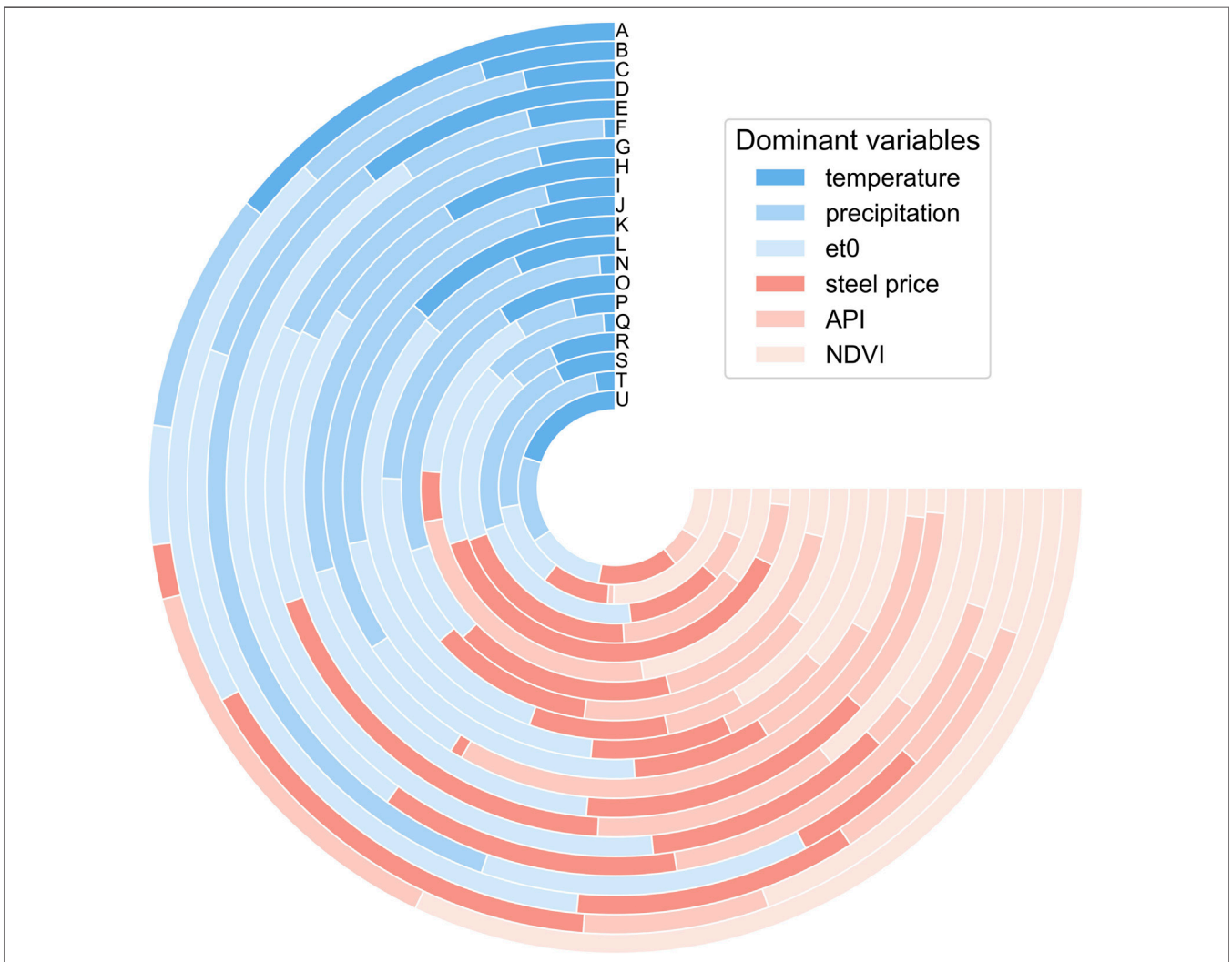

FIGURE 13 | Impact of dominant variables on water table depth at each site.

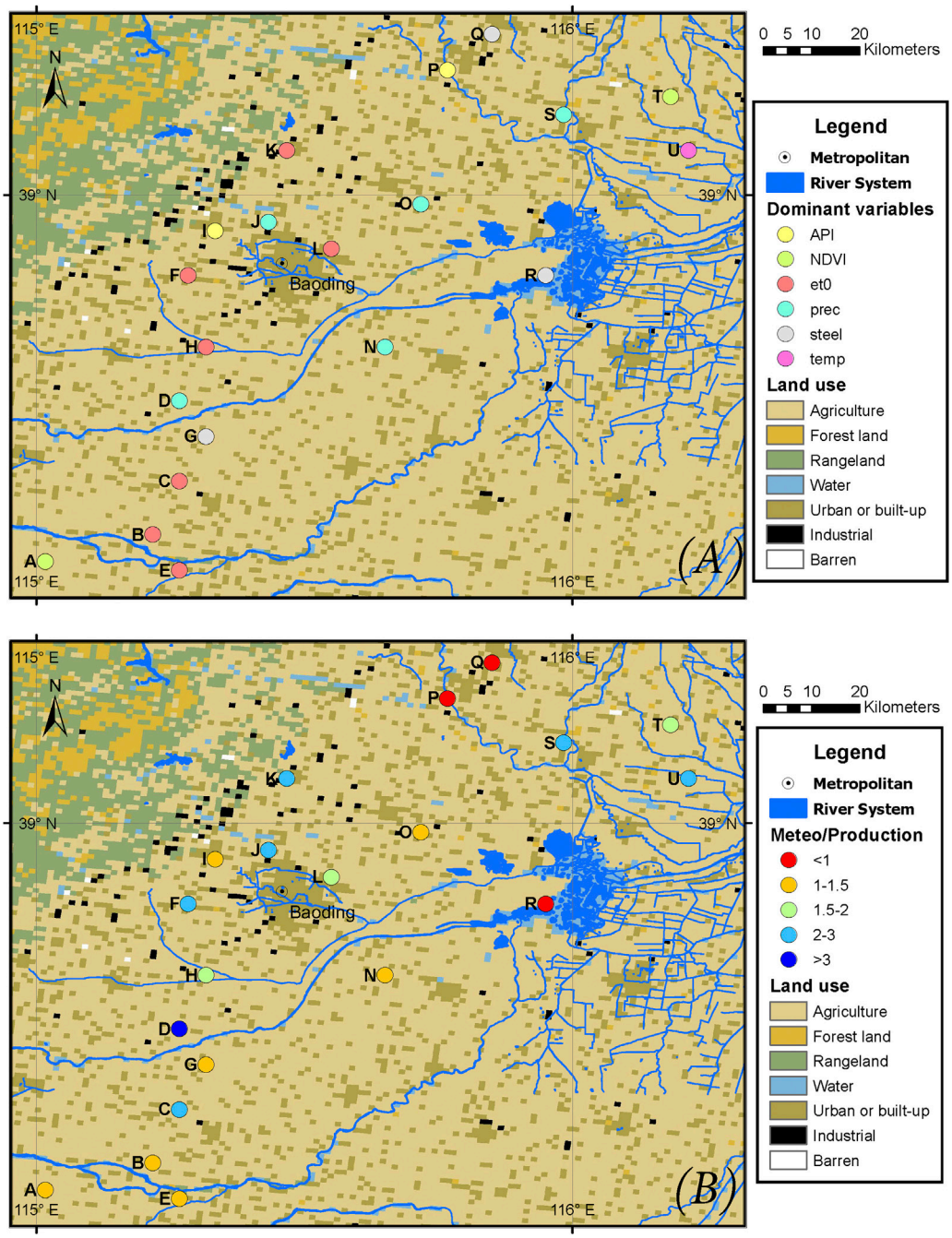

FIGURE 14 | The relationship between land use and importance of variables: **(A)** Dominant variables at each site **(B)** The ratio of meteorological factors to production activity factors.

predicted and produced simultaneously with the water table depth data. It can be seen that the losses of most sites have not increased significantly over time. Mainly at sites B, E, and L, the RMSE values reached 0.090, 0.059, and 0.218 m, respectively. The prediction curve and the observation curve can be matched precisely. The results show the effectiveness of this method for predicting and evaluating the water table.

3.4 Importance Evaluation of Each Variable

To evaluate the impact of each variable on the simulation effect, we calculate the contribution of each node in the following equation:

$$\begin{pmatrix} C_1 & \dots & C_j & \dots & C_n \end{pmatrix} = \begin{pmatrix} R_1 & \dots & R_j & \dots & R_m \end{pmatrix} \begin{pmatrix} w_{11} & \dots & w_{1j} & \dots & w_{1n} \\ \dots & & & & \dots \\ w_{i1} & \dots & w_{1j} & \dots & w_{in} \\ \dots & & & & \dots \\ w_{m1} & \dots & w_{mj} & \dots & w_{mn} \end{pmatrix} \quad (14)$$

where C_j represents the contribution of the j node to the results; R_i represents the correlation coefficient of prediction value and the measured value at i-th site (**Table 1**); w_{ij} represents the input layer weight of the i-th site, the j-th node.

As shown in **Figure 12**, a_3, d_1, d_2, d_3 represent wavelet decomposition sequence; L_{t-1} represents the past water table; temp represents temperature; et0 represents evapotranspiration; prec represents precipitation. The approximate component (a_3) of the wavelet has the greatest impact, accounting for 18.4% of the total contribution; then, the past water level (L_{t-1}) can explain 13.4% of the result. Among the external variables, precipitation and evapotranspiration have the greatest impact on the results through recharge, vegetation and soil evapotranspiration. The steel price contribution rate is 7.3%, slightly higher than NDVI and API. It fully shows that agricultural irrigation and climate change will affect groundwater, but the steel industry, the mainstay industry in the study area, also has a big impact on groundwater. The prices of agricultural products are also affected by meteorological conditions. For example, precipitation can increase the yield of crops such as corn, but it is harmful to cotton (Eck et al., 2020). However, increased agricultural production can also lead to a drop in the water table due to increased irrigation. Consequently, the contribution rate of agriculture is lower than that of industry.

Since the wavelet components and past water table depth data accounted for more than 50% percent of the weights. If the remaining variables are considered "external variables," the weights of the socio-economic factors (price of steel, API and NDVI) represent almost half of the external variables. **Figure 13** shows the impact of the temperature, precipitation, evaporation, API, steel price and NDVI on the water table in 20 sites, respectively. For most sites, precipitation and evaporation contribute to changes in the water table, and evaporation at site E and precipitation at site D was even more than half. While the weight of the price of steel and API is not as great as precipitation and evaporation, it is still considerable. Site D has the lowest socio-economic impact, and the weight is less than 1/3. However, the socio-economic ratio of most sites is in the range of 1/2–1/3.

It should be noted that it is unavoidable for socio-economic data to exhibit extreme price swings caused by emotional investment decisions. For example, due to the impact of the 2008 financial crisis, the steel price index fell sharply. In this case, water table fluctuation cannot accurately reflect the relationship between supply and demand. As a result, we strive to reflect the degree of influence of each element using the model.

In addition, this study also analysed the dominant factors affecting the water table by land use distribution. Most of the study area is occupied by agricultural land, forest and pastures. A large portion of industrial land is distributed northwest of Baoding city, close to the forest. According to surveys, the leading industry in northwest Baoding is papermaking, which consumes a lot of water and wood. Since the variables are not independent, we also used anthropogenic activities and meteorological ratios as much as possible to describe the relationship of each variable. The lower the ratio, the greater the impact of anthropogenic activities (**Figure 14**). It can be seen that except for P, Q, and R, the ratio of all sites are greater than 1. Since the R is close to Baiyangdian Lake, its water table is heavily influenced by human activities, fluctuates erratically and the simulation impact is weak. This outcome is also consistent with the study of Gorgij et al. (2017). They found that the sites located on the river may be affected by the fluctuations in the river water and that the simulation effect of these sites is not as good as that of other sites. In addition, the water table depth of P and Q are strongly affected by anthropogenic activities. Site A, B, C, D, E, and G in the southwest of the study region are the central over mining areas. Except for the points near C and D, fewer industries and the proportion of agricultural land is relatively large. The water table of these sites have trended downward and are greatly affected by anthropogenic activities. In this regard, Dong et al. (2019) concluded that the water table dropped most significantly in the place with the highest proportion of agricultural land.

The water table in the study area shows a downward trend from A, B, C, D, E, G, H, K, N, and O, located in the southwestern part of the study region, while this is where the groundwater funnel area is located. It may be because the southwest of the study area is dominated by agricultural land and far away from industrial areas and lakes. As the main crop in this region, the price of wheat per unit of yield is relatively stable, and the water table trend has not changed significantly. At these sites, the ratio of meteorological to human activity weights for A, B, E, G, N, and O is 1–1.5, and D is 3. Other sites (F, I, J, K, l, P, Q, R, S, T, U) showed a decrease then and increase, or complicated fluctuations. These sites are mainly distributed in the north of the study area. Among these sites, the ratio of meteorological to human activity weights for sites I, P, Q, and R are relatively low, while F, K, J, L, S, T, and U sites show high ratios. Since the effects of various variables on groundwater are not independent, agricultural production is also affected by meteorological changes. Therefore, these two regions consist of sites with higher meteorological weights and sites with lower meteorological weights. However, in general, the sites with lower meteorological weight are mainly distributed in the groundwater funnel area.

4 CONCLUSION

This study evaluated the predictive performance of the LSTM combined with wavelet transform in the groundwater over-

exploitation region. The results show that the NN can be used as an efficient model for prediction. Moreover, due to anthropogenic activities, the data is rich in noise and non-stationary in the groundwater over-exploitation area. The original sequence is decomposed into three levels by Meyer wavelet, which can significantly improve the simulation effect of LSTM. Using the wavelet transform combined with LSTM and FNN to predict the water table depth over the next 1–12 months, it can be concluded that the long-term prediction effect of LSTM-NN in areas of over-exploitation of groundwater is better than FNN, indicating that LSTM can memorise long-term information and effectively understand bit trend changes in water table. Furthermore, by using meteorological and socio-economic data, the proposed model can forecast future changes in the water table through a recursive method, providing a benchmark for rational utilisation planning of groundwater.

In addition, the contribution of various variables on the water table can be analysed through the LSTM-NN. The results show that Baoding's steel industry has a greater impact on water table changes. Moreover, the contribution of anthropogenic activities is higher in the sites close to the surface water. It shows that agricultural irrigation water can affect the water table. However, industrial production contributes to lowering the water table, especially in the study area where secondary industry represents a relatively large proportion. The simulation results can provide scientific guidance for the rational development and utilisation of groundwater resources in the study area.

However, we can still find that our interpretation of the variables is vague due to the nature of the NN black-box model. Therefore, more parameters, such as groundwater pumping data, should be considered in future research. If possible, in the subsequent application of the model, the amount of data should be further increased. Data that directly affects the water table should be collected, such as water pumping, crop yields, etc.

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because legal and policy restrictions. Requests to access the datasets should be directed to HH, huhongchang@tsinghua.edu.cn.

AUTHOR CONTRIBUTIONS

ZL performed the study and wrote the manuscript. YL and HH designed the study and revised the manuscript. HL, YM, and MK revised the manuscript. All authors approved the publication of the final manuscript.

FUNDING

The research was funded by the National Key R&D Program of China (2018YFC1508103, 2018YFC1508102), the National Natural Science Foundation of China (51879136, 51809173, 41930109).

REFERENCES

Adamowski, J., and Sun, K. (2010). Development of a Coupled Wavelet Transform and Neural Network Method for Flow Forecasting of Non-perennial Rivers in Semi-arid Watersheds. *J. Hydrol.* 390, 85–91. doi:10.1016/j.jhydrol.2010.06.033

Altunkaynak, A. (2007). Forecasting Surface Water Level Fluctuations of lake Van by Artificial Neural Networks. *Water Resour. Manage.* 21, 399–408. doi:10.1007/s11269-006-9022-6

Chang, F.-J., Chen, P.-A., Lu, Y.-R., Huang, E., and Chang, K.-Y. (2014). Real-time Multi-Step-Ahead Water Level Forecasting by Recurrent Neural Networks for Urban Flood Control. *J. Hydrol.* 517, 836–846. doi:10.1016/j.jhydrol.2014.06.013

Chen, L.-H., Chen, C.-T., and Pan, Y.-G. (2010). Groundwater Level Prediction Using SOM-RBFN Multisite Model. *J. Hydrologic Eng.* doi:10.1061/(asce)he.1943-5584.0000218

Chen, X., Wang, P., Muhammad, T., Xu, Z., and Li, Y. (2020). Subsystem-level Groundwater Footprint Assessment in North China Plain - the World's Largest Groundwater Depression Cone. *Ecol. Indicators* 117, 106662. doi:10.1016/j.ecolind.2020.106662

Cohen, A., and Kovacevic, J. (1996). Wavelets: The Mathematical Background. *Proc. IEEE* 84, 514–522. doi:10.1109/5.488697

Coppola, E., Szidarovszky, F., Poulton, M., and Charles, E. (2003). Artificial Neural Network Approach for Predicting Transient Water Levels in a Multilayered Groundwater System under Variable State, Pumping, and Climate Conditions. *J. Hydrologic Eng.* doi:10.1061/(asce)1084-0699(2003)8:6(348)

Coulibaly, P., Anctil, F., Aravena, R., and Bobée, B. (2001). Artificial Neural Network Modeling of Water Table Depth Fluctuations. *Water Resour. Res.* 37, 885–896. doi:10.1029/2000WR900368

Daliakopoulos, I. N., Coulibaly, P., and Tsanis, I. K. (2005). Groundwater Level Forecasting Using Artificial Neural Networks. *J. Hydrol.* 309, 229–240. doi:10.1016/j.jhydrol.2004.12.001

Daubechies, I. (1990). The Wavelet Transform, Time-Frequency Localization and Signal Analysis. *IEEE Trans. Inform. Theor.* 36, 961–1005. doi:10.1109/18.57199

de Vos, N. J., and Rientjes, T. H. M. (2005). Constraints of Artificial Neural Networks for Rainfall-Runoff Modelling: Trade-Offs in Hydrological State Representation and Model Evaluation. *Hydrol. Earth Syst. Sci.* 9, 111–126. doi:10.5194/hess-9-111-2005

Dong, Y., Jiang, C., Suri, M. R., Pee, D., Meng, L., and Rosenberg Goldstein, R. E. (2019). Groundwater Level Changes with a Focus on Agricultural Areas in the Mid-Atlantic Region of the United States, 2002-2016. *Environ. Res.* 171, 193–203. doi:10.1016/j.envres.2019.01.004

Ebrahimi, H., and Rajaee, T. (2017). Simulation of Groundwater Level Variations Using Wavelet Combined with Neural Network, Linear Regression and Support Vector Machine. *Glob. Planet. Change* 148, 181–191. doi:10.1016/j.gloplacha.2016.11.014

Eck, M. A., Murray, A. R., Ward, A. R., and Konrad, C. E. (2020). Influence of Growing Season Temperature and Precipitation Anomalies on Crop Yield in the southeastern United States. *Agric. For. Meteorology* 291, 108053. doi:10.1016/j.agrformet.2020.108053

Freire, P. K. d. M. M., Santos, C. A. G., and Silva, G. B. L. d. (2019). Analysis of the Use of Discrete Wavelet Transforms Coupled with ANN for Short-Term Streamflow Forecasting. *Appl. Soft Comput.* 80, 494–505. doi:10.1016/j.asoc.2019.04.024

Gong, H., Pan, Y., Zheng, L., Li, X., Zhu, L., Zhang, C., et al. (2018). Long-term Groundwater Storage Changes and Land Subsidence Development in the North China Plain (1971-2015). *Hydrogeol J.* 26, 1417–1427. doi:10.1007/s10040-018-1768-4

Gorgij, A. D., Kisi, O., and Moghaddam, A. A. (2017). Groundwater Budget Forecasting, Using Hybrid Wavelet-ANN-GP Modelling: A Case Study of Azarshahr Plain, East Azerbaijan, Iran. *Hydrol. Res.* 48, 455–467. doi:10.2166/nh.2016.202

Hewage, P., Trovati, M., Pereira, E., and Behera, A. (2021). Deep Learning-Based Effective fine-grained Weather Forecasting Model. *Pattern Anal. Applic* 24, 343–366. doi:10.1007/s10044-020-00898-1

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

Huang, J., Zhou, Y., Wenninger, J., Ma, H., Zhang, J., and Zhang, D. (2016). How Water Use of Salix Psammophila bush Depends on Groundwater Depth in a Semi-desert Area. *Environ. Earth Sci.* 75. doi:10.1007/s12665-016-5376-0

Khan, M. M. H., Muhammad, N. S., and El-Shafie, A. (2020). Wavelet Based Hybrid ANN-ARIMA Models for Meteorological Drought Forecasting. *J. Hydrol.* 590, 125380. doi:10.1016/j.jhydrol.2020.125380

Kim, T.-W., and Valdés, J. B. (2003). Nonlinear Model for Drought Forecasting Based on a Conjunction of Wavelet Transforms and Neural Networks. *J. Hydrologic Eng.* doi:10.1061/(asce)1084-0699(2003)8:6(319)

Kratzert, F., Klotz, D., Brenner, C., Schulz, K., and Herrnegger, M. (2018). Rainfall-runoff Modelling Using Long Short-Term Memory (LSTM) Networks. *Hydrol. Earth Syst. Sci.* 22, 6005–6022. doi:10.5194/hess-22-6005-2018

Lachaal, F., Mlayah, A., Bédir, M., Tarhouni, J., and Leduc, C. (2012). Implementation of a 3-D Groundwater Flow Model in a Semi-arid Region Using MODFLOW and GIS Tools: The Zéramdine-Béni Hassen Miocene Aquifer System (East-central Tunisia). *Comput. Geosciences* 48, 187–198. doi:10.1016/j.cageo.2012.05.007

Li, J., Wang, Y., Zhu, C., Xue, X., Qian, K., Xie, X., et al. (2020). Hydrogeochemical Processes Controlling the Mobilization and Enrichment of Fluoride in Groundwater of the North China Plain. *Sci. Total Environ.* 730, 138877. doi:10.1016/j.scitotenv.2020.138877

Liu, P., Zhang, H., and Eom, K. B. (2017). Active Deep Learning for Classification of Hyperspectral Images. *IEEE J. Sel. Top. Appl. Earth Observations Remote Sensing* 10, 712–724. doi:10.1109/JSTARS.2016.2598859

Liu, Z., Peng, C., Xiang, W., Deng, X., Tian, D., Zhao, M., et al. (2012). Simulations of Runoff and Evapotranspiration in Chinese Fir Plantation Ecosystems Using Artificial Neural Networks. *Ecol. Model.* 226, 71–76. doi:10.1016/j.ecolmodel.2011.11.023

Maheswaran, R., and Khosa, R. (2012). Comparative Study of Different Wavelets for Hydrologic Forecasting. *Comput. Geosciences* 46, 284–295. doi:10.1016/j.cageo.2011.12.015

Maihemuti, B., Simayi, Z., Alifujiang, Y., Aishan, T., Abliz, A., and Aierken, G. (2021). Development and Evaluation of the Soil Water Balance Model in an Inland Arid delta Oasis: Implications for Sustainable Groundwater Resource Management. *Glob. Ecol. Conservation* 25, e01408. doi:10.1016/j.gecco.2020.e01408

Mohanty, S., Jha, M. K., Kumar, A., and Panda, D. K. (2013). Comparative Evaluation of Numerical Model and Artificial Neural Network for Simulating Groundwater Flow in Kathajodi-Surua Inter-basin of Odisha, India. *J. Hydrol.* 495, 38–51. doi:10.1016/j.jhydrol.2013.04.041

Nanda, T., Sahoo, B., Beria, H., and Chatterjee, C. (2016). A Wavelet-Based Non-linear Autoregressive with Exogenous Inputs (WNARX) Dynamic Neural Network Model for Real-Time Flood Forecasting Using Satellite-Based Rainfall Products. *J. Hydrol.* 539, 57–73. doi:10.1016/j.jhydrol.2016.05.014

Nanda, T., Sahoo, B., and Chatterjee, C. (2019). Enhancing Real-Time Streamflow Forecasts with Wavelet-Neural Network Based Error-Updating Schemes and ECMWF Meteorological Predictions in Variable Infiltration Capacity Model. *J. Hydrol.* 575, 890–910. doi:10.1016/j.jhydrol.2019.05.051

Nayak, P. C., Rao, Y. R. S., and Sudheer, K. P. (2006). Groundwater Level Forecasting in a Shallow Aquifer Using Artificial Neural Network Approach. *Water Resour. Manage.* 20, 77–90. doi:10.1007/s11269-006-4007-z

Nourani, V., Alami, M. T., and Aminfar, M. H. (2009). A Combined Neural-Wavelet Model for Prediction of Ligvanchai Watershed Precipitation. *Eng. Appl. Artif. Intelligence* 22, 466–472. doi:10.1016/j.engappai.2008.09.003

Nourani, V., and Mousavi, S. (2016). Spatiotemporal Groundwater Level Modeling Using Hybrid Artificial Intelligence-Meshless Method. *J. Hydrol.* 536, 10–25. doi:10.1016/j.jhydrol.2016.02.030

Quilty, J., and Adamowski, J. (2018). Addressing the Incorrect Usage of Wavelet-Based Hydrological and Water Resources Forecasting Models for Real-World Applications with Best Practices and a New Forecasting Framework. *J. Hydrol.* 563, 336–353. doi:10.1016/j.jhydrol.2018.05.003

Rajaee, T., Ebrahimi, H., and Nourani, V. (2019). A Review of the Artificial Intelligence Methods in Groundwater Level Modeling. *J. Hydrol.* 572, 336–351. doi:10.1016/j.jhydrol.2018.12.037

Roy, P. K., Roy, S. S., Giri, A., Banerjee, G., Majumder, A., and Mazumdar, A. (2015). Study of Impact on Surface Water and Groundwater Around Flow fields Due to Changes in River Stage Using Groundwater Modeling System. *Clean. Techn Environ. Pol.* 17, 145–154. doi:10.1007/s10098-014-0769-9

Tang, Q., Zhang, X., and Tang, Y. (2013). Anthropogenic Impacts on Mass Change in North China. *Geophys. Res. Lett.* 40, 3924–3928. doi:10.1002/grl.50790

Taormina, R., Chau, K.-w., and Sethi, R. (2012). Artificial Neural Network Simulation of Hourly Groundwater Levels in a Coastal Aquifer System of the Venice Lagoon. *Eng. Appl. Artif. Intelligence* 25, 1670–1676. doi:10.1016/j.engappai.2012.02.009

Tiwari, M. K., and Chatterjee, C. (2010). Development of an Accurate and Reliable Hourly Flood Forecasting Model Using Wavelet-Bootstrap-ANN (WBANN) Hybrid Approach. *J. Hydrol.* 394, 458–470. doi:10.1016/j.jhydrol.2010.10.001

Xiang, Z., Bailey, R. T., Nozari, S., Husain, Z., Kisekka, I., Sharda, V., et al. (2020). DSSAT-MODFLOW: A New Modeling Framework for Exploring Groundwater Conservation Strategies in Irrigated Areas. *Agric. Water Manag.* 232, 106033. doi:10.1016/j.agwat.2020.106033

Xiao, D., Shen, Y., Qi, Y., Moiwo, J. P., Min, L., Zhang, Y., et al. (2017). Impact of Alternative Cropping Systems on Groundwater Use and Grain Yields in the North China Plain Region. *Agric. Syst.* 153, 109–117. doi:10.1016/j.agsy.2017.01.018

Xu, X., Huang, G., Zhan, H., Qu, Z., and Huang, Q. (2012). Integration of SWAP and MODFLOW-2000 for Modeling Groundwater Dynamics in Shallow Water Table Areas. *J. Hydrol.* 412-413, 170–181. doi:10.1016/j.jhydrol.2011.07.002

Zealand, C. M., Burn, D. H., and Simonovic, S. P. (1999). Short Term Streamflow Forecasting Using Artificial Neural Networks. *J. Hydrol.* doi:10.1016/S0022-1694(98)00242-X

Zhang, J., Zhu, Y., Zhang, X., Ye, M., and Yang, J. (2018). Developing a Long Short-Term Memory (LSTM) Based Model for Predicting Water Table Depth in Agricultural Areas. *J. Hydrol.* 561, 918–929. doi:10.1016/j.jhydrol.2018.04.065

Zhang, L., Liu, P., Zhao, L., Wang, G., Zhang, W., and Liu, J. (2021). Air Quality Predictions with a Semi-supervised Bidirectional Lstm Neural Network. *Atmos. Pollut. Res.* 12, 328–339. doi:10.1016/j.apr.2020.09.003

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.

frontiers in Environmental Science

ORIGINAL RESEARCH
published: 07 January 2022
doi: 10.3389/fenvs.2021.780069

Data-Driven Forest Cover Change and Its Driving Factors Analysis in Africa

Han Xiao[1,2], Jianbo Liu[1], Guojin He[1], Xiaomei Zhang[1], Hua Wang[1], Tengfei Long[1], Zhaoming Zhang[1], Wei Wang[1], Ranyu Yin[1], Yantao Guo[1], Bo Cheng[1] and Qianqian Cao[1,2]*

[1]*Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, China,* [2]*School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Science, Beijing, China*

Forest cover plays an important role in sustaining ecological security to realize Sustainable Development Goals (SDGs). The research target area is composed of the African region which is experiencing unprecedented deforestation based on the data collection from 54 countries and regions between 2000 and 2020. Spatial autocorrelation analysis, global principal component analysis, and geographic detector model have been used as the core research tool. The temporal and spatial patterns of forest cover change in Africa and the driving effects of population growth, economic and trade, social development, arable land expansion, and other factors on forest cover change in different periods have been demonstrated. The findings are as follows: 1) extremely unequal distribution of Africa forest has caused forest area reduction in 20 years. The reduction quantity of forest has been illustrated from strong to weak: Central Africa (strongest), East Africa (higher strong), West Africa (medium), South Africa (higher weak), and North Africa (weakest). However, the forest reduction area in West Africa with the original ratio is the most significant. More than 80% of the forest area reduction in Africa has occurred in 14 countries, just five national forest areas to achieve the net growth, but the increase amount was only 1% of loss amount. 2) The spatial pattern of forest cover change in Africa contracted and clustered gradually, especially after 2012. Algeria was the hotspot cluster of Morocco and Tunisia, forming the increase area of forest cover in North Africa. Zambia, the coldest point, gathers Angola significantly, while the Democratic Republic of the Congo and Tanzania form a significantly reduced forest cover area. 3) Total population, land area, cultivated land, urban population, consumer price index, and birth rate are the main factors influencing the temporal evolution of forest cover change in Africa. It can be divided into four stages to interpret the different explanations and significance of each factor for forest cover change in the study area.

Keywords: data-driven, forest cover change, driving factors, spatio-temporal features, Big Earth Data, environmental modeling

OPEN ACCESS

Edited by:
Yizhe Zhan,
MetService, New Zealand

Reviewed by:
Tianhai Cheng,
Institute of Remote Sensing and Digital Earth (CAS), China
Shi Shen,
Beijing Normal University, China

****Correspondence:***
Han Xiao
xiaohan@aircas.ac.cn

Specialty section:
This article was submitted to Environmental Informatics and Remote Sensing, a section of the journal Frontiers in Environmental Science

Received: *20 September 2021*
Accepted: *06 December 2021*
Published: *07 January 2022*

Citation:
Xiao H, Liu J, He G, Zhang X, Wang H, Long T, Zhang Z, Wang W, Yin R, Guo Y, Cheng B and Cao Q (2022) Data-Driven Forest Cover Change and Its Driving Factors Analysis in Africa. Front. Environ. Sci. 9:780069. doi: 10.3389/fenvs.2021.780069

1 INTRODUCTION

Forest is not only one of the important sources of human production and life but also one of the important indicators to maintain ecological security to achieve the Sustainable Development Goals (SDGs) (Glaser, 2012; Giam, 2017; Macdicken, 2015). After human beings frantically plundered forest resources, they gradually realized that large-scale deforestation would cause adverse consequences such as climate change, soil erosion, and biodiversity loss, so the net loss rate of

global forest cover has been slowed down through various appeals and efforts (Barlow et al., 2007; Edwards et al., 2012; Bonan, 2008; Hansen et al., 2013; Harris et al., 2021). However, from 2010 to 2020, Africa's forest cover will be reduced by 3.9 million hectares, surpassing South America, and the net loss rate is the highest in the world (FAO, 2020; Hoang and Kanemoto, 2021). It is expected that Africa will experience large-scale population growth and climate change, and Africa has a typical disturbance-driven ecosystem (Diffenbaugh and Giorgi, 2012; United Nations, Department of Economic and Social Affairs, Population Division, 2017). Once the forests in most areas are affected, the entire ecological environment will become highly vulnerable (Midgley and Bond, 2015; Réjou-Méchain et al., 2021). Analyzing and digging out the main driving factors that affect the changes in forest cover in Africa is essential to promote the protection of global forest protection and help regions achieve Sustainable Development Goals.

Many scholars have conducted research on the relationship between temporal and spatial characteristics based on the changes in forest cover in a specific study area and possible influencing factors, such as cocoa migration and deforestation in West Africa are positively correlated (Ruf et al., 2015), deforestation in Cameroon in relation to food security of poor populations (Gbetnkom, 2009; Pendrill et al., 2019), and road accessibility, agricultural product prices, household income levels, and catchment water balance . are all related to deforestation (Mainardi, 1998; Beckman et al., 2017; Kaimowitz, 1999; Assogba and Zhang, 2018; Mushi et al., 2020; Wilson et al., 2001; A et al., 2004; Garzuglia and Saket, 2003). Research shows that the aforementioned factors can be the root cause of deforestation to a certain extent, which in turn promotes deforestation (farmers, poor populations, animal husbandry, and loggers) to become the main source of deforestation (Gbetnkom, 2009; Afdb and Kouakou, 2010). However, causal attribution of natural and ecosystem system evolution is complicated because exploring forest cover dynamics is being affected by the interaction of many factors such as ecological environment, political, economic, and social factors, and have a wide range of impacts on time and space scales (Geist and Lambin, 2002; Defries et al., 2010; Sandel and Svenning, 2013; Rudel and Roper, 1997; Sannigrahi et al., 2018; Kim and Kim, 2010). Forzieri et al. quantify the vulnerability of European forests to fires, windthrows, and insect outbreaks by random forest (RF) regression (Forzieri et al., 2021). Curtis developed a forest loss classification model, to determine a spatial attribution of forest disturbance, was attributed to commodity production, forestry, shifting agriculture, and wildfire (Curtis et al., 2018; Goodchild and Glennon, 2010). Hoang et al. used remote sensing data and multi-regional input–output models to provide correlation analysis of the spatial pattern of international trade and deforestation (Hoang and Kanemoto, 2021). Beckman used economic models to simulate the reference scenarios of land use in 14 regions around the world to determine the potential impact of policies on relative forest losses (Beckman et al., 2017). DeFries used forest loss data from 2000 to 2005 to estimate the economic, agricultural, and population correlations of 41 countries and found that forest loss was positively correlated with urban population growth and agricultural exports during this period (Defries et al., 2010).

However, different countries have different social and economic development stages and different forest protection strategies, which may result in significant differences in regional deforestation (Deacon, 1994; Hosonuma et al., 2012; Heitzig et al., 2016). Considering Tobler's first law of geography, severe deforested areas can also be a key driver of deforestation in neighboring areas. Currently, most research in the context of the driving factors of forest cover change has used local decomposition analysis of a few specific factors but do not evolve the spatial heterogeneity of each factor on the intensity of forest cover. This research analysis has mainly evolved the following two aspects: 1) it analyzes the spatial correlation, heterogeneity, and evolution characteristics of the forest cover change pattern in African countries from 2000 to 2020; 2) the 21 driving factors have been opted based on four dimensions of economy, society, and population. The nature of the research analysis is to explore and detect the significant driving factors of forest cover change in different historical periods via geographic detectors and weighted regression geographical models' applications. The research results are intended to provide benchmark for balancing population, economy, and green sustainable development in Africa.

2 STUDY REGION

African area is about 30.2 million square kilometers (land area), accounting for 20.4% of the world's total land area. The plateau (60%) is the major part of the Africa terrain. The elevation is gradually lower from southeast to northwest. The climate factors present high temperature and arid due to the north–south symmetry distributed climate zone in the middle of the African continent (across equator). It is the second largest continent in the world and has the second largest population (about 1.286 billion). Africa is rich in forest resources, accounting for 13.85 percent of the world's forest area (about 51,400 hectares) in 2018 (National Remote Sensing Center of China, 2019). Africa also has one of the fastest growing populations in the world. The Africa's population is growing at an annual rate of 2.3 percent with sub-Saharan Africa in turn growing faster than North Africa. Sub-Saharan Africa is the only region in the world where the number of poor populations is increasing significantly, reaching 2 billion by 2050 (World Bank Group, 2020). The challenge of reducing widespread poverty in Africa can be the core of the continent's development model for improving their living conditions including their increased use of the natural resources around them (Kouakou, 2005). Population growth, the impact of poverty, and the exploitation of natural resources have inevitably restricted and influenced each other. It is difficult to obtain data of individual islands/regions due to the influence of economic development and political situation. Therefore, this study has selected 54 countries and regions in Africa which are represented in **Table 1**, and the population and area of the study area cover more than 99% of the African continent. (**Figure 1**). Refer to World Population Prospects

TABLE 1 | Geographical division of countries in the study area.

Research area	Countries & regions' Name	Datagaps excluded
Eastern Africa	Burundi, Comoros, Djibouti, Eritrea, Ethiopia, Kenya, Madagascar, Malawi, Mauritius, Mozambique, Rwanda, Seychelles, Somalia, South Sudan, Uganda, United Republic of Tanzania, Zambia, and Zimbabwe	Mayotte, Réunion
Middle Africa	Angola, Cameroon, Central African Republic, Chad, Congo, Democratic Republic of the Congo, Equatorial Guinea, Gabon, Sao Tome, and Principe	
Northern Africa	Algeria, Egypt, Libya, Morocco, Sudan, and Tunisia,	Western Sahara
Southern Africa	Botswana, Lesotho, Namibia, South Africa, and Swaziland	
Western Africa	Benin, Burkina Faso, Cabo Verde, Côte d'Ivoire, Gambia, Ghana, Guinea, Guinea-Bissau, Liberia, Mali, Mauritania, Niger, Nigeria, Senegal, Sierra Leone, and Togo	Saint Helena

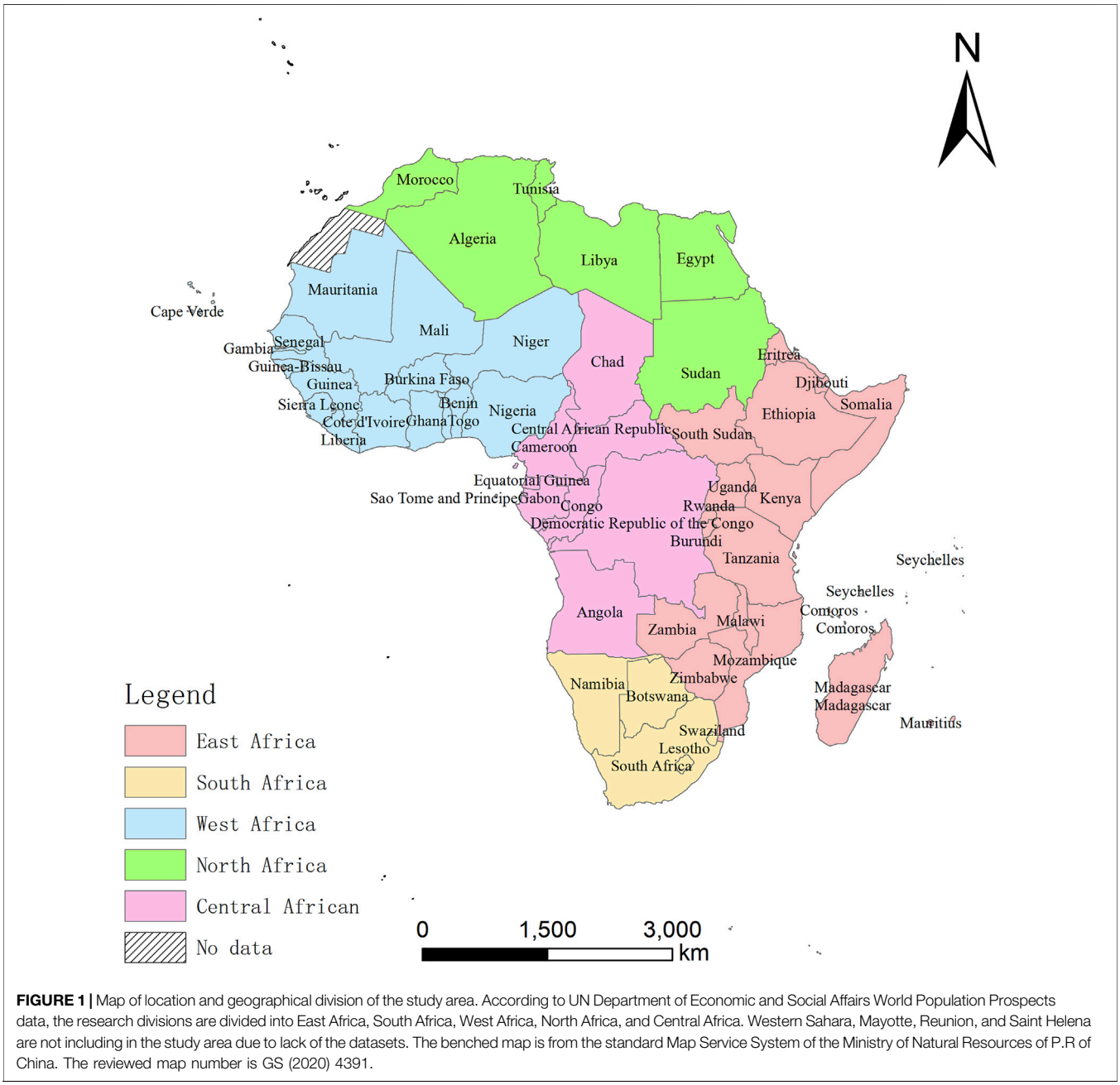

FIGURE 1 | Map of location and geographical division of the study area. According to UN Department of Economic and Social Affairs World Population Prospects data, the research divisions are divided into East Africa, South Africa, West Africa, North Africa, and Central Africa. Western Sahara, Mayotte, Reunion, and Saint Helena are not including in the study area due to lack of the datasets. The benched map is from the standard Map Service System of the Ministry of Natural Resources of P.R of China. The reviewed map number is GS (2020) 4391.

TABLE 2 | Variable description of driving indicators.

Research dimension	Indicators name	Unit	Mean	Standardized deviation	Forcasting
Economic factor	GDP	2010 US$	35901785382.50	75893129242.70	+
	GDP growth	%	2.50	29.91	+
	Foreign direct investment, net inflows	US$	739000886.48	1540164408.52	+
	Foreign direct investment, net inflows	%	4.37	7.82	+
	Goods exports	US$	7420032761.41	15841232780.62	+
	Goods imports	US$	7603241584.23	14153662572.90	+
Social factor	Forest rents	%	4.76	5.54	+
	Net ODA received	%	7.78	9.16	+
	Consumer price index	-	2069.05	35351.53	–
	Unemployment, total	%	8.66	7.05	–
	Access to clean fuels and technologies for cooking	%	26.12	31.61	+
	Compulsory education	Year	7.80	3.29	+
Population factor	Population, total	People	19476448.96	28256604.83	–
	Rural population	People	11628519.60	17045641.91	–
	Urban population	People	7682742.94	12373334.45	–
	Labor force, total	People	7280119.91	10046139.07	–
	Death rate	‰	10.26	3.63	+
	Birth rate	‰	34.84	8.42	–
Natural factor	Land area	Km^2	557709.49	621698.78	+
	Terrestrial protected areas	%	16.18	11.67	+
	Arable land	Hm^2	4067045.04	5830099.69	–

data for the classification of countries and regions (United Nations, Department of Economic and Social Affairs, Population Division, 2019).

3 DATASETS AND METHODOLOGY

3.1 Datasets

Forest cover change has been influenced by multiple factors (Curtis et al., 2018; Geist and Lambin, 2002; FAO, 2015). We selected 21 indicators from four dimensions of those are economy, society, population, and nature as explained variables for analysis. Economic dimension indicators reflect the size of a country's economic development including GDP, GDP growth, foreign direct investment (net inflows), Foreign Direct Investment (net inflows), goods exports, and goods imports. The social dimension reflects the general state of development including forest rents, net ODA received, and consumer price index, unemployment, access to clean fuels and technologies for cooking (% of population) and compulsory education. Population factor indicators reflect the possible consumption of forest resources, including population, total, rural population, urban population, labor force, death rate, and birth rate. The natural dimension index reflects natural resource abundance, land area, terrestrial protected areas (% of total land area), and arable Land. The data cover the period from 2000 to 2020, and the data sources include World Development Indicators (WDI), and it is consistent (**Table 2**) in the Center for International Prospective Studies (CEPII) database (The World Bank, 2021). The number of years of compulsory education is based on UNESCO statistics for 2020 (UNESCO, 2021). Since most of the other population, economic, and social data selected in the subsequent analysis originated from the official statistical data of the World Bank, the forest cover data of the research area from 2000 to 2018 inevitably refer to "Ag.lnd.frst.k2-Forest Area (SQ.km)" in the WDI dataset. A small part of the missing forest cover data is supplemented via data from FRA 2020 published by FAO and remote-sensing data products (Hansen et al., 2013; Shimada et al., 2014; National Remote Sensing Center of China, 2019; Zhang et al., 2020).

3.2 Spatial Autocorrelation

The purpose of spatial autocorrelation analysis is to analyze whether a certain variable can show a certain spatial correlation and the degree of correlation. Relevant studies indicate that forest distribution will be deeply affected via the Tobler's first law of geography, forest protection, logging, and trade in neighboring countries in terms of the same natural growth environment affected by neighboring regions with a certain spatial autocorrelation (Jha et al., 2016; Kaimowitz et al., 2002). The global Moran's I index can be used to describe the spatial dependence of forest cover, growth, or deforestation in African countries and regions quantitatively (Moran and Ap, 1950). When I∈ (0, 1], the variables become more similar with the shortening of the measured distance, presenting a positive spatial correlation. When I∈ [-1, 0), it indicates that the overall spatial agglomeration is negatively correlated; when I = 0, it indicates that there is no spatial dependence, and the spatial pattern is randomly distributed. The calculation equation is as follows:

$$Moran's\ I = \frac{N}{\sum_{ij} w_{ij}} \frac{\sum_i \sum_j w_{ij}(x_i - \bar{x})}{\sum_i (x_i - \bar{x})^2}, \quad (1)$$

where N is the number of countries along the Belt and Road, I and j represent country numbers, W_{ij} is the spatial weight matrix, X_i is

the forest cover value of country I, and $\bar{x}$ is the average of the research area. The global Moran's I index based on Z scores has been acquired to test its significance, when the $|Z| > 1.96$, and it indicates that there is spatial autocorrelation with 95% probability. The equation is as follows:

$$Z = \frac{I - E(I)}{\sqrt{v(I)}} \sim N(0,1), \tag{2}$$

where E (I) represents the mathematical expectation of global Moran's I, v (I) represents the variance of global Moran's I, and N (0, 1) represents the normal distribution.

For Moran's I value, if $Z \geq 1.96$ or ≤ -1.96, it is considered that space has spatial autocorrelation (Moran's I >0 represents positive spatial correlation. The larger the value is, the more significant the spatial correlation is. 0 means spatial negative correlation; the smaller its value is, the greater the spatial difference is; If not, Moran's I = 0 represents that the space is random).

The global Moran's I index can count the significant characteristics of overall agglomeration, while the local indicators of spatial association (LISA) can identify the local characteristics of spatial agglomeration, thus revealing the heterogeneity of local regional spatial association (Anselin, 1995). The LISA index identifies four spatial association types of "high–high cluster," "high–low cluster," "low–high cluster," and "low–low cluster," representing different types of spatial agglomeration patterns.

3.3 Geo-Detector

Geographical Detector is a popular spatial statistical tool developed by the Prof. WangJinfeng team of LREIS, IGNRSS of Chinese Academy of Sciences, based on Excel to measure spatial differentiation and reveal driving factors. By calculating the similarity of the spatial distribution pattern between the independent variable X and dependent variable Y, the geographic detector model detects the explanatory effect intensity of the independent variable on the dependent variable (Wang and Xu, 2017). The geographic detector method was used to analyze the relationship strength of driving factors among the 6 principal component variables obtained from the original 21 variables via global principal component analysis dimension reduction. The main driving factors affecting the formation of the spatial pattern of forest cover change in the study area have been identified in different timescales. The equation is as follows:

$$q = 1 - \frac{\sum_{i=1}^{N} N_i \sigma_i^2}{N\sigma^2}, \tag{3}$$

where q value is the explanatory degree of the independent variable to the dependent variable, and the larger the q value is, the stronger the explanatory power is. I is the stratification number of the independent variable X and dependent variable Y. The Jenks natural discontinuity point grading method is adopted for stratification to discretize continuous variables. N_i and N are the number of the layer I and countries in the research area, respectively. σ_i^2 and σ^2 are the variances of the layer I and independent variables within the research area, respectively.

3.4 Global Principal Component Analysis

Global principal component analysis (GPCA) is to construct the index system of the three-dimensional time-series data table based on the use of time series and spatial sequence mutual integration, seeking a unified for all cross-section data tables and comparable to simplify subspace. Each data table on which the projection to approximate the best comprehensive effect based on the global data view is constructed.

The indicators to influence forest cover change are very diverse. Human activities are likely to be more rapid and direct, while geographical and climatic factors can also influence over long time series (with the exception of sudden natural disasters, pest disasters, or forest fires). Therefore, the analysis of 20-year time series index data of different countries and regions is constituted into sequential stereoscopic data tables for global principal component analysis. A few of "factors" are deduced from numerous observable "variables" via the analysis of three-dimensional time series data, and the minimum "factors" are used to summarize and explain the largest number of observed facts so as to establish the simplest and most basic conceptual system and reveal the essential relationship between indicators affecting forest cover change.

The steps for extracting principal factors based on global principal component analysis are described as follows:

Step 1. Construct a data table: if n regions are counted and p variables are used to describe, the "global data table" of T years is

$$X = \left[(X1)' (X2)' \ldots (XT)' \right] \in R^{T\,N \times P}. \tag{4}$$

Step 2. Normalize the elements in the table X by moving the origin of coordinates to the data center and compress and transform them to eliminate dimensional influence.

$$x_{ij}^t = \frac{\left(x_{ij}^t - x_j\right)}{S_j}, \ (i = 1, 2, \cdots, n; j = 1, 2, \cdots, p); \tag{5}$$

$$x_j = \frac{1}{T_n} \sum_{t=1}^{T} \sum_{i=1}^{n} x_{ij}^t; \tag{6}$$

$$S_j^2 = \frac{1}{T_n} \sum_{t=1}^{T} \sum_{i=1}^{n} \left(x_{ij}^t - x_j\right)^2; \tag{7}$$

$$X = \left(X_{ij}\right)_{T_n \times p} \in R^{T_n \times p}. \tag{8}$$

Step 3. Calculate the covariance matrix $V \in R^{p \times p}$ of the X matrix.

Step 4. Find the first m eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_m$ of the matrix V and the corresponding eigenvectors $\mu_1, \mu_2 \ldots \mu_m$; they are orthonormal, and $\mu_1, \mu_2 \ldots \mu_m$ is called the global principal axis.

Step 5. Calculate the global principal fraction and calculate the representation accuracy of the original data: $Q = \frac{\sum_{i=1}^{m} \lambda_i}{p}$, generally $Q \geq 85\%$.

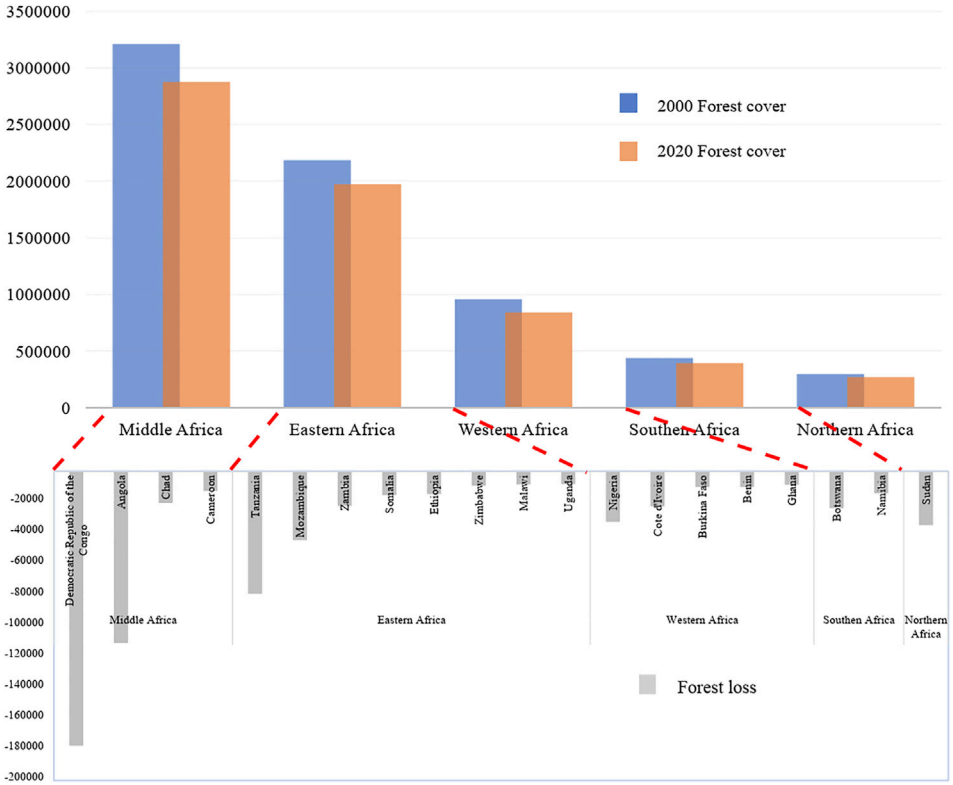

FIGURE 2 | Largest forest loss is in Central Africa, the second one is in East Africa, the third one is in West Africa, the fourth one is in South Africa, and the fifth one is in North Africa. The amount of forest reduction in the original area is more significant than that in West Africa. The most significant loss of forest in the original area was in West Africa. The top 20 countries with reduced forest area in Africa are the Democratic Republic of the Congo, Angola, Tanzania, Mozambique, Sudan, Nigeria, Botswana, Cote d'Ivoire, Zambia, Chad, Somalia, Ethiopia, Namibia, Cameroon, Burkina Faso, Benin, Zimbabwe, Ghana, Malawi, and Uganda.

4 RESULTS

4.1 Forest Cover and Its Change in 21 Years

The distribution of forests in Africa is extremely uneven in terms of the differences in precipitation, sunshine, soil properties, geographical location, natural conditions, and other factors. More than 45% of forests are distributed in Central Africa, 30% in East Africa, 13% in West Africa, 6.2% in South Africa, and 4.2% in North Africa. The forest area has decreased in recent 20 years. The largest loss is in Central Africa, the second one is in East Africa, the third one is in West Africa, the fourth one is in South Africa, and the fifth one is in North Africa. However, the most significant loss of the forest area was in West Africa, which has a loss percentage more than 12%. The percentage of forest area loss has decreased more than 10% in Central Africa and South Africa and 9.6% in East Africa and North Africa. More than 80 percent of the forest loss occurred in these 14 countries: the Democratic Republic of the Congo, Angola, Tanzania, Mozambique, Sudan, Nigeria, Botswana, Cote d'Ivoire, Zambia, Chad, Somalia, Ethiopia, Namibia, and Cameroon (**Figure 2**).

The Democratic Republic of the Central African Republic has experienced the largest loss of forest, accounting for about a quarter of the total in Africa, more than half of the loss in Central Africa, and equal to the sum of the loss in West, South, and North Africa. Middle Africa had the second largest loss of Chad forests, losing nearly a third of its total forest area in 20 years. Eastern Africa's Tanzania has the third largest forest in Africa and never changed for 20 years. Cote d'Ivoire and Benin forests in West Africa have lost 44 and 24% more of their original forest area. Gambia in Western Africa, which was not originally rich in forest resources, has lost nearly a third of its original forest area. Eastern Africa's Malawi and Uganda have both lost more than a quarter of their forests in 20 years.

Just seven countries have shown a net increase in the forest area over the past 20 years. Only five countries showed significant growth (more than 200 square kilometers). Algeria, Morocco, and Tunisia in North Africa, East Africa, and Swaziland in South Africa have shown forest improvement, but the forest growth area just accounts for only 1% of the loss area.

It is worth mentioning that more countries have realized the importance of forest protection and put into actions gradually. They have slowed down the rate of deforestation and actively regulated in Rwanda, Kenya, and Ghana. It experienced a low point toward growing year by year

TABLE 3 | Moran's I estimate of national and regional forest cover change from 2000 to 2019.

Indicator	Index	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
Net change in forest cover	Moran's I	0.08	0.08	0.08	0.101	0.101	0.101	0.101	0.084	0.084	0.080
	Z-value	1.2884	1.2888	1.2305	1.529	1.4855	1.5446	1.5534	1.3853	1.3210	1.2726
	p-value	0.101	0.101	0.11	0.088	0.091	0.072	0.068	0.089	0.104	0.101
	Indexes	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
	Moran's I	0.112	0.112	0.117	0.112	0.112	0.121	0.119	0.119	0.123	0.124
	Z-value	1.9178	1.954	1.9664	1.9664	1.9664	1.9579	1.970	1.990	2.003	2.0481
	p-value	0.049	0.056	0.043	0.043	0.043	0.048	0.045	0.042	0.049	0.041

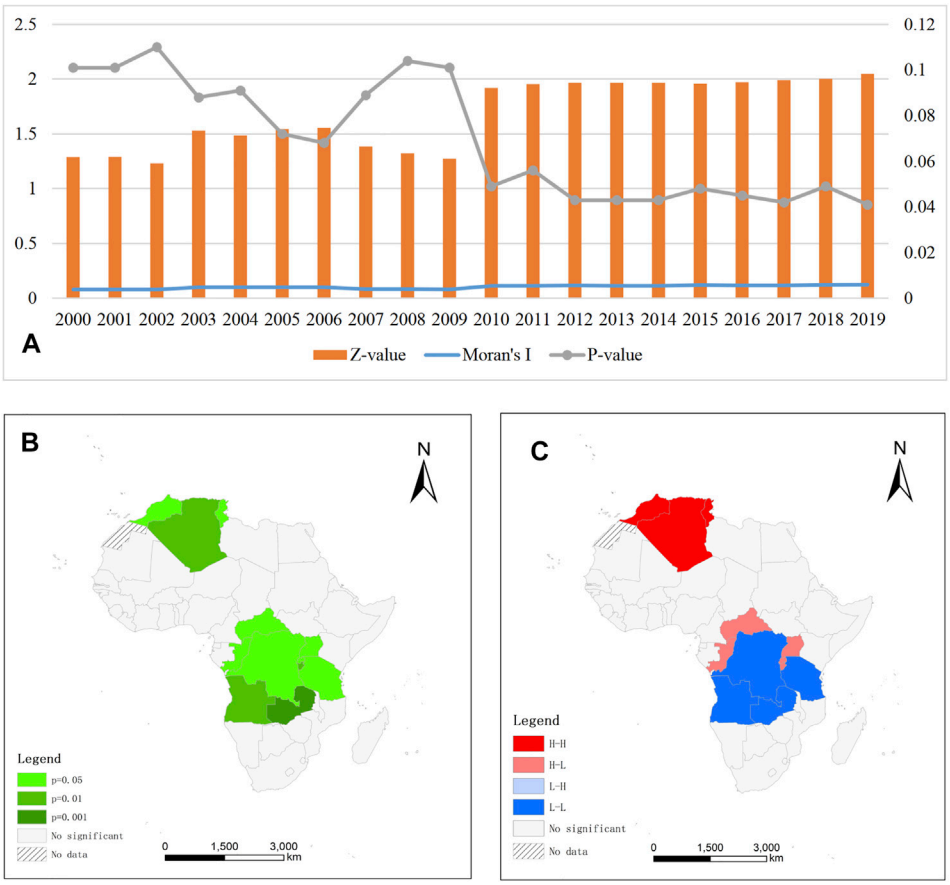

FIGURE 3 | Spatial correlation map. **(A)** Global Moran's | Z ≥1.96 after 2020 is considered that the forest cover change in the study area shows spatial autocorrelation, | ϵ (0, 1] forest cover change is positively correlated with spatial distance, and $p \leq 0.05$ after 2012 shows significant spatial aggregation characteristics in figures **(B,C)**. **(B)** Significant Map (2012–2019). The forest cover change in the study area has begun to cluster in 2010 and more significant in 2012, and the degree of spatial agglomeration increased annually. **(C)** Clustered map (2012–2019). The spatial clustering of the hot spots has been divided into two regions: Morocco and Tunisia forest cover expanding area based on Algeria as the hot spot agglomeration; Angola, the Democratic Republic of the Congo, and Tanzania have been reduced low-forest coverage in Angola, Rwanda, Uganda, Central African Republic, and Congo high–low zone.

although the forest area in Rwanda has been still decreasing in the past 20 years.

4.2 Spatial Correlation Features of Forest Cover Change

Global spatial autocorrelation analysis of net forest cover change values from 2000 to 2019 has been conducted in order to analyze the spatial correlation characteristics of forest cover change in different countries and regions in Africa. The results have illustrated (**Table 3**) that Moran's I indexes of forest cover change in the past 20 years were all positive, indicating that forest cover change presented a positive correlation of spatial aggregation. However, $0 \leq Z \leq 1.65$ and $p \geq 0.05$ in the first decade of the 21st century did not adapt the significance test. Forest cover reduction in the study area was judged to be random in that period

(of course, subjected to technical means, early statistical data may not be accurate enough to affect the calculation results). Until 2010 and 2012, when $p \leq 0.05$, only less than 5% of the data distribution in the study area is likely to be random, and the probability of data aggregation is greater than that of random distribution. $Z \geq 1.65$ and 1.96 can significantly reject the null hypothesis. The results have demonstrated that the forest cover change in the study area has begun to cluster in 2010 and more significant in 2012, and the degree of spatial agglomeration increased annually. The spatial correlation of forest cover change amongst neighboring countries or regions has been increasingly strengthened.

Moran's I estimate of national and regional forest cover change in the study area from 2000 to 2019 was analyzed (**Figure 3A**). A Global Moran's I $Z \geq 1.96$ after 2020 indicates that the forest cover change in the study area shows spatial autocorrelation, $I \in (0, 1]$ forest cover change is positively correlated with spatial distance, and $p \leq 0.05$ after 2012 shows significant spatial aggregation characteristics in **Figures 3B,C**. The LISA index has been used to study the local agglomeration characteristics of forest cover change in the region further. The LISA significance map and agglomeration map have passed the 95% significance test via ArcGIS software analysis. The spatial clustering of the hot spots has been divided into two regions: Morocco and Tunisia forest cover expanding area based on Algeria as the hot spot agglomeration; Angola, the Democratic Republic of the Congo, and Tanzania have been reduced low-forest coverage in Angola, Rwanda, Uganda, Central African Republic and Congo high–low zone.

4.3 GPCA Calculation Results and Analysis

A 54×21×21 dimensional sequential stereo data table was constructed, including sample points from 54 countries and regions, 21 driving index variables, and 21 data tables from 2000 to 2020. According to the analysis steps of GPCA, the global principal component factors were extracted by SPSS 26. In the analysis, the maximum variance transaction rotation method is adopted. After 11 iterations, the results tend to converge. According to the principle that the eigenvalue is greater than 1, a total of 6 principal components are extracted. The cumulative contribution rate of the first principal component is 33.189%; the cumulative contribution rate of the second principal component is 17.490%; the third to sixth principal components play a similar role, with accumulative contribution rates of 6.966%, 6.052%, 5.355% and 5.023%, respectively. In the process of reducing the dimension of the data table from 21 to 6, 74.075% of the information in the original data table is retained, and the original features are basically completely preserved (See **Table 4**). The six main components are population, rural population, urban population, labor force, GDP, unemployment, death rate, foreign direct investment (net inflows), consumer price index, and terrestrial protected areas (% of the total land area).

In order to investigate whether there is a certain linear relationship among economic, social, demographic, and geographical evaluation indexes affecting forest cover and whether GPCA analysis method is suitable, the KMO test and Bartlett sphericity test are selected for analysis, and the results are shown in **Table 5**. The KMO statistic is 0.783 and Sig. < 0.005, indicating that there is correlation for each variable, and the factor analysis is effective and got better results.

TABLE 4 | Generalized eigenvalues of principal components and accumulative percentage.

Principal component	Eigenvalues	Variance percentage (%)	Variance contribution rate (%)
C1	6.970	33.189	33.189
C2	3.673	17.490	50.679
C3	1.463	6.966	57.645
C4	1.271	6.052	63.697
C5	1.125	5.355	69.052
C6	1.055	5.023	74.075

TABLE 5 | KMO and Bartlett's test.

KMO and Bartlett's test		
Kaiser–Meyer–Olkin measure of sampling adequacy		0.783
Bartlett's test of sphericity	Approx. Chi-square	30845.267
	df	210
	sig	0.000

4.4 Forest Cover Change and Driving Factors on the Geographical Detector

The geographical detector is an emerging statistical method for driving force and factor analysis of spatial data (Wang, 2010), which is mainly composed of four detectors: factor detector, interactive detector, ecological detector, and risk detector. The first three detection functions are mainly adopted and conducted. Factor detection: The influence of economic, social, demographic, and natural factors (**Table 2**) on the spatial distribution of forest cover change can be calculated via factor detection analysis (q value). The larger the q value is, the greater the influence of this factor on forest cover change is. Ecological detection: Ecological detection is demonstrated to compare the significant differentiation in the spatial distribution of forest cover amongst economic, social, demographic, and natural factors. Interaction detection: The interactions are implemented amongst different influence factors. Based on the single factor q value comparative analysis and the q value sum of the two factors and the q value following the two interactive factors, we can identify whether the two post-interactive factors increase or decline their influence on the spatial distribution of forest cover or whether play an independent role themselves.

Factor detectors are first used to obtain the q values for five targeted years of 2000, 2005, 2010, 2015, and 2019 in the timescale (shown in **Table 6**). The explanatory power of the bold numerical factors all passed the 95% significance level. Overall, some significant factors have a great impact on the spatial differentiation analysis of forest cover change, such as

TABLE 6 | Factor detection results of spatial differentiation of forest cover change.

Factor		Year 2000			Year 2005			Year 2010			Year 2015			Year 2019		
Driver	**Symbolic representation**	**q**	**p**	**Rank**	**q**	**p**	**Rank**	**q**	**p**	**Rank**	**q**	**p**	**Rank**	**q**	**p**	**Rank**
GDP	X1	0.06	0.74	18	0.10	0.39	15	0.04	0.81	18	0.03	0.82	19	0.03	0.84	18
Foreign direct investment (BoP, current US$)	X2	0.15	0.27	10	0.46	0.07	1	0.26	0.33	3	0.36	0.08	1	0.26	0.28	5
Goods exports	X3	0.02	0.94	20	0.16	0.47	10	0.10	0.42	13	0.10	0.56	12	0.09	0.51	12
Goods imports	X4	0.07	0.59	16	**0.29**	**0.02**	4	0.19	0.11	7	0.20	0.08	8	0.15	0.20	9
Foreign direct investment, net inflows (% of GDP)	X5	0.16	0.45	9	0.05	0.82	19	0.06	0.57	17	0.06	0.78	16	0.01	0.99	21
Net ODA received	X6	0.02	0.87	19	0.01	0.97	21	0.09	0.42	15	0.02	0.90	21	0.01	0.95	20
Consumer price index	X7	**0.48**	**0.00**	1	0.12	0.27	11	0.14	0.17	10	0.09	0.41	14	0.03	0.87	17
Unemployment	X8	0.10	0.30	15	0.10	0.31	16	0.01	0.95	20	0.06	0.60	17	0.06	0.59	15
Forest rents	X9	0.02	0.93	91	0.12	0.43	12	0.15	0.27	9	0.18	0.25	9	0.14	0.22	10
Access to clean fuels and technologies for cooking	X10	0.11	0.38	14	0.06	0.55	18	0.04	0.77	19	0.05	0.67	18	0.06	0.59	14
GDP growth	X11	0.20	0.32	7	0.03	0.81	20	0.01	0.99	21	0.03	0.83	20	0.04	0.79	16
Population, total	X12	**0.36**	**0.04**	2	**0.31**	**0.04**	3	**0.31**	**0.04**	2	**0.29**	**0.05**	6	0.26	0.11	6
Rural population	X13	0.29	0.06	4	0.27	0.16	6	0.23	0.10	6	0.33	0.09	3	0.35	0.08	2
Urban population	X14	**0.29**	**0.04**	3	0.26	0.08	7	0.25	0.11	4	0.21	0.13	7	0.29	0.07	4
Labor force	X15	0.25	0.08	6	0.27	0.16	5	0.23	0.09	5	0.33	0.09	4	0.24	0.08	8
Death rate	X16	0.12	0.21	11	0.11	0.28	14	0.09	0.34	14	0.10	0.32	13	0.03	0.87	19
Birth rate	X17	0.11	0.27	12	0.18	0.10	9	0.16	0.14	8	0.17	0.11	10	**0.25**	**0.04**	7
Compulsory education	X18	0.11	0.28	13	0.11	0.25	13	0.10	0.28	12	0.12	0.23	11	0.11	0.24	11
Land area	X19	0.28	0.06	5	0.24	0.09	8	**0.31**	**0.04**	1	**0.30**	**0.05**	5	**0.30**	**0.05**	3
Terrestrial protected areas	X20	0.06	0.58	17	0.08	0.41	17	0.08	0.41	16	0.08	0.42	15	0.08	0.47	13
Arable land	X21	0.17	0.22	8	**0.33**	**0.04**	2	0.11	0.35	11	**0.34**	**0.02**	2	**0.45**	**0.01**	1

The q statistics of significant drivers ($P \leq 0.05$) are displayed in bold.

consumer price index, agricultural land area, goods import, urbanization population, goods export, and human birth rate. These six factors can be regarded as the leading factors affecting forest cover change this year in terms of their Q values which are all greater than 0.25. However, each driving factor showed a trend of fluctuation in the temporal scale. Consumer price index and urban population had significant effects in 2000 only, with q values of 0.48 and 0.29, respectively. The effect of the human birth rate on forest cover increased annually until it reached a significant level in 2019, with a value of 9 of 0.25. Total population, land area, and agricultural land area played important roles in the disturbance of forest cover all in this time span. The weight of the total population decreased annually, while the proportion of agricultural land increased. The former decreased from 0.36 down to 0.26 between 2019 and 2020, and the latter increased from 0.17 to 0.45 from 2019 to 2000.

To further explore the changes of interpreting power of forest disturbance based on the interaction of different driving factors, the sample analysis is implemented based on the demonstrated illustration of 2000, 2005, 2010, 2015, and 2019. The interactive factor detection mechanism of the leading factors picked on forest cover change was analyzed (shown in **Table 7**): The demonstration illustrated that each factor was significantly correlated and independent in rare cases. About 76.67% of the interaction categories among several leading factors were non-linear enhancement, more than 18.33% were dual factors enhancement, and only 5% were single factor independence. The explanatory power of factor interaction illustrates differentiation and shows the characteristics of fluctuation in timescale excluded that the interaction of X4 (goods import) and X14 (urban population) is increasing from 0.71 to 0.88 between 2019 and 2020.

5 DISCUSSION

Many scholars have been sorting out, improving, and verifying the development and evolution results of various direct driving factors under the UN SDGs constantly (Griggs et al., 2013; Weitz et al., 2019a; Weitz et al., 2019b; Guo, 2020; Guo et al., 2020; Guo et al., 2021). The IPBES global assessment provides a global review of the scientific evidence on major environmental and secondary social impacts of drivers of change, including the drivers, pressures, and responses sections of the DPSIR framework (Merino et al., 2019). The impact of geographical and climatic factors on forest cover change is likely to occur over a longer time series, but human activities are likely to be more rapid and straightforward, excluding accidental natural disasters, pest, and disease disasters or forest fires. The potential synergistic effect research method between the forest cover index in UN SDGs and human social and economic development has been illustrated further in terms of current research contributions.

Multiple driven factors play a role of different structure of the 6 principal components via global principal component analysis. The significant factors are influencing the forest cover change via the geographical detector monitoring all drivers annually. The two results of each analysis have been demonstrated slightly different subject to the inconsistency of research

TABLE 7 | Interactive detection results of spatial differentiation of forest cover change.

Factor interaction	Year 2000			Year 2005			Year 2010			Year 2015			Year 2019		
	q (Xi∩Xj)	q (Xi+Xj)	Type	q (Xi∩Xj)	q (Xi+Xj)	Type	q (Xi∩Xj)	q (Xi+Xj)	Type	q (Xi∩Xj)	q (Xi+Xj)	Type	q (Xi∩Xj)	q (Xi+Xj)	Type
X4∩X7	0.55	0.55	I	0.43	0.41	NE	0.78	0.32	NE	0.31	0.30	NE	0.30	0.18	NE
X4∩X12	0.68	0.43	NE	0.57	0.60	DE	0.48	0.49	DE	0.45	0.50	DE	0.50	0.40	NE
X4∩X14	0.71	0.36	NE	0.75	0.55	NE	0.77	0.44	NE	0.79	0.42	NE	0.88	0.43	NE
X4∩X19	0.73	0.35	NE	0.78	0.53	NE	0.63	0.50	NE	0.85	0.50	NE	0.79	0.45	NE
X4∩X21	0.43	0.24	NE	0.54	0.62	DE	0.42	0.30	NE	0.49	0.55	DE	0.83	0.60	NE
X7∩X12	0.87	0.84	NE	0.47	0.43	NE	0.70	0.45	NE	0.42	0.39	NE	0.80	0.30	NE
X7∩X14	0.93	0.77	NE	0.55	0.38	NE	0.82	0.39	NE	0.46	0.31	NE	0.54	0.32	NE
X7∩X19	0.83	0.76	NE	0.73	0.36	NE	0.83	0.45	NE	0.81	0.39	NE	0.46	0.34	NE
X7∩X21	0.75	0.65	NE	0.63	0.45	NE	0.81	0.25	NE	0.68	0.44	NE	0.54	0.48	NE
X12∩X14	0.72	0.66	NE	0.49	0.57	DE	0.79	0.56	NE	0.47	0.51	DE	0.43	0.54	DE
X14∩X19	0.54	0.57	DE	0.53	0.50	NE	0.50	0.57	DE	0.51	0.51	I	0.88	0.59	NE
X19∩X21	0.38	0.45	DE	0.66	0.56	NE	0.43	0.43	I	0.77	0.64	NE	0.80	0.75	NE

DE: Double enhancement q (Xi+Xj)>q (Xi∩Xj)>Max (q (Xi), q (Xj)); NE: Non-linear enhancement q (Xi∩Xj)>q (Xi+Xj); I: Independence q (Xi∩Xj) = q (Xi+Xj)

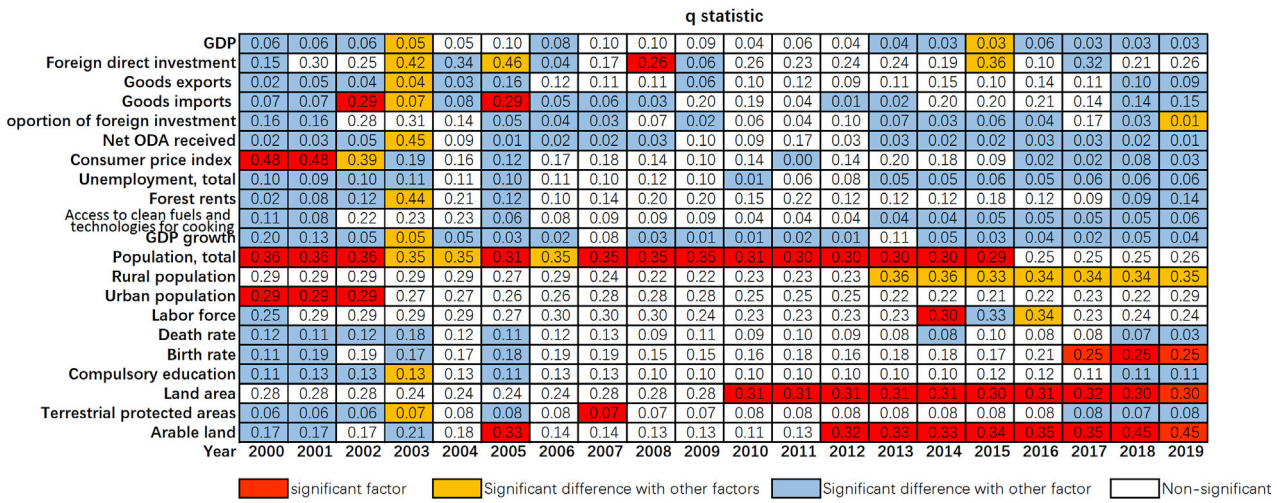

FIGURE 4 | Sketch map of differential superposition analysis and identification of significance, which of driving factors affecting forest cover change in the study area. Note: q value is the explanatory degree of driving factors to forest cover change in the study area. Red represents the driving factors with an explanatory degree of more than 95% confidence. Yellow represents the significant difference between the effects of the linear factor and other factors (≥3) on the spatial distribution of forest cover. Blue indicates that there is a significant between the effect of one factor and another or two factors on the spatial distribution of forest cover. White indicates no significant difference compared with other factors (95% confidence).

purpose, timescale, and precision. Global principal component analysis has placed all the data in the last 20 years on the same 3D observation model to abstract dimension reduction. The results of driving factors influencing the last 20 years have been extracted. The geographical detector has detected the significant drivers of forest cover change each year to verify and summarize the significant common factors over the past 20 years. The both extracted analyses of driving factors of forest cover change can be in support of making decision further.

A comprehensive analysis of 21 driving factors in 54 countries and regions in the target area for 20 consecutive years was completed by using geographical detectors (see **Figure 4**), and forest cover change and driving factors can be summarized into four stages: in the first stage, the factors causing the reduction of forest cover are complicated in the first 4 years of the 21st century (2000–2003). The consumer price index (CPI) has the significant and highest explanatory power, which can reflect the degree of inflation or deflation. The second is the total population. The pressure caused by the total and urban population also shows a significant impact. In addition, almost all other factors showed significantly different trends in the effects of two factors at the same time. In the second stage, the rapid growth of the total population in the study area was the most important and significant explanatory factor driving the rapid decrease of forest cover from 2004 to 2009. Meanwhile, the factors promoting regional economic development were different from population pressure and played a moderating role in influencing the change of forest cover gradually. In the third stage, the

expansion of cultivated land in the study area has a significant impact year by year from 2010 to 2015, especially in the area with a larger land area, the amount of deforestation and land conversion is greater. The growth of total population still has a relatively significant explanation for forest change, but the impact value decreases year by year. In Stage 4, the growth of the cultivated land for forest cover to reduce the influence of the significant increase from 2016 to 2019 (Williams et al., 2021), the chance of vast country in deforestation is greater, and with the development of economy, the total population bearing pressure on forest change is no longer significant; the neo-birth rate can significantly represent more of a certain explanation on forest cover change; in addition, changes in the number of rural population have significant differences with other factors (≥3) on the spatial distribution of forest cover, which becomes extremely complicated.

6 CONCLUSION

Forest cover change assessment has become one of the key indicators of Sustainable Development Goals (SDGs). Africa is experiencing unprecedented deforestation of forest resources. The study area is composed of data from 54 countries and regions from 2000 to 2020. Spatial autocorrelation analysis, global principal component analysis, and geographic detector model are used to illustrate the spatio-temporal patterns of forest cover change in Africa. The driving effects of population growth, economic and trade, social development, and cultivated land expansion on forest cover change in different timescale were analyzed in geographical spatial heterogeneity. Relevant conclusions are as follows:

1) The distribution of forests in Africa is extremely uneven. Based on the distributed quantities, Africa has the largest forest area in Central Africa, the second in East Africa, the third in West Africa, the fourth in South Africa, and the fifth in North Africa. In the past 20 years, the forest area decreased significantly, and the forest area decreased more significantly than that in West Africa including the Democratic Republic of the Congo, Angola, Tanzania, Mozambique, Sudan, Nigeria, Botswana and Cote d 'Ivoire, Zambia, Chad, Somalia, Ethiopia, Namibia, and Cameroon. Only 7 countries showed a net increase in the forest area during the 20 years, and only 5 countries showed a significant increase (over 200 square kilometers). Algeria, Morocco, and Tunisia in North Africa, East Africa, and Swaziland in South Africa have presented an increase status, but forest growth accounts for only 1% of the loss.
2) The forest cover change in the study area began to show the clustering from 2010 to 2012 and became more significant. The degree of spatial agglomeration increased year by year, and the spatial correlation of forest cover change between neighboring countries or regions became increasingly stronger. The LISA index was used to study the local agglomeration characteristics of forest cover change. The spatial clustering of hot spots was divided into two regions: first, Algeria as a hot spot clustering in Morocco and Tunisia increased forest cover area; in the other place, Zambia was the coldest spot with a significant clustering in Angola, the Democratic Republic of the Congo, and Tanzania have reduced their low-forest coverage to Angola, Rwanda, Uganda, Central African Republic, and Congo high–low zone.
3) Total population, land area, cultivated land, urban population, consumer price index, and birth rate are the main factors influencing the temporal evolution of forest cover change in Africa. It can be divided into four stages to interpret the different explanations and significance of each factor for forest cover change in the study area. In the first stage, the factors causing the reduction of forest cover are complicated in the first 4 years of the 21st century (2000–2003). The consumer price index (CPI) has the significant and highest explanatory power, which can reflect the degree of inflation or deflation. The second is the total population. The pressure caused by the total and urban population also shows a significant impact. In addition, almost all other factors showed significantly different trends in the effects of two factors at the same time. In the second stage, the rapid growth of the total population in the study area was the most important and significant explanatory factor driving the rapid decrease of forest cover from 2004 to 2009. Meanwhile, the factors promoting regional economic development were different from population pressure and played a moderating role in influencing the change of forest cover gradually. In the third stage, the expansion of cultivated land in the study area has a significant impact year by year from 2010 to 2015, especially in the area with a larger land area, the amount of deforestation and land conversion is greater. The growth of total population still has a relatively significant explanation for forest change, but the impact value decreases year by year. In Stage 4, the growth of the cultivated land for forest cover to reduce the influence of the significant increase from 2016 to 2019, the chance of vast country in deforestation is greater, and with the development of economy, the total population bearing pressure on forest change is no longer significant; the neo-birth rate can significantly represent more of a certain explanation on forest cover change; in addition, changes in the number of rural population have significant differences with other factors (≥3) on the spatial distribution of forest cover, which becomes extremely complicated.

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

HX, JL, NF, and GH conceived the research project. HX and HW performed simulations and wrote the manuscript. XZ, TL, ZZ, WW, RY, YG, BC, and QC contributed to model development. All other authors contributed to data and commented on or approved the manuscript.

REFERENCES

Afdb, A., and Kouakou, K. L. (2010). Gender, Poverty and Environmental Indicators on African Countries 2010. Available at: https://www.afdb.org/fileadmin/uploads/afdb/Documents/Publications/Gender%20Poverty%20and%20Environmental%20Indicators%20on%20African%20Countries%202010.pdf (Accessed March 20, 2021).

Anselin, L. (1995). Local Indicators of Spatial Association-LISA. *Geographical Anal.* 27 (2), 93–115.

Assogba, N. P., and Zhang, D. (2018). *Factors Influencing the Conservation of Tropical Forest Resources in West Africa. 2018 Conference*. Vancouver, British Columbia: International Association of Agricultural Economists.

Barlow, J., Gardner, T. A., Araujo, I. S., Avila-Pires, T. C., Bonaldo, A. B., Costa, J. E., et al. (2007). Quantifying the Biodiversity Value of Tropical Primary, Secondary, and Plantation Forests. *Proc. Natl. Acad. Sci.* 104, 18555–18560. doi:10.1073/pnas.0703333104

Beckman, J., Sands, R., Riddle, A., Lee, T., and Walloga, J. M. (2017). *International Trade and Deforestation: Potential Policy Effects via a Global Economic Model*. Washington: Economic Research Report.

Bonan, G. B. (2008). Forests and Climate Change: Forcings, Feedbacks, and the Climate Benefits of Forests. *Science* 320 (5882), 1444–1449. doi:10.1126/science.1155121

Curtis, P. G., Slay, C. M., Harris, N. L., Tyukavina, A., and Hansen, M. C. (2018). Classifying Drivers of Global forest Loss. *Science* 361 (6407), 1108–1111. doi:10.1126/science.aau3445

Deacon, R. T. (1994). Deforestation and the Rule of Law in a Cross-Section of Countries. *Land Econ.* 70, 414–430. doi:10.2307/3146638

Defries, R. S., Rudel, T., Uriarte, M., and Hansen, M. (2010). Deforestation Driven by Urban Population Growth and Agricultural Trade in the Twenty-First century. *Nat. Geosci* 3 (3), 178–181. doi:10.1038/ngeo756

Diffenbaugh, N. S., and Giorgi, F. (2012). Climate Change Hotspots in the Cmip5 Global Climate Model Ensemble. *Climatic Change* 114 (3-4), 813–822. doi:10.1007/s10584-012-0570-x

Edwards, D. P., Koh, L. P., and Laurance, W. F. (2012). Indonesia's Redd+ Pact: Saving Imperilled Forests or Business as Usual? *Biol. Conservation* 151 (1), 41–44. doi:10.1016/j.biocon.2011.10.028

FAO (2015). Global Forest Resources Assessment 2015. Available at: www.fao.org/forest-resources-assessment/past-assessments/fra-2015/en/(Accessed May 1, 2021).

FAO (2020). *Global Forest Resources Assessment 2020 – Key Findings*. Rome: FAO. doi:10.4060/ca8753en

Forzieri, G., Girardello, M., Ceccherini, G., Spinoni, J., Feyen, L., Hartmann, H., et al. (2021). Emergent Vulnerability to Climate-Driven Disturbances in European Forests. *Nat. Commun.* 12 (1), 1081. doi:10.1038/s41467-021-21399-7

Garzuglia, M., and Saket, M. (2003). *Wood Volume and Woody Biomass: Review of FRA 2000 Estimates*. Rome: Forest Resources Assessment Programme Working Paper. Available at: https://agris.fao.org/agris-search/search.do?recordID=XF2016031581 (Accessed March 15, 2021).

Gbetnkom, D. (2009). Forest Depletion and Food Security of Poor Rural Populations in Africa: Evidence from cameroon. *J. Afr. Economies* 18 (2), 261–286. doi:10.1093/jae/ejn012

Geist, H. J., and Lambin, E. F. (2002). Proximate Causes and Underlying Driving Forces of Tropical Deforestation. *Bioscience* 52 (2), 143–150. doi:10.1641/0006-3568(2002)052[0143:pcaudf]2.0.co;2

Giam, X. (2017). Global Biodiversity Loss from Tropical Deforestation. *Proc. Natl. Acad. Sci. USA* 114, 5775–5777. doi:10.1073/pnas.1706264114

Glaser, G. (2012). Policy: Base Sustainable Development Goals on Science. *Nature* 491 (7422), 35. doi:10.1038/491035a

Goodchild, M. F., and Glennon, J. A. (2010). Crowdsourcing Geographic Information for Disaster Response: a Research Frontier. *Int. J. Digital Earth* 3 (3), 231–241. doi:10.1080/17538941003759255

Griggs, D., Stafford-Smith, M., Gaffney, O., Rockström, J., Öhman, M. C., Shyamsundar, P., et al. (2013). Sustainable Development Goals for People and Planet. *Nature* 495 (7441), 305–307. doi:10.1038/495305a

Guo, H. D. (2020). *Big Earth Data in Support of the Sustainable Development Goals*. Beijing: Science Press, EDP Sciences.

Guo, H., Chen, F., Sun, Z., Liu, J., and Liang, D. (2021). Big Earth Data: a Practice of Sustainability Science to Achieve the Sustainable Development Goals. *Sci. Bull.* 66 (11), 1. doi:10.1016/j.scib.2021.01.012

Guo, H., Nativi, S., Liang, D., Craglia, M., Wang, L., Schade, S., et al. (2020). Big Earth Data Science: an Information Framework for a Sustainable Planet. *Int. J. Digital Earth* 13 (7), 743–767. doi:10.1080/17538947.2020.1743785

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., et al. (2013). High-resolution Global Maps of 21st-century forest Cover Change. *SCIENCE* 342, 850–853. doi:10.1126/science.1244693

Harris, N. L., Gibbs, D. A., Baccini, A., Birdsey, R. A., de Bruin, S., Farina, M., et al. (2021). Global Maps of Twenty-First century forest Carbon Fluxes. *Nat. Clim. Chang.* 11, 234–240. doi:10.1038/s41558-020-00976-6

Heitzig, J., Kittel, T., Donges, J. F., and Molkenthin, N. (2016). Topology of Sustainable Management of Dynamical Systems with Desirable States: from Defining Planetary Boundaries to Safe Operating Spaces in the Earth System. *Earth Syst. Dynam.* 7, 21–50. doi:10.5194/esd-7-21-2016

Hoang, N. T., and Kanemoto, K. (2021). Mapping the Deforestation Footprint of Nations Reveals Growing Threat to Tropical Forests. *Nat. Ecol. Evol.* 5, 845–853. doi:10.1038/s41559-021-01417-z

Hosonuma, N., Herold, M., De Sy, V., De Fries, R. S., Brockhaus, M., Verchot, L., et al. (2012). An Assessment of Deforestation and forest Degradation Drivers in Developing Countries. *Environ. Res. Lett.* 7 (4), 044009. doi:10.1088/1748-9326/7/4/044009

Jha, C., S., Manaswini, G., Singh, S., and Sathish, K. V. (2016). Conservation Priorities of forest Ecosystems: Evaluation of Deforestation and Degradation Hotspots Using Geospatial Techniques. *Ecol. Eng. J. Ecotechnology* 91, 333–342. doi:10.1016/j.ecoleng.2016.03.007

Kaimowitz, A. D. (1999). Rethinking the Causes of Deforestation: Lessons from Economic Models. *World Bank Res. Observer* 14 (1), 73–98. doi:10.2307/3986539

Kaimowitz, D., Mendez, P., Puntodewo, A., and Vanclay, J. K. (2002). *Spatial Regression Analysis of Deforestation in Santa Cruz, Bolivia*. Gainesville: OAI.

Kim, M. C., and Kim, J.-K. (2010). "Digital divide: Conceptual Discussions and prospect," in *The Human Society and the Internet-Related Socio-Economic Issues* (Springer), 78–91.

Kouakou, K. L. (2005). *Gender, Poverty and Environmental Indicators on African Countries[J]*. Tunis, Tunisia: Gender, Poverty and Environmental Indicators on African Countries.

Macdicken, K. G. (2015). Global forest Resources Assessment 2015: what, Why and How? *For. Ecol. Manag.* 352, 3–8. doi:10.1016/j.foreco.2015.02.006

Mainardi, S. (1998). An Econometric Analysis of Factors Affecting Tropical and Subtropical Deforestation/'N Ekonometriese Analise Oor Faktore Wat Tropiese En Subtropiese Ontbossing Beïnvloed. *Agrekon* 37 (1), 23–65. doi:10.1080/03031853.1998.9523494

Merino, L., Viña, A., García Frapolli, E., Ainul Hussain, S., AkongMinang, P., and Nagabhatla, N. (2019). "Chapter 2. Status and Trends; Indirect and Direct Drivers of Change," in *IPBES Global Assessment on Biodiversity and Ecosystem Services* (Bonn, Germany: IPBES secretariat).

Midgley, G. F., and Bond, W. J. (2015). Future of African Terrestrial Biodiversity and Ecosystems under Anthropogenic Climate Change. *Nat. Clim Change* 5, 823–829. doi:10.1038/NCLIMATE2753

Moran, P. A., and Ap, P. (1950). Notes on Continuous Stochastic Phenomena. *Biometrika* 37 (1-2), 17–23. doi:10.1093/biomet/37.1-2.17

Mushi, H., Yanda, P. Z., and Kleyer, M. (2020). Socioeconomic Factors Determining Extraction of Non-timber forest Products on the Slopes of Mt. Kilimanjaro, tanzania. *Hum. Ecol.* 48 (1), 695–707. doi:10.1007/s10745-020-00187-9

National Remote Sensing Center of China (2019). Global Ecosystem and Environment Observation Analysis Research Cooperation (GEOARC) Global forest Cover and Changes. Available at: http://www.chinageoss.cn/geoarc/en/news/2019A.html (Accessed March 15, 2021).

Parmesan, C., and Yohe, G. (2003). A Globally Coherent Fingerprint of Climate Change Impacts across Natural Systems. *Nature* 421 (6918), 37–42. doi:10.1038/nature01286

Pendrill, F., Persson, U. M., Godar, J., and Kastner, T. (2019). Deforestation Displaced: Trade in forest-risk Commodities and the Prospects for a Global forest Transition. *Environ. Res. Lett.* 14 (5), 055003. doi:10.1088/1748-9326/ab0d41

Réjou-Méchain, M., Mortier, F., Bastin, J.-F., Cornu, G., Barbier, N., Bayol, N., et al. (2021). Unveiling African Rainforest Composition and Vulnerability to Global Change. *Nature* 593 (7857), 90–94. doi:10.1038/s41586-021-03483-6

Rudel, T., and Roper, J. (1997). The Paths to Rain forest Destruction: Crossnational Patterns of Tropical Deforestation, 1975-90. *World Dev.* 25 (1), 53–65. doi:10.1016/S0305-750X(96)00086-1

Ruf, F., Schroth, G., and Doffangui, K. (2015). Climate Change, cocoa Migrations and Deforestation in West Africa: what Does the Past Tell Us about the Future? *Sustain. Sci.* 10 (1), 101–111. doi:10.1007/s11625-014-0282-4

Sandel, B., and Svenning, J.-C. (2013). Human Impacts Drive a Global Topographic Signature in Tree Cover. *Nat. Commun.* 4, 1. doi:10.1038/ncomms3474

Sannigrahi, S., Bhatt, S., Rahmat, S., Paul, S. K., and Sen, S. (2018). Estimating Global Ecosystem Service Values and its Response to Land Surface Dynamics during 1995–2015. *J. Environ. Manage.* 223, 115–131. doi:10.1016/j.jenvman.2018.05.091

Shimada, M., Itoh, T., Motooka, T., Watanabe, M., Shiraishi, T., Thapa, R., et al. (2014). New Global forest/non-forest Maps from ALOS PALSAR Data (2007-2010). *Remote Sensing Environ.* 155, 13–31. doi:10.1016/j.rse.2014.04.014

Stévart, T., Dauby, G., Lowry, P. P., Blach-Overgaard, A., Droissart, V., Harris, D. J., et al. (2019). A Third of the Tropical African flora Is Potentially Threatened with Extinction. *Sci. Adv.* 5 (11), eaax9444. doi:10.1126/sciadv.aax9444

The World Bank (2021). World Development Indicators. Available at: https://datacatalog.worldbank.org/search/dataset/0037712 (Accessed May 1, 2021).

Timothy, E. L., Mike, U., and Marks, D. (2004). The Dynamics of Rainfall Interception by a Seasonal Temperate Rainforest. *Agric. For. Meteorology* 124 (3–4), 171–191. doi:10.1016/j.agrformet.2004.01.010

UNESCO (2021). Education in Afica. Available at: http://uis.unesco.org/en/topic/education-africa (Accessed May 1, 2021).

United Nations, Department of Economic and Social Affairs, Population Division (2019). World Population Prospects 2019, Custom Data Acquired via Website. Available at: https://www.un.org/development/desa/pd/node/1114 (Accessed March 15, 2021).

United Nations, Department of Economic and Social Affairs, Population Division (2017). *World Population Prospects: The 2017 Revision, Key Findings and Advance Tables*. New York, NY: Working Paper No. ESA/P/WP/248.

Wang, J. F., and Xu, C. D. (2017). Geodetector:principle and Prospective. *Acta GeographicaSinica* 72 (1), 116–134. doi:10.1016/j.ecolind.2016.08.001

Weitz, N., Carlsen, H., Skanberg, K., Dzebo, A., and Viaud, V. (2019b). *SDGs and the Environment in the EU: A Systems View to Improve Coherence*. Stockholm, Sweden: Stockholm Environment Institute. Available at: https://www.sei.org/wp-content/uploads/2019/10/sei-2019-pr-weitz-sdg-synergies-eu-env.pdf (Accessed July 16, 2021).

Weitz, N., Carlsen, H., and Trimmer, C. (2019a). *SDG Synergies: An Approach for Coherent 2030 Agenda Implementation*. Stockholm, Sweden: Stockholm Environment Institute. Available at: https://www.sei.org/publications/sdg-synergies-factsheet/(Accessed July 16, 2021).

Williams, D. R., Clark, M., Buchanan, G. M., Ficetola, G. F., Rondinini, C., and Tilman, D. (2021). Proactive Conservation to Prevent Habitat Losses to Agricultural Expansion. *Nat. Sustain.* 4, 314–322. doi:10.1038/s41893-020-00656-5

Wilson, K. B., Hanson, P. J., Mulholland, P. J., Baldocchi, D. D., and Wullschleger, S. D. (2001). A Comparison of Methods for Determining forest Evapotranspiration and its Components: Sap-Flow, Soil Water Budget, Eddy Covariance and Catchment Water Balance. *Agric. For. Meteorology* 106 (2), 153–168. doi:10.1016/s0168-1923(00)00199-4

World Bank Group (2020). September 2020 PovcalNet Update: What's New (English). Available at: http://documents.worldbank.org/curated/en/144451599591697923/September-2020-PovcalNet-Update-What-s-New (Accessed September 1, 2020).

Zhang, X., Long, T., He, G., Guo, Y., Yin, R., Zhang, Z., et al. (2020). Rapid Generation of Global forest Cover Map Using Landsat Based on the forest Ecological Zones. *J. Appl. Rem. Sens.* 14 (2), 1. doi:10.1117/1.JRS.14.022211

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.

frontiers in Environmental Science

ORIGINAL RESEARCH
published: 14 January 2022
doi: 10.3389/fenvs.2021.756227

YOLOv4-Lite–Based Urban Plantation Tree Detection and Positioning With High-Resolution Remote Sensing Imagery

*Yueyuan Zheng and Gang Wu**

College of Information, Beijing Forestry University, Beijing, China

Automatic tree identification and position using high-resolution remote sensing images are critical for ecological garden planning, management, and large-scale environmental quality detection. However, existing single-tree detection methods have a high rate of misdetection in forests not only due to the similarity of background and crown colors but also because light and shadow caused abnormal crown shapes, resulting in a high rate of misdetections and missed detection. This article uses urban plantations as the primary research sample. In conjunction with the most recent deep learning method for object detection, a single-tree detection method based on the lite fourth edition of you only look once (YOLOv4-Lite) was proposed. YOLOv4's object detection framework has been simplified, and the MobileNetv3 convolutional neural network is used as the primary feature extractor to reduce the number of parameters. Data enhancement is performed for categories with fewer single-tree samples, and the loss function is optimized using focal loss. The YOLOv4-Lite method is used to detect single trees on campus, in an orchard, and an economic plantation. Not only is the YOLOv4-Lite method compared to traditional methods such as the local maximum value method and the watershed method, where it outperforms them by nearly 46.1%, but also to novel methods such as the Chan-Vese model and the template matching method, where it outperforms them by nearly 26.4%. The experimental results for single-tree detection demonstrate that the YOLOv4-Lite method improves accuracy and robustness by nearly 36.2%. Our work establishes a reference for the application of YOLOv4-Lite in additional agricultural and plantation products.

Keywords: YOLOv4-Lite, urban plantation, tree detection, tree position, high-resolution remote sensing image

OPEN ACCESS

Edited by:
Peng Liu,
Institute of Remote Sensing and Digital Earth (CAS), China

Reviewed by:
Konstantinos Demertzis,
International Hellenic University, Greece
Zhibin Ren,
Northeast Institute of Geography and Agroecology (CAS), China

****Correspondence:***
Gang Wu
wugang@bjfu.edu.cn

Specialty section:
This article was submitted to Environmental Informatics and Remote Sensing, a section of the journal Frontiers in Environmental Science

Received: *10 August 2021*
Accepted: *29 November 2021*
Published: *14 January 2022*

Citation:
Zheng Y and Wu G (2022) YOLOv4-Lite–Based Urban Plantation Tree Detection and Positioning With High-Resolution Remote Sensing Imagery. Front. Environ. Sci. 9:756227. doi: 10.3389/fenvs.2021.756227

1 INTRODUCTION

The smallest tree entity that makes up a terrestrial ecosystem is a single tree. Single-tree detection and positioning are critical components of precision forestry (Dimitrios and Azadeh, 2021; Dimitrios et al., 2021). Precision forestry is critical for reducing environmental pollution caused by catkins (BAILU, 2018) and monitoring the environmental impact of economic forests (Wang, 2019; Dong et al., 2020; Zhibin et al., 2020). With the advancement and popularization of the remote sensing image processing technology, the efficient collection of massive amounts of information about individual trees and the establishment of single-tree databases form the foundation for accurate and

intensive urban forestry management and are also the developmental goal of a new generation of intelligent forestry (Olli et al., 2017; Gomes et al., 2018). Indeed, some individual forestry fields in China have implemented intensive management at the single-tree level, such as managing ancient and famous trees, managing female willow trees precisely in Beijing, and managing fruit trees (Xiao et al., 2021). However, these are primarily accomplished through traditional ground surveys, which are inefficient and require a significant amount of time. Single-tree extraction from remote sensing images is a critical technology for efficiently constructing a single-tree database built on the foundation of single-tree detection from remote sensing images. Remote sensing platforms and associated processing technologies have advanced rapidly in recent years (Toth and Jozkow, 2016). Although active remote sensing technologies such as lidar have been used to detect single trees (including their position and height) and to build inversion models (Bouvier et al., 2015), their costs are prohibitively high, and they are not suitable for large-scale acquisition. However, single-tree extraction using high-resolution optical remote sensing images remains a research hotspot due to its strong technical foundation, large image scale, widespread availability, and ease of promotion and application (Wulder et al., 2000; Picos et al., 2020). Numerous studies on passive remote sensing demonstrate that the data source, season, tree growth (Liu et al., 2018), and the location method selection (Larsen et al., 2011) are all significant factors affecting the effect of a single-tree location. Regarding the problem of locating a single tree, scholars have successively proposed two types of methods, one is the direct location method, such as the local maximum method (Wulder et al., 2000) and threshold segmentation, for locating a single-tree tip. The other is individual tree crown detection and delineation (ITCD) (Yinghai et al., 2015), such as watershed methods (Wang et al., 2004; Zheng et al., 2017), the valley floor tracking method (Gougeon, Fran ois, 2014), the regional growth method (Jiang et al., 2016), and the marker process method (Gomes et al., 2018). Recently, scholars have experimented with a variety of novel positioning techniques, including peak climbing (Zhang et al., 2014), CV model (Nasor and Obaid, 2021), template matching (Hashim et al., 2020), and machine learning (Dong et al., 2019). Currently, the accuracy of the single-tree positioning method is insufficient to meet practical requirements, and no method is adaptable to a variety of forests (Zhang et al., 2019). ICD's primary challenges are as follows: 1) close spacing between single trees and canopy intersections results in over or under-segmentation within a single crown and between overlapping crowns, resulting in misdetection and missed detection of single-tree positioning. 2) The region background color is similar to that of a single tree, and the region background is complex and diverse. The background is incorrectly identified as the tree's crown, and the region exhibits the phenomenon of under-segmentation, resulting in the problem of error and missing detection. 3) It is unclear how the sunlight shadow affects the shape of the trees and canopy outline. There is no way to collect all single-tree templates. 4) When multiple trees of varying sizes cover the same area, a large number of saplings can easily escape detection. As a result, it is necessary to investigate a more efficient and stable method for a single-tree location.

Traditional object detection methods can be broadly classified into three categories: 1) region selection (Zhou et al., 2021), such as sliding window; 2) classifiers, such as support vector machines (SVMs) (Rau et al., 2021); 3) feature extraction, such as scale-invariant feature transform (SIFT) (Liang et al., 2020) and histogram of oriented gradient (HOG) (Syaputra et al., 2021). There are two significant issues with it. First, the region selection is not targeted, is time-consuming, and contains numerous redundant windows. On the other hand, the characteristics have a low degree of robustness, and many small objects are overlooked. Following the advent of deep learning, object detection has made a significant leap forward (Zhao et al., 2020). There are two primary areas of emphasis: 1) CNNs based on region proposal (R-CNN) (Ma et al., 2020), spatial pyramid pooling networks (SPP-Nets) (Wang et al., 2020), and fast R-CNN (Garcia-Ortiz et al., 2020). Faster R-CNN is used to represent deep learning object detection algorithms. 2) Regression-based deep learning object detection algorithms are represented as you only look once (YOLO) (Chaitanya et al., 2020). The former predicts the speeds of between 7 and 18 frames per second, which is too time-consuming. The latter makes use of the regression concept, which determines an input image and directly regresses the object boundary and object category of multiple locations in the image, significantly speeding up the prediction process. It has a long history of use in medicine, fishing, construction, and various other fields. This study performs the urban single-tree detection and positioning using the deep learning network model YOLOv4 (Richey and Shirvaikar, 2021). This model has a high detection speed and is capable of multi-object detection. A YOLOv4-Lite single-tree detection method is proposed to further integrate the YOLOv4 network model (Meneghetti et al., 2021), simplify the entire feature extraction network, optimize loss, and enhance sample data.

2 METHODOLOGIES

2.1 Data Enhancement

By referencing CutMix (Artés-hernández et al., 2017) data enhancement, the number of reconstructed images is increased from two to four using mosaic (Hofmann, 2000) data enhancement. First, four images from the dataset are chosen. Then, as illustrated in **Figure 1**, the four images are flipped, scaled, rotated, and gamut-adjusted, and other processing is applied. Finally, the four images are scaled and spliced together to create a single image containing the actual box. The blank spaces are filled with gray to align the training images with the network's characteristic size. One of the significant advantages of mosaic data enhancement approach is that it provides a rich background for object detection and can compute data from four images at once during the batch normalization calculation.

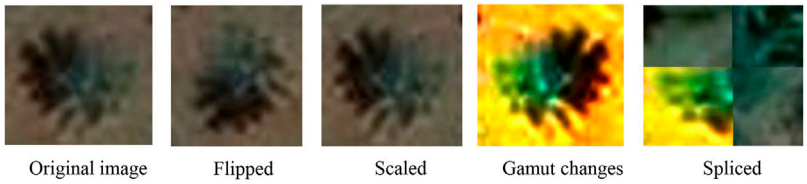

FIGURE 1 | Image of data enhancement.

TABLE 1 | Inverted residual with linear bottleneck.

The inverted residual with linear bottleneck
↓
Conv2d 1×1 Up sampling
Batch norm
ReLU6
↓
Depthwise Conv2d 3×3
Batch norm
ReLU6
↓
Conv2d 1×1 Down sampling
Batch norm
↓
Output

2.2 MobileNetv3

Deeply separable convolution is a factorized operation that can be decomposed into depthwise (***Dwise***) (Chollet, 2017) and pointwise convolution. For each input channel, depthwise convolution employs a unique convolution kernel. A convolution check should be an input channel. For each input channel, depthwise convolution employs a unique convolution kernel. Pointwise convolution changes the number of output channels by employing a 1 × 1 convolution kernel. While the overall effect is similar to that of standard convolution, this step operation significantly reduces the amount of computation and the number of model parameters. Three steps are required to compute the inverted residual with a linear bottleneck. As illustrated in **Table 1**, dimension is increased through the use of 1 × 1 convolution. Then, using deeply separable convolution, features are extracted. Finally, dimension is reduced using 1 × 1 convolution, and a residual edge is added directly from input to output.

The formula for ***mish*** activation function is shown in **(1)**.

$$\boldsymbol{mish} = \boldsymbol{x}\ \boldsymbol{tanh}\left(\boldsymbol{ln}\left(\mathbf{1} + \boldsymbol{e}^{\boldsymbol{x}}\right)\right), \tag{1}$$

where $\boldsymbol{x}$ represents the output of the previous network layer and the input of this network layer. The $\boldsymbol{h} - \boldsymbol{swish}$ activation function is shown in **formulas (2)** and **3**. The $\boldsymbol{h} - \boldsymbol{swish}$ activation function has the advantages of reducing the computation and improving performance, especially in a deep network.

$$\boldsymbol{ReLU6} = \boldsymbol{min}\left(\mathbf{6},\ \boldsymbol{max}\left(\mathbf{0}, \boldsymbol{x}\right)\right), \tag{2}$$

$$\boldsymbol{h} - \boldsymbol{swish}\left(\boldsymbol{x}\right) = \boldsymbol{x}\frac{\boldsymbol{ReLU6}\left(\boldsymbol{x} + \mathbf{3}\right)}{\mathbf{6}}. \tag{3}$$

The backbone network uses MobileNetv3, which employs a unique BNECK architecture. The structure of BNECK is shown in **Figure 2**, where ***NL*** stands for different activation functions. ***ReLU*** denotes the activation function of the fully convolutional (***FC***) layer, and ***Pool*** stands for the pooling layer. MobileNetv3 first adopts 1 × 1 convolution for dimension enhancement, and then carries out 3 × 3 depthwise convolution for feature extraction. It then adds an attention mechanism on the channel of a feature layer, which is two fully convolutional layers. This attention mechanism is to average pool the results of 3 × 3 deep separable convolution and then carries out two fully connected neural network processing. The feature layer resulting from this attention mechanism is multiplied by the feature layer resulting from the depthwise convolution of 3 × 3. Finally, 1 × 1 convolution is used to adjust the dimensions.

The detailed structural parameters of the entire MobileNetv3 are listed in **Table 2**. The first column, "Input," denotes the shape changes associated with each MobileNetv3 feature layer. The second column, "Operator," indicates the block structure through which each feature layer will pass. As it can be seen, feature extraction occurs *via* a variety of BNECK structures in MobileNetv3. The third and fourth columns indicate the number of channels after inverting the residual structure in BNECK and the number of channels at the characteristic layer when input to BNECK, respectively. The fifth column, "SE," indicates whether or not this level introduces attention mechanisms. The sixth column, "NL," represents the type of activation function, that is, "HS" for $\boldsymbol{h} - \boldsymbol{swish}$ and "RE" for ***ReLU***. The seventh column, "S," represents the step size used for each block structure.

2.3 Feature Pyramid

As illustrated in **Figure 3**, the feature pyramid is composed of two components: SPP and PANet. Their primary purpose is to improve feature extraction. "Concat + Conv a×a" is a deeply separable convolution, where "a×a" denotes depthwise convolution with aa's convolution kernel. The result of the backbone network is convolved by 3 × 3 in the SPP structure, and then the maximum pooling of 5, 9, and 13 is performed,

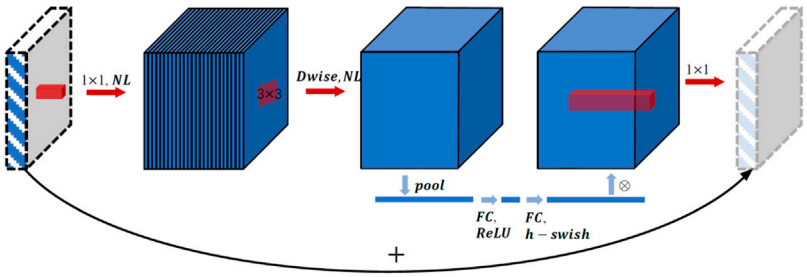

FIGURE 2 | Structure of BNECK.

TABLE 2 | Structure of MobileNetv3.

Input	Operator	Exp size	#Out	SE	NL	S
224 × 224 × 3	Conv2d	—	16	—	HS	2
112 × 112 × 16	BNECK, 3 × 3	16	16	—	RE	1
112 × 112 × 16	BNECK, 3 × 3	64	24	—	RE	2
56 × 56 × 24	BNECK, 3 × 3	72	24	—	RE	1
56 × 56 × 24	BNECK, 3 × 3	72	40	√	RE	2
28 × 28 × 40	BNECK, 3 × 3	120	40	√	RE	1
28 × 28 × 40	BNECK, 3 × 3	120	40	√	RE	1
28 × 28 × 40	BNECK, 3 × 3	240	80	—	HS	2
14 × 14 × 80	BNECK, 3 × 3	200	80	—	HS	1
14 × 14 × 80	BNECK, 3 × 3	184	80	—	HS	1
14 × 14 × 80	BNECK, 3 × 3	184	80	—	HS	1
14 × 14 × 80	BNECK, 3 × 3	480	112	√	HS	1
14 × 14 × 112	BNECK, 3 × 3	672	112	√	HS	1
14 × 14 × 112	BNECK, 5 × 5	672	160	√	HS	2
7 × 7 × 160	BNECK, 5 × 5	960	160	√	HS	1
7 × 7 × 160	BNECK, 5 × 5	960	160	√	HS	1
7 × 7 × 160	Conv2d, 1 × 1	—	960	—	HS	1
7 × 7 × 960	Pool, 7 × 7	—	—	—	—	1
1 × 1 × 960	Conv2d, 1 × 1, NBN	—	1,280	—	HS	1
1 × 1 × 1,280	Conv2d, 1 × 1, NBN	—	K	—	—	1

finally resulting in 3 × 3 profoundly separable convolutions. PANet's structure is depicted in **Figure 3**. By substituting deeply separable convolution blocks for standard convolution, the model's parameter count can be reduced.

2.4 YOLO Head and Parameter Controller

As illustrated in **Figure 3**, YOLO Head makes predictions using the extracted features from the backbone network and strengthened by the feature pyramid. YOLO Head is primarily composed of a convolution of 3 × 3 and a convolution of 1 × 1. Convolution of 3 × 3 is used to integrate all features, whereas convolution of 1 × 1 is used for prediction.

To reduce the number of parameters and improve the detection accuracy of the whole network, the α parameter is set at feature pyramid and Yolo Head, and the formula is as follows:

$$\gamma = \alpha \times \beta \tag{4}$$

where γ is the number of channels per convolution, β is quantitative and is the value each convolution needs to set in advance, and α is a global variable. The number of channels in a convolutional network can be controlled by adjusting the α parameter.

2.5 Loss Function

Due to the unbalanced and difficult-to-classify nature of the single-tree data collected in plantation monitoring, the YOLOv4-Lite uses focal loss to optimize the classification loss (Ruihuan hou, 2021; YuZ, 2021). In a one-stage model, the focal loss can be used to correct the background classification imbalance and the serious imbalance of positive and negative samples. The loss function in the YOLOv4 is divided into three components: position loss, confidence loss, and class loss. By combining the cross-entropy loss function, the sigmoid is calculated. Cross-entropy measures the distance between the actual output of the activation function and the expected output value in a multi-classification task. The lower the cross-entropy, the more compact the probability distribution. The following is the calculation formula:

$$L = -y \log \dot{y}, \tag{5}$$

where L represents the cross-entropy, y represents the expected output, and $\dot{y}$ denotes the actual output of the activated function. The calculation of improved focal loss is as follows:

$$L_f = -\dot{\alpha}\left[1 - \dot{y}\right]^{\dot{\beta}} \log \dot{y}, \tag{6}$$

where L_f is the improved cross-entropy, $\dot{\alpha}$ is 0.25, $\dot{\beta}$ is 2, the role of $\dot{\alpha}$ factor is to balance the number of samples, and the role of $\dot{\beta} > 0$ is to reduce easily classified objects.

3 EXPERIMENTS

3.1 Experimental Platform

This study's experimental model is built on the Keras = 2.1.5 framework, *Python* = 3.6, and the Spyder platform. The model is trained on an experimental 11th Gen Intel(R) Core (TM) i7-11800H at 2.30GHz, NVIDIA GeForce RTX 3060 Laptop GPU, and 16.0 GB of RAM running Windows x64.

3.2 Datasets

3.2.1 Dataset Description

There are specific management requirements for individual plants for urban plantation, such as catkin management and wind-fallen tree management. Furthermore, the technical foundation is sound. The urban plantation is classified into three types: campus, orchard, and economic plantation. To

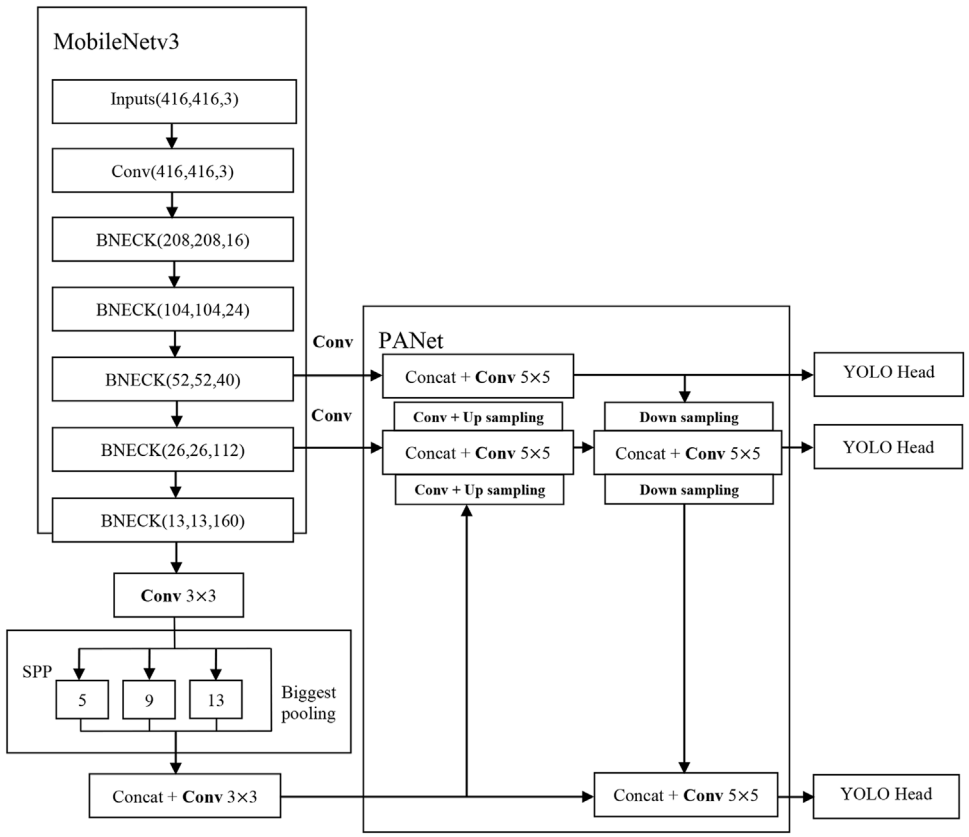

FIGURE 3 | Schematic diagram of YOLOv4-Lite.

verify the single-tree detection experiment, sample sites from these three types of urban plantations were chosen.

In June 2019, the green trees surrounding Beijing Forestry University's basketball court were located at 11,620′8.76″E, 400′6.52″N. The campus's primary tree species were broad-leaved *Sophora japonica* Linn and steamed bread willow (*Salix matsudana* var. *matsudana* F. umbraculifera Rehd.). In December 2017, the litchi orchard was located in Shenzhen's Bao 'a District at 11,353′26.34″E, 2,238′41.22″N. The image of a section of a palm plantation in Phang Nga, Thailand, was taken at 9,820′53.22″ E, 827′18.45″ N. Google Earth images with a spatial resolution of 0.27m, a scale of 800:1, and a visual field height of 1 km were used in our experiment.

Due to Thailand's extensive palm cultivation, the terrain is varied. We sampled palm trees from three different types of palm plantation: open plantation, complex background plantation, and dense plantation.

For each of the five sample locations, the entire image was predicted using local image features. The training data are divided into two sections. To begin, each sample site's image was divided into numerous small pieces. Then, from a large number of small pieces of training data, a small number of images were selected. The remaining training data were gathered in the vicinity of the sample sites. Data enhancement was used to increase the training dataset to 600 pieces. Individual trees of various tree species, backgrounds, and sizes are included in the training data. Diverse backgrounds are incorporated to increase the variety of elements and the training effect.

TABLE 3 | Training model parameters.

Training model	YOLOv4-Lite
Epochs	400
Batch sizes	8
Input image size	416×416
A	2
Score	0.2
Weight file size (MB)	47.5

3.2.2 Experiment Parameter

The experiment's primary parameters are listed in **Table 3**. The term "Epoch" refers to the process by which a complete dataset passes through a neural network once and returns once. When a large amount of data cannot be processed in a single pass through the neural network, the dataset must be divided into several "batch sizes." Each batch size corresponds to a new small dataset; the batch size parameter specifies the size of the new small dataset. The term "score" refers to a measure of confidence. The "weight file size (MB)" column indicates the amount of space consumed by the YOLOv4-Lite model.

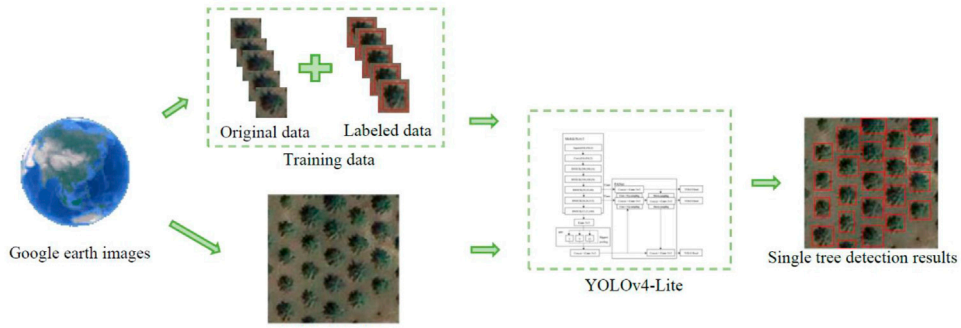

FIGURE 4 | Single tree detection process based on YOLOv4-Lite.

TABLE 4 | Single tree detection evaluation indicators.

Evaluation indicators	Formula
Recall	$N_{mat} = \frac{N_{match}}{N_r}$
Commission rate (1-precision)	$N_{om} = \frac{N_{error}}{N_a}$
Omission rate	$N_{com} = \frac{N_{leave}}{N_r}$
Accuracy	$M = \frac{N_{match}}{N_{match}+N_{leave}+N_{error}} \times 100$

3.3 Process

Figure 4 illustrates the overall process of single-tree detection. The acquisition of high-resolution remote sensing image data is divided into two stages. One part contains training data. Manual labeling of training data is required. The second part contains the test data, which are displayed in the article. The YOLOv4-Lite model extracted single-tree features from training data repeatedly. On the test data, the trained YOLOv4-Lite model performs single-tree detection.

3.4 Evaluation Criteria

A professional evaluation standard is critical when evaluating the effect of a variety of single-tree detection methods. There is no unified evaluation standard in place at the moment. Correct detection occurs when the difference between the ground reference single tree and the detected single tree is within a specified range. The location of a single tree defines the actual box's geometric center. The position of each detected single tree is represented by M_i. The position of each ground reference single tree is represented by E_j. There are three possibilities for the results of single-tree detection: correct detection, error detection, and omission. Setting threshold $\varepsilon > 0$, $d\ (M_i, E_j)$ is denoted as the distance between two points M_i and E_j. The experiment results are as follows:

1 When $d\ (M_i, E_j) < \varepsilon$, it is considered that the detected single tree matches the ground reference single tree. It is the correct detection.
2 If $d\ (M_i, E_j) > \varepsilon$ for any M_i, there is no ground reference single-tree matching with the detected single tree. The detected single tree is considered as error detection.
3 E_j is neither in case (1) nor in case (2), in which case E_j is omission.

Based on the above conditions, N_r is the number of reference single trees, N_a is the number of detected single trees, and N_{match} is the correct number of correctly detected single trees in the detected single trees. The calculation formula of all values is shown in **Table 4**, N_{leave} is the number of undetected reference single trees and is also the difference between N_r and N_{match}, N_{error} is the difference between N_{match} and N_a, the recall rate is represented by the symbol N_{mat}, the commission rate is represented by the symbol N_{om}, the omission rate is represented by the symbol N_{com}, and the accuracy is represented by symbol M.

4 RESULTS AND DISCUSSION

To assess the method's effectiveness, two classical traditional methods, namely, the local maximum method and the watershed method, are compared to two novel methods, namely, the CV model and template matching method (Peng et al., 2017; Dong and Zhou, 2018; Zhang et al., 2020). This study analyzes and interprets the experimental results.

We use 416 × 416 images as the input image in this experiment. Because the experiment area is too large, it is divided into numerous small images for detection. Partitioning divides a tree in half or even a quarter.

4.1 Campus and Orchard Detection

Figures 5, **6** illustrate the results of various methods used to test the green trees surrounding the basketball court and the litchi trees near the litchi orchard. Red circles and yellow dots denote the trees that have been detected.

As illustrated in **Figure 5**, the campus experiment site resembles a plantation scene. The sliding window size is set to 23 for the local maximum method. The pixel difference threshold between the template and the sample is set to between 0.28 and 0.35 for the template matching method. The watershed method's marker parameter is set to 24. The treetop probe window and the background label parameter of the CV model are both set to 23. Because the local maximum value extracts only the maximum

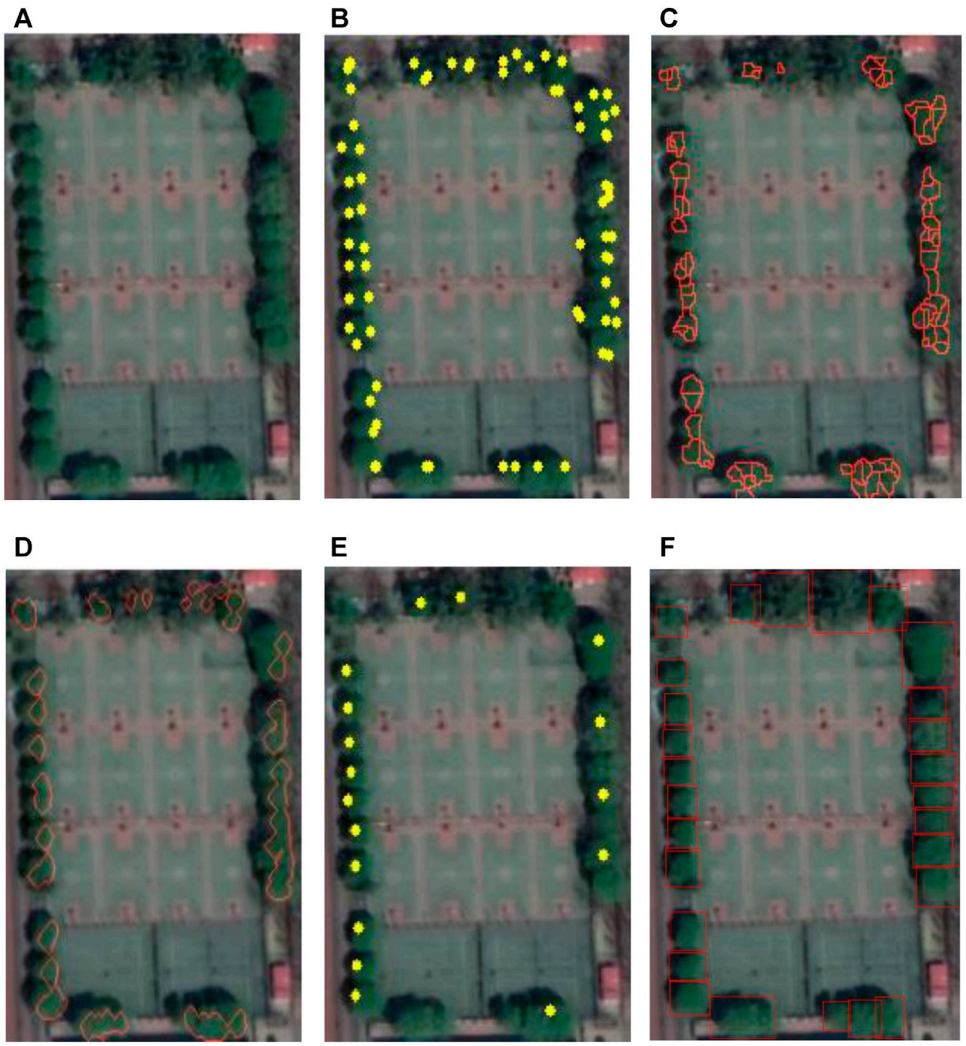

FIGURE 5 | Results of different methods in campus: **(A)** original image, **(B)** local maximum, **(C)** watershed, **(D)** CV model, **(E)** template matching, and **(F)** YOLOv4-Lite.

value, the watershed method extracts the contour of the puddle while easily being influenced by the tree branches. Several trees on the campus site have two large branches that reach the tree's top. As illustrated in **Figure 5B** and **Table 5**, both the local maximum and watershed methods incorrectly identify the tree branch as a tree, resulting in significant error detection. To minimize error detection, the CV model combines the advantages of the watershed and local maximum methods. On the other hand, the CV model considers two trees with a close crown connection to be a single tree, resulting in numerous detection omissions. Japonica *Sophora* Linn and *Salix* are the campus experimental site's dominant tree species. The two trees have similar characteristics, as do the tree canopy characteristics of the same trees. Additionally, the template matching method cannot extract all tree templates for trees with close crown connections.

Although the template matching method has a zero-commission rate, its accuracy, recall rate, and omission rate are lower than those of the YOLOv4-Lite method.

The YOLOv4-Lite method has a single tree detection accuracy of 96.3%. In conclusion, when compared to other algorithms, the proposed YOLOv4-Lite algorithm achieves the highest detection success rate and recall rate. Due to the small tree population on the campus experimental site, the experimental effect of YOLOv4-Lite is significantly greater than that of other sample sites.

As illustrated in **Figure 6**, the litchi garden's experimental site resembles an urban orchard. The sliding window size is set to 18 for the local maximum method. The template matching method sets the pixel threshold between the template and the sample to be between 0.32 and 0.38. The watershed method's marker parameter is set to 25. The CV method's treetop probe window is set to 20 and the background label parameter to 24.

As illustrated in **Figure 6F**, the single-tree detection method proposed by YOLOv4-Lite appears to have a high degree of error detection. Indeed, this phenomenon demonstrates Yolov4-Lite's

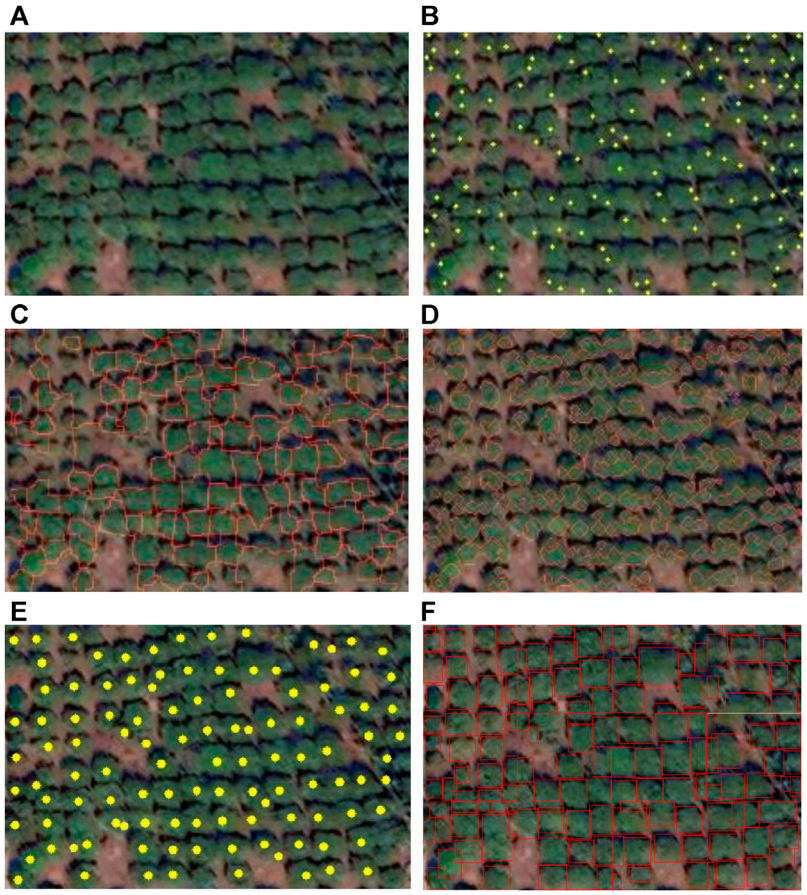

FIGURE 6 | Results of different methods in litchi garden: **(A)** original image, **(B)** local maximum, **(C)** watershed, **(D)** CV model, **(E)** template matching, and **(F)** YOLOv4-Lite.

TABLE 5 | Campus single-tree detection results' statistics.

Image number	Method	N_r	N_a	N_{match}	N_{leave}	N_{error}	M, %	N_{mat}, %	N_{com}, %	N_{om}, %
Campus	YOLOv4-Lite	26	27	26	0	1	96.3	100	0.0	3.7
	Watershed	26	73	25	1	48	33.8	96.2	3.8	65.8
	Local maximum	26	71	26	0	45	36.6	100	0.0	63.4
	Template matching	26	17	17	9	0	65.4	65.4	34.6	0.0
	CV model	26	21	17	9	4	56.7	65.4	34.6	19.0

extremely low omission rate and ability to detect up to a quarter of a tree. The YOLOv4-Lite single-tree detection method successfully detects each half-crown or quarter crown in the small image. Finally, the detected images were spliced together. The tree crown shape is irregular in the experimental site of the litchi garden. The crown difference is obvious, the connection is relatively close, and some background areas have higher values. As illustrated in **Table 6**, the detection efficacy of several methods is quite low. The local maximum method extracts only the tree with the greatest value and performs numerous error and omission detections. The experimental site of litchi garden contains a large number of saplings, and the contours of these saplings were not detectable using the watershed method. Additionally, the CV model incorporates extensive error and omission detection. Not all templates for single trees are available in advance. Because the template matching algorithm can only match regions of the same size, a small number of trees are labeled as missing errors, and numerous connected tree crowns are misidentified as a single tree. The YOLOv4-Lite method was found to have the best anti-interference ability in this study. It is capable of distinguishing a limited number of difficult-to-distinguish samples while maintaining a 93.8% accuracy, a 95.5% recall rate, and a 1.9% omission rate.

TABLE 6 | Litchi garden single-tree detection results' statistics.

Image number	Method	N_r	N_a	N_{match}	N_{leave}	N_{error}	M, %	N_{mat}, %	N_{com}, %	N_{om}, %
Litchi garden	YOLOv4-Lite	111	108	106	5	2	93.8	95.5	4.5	1.9
	Watershed	111	127	91	20	36	61.9	82.0	18.0	28.3
	Local maximum	111	204	76	35	128	31.8	68.5	31.5	62.7
	Template matching	111	116	104	7	12	84.6	93.7	6.3	10.3
	CV model	111	149	90	21	59	52.9	81.1	18.9	39.6

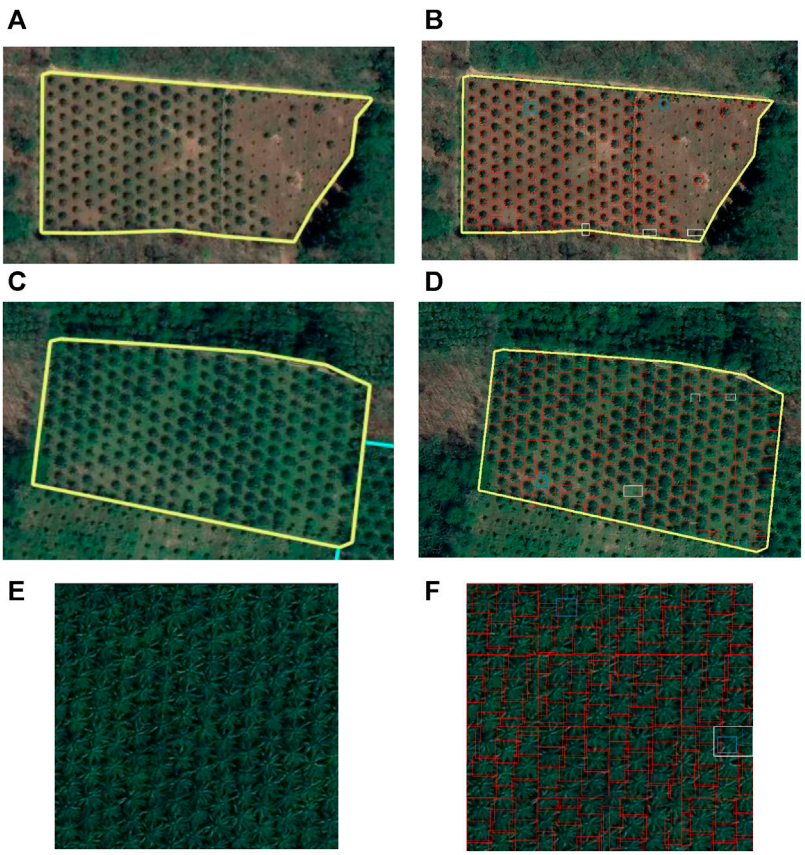

FIGURE 7 | Results of different methods in palm: **(A)** original image of sample site 1, **(B)** YOLOv4-Lite of sample site 1, **(C)** original image of sample site 2, **(D)** YOLOv4-Lite of sample site 2, **(E)** original image of sample site 3, and **(F)** YOLOv4-Lite of sample site 3.

4.2 Economic Plantation Detection

Palm trees are economic plantations, and many countries have a sizable artificial planting base. Palm tree monitoring and management are critical for plantation production. A stand of the open plantation has a canopy density of between 0.4 and 0.6, while a dense plantation has a canopy density of between 0.7 and 1. As illustrated in **Figure 7**, this study examines three distinct scenarios involving a single palm tree. The three scenes are sparse plantation with a simple background, sparse plantation with a complex background, and dense plantation. Red circles indicate correctly detected single trees. The trees that have been identified as having an error are circled in white. Blue circles indicate undetected single trees.

Sample site one depicts an open plantation with a simple background. The sliding window size is set to 23 for the local maximum method. The template matching method sets the pixel threshold between the template and the sample to between 0.31 and 0.35. The watershed method's marker parameter is set to 22. The CV model's treetop probe window is set to 23 and the background label parameter to 22. The background value of sample site one is high, which results in a high number of detected errors using the local maximum method and a high number of missed single trees using the watershed method. Many saplings are present at experimental site 1. Not all templates for single trees are available. The template matching method has a low recall rate and accuracy. As shown in **Table 7**, the

TABLE 7 | Sample site single-tree detection results' statistics.

Image number	Method	N_r	N_a	N_{match}	N_{leave}	N_{error}	M, %	N_{mat}, %	N_{com}, %	N_{om}, %
Sample site 1	YOLOv4-Lite	164	165	162	2	3	97.0	98.8	1.2	1.8
	Watershed	164	117	108	56	9	62.4	65.9	34.1	7.7
	Local maximum	164	213	132	32	81	53.9	80.5	19.5	38.4
	Template matching	164	157	155	9	2	93.4	94.5	5.5	1.32
	CV model	164	143	119	45	24	63.3	72.6	27.4	16.8
Sample site 2	YOLOv4-Lite	224	226	222	2	4	97.4	99.1	0.9	1.8
	Watershed	224	123	118	106	5	51.5	52.7	47.3	4.1
	Local maximum	224	162	93	131	69	31.4	41.5	58.5	42.6
	Template matching	224	217	203	21	14	85.3	90.6	9.4	6.5
	CV model	224	202	111	113	91	35.2	49.6	50.4	45.0
Sample site 3	YOLOv4-Lite	148	147	146	2	1	98.0	98.6	1.4	0.7
	Watershed	148	183	136	0	47	74.3	91.8	0.0	31.7
	Local maximum	148	162	124	24	38	66.7	83.8	16.2	23.5
	Template matching	148	152	147	1	5	96.1	99.3	0.7	3.3
	CV model	148	173	130	18	43	68.1	87.9	12.1	24.9

YOLOv4-Lite method has a higher accuracy and recall rate than the template matching method.

The background has the greatest influence on the experimental results when it comes to single-tree detection. The scene in which the background color matches the color of the crown of a single tree has the most interference. Sample site two is an open plantation with a complex background plantation. **Table 7** summarizes the experimental results for sample site 2. A single tree's canopy color in site two is very similar to the background color. The sliding window size is set to 23 for the local maximum method. The pixel difference threshold between the template and the sample is set to between 0.22 and 0.26 for the template matching method. The watershed method's marker parameter is set to 25. The CV model's treetop probe window is set to 23 and the background label parameter to 22. Local maximum only extracts the maximum value. When the crown color of a single tree is similar to the background color, the background color is brighter. The background is frequently mistaken for the crown apex, resulting in many missed and error detection data. Watershed is also difficult to extract the correct seed points, which results in numerous missed and incorrect detections in this case. In sample site 2, there are single trees of varying sizes, and not all single-tree templates can be extracted. As a result, the template matching method's recall rate and accuracy are low. In the scenario of sample site 2, the YOLOv4-Lite method maintains a 97.4% accuracy and a 99.1% recall rate.

Sample site three represents a dense plantation. The sliding window size is set to 35 for the local maximum method. The template matching method sets the pixel threshold between the template and the sample to between 0.16 and 0.22. The watershed method's marker parameter is set to 34. The CV model's treetop probe window has been set to 34, and the background label parameter has been set to 33. Dense plantations are typically mature plantations with large trees. When one tree crown is near another, the background cannot interfere with the single-tree detection effect. Because sample site three is the same age plantation as sample site 1, there are no saplings. As illustrated in **Table 7**, sample site 3 has the best effect for single-tree detection. If a tree has a large number of branches, multiple maximums will occur, and the local maximum method will result in a high rate of missed and error detection. Additionally, the watershed method generates a large amount of error detection. Additionally, the morphological characteristics of palm tree crowns are evident, and the template matching method detects sample site 3 with a 96.1% accuracy. The template matching method, on the other hand, is ineffective at detecting half and quarter crowns. This results in some missing tests. The accuracy of the YOLOv4-Lite method achieves 98.0%.

5 CONCLUSION

A YOLOv4-Lite method for single-tree detection is proposed in this study. Although it is not revolutionary in terms of object detection, it offers an excellent balance of speed and accuracy. Not only does the YOLOv4-Lite method overcome the four difficulties discussed in Chapter 1 but it also has a stronger anti-interference capability. Moreover, it performs well in various plantation types, including campus, orchards, and economic plantation, with extremely stable single-tree detection and location performance. The research presented in this article serves as a reference for the YOLOv4 model's application in its field.

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

YZ contributed to conceptualization, methodology, software, formal analysis, investigation, resources, data curation, writing—original draft, writing—review and editing, and visualization. GW contributed to validation and funding acquisition.

REFERENCES

Artés-hernández, F., Formica-Oliveira, A. C., Artés, F., and Martínez-Hernández, G. B. (2017). Improved Quality of a Vitamin B12-Fortified 'ready to Blend' Fresh-Cut Mix Salad with Chitosan. *Food Sci. Technol. Int.* 23, 513–528. doi:10.1177/1082013217705036

Bouvier, M., Durrieu, S., Fournier, R. A., and Renaud, J.-P. (2015). Generalizing Predictive Models of forest Inventory Attributes Using an Area-Based Approach with Airborne LiDAR Data. *Remote Sensing Environ.* 156, 322–334. doi:10.1016/j.rse.2014.10.004

Chaitanya, Sarath, S., Malavika, Prasanna, and Karthik. (2020). "Human Emotions Recognition from Thermal Images Using Yolo Algorithm," in *2020 International Conference on Communication and Signal Processing (ICCSP)*, 28-30 July 2020Chennai, India (IEEE). doi:10.1109/iccsp48568.2020.9182148

Chollet, F. (2017). "Xception: Deep Learning with Depthwise Separable Convolutions," in 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). doi:10.1109/cvpr.2017.195

Dimitrios, P., and Azadeh, A. (2021). Accuracy Assessment of Total Stem Volume Using Close-Range Sensing: Advances in Precision Forestry. *Forests* 12 (6).

Dimitrios, P., Azadeh, A., and Martin, S. (2021). Assessment of Stem Volume on Plots Using Terrestrial Laser Scanner: A Precision Forestry Application. *Sensors* 21 (1).

Dong, T., Shen, Y., Zhang, J., Ye, Y., and Fan, J. (2019). Progressive Cascaded Convolutional Neural Networks for Single Tree Detection with Google Earth Imagery. *Remote Sensing* 11 (15), 1786. doi:10.3390/rs11151786

Dong, T. Y., and Zhou, Q. Z. (2018). Single Tree Detection in Remote Sensing Images Based on Morphological Snake Model. *ISPRS, Int. J. Geo-Inf.* Computer Science.

Dong, Y., Ren, Z., Fu, Y., Miao, Z., Yang, R., Sun, Y., et al. (2020). Recording Urban Land Dynamic and its Effects during 2000-2019 at 15-m Resolution by Cloud Computing with Landsat Series. *REMOTE SENSING* 12 (15). doi:10.3390/rs12152451

Garcia-ortiz, L. B., Sanchez-Perez, G., Hernandez-Suarez, A., Olivares-Mercado, J., and Portillo-Portillo, J. (2020). *A Fast-RCNN Implementation for Human Silhouette Detection in Video Sequences, Knowledge Innovation through Intelligent Software Methodologies, Tools and Techniques*. doi:10.3233/FAIA200553

Gomes, M. F., Maillard, P., and Deng, H. (2018). Individual Tree crown Detection in Sub-meter Satellite Imagery Using Marked Point Processes and a Geometrical-Optical Model. *Remote Sensing Environ.*, 211184–211195. doi:10.1016/j.rse.2018.04.002

Gougeon, F. A., and FranOis, A. (2014). A crown-following Approach to the Automatic Delineation of Individual Tree Crowns in High Spatial Resolution Aerial Images. *Can. J. Remote Sensing* 21 (3), 274–284. doi:10.1080/07038992.1995.10874622

Hashim, S. A., Daliman, S., Rodi, I., Aziz, N. A., and Rak, A. E. (2020). Analysis of Oil Palm Tree Recognition Using Drone-Based Remote Sensing Images. *IOP Conf. Ser. Earth Environ. Sci.* 596012070. doi:10.1088/1755-1315/596/1/012070

Hofmann, H. (2000). Exploring Categorical Data: Interactive Mosaic Plots. *Metrika* 51 (1), 11–26. doi:10.1007/s001840000041

Huiling, L., Xiaoli, Z., Ying, Z., Yunfeng, Z., Hui, L., and Longyang, W. (2018). Review on Individual Tree Detection Based on Airborne LiDAR. *[J]. Prog. Laser Optoelectronics* 55 (8), 082805. doi:10.3788/lop55.082805

Jiang, R., Wang, C., Shen, L., Wang, P., Koirala, A., Walsh, K., et al. (2016). "A Method for Lichee's Tree-crown Information Extraction Based on High Spatial Resolution Image," in *Transactions of the Chinese Society for Agricultural MachineryDeep Learning for Real-Time Fruit Detection and Orchard Fruit Load Estimation: Benchmarking of 'MangoYOLO*, 20, 1107–1135. doi:10.6041/j.issn.1000-1298.2016.09.003

Larsen, M., Eriksson, M., Descombes, X., Perrin, G., Brandtberg, T., and Gougeon, F. A. (2011). Comparison of Six Individual Tree crown Detection Algorithms Evaluated under Varying forest Conditions. *Int. J. Remote Sensing* 32 (20), 5827–5852. doi:10.1080/01431161.2010.507790

Liang, Y., Liao, J. C., and Pan, J. H. (2020). Mesh-Based Scale-Invariant Feature Transform-like Method for Three-Dimensional Face Recognition under Expressions and Missing Data. *J. Electron. Imaging* 29 (5), 053008.

Liu, H., Zhang, X., and Zhang, Y. (2018). Advances in Single Wood Identification of Airborne Lidar. *J. Laser Optoelectron. Prog.* 55 (8), 39–47. doi:10.3788/LOP55.082805

Ma, K., Cheng, Y., Ge, W., Zhao, Y., and Qi, Z. (2020). *Identification, Extraction and Three-Dimensional Building Model Reconstruction Though Faster R-CNN of Architectural Plans*. Geoinformatics in Sustainable Ecosystem and Society, Springer.

Meneghetti, D., Homem, T., Oliveira, J., Silva, I., and Bianchi, R. (2021). Detecting Soccer Balls with Reduced Neural Networks. *J. Intell. Robotic Syst.* 101 (3). doi:10.1007/s10846-021-01336-y

Nasor, M., and Obaid, W. (2021). Segmentation of Osteosarcoma in MRI Images by K-Means Clustering, Chan-Vese Segmentation, and Iterative Gaussian Filtering. *IET Image Process.* 15.

Olli, N., Eija, H., Sakari, T., Niko, V., Teemu, H., Yu, X., et al. (2017). Individual Tree Detection and Classification with UAV-Based Photogrammetric Point Clouds and Hyperspectral Imaging. *Remote Sensing*.

Peng, L., Hui, Z., and Eom, K. B. (2017). Active Deep Learning for Classification of Hyperspectral Images. *IEEE J. Selected Top. Appl. Earth Observations Remote Sensing* 10 (2), 712–724. doi:10.1109/jstars.2016.2621416

Picos, J., Bastos, G., Míguez, D., Martínez, L. A., and Armesto, J. (2020). Individual Tree Detection in a Eucalyptus Plantation Using Unmanned Aerial Vehicle (UAV)-LiDAR. *Remote Sensing* 12 (5). doi:10.3390/rs12050885

Rau, A., Kim, S., Yang, S., Reisert, M., and Egger, K. (2021). *SVM-based Normal Pressure Hydrocephalus Detection*. Heidelberg, Germany: Springer. Clinical Neuroradiology. No. 3 17, D-69121.

Richey, B., and Shirvaikar, M. V. (2021). "Deep Learning Based Real-Time Detection of Northern Corn Leaf Blight Crop Disease Using YoloV4," in *Real-Time Image Processing and Deep Learning 2021*. doi:10.1117/12.2587892

Ruihuan hou, X. Y. Z. W. (2021). A Real-Time Detection Methods for forest Pests Based on YOLO V4-TIA Algorithm. *Computer Eng.*, 1–8.

Syaputra, R., Syamsuar, D., and Negara, E. S. (2021). Multiple Smile Detection Using Histogram of Oriented Gradient and Support Vector Machine Methods. *IOP Conf. Ser. Mater. Sci. Eng.* 1071 (1), 012027. doi:10.1088/1757-899x/1071/1/012027

Toth, C., and Jóźków, G. (2016). Remote Sensing Platforms and Sensors: A Survey. *ISPRS J. Photogrammetry Remote Sensing* 115 (may), 22–36. doi:10.1016/j.isprsjprs.2015.10.004

Wang, L., Gong, P., and Biging, G. S. (2004). Individual Tree-Crown Delineation and Treetop Detection in High-Spatial-Resolution Aerial Imagery. *Photogramm Eng. Remote Sensing* 70 (3), 351–357. doi:10.14358/pers.70.3.351

Wang, M. (2019). *Multi-level Assessment of Mangrove Ecosystem Based on Remote Sensing Technology*. Xiamen, China: Xiamen University.

Wang, X., Wang, S., Cao, J., and Wang, Y. (2020). *Data-driven Based Tiny-YOLOv3 Method for Front Vehicle Detection Inducing SPP-Net*, 99. IEEE Access, 1.

Wulder, M., Niemann, K. O., and Goodenough, D. G. (2000). Local Maximum Filtering for the Extraction of Tree Locations and Basal Area from High Spatial Resolution Imagery. *Remote Sensing Environ.* 73 (1), 103–114. doi:10.1016/s0034-4257(00)00101-2

Xiao, Y., Wang, D., and Wu, Y. (2021). Investigation and Protection of Ancient and Famous Trees Resources of Daxiong Mountain. *IOP Conf. Ser. Earth Environ. Sci.* 692 (37pp), 032009. doi:10.1088/1755-1315/692/3/032009

Yinghai, K., Xiaojuan, L., and Huili, G. (2015). *Application of Remote Sensing Technology in Automatic forest Inventory*. Beijing, China: China Environmental Press.

Yuz, S. Y. S. C. (2021). A Real-Time Detection Approach for Bridge Cracks Based on YOLOv4-FPM. *Automation in Construction* 103514 (122). doi:10.1016/j.autcon.2020.103514

Zhang, L., Liu, P., Zhao, L., Wang, G., and Liu, J. (2020). Air Quality Predictions with a Semi-supervised Bidirectional LSTM Neural Network. *Atmos. Pollut. Res.* 12 (1).

Zhang, N., Zhang, X., and Ye, L. (2014). Tree crown Extraction Based on Segmentation of High-Resolution Remote Sensing Image Improved Peak-Climbing Algorithm. *Trans. Chin. Soc. Agric. Machinery* 45 (12), 294–300.

Zhang, S., Hu, X., Jie, X., Li, S., Wang, Z., Zhao, Z., et al. (2019). *Research on Image Detection Method for Assembly Failure of Monomer thermal Battery*. Beijing, China: Journal of Electronic Measurement and Instrumentation.

Zhao, L., Zeng, Y., Liu, P., and Su, X. (2020). Band Selection with the Explanatory Gradient Saliency Maps of Convolutional Neural Networks. *IEEE Geosci. Remote Sensing Lett.* 99, 1–5. doi:10.1109/lgrs.2020.3012140

Zheng, X., Wang, R., Jin, M., and Cas (2017). *Extraction of High-Resolution Images of Single Tree Crown Based on Watershed Algorithm with Morphological Threshold Mark*. Central South Forest Inventory and Planning.

Zhibin, R., Hongbo, Z., Yao, F., Lu, X., and Yulin, D. (2020). *Effects of Urban Street Trees on Human thermal comfort and Physiological Indices: A Case Study in Changchun City, China*. Harbin, China: JOURNAL OF FORESTRY 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.

frontiers in Environmental Science

ORIGINAL RESEARCH
published: 21 February 2022
doi: 10.3389/fenvs.2022.798336

An Optimality-Based Spatial Explicit Ecohydrological Model at Watershed Scale: Model Description and Test in a Semiarid Grassland Ecosystem

Lajiao Chen[1,2,3], *Zhidan Hu*[4], *Xiaoping Du*[5,6]*, *Mohd Yawar Ali Khan*[7], *Xiaojun Li*[8] and *Jie Wen*[9]

[1]Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, China, [2]University of Chinese Academy of Sciences, Beijing, China, [3]State Key Laboratory of Hydroscience and Engineering, Tsinghua University, Beijing, China, [4]Information Center (Hydrology Monitor and Forecast Center), Ministry of Water Resources, Beijing, China, [5]Key Lab of Digital Earth Sciences, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, China, [6]International Research Center of Big Data for Sustainable Development Goals, Beijing, China, [7]Department of Hydrogeology, Faculty of Earth Sciences, King Abdulaziz University, Jeddah, Saudi Arabia, [8]China Water Northeastern Investigation, Design and Research Co., Ltd, Beijing, China, [9]China Institute of Water Resources and Hydropower Research, Beijing, China

OPEN ACCESS

Edited by:
Yizhe Zhan,
MetService, New Zealand

Reviewed by:
Shi Hu,
Institute of Geographic Sciences and Natural Resources Research (CAS), China
Hongkai Gao,
East China Normal University, China

***Correspondence:**
Xiaoping Du
duxp@aircas.ac.cn

Specialty section:
This article was submitted to Environmental Informatics and Remote Sensing, a section of the journal Frontiers in Environmental Science

Received: 20 October 2021
Accepted: 05 January 2022
Published: 21 February 2022

Citation:
Chen L, Hu Z, Du X, Khan MYA, Li X and Wen J (2022) An Optimality-Based Spatial Explicit Ecohydrological Model at Watershed Scale: Model Description and Test in a Semiarid Grassland Ecosystem. Front. Environ. Sci. 10:798336. doi: 10.3389/fenvs.2022.798336

Optimality principles have been applied in ecohydrological modeling to derive optimal vegetation properties and describe co-evolution states of vegetation and water cycle. Unfortunately, most existing optimality-based models only consider vertical vegetation-soil-water interactions on plot scale, without considering the lateral hydrological processes. This work aims to extend the field-scale Vegetation Optimality Model (VOM) to the watershed scale. Lateral flow is incorporated to VOM through a hierarchical strategy, establishing the Distributed Vegetation Optimality Model (DisVOM). The model is tested with long-tem flux measurements in the Walnut Gulch watershed, a United States Agricultural Research Service (US-ARS) experimental watershed in southern Arizona. The results indicate the model performance is acceptable for most of years, especially for the growing season. The seasonal dynamic of ET, soil water, and GPP demonstrate good consistency with observations. The model provides reasonable spatial distribution of ET and GPP, suggesting the model can discriminate the effect of lateral flow on water redistribution, and consequently on root water uptake, as well as carbon assimilation. The model could be a useful tool assessing the impact of climate change and human activities on vegetation and water cycle.

Keywords: ecohydrology, ecohydrological modeling, vegetation optimality, lateral hydrological processes, spatial explicit

1 INTRODUCTION

Vegetation and the water cycle are intrinsically coupled through physical and biological processes, such as carbon, water, energy, and nutrients exchange (Rodriguez-Iturbe et al., 1999; Biederman et al., 2017; Wang et al., 2019; Rice and Emanuel, 2019; Xia et al., 2021). Over the past few decades, there has been a growing awareness that vegetation co-evolves with the environment achieving an equilibrium status and optimal use of resources (such as light, water) when adapting to the environment (Ball et al., 1987; Cowan, 2002; Berninger et al., 1996; Eagleson, 2002; Rodriguez-Iturbe and Porporato, 2005; Chen et al., 2019; Franklin et al., 2020).

Based on Darwin's theory of natural evolution, Eagleson (1978, 2002) proposed the optimality hypothesis that in water-limited ecosystem vegetation reaches a "growth equilibrium" density when stress is minimized in short term, and an "evolutionary equilibrium" at which the rate of production is maximized in long terms (Eagleson, 1978; Eagleson, 2002). Following Eagleson's pioneering work, different optimality principles have been established (Rodriguez-Iturbe and Porporato, 2005). Those optimality principles are applied to derive the optimal vegetation properties, such as optimal vegetation distribution (Caylor et al., 2004), optimal root properties (Schymanski et al., 2008; Gao et al., 2014; Carbon et al., 2018; Speich et al., 2018), optimal photosynthetic canopy properties (van der Tol et al., 2008; de Boer et al., 2011).

In recent years, optimality principles have been applied in ecohydrological modeling to derive optimal vegetation properties and describe co-evolution states of vegetation and water cycle (Van der Tol et al., 2008; Schymanski, et al., 2009; Caylor et al., 2009; Pauwels et al., 2007; Lei et al., 2009; Huang et al., 2020). In contrast to the traditional ecohydrological models (Chen et al., 2015; Fatichi et al., 2016), which heavily rely on calibration (Kuppel et al., 2018), optimality-based ecohydrological models do not need prior knowledge of vegetation, which make them powerful to predict the system's response to new conditions (Sutherland, 2005; Sivapalan, 2009). To our knowledge, most existing optimality-based models only consider vertical vegetation-soil-water interactions on plot scale, without considering the lateral hydrological processes (Hwang et al., 2009).

It is well acknowledged that lateral hydrological processes and redistribution of soil water contribute to the complex vegetation structure and patterns at watershed scale (Band et al., 1993; Govind et al., 2009; Chen et al., 2015). It is especially significant in mountainous and hilly terrain due to the complex variability of topography (Ivanov et al., 2008; Gao et al., 2019). Therefore, the lack of lateral flow in ecohydrological models will lead to inappropriate representation of soil water, which directly affects the description of ecohydrological processes, such as root water uptake, evapotranspiration.

Vegetation Optimality Model (VOM) (Schymanski et al., 2009) is an optimality-based model, which applies a principle of maximization of Net Carbon Profit (*NCP*) to acquire optimal vegetation properties. One key merit of this model is that it does not need prior knowledge about the vegetation to run the model. However, like the other optimality-based models, this model only considers vertical water and vegetation dynamics on field scale.

This work aims to extend the field-scale VOM to the watershed scale. We incorporate lateral hydrological processes to VOM, establishing the Distributed Vegetation Optimality Model (DisVOM). The model is tested in the Walnut Gulch watershed, a United States Agricultural Research Service (US-ARS) experimental watershed in southern Arizona. The model outputs are validated with the observed flux data. To evaluate the contribution of lateral flow to ecohydrological processes modeling, we compare the distribution of modeled evapotranspiration (ET) and Gross Primary Productivity (GPP) with the spatial pattern of topography. We further examine the problems associated with the model, which pointed out the orientation of model improvements in the future.

2 MATERIALS AND METHODS

2.1 Model Description

The model developed in this research is an optimality based, spatial explicit ecohydrological model at watershed scale. The watershed is delineated from the Digital Elevation Model (DEM) into grid cells, and each grid is considered as a unique vegetation-soil system. Vertically, each grid is subdivided into many layers and simulates ecohydrological processes such as photosynthesis, transpiration, soil evaporation, infiltration, etc. Horizontally, it describes the heterogeneities in topography, soil, vegetation and atmospheric forcing, and adequately describes the mutual interaction among grids utilizing grid-based routing algorithms.

2.2 Vertical Ecohydrological Processes

VOM (Schymanski et al., 2009) is implemented to describe vertical ecohydrological processes at each grid cell. VOM couples a multilayered physically based water balance model and an ecophysiological gas exchange model. The model represents vegetation as two "big leaves." One covering an invariant area fraction represented perennial vegetation (e.g., trees) and a varying area fraction representing seasonal vegetation (e.g., annual grasses). The model acquires optimal vegetation properties through the optimality principle that vegetation would maximise *NCP*. We list the most important equations as follow. For detail of the model, please see the reference of Schymanski et al (2009).

2.2.1 Photosynthesis

Leaf CO_2 assimilation is based on a biochemical model of photosynthesis (Schymanski et al., 2009).

$$A_g = \frac{1}{8}\left(4C_aG_s + 8\Gamma_*G_s + J_A - 4R_l\right) - \frac{1}{8}\sqrt{\left(-4C_aG_s + 8\Gamma_*G_s + J_A - 4R_l\right)^2 + 16G_s\Gamma_*\left(8C_aG_s + J_A + 8R_l\right)} \tag{1}$$

where C_a is the mole fraction of CO_2 in the air, G_s is stomatal conductivity, Γ_* is the CO_2 compensation point in the absence of mitochondrial respiration, and J_A is photosynthetic electron transport rate.

2.2.2 Stomatal Conductivity and Transpiration

Transpiration is treated as a diffusive processes controlled by stomatal conductivity, following the equation developed by Cowan et al. (1977):

$$E_t = aG_s\left(W_l - W_a\right) = aG_sD_v \tag{2}$$

where D_v is atmospheric vapor deficit, W_l and W_a denote the mole fraction of water vapor in air inside the leaf and in the atmosphere, which is approximate to D_v, and a is the molecular diffusion coefficient of CO_2 in the air, defined as 1.6.

Cowan and Farquhar (Cowan, et al., 1977) proposed the stomatal conductivity optimal hypothesis that leaves would maximize CO_2 uptake for any given amount of water in a period. It can be expressed as the slope of transpiration (E_t) and carbon assimilation (A_g), which can be maximized with a constant value of λ over the period.

$$\frac{\partial E_t / \partial G_s}{\partial A_g / \partial G_s} = \frac{\partial E_t}{\partial A_g} = \lambda \tag{3}$$

Combining **Eq. 1**, **Eq. 2**, and **Eq. 3**, vegetation transpiration can be calculated as

$$E_t = \frac{aD_v[C_a(J_A - 4R_l) - 4(J_A + 2R_l)\Gamma_*]}{4(C_a - 2\Gamma_*)^2} + \frac{\sqrt{3}\sqrt{aD_v J_A \Gamma_* (\lambda C_a - 2aD_v + 2\lambda\Gamma_*)^2 (\lambda C_a - aD_v + 2\lambda\Gamma_*)[C_a(J_A - 4R_l) - (J_A + 8R_l)\Gamma_*]}}{4(C_a + 2\Gamma_*)^2 (\lambda C_a - aD_v + 2\lambda\Gamma_*)} \tag{4}$$

where λ including λ_s and λ_p is a constant within 1 day, and is parameterized as a function of the average matric suction head of each soil layer (h_i) in the root zone.

2.2.3 Soil Evaporation

Soil evaporation includes evaporation from saturated zone and evaporation from unsaturated zone. Soil evaporation is determined by radiation, soil surface fraction, as well as soil water.

$$E_{su} = \frac{I_g(1 - 0.8(1 - M_A))\omega_u S_{u,1}}{\lambda_E \rho} \tag{5}$$

$$E_{ss} = \frac{I_g(1 - 0.8(1 - M_A))\omega_0}{\lambda_E \rho} \tag{6}$$

where I_g is global irradiance, M_A is fraction of area covered by vegetation, ω_u is unsaturated surface area fraction, $S_{u,1}$ is average saturation degree in the unsaturated zone, ω_0 is saturated surface area fraction.

2.2.4 Vegetation Optimality Principle

The vegetation optimality principle applied in VOM is maximization of *NCP*, which is defined as total CO_2 uptake of tree and grasses over the entire period, excluding all identified maintenance costs of the organs assisting photosynthesis, including foliage, roots, and water transport tissues:

$$NCP = \int_{t_{start}}^{t_{end}} \left(A_{g,tot}(t) - R_f(t) - R_r(t) - R_v(t)\right) dt \tag{7}$$

where $A_{g,tot}$ is the combined CO_2 uptake by trees and grasses, R_f is the foliage cost of grasses and trees combined, R_r is the root cost of grasses and trees combined, and R_v is the cost associated with the vascular systems of grasses and trees combined.

2.2.5 Vegetation Optimality Strategy

Vegetation optimality strategy consists of long-term optimization of vegetation properties adapted to environmental conditions and short-term optimization of vegetation properties adapted to daily changes of environment. Long-term vegetation properties include fraction of area covered by perennial vegetation ($M_{A,p}$), the thickness of root zone of perennial vegetation ($y_{r,p}$) and water use parameters of perennial and seasonal vegetation ($c_{\lambda f,p}$, $c_{\lambda e,p}$, $c_{\lambda f,s}$, $c_{\lambda e,s}$). Short-term vegetation properties include the fraction of area covered by seasonal vegetation ($M_{A,s}$), electron transport capacity of perennial and seasonal vegetation ($J_{\max 25,p}$, $J_{\max 25,s}$), root area depth distribution of perennial and seasonal vegetation ($S_{adr,i,p}$, $S_{adr,i,s}$). The Shuffle Complex Evolution (SCE) algorithm developed by Duan et al. (1994) is used to achieve the optimal vegetation parameters to maximize *NCP* over the entire period.

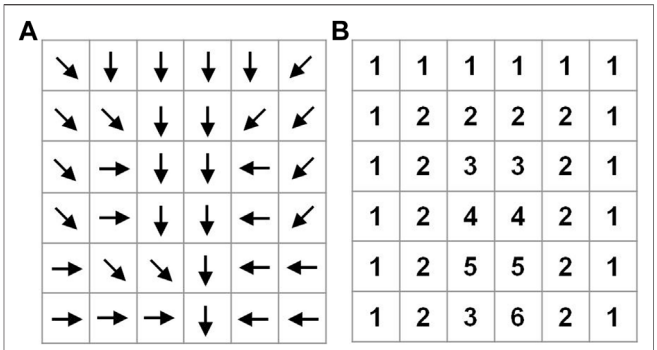

FIGURE 1 | Flow direction **(A)** and different level of priority **(B)**.

2.3 Explicit Flow Routing

This model calculates the overland flow with run-on infiltration determined by the excess water flow. It is assumed that all overland flow generated by a grid will rout to the downstream grid cell in a single time step. If the receiving grid is not saturated, the overland flow will infiltrate according to infiltration capacity, and the excess amount will be routed.

Overland flow can be described by dynamic wave equations, which is also known as St Venant equations. As St Venant equations are highly nonlinear and do not have analytical solutions, practical equations are derived, such as kinematic-wave and diffusion-wave. In this study, we used the 4-point implicit method to solve kinematic wave overland flow (Amein and Fang, 1969; Cevza et al., 2005). The governing equation for kinematic-wave is expressed as follow:

$$\frac{\partial y}{\partial t} + \alpha m y^{m-1} \frac{\partial y}{\partial x} = i - f \tag{8}$$

where y is the depth of water, t is distance, x is time, i is rainfall intensity, f is infiltration rate. α and m are coefficient, which can be obtained from Manning equations for fully turbulent flow:

$$\alpha = \frac{1}{n}\sqrt{S_0} \tag{9}$$

$$m = \frac{5}{3} \tag{10}$$

The implicit method gives the approximation of kinematic wave equation:

$$\frac{q_{i+1}^{j+1} - q_i^{j+1}}{\Delta x} + \frac{y_{i+1}^{j+1} - y_{i+1}^{j}}{\Delta t} = \left(i - f\right)_{i+1}^{j} \tag{11}$$

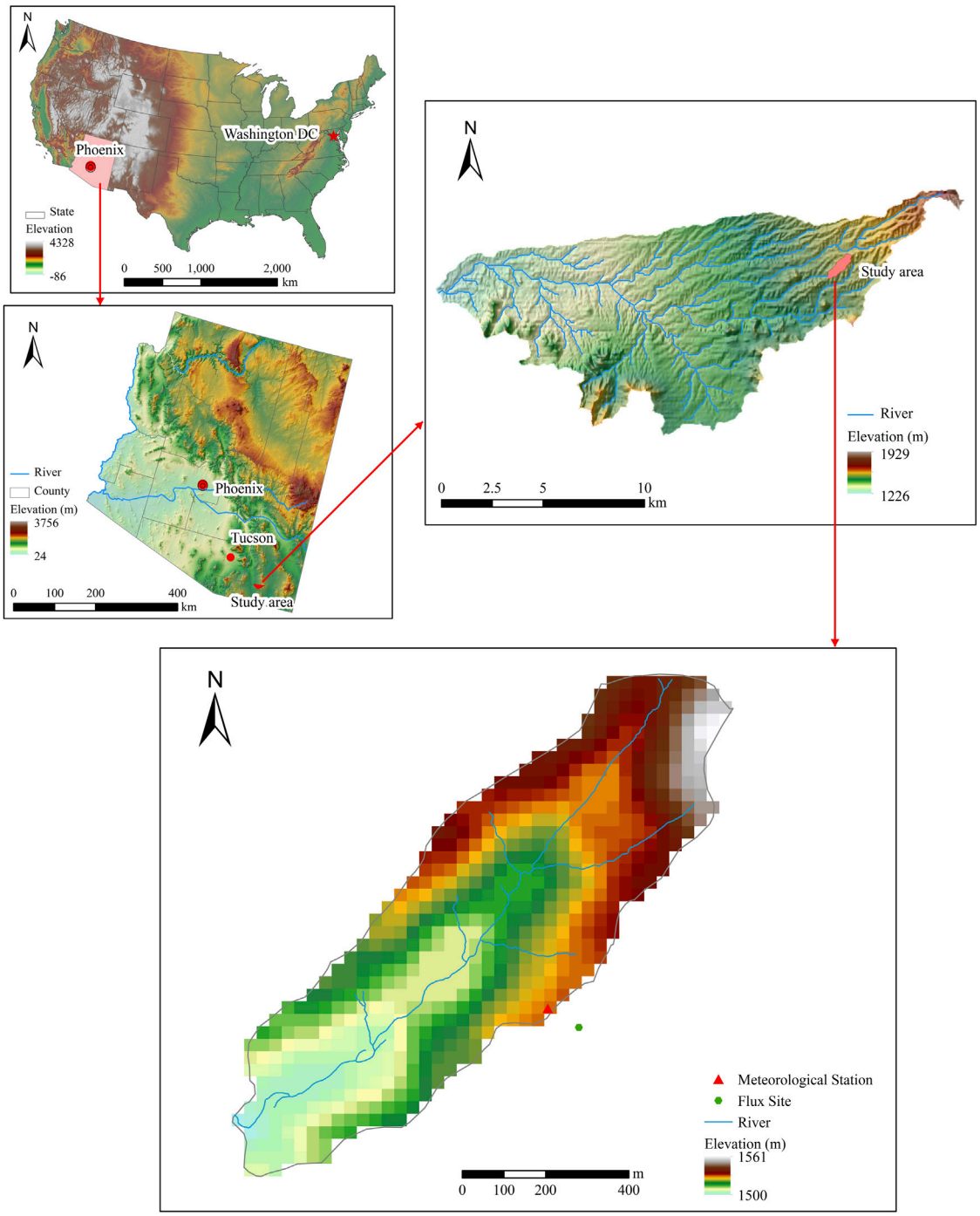

FIGURE 2 | Location of study area and observation station.

$$q_{i+1}^{j+1} = \alpha\left(y_{i+1}^{j+1}\right)^m \tag{12}$$

$$q_i^{j+1} = \alpha\left(y_i^{j+1}\right)^m \tag{13}$$

We can build a constructor equation as follow:

$$f\left(y_{i+1}^{j+1}\right) = \frac{\Delta t}{\Delta x}\alpha\left(y_{i+1}^{j+1}\right)^m + y_{i+1}^{j+1} - \left[\frac{\Delta t}{\Delta x}\alpha\left(y_i^{j+1}\right)^m + y_{i+1}^{j} + \Delta t\left(i - f\right)_{i+1}^{j+1}\right] \tag{14}$$

TABLE 1 | Data used in this study.

Data type		Data items	Time scale
Input data	Meteorological data	Solar radiation, temperature, precipitation, PAR, Soil temperature	1 h
	Topography	DEM	30 m
	Soil data	Soil type, soil physical properties including texture, saturated conductivity, etc	30 m
Validation data	Flux measurements	CO_2 flux, moisture flux	1 h
	Hydrological data	Soil water	1 h

Eq. 15 can be solved by Newton–Raphson iteration.

$$\left(y_{i+1}^{j+1}\right)_{k+1} = \left(y_{i+1}^{j+1}\right)_k - \frac{f\left(y_{i+1}^{j+1}\right)_k}{f'\left(y_{i+1}^{j+1}\right)_k} \tag{15}$$

where $f'(y_{i+1}^{j+1})_k$ is the derivative of $f(y_{i+1}^{j+1})$, which can be expressed as:

$$f'\left(y_{i+1}^{j+1}\right) = 1 + \alpha m \frac{\Delta t}{\Delta x}\left(y_{i+1}^{j+1}\right)^{m-1} \tag{16}$$

2.4 Hierarchical Incorporation Strategy

To account for the spatial variability of topography and consequently modify optimality principle in VOM, we incorporate lateral flow to VOM for each grid cell. According to the characteristic of grid-based flow routing solved by the 4-point implicit scheme as stated in **section 2.4**, for a given grid, the overland flow of a given grid at a time step is determined by two grids at two-time steps, including water depth of the upstream grid at current time step, water depth of upstream grid at last time step, water depth of the current grid at current time step, water depth of current grid at last time step. In such a circumstance, the downstream grid's water depth can only be calculated until the upstream grid has its outflow calculated. However, the upstream grid's outflow can only be obtained until the vegetation parameters optimization, which also depends on its upstream grid. This extremely complicates the incorporation of lateral flow to VOM.

To overcome the problems mentioned above, a hierarchical strategy is developed to integrate lateral flow to VOM. The hierarchical approach first prioritizes the grid cells of the watershed into different levels according to the flow dependency of grid cells. Different levels have different calculation priority. The highest level includes those grid cells that have no inflow from other grid cells. Vegetation optimization can be conducted by themselves as they do not need inputs from other grid cells. We define the grid cells of highest level as the first grid layer. The downstream grid cells of the first-layer grid cells are defined as the second grid layer. Calculations of those grid cells, including vegetation optimization and ecohydrogical processes modeling, rely on the outflow of the first-level grid cells. The downstream grid cells of the second-layer cells are defined as the third layer cells, and so on until all of the grid cells are ranked.

Prioritization of grid cells is conducted based on flow direction, which is obtained from DEM data. This study uses the D8 algorithm from ArcGIS to derive flow direction. **Figure 1A** shows a simple example of flow directions. Grid cells with no upstream grid are ranked as first level, labeled 1 in **Figure 1B**. The downstream grid cells of those grids are ranked as second level, which are labeled 2 in **Figure 1B**. By this means, all of the grid cells can be classified into different levels.

Following the prioritization of grid cells, ecohydrological modeling of different-level grid cells is implemented. The first level has the highest computation priority. Vegetation optimization is firstly conducted for the first-level grid cells and after vegetation parameters have been gained, ecohydrological processes of those grid cells are modeled. The outflow of those grid cells serves as the inflow of the downstream grid cells and accordingly participates in vegetation properties optimization of downstream grid cells. In this way, the lateral flow can be fully incorporated into vegetation optimality model.

2.5 Study Area

The model test and evaluation are conducted in USDA-ARS Walnut Gulch Experimental Watershed (WGEW) near Tombstone, Arizona (**Figure 2**). The watershed is located in the upper San Pedro River Basin, covering 7,600 km^2 in Sonora, Mexico and Arizona. The area of the WGEW is about 145 km^2. The average annual temperature is about 17–19°C, and the average annual precipitation is 322 mm, with 67% rainfall in summer (Goodrich, et al., 2008). The primary vegetation of the watershed is grass and shrubs and is a transition zone between the Chihuahuan and Sonoran Deserts.

We use a small watershed near the Kendall grass site (109°56′8″W, 31°44′10″N; elevation: 1,526 m) as the validation watershed (**Figure 2**). The watershed is small, with nearly unique climate, vegetation, and soil type, therefore, the spatial variability of modeled ET and GPP can be attributed to lateral flow.

2.6 Data

Data used in this study are obtained from the Southwest Watershed Research Center (SWRC) of USDA through the official governmental website (http://www.tucson.ars.ag.gov/dap/) (Emmerich and Verdugo, 2008). These data include spatial data, model driving data, and model validation data (**Table 1**). Meteorological measurements include precipitation (except for 2002), air temperature, air humidity, Photosynthetically Active Radiation (PAR), solar radiation are gained from flux data near the Kendall grass site. Flux data are acquired from Bowen Ratio Energy Balance System (BREB, Model 023/CO_2, Campbell Scientific Inc., Logan, UT). In the

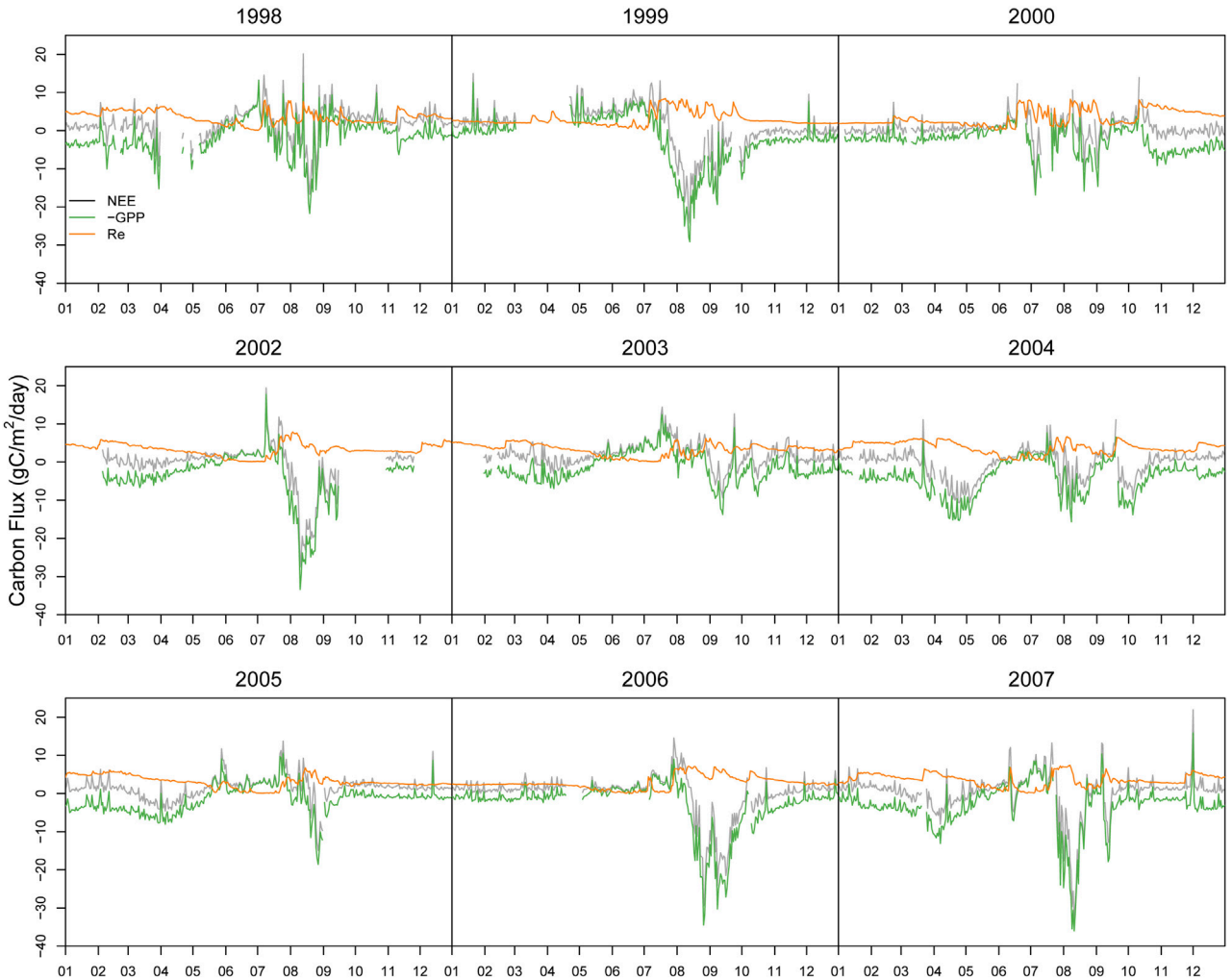

FIGURE 3 | Subdivision of measured NEE into GPP and ecosystem respiration.

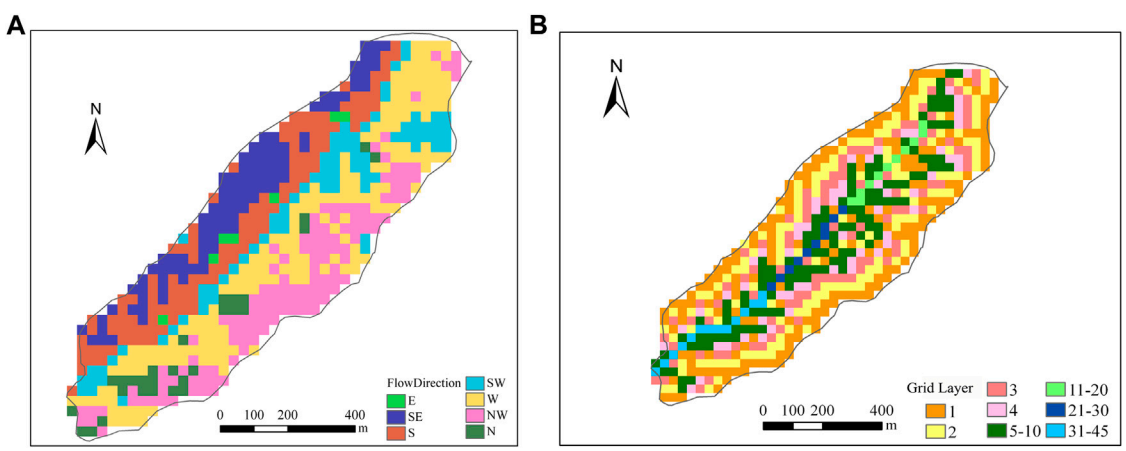

FIGURE 4 | Flow direction **(A)** and prioritization of grid cells **(B)** in Kendall watershed.

TABLE 2 | Soil and vegetation parameters.

Parameters	Value	Source
Soil depth (m)	1.5	USDA (2008)
Soil layer depth (m)	0.25	This study
Residual soil water θ_r (m^{-3} m^{-3})	0.065	Celia et al. (1990)
Saturated soil water θ_s (m^{-3} m^{-3})	0.30	This study
Van Genuchten parameter α (m^{-1})	7.5	Celia et al. (1990)
Van Genuchten parameter n (-)	1.89	Celia et al. (1990)
Saturated hydraulic conductivity K_{sat} (mm s^{-1})	1.228×10^{-2}	Celia et al. (1990)
Rate of exponential increase of Jmax with temperature (J mol^{-1})	88,900	Massad et al. (2007)
Rate of exponential decrease of Jmax with temperature above optimal temperature (J mol^{-1})	22,000	Massad et al. (2007)
Optimal temperature for electron transport (°C)	35	Schymanski et al. (2009)
Leaf respiration rate per volume (-)	0.07	Schymanski et al. (2009)

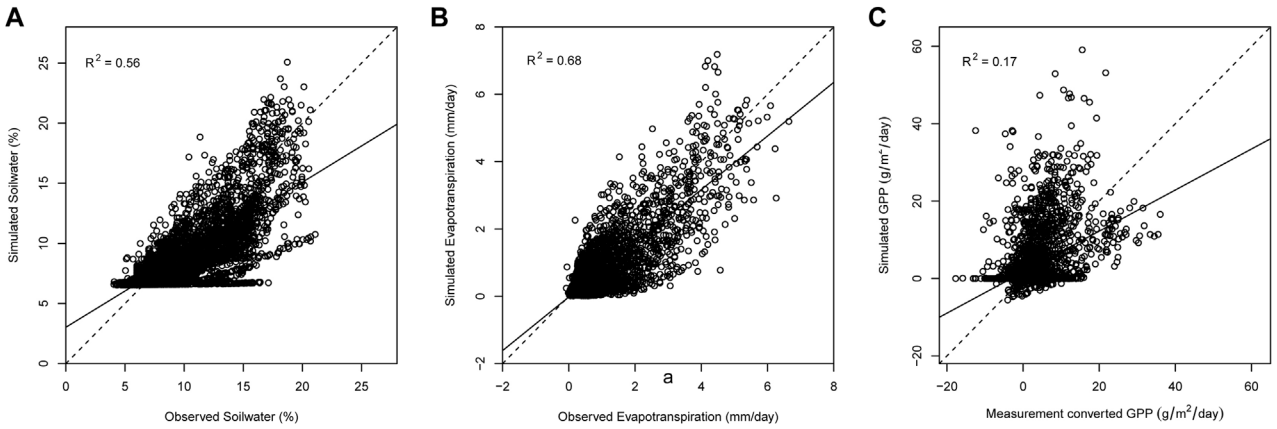

FIGURE 5 | Scatters of simulated and measured ET **(A)**, SW **(B)** and GPP **(C)**.

BREB system, atmospheric gradients of air temperature, moisture, and CO_2, are measured every 10 s and averaged to 20 min. These measurements are scaled up to 1 h for model input. Precipitation data of 2002 is gained from the precipitation database of Walnut Gulch Raingages because precipitation of 2002 from flux data is much lower than that from raingage indicating there is missing observation of this year.

The validation data including water vapor flux and CO_2 flux are also acquired from Bowen Ratio Energy Balance System. Considering the quality and continuity of the data, the meteorological data from 1998-01-01 to 2007-12-31 are selected to force the DisVOM model.

The model only has the outcome of leaf photosynthesis, which can not be directly measured. In this study, as grass is the main vegetation in the ecosystem, we convert the measured Net Ecosystem Exchange (NEE) of CO_2 through a flux-partition method into GPP and ecosystem respiration (R_e) by the following equation.

$$GPP = -NEE + R_e \tag{17}$$

As R_e is explicitly dependent on air or soil temperature, R_e can be estimated from observed temperature data using the exponential regression model (van't Hoff, 1884)

$$R_e = a * exp(b * T_s) \tag{18}$$

Where a and b are fitting parameters, T_s is the soil temperature. As nighttime NEE values are equal to R_e as GPP equals 0, therefore we can use the nighttime NEE (nighttime is defined as down solar radiation <1 W m^{-2}) and soil temperature to estimate the parameters. Once the parameters are fitted, the daytime NEE can be computed (**Figure 3**).

3 RESULTS

3.1 Grid Cells Prioritization

Based on the flow direction data (**Figure 4A**), prioritization of the grid cells are conducted as illustrated in **Figure 4B**. All the grid cells are partitioned into 45 layers. Grid cells of lower layers have higher priority, and grid cells in the same layer can model simultaneously.

3.2 Model Parameterization

Parameters for the optimality-based model consist of soil parameters (**Table 2**) and vegetation parameters. Soil parameters are specified according to Scott et al., 2000. In the VOM model, the soil profile is subdivided into sub-layers but the van Genuchten soil parameters of each layer are treated as one value. Therefore, we then average the soil parameters of each sub-layer from the work of Scott et al (2000) and obtain one value for the whole soil profile.

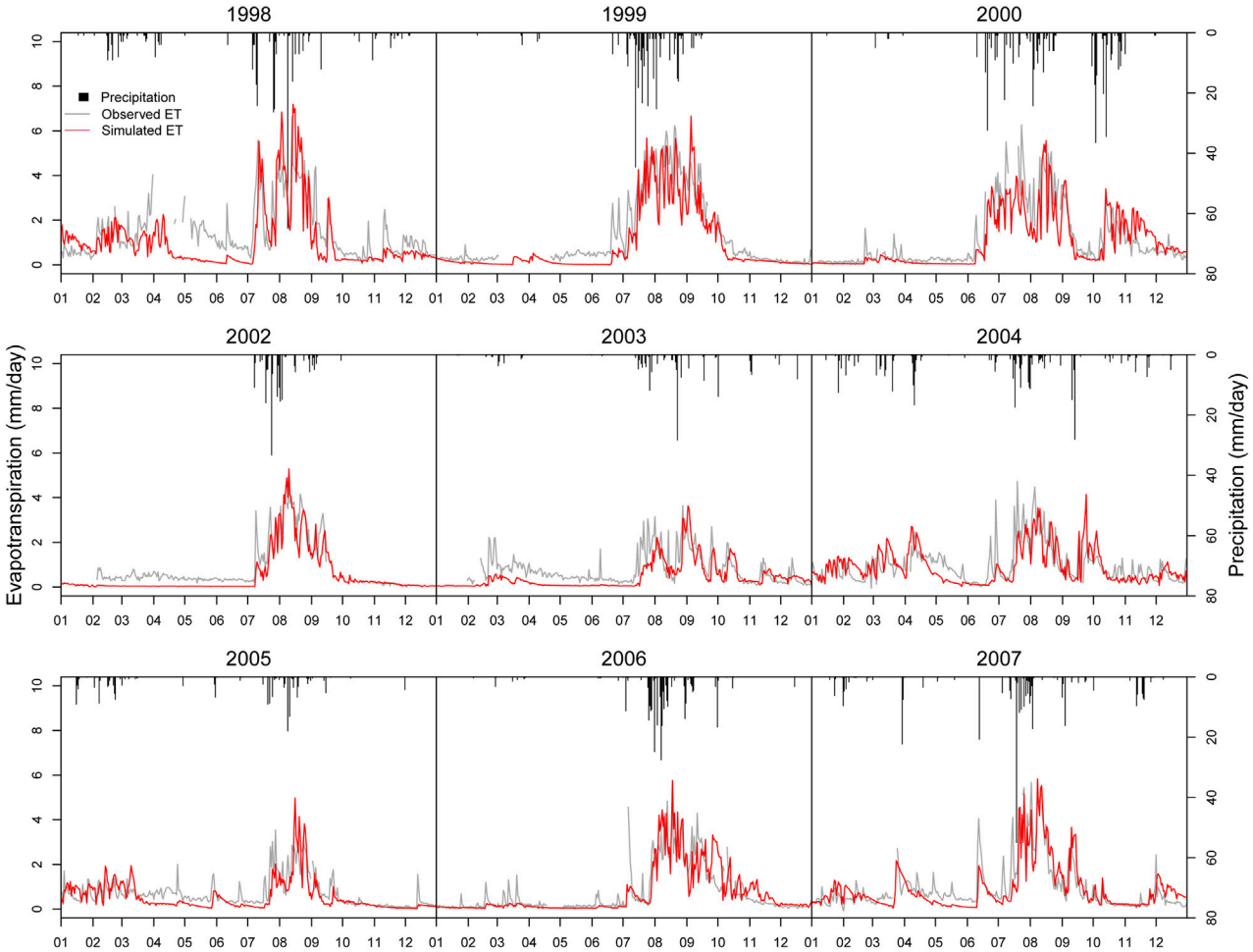

FIGURE 6 | Comparison of simulated and observed daily ET.

TABLE 3 | Accuracy of simulation.

Fluxes	RMSE	Standard deviation
Soil water	2.44 mm/day	3.67 mm/day
ET	0.57 mm/day	1.09 mm/day
GPP	5.72 g/m^2/day	4.84 g/m^2/day

Most of the vegetation parameters in this study are derived from vegetation optimization. Some of the parameters that cannot be optimized are obtained from existing studies, as shown in **Table 2**. These parameters are prescribed according to the average value of C3 shrub from previous studies (Schymanski et al., 2009; Massad et al., 2007; Lei et al., 2009).

3.3 Validation of Simulated Fluxes

The simulated daily ET, soil water and GPP are compared with the measured at the Kendall grass site. As the flux measurements of 2001 are seriously absent, this year is removed for validation. A one-to-one comparison of simulated values and measurements is shown in **Figure 5**. Most of the points are distributed along the 1:1 line, indicating a good agreement between the observed data and simulated values.

Figure 6 illustrates the measured and simulated ET from 1998-2007 (except 2001). The simulated ET shows a similar seasonal dynamic pattern with measured ET. It indicates that the model can explain the variability of measured ET. As compared with precipitation data, we can see that the model captures the ET dynamic reasonably well in response to rainfall events. When precipitation occurs ET occurs and simulated ET is mainly concentrated in monsoon when rainfall concentrates. Root Mean Square Error (RMSE) of ET simulation is 0.57 mm/day, which is lower than the standard deviation of ET, 1.09 mm/day (**Table 3**). **Figure 7** shows simulated daily evaporation and transpiration. Evaporation and transpiration both mainly occur in the monsoon following the precipitation events. Evaporation responds immediately to precipitation events, while transpiration shows a lagged response to those events. In spring or winter, ET is mainly dominated by evaporation, while in the monsoon, transpiration increases quickly after precipitation and dominates ET. These results are consistent with Emmerich and Verdugo (2008).

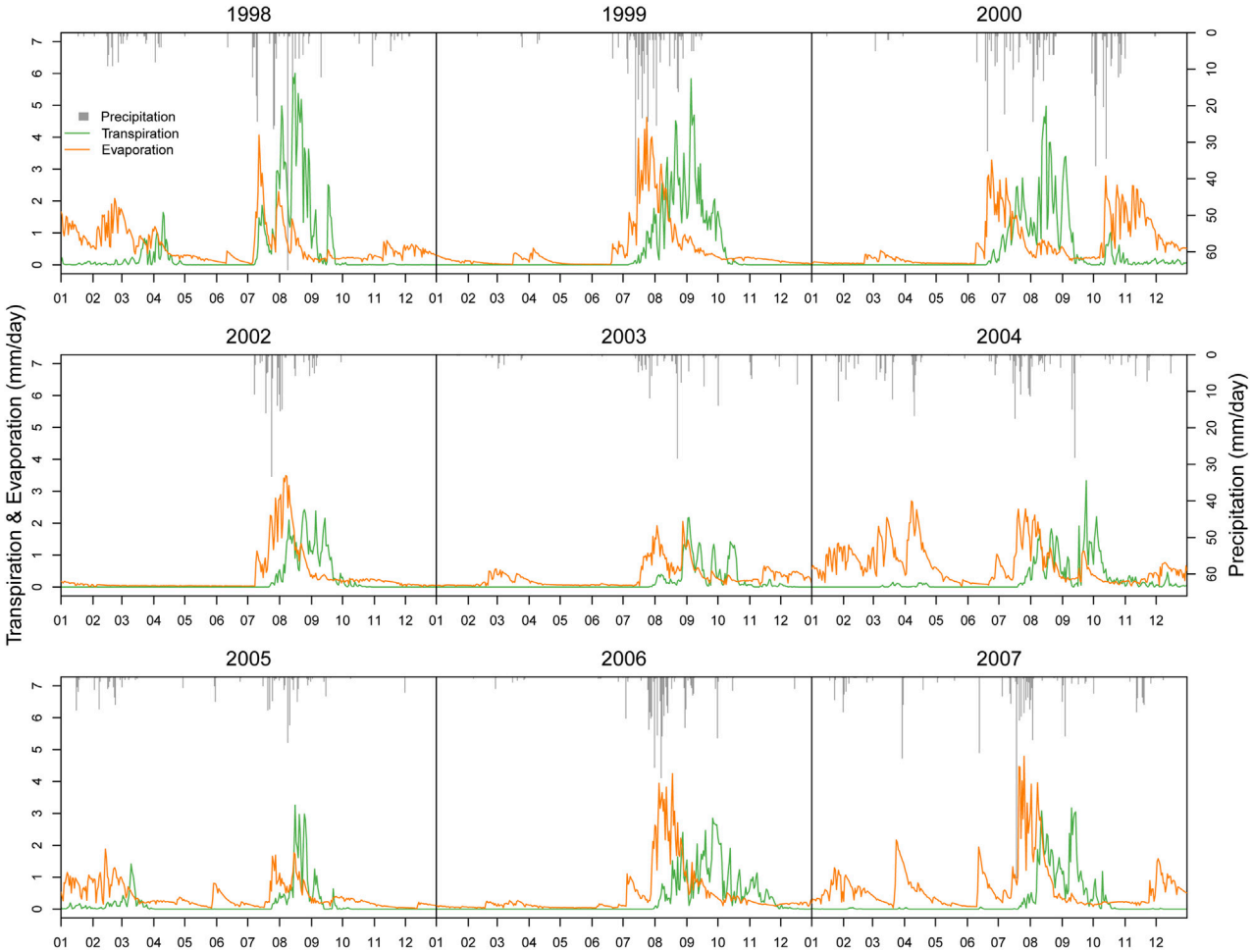

FIGURE 7 | Comparison of simulated soil evaporation and transpiration.

Figure 8 shows the seasonal dynamic of measured and simulated soil water. Generally, the performance of soil water simulation is well. Soil water in the monsoon season is simulated pretty well with nearly the same peak and corresponding seasonal variation. Root Mean Square Error (RMSE) of soil water simulation is 2.44 mm/day, lower than the standard deviation of soil water, 3.67 mm/day (**Table 3**). However, there is a tendency to underestimate soil water in the spring, especially in those years with low spring precipitation (like 2005 and 2006). It might be because, as precipitation is relatively small, the model assumes vegetation is inactive, and roots are not developed, leading to low soil water restored.

Figure 9 gives the simulated results of daily GPP and NEE-partitioned GPP. Generally, the performance of GPP simulation is acceptable. Simulated GPP shows the same patterns with NEE-partitioned GPP in most years. RMSE of GPP simulation is 5.72 g/m^2/day, which is slightly higher than the standard deviation of GPP, 4.84 g/m^2/day (**Table 3**). However, we also find a tendency to underestimate GPP in some years. It might be due to the unreasonable prediction of vegetation cover as it underestimates vegetation in the non-growing season.

3.4 Spatial Patterns of ET and GPP

Figure 10 shows the spatial distributions of annual ET and GPP of DisVOM (**Figure 10**) for 1999. The spatial variations of ET and GPP exhibit nearly identical patterns. Comparing the distribution of slope and Topographic Wetness Index (TWI) (**Figure 11**), we can find that the spatial distribution of annual ET and GPP of DisVOM demonstrates similar patterns with slope and TWI. Grid cells with high slopes tend to have low ET and GPP, while grid cells with low slopes tend to have high ET and GPP. Grid cells with high TWI are inclined to have high ET and GPP, while grid cells with low TWI appear to have low ET and GPP.

ET and GPP of all grid cells in the watershed are sorted and divided into eight groups to quantitatively analyze the correlation between simulated ET and GPP with slope and TWI. The averaged value of ET, GPP, slope and TWI of each group is estimated. As illustrated in **Figure 12**, ET and GPP show a significant negative correlation with slope and positive correlation with TWI. Annual ET and GPP decreased

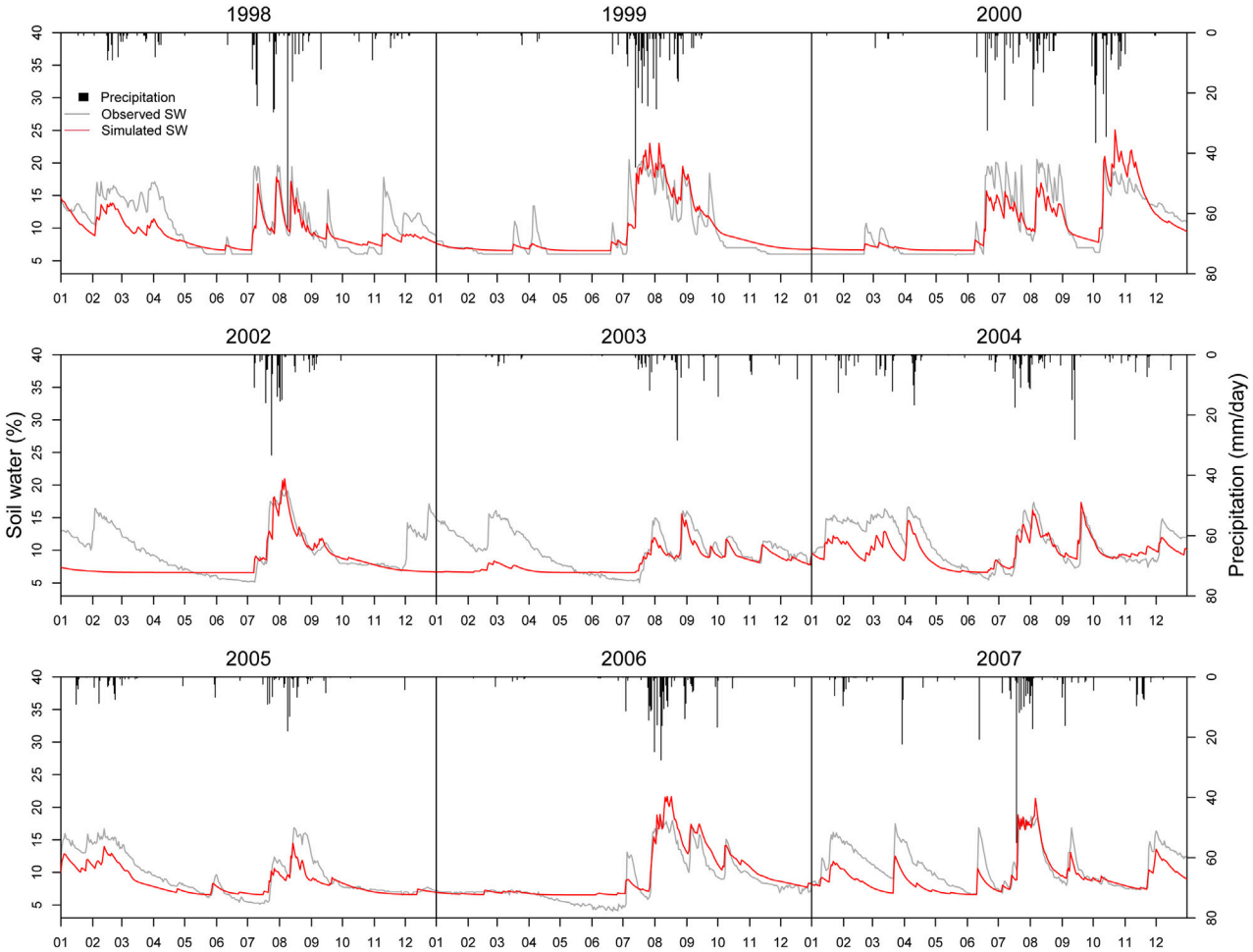

FIGURE 8 | Comparison of simulated and observed daily soil water dynamic.

with the increase of slope. Annual ET and GPP increase with the increase of TWI.

In general, the spatial pattern of annual ET and GPP reflect the fundamental impact of topographical controls on soil water and root water uptake. The lower portions of the watershed have higher soil water content due to the above recharge. Grass could develop high root density in those areas and consequent high transpiration and carbon assimilation. It is the truth, as in water-limited ecosystems, water is a scarce resource for plants and plays a fundamental role in vegetation dynamics, such as canopy property and root density. The DisVOM model can discriminate the effect of topography on redistribution of precipitation into soil water and the consequent variations in ET and carbon assimilation.

4 DISCUSSION AND CONCLUSIONS

This paper presents an optimality-based watershed ecohydrological model. The model subdivided watershed into grid cells. On each grid cell, an optimality-based ecohydrological model, VOM is applied to describe the vertical water and vegetation mutual interactions. Overland flow is incorporated to the VOM model to simulate soil water redistribution and subsequently modify vegetation properties optimization. A hierarchical strategy is developed to integrate lateral flow to VOM model, which prioritizes the grid cells of the watershed into different levels according to the dependency of grid cells. The model is tested in the Walnut Gulch watershed and demonstrates good consistency with site measurements. However, there are still some limitations of this study, which can be summarized as follows.

4.1 Discrepancy of Simulation in Non-growing Season

The model demonstrates good performance in growing season, however, evident discrepancies arise during winter and spring. This is probably, to a great degree, attributed to the underestimation of soil water.

As we look into the precipitation data and compare spring soil water of different years, one possible reason might be the undocumented precipitation. For 2002, we utilize the data from rain gauge, because precipitation measurement of flux is

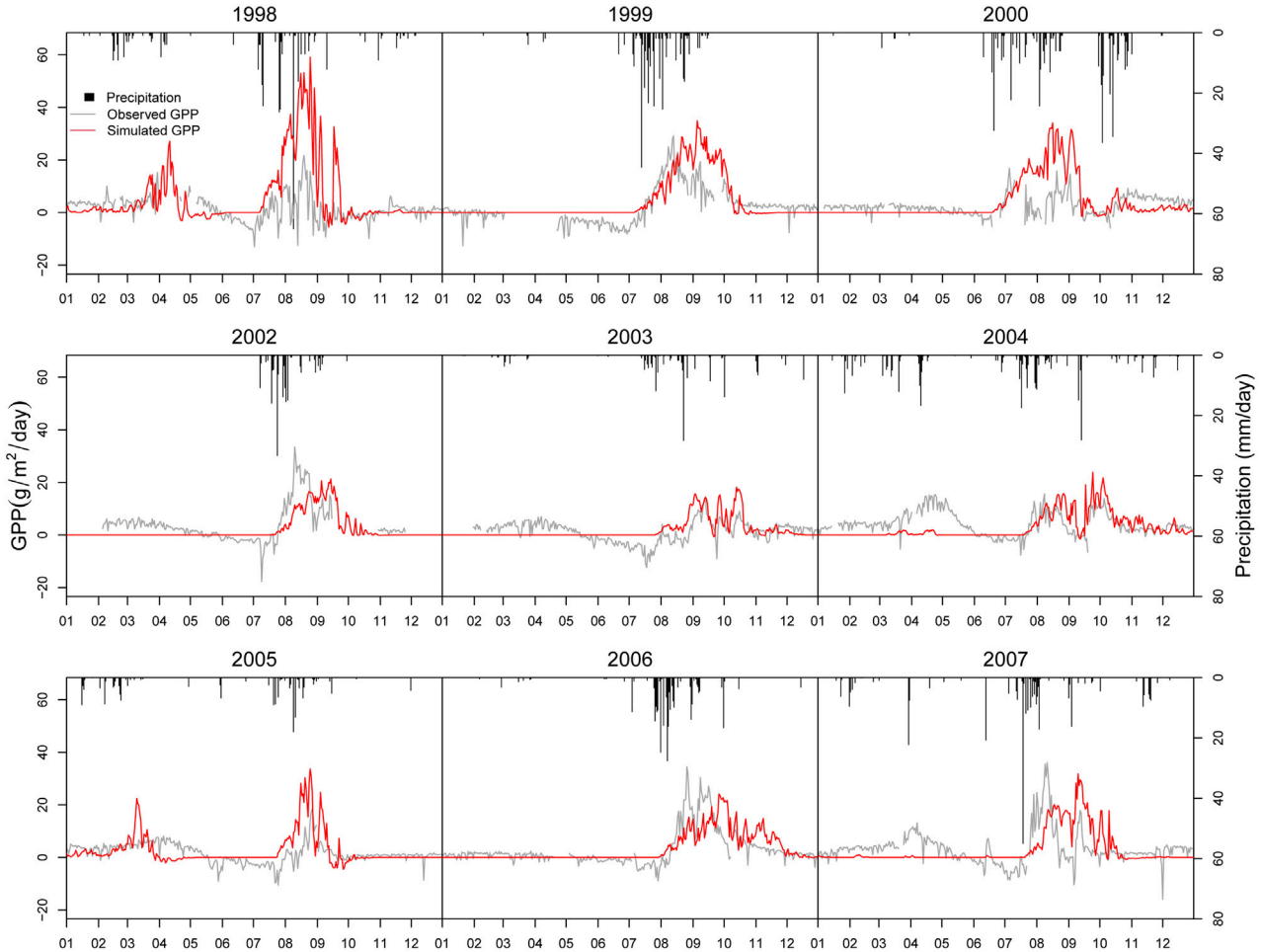

FIGURE 9 | Comparison of simulated daily GPP and NEE-partitioned GPP.

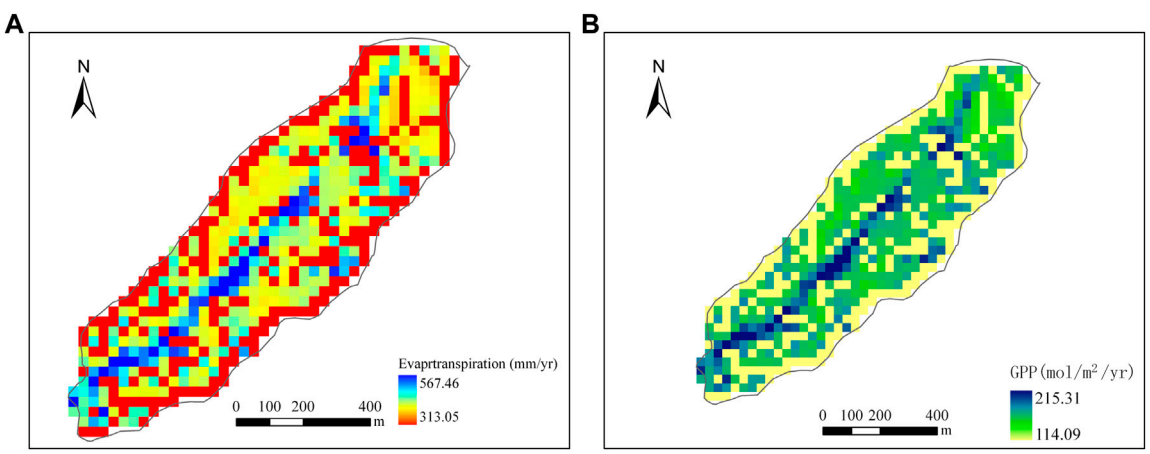

FIGURE 10 | Spatial distribution of ET **(A)** and GPP **(B)** simulated by DisVOM and VOM.

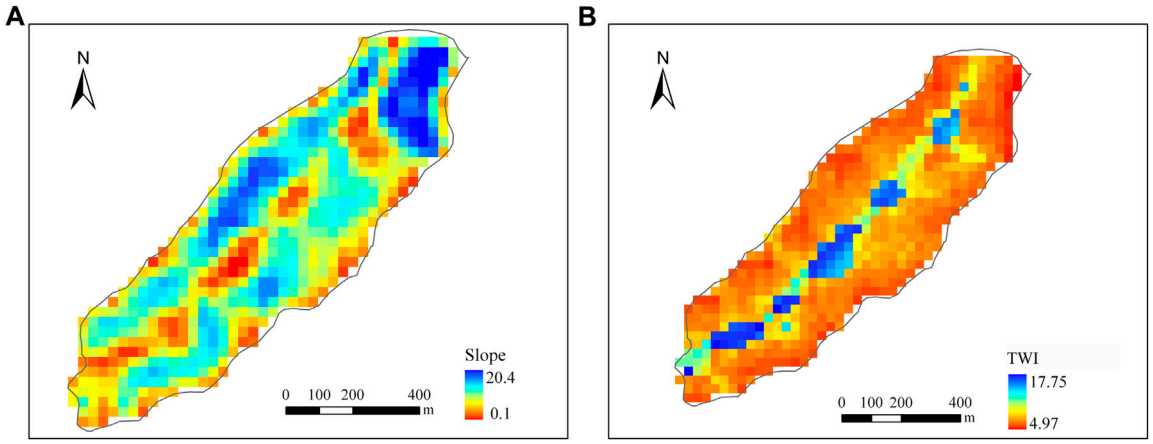

FIGURE 11 | Spatial distribution of slope **(A)** and TWI **(B)** of Kendall watershed.

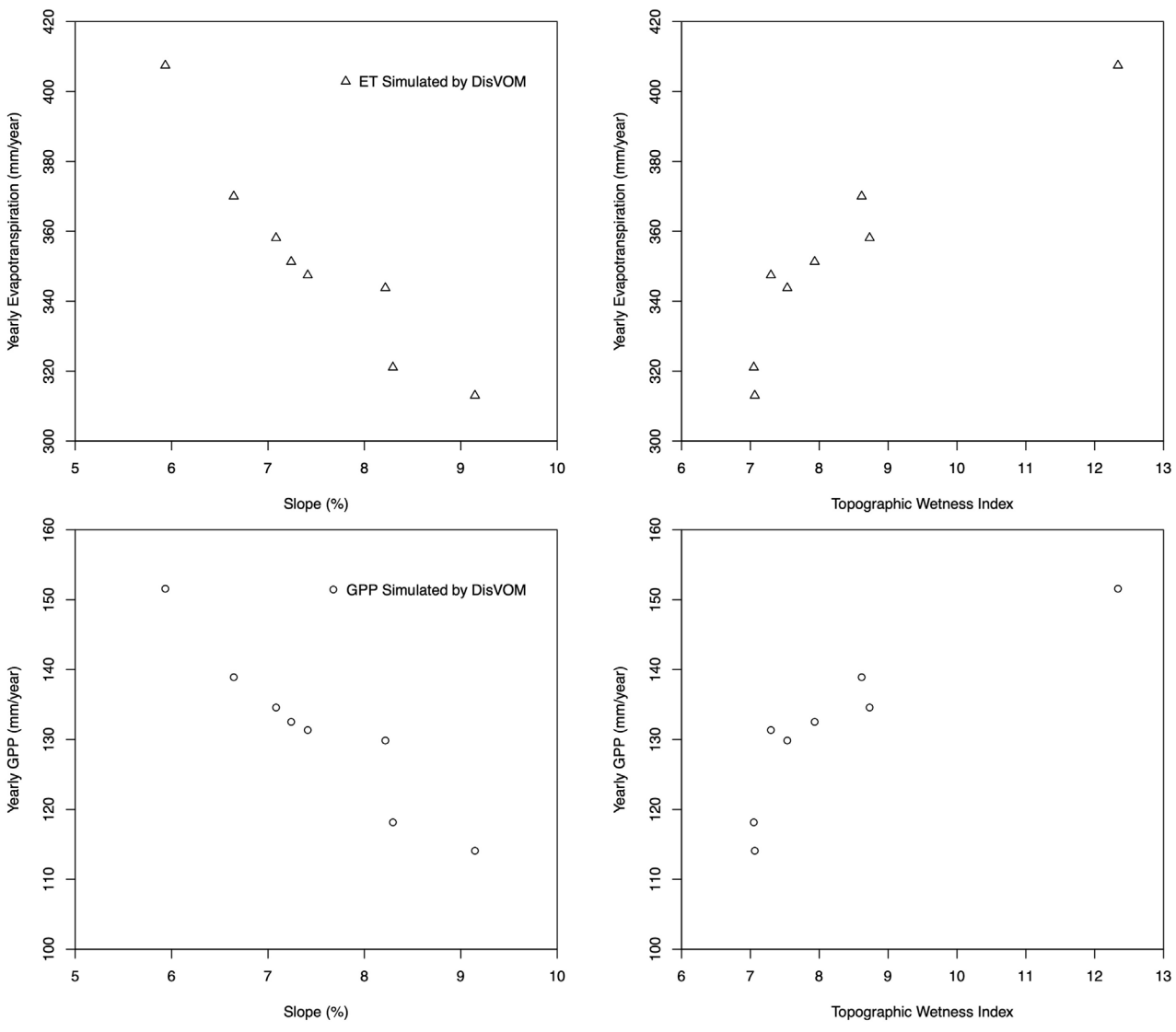

FIGURE 12 | Correlation of annual ET and GPP with slope and TWI.

lower. However, there was no rain in the spring of 2002, which is unrealistic. In contrast, from the flux data we find a small amount of precipitation. Besides, we compare observed soil water of two period 2002.12-2003.05 and 2007.01-2007.05. The two periods have quite similar soil water with respect to similar amount, shape and peak. However, measured precipitation in 2002.12–2003.05 is considerably lower than precipitation of 2007.01–2007.05. As there is no precipitation for nearly 2 months for both periods, antecedent soil moisture is nearly the same for the two periods. Hence, it is unlikely to have the same soil water with dramatically different amount of precipitation.

Another possible reason might be the uniform setting of soil profile. In the model, soil profile is subdivided into sub-layers with same depth (25 cm). According to the previous study in this area, vertical soil properties vary dramatically (Scott et al., 2000). For example, saturated hydraulic conductivity of the top layer (0–4 cm) is about 31.1 cm/day, 18.7 cm/day for the second layer (4–9 cm), 7.2 cm/day for the third layer (9–16 cm), 8.5 cm/day for the third layer (16–35 cm), 8.7 cm/day for the third layer (35–62 cm), and 36.9 cm/day for the bottom layer (62–150 cm). For the layers near surface, a depth of 25 cm is unable to characterize the variation of soil properties, which could attributed to the underestimation of soil water in the non-growing season when root uptake of vegetation is low.

4.2 Discrepancy of GPP Simulation

There are some inconsistent of GPP for some years especially in the non-growing season. For 2002, 2003, 2004, and 2007, grass grew in the spring as there is GPP partitioned from NEE, but from the model there is no grass grew as GPP is little. This is possibly caused by the underestimation of soil water. As stated in 4.1, soil water is dramatically underestimated, so the available water for root water uptake is low. This might directly influence root development and leaf growth, leading to low simulated carbon assimilation.

Another reason is that there is no directed measured GPP. GPP separated from NEE measurements are used for validation. Not only the NEE measurements are highly associated with measurement error, but also the NEE separation is associated with uncertainty (Raj, et al., 2016). This might contributed partly to the inconsistent of simulated and observed GPP comparison.

4.3 Simple Representation of Lateral Flow

Another problem of this study is that the model only considers overflow routing, subsurface flow and ground water is not simulated. Therefore it can't give stream flow as the other two components are neglected. The main concern of this study is to explore whether incorporated a kind of lateral flow could alter the soil water redistribution and influence vegetation-soil-water mutual interaction and coevolution. Incorporation of subsurface flow and ground water would otherwise require much more parameters which need to be calibrated. Although VOM model has the unique merit that it does not need prior vegetation properties as input, we still find environment parameters (such as topography, soil), associated with hydrological processes are difficult to determine. Calibration might need for hydrological processes. Nevertheless, the lateral flow of the current model is simple and it is unrealistic without simulating the other component of stream flow. We will integrate subsurface flow and ground water to the model in the future and learn the method from big data science for hydrological parameters estimation (Liu et al., 2017; Zhang et al., 2021).

Despite the above issues, the model presents in this study is capable of predicting ET, soil water, and GPP reasonably well, for most of years, especially for the growing season. The seasonal dynamic of ET, soil water, and GPP demonstrate good consistency with observations. The model produces reasonable spatial distribution of ET and GPP, indicating it is able to capture the influence of lateral flow on water redistribution, vegetation dynamic and so on. It can be concluded that the model effectively extends VOM from field scale to watershed scale and could be a useful tool assessing the impact of climate change and human activities on vegetation and water cycle.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: http://www.tucson.ars.ag.gov/dap/.

AUTHOR CONTRIBUTIONS

LC carried out model development, tested the model, analyzed the results, and drafted the manuscript. ZH participated in collecting and processing some of the data for model validation. XD provided fundamental suggestions for this study. MYAK reviewed and revised the manuscript. XL and JW processed some of the figures, and revised the manuscript.

FUNDING

This study is supported by Open Research Fund Program of State key Laboratory of Hydroscience and Engineering (sklhse-2021-A-01), Water Conservancy Science and Technology Innovation Project in Guangdong Province (No. 202012), National Key R&D Program of China (No. 2018YFC1508103), Strategic Priority Research Program of the Chinese Academy of Sciences (XDA 19080103), Innovation Drive Development Special Project of Guangxi (GuikeAA20302022).

REFERENCES

Amein, M., and Fang, C. S. (1969). *Stream Flow Routing (With Applications to North Carolina Rivers), Report No. 17*. Raleigh, NC: Water Resources Research Institute of the University of North Carolina.

Ball, J. T., Woodrow, I. E., and Berry, J. A. (1987). "A Model Predicting Stomatal Conductance and its Contribution to the Control of Photosynthesis under Different Environmental Conditions," in *Progress in Photosynthesis Research* (Dordrecht: Martinus Nijhoff Publishers), 221–224. doi:10.1007/978-94-017-0519-6_48

Band, L. E., Patterson, P., Nemani, R., and Running, S. W. (1993). Forest Ecosystem Processes at the Watershed Scale: Incorporating Hillslope Hydrology. *Agric. For. Meteorology* 63, 93–126. doi:10.1016/0168-1923(93)90024-c

Berninger, F., Mäkelä, A., and Hari, P. (1996). Optimal Control of Gas Exchange during Drought: Empirical Evidence. *Ann. Bot.* 77, 469–476. doi:10.1006/anbo.1996.0057

Biederman, J. A., Scott, R. L., Bell, T. W., Bowling, D. R., Dore, S., Garatuza-Payan, J., et al. (2017). CO 2 Exchange and Evapotranspiration across Dryland Ecosystems of Southwestern North America. *Glob. Change Biol.* 23 (10), 4204–4221. doi:10.1111/gcb.13686

Cabon, A., Martínez-Vilalta, J., Martínez de Aragón, J., Poyatos, R., and De Cáceres, M. (2018). Applying the Eco-Hydrological Equilibrium Hypothesis to Model Root Distribution in Water-Limited Forests. *Ecohydrology* 11 (7), e2015. doi:10.1002/eco.2015

Caylor, K. K., Scanlon, T. M., and Rodriguez-Iturbe, I. (2009). Ecohydrological Optimization of Pattern and Processes in Water-Limited Ecosystems: A Trade-Off-Based Hypothesis. *Water Resour. Res.* 45, W08407. doi:10.1029/2008WR007230

Caylor, K. K., Scanlon, T. M., and Rodriguez-Iturbe, I. (2004). Feasible Optimality of Vegetation Patterns in River Basins. *Geophys. Res. Lett.* 31, a–n. doi:10.1029/2004GL020260

Celia, M. A., Bouloutas, E. T., and Zarba, R. L. (1990). A General Mass-Conservative Numerical Solution for the Unsaturated Flow Equation. *Water Resour. Res.* 26 (7), 1483–1496. doi:10.1029/wr026i007p01483

Cevza, M. K., Miguel, A., Medina, J., and Rao, P, (2005). On Numerical Modeling of Overland Flow. *Appl. Maths. Comput.* 166, 724–740. doi:10.1016/j.amc.2004.06.063

Chen, L., Sun, L., Liu, W., Wang, L., Wu, H., Zhu, A.-X., et al. (2019). Evapotranspiration Partitioning Using an Optimality-Based Ecohydrological Model in a Semiarid Shrubland. *Int. J. Digital Earth* 12 (12), 1423–1440. doi:10.1080/17538947.2018.1503741

Chen, L., Wang, L., Ma, Y., and Liu, P. (2015). Overview of Ecohydrological Models and Systems at the Watershed Scale. *IEEE Syst. J.* 9 (3), 1091–1099. doi:10.1109/jsyst.2013.2296979

Cowan, I. R., Farquhar, G. D., and Jennings, D. H. (1977). "Stomatal Function in Relation to Leaf Metabolism and Environment," in *Integration of Activity in the Higher Plant* (Cambridge: Cambridge University Press).

Cowan, I. R. (2002). Fit, Fitter, Fittest: Where Does Optimization Fit in? *Silva Fennica* 36 (3), 744–754. doi:10.14214/sf.536

De Boer, H. J., Lammertsma, E. I., Wagner-Cremer, F., Dilcher, D. L., Wassen, M. J., and Dekker, S. C. (2011). Climate Forcing Due to Optimization of Maximal Leaf Conductance in Subtropical Vegetation under Rising CO2. *Proc. Natl. Acad. Sci.* 108 (10), 4041–4046. doi:10.1073/pnas.1100555108

Duan, Q. Y., Sorooshian, S., and Gupta, V. K. (1994). Optimal Use of the SCE-UA Global Optimization Method for Calibrating Watershed Models. *J. Hydrol.* 158 (3-4), 265–284. doi:10.1016/0022-1694(94)90057-4

Eagleson, P. S. (2002). *Ecohydrology: Darwinian Expression of Vegetation Form and Function*. Cambridge: Cambridge University Press.

Eagleson, P. S. (1978). Climate, Soil, and Vegetation: 1. Introduction to Water Balance Dynamics. *Water Resour. Res.* 14 (5), 705–712. doi:10.1029/wr014i005p00705

Emmerich, W. E., and Verdugo, C. L. (2008). Long-term Carbon Dioxide and Water Flux Database, Walnut Gulch Experimental Watershed, Arizona, United States. *Water Resour. Res.* 44, W05S09. doi:10.1029/2006WR005693

Fatichi, S., Pappas, C., and Ivanov, V. Y. (2016). Modeling Plant-Water Interactions: an Ecohydrological Overview from the Cell to the Global Scale. *WIREs Water* 3 (3), 327–368. doi:10.1002/wat2.1125

Franklin, O., Harrison, S. P., Dewar, R., Farrior, C. E., Brännström, Å, Dieckmann, U., et al. (2020). Organizing Principles for Vegetation Dynamics. *Nat. Plants* 6 (5), 444–453. doi:10.1038/s41477-020-0655-x

Gao, H., Birkel, C., Hrachowitz, M., Tetzlaff, D., Soulsby, C., and Savenije, H. H. G. (2019). A Simple Topography-Driven and Calibration-free Runoff Generation Module. *Hydrol. Earth Syst. Sci.* 23, 787–809. doi:10.5194/hess-23-787-2019

Gao, H., Hrachowitz, M., Schymanski, S. J., Fenicia, F., Sriwongsitanon, N., and Savenije, H. H. G. (2014). Climate Controls How Ecosystems Size the Root Zone Storage Capacity at Catchment Scale. *Geophys. Res. Lett.* 41 (22), 7916–7923. doi:10.1002/2014gl061668

Goodrich, D. C., Keefer, T. O., and Unkrich, C. L., (2008). Smith Long-Term Precipitation Database, Walnut Gulch Experimental Watershed, Arizona, United States. *Water Resour. Res.* 44, W05S04. doi:10.1029/2006R00578210.1029/2006wr005782

Govind, A., Chen, J. M., Margolis, H., Ju, W., Sonnentag, O., and Giasson, M.-A. (2009). A Spatially Explicit Hydro-Ecological Modeling Framework (BEPS-TerrainLab V2.0): Model Description and Test in a Boreal Ecosystem in Eastern North America. *J. Hydrol.* 367, 200–216. doi:10.1016/j.jhydrol.2009.01.006

Huang, F., Mo, X., Hu, S., and Li, L. (2020). Agricultural Water Optimization Coupling with a Distributed Ecohydrological Model in a Mountain-plain basin. *J. Hydrol.* 590, 125336. doi:10.1016/j.jhydrol.2020.125336

Hwang, T., Band, L., and Hales, T. (2009). Ecosystem Processes at the Watershed Scale: Extending Optimality Theory from Plot to Catchment. *Water Resour. Res.* 45, W11425. doi:10.1029/2009WR007775

Ivanov, V. Y., Bras, R. L., and Vivoni, E. R. (2008). Vegetation-hydrology Dynamics in Complex Terrain of Semiarid Areas: 1. A Mechanistic Approach to Modeling Dynamic Feedbacks. *Water Resour. Res.* 44, W03429. doi:10.1029/2006WR00558810.1029/2006wr005588

Keefer, T. O., Moran, M. S., and Paige, G. B. (2008). Long-term Meteorological and Soil Hydrology Database, Walnut Gulch Experimental Watershed, Arizona, United States. *Water Resour. Res.* 44, W05S07. doi:10.1029/2006WR00570210.1029/2006wr005702

Kuppel, S., Tetzlaff, D., Maneta, M. P., and Soulsby, C. (2018). What Can We Learn from Multi-Data Calibration of a Process-Based Ecohydrological Model? *Environ. Model. Softw.* 101, 301–316. doi:10.1016/j.envsoft.2018.01.001

Lei, H., Yang, D., Schymanski, S. J., and Sivapalan, M. (2009). Modeling the Crop Transpiration Using an Optimality-Based Approach. *Sci. China Ser. E-technol. Sci.* 51 (S2), 60–75. doi:10.1007/s11431-008-6008-z

Liu, P., Zhang, H., and Eom, K. B. (2017). Active Deep Learning for Classification of Hyperspectral Images. *IEEE J. Sel. Top. Appl. Earth Observations Remote Sensing* 10 (2), 712–724. doi:10.1109/jstars.2016.2598859

Massad, R. S., Tuzet, A., and Bethenod, O. (2007). The Effect of Temperature on C(4)-type Leaf Photosynthesis Parameters. *Plant Cel Environ* 30 (10), 1191–1204. doi:10.1111/j.1365-3040.2007.01691.x

Pauwels, V. R. N., Verhoest, N. E. C., De Lannoy, G. J. M., Guissard, V., Lucau, C., and Defourny, P. (2007). Optimization of a Coupled Hydrology-Crop Growth Model through the Assimilation of Observed Soil Moisture and Leaf Area index Values Using an Ensemble Kalman Filter. *Water Resour. Res.* 43. doi:10.1029/2006WR004942

Raj, R., Hamm, N. A. S., Tol, C. v. d., and Stein, A. (2016). Uncertainty Analysis of Gross Primary Production Partitioned from Net Ecosystem Exchange Measurements. *Biogeosciences* 13 (5), 1409–1422. doi:10.5194/bg-13-1409-2016

Rice, J. S., and Emanuel, R. E. (2019). Ecohydrology of Interannual Changes in Watershed Storage. *Water Resour. Res.* 55 (10), 8238–8251. doi:10.1029/2019wr025164

Rodriguez-Iturbe, I., D'Odorico, P., Porporato, A., and Ridolfi, L. (1999). On the Spatial and Temporal Links between Vegetation, Climate, and Soil Moisture. *Water Resour. Res.* 35, 3709–3722. doi:10.1029/1999wr900255

Rodriguez-Iturbe, I., and Porporato A. (Editors) (2005). "*Ecohydrology of Water Controlled Ecosystems: Soil Moisture and Plant Dynamics*," (London: Cambridge University Press).

Schymanski, S. J., Sivapalan, M., Roderick, M. L., Beringer, J., and Hutley, L. B. (2008). An Optimality-Based Model of the Coupled Soil Moisture and Root Dynamics. *Hydrol. Earth Syst. Sci.* 12 (3), 913–932. doi:10.5194/hess-12-913-2008

Schymanski, S. J., Sivapalan, M., Roderick, M. L., Hutley, L. B., and Beringer, J. (2009). An Optimality-Based Model of the Dynamic Feedbacks between Natural Vegetation and the Water Balance. *Water Resour. Res.* 45. doi:10.1029/2008WR006841

Scott, R. L., Shuttleworth, W. J., Keefer, T. O., and Warrick, A. W. (2000). Modeling Multiyear Observations of Soil Moisture Recharge in the Semiarid American Southwest. *Water Resour. Res.* 36 (8), 2233–2247. doi:10.1029/2000wr900116

Sivapalan, M. (2009). The Secret to 'doing Better Hydrological Science': Change the Question! *Hydrol. Process.* 23 (9), 1391–1396. doi:10.1002/hyp.7242

Speich, M. J. R., Lischke, H., and Zappa, M. (2018). Testing an Optimality-Based Model of Rooting Zone Water Storage Capacity in Temperate Forests. *Hydrol. Earth Syst. Sci.* 22 (7), 4097–4124. doi:10.5194/hess-22-4097-2018

Sutherland, W. J. (2005). The Best Solution. *Nature* 435, 569. doi:10.1038/435569a

USDA (2008). *Soil Survey of Walnut Gulch Experimental Watershed*. Arizona.

Van der Tol, C., MeestersDolman, A. G. C. A. A. J., Dolman, A. J., and Waterloo, M. J. (2008). Optimum Vegetation Characteristics, Assimilation, and Transpiration during a Dry Season: 1. Model Description. *Water Resour. Res.* 44 (3), W03421. doi:10.1029/2007WR006241

Van't Hoff, J. H. (1884). *Etudes de dynamique chimique (studies of Chenmical dynamics)*. Amsterdam, Netherlands: Frederik Muller and Co.

Wang, C., Fu, B., Zhang, L., and Xu, Z. (2019). Soil Moisture-Plant Interactions: an Ecohydrological Review. *J. Soils Sediments* 19 (1), 1–9. doi:10.1007/s11368-018-2167-0

Xia, J., Zhang, Y., Mu, X., Zuo, Q., Zhou, Y., and Zhao, G. (2021). A Review of the Ecohydrology Discipline: Progress, Challenges, and Future Directions in China. *J. Geogr. Sci.* 31, 1085–1101. doi:10.1007/s11442-021-1886-0

Zhang, L., Liu, P., Zhao, L., Wang, G., Zhang, W., and Liu, J. (2021). Air Quality Predictions with a Semi-supervised Bidirectional LSTM Neural Network. *Atmos. Pollut. Res.* 12 (1), 328–339. doi:10.1016/j.apr.2020.09.003

Conflict of Interest: Author XL was employed by company China Water Northeastern Investigation, Design and Research Co., Ltd.

The remaining 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.

frontiers in Environmental Science

ORIGINAL RESEARCH
published: 25 February 2022
doi: 10.3389/fenvs.2022.799027

Remote Sensing Image-Based Comprehensive Monitoring Detection Platform for Coastal Tidal Mudflat Ecological Development

*Shengjun Xiao[1], Lin Yi[2], Zengjie Wang[1], Huiyu Liu[1], Hong Gao[3] and Zhaoyuan Yu[1]**

[1]School of Geography, Nanjing Normal University, Nanjing, China, [2]College of Earth Sciences, Chengdu University of Technology, Chengdu, China, [3]Faculty of Geomatics, Lanzhou Jiaotong University, Lanzhou, China

Considering the embarrassments related to inadequate safety monitoring detection and chaotic development order during the current ecological development of coastal tidal mudflat, we put forward a remote sensing image-based comprehensive monitoring detection platform integrating image data acquisition, mechanism analysis, comprehensive evaluation, and intelligent monitoring detection for coastal tidal mudflat ecological development. The platform includes imaging and data fusion technology, comprehensive evaluation and spatial layout optimization technology, and dynamic safety monitoring and early warning technology. A case study on Dafeng coastal tidal mudflat was conducted. The results show that our research can provide substantial reference for the coordination, dynamic management, and comprehensive regulation for the ecological development of coastal tidal mudflat.

Keywords: remote sensing image, ecological development, monitoring detection platform, coastal tidal mudflat, data analysis

OPEN ACCESS

Edited by:
Peng Liu,
Institute of Remote Sensing and Digital Earth (CAS), China

Reviewed by:
Qinjun Qiu,
China University of Geosciences Wuhan, China
Fei Li,
National Marine Environmental Monitoring Center, China

****Correspondence:***
Zhaoyuan Yu
yuzhaoyuan@njnu.edu.cn

Specialty section:
This article was submitted to Environmental Informatics and Remote Sensing, a section of the journal Frontiers in Environmental Science

Received: *21 October 2021*
Accepted: *19 January 2022*
Published: *25 February 2022*

Citation:
Xiao S, Yi L, Wang Z, Liu H, Gao H and Yu Z (2022) Remote Sensing Image-Based Comprehensive Monitoring Detection Platform for Coastal Tidal Mudflat Ecological Development. Front. Environ. Sci. 10:799027. doi: 10.3389/fenvs.2022.799027

1 INTRODUCTION

The ecosystem of coastal tidal mudflat is characterized by diversification in influencing element, complex processes, strong environmental sensitivity, and vulnerability and has a wide range of impacts derived from multiple channels (Turner, 2000; Dauvin, 2008; Wang, 2013). The existing development process for coastal tidal mudflat has many problems, such as the chaotic development pattern, the lack of comprehensive integration, and dynamic monitoring (Zhao et al., 2009; Liang et al., 2011). Furthermore, high-intensity development has also led to the aggravation of regional habitat destruction, biodiversity reduction, pollution, and resource degradation (Paoli et al., 2008; Hua et al., 2012). Considering the rational and coordinated development for coastal tidal mudflat, a batch of investigators have carried out multidimensional studies on the aspects of the comprehensive investigation and management of coastal tidal mudflat, relevant monitoring methods and technologies, analysis models and evaluation systems, regulation mechanisms and early warning platforms, etc. (Lee, 1999; Wang and Zou, 2009). With the transformation from an early single-objective management to a multi-objective, multi-scale, and multilevel coordinated and unified comprehensive management in coastal zones (Zhang, 1999; Pickaver et al., 2004; Hou et al., 2005), multiple elements including ecological, environmental, population, social, economic, and legal factors should also be considered in the dynamic regulation of coastal tidal mudflat. Under the constraints of ecology, environment, security, and other comprehensive factors, high-precision and high-resolution dynamic continuous monitoring of the development process should be realized, and

the spatiotemporal pattern characteristics and dynamic evolution mechanism of coastal tidal mudflat should also be explored so as to carry out suitability evaluation for ecological development (Brommer and Bochev-van der Burgh, 2009; Li et al., 2011; Harvey et al., 2012). In view of urgent requirement for environmental safety and economic development, dynamic monitoring, evaluation, and regulation analysis in the development process of coastal tidal mudflat should be appreciated and the transformation from traditional static management to dynamic comprehensive monitoring detection should be realized (Zhang, 1999; Henocque, 2003; Wang and Li, 2006). Besides the community of scholars, supervision departments of government also emphasized on the trend that the management method of coastal ecosystem should transform to dynamic management and comprehensive regulation (Long et al., 2016).

Currently, there are many reports on land use monitoring detection platform, such as the mobile law enforcement and online analysis (Chen et al., 2009), regional land information-integrated management model (RLI2M) (Zhu and Huang, 2010), post-processing monitoring for land use (Xu et al., 2010), and "one map" for the land resources (Zhang, 2010). Despite these, there is still a lack of system platform that can support the integrated management and integration of ecology, environment, land, society, and economy, in terms of comprehensive monitoring detection of ecological development in coastal tidal mudflat. Moreover, the existing platforms of information management for coastal tidal mudflat are mostly targeted at a single-user group, therefore, lacking the cooperative participation management mechanism between the public and government departments and industries. It is necessary to construct the comprehensive coordinated information dynamic management mode of coastal zones with multi-scale, multi-objective and multilevel, in order to develop the information management of coastal tidal mudflat from the single-field to multi-field comprehensive open service stage. Subsequently, the network sharing intelligent regulation and real-time warning mechanism driven by the analysis results of the coupling effect and evolution law need to be formed (Liu et al., 2017; Zhang et al., 2021). Finally, a dynamic monitoring detection system and platform should be established to support the demand of mass data transmission, scheduling and analysis, realize the comprehensive integration of monitoring detection system and business process, dynamically supervise the development process of coastal tidal mudflat, and comprehensively coordinate the interests of all parties.

Based on the research approach of "data acquisition–process mechanism–comprehensive evaluation–intelligent monitoring detection," we constructed a technical system for the comprehensive supervision of ecological development of coastal tidal mudflat. Furthermore, we designed a comprehensive analysis model, involving a business process-driven data collection and integration system, security evaluation of coastal tidal mudflat ecological development, identification and analysis of sensitivity factors of ecological security, and the landscape optimization method of coastal tidal mudflat ecological development. Based on the this, a system platform for comprehensive supervision of coastal tidal mudflat development was further put forward. Our study aims to break through the traditional single-perspective and single-objective framework, realizing the conversion from the management mode based on remote sensing resources monitoring and summary evaluation to the intelligent management mode based on internal driving mechanism mining and simulation reasoning research. Finally, an integrated monitoring and analysis platform with multi-scale, multi-objective, multilevel, and comprehensive research was constructed.

2 METHODS

2.1 Research Scheme of Detection Platform for Ecological Development

It is of great significance to establish an information monitoring detection platform on coastal tidal mudflat development for realizing overall coordination, dynamic management, and comprehensive control of the ecological development process of coastal tidal mudflat (Christensen et al., 2008). In view of the demand for comprehensive supervision of the ecological development of coastal tidal mudflat, the comprehensive supervision of coastal tidal mudflat needs to integrate the data of multisource coastal tidal mudflat ecological environment and the current situation of development and utilization. From the perspective of coordinated development of resources, environment, industrial development, population, and output, the comprehensive evaluation index system and evaluation model for the ecological development of coastal shoal flats were established. Based on the this, we can identify the different development modes such as the stress index, ecological environment influence mechanism, analysis of the ecological development of safety factors, and different types of tidal flat-suitable development intensity, and solve the different development needs and constraints, to establish an ecology suitable for the coastal shoal development of the early warning and decision-making mechanism.

Based on methods mentioned previously, the ideas for constructing the monitoring detection platform for the ecological development of coastal tidal mudflat are shown in **Figure 1**. With respect to the conceptual design, it is mainly guided by the national coastal development strategy, marine spatial planning, and ecological management and based on the development concepts of ecological development, sustainable development, pattern optimization, and security development. The research is based on multilevel, high-resolution, and high-precision status perception systems and data analysis technology, comprehensive evaluation technology of coastal tidal mudflat ecological development, analysis technology of the coastal tidal mudflat development process and mechanism, and optimization technology of the spatial layout of coastal tidal mudflat ecological development and utilization. Subsequently, through the setting and implementation of relevant data standards, standardization, business processes, and analysis models, the system implementation and business integration of the

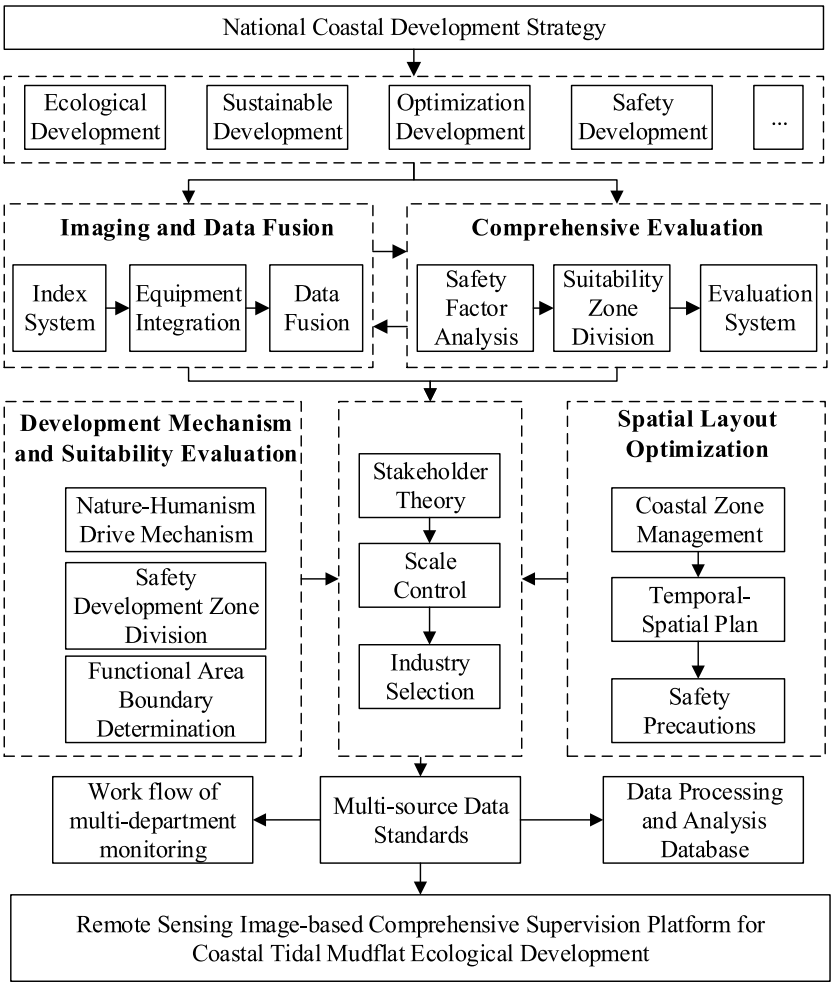

FIGURE 1 | Research scheme of ecological development and monitoring detection of coastal tidal mudflat.

aforementioned key technologies were carried out, and a comprehensive monitoring detection platform for the ecological development of coastal tidal mudflat was finally formed.

2.2 Platform Architecture and Functional System

Based on the existing research results, we summarized the general ideas for establishing a comprehensive coordination and comprehensive monitoring detection platform for the ecological development of coastal tidal mudflat. To be specific, it is the path of differentiation, evolution, and reconfiguration of the goal of comprehensive monitoring detection. This process is similar to the continuous division of human embryonic cells in the process of gestation, differentiation into skin, viscera, bone, hair, and other different directions of evolution of the basic cells, each of the basic cells independently evolving into a fully functional organ, and finally all organs work together to form a complete human body. Concretely, the differentiation refers to the process of decomposing various business units of comprehensive supervision based on the boundaries of various business objectives involved in comprehensive supervision; evolution refers to the process of further development and strengthening of the functions of individual business units; and reconstruction refers to the process of coordination and integration of fully evolved business units on a unified platform. Among them, the evolution process of larger and more complex business units can be divided into the process of differentiation, evolution, and reconstruction. Through the nesting and iteration of this process, the construction of a comprehensive regulatory platform can be realized through the coupling of layers of underlying technologies. Underlying technologies with repeatability or generality need to be implemented only once, while functional implementations with expertise can evolve independently. Compared with the construction of the traditional monitoring detection platform for individual businesses, the functions and objectives of a specific business can be realized in the comprehensive monitoring detection platform, but its realization process relies on the power of the

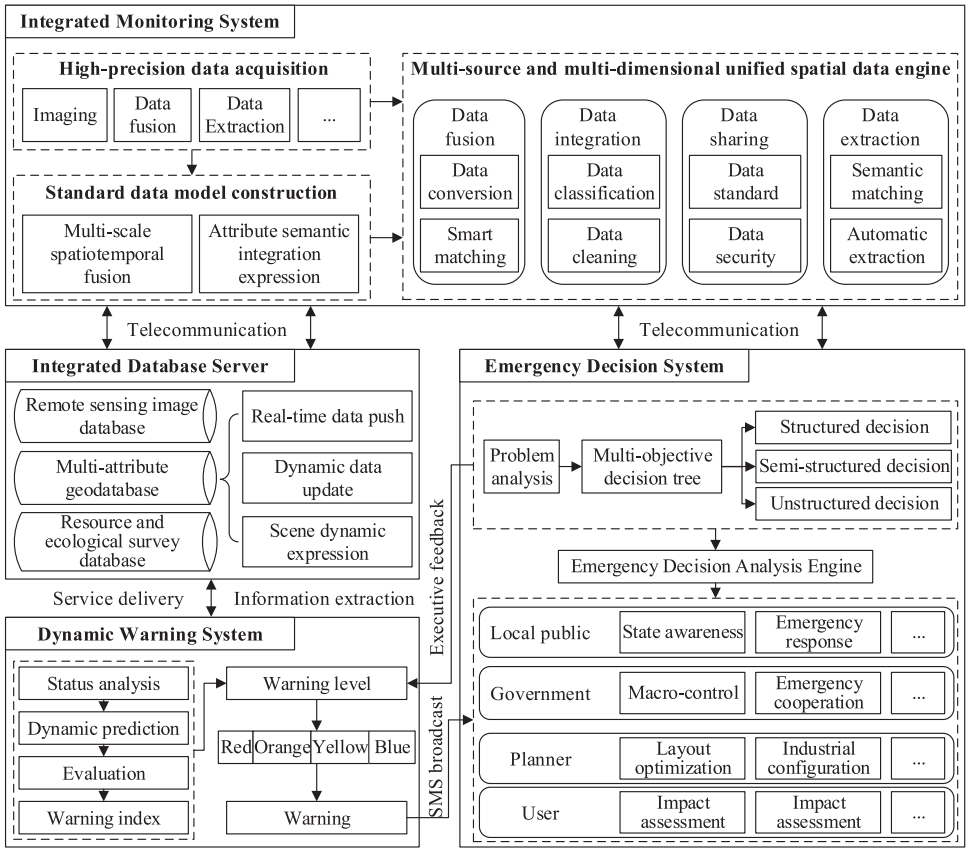

FIGURE 2 | Functional system of the dynamic monitoring and warning platform for ecological development, monitoring and detection of coastal tidal mudflat.

comprehensive monitoring detection platform to mobilize other functions. Overall, there are two advantages to this approach. First, the implementation of a single business can focus more on its essence, eliminating the construction of functions supporting its operation. Second, each business can be made compatible with the platform and presented in a unified, comprehensive, and coordinated way.

According to the dynamic monitoring detection requirements of the ecological development process and business of coastal tidal mudflat, a dynamic monitoring and early warning platform system of coastal tidal mudflat development was constructed based on service-oriented architecture (SOA) and component-based development ideas (**Figure 2**). Due to the problems of data decentralization and ineffective cooperative supervision in multi-department management before, during, and after the ecological development of coastal tidal mudflat, it is necessary to base on the idea of "one map" for coastal areas. Based on the business flow model, module development, and network development technology, the comprehensive regulatory business process of coastal resource security-influencing factor (natural disasters, shoreline changes, and reclamation project security) supervision, development process monitoring and detection, business monitoring and detection, and forecast and early warning were studied. Then, a comprehensive monitoring detection platform for information visualization of coastal beach ecological development was developed, which integrates on-site monitoring, comprehensive information management, decision and early warning supports, and network sharing for coastal beach ecological development. In addition, we have also realized the network space information dissemination and formed a visual and comprehensive monitoring detection system that integrates on-site monitoring, integrated information management, decision and early warning supports, and network sharing. We finally achieved the goal of fast, real-time, and dynamic supervision of "painting by drawing pipe and dynamic monitoring."

2.3 Key Technologies

The comprehensive monitoring detection platform is composed of three subsystems: stereo monitoring system, dynamic early warning system, and emergency decision system, which has the functions of dynamic updating, monitoring detection and management, searching and querying, statistical analysis, online publishing, and system maintenance of the supervision data of coastal shoal development and construction projects. In view of the basic concept and application requirements of comprehensive supervision for ecological development of

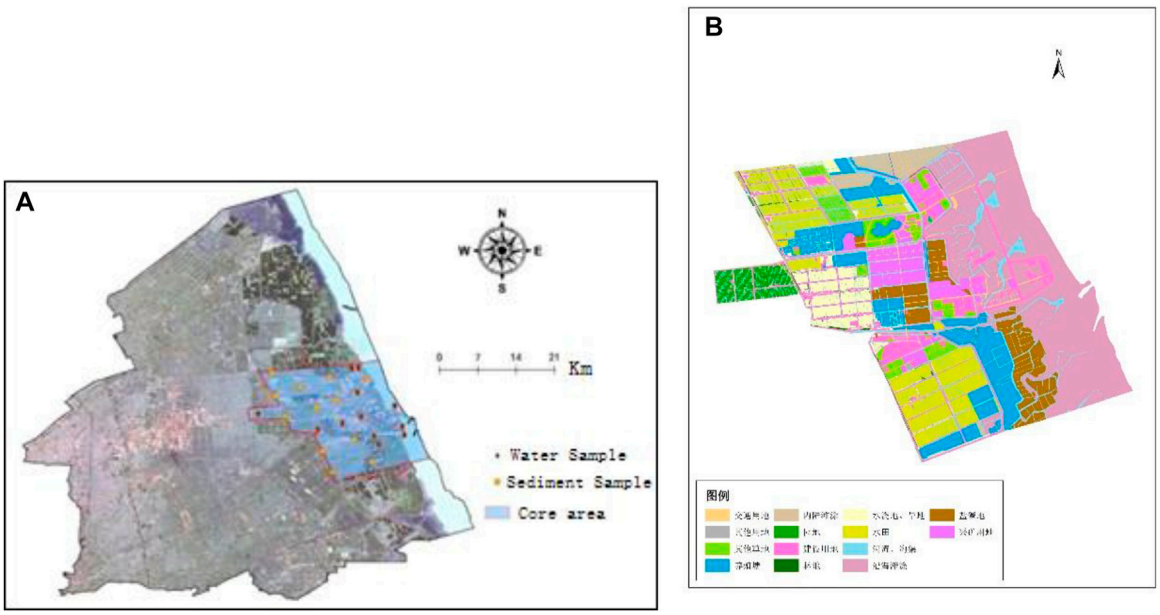

FIGURE 3 | Location of the evaluation area **(A)** and land use type **(B)**.

coastal tidal mudflat, some breakthroughs should be made in the following key technologies.

2.3.1 Multisource Data Adaptive Integration and Data Push Technology

The ecological development of coastal tidal mudflat is characterized by numerous influencing factors and the obvious coupling effect of multiple layers. In recent years, it has developed and formed a high-precision continuous observation system, which covers the ecological platform, airborne remote sensing, medium- and short-range ground wave radar observation, offshore array buoy observation, and underwater monitoring and observation. How to carry out unified management of the previous data and push the data for business process has become the primary problem that must be solved for the ecological supervision of coastal tidal mudflat. Therefore, a process of adaptive integration and data push of multisource data for coastal tidal flat was constructed, which is based on the geometric algebra spatiotemporal unified data model, and a multisource data integration and transformation engine was built by using GDAL/OGR. Based on the business process of ecological supervision of coastal tidal mudflat, the monitoring data are classified and supervised. By formulating unified data metadata standards for the coordinate system, data format, attribute semantics, and other information, a multisource, multi-scale, and heterogeneous data fusion technology based on data transformation and intelligent matching was constructed. Aiming at the problems of coupling between the existing geographical analysis model, the multisource complex geological data, and the complexity of the processing process of the model operation data, an automatic processing and pushing method of multisource geographical data based on model-demand template

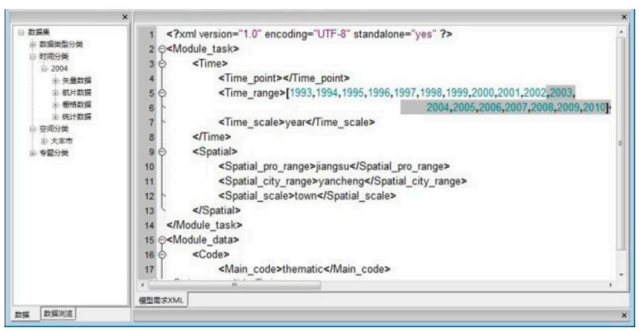

FIGURE 4 | Template of models and data manipulating flows.

matching was constructed. In addition, we used the data push template to achieve automatic matching and extraction of observation data and the analysis model. According to different needs, two different ways of service transmission and information extraction were adopted to realize the real-time push of monitoring data of natural disasters and reclamation projects, as well as the dynamic update of relevant data of the ecological environment, development, and utilization.

2.3.2 Ecological Security Evaluation and the Sensitivity Factor Analysis Model

A coastal zone is an area with strong human-land interaction, which is strongly influenced by natural and human activities. Therefore, ecological security assessment should take into account all kinds of resources and ecosystems in the ecological environment to comprehensively evaluate and quantitatively express the regional ecological security state.

FIGURE 5 | Results of the automatic push.

Remote sensing data have the advantages of rapid acquisition, wide range of searchable, sustainable updates, and so on. The collection of remote sensing data and other data to build a comprehensive ecological security assessment system can enhance the integrity, comprehensiveness, and accuracy of ecological security assessment. Based on the characteristics of ecology, environment, security and high-precision remote sensing data, the ecological security evaluation index system of comprehensive development needs to be established. The relevant indicators cover the resources combination, natural conditions, geological conditions, socio-economic, technology and other background elements, associated with the interaction of the development and utilization of the key resources. On this basis, we integrated a variety of comprehensive evaluation models, such as TOPSIS, gray system, fuzzy, and situation decision, to realize the impact evaluation of land carrying capacity, ecological carrying capacity, land resource quantity, land resource quality, land environmental pressure, and other factors under different development strategies. Then, the economic benefits and environmental effects of development and utilization were evaluated by combining the pressure–state–response, DEA, and econometric models. This study analyzes the ecological and environmental impact caused by the development strategy of coastal areas, its strategic space, and the choice of leading industries.

2.3.3 Process Mechanism and Pattern Optimization Model

The long-term goal of the ecological development of coastal tidal mudflat is to realize the optimal allocation of resources under the constraints of comprehensive factors such as ecology, environment, and security. By evaluating the suitability of different regional card methods, different development modes were extracted, so as to provide technical guarantee for the sustainable development of social economy in coastal tidal mudflat. In this system, the pattern optimization and emergency decision of coastal tidal mudflat were built on the basis of the analysis of the development process and mechanism of coastal tidal mudflat. Starting from the process of planning, development, protection, and monitoring, we established an identification and sensitivity analysis model of influence factors. In addition, we used spatial statistics, spatiotemporal Markov chain, and multivariate stochastic simulation methods to reveal the coupling mechanism of processes related to the ecological environment and human survival and safety, and to explore the overall pattern, structure, and change the driving forces on coastal tidal flat. On this basis, decision tree, situation model, projection pursuit regression, and other methods were introduced to express and describe the spatial and temporal continuous process of coastal tidal flat development, to study its natural evolution process, human activity-driven evolution law, and internal response mechanism, and to analyze the process and evolution mechanism of coastal tidal flat development. Based on the analysis of land use change pattern by using the CLUE-S model, the key areas, ecological corridors, and ecological nodes of landscape ecological function were identified and evaluated. Based on the multi-objective optimization approach, resistance surface, and cumulative cost path analysis method, the ecological circulation index (F), overall connectivity index (IIC), and possible

TABLE 1 | Index system of comprehensive assessment of ecological security.

Target layer	Criterion layer	Index layer	Index quantification	Index value function
Ecological security impact	Population and economic pressure	Population	Population base	Negative
		Economic	GDP	Negative
	Ecological and environmental status	Tidal flat changes	Tidal flat buffer	Negative
		Soil type	AHP method	Negative
		Sediment heavy metals	Cr	Negative
			Cu	
			Hg	
		Water heavy metal	As	Negative
			Cr	
			Cu	
		Ecosystem	NDVI	Negative
			GVI	
			WI	
	Human–environment interaction	Road	Road buffer	Positive
		Industry area	Industry area buffer	Negative
		Reclamation	Reclamation buffer	Negative

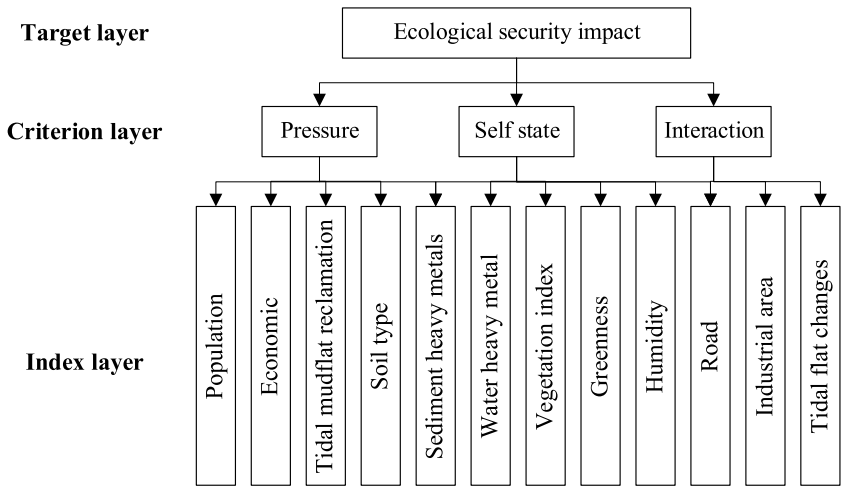

FIGURE 6 | AHP weight structure hierarchy.

connectivity index (PC) were used to analyze the impact and optimization effect of landscape connectivity, and then the spatial allocation was optimized.

2.3.4 Emergency Decision-Making and Dynamic Early Warning Model

The development of coastal tidal flat needs to deal with the possible problems such as ecological security and environmental events at the same time, and it needs to study the early warning and decision-making mechanism for different user groups under different industrial structure types, different environmental background types, different land use types, and different development modes. For this reason, we analyzed the economic, social, ecological, resource, environmental, and human factors in the process of ecological development of coastal tidal mudflat, and established a data set of influencing factors for development security. On the basis of studying the process rules of shoreline evolution, ecological evolution, reclamation engineering construction, industrial structure allocation, spatial pattern evolution, and the disaster response process and damage degree of large-scale, long-period, and strong disturbance coastal shoreline development, combined with the econometric model and prediction method, the future development trend of shoreline development was predicted. Based on the ecology, resource, environment, economy, and society of the index system of comprehensive assessment of risk, development of the ecological development standard of disaster-warning hierarchy and clear warning category of different stakeholders evaluates the warning issued strong pertinence, storm surge, and reclamation to the danger of short-term destructive disasters, such as spontaneous preparation of an emergency plan. According to the early warning needs of different user groups, the emergency decision model analysis engine will act on structured, unstructured, and semi-structured decision services based on the intelligent network.

TABLE 2 | Result of the weight index.

Index	PCA	AHP	Comprehensive
	Weight	Weight	Weight
Soil type	0.060141	0.045153	0.051148
Reclamation	0.060142	0.065747	0.063505
Tidal flat changes	0.060141	0.042316	0.049446
Water Cu	0.060141	0.020339	0.03626
Water Cr	0.060141	0.020339	0.03626
Water As	0.060141	0.020339	0.03626
Population	0.071759	0.063133	0.066583
Economic	0.072112	0.07836	0.075861
Road	0.060142	0.080599	0.072416
Industry	0.060142	0.115383	0.093287
Sediment Hg	0.059389	0.01706	0.033992
Sediment Cu	0.060141	0.01706	0.034292
Sediment Cr	0.060141	0.01706	0.034292
WI	0.07096	0.126172	0.104087
NDVI	0.059389	0.186146	0.135443
GVI	0.064978	0.084794	0.076868

3 CASE STUDY

3.1 Overview of the Research Area and Research Data

As a typical region with coastal ecological environment, Dafeng is located in the coastal areas north of the Yangtze River, east of the Jiangsu Province, and adjacent to the Yellow Sea of China in the east. This region has rich ecological resources of tidal flats, with a coastline of 112 km and a total area of 5,400 km^2 of sea area outside the coastline. The existing tidal flats cover a total area of 1,175,900 mu, making it one of the largest cultivated land reserve resources in China. Dafeng is one of the few high-quality ecological areas in eastern China, which has the largest nature reserve for elk and is also a national scenic spot. Dafeng is located in the zone where the ocean and land interact with each other, and the ecological environment composition is complicated. In recent years, the coastal area of Dafeng has developed intensively. A large number of tidal flats have been reclaimed for agricultural land, and harbors and industrial parks have been established along the coast one after another. The coastal ecological environment has greatly changed, and the research and evaluation of ecological security has become extremely urgent. It is worth emphasizing that human activities such as reclamation of tidal flats and construction of industrial zones have the strongest impact on regional ecological environment. Starting from the particularity of typical ecological environment and resources in Dafeng, the core demonstration area of ecological environment was selected as the representative area of coastal ecological security assessment. The selected evaluation area (**Figure 3**) is located in the eastern coast of Dafeng, including agricultural land, tidal flats, factories, roads, residential buildings, and other spatial units. The area is a typical coastal tidal flat with crispy soil and water network, which can well reflect the natural ecological environment and the interaction between the nature and humanity in the coastal area of Dafeng.

The remote sensing data employed in this study are the environmental satellite CCD remote sensing data (HJ1a-CCD2) that includes the following four bands: 0.43 ~ 0.52, 0.52 ~ 0.60, 0.63 ~ 0.69, and 0.76 ~ 0.90 μm. Remote sensing

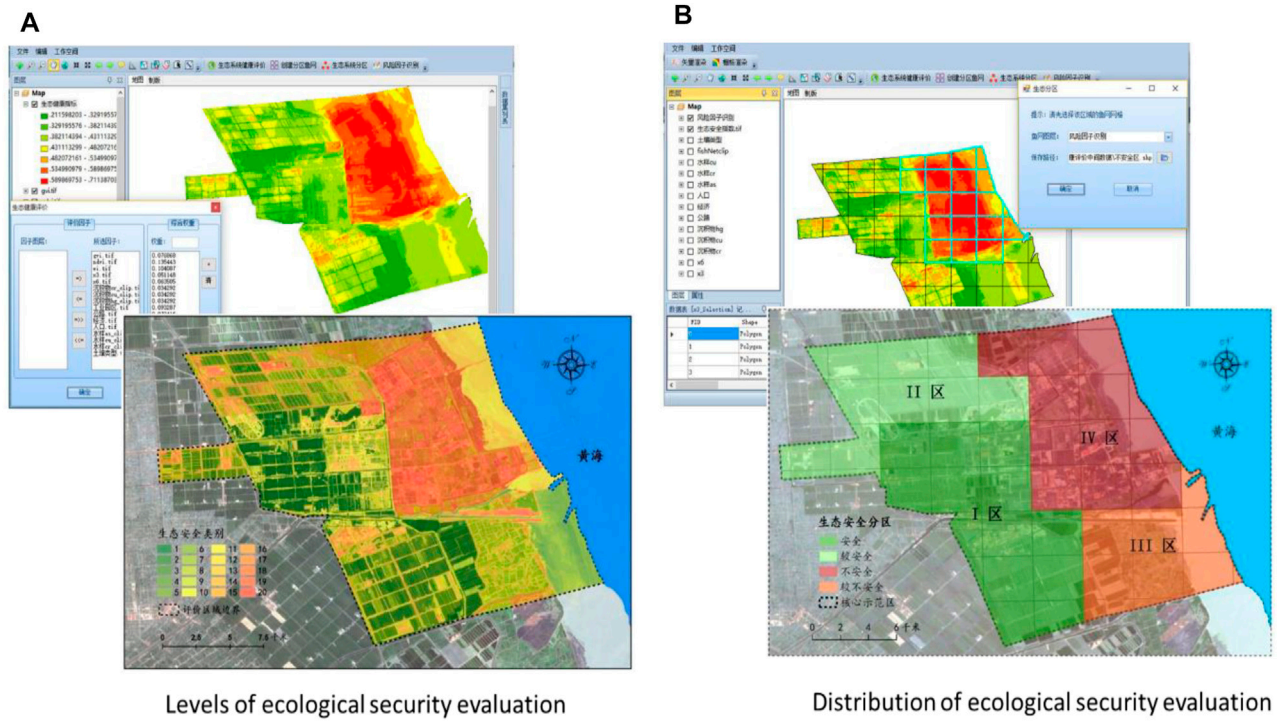

FIGURE 7 | Publication of ecological security evaluation illustrated as levels **(A)**, and levels can be transformed into zones **(B)**.

FIGURE 8 | Results of pattern optimization.

data are used to extract the spatial changes of coastal mudflats and mudflat farmlands and quantify them as evaluation indexes, with a spatial resolution of 30 m. In addition, Landsat ETM + remote sensing data were used to calculate *NDVI*, *GVI*, *WI*, and other evaluation indexes, and the spatial resolution was again 30 m. Furthermore, the measured data of heavy metals from the internal water and surface sediments sampled in the study area in 2014 were used to illustrate the pollution situation of heavy metals and quantify it as an evaluation index. The social and economic data of Dafeng in 2014 were used to reflect the degree of economic activities and the number of population within the research area, which were quantified as an evaluation index. Finally, the soil type distribution data were used to construct the soil type evaluation index.

3.2 Data Integration and Push

The multisource geographic data used in this case include vector data of tidal flat of Dafeng in 2000, and remote sensing image classification and interpretation data of land use, landscape pattern, population, and area of Dafeng in 1995, 2000, and 2005. The main analysis models used in this paper include the following: 1) the evolution of Dafeng tidal flat through the classification and interpretation of remote sensing data, and the changes of tidal flat structure over the years were obtained; 2) analysis of land use ratio per capita in Dafeng; and 3) landscape pattern analysis and the landscape index, using the classified landscape grid data and statistics of each unit value .

The model requirement template is shown clearly in **Figure 4**. The operation templates for different requirements are selected to construct the operation flow in line with the model data requirements by comparing the multisource data in the database (**Figure 4**). Subsequently, the vector data of tidal flat, remote sensing data, and attribute data were extracted from the original database by using the model demand template matching and processed by the model data operation flow (**Figure 5**).

3.3 Safety Evaluation

Based on the remote sensing images and field observation data of the Dafeng region in 2014, we constructed an evaluation index system for coastal ecological security assessment for the purpose of fine-scale ecological security assessment. The objective layer of the evaluation system was the impact degree of ecological security status, and the criterion layer includes three aspects: population and economic pressure, ecological environment, and human–environment interaction. The indicator layer contains 10 types of indicators with a total of 16 specific quantitative indicators (**Table 1**). Specifically, the population, economy, tidal flat change, highway, industrial park, and ecosystem are macroscopic indexes, while the soil type, heavy metals in surface sediments, and heavy metals are microscopic indexes. Based on the criteria of population and economic pressure, ecological environment status, and human–environment interaction, the analytic hierarchy process (AHP) was used to assign the subjective weight of each quantitative index (of which, heavy metals in water body and heavy metals in surface sediments were treated as different categories, and the overall weight was determined before the internal equalization was carried out) (**Figure 6**), and the finally determined AHP weights of each index are presented in **Table 2**. Based on the ecological evaluation model to calculate comprehensive weight of each indicator, it was found that the larger the final index number, the worse the ecological security status of the evaluation unit.

The comprehensive evaluation value of the ecological security degree of the evaluation unit (i.e., the pixel of remote sensing image) can be obtained by multiplying each quantitative index by the corresponding comprehensive weight after standardization. Considering that the difference between the evaluation value and the adjacent evaluation value is small, this article classifies the comprehensive evaluation value. We classified the evaluation value by a quantile based on R, set the quantile to 20, and assigned values from 1 to 20 to each category in order from the smallest to the largest, that is, the evaluation results were divided into 20 categories. The larger the assigned number, the worse the ecological security status of the evaluation unit (**Figure 7**).

3.4 Optimization of the Ecological Development Pattern

First, according to the ecological type and ecological value per unit area of the demonstration area, the Costanza model was selected to evaluate the ecosystem service value. Second, the patch density (PD), average patch area (AREA_MN), average shape index (SHAPE_MN), average proximity index (PROX_MN), connectivity degree (CONNECT), and so on were selected to dig deeply the relationship among various landscape types, and

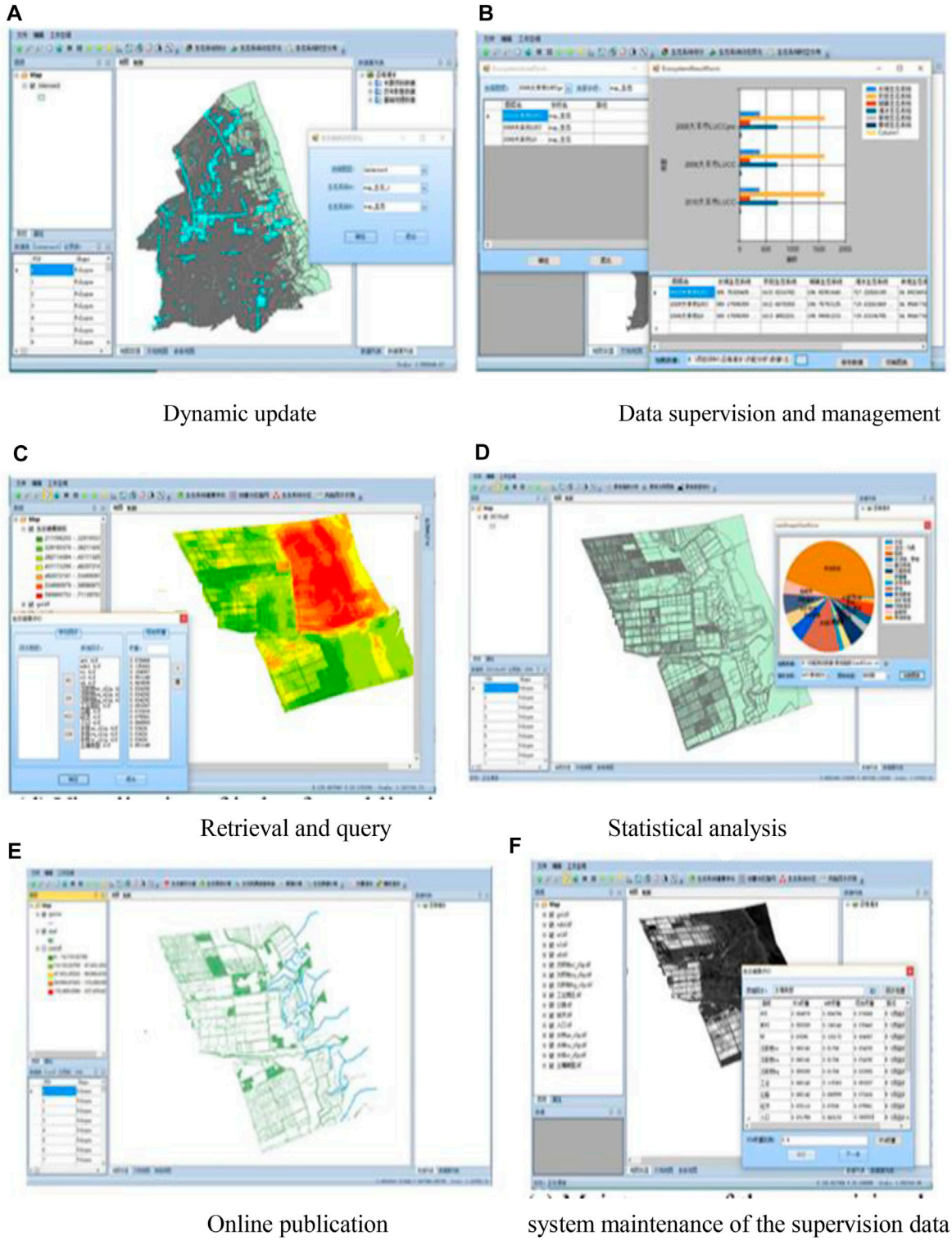

FIGURE 9 | Demonstration of dynamical monitoring detection platform. **(A)** Dynamic update, **(B)** data supervision and management, **(C)** retrieval and query, **(D)** statistical analysis, **(E)** online publication, and **(F)** system maintenance of the supervision data.

the spatial distribution of ecological function value of the demonstration area was consequently obtained. Furthermore, natural landscape types such as water and woodland with high service value per unit area and patches with an area greater than 40 hm^2 were selected as the ecological source areas. Ecological source and cost (ecological resistance, determined by the value proportion of ecosystem service functions) were determined to obtain the cumulative cost distance surface of landscape ecological functions. Finally, based on the above results, the adjacent areas of mining, construction, and traffic lands with high ecological resistance were allocated with grassland, garden land, woodland, and other landscape types. According to the characteristics of the soil around the aquaculture ponds, low cost and high efficiency salt-tolerant plants and protective woodland were allocated in the adjacent coastal tidal mudflat to prevent pollutants from discharging into the coastal tidal mudflat. In order to optimize the allocation of ecological resources in the whole study area, woodland and grassland should be allocated around the farmland, and salt-tolerant cash crops such as sea-buckthorn, Chinese wolfberry and jujube should be selected for ecological desalination (**Figure 8**).

3.5 Comprehensive Monitoring and Prescient Warning

The comprehensive monitoring detection platform is composed of three subsystems: monitoring, dynamic early warning, and emergency decision-making systems. It has the functions of dynamic update, supervision and management, retrieval and query, statistical analysis, online publication, and system maintenance of the supervision data of coastal tidal flat development and construction projects (**Figure 9**). Based on the idea of "one map" for coastal areas, a comprehensive supervision business flow coordinated by multiple departments was constructed, and a data model of ecological management and protection business flow serving coastal tidal mudflat was constructed using unified modeling language (UML). We integrated techniques to identify the relationship between changes in geomorphological, hydrological, and soil factors and the evolution of natural and artificial vegetation. According to different types of tidal flats wetland landscape reconstruction, land management, biodiversity maintenance and plant community building technology, as well as community stability and long-term management, the platform support series of functions: data acquisition and integration, dynamic update, query, statistical analysis, and publication, etc.

4 DISCUSSION

Due to the adoption of the adaptive integration technology of multisource data and based on the unified metadata standard, the comprehensive monitoring detection platform has already transformed, matched, and fused the multisource, multi-scale, and heterogeneous data in the data I/O and integration stage, which greatly shortened the data processing cycle. Through the adaptive data integration, the process of compatibility between old and new data was greatly shortened, so as to ensure that the updated data play a role more quickly and in real time. Meanwhile, the form of terminal data presented to users through data push technology has been combined with the timeliness of new data. Moreover, data push technology can not only be implemented according to the needs of the analysis model but also can be flexibly mastered according to the specific needs of the business so as to ensure the timeliness of the actual use of data at the client end. Application of the platform can help government departments to supervise and control the coastal ecosystem and analyze changes comprehensively.

Supported by data integration and push technology, the calculation and result presentation of the analysis model for monitoring, evaluation, optimized configuration, and early warning were all based on real-time updated data, and the presented analysis results are constantly changing over time. The comprehensive monitoring detection platform realized the integration of dynamic data update and supervision business process. The use of multiple sources of data in the process of supervision, coordination of various interests, ecological, environmental, land, social, and economic aspects was achieved.

On the basis of supporting data, we realize data collection and management and finally achieve the purpose of coordination and intelligent monitoring detection. The whole operation process from the bottom to the top is inter-associative. At the same time, due to the diversity of data and the multifaceted nature of the evaluation process, there are still some uncertainties in the analysis results in the process of mechanism analysis and comprehensive evaluation, such as the selection of some thresholds.

5 CONCLUSION

Integrated coastal zone management has always been one of the most important national development plans, but there are still issues that neglect the internal systematic research in various fields and neglect the multidisciplinary comprehensive research at the national macro level. The system structure and key technologies still need to be further improved and breakthrough. The ecological development and management of coastal tidal mudflat should be systematic and can be comprehensively improved from the perspectives of multidisciplinary comprehensive application, comprehensive data integration, model integration and sharing, and multifactor coupling analysis so as to establish relevant monitoring and supervision mechanisms and realize the overall improvement of the coastal zone management level. The construction of a comprehensive monitoring and detection platform for the ecological development of coastal tidal mudflat can provide support for the study of the development process of coastal tidal mudflat and supervision of human activities.

This article discusses the remote sensing image-based framework, functions, implementation ideas, and methods of the dynamic monitoring and analysis system, which can provide reference for the further development in the dynamic management of the overall coordination and comprehensive regulation of the coastal zone. The future highlights are as follows: 1) carrying out the research on the multidimensional unified data model and data assimilation based on six elements of location, attribute, semantics, geometry, time, and constraint, and solve the integrated dynamic management of multisource, multi-scale, multi-domain, and multi-resolution data;

2) model sharing and the integration mechanism oriented to resources, pattern structure, and process events support the construction of model coupling analysis platform integrating multidisciplines and multi-fields; and 3) study the warning and control mechanism of different levels for the public and government affairs, realizing the scientific guidance of social and government behaviors, and the coordination of interest conflicts (Brommer and Bochev-van der Burgh, 2009; Zhao et al., 2009; Li et al., 2011; Liang et al., 2011; Harvey et al., 2012).

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

SX and LY designed the ideas and methods. ZW, HL, and HG classified the data and drew the graphs. SX, LY, and ZY wrote the article with considerable input from ZW, HL, and HG. ZY revised and checked the language of the article.

FUNDING

This study is funded by the National Natural Science Foundation of China (No. 41971404). This work was also supported by the National Natural Science Foundation of China (No. 41801223) and the National Key Research and Development Foundation of China (No. 2017YFE0100700).

REFERENCES

Brommer, M. B., and Bochev-van der Burgh, L. M. (2009). Sustainable Coastal Zone Management: a Concept for Forecasting Long-Term and Large-Scale Coastal Evolution. *J. Coastal Res.* 251 (1), 181–188. doi:10.2112/07-0909.1

Chen, J., Guan, W., and Sun, Y. (2009). Research on Dynamic Monitoring System for Land Use with 3S Technology. *Geomatics World* 7 (4), 60–65. (in Chinese with English abstract). doi:10.3969/j.issn.1672-1586.2009.04.012

Christensen, S. M., Tarp, P., and Hjortsø, C. N. (2008). Mangrove forest Management Planning in Coastal Buffer and Conservation Zones, Vietnam: a Multimethodological Approach Incorporating Multiple Stakeholders. *Ocean Coastal Manage.* 51 (10), 712–726. doi:10.1016/j.ocecoaman.2008.06.014

Dauvin, J. C. (2008). The Main Characteristics, Problems, and Prospects for Western European Coastal Seas. *Mar. Pollut. Bull.* 57 (1-5), 22–40. doi:10.1016/j.marpolbul.2007.10.016

Harvey, N., Clarke, B., and Nursey-bray, M. (2012). Australian Coastal Management and Climate Change. *Geographical Res.* 50 (4), 356–367. doi:10.1111/j.1745-5871.2011.00734.x

Henocque, Y. (2003). Development of Process Indicators for Coastal Zone Management Assessment in France. *Ocean Coastal Manage.* 46 (3-4), 363–379. doi:10.1016/s0964-5691(03)00013-9

Hou, Y., Chen, X., and Li, Y. (2005). Application of 3s Technologies in Integrated Coastal Zone Management. *Hydrographic Surv. Charting* 25 (3), 24–27. (in Chinese with English abstract). doi:10.3969/j.issn.1671-3044.2005.03.007

Hua, Z., Geng, Y., and Gu, L. (2012). Research Progress on Environmental Impact and Ecological Effect of Tidal Flat Reclamation. *J. Econ. Water Resour.* 30 (3), 66–69. (in Chinese). doi:10.3969/j.issn.1003-9511.2012.03.017

Lee, P. P. V. (1999). Providing Public Access to the Shore: The Role of Coastal Zone Management Programs. *Coastal Manage.* 27 (2-3), 219–237. doi:10.1080/089207599263848

Li, S., Huang, J., Shao, X., and Zhang, Y. (2011). Regional Ecological Compensation and Land Ecological Security. *China Land Sci.* 25 (5), 39–41. (in Chinese with English abstract). doi:10.3969/j.issn.1001-8158.2011.05.008

Liang, J., Song, P., Cui, T., Ren, G., and Zhao, W. (2011). Design and Implementation of HY-1 B CZI Coastal Zone Dynamic Monitoring System. *Mar. Sci. Bull.* 30 (1), 94–99. (in Chinese with English abstract). doi:10.3969/j.issn.1001-6392.2011.01.015

Liu, P., Zhang, H., and Eom, K. B. (2017). Active Deep Learning for Classification of Hyperspectral Images. *IEEE J. Sel. Top. Appl. Earth Observations Remote Sensing* 10 (2), 712–724. doi:10.1109/jstars.2016.2598859

Long, X.-h., Liu, L.-p., Shao, T.-y., Shao, H.-b., and Liu, Z.-p. (2016). Developing and Sustainably Utilize the Coastal Mudflat Areas in China. *Sci. total Environ.* 569-570, 1077–1086. doi:10.1016/j.scitotenv.2016.06.170

Paoli, C., Vassallo, P., and Fabiano, M. (2008). An Emergy Approach for the Assessment of Sustainability of Small Marinas. *Ecol. Eng.* 33 (2), 167–178. doi:10.1016/j.ecoleng.2008.02.009

Pickaver, A. H., Gilbert, C., and Breton, F. (2004). An Indicator Set to Measure the Progress in the Implementation of Integrated Coastal Zone Management in Europe. *Ocean Coastal Manage.* 47 (9-10), 449–462. doi:10.1016/j.ocecoaman.2004.06.001

Turner, R. K. (2000). Integrating Natural and Socio-Economic Science in Coastal Management. *J. Mar. Syst.* 25 (3-4), 447–460. doi:10.1016/s0924-7963(00)00033-6

Wang, G. (2013). Conditions and Innovation of Environmental Management on Tidal Flat in Coastal China. *China Land Sci.* 27 (4), 46–52. (in Chinese with English abstract). doi:10.3969/j.issn.1001-8158.2013.04.008

Wang, J., and Zou, X. (2009). Study on Environment Early Warning System of Coastal Zone in Yancheng. *Adm. Tech. Environ. Monit.* 21 (3), 19–23. (in Chinese with English abstract). doi:10.3969/j.issn.1006-2009.2009.03.006

Wang, Y., and Li, R. (2006). Landscape Ecological Planning in Southeast Coastal Zone in Hainan Island. *Areal Res. Dev.* 25 (5), 103–107. (in Chinese with English abstract). doi:10.3969/j.issn.1003-2363.2006.05.023

Xu, S., Du, R., and Wang, X. (2010).Research of post-approval Supervision System Based on GIS, Proceeding of the 2010 18th International Conference on Geoinformatics, June 2010, Beijing, China. IEEE, 1–5. doi:10.1109/geoinformatics.2010.5568079

Zhang, L. (1999). Comprehensive Management System and Operation Mechanism of Regional Coastal Zone — Taking Yuhuan County as an Example. *Mar. Sci. Bull.* 18 (3), 68–76. (in Chinese with English abstract).

Zhang, A. (2010). Design and Research of Land Resources" A Map" Data Management and Service Platform. *Mod. Surv. Mapp.* 33 (4), 41–42. (in Chinese with English abstract). doi:10.3969/j.issn.1672-4097.2010.04.014

Zhang, L., Liu, P., Zhao, L., Wang, G., Zhang, W., and Liu, J. (2021). Air Quality Predictions with a Semi-supervised Bidirectional LSTM Neural Network. *Atmos. Pollut. Res.* 12 (1), 328–339. doi:10.1016/j.apr.2020.09.003

Zhao, Z., Hua, Y., Wu, C., and Zhao, J. (2009). Dynamic Monitoring Techniques for the Implementation of Township Land Use Plan: A Case Study of Tangxi Town, Yuhang District, Hangzhou. *China Land Sci.* 23 (7), 43–48. (in Chinese with English abstract).

Zhu, S., and Huang, Y. (2010). Research on Regional Land Information Integration Model. *Geo-information Sci.* 12 (4), 568–573. (in Chinese with English abstract). doi:10.3724/sp.j.1047.2010.00568

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.

frontiers in Environmental Science

ORIGINAL RESEARCH
published: 28 February 2022
doi: 10.3389/fenvs.2022.799250

Waterline Extraction for Artificial Coast With Vision Transformers

Le Yang, Xing Wang and Jingsheng Zhai*

School of Marine Science and Technology, Tianjin University, Tianjin, China

Accurate acquisition for the positions of the waterlines plays a critical role in coastline extraction. However, waterline extraction from high-resolution images is a very challenging task because it is easily influenced by the complex background. To fulfill the task, two types of vision transformers, segmentation transformers (SETR) and semantic segmentation transformers (SegFormer), are introduced as an early exploration of the potential of transformers for waterline extraction. To estimate the effects of the two methods, we collect the high-resolution images from the web map services, and the annotations are created manually for training and test. Through extensive experiments, transformer-based approaches achieved state-of-the-art performances for waterline extraction in the artificial coast.

Keywords: coastline extraction, vision transformer, SegFormer, sea–land segmentation, waterline extraction

OPEN ACCESS

Edited by:
Peng Liu,
Institute of Remote Sensing and Digital Earth (CAS), China

Reviewed by:
Zq Gao,
Chinese Academy of Sciences (CAS), China
Min Chen,
Southwest Jiaotong University, China

****Correspondence:***
Xing Wang
xing.wang@tju.edu.cn

Specialty section:
This article was submitted to Environmental Informatics and Remote Sensing, a section of the journal Frontiers in Environmental Science

Received: *21 October 2021*
Accepted: *10 January 2022*
Published: *28 February 2022*

Citation:
Yang L, Wang X and Zhai J (2022) Waterline Extraction for Artificial Coast With Vision Transformers. Front. Environ. Sci. 10:799250. doi: 10.3389/fenvs.2022.799250

INTRODUCTION

A coastline is the boundary between the dry and wet part in the coastal area when the high tide water is in the mean level Toure et al. (2018). The coastline is a critical geographic information source, and it is of great significance to autonomous navigation, coastal resource management, and protection of the environment Liu et al. (2013). Coastline extraction is a very challenging problem because it is obtained from a region not an instantaneous line. The waterline extraction is the precondition for computing the natural coastline, so the waterline extraction is very important and meaningful. The waterline is the instantaneous boundary between the land and sea. It can be extracted from the high-resolution images without other tools. In the artificial coast, the waterline can be considered as the coastline because the waterline is very slightly influenced by the tides.

With the development of satellite remote sensing technology, it supplies tons of high-resolution images of the coastal area, and they can be used for waterline extraction Roelfsema et al. (2013). Besides buying these remote sensing images directly from the remote sensing image providers, users can obtain many satellite map images freely from the web map services. All these data can be used for the waterline extraction.

The waterline extraction methods mainly include threshold segmentation methods, edge-based methods, object-oriented methods, active contour method, conventional machine learning methods, and deep learning methods. The threshold segmentation methods are intuitive methods that set a threshold value according to the image intensity to segment the land and water. Guo et al. (2016) proposed a method that utilized a normalized difference water index to segment water and land. Chen et al. (2019) used the components of the tasseled cap transformation to extract waterline information. Wernette et al. (2016) presented a threshold-based multi-scale relative relief method to extract the barrier island morphology from high-resolution DEM. These methods are handy and effective for the simple image segmentation task. In these methods, threshold selection is the key and difficult problem. In addition, the methods cannot deal with the images with occlusions or a complex background.

The edge-based methods utilize the distinctive edge feature from the abrupt transition. The common methods including Sobel, Roberts (Yang et al., 2018), Laplacian, and Canny operators (Lin et al., 2013; Paravolidakis et al., 2016; Ao et al., 2017; Widyantara et al., 2017; Paravolidakis et al., 2018) can be adopted to extract the waterline. Wang and Liu (2019) proposed a robust ridge-tracing method utilizing the statistical properties of the pixel intensities in the land and sea to detect the boundary. These methods are easy to detect clear, continuous boundaries. However, in the waterline images with a complex background, they are greatly affected by noise. The continuity of the extracted waterline is hardly guaranteed.

Object-oriented methods no longer use the pixel as the basic processing unit; instead, they use an object composed of homogeneous pixels (Gucluer et al., 2010; Rasuly et al., 2010; Bayram et al., 2017). Ge et al. (2014) presented an object-oriented multi-scale segmentation method using interpretation rule sets for automated waterline extraction from remote sensing imagery. Wu et al. (2018) used the object-oriented classification method to extract the waterline from Landsat images of Shenzhen city. These methods use higher level features to classify images, which can reduce the impact of fine texture characteristics on the results of image classification. However, in face of a large amount of information in high-resolution images, object-oriented classification methods may ignore some of the hidden useful information, and it is difficult to achieve the desired classification accuracy.

The active contour methods can achieve better results for remote sensing images of waterlines with simple backgrounds, strong contrast, and continuous boundaries. Cao et al. (2016) proposed a new geometric active contour model for waterline detection from SAR images, which is adaptive to the speckle noises. Fan et al. (2016) proposed a level set approach with a particle swarm optimization algorithm for waterline automatic detection in SAR images. Elkhateeb et al. (2021) adopted a modified Chan–Vese method for sea–land segmentation, which is initiated by a superpixel-based fuzzy c-means automatically. In the study by Modava and Akbarizadeh, (2017), a waterline extraction method–based active contour for SAR images is proposed, in which the initial contour is obtained from a fuzzy clustering with spatial constraints. In the study by Liu C et al., (2016), the waterline is extracted hierarchically by the level set techniques from single-polarization and four-polarization SAR images. Liu et al. (2017) integrated an edge-based and a region-based active contour model in different scales to fulfill the waterline detection from SAR Images. Due to the characteristics of the active contour model method, the application of this method is feasible for waterline images with a simple background, strong contrast, and continuous boundaries. However, the iterative method inevitably produces a large amount of calculation, which restricts its efficiency.

Conventional machine learning methods distill useful information and hidden knowledge based on a variety of data to extract the waterline. Rigos et al. (2016) and Vos et al. (2019) used a shallow neural network to extract the shoreline from satellite images and video images, separately. Sun et al. (2019) built a superpixel-based conditional random field model to segment the sea and land area. Dewi et al. (2016) presented fuzzy c-means methods to detect positions of the coastline and estimate the uncertainty of the coastline change. Cheng et al. (2016) proposed a graph cut method to segment the sea and land, in which the seed points are achieved by a probabilistic support vector machine. Compared with the traditional waterline extraction method, the shallow neural network, clustering analysis technology, fuzzy logic technology, and support vector machine use intelligent means to find out frequent regular things from a large number of data information effectively. These methods can automatically and efficiently extract regular objects. However, for more complex objects in high-resolution images, the extraction accuracy is unsatisfactory. Some other traditional methods, such as the polarization method (Nunziata et al., 2016), wavelet transform method (Toure et al., 2018), region growing method (Liu Z et al., 2016), and decision tree algorithm (Wang et al., 2020) , are all influenced by noise and cannot process the high-resolution images easily.

In these years, deep learning methods have been rapidly developing with the quickly growing performance of computer hardware. Different from traditional machine learning, it can learn the characteristics of the target more accurately. Some convolutional neural network (CNN) methods are naturally introduced in waterline extraction by segmenting the land and sea. In the study by Liu et al., (2019), a simple CNN with multi-scale features and leaky rectified linear unit (leaky-ReLU) activation function is used for waterline extraction. Liu W et al. (2021) proposed an end-to-end lightweight multitask CNN without downsampling to obtain lakes and shorelines from remote sensing images. Shamsolmoali et al. (2019) adopted a residual dense UNet to facilitate the hierarchical features from the original images for sea–land segmentation. Tsekouras et al. (2018) presented a novel Hermite polynomial neural network to detect the shoreline at a reef-fronted beach. Cheng et al. (2017a) proposed a local smooth regularized deep CNN that can obtain the segmentation and edge results of the sea and land simultaneously. Cheng et al. (2017b) employed a multitasking edge–aware CNN for sea–land segmentation and edge detection simultaneously. Cui et al. (2021) presented a scale-adaptive CNN for sea–land segmentation, which fused multiscale information and emphasized the boundaries' features actively. A sea–land segmentation approach utilizing the fast structured edge network and the waterline database was taken from the study by He et al., (2018). A novel UNet-like CNN was proposed for sea–land segmentation, and the network can be deeper, and the convergence can be faster based on local and global information (Li et al., 2018). Erdem et al. (2021) proposed a majority voting method based on different deep learning architectures to obtain shorelines automatically. Lin et al. (2017) presented a multi-scale end-to-end CNN for sea–land segmentation and ship detection, which can increase the receptive field while maintaining fine details. Even though the CNNs have achieved great performances, the limited receptive field affected the performance because of the structure of the CNN.

Transformers, as the most advanced methods in the semantic segmentation, are migrated to compute vision tasks to solve the problem of long-distance dependence by the self-attention

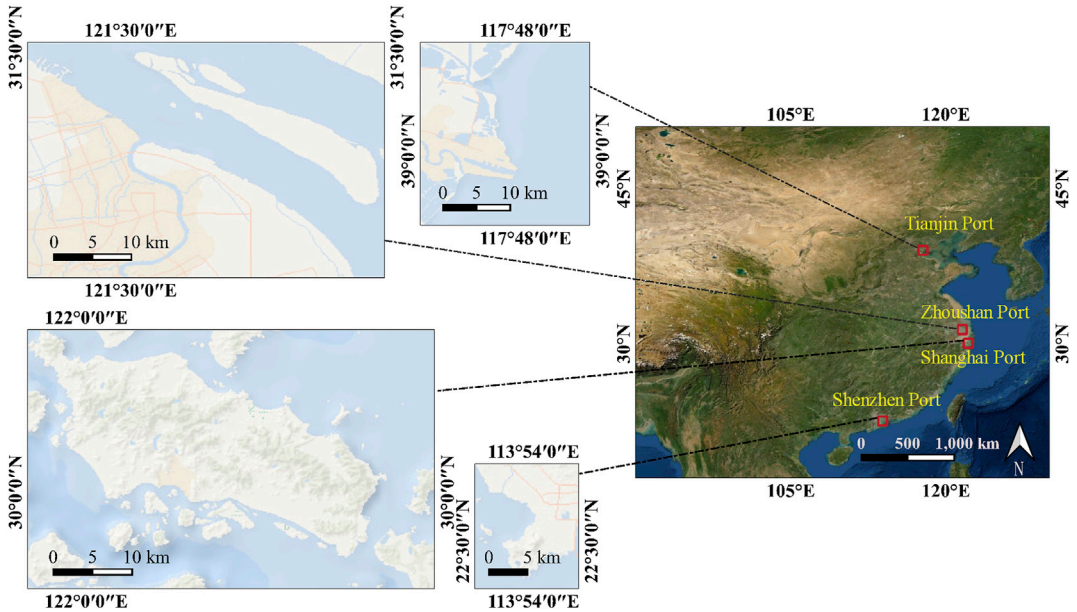

FIGURE 1 | Research area.

mechanism, which is the core of transformers. It determines the global contextual information of each item by capturing its interaction amongst all items. A vision transformer (ViT) is the first work that uses a pure transformer for image classification, which proves that the transformer can achieve the state-of-the-art (Dosovitskiy et al., 2021). It treats each image as a sequence of tokens and then feeds them to multiple transformer layers to make the classification. Subsequently, the dual intent and entity transformer (DeiT) (Touvron et al., 2021) further explores a data-efficient training strategy and a distillation approach for ViTs. The pyramid vision transformer (PVT) is the first work to introduce a pyramid structure in a transformer, demonstrating the potential of a pure transformer backbone compared to CNN counterparts in dense prediction tasks (Wang et al., 2021). After that, methods such as shifted windows (Swin) transformer (Liu Z et al., 2021), convolution transformers (CvT) (Xu et al., 2021), and twin transformer (Chu et al., 2021) enhance the local continuity of features and remove fixed size position embedding to improve the performance of transformers in dense prediction tasks. Segmentation transformers (SETRs) adopt the ViT as a backbone to extract features, achieving impressive performance in segmentation (Zheng et al., 2021). Following it, semantic segmentation transformers (SegFormer) achieved even better results later (Xie et al., 2021).

Therefore, we use the most advanced transformer methods to extract the waterline as an early exploration. This study mainly focuses on the process of extraction of the waterlines for artificial coasts and presents the early research for investigating the potential of transformers in waterline extraction from very high-resolution images.

The rest of the study has the following sections. **Materials and Methods** suggests details about the dataset and methodology. Results reports experimental results with a discussion. Finally, the conclusion section concludes and discusses future research directions.

MATERIALS AND METHODS

Dataset

For this research, we selected Tianjin, Zhoushan, Shanghai, and Shenzhen four ports as research areas, which are shown in **Figure 1**. The waterline images are collected from Mapbox (Mapbox, 2021), Google Maps (Google Maps, 2021), and Bing Maps (Microsoft, 2021) guided by OpenStreetMap (OSM) tiles. The images are in 18 levels in the map, and the initial resolution is 256 × 256. The ground sampling distance (GSD) is about 0.48 m. We combine each neighboring four tiles into a 512 × 512 image. A total of 600 images are chosen as the initial data, and the ground truths of waterlines are created by hand. We also augment it with the random rotation, flip, scale, contrast, brightness, and saturation to 6000 images. Among them, 3600 images are considered as the training set, 1200 images and 1200 images for validation and test, respectively. The images and corresponding annotations are indicated in **Figure 2**. To evaluate the effects of transformers, six CNN segmentation methods are introduced in the experiments.

Methodology

SETR

SETR is an Encoder–Decoder architecture, as seen in **Figure 3**. SETR adopted a high resolution of local features extracted by a CNN and the global information encoded by transformers to segment pixels in an image. Because of quadratic model complexity of the transformer, flattening the whole image as a

FIGURE 2 | Dataset.

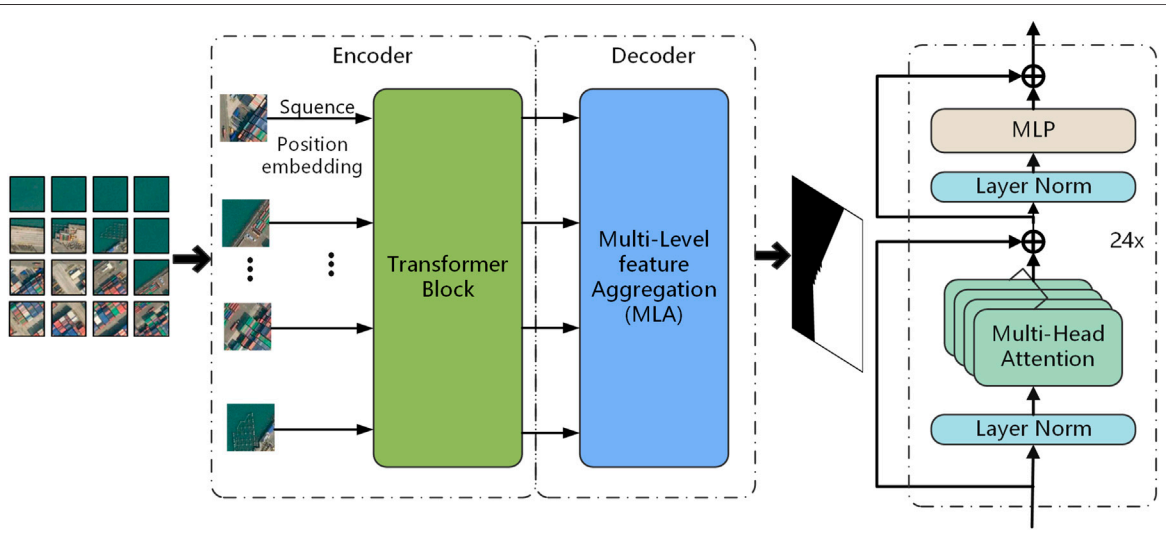

FIGURE 3 | Architecture of the SETR.

sequence makes a huge amount of computation. To speed up, an image is divided into 256 even patches, and then, each patch is flattened into a sequence for input separately.

All the sequences are entered into the pure transformer-based encoder. Therefore, all the transformer layers have a global receptive field, which improve the limited receptive field problem from the CNN. There are 24 layers of transformers in the encoder, in which there are multi-head self-attention (MSA), multilayer perceptron (MLP), and layer normalization blocks residually connected.

The decoder is called the multi-level feature aggregation (MLA). Some feature representations from the transformer layers are first reshaped from 2D to 3D and then aggregated. A 3-layer convolution network downsamples the features at the first and third layer. To enhance the interactions of different levels of features, a top-down aggregation design is introduced. The fused feature is obtained *via* channel-wise concatenation after the third layer. At last, the outputs are upscaled by bilinear operation to the original resolution.

SegFormer

The architecture of SegFormer is depicted in **Figure 4**. The SegFormer consists of two main modules, encoder and decoder. An image as the input is first divided into patches in 4×4. Then, these patches are imported to the hierarchical transformer encoder to obtain multi-level features. These multi-level features are passed to the MLP decoder to predict the segmentation mask at a $H/4 \times W/4 \times N_{cls}$ resolution, where H, W, N_{cls} are the height, width of the image, and the number of categories in the image, respectively.

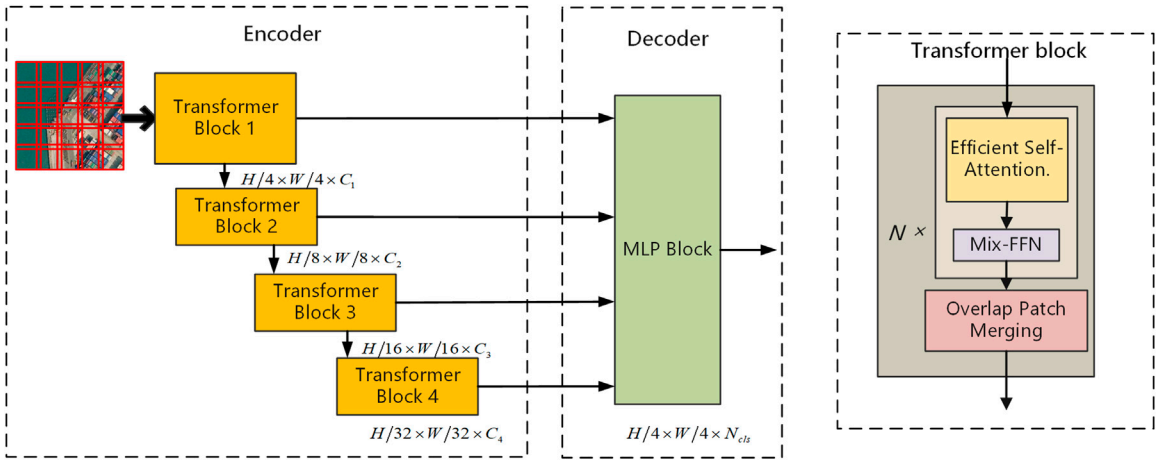

FIGURE 4 | Architecture of the SegFormer.

In the encoder, an input image is one with a resolution of $H \times W \times 3$, and C_i is the channel number in the feature map F_i. A hierarchical feature map F_i with a resolution of $H/2^{i+1} \times W/2^{i+1} \times C_i$ is obtained after each transformer block, where $i \in \{1, 2, 3, 4\}$, and C_{i+1} is larger than C_i.

The transformer consists of efficient self-attention, Mix-FFN, and overlap patch merging blocks. Efficient self-attention improves the computational efficiency of the self-attention. In the original multi-head self-attention process, each of the heads has the same dimension $N \times C$, where $N = H \times W$ is the length of the sequence, and C is the channel number. The self-attention is expressed as follows:

$$Attention(Q, K, V) = \text{Softmax}\left(\frac{QK^{\top}}{\sqrt{d_h}}\right)V. \quad (1)$$

In the equation, d_h is the dimension of the head. The computational complexity of this process is $O(N^2)$. To alleviate it, the sequence K is reduced with a reduction ratio R. It is first reshaped into $N/R \times C \cdot R$ and then simplified by a fully connected layer. Therefore, the new K has dimensions $N/R \times C$. As a result, the complexity of the self-attention mechanism is reduced from $O(N^2)$ to $O(N^2/R)$.

ViT uses positional encoding (PE) to express the location information. It influences the test accuracy when the image resolution is not the same with that in the training because the positional code needs to be interpolated. To address it, Mix-FFN considers the effect of zero padding to leak location information, and a 3×3 convolution is used in the feed-forward network (FFN). Mix-FFN can be formulated as follows:

$$x_{out} = \text{MLP}(\text{GELU}(\text{Conv}_{3\times3}(\text{MLP}(x_{in})))) + x_{in}, \quad (2)$$

where x_{in} is the feature from the self-attention module. Mix-FFN mixes a 3×3 convolution and an MLP into each FFN. The Gaussian error linear unit (GELU) (Hendrycks and Gimpel, 2020) is an activation function. x_{out} is the output of the Mix-FFN.

To preserve the local continuity around those patches, an overlapping patch merging process is used. K is the patch size, S is the stride between two adjacent patches, and P is the padding size. K = 7, S = 4, P= 3, and K = 3, S = 2, P= 1 are set to perform overlapping patch merging to produce features with the same size as the non-overlapping process.

The SegFormer incorporates a lightweight decoder consisting only of MLP layers. The proposed All-MLP decoder consists of four main steps. First, multi-level features from the encoder go through an MLP layer to unify the channel dimension. Then, features are upsampled to 1/4th and concatenated together. Third, an MLP layer is adopted to fuse the concatenated features. Finally, another MLP layer takes the fused feature to predict the segmentation mask.

The Proposed Method

In this section, we demonstrated a method used in the task of waterline extraction. The workflow is shown in **Figure 5**. The transformer first learns the coast features from the training samples. This step is the most time-consuming since most layers of the network are trained in this step. After the learning step, parameters of the model are convergent, and it can infer other new coast images for waterline extraction. Then, a binary mask of the coast is obtained from each input image; the waterline can be extracted from the mask easily. It is worth noting that the contours of the coast at the edges of the image should be excluded because this part is truncated when slicing image tiles.

Metrics

The proposed approaches are evaluated by precision, recall, F1-score, and IoU.

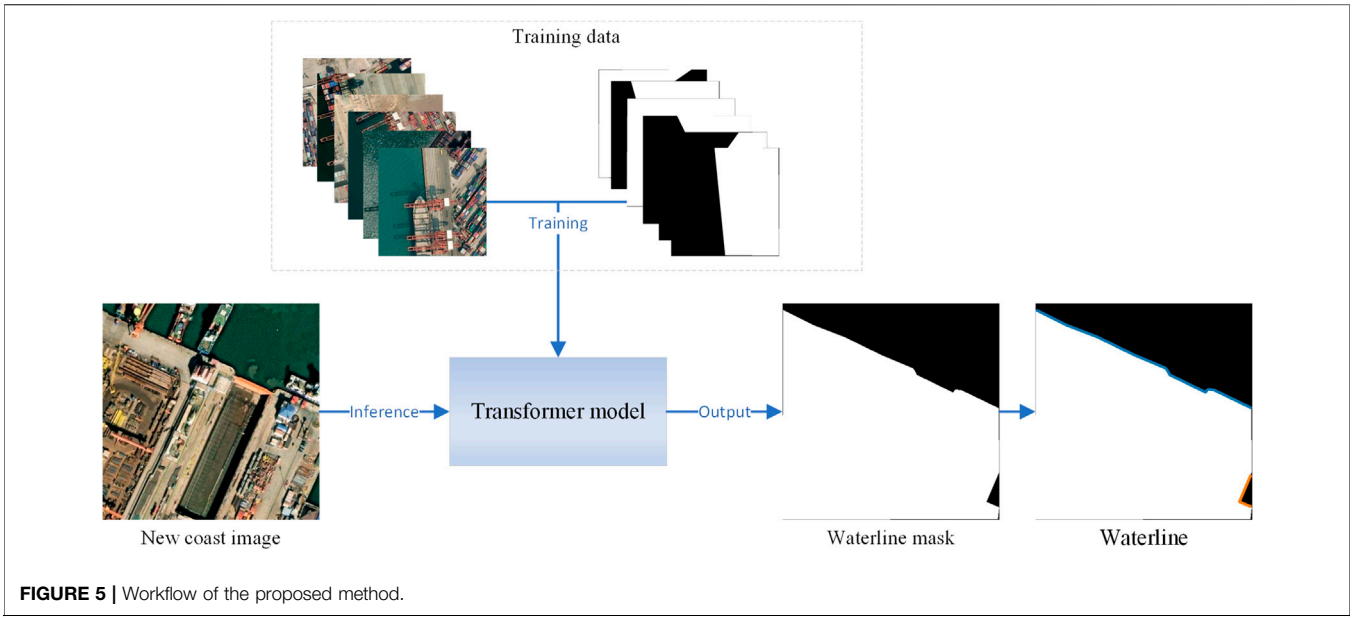

FIGURE 5 | Workflow of the proposed method.

$$P = \frac{T_P}{T_P + F_P}, \tag{3}$$

$$R = \frac{T_P}{T_P + F_N}, \tag{4}$$

$$F_1 = \frac{2 \times P \times R}{P + R}, \tag{5}$$

$$I_U = \frac{T_P}{T_P + F_P + F_N}. \tag{6}$$

Among them, the P and R stand for precision and recall, respectively; the true positive (T_P) stands for the rightly extracted land area; the false positive (F_P) represents the area mistaken as the land; the false negative (F_N) means omitted land pixels. In our study, the reference land images are drawn manually. Precision and recall are contradictory in most cases. To address this, comprehensive metrics F1 (F_1) and IoU (I_U) are employed commonly. Inference time is defined as the average segmentation's time using our test data. The floating-point operations (Flops) represent the computation of the model, and it is a metric for the computational complexity.

RESULTS

Experiment Setting

The proposed transformers were developed under MMsegmentation (MMSegmentation, 2020) by PyTorch (Paszke et al., 2017). Training and testing were performed with eight NVIDIA TITAN Xp GPUs and one NVIDIA TITAN Xp GPU, respectively. In our experimental dataset, there are 3600 images for training, 1200 images for validation, and 1200 images for testing. All the annotations are manually annotated. The resolution of all images is 512 × 512. The SETR uses a learning rate value of 10^{-3}, the number of iterations is 160,000, and the weights are pretrained on ImageNet-21K. The SegFormer was tested using a learning rate value of 10^{-6}, the number of iterations is 40,000, and the weights are pretrained on ImageNet-1K. The other compared methods are all run in 40,000 iterations.

Experimental Results

Qualitative Results

The results of the eight methods are displayed in **Figure 6**. From the results, we can see that PSPNet-UNet, DeepLabV3-UNet, and SETR cannot obtain good results in Image 1, Image 4, and Image 6. A large area in the land is missed, and the fine dock structure is not extracted in all the three methods. For the CNN methods, the methods with the ResNet101 backbone are better than the methods with HRNet, and the methods with the UNet backbone achieve the cheapest results. Only the methods with ResNet101 and HRNet extract the small striped object in Image 6, but no methods can avoid the influence of the ship. In Image 8, only the FCN-ResNet101 and DeepLabV3-ResNet101 gain terrific results. Other CNNs get a lot of false-positive or false-negative parts. For the transformer methods, the SegFormer achieves very nice results in all the images, especially for the fine structures. In contrast, the SETR can also extract the large object effectively in Image 3, but it struggles to the small and thin objects in Image 6 and Image 8. Overall, DeepLabV3-ResNet101, FCN-ResNet101, FPN-ResNet101, and SegFormer are all outstanding, and PSPNet-UNet, DeepLabV3-UNet, and SETR are relatively weak.

We can see in **Figure 7**, the SETR cannot extract the fine objects in Image1, Image 3, and Image 4. The dock and infrastructure are all not complete in the three images. It

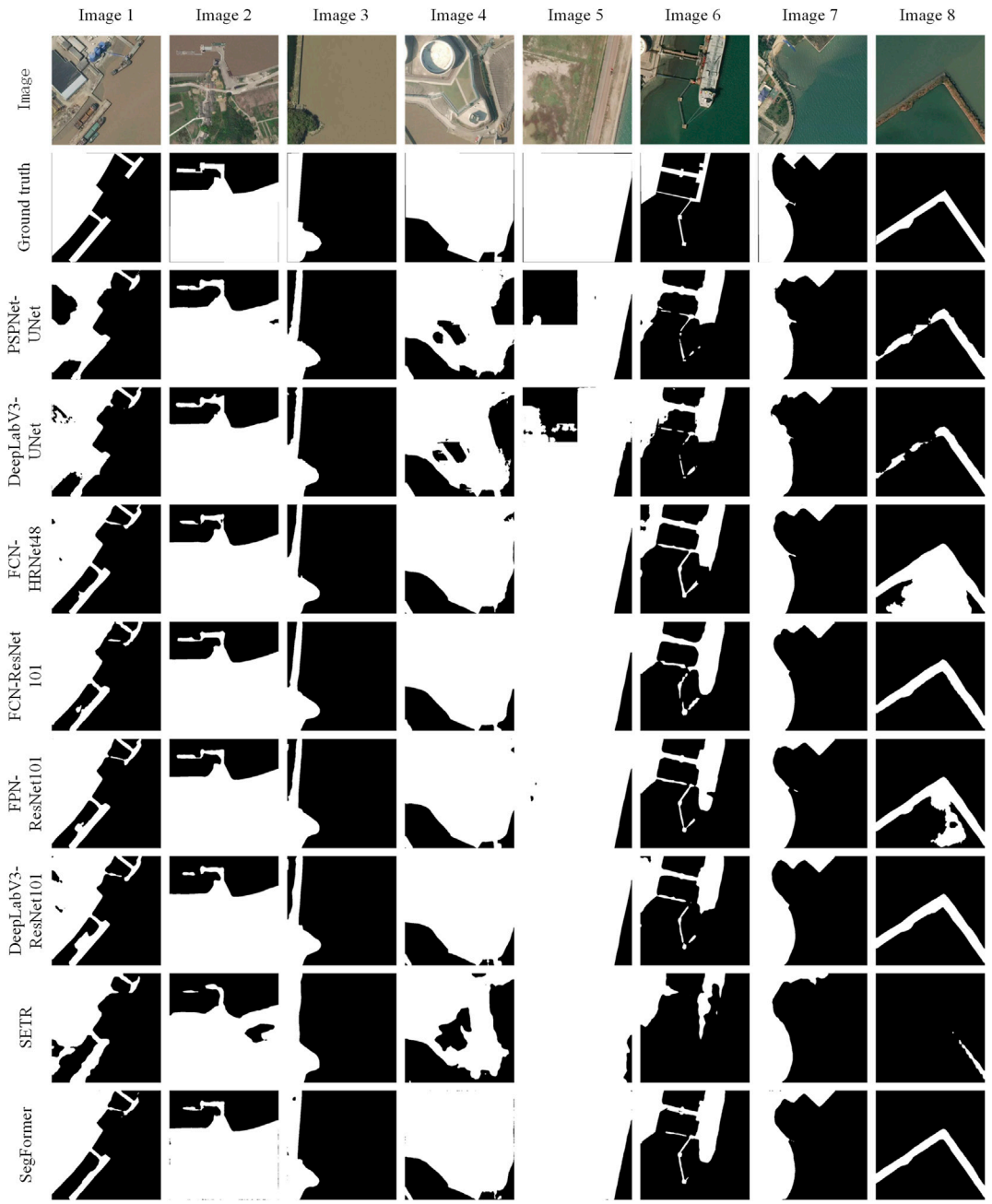

FIGURE 6 | Results of all methods in comparison.

hardly finds the land area near the boundary in Image 2 and Image 3. The connection part in the dock is neglected, and the shape of the harbor is not regular due to the incomplete extraction in Image 2. There are missed land pixels near the frame in Image 3. For large objects, it can perform well, although the edges are not kept fine in Image 1, Image 3, and Image 4.

Figure 8 depicts the results of SegFormer results. We can see that it correctly segments nearly all the pixels. It can even keep the details of objects well, especially in Image 1 and Image 3. The spindly parts in Image 1 and Image 3 are all fine and unmistakable. The integrity and differentiation are impressive. The minor complaints are the small leaks near the edges in Image 1, Image 3, and Image 4 and the small holes in Image 2. The SegFormer can extract the land features so good that the waterline can be extracted completely and accurately.

Quantitative Results

All the experimental methods are reported in **Table 1**. For the CNN methods, the models with ResNet101 achieved best results. Among them, the DeepLabV3 is the best with 0.9056 in precision,

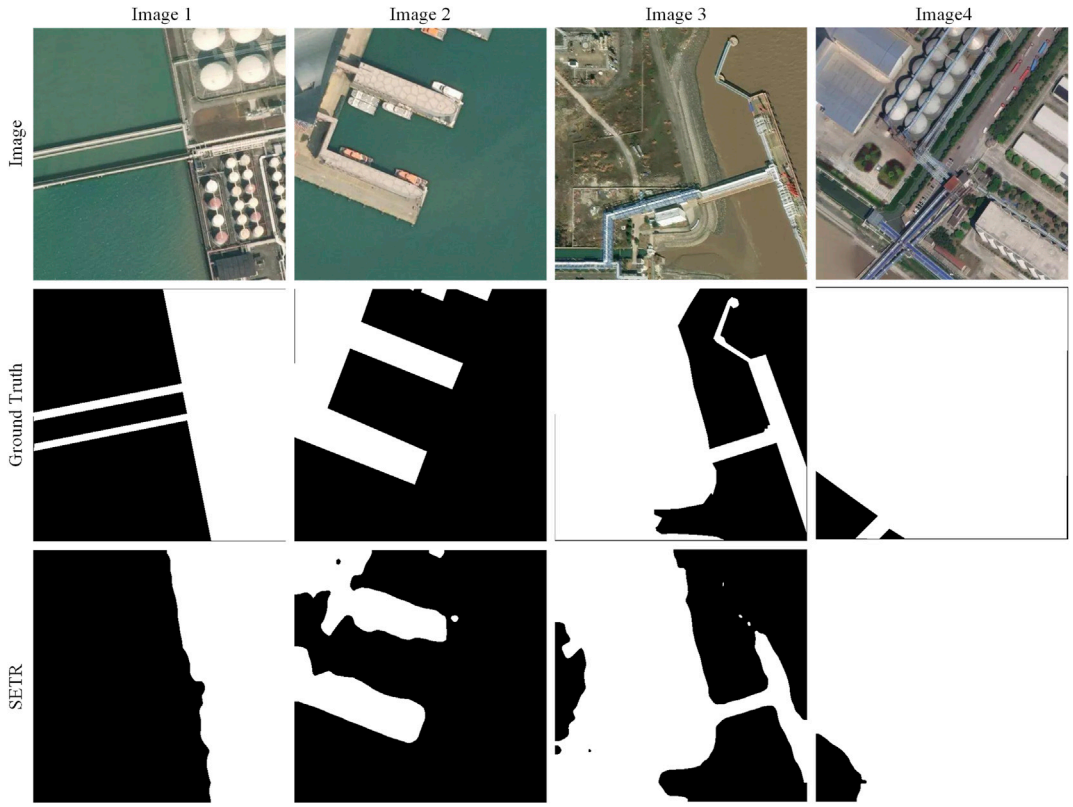

FIGURE 7 | Results of SETR.

0.8814 in recall, 0.8674 in F1, and 0.8169 in IoU. The FCN and FPN with the backbone of ResNet101 also reach or approach 0.9 in precision, 0.88 in recall, 0.86 in F1, and 0.81 in IoU. By comparison, the methods with UNet get lowest scores. PSPNet with UNet has the least scores, with the precision 0.8298, F1 0.8333, and IoU 0.7632. The performances of DeepLabV3 with UNet are slightly higher than that of PSPNet with UNet. The FCN with HRNet48 is moderate, which achieves 0.8964 in precision, 0.8766 in recall, 0.8581 in F1, and 0.8052 in IoU.

For the vision Transformer methods, we can see that the SegFormer reaches the precision 0.9121, recall 0.9104, F1-score 0.8883, and IoU 0.8439, respectively, which prove its accurate and robust performance to segment the land and sea. SETR gets the lowest scores in all metrics. The scores match the results in **Figures 6**, **7**; it cannot acquire the ideal land area, and the shapes are very incomplete.

The floating-point operations (Flops) represent the model complexity. **Table 1** shows that the DeepLabV3 models occupy more computing power, followed by SETR and PSPNet. The FCN, FPN, and SegFormer use the least resource in all the methods. Specially, the DeepLabV3-ResNet101 consume the largest computing units, and the SegFormer is the most resource-saving. For the inference time, except the FCN-HRNet48 and SETR with 0.77 and 0.32, other methods are all under 0.3 s. The FCN with ResNet101, FPN with ResNet101, and SegFormer can even infer an image in 0.2 s. FCN-ResNet101 is the fastest method in inference.

DISCUSSION

Performance Analysis of the Superior CNNs

In the six CNN methods, the networks with the backbone of ResNet101 are the best extractors, and they occupy the top three for accuracy. It is followed by the HRNet48, and the UNet is the last. In ResNet101, the convolution layers are very deep, and the features are connected with residual blocks. It can keep more detail features and avoid gradient vanishing by this way. Meanwhile, this ResNet101 uses dilated convolution to increase the receptive field, which makes it more powerful. The HRNet generates high-resolution and low-resolution parallel subnetworks. It can merge the high-resolution features and low-resolution features through the different stages by connecting the multi-resolution parallel subnetworks. Therefore, it can obtain rich high-resolution and low-resolution representations. The best header is DeepLabV3 because it achieves the best scores when different methods use the same backbone of UNet or ResNet101. In DeepLabV3, the atrous spatial pyramid pooling (ASPP) adds a series of atrous convolutions with different dilated rates to increase global contextual information. Global average pooling (GAP) also

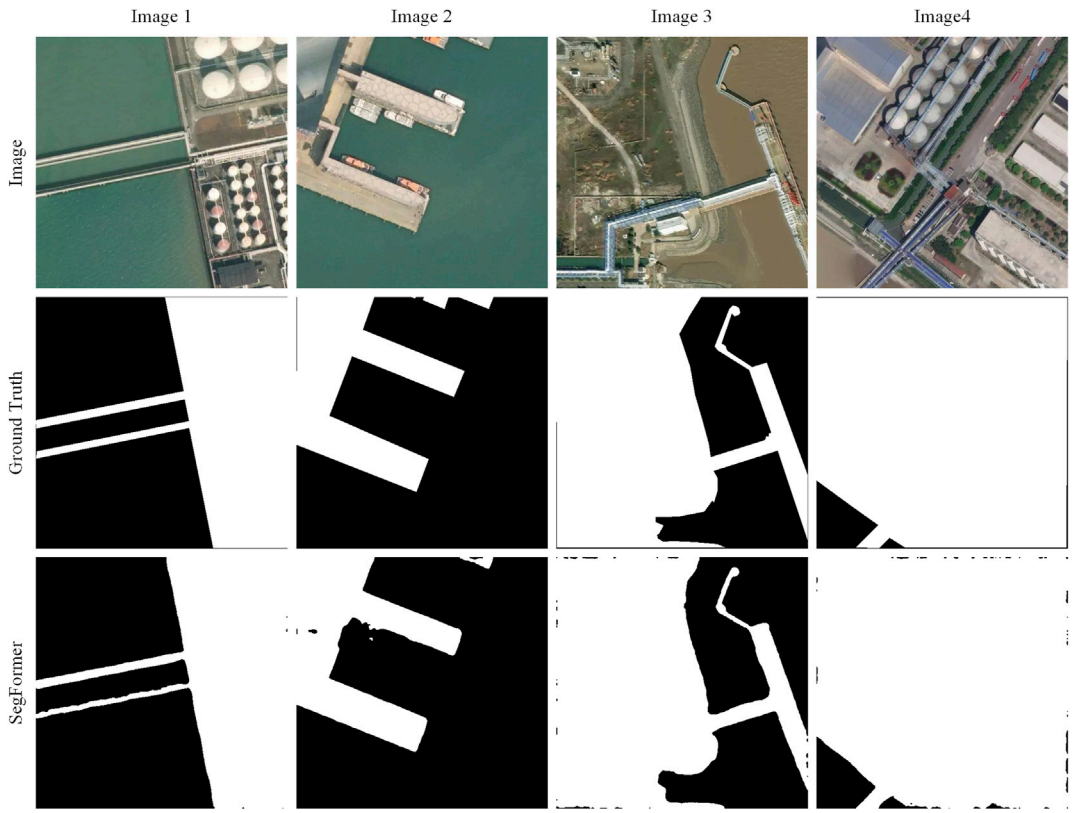

FIGURE 8 | Results of SegFormer.

TABLE 1 | Comparison for all the methods in metrics.

Method	Backbone	Flops (GFLOPs)	Inference Time(s)	Precision	Recall	F1	IoU
PSPNet	UNet	197.76	0.25	0.8298	0.8782	0.8333	0.7632
DeepLabV3	UNet	203.43	0.25	0.8518	0.8703	0.8406	0.7719
FCN	HRNet48	93.38	0.77	0.8964	0.8766	0.8581	0.8052
FCN	ResNet101	76.07	0.08	0.8995	0.8814	0.8636	0.8113
FPN	ResNet101	64.73	0.09	0.9010	0.8793	0.8637	0.8096
DeepLabV3	ResNet101	347.33	0.23	0.9056	0.8814	0.8674	0.8169
SETR	T-Large	212.4	0.32	0.8244	0.8397	0.8018	0.7268
SegFormer	MiT-B5	51.83	0.15	0.9121	0.9104	0.8883	0.8439

combines image-level features. These all make the DeepLabV3 outstanding.

Performance Analysis of the Superior Transformer

In the two vision transformers, the SETR obtain F1 with 0.8018 and IoU with 0.7268, and the SegFormer achieves 0.8883 in F1-score and 0.8439 in IoU. The SegFormer wins the SETR in accuracy completely. It can also be seen from the **Figures 7**, **8**, the SETR cannot extract the integrated and continuous structures in Image 2 and Image 3, and the SegFormer can extract nearly the whole and accurate structures. In SETR, the feature maps after the transformer layers are in the same size, and in the SegFormer, it generates multi-level feature maps. The different scales of feature maps include the high-resolution coarse features and low-resolution fine-grained features, so it can adapt to large and small object extractions. At the same time, the decoder in the SegFormer is made up of only MLP, which is lighter and has a larger effective field than traditional CNN encoders. These all make the SegFormer perform better than the SETR. On the other side, the SETR has more computational complexity with 212.4 GFLOPs than the SegFormer with just 51.83 GFLOPs. The huge amount of computation of the SETR is from the self-attention in the transformer. Because the computational complexity of self-attention is $O(N^2)$, N is the length of the input sequence. The SegFormer uses the efficient self-attention

which reduces the computational complexity by some transforms. Therefore, the SegFormer is much easier to compute.

The Most Robust Method

The top CNN DeepLabV3 and the transformer SegFormer are all very competitive. However, the vision transformer SegFormer is superior to DeepLabV3 in precision, recall, F1, and IoU. It also has a smaller complexity and shorter inference time. The limited receptive field in DeepLabV3 requires the ASPP module to enlarge the receptive field, but the model inevitably becomes heavy. The SegFormer benefits from the non-local attention in transformers and enjoys a larger receptive field. The transformer integrates with the MLP decoder an can produce both highly local and non-local attention by adding fewer parameters. These all make the SegFormer more efficient and lighter in waterline extraction.

CONCLUSION

We propose a new method based on the vision transformers for the waterline extraction by sea–land segmentation. Two transformers, the SegFormer and SETR, are adapted to segment and identify land pixels by a custom dataset from satellite maps. The performances of the two transformers are compared with other state-of-the-art CNN methods, PSPNet, DeepLabV3, FCN, and FPN. The SETR with a pure transformer structure, as an early comer to image segmentation, achieves a nearly equivalent performance compared with the developed CNN methods. More surprisingly, the latter method, the SegFormer outperforms state-of-the-art CNN methods and demonstrates an extraordinary ability to segment land pixels under different conditions. For future work, we hope to improve the method in accuracy and robustness, though it has achieved a fairly good performance.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

LY conceived and designed the analysis, collected the data, performed the analysis, wrote the original draft, and discussed the results. XW verified the analytical methods, discussed the results, and reviewed and edited the manuscript. JZ supervised the work and reviewed and edited the manuscript.

FUNDING

This study was supported by the National Natural Science Foundation of China (Grant nos. 41701480, 42006168, and 41806116) and the Key Research and Development Program of Tianjin (Grant no. 20YFZCSN01040) and the Natural Science Foundation of Tianjin (Grant nos. 20JCQNJC01230). Tianjin Philosophy and Social Science Planning Project of China (Grant no. TJKS20XSX-015).

REFERENCES

Ao, D., Dumitru, O., Schwarz, G., and Datcu, M. P. (2017). "Coastline Detection with Time Series of SAR Images," in Remote Sensing of the Ocean, Sea Ice, Coastal Waters, and Large Water Regions 2017, San Diego, CA, August 6–10, 2017 (SPIE), 70–78. doi:10.1117/12.2278318

Bayram, B., Avsar, O., Seker, D. Z., Kayi, A., Erdogan, M., Eker, O., et al. (2017). The Role of National and International Geospatial Data Sources in Coastal Zone Management. *Fresenius Environ. Bull.* 26, 383–391.

Cao, K., Fan, J., Xinxin Wang, X., Xiang Wang, X., Jianhua Zhao, J., and Fengshou Zhang, F. (2016). "Coastline Automatic Detection Based on High Resolution SAR Images," in 2016 4th International Workshop on Earth Observation and Remote Sensing Applications, Guangzhou, China, July 4–6, 2016 (EORSA), 43–46. doi:10.1109/EORSA.2016.7552763

Chen, C., Fu, J., Zhang, S., and Zhao, X. (2019). Coastline Information Extraction Based on the Tasseled Cap Transformation of Landsat-8 OLI Images. *Estuarine Coastal Shelf Sci.* 217, 281–291. doi:10.1016/j.ecss.2018.10.021

Cheng, D., Meng, G., Xiang, S., and Pan, C. (2016). Efficient Sea-Land Segmentation Using Seeds Learning and Edge Directed Graph Cut. *Neurocomputing* 207, 36–47. doi:10.1016/j.neucom.2016.04.020

Cheng, D., Meng, G., Cheng, G., and Pan, C. (2017a). SeNet: Structured Edge Network for Sea-Land Segmentation. *IEEE Geosci. Remote Sensing Lett.* 14, 247–251. doi:10.1109/LGRS.2016.2637439

Cheng, D., Meng, G., Xiang, S., and Pan, C. (2017b). FusionNet: Edge Aware Deep Convolutional Networks for Semantic Segmentation of Remote Sensing Harbor Images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sensing* 10, 5769–5783. doi:10.1109/JSTARS.2017.2747599

Chu, X., Tian, Z., Wang, Y., Zhang, B., Ren, H., Wei, X., et al. (2021). Twins: Revisiting the Design of Spatial Attention in Vision Transformers. arXiv [Preprint]. Available at: http://arxiv.org/abs/2104.13840 (Accessed October 20, 2021).

Cui, B., Jing, W., Huang, L., Li, Z., and Lu, Y. (2021). SANet: A Sea-Land Segmentation Network via Adaptive Multiscale Feature Learning. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sensing* 14, 116–126. doi:10.1109/JSTARS.2020.3040176

Dewi, R., Bijker, W., Stein, A., and Marfai, M. (2016). Fuzzy Classification for Shoreline Change Monitoring in a Part of the Northern Coastal Area of Java, Indonesia. *Remote Sensing* 8, 190. doi:10.3390/rs8030190

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., et al. (2021). An Image Is Worth 16x16 Words: Transformers for Image Recognition at Scale. arXiv [Preprint]. Available at: http://arxiv.org/abs/2010.11929 (Accessed October 20, 2021).

Elkhateeb, E., Soliman, H., Atwan, A., Elmogy, M., Kwak, K.-S., and Mekky, N. (2021). A Novel Coarse-To-Fine Sea-Land Segmentation Technique Based on Superpixel Fuzzy C-Means Clustering and Modified Chan-Vese Model. *IEEE Access* 9, 53902–53919. doi:10.1109/ACCESS.2021.3065246

Erdem, F., Bayram, B., Bakirman, T., Bayrak, O. C., and Akpinar, B. (2021). An Ensemble Deep Learning Based Shoreline Segmentation Approach (WaterNet) from Landsat 8 OLI Images. *Adv. Space Res.* 67, 964–974. doi:10.1016/j.asr.2020.10.043

Fan, J., Cao, K., Zhao, J., Jiang, D., and Tang, X. (2016). "A Hybrid Particle Swarm Optimization Algorithm for Coastline SAR Image Automatic Detection," in 2016 12th World Congress on Intelligent Control and Automation, Guilin, China, June 12–15, 2016 (WCICA), 822–825. doi:10.1109/WCICA.2016.7578256

Ge, X., Sun, X., and Liu, Z. (2014). "Object-oriented Coastline Classification and Extraction from Remote Sensing Imagery," in Remote Sensing of the Environment: 18th National Symposium on Remote Sensing of China, Wuhan, China, October 20–23, 2012 (SPIE), 131–137. doi:10.1117/12.2063845

Google Maps (2021). Available at: https://www.google.com/maps/ (Accessed August 15, 2020).

Gucluer, D., Bayram, B., and Maktav, D. (2010). "Land Cover and Coast Line Change Detection by Using Object Oriented Image Processing in Alacati, Turkey," in Imagin[e,g] Europe, Chania, Greece, 158–165. doi:10.3233/978-1-60750-494-8-158

Guo, Q., Pu, R., Zhang, B., and Gao, L. (2016). "A Comparative Study of Coastline Changes at Tampa Bay and Xiangshan Harbor During the Last 30 Years," in 2016 IEEE International Geoscience and Remote Sensing Symposium, Beijing, China, July 10–15, 2016 (IGARSS), 5185–5188. doi:10.1109/IGARSS.2016.7730351

He, L., Xu, Q., Hu, H., and Zhang, J. (2018). "Fast and Accurate Sea-Land Segmentation Based on Improved SeNet and Coastline Database for Large-Scale Image," in 2018 Fifth International Workshop on Earth Observation and Remote Sensing Applications, Xi'an, China, June 18–20, 2018 (EORSA), 1–5. doi:10.1109/EORSA.2018.8598546

Hendrycks, D., and Gimpel, K. (2020). Gaussian Error Linear Units (GELUs). arXiv [Preprint]. Available at: http://arxiv.org/abs/1606.08415 (Accessed December 6, 2021).

Li, R., Liu, W., Yang, L., Sun, S., Hu, W., Zhang, F., et al. (2018). DeepUNet: A Deep Fully Convolutional Network for Pixel-Level Sea-Land Segmentation. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sensing* 11, 3954–3962. doi:10.1109/JSTARS.2018.2833382

Lin, L., Pan, Z., Xiao, K., and Ye, N. (2013). "The Coastline Extraction for Fujian Province Based on Long Time Series of Remote Sensing Image," in Proceedings of the 2013 International Conference on Remote Sensing,Environment and Transportation Engineering, Nanjing, China, July 26–28, 2013, 63–66.

Lin, H., Shi, Z., and Zou, Z. (2017). Maritime Semantic Labeling of Optical Remote Sensing Images with Multi-Scale Fully Convolutional Network. *Remote Sensing* 9, 480. doi:10.3390/rs9050480

Liu, Y., Huang, H., Qiu, Z., and Fan, J. (2013). Detecting Coastline Change from Satellite Images Based on beach Slope Estimation in a Tidal Flat. *Int. J. Appl. Earth Obs. Geoinf.* 23, 165–176. doi:10.1016/j.jag.2012.12.005

Liu C, C., Yang, J., Yin, J., and An, W. (2016). Coastline Detection in SAR Images Using a Hierarchical Level Set Segmentation. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sensing* 9, 4908–4920. doi:10.1109/JSTARS.2016.2613279

Liu Z, Z., Li, F., Li, N., Wang, R., and Zhang, H. (2016). A Novel Region-Merging Approach for Coastline Extraction from Sentinel-1A IW Mode SAR Imagery. *IEEE Geosci. Remote Sensing Lett.* 13, 1–5. doi:10.1109/LGRS.2015.2510745

Liu, C., Xiao, Y., and Yang, J. (2017). A Coastline Detection Method in Polarimetric SAR Images Mixing the Region-Based and Edge-Based Active Contour Models. *IEEE Trans. Geosci. Remote Sensing* 55, 3735–3747. doi:10.1109/TGRS.2017.2679112

Liu, X.-Y., Jia, R.-S., Liu, Q.-M., Zhao, C.-Y., and Sun, H.-M. (2019). Coastline Extraction Method Based on Convolutional Neural Networks-A Case Study of Jiaozhou Bay in Qingdao, China. *IEEE Access* 7, 180281–180291. doi:10.1109/ACCESS.2019.2959662

Liu W, W., Chen, X., Ran, J., Liu, L., Wang, Q., Xin, L., et al. (2021). LaeNet: A Novel Lightweight Multitask CNN for Automatically Extracting Lake Area and Shoreline from Remote Sensing Images. *Remote Sensing* 13, 56. doi:10.3390/rs13010056

Liu Z, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., et al. (2021). Swin Transformer: Hierarchical Vision Transformer Using Shifted Windows. arXiv [Preprint]. Available at: http://arxiv.org/abs/2103.14030 (Accessed October 20, 2021).

Mapbox (2021). Available at: https://www.mapbox.com/ (Accessed December 7, 2021).

Microsoft (2021). Bing Maps Imagery API. Available at: https://docs.microsoft.com/en-us/bingmaps/rest-services/imagery (Accessed December 7, 2021).

MMSegmentation (2020). MMSegmentation: OpenMMLab Semantic Segmentation Toolbox and Benchmark. Available at: https://github.com/open-mmlab/mmsegmentation (Accessed December 7, 2021).

Modava, M., and Akbarizadeh, G. (2017). Coastline Extraction from SAR Images Using Spatial Fuzzy Clustering and the Active Contour Method. *Int. J. Remote Sens.* 38, 355–370. doi:10.1080/01431161.2016.1266104

Nunziata, F., Buono, A., Migliaccio, M., and Benassai, G. (2016). Dual-Polarimetric C- and X-Band SAR Data for Coastline Extraction. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sensing* 9, 4921–4928. doi:10.1109/JSTARS.2016.2560342

Paravolidakis, V., Moirogiorgou, K., Ragia, L., Zervakis, M., and Synolakis, C. (2016). "Coastline Extraction from Aerial Images Based on Edge Detection," in XXIII Congress of International Society for Photogrammetry and Remote Sensing (ISPRS 2016), Prague, Czech Republic, July 12–19, 2016, 153–158. doi:10.5194/isprsannals-III-8-153-2016

Paravolidakis, V., Ragia, L., Moirogiorgou, K., and Zervakis, M. (2018). Automatic Coastline Extraction Using Edge Detection and Optimization Procedures. *Geosciences* 8, 407. doi:10.3390/geosciences8110407

Paszke, A., Gross, S., Chintala, S., Chanan, G., Yang, E., DeVito, Z., et al. (2017). "Automatic Differentiation in PyTorch," in 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA.

Rasuly, A., Naghdifar, R., and Rasoli, M. (2010). Monitoring of Caspian Sea Coastline Changes Using Object-Oriented Techniques. *Proced. Environ. Sci.* 2, 416–426. doi:10.1016/j.proenv.2010.10.046

Rigos, A., Tsekouras, G. E., Vousdoukas, M. I., Chatzipavlis, A., and Velegrakis, A. F. (2016). A Chebyshev Polynomial Radial Basis Function Neural Network for Automated Shoreline Extraction from Coastal Imagery. *ICA* 23, 141–160. doi:10.3233/ICA-150507

Roelfsema, C., Kovacs, E. M., Saunders, M. I., Phinn, S., Lyons, M., and Maxwell, P. (2013). Challenges of Remote Sensing for Quantifying Changes in Large Complex Seagrass Environments. *Estuarine Coastal Shelf Sci.* 133, 161–171. doi:10.1016/j.ecss.2013.08.026

Shamsolmoali, P., Zareapoor, M., Wang, R., Zhou, H., and Yang, J. (2019). A Novel Deep Structure U-Net for Sea-Land Segmentation in Remote Sensing Images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sensing* 12, 3219–3232. doi:10.1109/JSTARS.2019.2925841

Sun, B., Li, S., and Xie, J. (2019). "Sea-Land Segmentation for Harbour Images with Superpixel CRF," in IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, July 28–August 2, 2019, 3899–3902. doi:10.1109/IGARSS.2019.8899001

Toure, S., Diop, O., Kpalma, K., and Maiga, A. S. (2018). "Coastline Detection Using Fusion of over Segmentation and Distance Regularization Level Set Evolution," in *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* (Istanbul, Turkey: Copernicus GmbH), 513–518. doi:10.5194/isprs-archives-XLII-3-W4-513-2018

Touvron, H., Cord, M., Douze, M., Massa, F., Sablayrolles, A., and Jegou, H. (2021). "Training Data-Efficient Image Transformers & Distillation Through Attention," in Proceedings of the 38th International Conference on Machine Learning (PMLR), Virtual Event, July 18–24, 2021, 10347–10357. Available at: https://proceedings.mlr.press/v139/touvron21a.html (Accessed October 20, 2021).

Tsekouras, G. E., Trygonis, V., Maniatopoulos, A., Rigos, A., Chatzipavlis, A., Tsimikas, J., et al. (2018). A Hermite Neural Network Incorporating Artificial Bee Colony Optimization to Model Shoreline Realignment at a Reef-Fronted beach. *Neurocomputing* 280, 32–45. doi:10.1016/j.neucom.2017.07.070

Vos, K., Splinter, K. D., Harley, M. D., Simmons, J. A., and Turner, I. L. (2019). CoastSat: A Google Earth Engine-Enabled Python Toolkit to Extract Shorelines from Publicly Available Satellite Imagery. *Environ. Model. Softw.* 122, 104528. doi:10.1016/j.envsoft.2019.104528

Wang, D., and Liu, X. (2019). Coastline Extraction from SAR Images Using Robust Ridge Tracing. *Mar. Geodesy* 42, 286–315. doi:10.1080/01490419.2019.1583147

Wang, C., Yang, J., Li, J., and Chu, J. (2020). Deriving Natural Coastlines Using Multiple Satellite Remote Sensing Images. *J. Coastal Res.* 102, 296–302. doi:10.2112/SI102-036.1

Wang, W., Xie, E., Li, X., Fan, D.-P., Song, K., Liang, D., et al. (2021). Pyramid Vision Transformer: A Versatile Backbone for Dense Prediction without Convolutions. arXiv [Preprint]. Available at: http://arxiv.org/abs/2102.12122 (Accessed October 20, 2021).

Wernette, P., Houser, C., and Bishop, M. P. (2016). An Automated Approach for Extracting Barrier Island Morphology from Digital Elevation Models. *Geomorphology* 262, 1–7. doi:10.1016/j.geomorph.2016.02.024

Widyantara, I. M. O., Wirastuti, N. M. A. E. D., and Asana, I. M. D. P. (2017). "Gamma Correction-Based Image Enhancement and Canny Edge Detection for Shoreline Extraction from Coastal Imagery," in 2017 1st International Conference on Informatics and Computational Sciences (ICICoS), Semarang, Central Java, Indonesia, November 15–16, 2017, 17–22. doi:10.1109/icicos.2017.8276331

Wu, X., Liu, C., and Wu, G. (2018). Spatial-Temporal Analysis and Stability Investigation of Coastline Changes: A Case Study in Shenzhen, China. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sensing* 11, 45–56. doi:10.1109/JSTARS.2017.2755444

Xie, E., Wang, W., Yu, Z., Anandkumar, A., Alvarez, J. M., and Luo, P. (2021). SegFormer: Simple and Efficient Design for Semantic Segmentation with Transformers. arXiv [Preprint]. Available at: http://arxiv.org/abs/2105.15203 (Accessed October 20, 2021).

Xu, W., Xu, Y., Chang, T., and Tu, Z. (2021). Co-Scale Conv-Attentional Image Transformers. arXiv [Preprint]. Available at: http://arxiv.org/abs/2104.06399 (Accessed October 20, 2021).

Yang, C.-S., Park, J.-H., and Rashid, A. H.-A. (2018). An Improved Method of Land Masking for Synthetic Aperture Radar-Based Ship Detection. *J. Navigation* 71, 788–804. doi:10.1017/S037346331800005X

Zheng, S., Lu, J., Zhao, H., Zhu, X., Luo, Z., Wang, Y., et al. (2021). Rethinking Semantic Segmentation from a Sequence-To-Sequence Perspective with Transformers. arXiv [Preprint]. Available at: http://arxiv.org/abs/2012.15840 (Accessed August 12, 2021).

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.

frontiers in Environmental Science

ORIGINAL RESEARCH
published: 04 March 2022
doi: 10.3389/fenvs.2022.794028

Active Fire Detection Using a Novel Convolutional Neural Network Based on Himawari-8 Satellite Images

Zhonghua Hong[1], Zhizhou Tang[1], Haiyan Pan[1], Yuewei Zhang[2]*, Zhongsheng Zheng[1], Ruyan Zhou[1], Zhenling Ma[1], Yun Zhang[1], Yanling Han[1], Jing Wang[1] and Shuhu Yang[1]*

[1]College of Information Technology, Shanghai Ocean University, Shanghai, China, [2]Meteorological Satellite Ground Station, Guangzhou, China

Fire is an important ecosystem process and has played a complex role in terrestrial ecosystems and the atmosphere environment. Sometimes, wildfires are highly destructive natural disasters. To reduce their destructive impact, wildfires must be detected as soon as possible. However, accurate and timely monitoring of wildfires is a challenging task due to the traditional threshold methods easily be suffered to the false alarms caused by small forest clearings, and the omission error of large fires obscured by thick smoke. Deep learning has the characteristics of strong learning ability, strong adaptability and good portability. At present, few studies have addressed the wildfires detection problem in remote sensing images using deep learning method in a nearly real time way. Therefore, in this research we proposed an active fire detection system using a novel convolutional neural network (FireCNN). FireCNN uses multi-scale convolution and residual acceptance design, which can effectively extract the accurate characteristics of fire spots. The proposed method was tested on dataset which contained 1,823 fire spots and 3,646 non-fire spots. The experimental results demonstrate that the FireCNN is fully capable of wildfire detection, with the accuracy of 35.2% higher than the traditional threshold method. We also examined the influence of different structural designs on the performance of neural network models. The comparison results indicates the proposed method produced the best results.

Keywords: active fire detection, deep learning, active fire dataset, wildfire, himawari-8 imagery, fireCNN

OPEN ACCESS

Edited by:
Peng Liu,
Institute of Remote Sensing and Digital Earth (CAS), China

Reviewed by:
Yiyun Chen,
Wuhan University, China
Bijoy Vengasseril Thampi,
Science Systems and Applications, Inc., United States
Osama M. Bushnaq,
King Abdullah University of Science and Technology, Saudi Arabia

****Correspondence:***
Haiyan Pan
hy-pan@shou.edu.cn
Yuewei Zhang
weixing1132@126.com

Specialty section:
This article was submitted to Environmental Informatics and Remote Sensing, a section of the journal Frontiers in Environmental Science

Received: *13 October 2021*
Accepted: *31 January 2022*
Published: *04 March 2022*

Citation:
Hong Z, Tang Z, Pan H, Zhang Y, Zheng Z, Zhou R, Ma Z, Zhang Y, Han Y, Wang J and Yang S (2022) Active Fire Detection Using a Novel Convolutional Neural Network Based on Himawari-8 Satellite Images. Front. Environ. Sci. 10:794028. doi: 10.3389/fenvs.2022.794028

INTRODUCTION

Fire is an important ecosystem process and has played a complex role in shaping landscapes, biodiversity and terrestrial ecosystems and the atmosphere environment (Bixby et al., 2015; Ryu et al., 2018; McWethy et al., 2019; Tymstra et al., 2020). It provide nutrients and habitat for vegetation and animals, and plays multiple important roles in maintaining healthy ecosystems (Ryan et al., 2013; Brown et al., 2015; Harper et al., 2017). However, wildfires are also destructive forces—it cause great loss of human life and damage to property, atmospheric pollution, soil damage and so on. The existing studies showing an estimated global annual burning area of approximately 420 million hectares (Giglio et al., 2018). Therefore, to reduce the negative impact of fire, real-time detection of active fires should be carried out, which can provide timely and valuable information for fire management department.

With the continuous development of satellite remote sensing technology, an increasing number of researchers have chosen to use satellite multispectral images to detect forest wildfires (Allison et al., 2016; Kaku, 2019; Barmpoutis et al., 2020). The common features of fires are bright flames and smoke produced during combustion, as well as high temperatures on fire surfaces that are different from the surrounding environment. Smoke and flames produced during combustion can be detected in the visible light bands of remote sensing images, and high temperatures on the surface of fires are easily detected in the mid-infrared, shortwave infrared and thermal infrared bands (Leblon et al., 2012). In moderate or low spatial resolution images, the fire is represented as a fire spot with extremely high temperature, which also called thermal anomalies on a per-pixel basis (Xie et al., 2016). For instance, MOD14 monitors fire actively at a 1 km spatial resolution. Satellite remote sensing has the advantages of strong timeliness, wide observation range and low cost, which provides great convenience for fire detection (Coen and Schroeder, 2013; Xie et al., 2018).

Active fire detection methods can be divided into two types: those that are based on a manual design algorithm, primarily the threshold method, and the alternative approach, based on deep learning, including shallow neural networks and image-level deep networks.

The threshold-based method sets one or more thresholds for a specific imager channel, or the combination of different spectral channels, checks each pixel one by one, and classifies the pixels that meet the threshold as fire spots; otherwise, they are classified as non-fire spots. Spectral, spatial or contextual information usually involved. The major satellite remote sensing for active fire detection are: the Moderate-resolution Imaging Spectro radiometer (MODIS) sensor that equips the NASA Terra and Aqua satellites, with the spatial resolution of 250m to 1 km (Justice et al., 2002; Morisette et al., 2005; Giglio et al., 2008; Maier et al., 2013; Xie et al., 2016; Giglio et al., 2016; Earl and Simmonds, 2018); the AVHRR sensor on board NOAA satellite, with the spatial resolution of 1 km (Baum and Trepte, 1999; Boles and Verbyla, 2000); the Visible Infrared Imaging Radiometer Suite (VIIRS) on board the joint NASA/NOAA Suomi National Polar-orbiting Partnership (Suomi NPP) and NOAA-20 satellites (Schroeder et al., 2014; Li et al., 2018). In addition, the Landsat series, Sentinel-2 remote sensing images have also been used for research on this filed due to they are relatively high spatial and temporal resolution (Schroeder et al., 2008; Murphy et al., 2016; Schroeder et al., 2016; Malambo and Heatwole, 2020; Hu et al., 2021; van Dijk et al., 2021). The major improvements of these methods are concentrate on integrating contextual or temporal information (Schroeder et al., 2008; Murphy et al., 2016; Lin et al., 2018; Kumar and Roy, 2018), setting more accurate threshold (Baum and Trepte, 1999), and improving the disturbance factors algorithm, such as cloud, smoke, and snow (Giglio et al., 2016).

In October 2014, a new geostationary meteorological satellite Himawari-8 was launched by the Japan Meteorological Agency (JMA). The satellite is equipped with an Advanced Himawari Imager (AHI) 16-channel multispectral sensor with a spatial resolution of 2 km (Xu et al., 2017). AHI can collect a full-disk of data every 10 min, covering East Asia (Da, 2015). The high temporal resolution of the Himawari-8 satellite makes it more suitable for time-sensitive tasks, e.g., fire monitoring. Wickramasinghe et al. (2016) proposed a new AHI-FSA algorithm that uses the Advanced Himawari Imager (AHI) data to detect burning and unburned vegetation, and the edge between smoke-covered and non-smoke-covered areas, respectively. Xie et al. (2018) proposed a spatial and temporal context model to detect fires based on the high temporal resolution of the Himawari-8 satellite images and applied it to real fire scenarios. Na et al. (2018) used the 7, 4, and 3 bands of Himawari-8 data to monitor grassland fires in the border areas between China and Mongolia. The results show that the detected fires are highly consistent with the actual situation on the ground. More studies of fire detection algorithms can be found in Cocke et al. (2005), French et al. (2008), Boschetti et al. (2015).

As can be seen from the above studies, the traditional threshold methods have been widely used in active fire detection tasks. However, due to the different design of spectral bands and central wavelength of different sensors, the threshold method is usually applicable to specific satellites, which makes it difficult to apply to multiple satellites. Moreover, the threshold is determined according to the statistical data of the surrounding areas, where fires under different landforms, climates, and seasons have diverse characteristics. This imply that the threshold changes dynamically according to the study area and the data. Furthermore, the threshold methods are easily affected by cloud, thick smoke, when it used to assess large areas, is prone to false positives and omissions.

Deep learning techniques have achieved excellent results in the field of machine vision (LeCun et al., 2015). Deep learning has the characteristics of strong learning ability, strong adaptability and good portability. It can discover the intricate patterns in massive data by using a series of processing layers. Therefore, an increasing number of researchers have tried to use deep learning technology in the field of fire or smoke detection and have developed and designed many algorithms. These algorithms can be divided into neural networks at the image level and pixel level.

In the field of fire detection at image level, semantic segmentation models generally involved. Langford et al. (2018) applied the deep neural networks (DNN) to detect the wildfires. To solve the problem of imbalanced training samples, a weight-selection strategy was adopted during the DNN training process. The results showed that the weight-selection strategy was able to map wildfires more accurate compared to the normal DNN. Ba et al. (2019) designed a new convolutional neural network (CNN) model, SmokeNet, which integrates the attention of space and channel direction into CNN to enhance the feature representation of scene classification. The model was tested using the MODIS data. The experimental results indicate high consistency between model predictions and actual classification results. Vani et al. (2019) designed a convolutional neural network Inception-v3 method based on transfer learning to classify the fire and non-fire. Gargiulo et al. (2019) suggested a CNN-based super-resolution technique for active fire detection using Sentinel-2 images. Pinto et al. (2020) first use convolutional neural networks and Long Short-Term Memory (LSTMs) with an

architecture based on U-net. The red, near-infrared and middle-infrared (MIR) bands from the VIIRS sensor, combined with the VIIRS 375 m active fire product as inputs to train the model. de Almeida Pereira et al. (2021) created training and testing images and labels using high-resolution images collected by Landsat-8 to train the improved U-Net networks. Different from most of the existing studies that use optical images, Ban et al. (2020) used CNN to detect burnt areas from Sentinel-1 SAR time series images. By analysing the temporal backscatter variations, the CNN-based deep learning method can better distinguish burnt areas with higher accuracy to traditional method. Larsen et al. (2021) and Guede-Fernández et al. (2021) adopted deep learning method to identify the fire smoke. However, the location of the fire can not be directly determined. In addition, researches are also using deep learning method to detect the fire using unmanned aerial vehicle (UAV) images or videos (Yuan et al., 2017; Jiao et al., 2019; Kinaneva et al., 2019; Bushnaq et al., 2021; Guede-Fernández et al., 2021). For instance, Muhammad et al. (2018) suggests a convolutional neural network using surveillance videos. The UAV can provide timely images of the fire. However, it may not suitable for large area forest fire detection. In terms of fire detection based on pixel level, according to the knowledge of the authors, there are only two related literatures. The first literature is that Zhanqing et al. (2001) integrated a back-propagation neural network (BPNN) and the threshold methods for extracting smoke based on AVHRR imagery. The BPNN can discover and learn complex linear and nonlinear relationships from radiation measurements between smoke, cloud, and land, it is can identify the potential area covered by smoke. To remove the misclassified pixels and improve the precision, multi-threshold testing also incorporated. In 2015, Li et al. proposed an improved algorithm based on their earlier model (Zhanqing et al., 2001). In the improved algorithm, all bands were regarded as the input vectors of the BPNN, and the training dataset was established using the multi-threshold method to train the BPNN to identify smoke.

According to the above-mentioned studies, some open problems still exist. First, the traditional threshold methods are easily affected by cloud, thick smoke, which lead to false positives and omission errors. Second, most of the existing deep learning methods use polar-orbiting satellites images which could provide fine spatial resolution, but the temporal resolution is relatively low. High temporal is critical for fire monitoring. The Himawari-8 satellite which has the temporal resolution of 10 min, can continuously monitor fire and are thus conducive to early fire detection and adopt aggressive measures. However, few deep learning methods use Himawari-8 satellite to detect the fires. Third, most of the existing deep learning methods are conducted at image level. Over the past decades, deep learning methods have been promoting major advances in artificial intelligence, and a variety of new models have been proposed, such as generative adversarial networks (GAN), deep Convolutional Neural Networks, Recurrent Neural Networks, Long Short-Term Memory. However, they are difficult to detect the fires at pixel level, due to the low spatial resolution of Himawari-8 images and the subtle target of the fire. The existing models need to be improved.

TABLE 1 | Band information of Himawari-8.

Band	Centre wavelength (μm)	Notation	Unit
1	0.47	*A*_1	Unitless
2	0.51	*A*_2	Unitless
3	0.64	*A*_3	Unitless
4	0.86	*A*_4	Unitless
5	1.6	*A*_5	Unitless
6	2.3	*A*_6	Unitless
7	3.9	*T*_7	K
8	6.2	*T*_8	K
9	6.9	*T*_9	K
10	7.3	*T*_10	K
11	8.6	*T*_11	K
12	9.6	*T*_12	K
13	10.4	*T*_13	K
14	11.2	*T*_14	K
15	12.4	*T*_15	K
16	13.3	*T*_16	K

Therefore, the objective of the study is to propose an active fire detection system using a novel convolutional neural network (FireCNN) based on Himawari-8 satellite imageries, to fill the research gap of this area. The presented FireCNN uses multi-scale convolution and residual acceptance design, which can effectively extract the accurate characteristics of fire spots, and to improve the fire detection accuracy. The main contributions of our study are as follows. 1) We developed a novel active fire detection convolutional neural network (FireCNN) based on Himawari-8 satellite images. The new method utilizes multi-scale convolution to comprehensively assess the characteristics of fire spots and uses residual structures to retain the original characteristics, which makes it able to extract the key features of the fire spots. 2) A new Himawari-8 active fire detection dataset was created, which includes a training set and a test set. The training set includes 654 fire spots and 1,308 non-fire spots, and the test set includes 1,169 fire spots and 2,338 non-fire spots.

The remainder of the article is organised as follows. In the Data section, we explain the source and composition of the data and pre-processing steps and provide basic information regarding the study area as well as a detailed description of the database established in this study. In the Methodology section, the proposed algorithm is described in detail, and both the traditional threshold method and deep learning method used in the experiment are introduced. In the Experiment section, the relevant settings of the experiment, the parameters used for evaluation, and the analysis of the results are described. Finally, the key findings of the study are summarized, and possible future research is briefly discussed.

DATA

Data and Pre-Processing

The fire location data (Label) and multispectral image data used in this study were obtained from the Meteorological Satellite Ground Station, Guangzhou, Guangdong, China and Himawari-8. Specifically, the Himawari-8 product used in this article is full

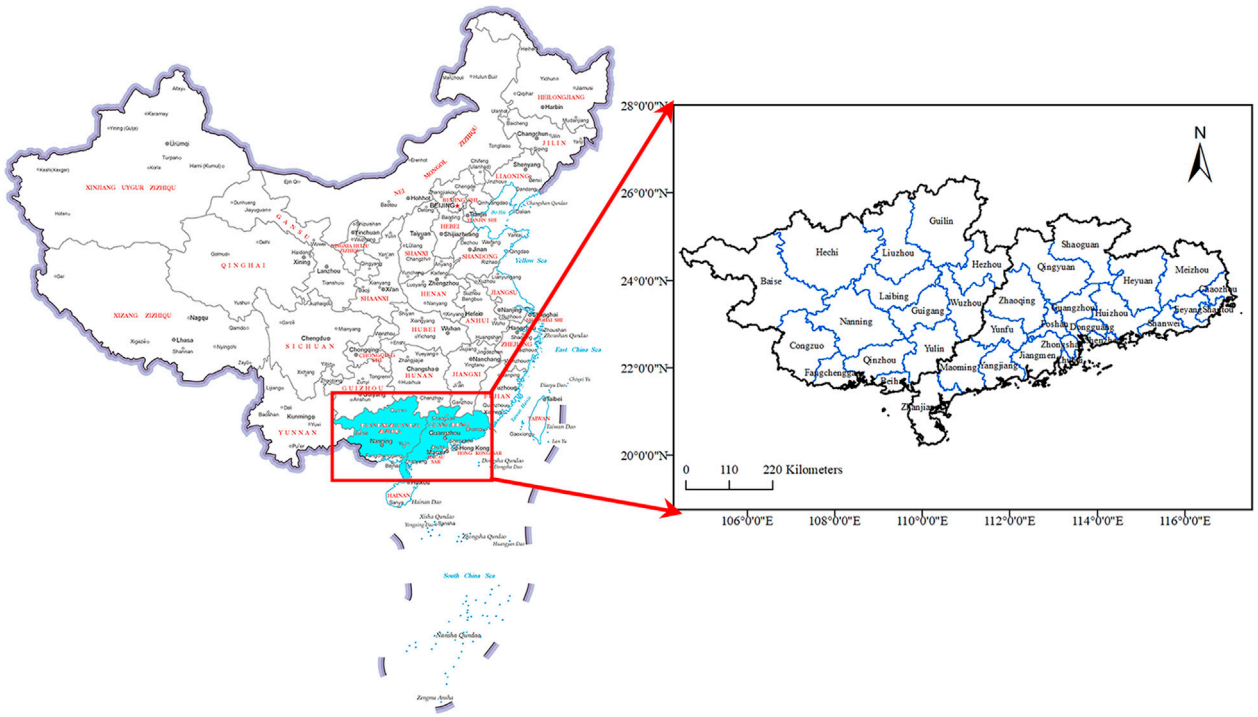

FIGURE 1 | Study area.

disk Available Himawari L1 Data (Himawari L1 Gridded data), The detailed information is available on website https://www.eorc.jaxa.jp/ptree/userguide.html. The fire location data include the longitude, latitude, and time of the fire. The data are presented in the form of mask, from which the fire and non-fire points can be extracted from the multispectral image. The data have high reliability after manual identification and algorithm inversion.

The multispectral image data were obtained from the Himawari-8 satellite Himawari-8 is comprised of 16 bands; information on each band is provided in **Table 1**. The spatial resolution of the visible light bands is 0.5–1 km and that of the near-infrared and infrared bands is 1–2 km. The temporal resolution is 10 min. The entire range covers the earth, from 60° N to 60° S and from 80° E to 160° W. Although the spatial resolution of the geostationary satellite is less than that of a polar orbit satellite, the geostationary satellite has the characteristics of wide coverage, time synchronisation of data acquisition, fixed observation position, and high temporal resolution, all of which make it well suited for real-time monitoring of wildfires. Moreover, the repeated visit once every 10 min can alleviate the problem of the blank monitoring period caused by the low temporal resolution of a polar orbit satellite. Among the Himawari-8 satellite multispectral images, bands 1-6 are albedo data, $albedo = reflectance * \cos(solar\ zenith\ angle)$, and bands 7–16 are brightness temperature data.

Study Area

A map of Guangdong Province, in southern China, is shown in **Figure 1** The whole region is located between 20° 13′ and 25° 31′ N and between 109° 39′ and 117° 19′ E. The terrain is hillier in the south than in the north. The study area is located in the East Asian monsoon region, primarily in a subtropical monsoon climate. Guangxi is adjacent to Guangdong Province, as shown in **Figure 1**, located between 104° 28′ and 112° 04′ E and 20° 54′ and 26° 24' N. In Guangxi, the terrain tends to be hilly in the northwest and less so in the southeast. The main climate is subtropical monsoon and tropical monsoon climates. The two provinces have high forest coverage rates, and both lie close to the equator, making these areas prone to forest fires during dry periods. For these reasons, they were selected as the study area.

Establishing the Database

In assembling the data, the first consideration is that the fire location data should correspond to the multispectral image data in terms of position and time. A part of the study area was cut out from the multispectral image data, and a grid of M × M size was set up at the centre of each pixel. The average and standard deviation of each band in the grid were calculated as the surrounding environment information of the pixels. To ensure that the pixels at the edge of the image can also set a sufficient window size, a sufficient width of the mirror edge was added to the image before processing. The training data is provided by Meteorological Satellite Ground Station, Guangzhou, Guangdong, China, which use combination of traditional algorithm and field survey.

According to the time and latitude information of the fire spot, the information of each band and the surrounding environment information of the fire spot were taken from the corresponding

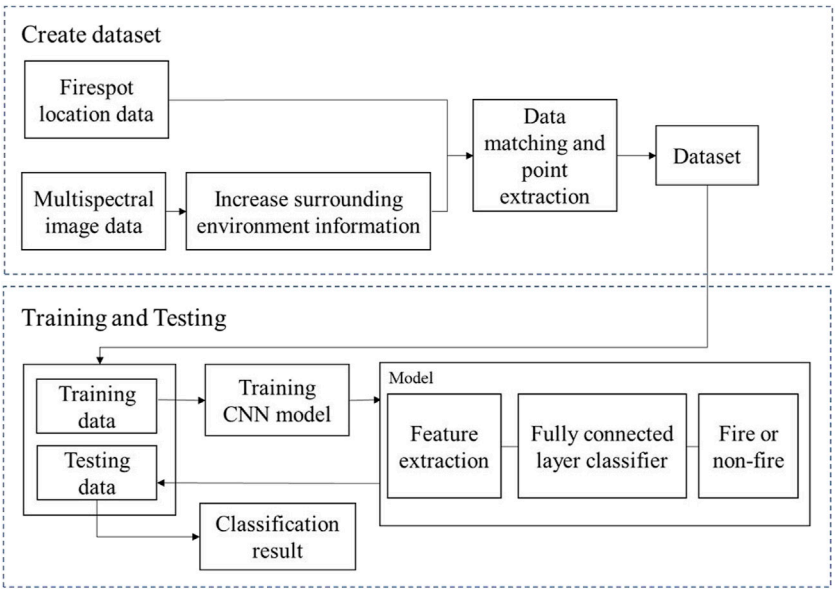

FIGURE 2 | The flow chartof the fire detection solution.

Himawari-8 image as the original characteristics of the fire spot. At the same time, the original features of non-fire spots were extracted randomly according to a certain proportion on the same scene image, where the fire spots were marked as 1 and the non-fire spots were marked as 0.

The training data set included the data of Guangdong and Guangxi provinces from January to December 2020, with the data collected at 3:00 a.m. and 7:00 p.m. (UTC) every day. Due to the unbalance number of fire and non-fire points, the proportion of fire and non-fire training points was set by comparison experiment, and the result indicates that the network can fully learns the characteristics of fires and correctly distinguishes between fires and non-fires with the proportion of 1:2. A total of 654 fire spots and 1,308 non-fire spots were included in the training set, and 40% of the training set was randomly selected as the validation set, which was not involved in training and was only used to adjust the hyper-parameters of the model and preliminarily evaluate the ability of the model to determine whether continuous training can be stopped.

METHODOLOGY

Active Fire Detection With Traditional Threshold Method

In this section, we mainly introduce the fire detection algorithm proposed by Xu et al. (2017). The algorithm first uses the 3.9- and 11.2-μm bands of Himawari-8 to identify potential fire spots. The 2.3-μm band is then used to identify water, and the 0.64-, 0.86-, and 12.4-μm bands are used to identify clouds. Water pixels and cloud pixels are removed from potential fire spots to reduce false alarms. As the final fire detection results, the experimental results

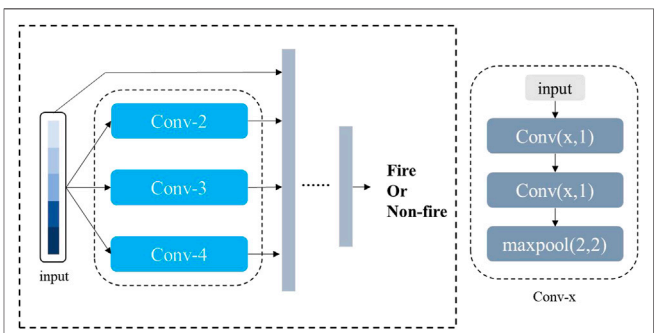

FIGURE 3 | Framework of the FireCNN.

show that the fire detection method is robust in situations of smoke and thin clouds and is very sensitive to small fires. It can provide valuable real-time fire information for wildfire management. The conditions for the algorithm to identify potential fires during the day are as follows.

$$(Z_{T_{3.9}} > 0.8) AND (Z_{T_{3.9}-T_{11.2}} > 1.5) \quad (1)$$

where $Z_{(\lambda)} = \frac{(\lambda)-mean(\lambda)}{std(\lambda)}$, $mean(\lambda)$ and $std(\lambda)$ represents the mean and standard deviation of the band in the study area.

The conditions for non-water pixels are:

$$A_{2.3} > 0.05 \quad (2)$$

where, A_λ represents the albedo value in this band.

The conditions for non-cloud pixels are:

$$(A_{0.64} + A_{0.86} < 1.2) AND (T_{12.4} > 265K) AND ((A_{0.64} + A_{0.86} < 0.7) or (T_{12.4} > 285K)) \quad (3)$$

Active Fire Detection Based on Convolutional Neural Network

In **Figure 2** we present a flow chart of the fire detection model used in this study. First, we create the training and testing data sets. The specific steps refer to the Data section. This is followed by the use of the training set to train the model. The training model is then tested and the classification results are generated.

The active fire detection problem can be transformed into a two-classification problem; that is, the pixels on the satellite image are classified as fire or non-fire. The active fire detection framework based on the CNN proposed in this study (**Figure 3**) is composed primarily of a feature extraction component and a fully connected layer classification component. The feature extraction component performs feature extraction and feature fusion on the input samples, and then the extracted features are inputted into the fully connected layer component, finally outputting the probability that the point is a fire/non-fire spot.

Feature Extraction

The feature extraction component includes three convolution modules of different scales and residual edges. The convolution modules are Conv-2, Conv-3, and Conv-4; that is, the size of the convolution kernel is 2, 3, and 4. Each convolution module includes two convolutional layers and a maximum pooling layer, and each convolutional layer is followed by a rectified linear unit (ReLU) activation function. In this study, convolutional neural networks were used in the convolution module to select features. Through convolutional layers of different scales, feature selection and extraction can be performed in different ranges, which is not only beneficial to reduce the weight of the features with poor correlation with wildfire in the original feature, but also a more comprehensive analysis of the relationship between different quantitative features and extract the key features. In the pooling layer, we chose to use the maximum pooling to retain the key features to the greatest extent, while reducing the dimension of the features to facilitate subsequent calculations. The residual edge in the convolution module prevents the loss of original features and effectively solves the problem of neural network degradation. The feature extraction component fuses the features extracted by the three convolution modules of different scales with the original features as the output.

Fully Connected Layer Classifier

The fully connected layer takes the fused features as input, taking into account all the features, and finally outputs the probability that the sample point is a fire/non-fire spot through the Softmax function. Because the problem is finally transformed into a binary classification problem, we chose the binary cross-entropy function as the loss function. The binary cross-entropy function is defined as follows:

$$Loss_{CELF}(y, \hat{y}) = y \cdot \log\hat{y} + (1 - y)\log(1 - \hat{y}) \quad (4)$$

where y is the predicted value, and $\hat{y}$ is the true value. If the point is a fire spot, $\hat{y} = 1$, if the point is a non-fire spot, $\hat{y} = 0$, and y is the probability that the point is a fire spot, $0 \leq y \leq 1$.

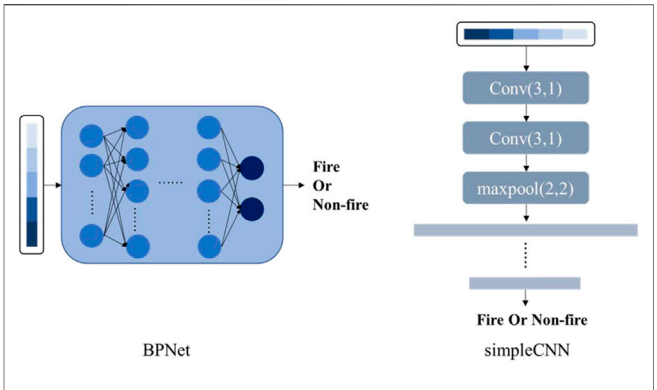

FIGURE 4 | The structure of BPNet and simpleCNN.

To verify that CNN has the potential to be suitable for thermal power detection tasks at fire spots, we compared our model with the threshold-based algorithm proposed by Xu et al. (2017). At the same time, we also compared FireCNN with BPNN (BPNet) and CNN (simpleCNN), removing multi-scale convolution and residual edges on the basis of FireCNN. The BP neural network includes five hidden layers; the number of neurons is 44, 22, 11, 6, and 2. Except for the last layer, each hidden layer uses the ReLU activation function, and the last layer uses the Softmax function. SimpleCNN includes two convolution kernels with a size of three convolutional layers, a maximum pooling layer, and a fully connected layer. The network structure of BPNet and simpleCNN is shown in **Figure 4**.

EXPERIMENT

Experimental Setup

In this study, there was no overlap between the training set (used to train the CNN model) and the test set (used to test the performance of the CNN model). The code used in this study was written using Python 3.6, and the deep learning framework used was Pytorch1.2. In terms of hardware, experiments were conducted on an Intel CoreI i5-8300H CPU at 2.30 GHz, 8 GB of RAM, running Windows, with an NVIDIA GeForce GTX 1060. In the CNN model, the Adam optimizer was selected as the parameter optimizer, the EPOCH level was 500, the batch size was 100, and the learning rate was 10^{-6}.

Evaluation of Indicators

Precision rate, misclassification error (ME), recall rate, omission error (OE), accuracy rate, and F-measure were used to evaluate the performance of the model. TP denotes true positive (correctly classified as fire point), TN denotes false positives (non-fire point), FP denotes false positives (pixels misclassified as fire point), FN denotes false negatives (pixels incorrectly classified as non-fire point).

The precision rate (P) refers to the number of fire spots predicted by the model that are actually fire spots. The higher

TABLE 2 | Comparison of five models, including the proposed FireCNN.

Method	Precision	CE	Recall	OE	Accuracy	F-measure
Xu	0.483	0.517	0.800	0.200	0.648	0.602
SVM	0.144	0.856	0.336	0.664	0.112	0.202
RF	0.281	0.719	0.776	0.224	0.264	0.413
BPNet	0.000	1.000	0.000	1.000	0.667	—
SimpleCNN	0.333	0.667	1.000	0.000	0.333	0.165
FireCNN	1.000	0.000	1.000	0.000	1.000	1.000

TABLE 3 | Additional results from the comparison of models.

Method	TP	FP	TN	FN
Xu	935	1,001	234	234
SVM	393	2,338	0	776
RF	907	2,318	20	262
BPNet	0	0	2,338	1,169
simpleCNN	1,169	2,338	0	0
FireCNN	1,169	0	2,338	0

the value, the higher is the reliability of the fire spots predicted by the model. The formula is as follows:

$$Precision = \frac{TP}{TP + FP} \cdot 100\% \tag{5}$$

The commission error (CE) refers to how many of the predicted fire spots are erroneous, and the higher the value, the more unreliable the fire spots predicted by the model. The formula is as follows:

$$CE = \frac{FP}{TP + FP} \cdot 100\% = 1 - Precision \tag{6}$$

The recall rate refers to the fire spots in the original data, how many are correctly predicted by the model, the higher the value is, the fewer the fire spots missed by the model. Recall rate is calculated, as follows:

$$Recall = \frac{TP}{TP + FN} \cdot 100\% \tag{7}$$

The omission error (OE) refers to the fire spot in the original data and the extent to which it is omitted. The higher the value, the lower the comprehensiveness of the model. The formula is as follows:

$$OE = FN/(TP + FN) = 1 - Recall \tag{8}$$

Accuracy is the ratio of the number of correct predictions in all categories to the total number of predictions. The formula is as follows:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \cdot 100\% \tag{9}$$

The F-measure was used to comprehensively evaluate the performance of the model. The formula is as follow:

$$F\,measure = (2{*}Recall{*}Precision)/(Recall + Precision) \tag{10}$$

Analysis of Results

In order to test the effectiveness of the proposed FireCNN model, five related methods are selected for comparison. They are can be divided into three types, that traditional threshold method (Xu and Zhong, 2017), machine learning methods (Support Vector Machine and Random Forest), and the deep learning methods which BP neural network and simpleCNN are involved (Zhanqing et al., 2001; Li et al., 2015). The Precision, CE, Recall, OE, Accuracy and F-measure are selected as the indicators of model performance. The results are presented in **Table 2**. In each of the methods, the training of the deep learning network uses the Adam optimizer, EPOCH level is 500, the batch size is 200, and the learning rate is 10^{-6}. More specific data are presented in **Table 3**. In particular, SVM and RF algorithms used in this study are implemented by sklrean library, where SVM kernel function is set to Gaussian kernel function; RF has 100 trees with a maximum depth of 5.

It can be observed from **Table 2** that the best performance indicators are obtained with FireCNN. The algorithm proposed by Xu does not perform as well as FireCNN in each indicator, indicating the superiority of the FireCNN to the traditional threshold method. In machine learning methods, the comprehensive performance of RF is better than SVM. SVM identifies most of the points as fire points, resulting in low accuracy. At the same time, it failed to find most of the fire points. Therefore, a low recall rate is obtained. Although RF performs better than SVM in recall rate, it still has the problems of low accuracy and more false positives. Consequently, a low recall rate is obtained. In addition, it still has the problem of low precision and more false alarms. The main reason responsible for this is that in Himawari-8 images, although abundant spectral information is provided, spatial contextual information is difficult to integrate in machine learning method due to the low spatial resolution of the images. An important character of a fire point is that its temperature is higher than the surrounding temperature. Without the surrounding environment information, the machine learning method cannot completely learn the features of the fire point. In contrast, the multi-scale convolution in FireCNN can consider and analyse the hidden relationship between various features from different scales, and the residual structure prevents the loss of original features, so it has better learning ability.

Compared with deep learning method, under the same training rounds, BPNet identifies all points as non-fire points and simpleCNN identifies all points as fire points. This indicates that these two networks have not yet learned the complete fire/non-fire characteristics. This is because the main structure of BPNet is multiple perceptron (MLP). The learning efficiency of this simple network structure is relatively low, and with the deepening of the network depth, it is prone to the problem of gradient dispersion, which leads to the network unable to further learn. SimpleCNN only uses convolution structure while multi-scale convolution is not included. When convolution is carried out on the original feature, the form is single, so the learning efficiency is low. In addition, simpleCNN does not use residual structure, which is easy to cause the loss of the original feature in the learning process. In contrast, FireCNN was able to distinguish fire from non-fire spots. In general, FireCNN performs best in all indicators, indicating that FireCNN which use multi-scale convolution and

TABLE 4 | The results of deep learning methods.

Method	EPOCH	TP	FP	TN	FN	Precision	Recall	Accuracy	Time(s)
BPNet	16,605	11,677	4	23,376	13	0.999	0.998	0.999	676.61
simpleCNN	15,905	11,688	4	23,376	2	0.999	0.999	0.999	1,512.65
FireCNN	7,945	11,683	22	23,358	7	0.998	0.999	0.999	571.42

residual edge structure, can reliably and comprehensively identify fire spots in this dataset, with fewer false positives and few omitted fire spots. This also shows that a CNN with a reasonable structure can fully adapt to fire-detection tasks.

To further explore the differences between FireCNN and BPNet and simpleCNN, we devised an additional set of experiments. Using the training set and test set used in the above experiments, EPOCH trained by BPNet and simpleCNN was added, and the EPOCH required for BPNet and simpleCNN to reach 1.000 accuracy on the training set was recorded. The higher the EPOCH level required, the slower the convergence speed of the algorithm. Because the initial parameters of the deep learning network are random, to reduce the impact of randomness, we take the sum of ten effective experiments as the final result. At the same time, the model trained by these ten experiments was tested on the test set, and the precision, recall, and accuracy were recorded, and the average value of ten effective experiments was taken. To show the training speed of each network more intuitively, we recorded the total time (unit: s) spent by ten trainings. The results are presented in **Table 4**. We also tested the results of our model against data from Guangdong and Guangxi and obtained an accuracy of 0.999 and a recall rate of 0.999. The data used for this test included all pixels in the study area, and all pixels in the study area were classified without artificially setting the ratio of fire to non-fire. On average, FireCNN spends only 4.784 s for each prediction on the data of Guangdong, and 3.659 s for Guangxi, which is far less than the 10-min time resolution of the Himawari-8 satellite. Using FireCNN, managers can fully realise the real-time monitoring of fire in Guangdong and Guangxi.

It can be seen from **Table 4** that BPNet requires the highest level of EPOCH, followed by simpleCNN, and FireCNN requires the least EPOCH. The EPOCH level of simpleCNN and FireCNN is less than that of BPNet, indicating that under the same setting, the convolution network can extract the characteristics of fire/non-fire spots more efficiently. The EPOCH level required by FireCNN is only 49.9% of that required by simpleCNN, indicating that multi-scale convolution is more effective than single-scale convolution, and in the design of the network, we set multi-scale convolution for FireCNN to ensure that the network can integrate different numbers of initial features for consideration. At the same time, there may be connections between different initial features. The convolution of different scales enables the network to consider the relationship between different numbers of features and ultimately extract the more essential features of fire/non-fire. At the same time, the existence of residual edges prevents the network from losing its original features. In terms of accuracy and recall rate, the gap between the three was not more than 0.001, indicating that the three could be applied to the fire detection task. In terms of time, FireCNN spends the least time in training, BPNet takes the second place, and simpleCNN spends the most time. In fact, the BPNet network is the simplest, and the time of one EPOCH is very short. However, because of the simple network, a higher EPOCH level is needed to train the network, and simpleCNN spends the most time. This is because its network structure is more complex than that of BPNet and it has more time to train an EPOCH. In addition, it simply uses convolution to extract features, and its efficiency is not high. The results show that the network needs to be complex enough to extract features efficiently, and a reasonable network will make the training more effective.

TABLE 5 | Statistical results.

	January	February	March	April	May	June
Guangdong	300	211	27	46	40	2
Guangxi	283	140	55	30	28	7
summation	583	351	82	76	68	9

We also recorded the test time, and the average prediction time for each point did not exceed 0.00003 s. It takes no more than 4 s to make a prediction for all data points in Guangdong Province, and no more than 5 s to make a prediction for all data points in Guangxi Province. Compared with the 10-min time resolution of Himawari-8, FireCNN is fully capable of real-time monitoring.

To place these results in context, we have provided simple statistics of the fires in Guangdong and Guangxi provinces from January 2021 to June 2021 (**Table 5**).

As shown in **Table 5**, the number of fires in Guangdong and Guangxi provinces decreased gradually from January and then again sharply in June, when the number of fires decreased to a single digit. Based on the preliminary analysis of the climate and geographical environment of Guangdong and Guangxi, the early spring, autumn, and winter rains in Guangdong and Guangxi have decreased, and the wind is dry, which leads to frequent forest fires, as evidenced by the higher number of fires in January and February. Over time, in late spring and summer, although the temperature gradually increased, it was affected by the monsoon. At this time, the rainfall was abundant, and the air humidity was high; accordingly, the number of fires decreased sharply.

CONCLUSION

To reduce the destructive impact of wildfires, it is crucial to detect the active fires accurately and quickly in the early stage. However, the most widely used threshold methods are confronted with the

problem of relative large commission and omission errors, the thresholds are varied with the study areas and so on. There are relatively few researches focus on monitoring active fires using deep learning method in a nearly real time way. In this article, we presented an active fire detection system using a novel FireCNN. FireCNN uses multi-scale convolution structure, which can consider the relationship between features from different scales, so that the network can efficiently extract features from non-fire points, and the residual structure prevents the loss of original features. These structures improve the network learning ability and learning speed. In order to evaluate the effectiveness of the proposed algorithm, it was test on Himawari-8 satellite images and the presented algorithm is compared with threshold method and the state-of-the-art deep learning models. Finally, we explored the influence of different structural designs on the deep neural network. A number of conclusions can be made as follows:

1) The FireCNN is fully capable of wildfire detection, with the accuracy of 35.2% higher than the traditional threshold method.
2) By using combination of FireCNN and Himawari-8 satellite images, the active fires can be accurately detected in nearly real time way, which is critical important to reduce the destructive impact of the active fires.
3) Reasonable network design can make the algorithm converge faster and shorten the training time.

However, the limitation of the proposed method is that the training and testing sets are relatively small. The effectiveness of the proposed method under large amount of data sets remains to be studied. In addition, the environmental information is artificially added to the original data to strengthen the feature representation. In the future research, we will try to present a more effective and robust method under large data sets.

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

Conceptualisation and design of study: ZH, ZT, HP, and YWZ; data collection: YWZ and ZT; data analysis and interpretation: ZH, ZT, HP, YWZ, ZZ, RZ, ZM, YZ, YH, JW, and SY; Writing and preparation of original draft: ZH, ZT, HP, and YWZ; Funding acquisition: ZH All authors contributed to the article revision and read and approved the submitted version.

FUNDING

This work was supported by the National Natural Science Foundation of China under Grant 41871325.

ACKNOWLEDGMENTS

Many thanks to the reviewers for their valuable comments.

REFERENCES

Allison, R., Johnston, J., Craig, G., and Jennings, S. (2016). Airborne Optical and thermal Remote Sensing for Wildfire Detection and Monitoring. *Sensors.* 16 (8), 1310. doi:10.3390/s16081310

Ba, R., Chen, C., Yuan, J., Song, W., and Lo, S. (2019). Smokenet: Satellite Smoke Scene Detection Using Convolutional Neural Network with Spatial and Channel-wise Attention. *Remote Sensing.* 11 (14), 1702. doi:10.3390/rs11141702

Ban, Y., Zhang, P., Nascetti, A., Bevington, A. R., and Wulder, M. A. (2020). Near Real-Time Wildfire Progression Monitoring with sentinel-1 Sar Time Series and Deep Learning. *Sci. Rep.* 10, 1322. doi:10.1038/s41598-019-56967-x

Barmpoutis, P., Papaioannou, P., Dimitropoulos, K., and Grammalidis, N. (2020). A Review on Early forest Fire Detection Systems Using Optical Remote Sensing. *Sensors.* 20 (22), 6442. doi:10.3390/s20226442

Baum, B. A., and Trepte, Q. (1999). A Grouped Threshold Approach for Scene Identification in AVHRR Imagery. *J. Atmos. Oceanic Technol.* 16 (6), 793–800. doi:10.1175/1520-0426(1999)016<0793:agtafs>2.0.co;2

Bixby, R. J., Cooper, S. D., Gresswell, R. E., Brown, L. E., Dahm, C. N., and Dwire, K. A. (2015). Fire Effects on Aquatic Ecosystems: an Assessment of the Current State of the Science. *Freshw. Sci.* 34 (4), 1340–1350. doi:10.1086/684073

Boles, S. H., and Verbyla, D. L. (2000). Comparison of Three AVHRR-Based Fire Detection Algorithms for interior Alaska. *Remote Sensing Environ.* 72 (1), 1–16. doi:10.1016/s0034-4257(99)00079-6

Boschetti, L., Roy, D. P., Justice, C. O., and Humber, M. L. (2015). MODIS-landsat Fusion for Large Area 30 M Burned Area Mapping. *Remote sensing Environ.* 161, 27–42. doi:10.1016/j.rse.2015.01.022

Brown, L. E., Holden, J., Palmer, S. M., Johnston, K., Ramchunder, S. J., and Grayson, R. (2015). Effects of Fire on the Hydrology, Biogeochemistry, and Ecology of Peatland River Systems. *Freshw. Sci.* 34 (4), 1406–1425. doi:10.1086/683426

Bushnaq, O. M., Chaaban, A., and Al-Naffouri, T. Y. (2021). The Role of UAV-IoT Networks in Future Wildfire Detection. *IEEE Internet Things J.* 8 (23), 16984–16999. doi:10.1109/jiot.2021.3077593

Cocke, A. E., Fulé, P. Z., and Crouse, J. E. (2005). Comparison of Burn Severity Assessments Using Differenced Normalized Burn Ratio and Ground Data. *Int. J. Wildland Fire.* 14 (2), 189–198. doi:10.1071/wf04010

Coen, J. L., and Schroeder, W. (2013). Use of Spatially Refined Satellite Remote Sensing Fire Detection Data to Initialize and Evaluate Coupled Weather-wildfire Growth Model Simulations. *Geophys. Res. Lett.* 40 (20), 5536–5541. doi:10.1002/2013gl057868

Da, C. (2015). Preliminary Assessment of the Advanced Himawari Imager (AHI) Measurement Onboard Himawari-8 Geostationary Satellite. *Remote sensing Lett.* 6 (8), 637–646. doi:10.1080/2150704x.2015.1066522

de Almeida Pereira, G. H., Fusioka, A. M., Nassu, B. T., and Minetto, R. (2021). Active Fire Detection in Landsat-8 Imagery: A Large-Scale Dataset and a Deep-Learning Study. *ISPRS J. photogrammetry remote sensing.* 178, 171–186. doi:10.1016/j.isprsjprs.2021.06.002

Earl, N., and Simmonds, I. (2018). Spatial and Temporal Variability and Trends in 2001-2016 Global Fire Activity. *J. Geophys. Res. Atmospheres.*, 2524–2536. doi:10.1002/2017jd027749

French, N. H. F., Kasischke, E. S., Hall, R. J., Murphy, K. A., Verbyla, D. L., Hoy, E. E., et al. (2008). Using Landsat Data to Assess Fire and Burn Severity in the North American Boreal forest Region: an Overview and Summary of Results. *Int. J. Wildland Fire.* 17 (4), 443–462. doi:10.1071/wf08007

Gargiulo, M., Dell'Aglio, D. A. G., Iodice, A., Riccio, D., and Ruello, G. (2019). "A CNN-Based Super-resolution Technique for Active Fire

Detection on Sentinel-2 Data," in PhotonIcs Electromagnetics Research Symposium (Spring), 418–426. doi:10.1109/piers-spring46901.2019.9017857

Giglio, L., Boschetti, L., Roy, D. P., Humber, M. L., and Justice, C. O. (2018). The Collection 6 MODIS Burned Area Mapping Algorithm and Product. *Remote sensing Environ.* 217, 72–85. doi:10.1016/j.rse.2018.08.005

Giglio, L., Csiszar, I., Restás, Á., Morisette, J. T., Schroeder, W., Morton, D., et al. (2008). Active Fire Detection and Characterization with the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER). *Remote Sensing Environ.* 112, 3055–3063. doi:10.1016/j.rse.2008.03.003

Giglio, L., Schroeder, W., and Justice, C. O. (2016). The Collection 6 MODIS Active Fire Detection Algorithm and Fire Products. *Remote sensing Environ.* 178, 31–41. doi:10.1016/j.rse.2016.02.054

Guede-Fernández, F., Martins, L., Almeida, R. V. D., Gamboa, H., and Vieira, P. (2021). A Deep Learning Based Object Identification System for Forest Fire Detection. *Fire.* 4 (4), 75. doi:10.3390/fire4040075

Harper, A. R., Doerr, S. H., Santin, C., Froyd, C. A., and Sinnadurai, P. (2017). Prescribed Fire and its Impacts on Ecosystem Services in the UK. *Sci. Total Environ.* 624, 691–703. doi:10.1016/j.scitotenv.2017.12.161

Hu, X., Ban, Y., and Nascetti, A. (2021). Sentinel-2 MSI Data for Active Fire Detection in Major Fire-Prone Biomes: a Multi-Criteria Approach. *Int. J. Appl. Earth Observation Geoinformation.* 101, 102347. doi:10.1016/j.jag.2021.102347

Jiao, Z., Zhang, Y., Xin, J., Mu, L., Yi, Y., Liu, H., et al. (2019). "A Deep Learning Based forest Fire Detection Approach Using UAV and YOLOv3," in 2019 1st International Conference on Industrial Artificial Intelligence (IAI) (IEEE), 1–5. doi:10.1109/iciai.2019.8850815

Justice, C., Giglio, L., Korontzi, S., Owens, J., Morisette, J., Roy, D., et al. (2002). The MODIS Fire Products. *Remote sensing Environ.* 83 (1-2), 244–262. doi:10.1016/s0034-4257(02)00076-7

Kaku, K. (2019). Satellite Remote Sensing for Disaster Management Support: A Holistic and Staged Approach Based on Case Studies in Sentinel Asia. *Int. J. Disaster Risk Reduction.* 33, 417–432. doi:10.1016/j.ijdrr.2018.09.015

Kinaneva, D., Hristov, G., Raychev, J., and Zahariev, P. (2019). "Early forest Fire Detection Using Drones and Artificial Intelligence," in 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO) (IEEE), 1060–1065. doi:10.23919/mipro.2019.8756696

Kumar, S. S., and Roy, D. P. (2018). Global Operational Land Imager Landsat-8 Reflectance-Based Active Fire Detection Algorithm. *Int. J. Digital Earth.* 11 (2), 154–178. doi:10.1080/17538947.2017.1391341

Langford, Z., Kumar, J., and Hoffman, F. (2018). "Wildfire Mapping in Interior Alaska Using Deep Neural Networks on Imbalanced Datasets," in IEEE International Conference on Data Mining Workshops, 770–778. doi:10.1109/icdmw.2018.00116

Larsen, A., Hanigan, I., Reich, B. J., Qin, Y., Cope, M., Morgan, G., et al. (2021). A deep learning approach to identify smoke plumes in satellite imagery in near-real time for health risk communication. *J. Exposure Sci. Environ. Epidem.* 31, 170–176.

Leblon, B., Bourgeau-Chavez, L., and San-Miguel-Ayanz, J. (2012). *Use of Remote Sensing in Wildfire Management*. London: Sustainable development-authoritative and leading edge content for environmental management, 55–82.

LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep Learning. *Nature.* 521 (7553), 436–444. doi:10.1038/nature14539

Li, F. J., Zhang, X. Y., Kondragunta, S., and Csiszar, I. (2018). Comparison of Fire Radiative Power Estimates from VIIRS and MODIS Observations. *J. Geophys. Res. Atmospheres.* 123 (9), 4545. doi:10.1029/2017jd027823

Li, X., Song, W., Lian, L., and Wei, X. (2015). Forest Fire Smoke Detection Using Back-Propagation Neural Network Based on MODIS Data. *Remote Sensing.* 7 (4), 4473–4498. doi:10.3390/rs70404473

Lin, Z., Chen, F., Niu, Z., Li, B., Yu, B., Jia, H., et al. (2018). An Active Fire Detection Algorithm Based on Multi-Temporal Fengyun-3c VIRR Data. *Remote Sensing Environ.* 211, 376–387. doi:10.1016/j.rse.2018.04.027

Maier, S. W., Russell-Smith, J., Edwards, A. C., and Yates, C. (2013). Sensitivity of the MODIS Fire Detection Algorithm (MOD14) in the savanna Region of the Northern Territory, Australia. *ISPRS J. photogrammetry remote sensing.* 76, 11–16. doi:10.1016/j.isprsjprs.2012.11.005

Malambo, L., and Heatwole, C. D. (2020). Automated Training Sample Definition for Seasonal Burned Area Mapping. *ISPRS J. Photogrammetry Remote Sensing.* 160, 107–123. doi:10.1016/j.isprsjprs.2019.11.026

McWethy, D. B., Schoennagel, T., Higuera, P. E., Krawchuk, M., Harvey, B. J., Metcalf, E. C., et al. (2019). Rethinking Resilience to Wildfire. *Nat. Sustain.* 2 (9), 797–804. doi:10.1038/s41893-019-0353-8

Morisette, J. T., Giglio, L., Csiszar, I., and Justice, C. O. (2005). Validation of the MODIS Active Fire Product over Southern Africa with ASTER Data. *Int. J. Remote Sensing.* 26 (19), 4239–4264. doi:10.1080/01431160500113526

Muhammad, K., Ahmad, J., Mehmood, I., Rho, S., Baik, S. W., Na, L., et al. (2018). Convolutional Neural Networks Based Fire Detection in Surveillance videos. *IEEE Access* 6, 18174–18183. doi:10.1109/access.2018.2812835

Murphy, S. W., de Souza Filho, C. R., Wright, R., Sabatino, G., and Correa Pabon, R. (2016). HOTMAP: Global Hot Target Detection at Moderate Spatial Resolution. *Remote sensing Environ.* 177, 78–88. doi:10.1016/j.rse.2016.02.027

Pinto, M. M., Libonati, R., Trigo, R. M., Trigo, I. F., and DaCamara, C. C. (2020). A Deep Learning Approach for Mapping and Dating Burned Areas Using Temporal Sequences of Satellite Images. *ISPRS J. Photogrammetry Remote Sensing.* 160, 260–274. doi:10.1016/j.isprsjprs.2019.12.014

Ryan, K. C., Knapp, E. E., and Varner, J. M. (2013). Prescribed Fire in North American Forests and Woodlands: History, Current Practice, and Challenges. *Front. Ecol. Environ.* 11, e15–e24. doi:10.1890/120329

Ryu, J.-H., Han, K.-S., Hong, S., Park, N.-W., Lee, Y.-W., and Cho, J. (2018). Satellite-based Evaluation of the post-fire Recovery Process from the Worst forest Fire Case in South Korea. *Remote Sensing.* 10 (6), 918. doi:10.3390/rs10060918

Schroeder, W., Oliva, P., Giglio, L., Quayle, B., Lorenz, E., and Morelli, F. (2016). Active Fire Detection Using Landsat-8/OLI Data. *Remote sensing Environ.* 185, 210–220. doi:10.1016/j.rse.2015.08.032

Schroeder, W., Oliva, P., Giglio, L., and Csiszar, I. A. (2014). The New VIIRS 375 m active fire detection data product: algorithm description and initial assessment. *Remote Sens. Environ.* 143, 85–96.

Schroeder, W., Prins, E., Giglio, L., Csiszar, I., Schmidt, C., Morisette, J., et al. (2008). Validation of GOES and MODIS Active Fire Detection Products Using ASTER and ETM+ Data. *Remote Sensing Environ.* 112, 2711–2726. doi:10.1016/j.rse.2008.01.005

Tymstra, C., Stocks, B. J., Cai, X., and Flannigan, M. D. (2020). Wildfire Management in Canada: Review, Challenges and Opportunities. *Prog. Disaster Sci.* 5, 100045. doi:10.1016/j.pdisas.2019.100045

van Dijk, D., Shoaie, S., van Leeuwen, T., and Veraverbeke, S. (2021). Spectral Signature Analysis of False Positive Burned Area Detection from Agricultural Harvests Using sentinel-2 Data. *Int. J. Appl. Earth Observation Geoinformation.* 97 (80), 102296. doi:10.1016/j.jag.2021.102296

Vani, K., Wickramasinghe, C. H., Jones, S., Reinke, K., and Wallace, L. (2019). Deep Learning Based forest Fire Classification and Detection in Satellite Images. 11th International Conference on Advanced Computing (ICoAC). *Remote Sensing.* 8 (11), 932. doi:10.1109/ICoAC48765.2019.246817

Xie, H., Du, L., Liu, S. C., Chen, L., Gao, S., Liu, S., et al. (2016). Dynamic Monitoring of Agricultural Fires in china from 2010 to 2014 Using MODIS and Globeland30 Data. *Int. J. Geo-Information.* 5 (10), 172. doi:10.3390/ijgi5100172

Xie, Z., Song, W., Ba, R., Li, X., and Xia, L. (2018). A Spatiotemporal Contextual Model for forest Fire Detection Using Himawari-8 Satellite Data. *Remote Sensing* 10 (12), 1992. doi:10.3390/rs10121992

Xu, G., and Zhong, X. (2017). Real-Time Wildfire Detection and Tracking in Australia Using Geostationary Satellite: Himawari-8. *Remote sensing Lett.* 8 (11), 1052–1061. doi:10.1080/2150704x.2017.1350303

Xu, W., Wooster, M. J., Kaneko, T., He, J., Zhang, T., and Fisher, D. (2017). Major Advances in Geostationary Fire Radiative Power (FRP) Retrieval over Asia and Australia Stemming from Use of Himarawi-8 AHI. *Remote Sensing Environ.* 193, 138–149. doi:10.1016/j.rse.2017.02.024

Yuan, C., Liu, Z., and Zhang, Y. (2017). Fire Detection Using Infrared Images for UAV-Based forest Fire Surveillance. *Proc. Int. Conf. Unmanned Aircr. Syst. (Icuas).*, 567–572. doi:10.1109/icuas.2017.7991306

Zhanqing, L., Khananian, A., Fraser, R. H., and Cihlar, J. (2001). Automatic Detection of Fire Smoke Using Artificial Neural Networks and Threshold

Approaches Applied to AVHRR Imagery. *IEEE Trans. Geosci. Remote Sensing.* 39 (9), 1859–1870. doi:10.1109/36.951076

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, orclaim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Environmental Science

ORIGINAL RESEARCH
published: 02 May 2022
doi: 10.3389/fenvs.2022.765604

A High-Resolution Remote-Sensing-Based Method for Urban Ecological Quality Evaluation

Huiping Huang[1,2], *Qiangzi Li*[1,2] *and Yuan Zhang*[1]*

[1]Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, China, [2]University of Chinese Academy of Sciences, Beijing, China

Urban ecological quality evaluation attracts more and more attention in urban land use and ecosystem planning optimization due to continuity problems from rapid urbanization and population growth. Remote sensing was always considering contribute to the evaluation. However, accurate and efficient evaluation of urban ecological quality is being challenged, as traditional remote-sensing-based methods were mainly based on low spatial resolution data, pixel-based land cover classification, and vegetation condition factors, and ignored object-oriented high spatial resolution classification and urban landscape pattern. Thus, method for urban ecological quality evaluation based on high-resolution remote sensing is greatly needed to support spatially explicit decision-making in urban planning. In this study, a novel high-resolution remote-sensing-based method based on six ecological indicators from vegetation conditions and landscape patterns was proposed to evaluate urban ecological quality. The six ecological indicators were derived from high-resolution remote sensing data using an object-oriented land cover classification. Factor analysis indicated that the sensitivity of landscape patterns to ecological quality is relatively weaken. Therefore, vegetation conditions and landscape patterns were used as two respective variables to generate a linear evaluation model, with their weights calculated from the loadings of factor analysis, to evaluation urban ecological quality. The results showed that the proposed linear model, considering both vegetation conditions and landscape patterns, is effective and trustworthy, and can provide more suitable support to urban land use and ecological planning.

Keywords: urban, ecological quality, vegetation condition, landscape pattern, remote sensing

OPEN ACCESS

Edited by:
Xuan Zhu,
Monash University, Australia

Reviewed by:
Tianhai Cheng,
Institute of Remote Sensing and Digital Earth (CAS), China
Pedzisai Kowe,
Midlands State University, Zimbabwe

***Correspondence:**
Yuan Zhang
zhangyuan@radi.ac.cn

Specialty section:
This article was submitted to Environmental Informatics and Remote Sensing, a section of the journal Frontiers in Environmental Science

Received: *27 August 2021*
Accepted: *21 March 2022*
Published: *02 May 2022*

Citation:
Huang H, Li Q and Zhang Y (2022) A High-Resolution Remote-Sensing-Based Method for Urban Ecological Quality Evaluation. Front. Environ. Sci. 10:765604. doi: 10.3389/fenvs.2022.765604

1 INTRODUCTION

Urbanization in China has been an unforeseen major historical event over the past 40 years. The urban population increased from 172.45 million in 1978 to 914.25 million in 2021, and the urbanization rate increased from 17.9 to 64.7% during the same period (National Bureau of Statistics of China, 2021). Due to the relaxation of population policies in China, the urbanization rate has shown an accelerating tendency, and is predicted to surpass 70% by 2035 (Chinese Academy of Social Sciences, 2019). However, the dramatic urbanization process poses great challenges to the ecological environment and resources (Zhang et al., 2019), leading to various ecological issues including biodiversity reduction (Elmqvist, 2013), lengthy drought (Kaufmann et al., 2007), heat island effect (Du et al., 2020), water quality deterioration (Owen, 2010), atmospheric pollution (Sarrat et al., 2006), and disease spreading (Allender et al., 2010). Shrinkage of ecological land and the resulting water shortage have become

outstanding issues in highly urbanized cities such as Beijing, China (Zhang et al., 2009). China implemented the "Green GDP" project in 2007 to deal with the challenging ecological situation. Green GDP suggests less urbanization and industrialization, and more green space expansion in urban areas. Accordingly, Beijing municipal government decided to limit residential and business land, and develop more green space in the urban area to strengthen the ecological conservation. Both growing ecological issues and the effectiveness of policy optimization need to be considered for understanding the impacts on urban residential suitability. Therefore, the analysis of urban ecological quality is of greater importance than ever in the rapid urbanization region.

A well-established approach is to use ecological indicators as standardized tools to provide comparable and comprehensive information of an urban area for urban ecological quality analysis (Lakes and Kim, 2012). Urban ecological indicators are quantitative and spatially continuous descriptions of ecological conditions in urban environments. They reflect ecological functions and ecosystem services such as groundwater recharge, retention of contaminants, air purification, and urban climate regulation (Henry and Dicks, 1987; Bolund and Hunhammar, 1999; Eliasson and Svensson, 2003; Arlt and Lehmann, 2005; Gómez-Baggethun and Barton, 2013).

Many urban ecological indicators have been developed for evaluating urban ecological quality, ranging from qualitative to quantitative, from physical based to remote sensing based, from point to surface monitoring, and from intermittent monitoring to continuous monitoring in the last 2 decades (Wang et al., 2017). However, most of the existing indicators were developed using local statistical data, labor-intensive field surveys, or visual interpretation of aerial photographs, which are difficult to collect, and often lack high quality spatial context (Cadenasso et al., 2007). Therefore, there is a need to develop a high-resolution cyberinfrastructure-based ecological indicator for effective ecological planning (Rose et al., 2015).

A remarkable array of ecological measurements can be derived from remotely sensed images that include habitats (or land cover classifications) and their biophysical properties (integrated ecosystem measurements) as well as natural and human-induced changes across the landscape (Pettorelli et al., 2014). Various automatic remote-sensing-based methods have been developed to improve the efficiency of urban ecological indicators. Behling et al. presented an automatic remote sensing and GIS based system to generate flexible and user-defined urban ecological indicators. In their work, fourteen indicators were developed based on hyperspectral remote sensing data and its corresponding height information (Behling et al., 2015). Xu developed a remote sensing based ecological index (RSEI), which takes greenness, wetness, dryness and heat into consideration. The four aspects were quantified by four remote sensing indices: normalized difference vegetation index (NDVI), normalized difference built-up and soil index (NDBSI), wetness component of the tasseled cap transformation (Wet), and land surface temperature (LST) (Xu, 2013). However, a common issue exists in the current remote-sensing-based ecological evaluation: the use of coarse indicators from low-resolution remote sensing data, which hindered the development of spatially detailed urban ecological indicators.

On the other hand, current ecological indicators from remotely sensed images mainly focused on the quantity of urban ecological land and ignored their spatial patterns. However, the spatial pattern is considered as one of the most important factors that affect urban ecological quality (Su et al., 2012; Zhou et al., 2012; Estoque and Murayama, 2013). For instance, landscape pattern strongly influences ecological processes with respect to population persistence, biodiversity, and ecosystem health. The ecological consequences of urbanization can be observed and described by the dynamic changes of regional landscape through landscape metrics (Li et al., 2010; Peng et al., 2016). Some studies have expounded the significant functions of Urban Green Space (UGS) in urban life (Uy and Nakagoshi, 2008). Landscape pattern assessment could also infer potential ecological processes (Turner et al., 2001; Botequilha Leitão and Ahern, 2002). These factors express the ecological quality of UGS from different aspects, such as biodiversity, physical and mental health, and visual and amenity benefits (Harper et al., 2005; Fuller et al., 2007; Gonzalez et al., 2010). Therefore, landscape pattern assessment should be considered to derive more comprehensive ecological quality assessment.

In light of the issues of existing studies on remote-sensing-based urban ecological quality evaluation, the objective of this article is to propose an innovative method to develop urban ecological indicators using high resolution remote sensing images. The proposed method takes both vegetation condition and landscape pattern into consideration to evaluate the urban ecological quality. The results are expected to greatly support urban managers for better understanding of the importance of urban ecological quality, and for more objective decision making in urban planning.

2 STUDY AREA, DATA COLLECTION AND PRE-PROCESSING

2.1 Study Area

In this study, Haidian District in Beijing, China (**Figure 1**), was selected as the study area. The area is located in the northwest part of Beijing city. It covers an area of 430.8 km^2 and is divided into 30 administrative sub-districts. The population of the area amounted to 3.24 million by the end of year 2019, while the production value reached 113.23 billion US dollars, taking up 22.4% of the total value of Beijing (Bureau of Statistics of Haidian District, 2020). Generally, the built-up area occupies half of the southeast area while mountains distribute in the west margin, with urban-rural fringe located in between. Because of the tremendous population, high industrial output and fragmented land use patches, Haidian District becomes a suitable area to develop an ecological quality evaluation method with both vegetation conditions and landscape patterns considered.

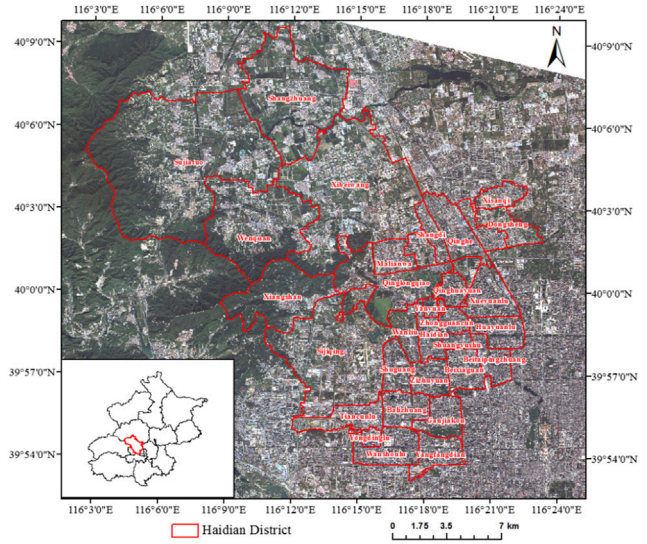

FIGURE 1 | The study area of Haidian district overlaid on a mosaicked GaoFen-2 true-color image and the location of Haidian District within Beijing, China.

2.2 Remote Sensing Data and Preprocessing

GaoFen-2 (GF-2) is a high spatial resolution remote sensing satellite that was launched by China, on 19 August 2014. It contains Panchromatic and Multi-Spectral (PMS) sensors (PMS-1 and PMS-2 with the same band designations). The band designation of GF-2 PMS is shown in **Table 1**. The spatial resolution of the multispectral and panchromatic bands is 4 and 1 m, respectively. Four GF-2 multispectral images, acquired on 12 September 2015, were used to develop ecological indicators in this study.

The multispectral data preprocessing includes radiometric calibration, atmospheric correction, ortho-rectification and mosaic. Radiometric calibration was used to convert digital numbers to Top-Of-Atmosphere (TOA) reflectance using parameters developed by the China Centre for Resources Satellite Data and Application (http://www.cresda.com.cn). Atmospheric correction is used to convert the TOA reflectance to surface reflectance using a Fast Line-of-sight Atmospheric Analysis of Spectral Hypertube (FLAASH) module. After that, the images were visualized and orthorectified using ENVI 5.1 software. Lastly, four images were mosaicked.

3 METHODS

We propose a novel high-resolution remote-sensing-based methodological framework to evaluate ecological quality of Haidian District (**Figure 2**). The proposed framework consists of four modules: 1) Generation of a fine-scaled ecological land cover map using an object-oriented image classification method; 2) Derivation of vegetation parameters and landscape metrics based on GF-2 multi-spectral data and land cover mapping; 3) Evaluation of the ecological quality (EQ1) using factor analysis and analysis of the contribution of vegetation and landscape to ecological quality based on the loadings of six indicators. 4) Development of a linear model combining both vegetation condition and landscape pattern to evaluate the ecological quality (EQ2).

TABLE 1 | Spectral bands of GaoFen-2 (GF-2) Panchromatic/Multi-Spectral (PMS).

GF–2 PMS		
Band	**Wavelength (μm)**	**Resolution (m)**
1. Blue	0.450–0.520	4
2. Green	0.520–0.590	4
3. Red	0.630–0.690	4
4. Near Infrared	0.770–0.890	4
5. Pan	0.450–0.900	1

3.1 Object-Oriented Land Cover Classification

Object-oriented image classification is one of the most effective methods to conduct land use/cover mapping using high spatial resolution remotely sensed images. As compared with pixel-based image classification, this method can remove salt-and-pepper noise and generate more reliable and accurate results. In this study, the GF-2 images were classified into various land cover types. The results were then used as a basis for landscape metrics calculation.

Object-oriented image classification involved three steps: multi-scale segmentation, general classes creation, and classification rules (Ramakrishnan, 2014). The segmentation parameters were defined as follows: layer weights were all set to equal, scale to 20, shape factor to 0.3, color to 0.7, and compactness and smoothness to 0.5. Four classes (water, vegetation, soil, and impervious surface) were created to form class hierarchy. We chose the nearest neighbor method as the classifier. Urban area surface is complex and heterogeneous, so it is very difficult to identify all land cover classes simultaneously. In this research, based on the difference of pixel spectral heterogeneity, four classes were identified. After that, water and vegetation were subdivided further. Water was further divided into two sub-types, clean water and turbid water. Vegetation was divided into two sub-types: grassland, and forest/shrub. 50 sample regions of 100 m × 100 m in size were randomly selected, and the land cover type of these pixels within these sample regions were identified using visual interpretation in order to assess the accuracy of land cover classification.

3.2 Landscape Metrics Calculation

3.2.1 Definition of Urban Ecological Land

Ecological land patches are the basic units for landscape metric calculation. Urban ecological land refers to the land-use type that can provide ecosystem services. Its

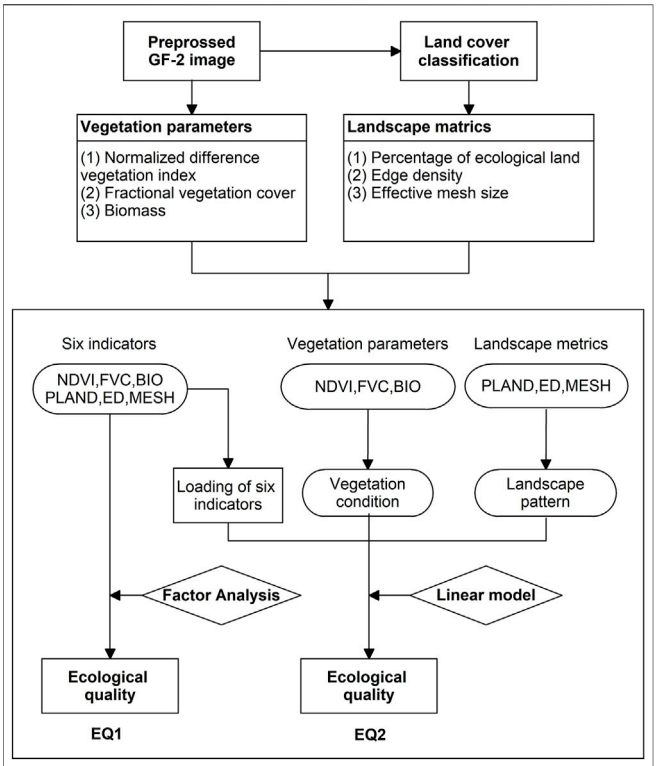

FIGURE 2 | The framework of ecological quality evaluation using GF-2 remote sensing data.

definition has been involved in the land use classification system, though has not yet been proposed as an explicit concept in ecological planning (Peng et al., 2017). Li et al. has defined urban ecological land as land "aimed at improving the quality of life of people in cities, protecting important ecosystems and habitats, maintaining and improving the natural and urban artificial ecological unit, and stabilizing the urban ecosystem services at a certain level." (Li et al., 2009).

Vegetation areas and water bodies are conventionally recognized as ecological lands, which are important to ecosystem service. Barren land is inferior ecological land that has potential to convert into vegetation, and is therefore also considered part of ecological land in this study. The impervious surface is not regarded as urban ecological land. Ecological land patches were extracted from land cover maps based on GF-2 image in this study. Landscape metrics were calculated at the class level for only urban ecological land.

3.2.2 Landscape Metrics Selection and Measurement

Landscape metrics have been extensively used for analyzing spatial patterns and differentiating urban land uses (O'Neill et al., 1991; Listopad et al., 2015). They can be calculated at various levels of patch, class or landscape, quantifying different aspects such as landscape composition and configuration. Many landscape metrics are highly correlated; therefore, correlation analysis should be performed in order to identify suitable metrics which are not significantly correlated with each other. At the same time, those metrics should represent as many aspects of landscape pattern as possible. From the perspective of heterogeneity level, landscape metrics are typically grouped according to the aspect of landscape pattern measured, such as area and edge metrics, shape metrics, core area metrics, contrast metrics, aggregation metrics, and diversity metrics (McGarigal, 2015). On the basis of correlation analysis, diversity and heterogeneity consideration, three landscape metrics, i.e., percentage of Landscape (PLAND), edge density (ED), and effective mesh size of ecological land (MESH), were selected as ecological quality evaluation indices, and calculated at grid level. Here, the grid size is defined as 100 m × 100 m, which is suitable to analyze the landscape pattern for ecological quality evaluation.

PLAND is a landscape composition metric, and quantifies the proportional abundance of ecological type in the landscape (McGarigal, 2015). Higher PLAND values represent better ecological quality. The PLAND is calculated as follows:

$$\mathrm{PLAND} = \mathrm{P_i} = \frac{\sum_{j=1}^{n} \mathrm{a_{ij}}}{\mathrm{A}} (100) \quad (1)$$

Pi = proportion of the landscape occupied by urban ecological land patches type (class) i. a_{ij} = area (m^2) of urban ecological land patch ij. A = total urban ecological land area (m^2).

ED was calculated as the sum of the lengths of all edge segments of urban ecological land patches, divided by the total landscape area (McGarigal, 2015). ED quantifies the fragmentation and shape complexity of the ecological land patches. Higher values of ED represent lower ecological quality. The ED is calculated as follows:

$$\mathrm{ED} = \frac{\sum_{k=1}^{m} e_{ik}}{A} (10,000) \quad (2)$$

e_{ik} = total length (m) of edge in landscape involving urban ecological land patches type (class) i, includes landscape boundary and background segments involving patch type i. A = total urban ecological land area (m^2).

MESH focuses on the patch area of ecological land (McGarigal, 2015). It represents the fragmentation degree, patch area and segmentation. MESH is calculated as the sum of the squared area of patches, divided by the total landscape area (i.e., 10,000 m^2). Higher values of MESH represent better ecological quality. The MESH is calculated as follows:

$$MESH = \frac{\sum_{j=1}^{n} a_{ij}^2}{A}\left(\frac{1}{10,000}\right) \quad (3)$$

a_{ij} = area (m^2) of urban ecological land patch ij. A = total urban ecological land area (m^2).

3.3 Vegetation Parameters Calculation

Three ecological indicators, i.e., Normalized Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI) and Biomass (BIO) from GF-2 images were involved to quantify vegetation condition.

NDVI is used to determine the density of green vegetation on the land by observing distinct colors (wavelengths) of the visible and near infrared (NIR) sunlight reflected by the plants. The NDVI is calculated as follows:

$$\mathrm{NDVI} = \frac{NIR - Red}{NIR + Red} \tag{4}$$

where Red and NIR refer to the reflectance of red and near-infrared bands in the GF-2 remote sensing data. NDVI values range from minus one to plus one, and no green leaves gives a value close to zero (NASA, 2017).

EVI improves sensitivity to high biomass regions and improved vegetation monitoring capability through a de-coupling of the canopy background signal and a reduction in atmospheric influences (Huete et al., 1999). It has shown significant improvements related to the analysis of environments composed by dense vegetation (Matsushita et al., 2007). According to the work of Liu and Huete, the EVI is defined as (Liu and Huete, 1995):

$$\mathrm{EVI} = G \times \frac{NIR - Red}{NIR + (C_1 \times Red - C_2 \times Blue) + \mathrm{L}} \tag{5}$$

where L is a soil adjustment factor, C_1 and C_2 are coefficients used to correct aerosol scattering in the red band using the blue band, Blue, Red, and NIR represent reflectance at the blue, red, and NIR bands, respectively. In general, $G = 2.5$, $C1 = 6.0$, $C2 = 7.5$, and $L = 1$ (Huete et al., 1997). The formula can be written as follow:

$$\mathrm{EVI} = 2.5 \times \frac{NIR - Red}{NIR + (6.0 \times Red - 7.5 \times Blue) + 1} \tag{6}$$

BIO is the mass of living biological organisms in an area or ecosystem at a given time. In remote sensing based methods, empirical algorithms have been widely used to explore the relationships between BIO and various vegetation indices (Zhang, 2007). Kong et al. (2016) developed a single curve regression model and multiple linear regression models for estimating the BIO value of the grassland, and for forest and shrub with R squared values of 71.9 and 52.8%, respectively (Kong et al., 2016). Their work has been applied to the study area of Fengning county of China, which has the similar vegetation cover to Haidian district of Beijing. Therefore, Kong et al.'s biomass estimation model was used in this research. Kong et al.'s model for estimating BIO value of the grassland was defined as:

$$\mathrm{BIO} = 0.272 \times RVI + 0.083 \tag{7}$$

where RVI refers to the Ratio Vegetation Index, the reflectance ratio of NIR and Red band. The model for estimating biomass value of forest and shrub is:

$$\mathrm{BIO} = 3683.07 \times GNDVI - 2254.634 \times GBNDVI - 1222.285 \tag{8}$$

where GNDVI refers to Green Normalized Difference Vegetation Index, and GBNDVI refers to Green and Blue Normalized Difference Vegetation Index. GNDVI and GBNDVI can be calculated using **Equation 9**, **10**.

$$\mathrm{GNDVI} = (NIR - Green)/(NIR + Green) \tag{9}$$

$$\mathrm{GBNDVI} = (NIR - (Green + Blue))/(NIR + (Green + Blue)) \tag{10}$$

In order to quantify the ecological quality in the study area, all selected vegetation parameters and landscape metrics were standardized to the range of 0 and 1 based on the standard deviation model.

3.4 Factor Analysis

In this study, factor analysis was used to determine the relationships between the ecological indicators (input variables) and output unobserved factors (urban ecological quality/vegetation condition/landscape pattern). The important output of the analysis is a table of factor loadings. Each item's loading represents how strongly the item is associated with the underlying factor. The absolute value of loadings should be 0.7 or higher to indicate the independent variables identified a priori are represented by a factor. The mathematical form of the factor analysis can be described as:

$$\begin{aligned} Y_1 &= \beta_{11}F_1 + \beta_{12}F_2 + \ldots + \beta_{1n}F_n + \varepsilon_1 \\ Y_2 &= \beta_{21}F_1 + \beta_{22}F_2 + \ldots + \beta_{2n}F_n + \varepsilon_2 \\ &\ldots \\ Y_l &= \beta_{l1}F_1 + \beta_{l2}F_2 + \ldots + \beta_{ln}F_n + \varepsilon_l \end{aligned} \tag{11}$$

where $Y_1, Y_2, \ldots, Y_l$ are the observed variables, $F_1, F_2, \ldots, F_n$ are the underlying factors, and β_{ln} are loadings. For instance, β_{11} is called the loading of variable Y_1 on factor F_1. The error ε serve to indicate that the hypothesized relationships are not exact. The maximum likelihood estimation, one of the most commonly used factor analysis procedures, was used in this paper to estimate factor loadings.

In this study, the factor analysis was applied three times to evaluation urban ecological quality from three aspects. First, the factor analysis was conducted to assess the sensitivities of six indicators to the overall ecological quality, and ecological quality was used as an unobserved factor. A loading matrix of six variables and a map of ecological quality were derived. Then, vegetation condition and landscape patterns were optimized using factor analysis, based on three vegetation parameters and three landscape metrics respectively. Lastly, based on the loading matrix, a linear function with two variables (vegetation condition and landscape pattern) was developed to provide more reliable evaluation of the ecological quality of the study area. For easy comparison, all values derived from factor analysis and linear model were normalized to the range of 0 and 1.

3.5 Linear Model

In order to present the contribution of landscape patterns more rationally without neglecting the influence of both aspects, a linear model was built to combine vegetation condition and landscape pattern as two variables, with their weights calculated from loadings of factor analysis.

Linear functions commonly arise from practical problems involving variables x and y with a linear relationship, that is,

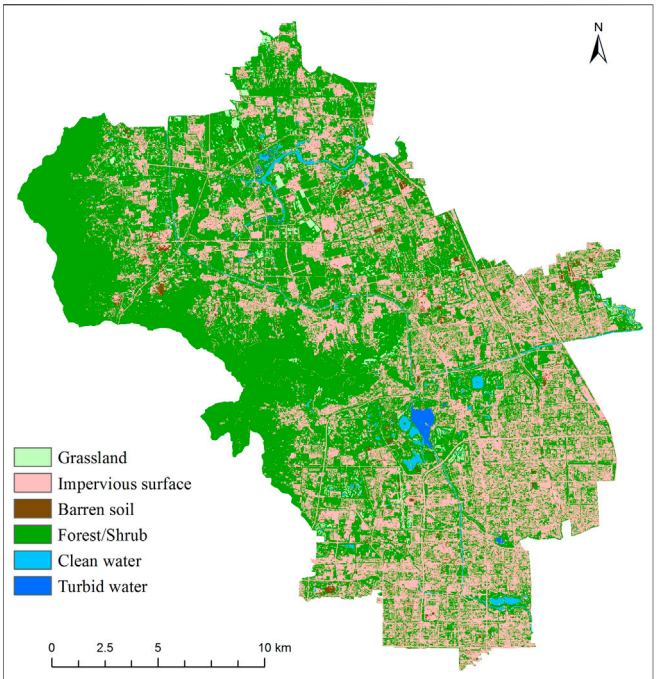

FIGURE 3 | Land cover classification of Haidian District, Beijing.

obeying a linear equation. The linear model used in this paper is deduced from the results of factor analysis, taking the form:

$$F = \frac{a \cdot X + b \cdot Y}{a + b} \tag{12}$$

where a and b are mean factor loadings of vegetation parameters and landscape metrics respectively, X and Y are two factors representing vegetation condition and landscape pattern, and F is the urban ecological quality.

According to the discussions on factor loadings of every variable for the effect of ecological quality, a linear model was built based on the six variable loadings and the output of vegetation condition and landscape pattern. The result of model calculation represents the more reasonable ecological quality.

4 RESULTS AND ANALYSIS

4.1 Land Cover Classification

Land cover types in Haidian District, involving grassland, forest, shrub, clean water, turbid water, barren soil, and impervious surface were identified from GF-2 images through object-oriented method (**Figure 3**). Accuracy evaluation using randomly selected samples showed that the overall accuracy of the classification reaches 84.1%, and the Kappa coefficient (Consistency check index) is 0.72. Large patches of forest/shrub and grassland are located in the northwest, while patches of impervious surface locate mostly in the highly urbanized east-southern regions.

4.2 Spatial Distribution of the Six Ecological Indicators

The spatial distribution of the three vegetation parameters showed similar patterns (**Figure 4**). Lower value patches indicated lower productivities and poor ecological effects due to poor vegetation coverage and low-level photosynthesis ratio. Spatial distribution patterns of NDVI, EVI, and BIO are also highly consistent among sub-districts (**Figure 4D**).

Landscape metrics of the ecological land in Haidian District were also calculated (**Figure 5**). PLAND (a) and MESH (c) showed lower values in the southeastern area, indicating higher level fragmentation of ecological land. Higher values of both metrics can be observed in the northwest, indicating better landscape pattern. ED (b) showed the opposite distribution of PLAND (a) and MESH (c). Mean metrics at sub-district scale reveal that PLAND and MESH values are highly consistent while ED values are negatively correlated with them (**Figure 5D**).

4.3 Urban Ecological Quality Evaluation

4.3.1 Results of Factor Analysis

Factor analysis was used in this study to analyze the relationship between input variables (ecological indicators) and the output unobserved factor. The results of three factor analyses correspond respectively to levels of ecological quality, vegetation condition and landscape pattern.

First, the output of unobserved factor was generated through factor analysis with all six input variables. The weight of each indicator was multiplied by its corresponding indicator value to generate an overall ecological quality value (**Figure 6**). Here, the urban ecological quality based on factor analysis with six indicators is abbreviated as EQ1.

Likewise, two other maps were generated using factor analysis (**Figure 7**), including the output unobserved factor of vegetation condition based on three vegetation parameters, and the output of landscape pattern using three landscape metrics as input. **Figure 7A** is the map of vegetation condition levels. **Figure 7B** is the map of landscape pattern levels. The hot spots in **Figure 7A**, in red, show the areas that have poor vegetation condition, and the hot spots in **Figure 7B**, in red, refer to the areas that have the lower-level landscape patterns.

Figure 7C showed two curves with green and red representing levels of vegetation condition and landscape pattern at sub-district scale. Most sub-districts have higher vegetation condition value than landscape pattern value, and only nine are in opposite situation. It is also indicated that Qinglongqiao has the biggest difference of 0.36 between values of vegetation condition and landscape pattern.

Figure 8 illustrates the contribution of vegetation condition and landscape pattern to urban ecological quality. The comparison revealed that EQ1 values were more likely to be consistent with those of vegetation condition, regardless of the values of landscape patter. This indicates that EQ1 from the 6-indicator factor analysis overlooked the contribution of landscape pattern, and may represent biased ecological quality of the study area.

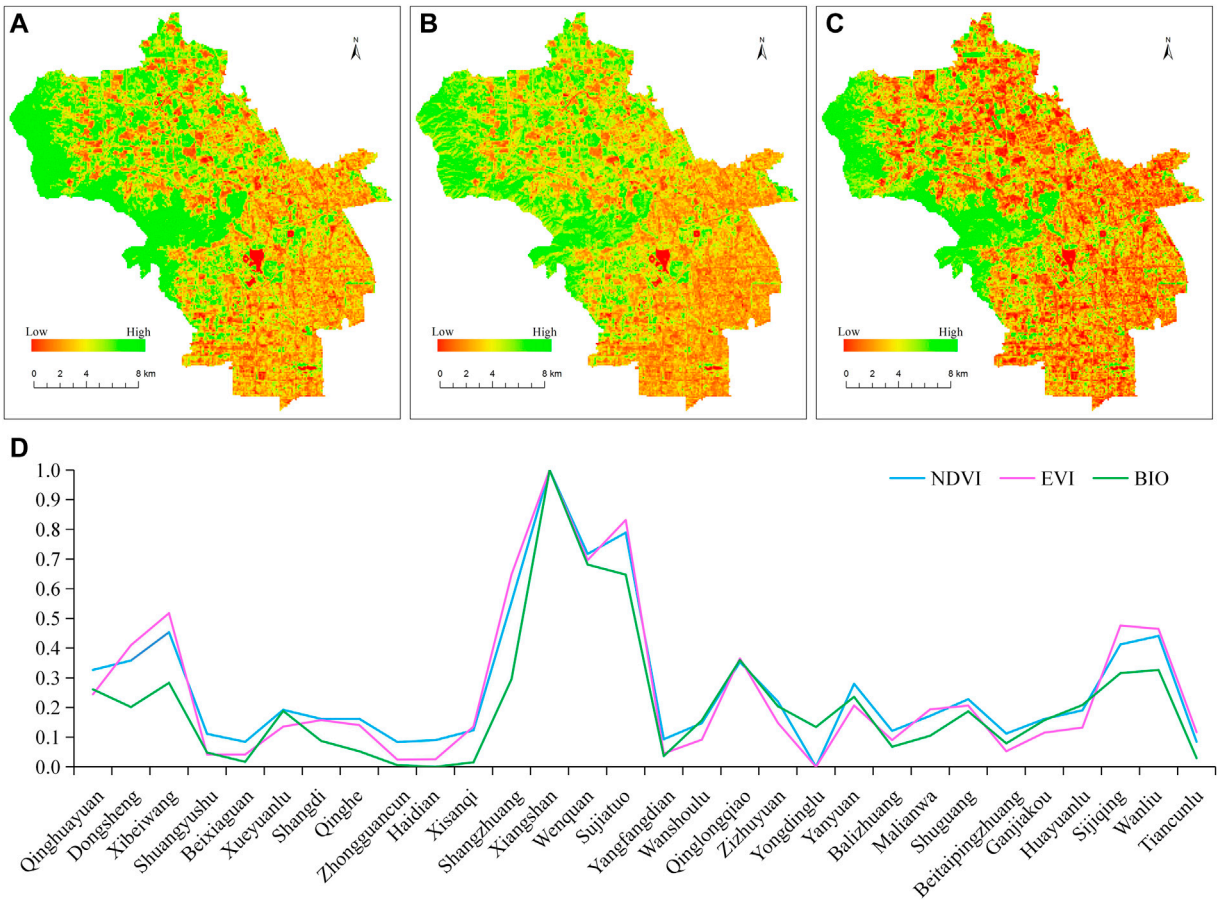

FIGURE 4 | The spatial distribution of NDVI **(A)**, EVI **(B)**, and BIO **(C)** values, and sub-district-wise normalized indices **(D)**.

4.3.2 Results of the Linear Model

It is indicated by the factor analysis results that, landscape pattern, one of the most important factors that affect urban ecological quality, should be incorporated in a different way in the ecological evaluation process. Therefore, this study adopted a linear model integrating vegetation condition and landscape pattern with their respective weights, instead of the 6-factor analysis. We take the average of the absolute loading values as a weight. Then, ecological quality was calculated based on the linear model which considers vegetation and landscape with their respective weights.

4.3.2.1 The Weights of Vegetation and Landscape

Equation 13 in factor analysis gives the relationship between ecological quality and the six observed variables (PLAND, ED, MESH, NDVI, EVI, and BIO) with indicator loadings shown.

$$\begin{aligned} PLAND &= 0.884 \times F_1 + \varepsilon_1 \\ ED &= -0.508 \times F_1 + \varepsilon_2 \\ MESH &= 0.790 \times F_1 + \varepsilon_3 \\ NDVI &= 0.993 \times F_1 + \varepsilon_4 \\ EVI &= 0.955 \times F_1 + \varepsilon_5 \\ BIO &= 0.878 \times F_1 + \varepsilon_6 \end{aligned} \tag{13}$$

These loadings represent how strongly this indicator is associated with the unobserved factor (EQ1). The importance of various ecological indicators can be determined through comparison of their loadings on ecological quality. PLAND, MESH, NDVI, EVI, and BIO had positive loading values higher than 0.7, indicating that those variables were positively correlated with EQ1 and therefore can be used as important indicators to evaluate the ecological quality. On the other hand, ED has a negative loading value, which indicates a negative correlation between ED and ecological quality. The absolute loading value of ED is lower than 0.7, showing weaker relationship between ED and the factor compared to other indicators.

According to **Equation 14**, the weight of vegetation condition 1) is 0.942, derived as the average of absolute loadings of three vegetation parameters. The weight of landscape pattern 2) is 0.727, the average of the absolute loadings of three landscape metrics.

$$\begin{aligned} a &= (0.878 + 0.993 + 0.955)/3 = 0.942 \\ b &= (0.884 + 0.508 + 0.790)/3 = 0.727 \end{aligned} \tag{14}$$

In the study, a linear combination model was adopted to give ecological quality evaluation considering both vegetation condition and landscape pattern. Here, the

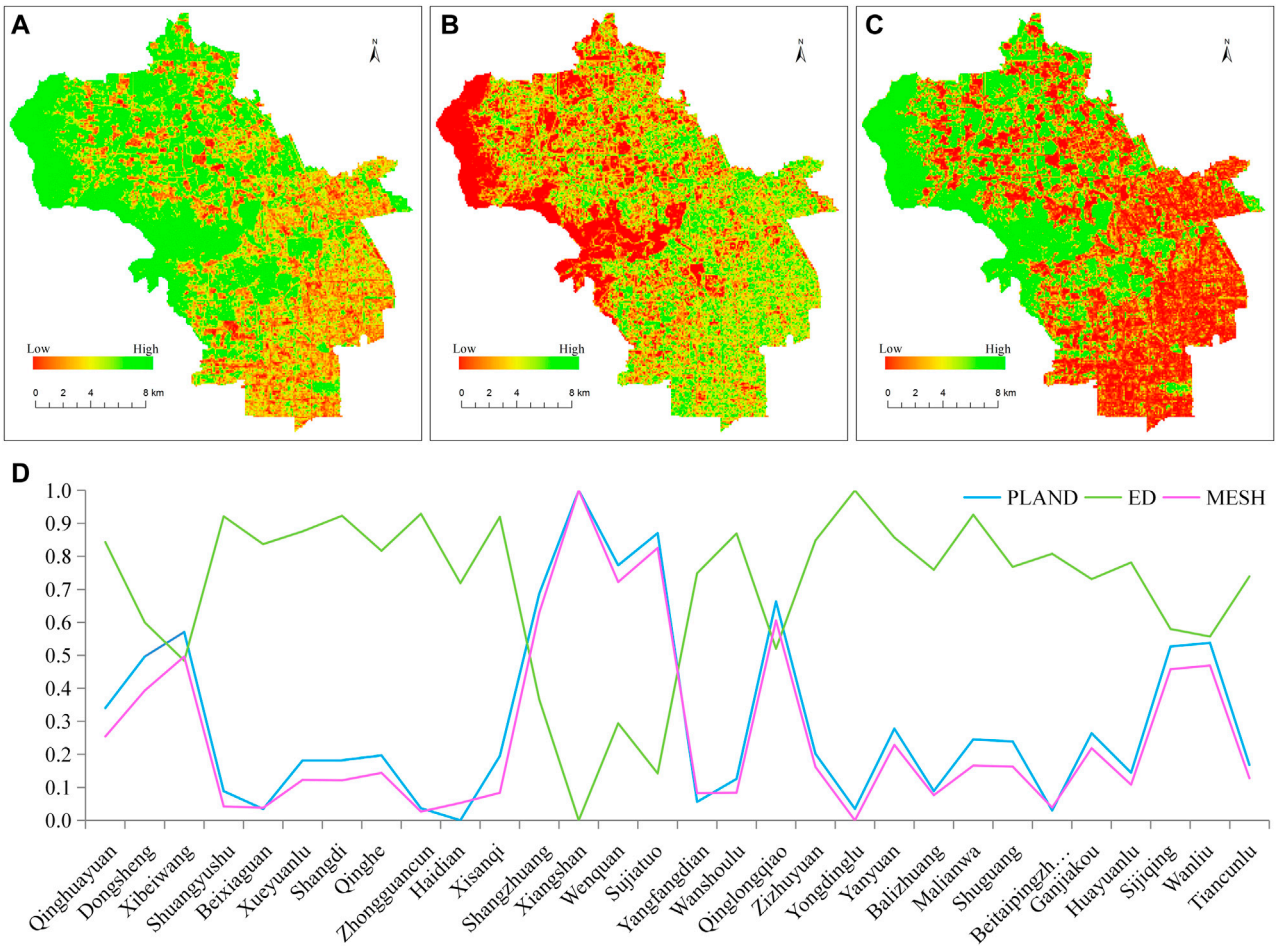

FIGURE 5 | The spatial distribution of landscape metrics PLAND **(A)**, ED **(B)**, and MESH **(C)**, and sub-district-wise normalized landscape metrics **(D)**.

urban ecological quality based on linear model with the wights of vegetation condition and landscape pattern is abbreviated as EQ2. Vegetation condition and landscape pattern from factor analysis (**Figure 7**) were used as two separate sets in the model. **Equation 15** shows the relationship between ecological quality and the two observed variables (vegetation condition X and landscape pattern Y).

$$EQ2 = \frac{a \cdot X + b \cdot Y}{a + b} = \frac{0.942 \cdot X + 0.727 \cdot Y}{1.669} \qquad (15)$$

4.3.2.2 Ecological Quality Map

Figure 9A showed the result of EQ2 in Haidian District using the linear model. Compared to **Figures 7A,B**, there exists good consistency between higher ecological quality and higher vegetation level and reasonable landscape patterns. The areas with high levels of ecological quality mostly present good vegetation conditions and reasonable landscape patterns. On the other hand, those areas with low levels of ecological quality are mainly caused by unreasonable landscape patterns combined with average levels of vegetation conditions. The hot spots, in red, in **Figure 9A** are the areas of poor ecological quality.

Figure 9B showed EQ1 from the factor analysis and EQ2 from the linear model at sub-districts scale. EQ1 indicated that the levels were more likely consistent with that of vegetation condition. Comparisons found the levels of EQ2, which considered both vegetation conditions and landscape patterns is effective in evaluating urban ecological quality.

From the perspective of spatial variation of ecological quality, it can be concluded that there exists a gradual change from better quality in the northwest to worse in the southeast (**Figure 9A**). The northwestern part of Haidian, including Xiangshan, Sujiaduo and Wenquanzhen sub-districts has higher ecological quality. Then the sub-districts to the east of the top three high level ecological quality sub-districts, i.e., Shangzhuang, Xibeiwang and Sijiqing, ranked in the second tier of higher ecological quality.

In the middle-eastern part, sub-districts have a medium level of ecological quality. Southeastern part of Haidian district, especially Yongdinglu, Zhongguancun, Beixiaguan, Haidian

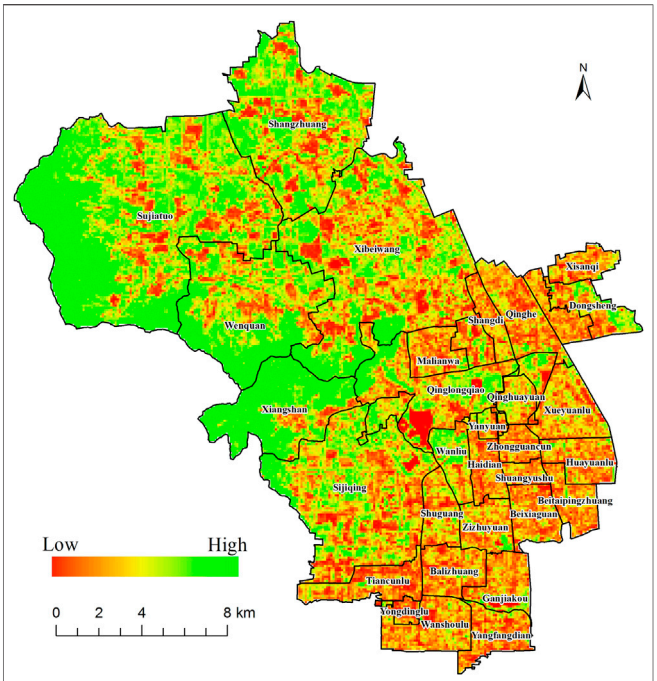

FIGURE 6 | Map of ecological quality (EQ1) using six-indicator-based factor analysis in Haidian district.

and Beitaipingzhuang sub-districts, has lower level of ecological quality. The other areas exhibited mixed ecological quality patterns.

With that, Beijing government released a general city plan for 2016 to 2035. The plan emphasizes on removing non-capital functions and solving "big city disease". It was also mentioned that the green area of Beijing will grow from 41.6 to 44% by year 2020 (Beijing general plan (2016-2035)). Northern region with the rich natural and human resources is the expanding area of Zhongguancun Science Park. In Haidian district overall development and planning, northern region is considered as a tourist area and ecological barrier, therefore urbanization is strictly under control and hardly permitted in these areas.

5 DISCUSSION

5.1 Vegetation Plays the Decisive Role for Ecological Quality

Over the past several decades, vegetation has been identified as an important contributor to urban environment and ecological service. Urban vegetation can contribute to quality of life at the most fundamental level through biodiversity protection, water quality levels, and maintenance of ecological processes and life-support systems (Carne, 1994). Interactions between vegetation and ecological quality were described by experiments (Fennessy et al., 2002; Pu et al., 2008; Giménez et al., 2017) or by models (Rajabov, 2009; Giménez et al., 2016; Kuipers et al., 2016). In this research, vegetation played a key role in ecological evaluation as shown by factor analysis results. **Equation 14** showed that the three vegetation parameters (NDVI, EVI, and BIO) have positive and much higher loading values. This is consistent with the work that other researchers have done (Aksoy, 2010; Kuipers et al., 2016). Among the three vegetation parameters, NDVI represents the richness of vegetation and plays the decisive role for ecological quality, resulting in a push to improve urban environments by planting more trees and grass.

5.2 Landscape Metrics Contribute Less but Are Indispensable to Ecological Quality

Landscape structure has an important influence on a wide range of ecological patterns and processes, and landscape metrics are common tools to assess these relations under the matrix-corridor-patch model (Forman, 1995; Turner et al., 2001). Landscape patterns become increasingly crucial in ecological quality evaluation in urban areas because of the gradually fragmented ecological land. The importance of landscape patterns was underestimated as shown in the 6-indicator factor analysis results, while actually landscape metrics such as patch size, patch shape and distribution of urban ecological land play a decisive role in defining their ecological and landscape functions (Kong et al., 2007).

In light of this issue, the study evaluated ecological quality based on two separate groups of indicators, e.g., vegetation parameters and landscape pattern. Factor analysis was performed separately to generate their ecological effects, and then a linear combination model was used to evaluate ecological quality integrating the effects of both indicator groups. In this way can landscape metrics provide proper contribution to ecological quality evaluation. The results are expected to emphasize the contribution of landscape pattern to urban ecological quality in future studies. These can help decision-makers better understand cause and effect relationships between the influencing factors and urban ecological quality.

Figure 10 shows three maps of vegetation condition (a), landscape pattern (b), and EQ2 (c) of Qinglongqiao sub-district. In **Figure 10A**, the hot spots in the left-bottom indicated the area with extremely poor vegetation condition. However, **Figure 10C** showed that the ecological quality is not in the lowest level in the same area. This result was caused by the quite good landscape pattern with green color (**Figure 10B**). Therefore, the method proposed in this study synthesizes together landscape metrics and vegetation parameters in a reasonable way, and is expected to result in more convinced ecological results.

Furthermore, landscape metrics have been used to assess ecological patterns and processes. In this study, higher PLAND and MESH, and lower ED represent better ecological quality. PLAND and MESH have positive influence while ED has negative one. Thus, patch size and shape present the fundamental landscape indicator in assessing the urban ecological quality. For example, a connected green network has higher amenity value than smaller and fragmented ones. It can enhance the amenity values of green spaces and provide more choices to residents (Jim and Chen, 2003). Although ED

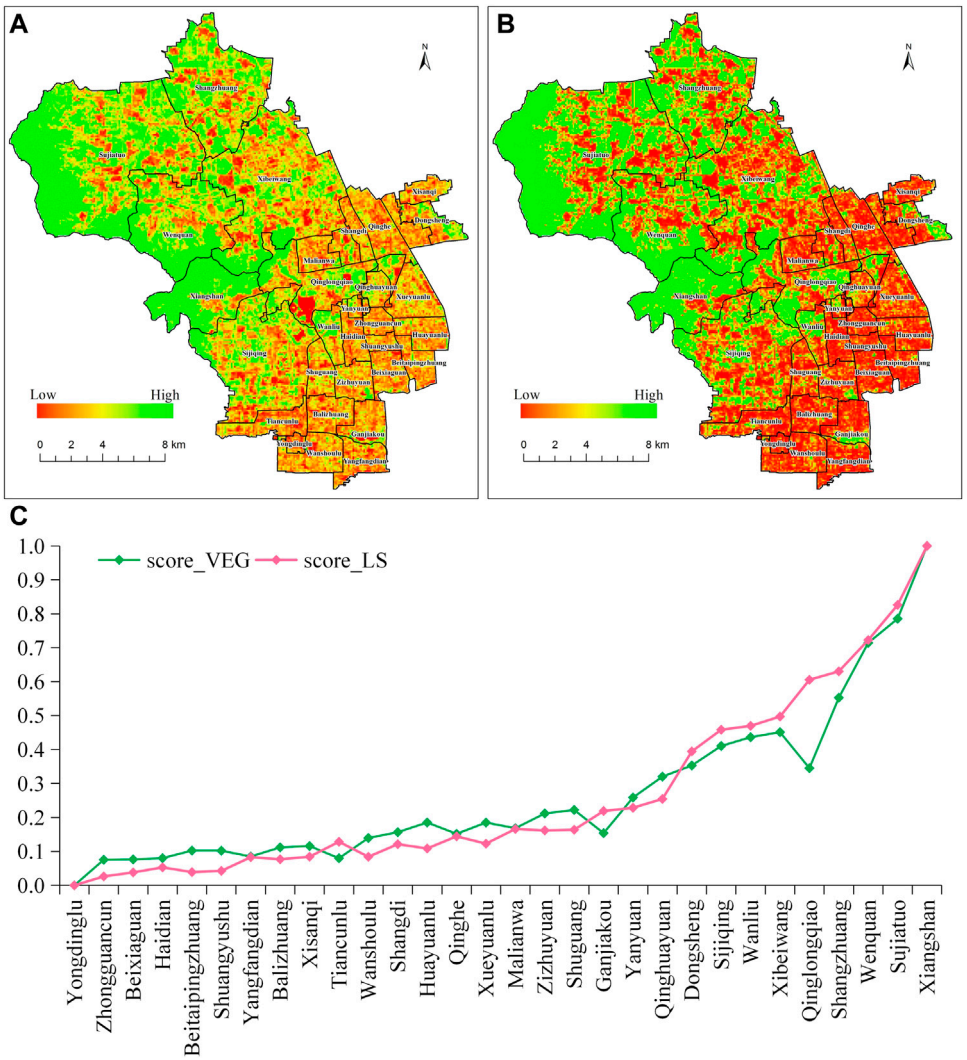

FIGURE 7 | Map of Vegetation condition **(A)**, Landscape pattern **(B)**, and the sub-district-wise values in Haidian district **(C)**.

has the least weight, it also can contribute to the complexity of patch edge and consequently the interactions between green spaces and ecological quality according to edge effect (Harper et al., 2005).

5.3 Advantages and Limitations of the Proposed Method

Remotely sensed images, especially high spatial resolution images provide several advantages in ecological land mapping and quantitative evaluation of ecological quality. Object-oriented image classification can provide highly accurate land cover map that can be used to generate ecological indicators with reliable accuracy, especially when meter level spatial resolution images were involved. The proposed method described the detailed differences of ecological quality at regional scales, which account for the spatial heterogeneity in evaluating urban ecological quality.

The shortage of the method lies in the spatial scale of the evaluation. Vegetation parameters can be derived from high spatial resolution images and UAV (Unmanned Aerial Vehicle) based remote sensing techniques. However, landscape metrics were calculated at a grid pixel level, which causes a coarser spatial resolution of the landscape metrics than vegetation parameters. Due to the availability of data, remote sensing-based vegetation and landscape indicators were considered in this study. Other indicators, e.g., air quality, vegetation diversity could be discussed in future work.

5.4 Policy Optimization by Combining Ecological Land Spatial Pattern

Increasing the area of ecological land (such as water bodies and especially vegetation) is a traditional way to improve ecological quality. With the rapid increase of population and expansion of built-up area, ecological land faces a growing risk of

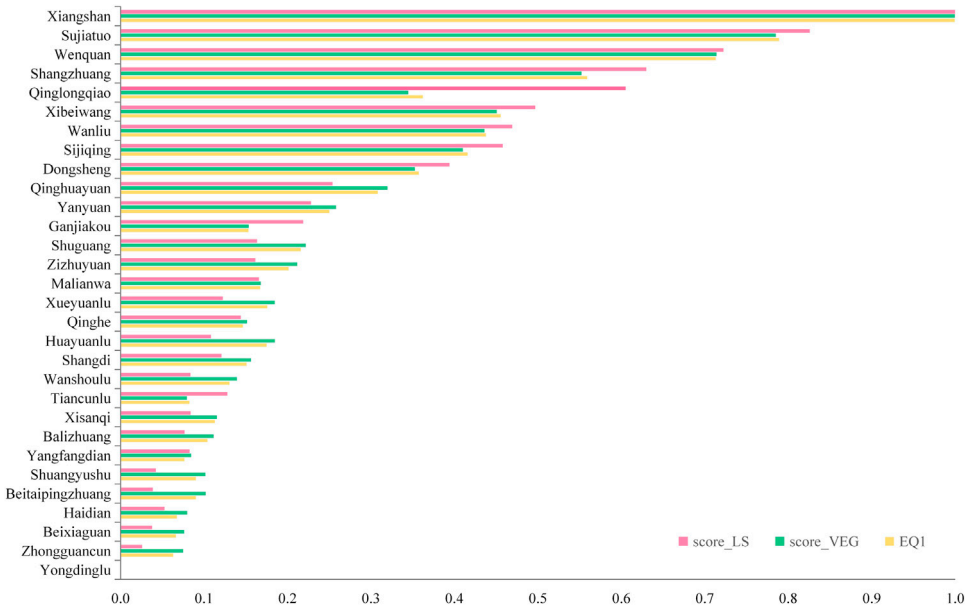

FIGURE 8 | Ecological quality (EQ1), vegetation condition and landscape pattern derived from factor analysis at sub-district scale.

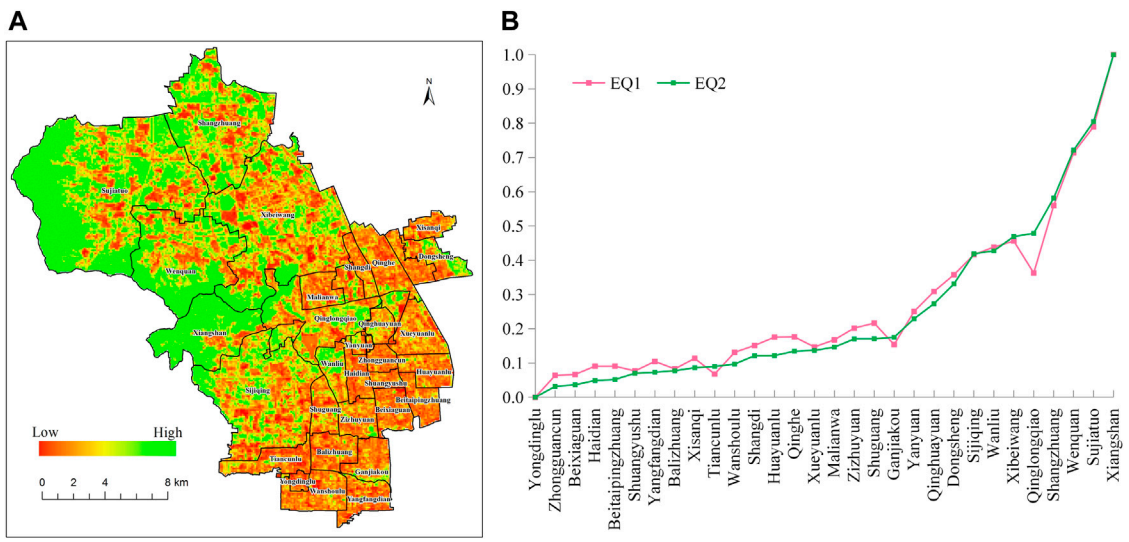

FIGURE 9 | Results of ecological quality evaluation (EQ2) with the linear model **(A)**, comparison to results with two methods **(B)**.

fragmentation. Therefore, the ecological quality evaluation method proposed in this paper, could provide decision-makers and the general public with specific information on the current status of vegetation condition and their spatial pattern, from the ecological quality view. Urban ecological quality could be enhanced through spatial pattern optimization as well as ecological land area expansion.

Figure 11 showed three lines with different colors to represent the levels of EQ2, vegetation condition and landscape pattern for each sub-district in Hiadian District. Xiangshan has the highest ecological quality level due to the most reasonable landscape pattern and the best vegetation condition, while Yongdinglu has the lowest. Either poor vegetation or unreasonable landscape patterns can result in low levels of ecological quality. The sub-districts with lower ecological quality can be potentially improved through optimization of the landscape pattern of ecological land or improvement of vegetation vigor. Hence, efficient policies to either improve vegetation vigor, increase vegetation area, or optimize spatial pattern could be drawn from **Figure 11**. For

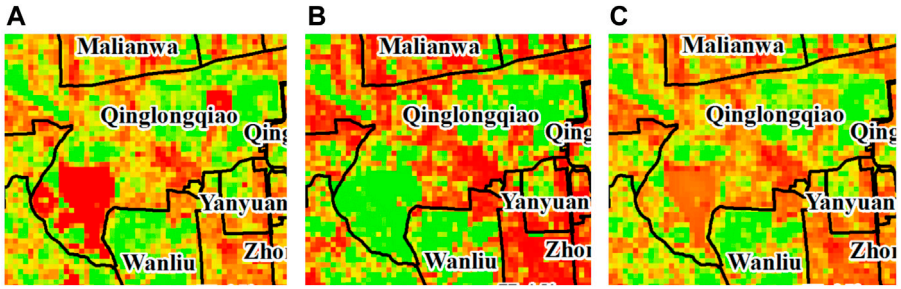

FIGURE 10 | Maps of vegetation condition **(A)**, landscape pattern **(B)**, and EQ2 **(C)** of Qinglongqiao sub-district.

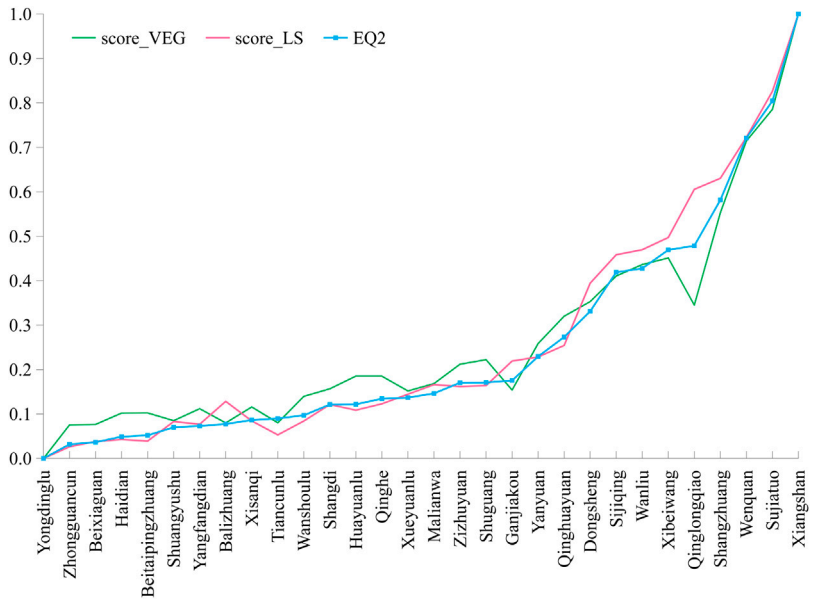

FIGURE 11 | The normalized results of landscape pattern levels, vegetation condition levels and ecological quality.

19 sub-districts including Zhongguancun, which have higher vegetation condition value than landscape pattern value, the better way to improve the ecological quality of those areas is to make the landscape pattern more reasonable. Correspondingly, for the other nine sub-districts including Qinglongqiao, which have higher landscape pattern value than vegetation condition value, improving the vegetation condition of ecological land is considered a better solution to achieve higher ecological quality of this area.

Another result shown in **Figure 11** is that ecological quality is greatly affected by landscape pattern in those sub-districts with lower levels (score_VEG <0.4) of vegetation condition. Due to building protection and cultural tradition, it is difficult to improve vegetation condition in the above-mentioned sub-districts. Therefore, improving the level of landscape pattern is an effective way to make ecological quality better.

With the "ecological city" initiative launched as an integral part of Beijing and local government strategy, a plan is developed to build a compound ecological security pattern, according to the actual situation of Haidian district, The revealed performances and pattern of ecological quality will be useful to understand vegetation condition and landscape pattern as the sensitive influencing factors of ecological environment, and to improve urban land use and ecology management decisions. In the northern regions with mountains and wetlands, vegetation conditions need to be maintained and improved. However, in the southern regions with built-up land, ecological land patches are smaller and more fragmented. The better way to improve ecological quality is to optimize the landscape pattern instead of vegetation condition.

6 CONCLUSION

This paper proposed a novel method to evaluate urban ecological quality integrating vegetation condition and landscape pattern metrics from remotely sensed images. The paper gives more reliable ecological quality mapping in Haidian

District, Beijing, China. The following conclusions can be drawn in the study.

- High-resolution remote-sensing-based method enables the development of ecological indicators with high spatial accuracy, thus better describing the detailed differences of ecological quality at regional scales, and could account for the spatial heterogeneity in evaluating urban ecological quality.
- Landscape patterns should be considered in the ecological evaluation process. Ecological quality evaluation based on only vegetation condition factors, e.g., NDVI, EVI, and BIO, cannot reflect the reality in Haidian District, Beijing, leading to the situation that areas with high levels of vegetation condition are more likely to be associated with higher level ecological quality, regardless of the landscape pattern level. This revealed the big shortcoming of the evaluation for neglecting the fragmentation of the ecological land and against the common sense of ecological quality.
- Ecological quality evaluation integrating vegetation condition factors and landscape metrics, e.g., PLAND, ED, and MESH, could result in more reliable and effective estimation. Areas with high levels of vegetation condition and reasonable landscape pattern lead to higher ecological quality results. On the other hand, those areas with low levels of ecological quality are mainly caused by unreasonable landscape patterns combined with below-average levels of vegetation condition.
- Ecological quality evaluation based on vegetation condition factors and landscape metrics could result in more spatially specific ecological quality level, and reflect the spatial variation in the study area. Analysis of the ecological quality evaluation results found that there exists a gradual degradation of ecological quality from the northwest to the southeast along, consistent with the intensified urbanization.
- Ecological quality is greatly affected by landscape pattern when the levels of vegetation condition is below 0.4. For optimization of ecological quality in those sub-districts, the government could make great efforts to improve landscape pattern such as creating more small patches of green space or water bodies, as well as strengthening the greening pattern of the built-up area.

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because Only used for scientific research where the study area is Beijing, and people are restricted to students and researchers in China. Requests to access the datasets should be directed to HH, huanghp@aircas.ac.cn.

AUTHOR CONTRIBUTIONS

HH conceived and designed the experiments, conducted the data analysis and wrote the manuscript; QL developed the methodology; YZ performed the experiments; QL and YZ edited the manuscript.

FUNDING

This research was funded by National Key R&D Program of China (2017YFB0503800).

REFERENCES

Aksoy, Y. (2010). Examining the Ecological Quality of Küçükçekmece District Parks in Istanbul in Terms of Permeability and Natural Vegetation. *Ekoloji* 19, 181–189. doi:10.5053/ekoloji.2010.7422

Allender, S., Lacey, B., Webster, P., Rayner, M., Deepa, M., Scarborough, P., et al. (2010). Level of Urbanization and Noncommunicable Disease Risk Factors in Tamil Nadu, India. *Bull. World Health Organ.* 88 (4), 297–304. doi:10.2471/blt.09.065847

Arlt, G., and Lehmann, I. (2005). *Ecosystem Land Service Method System–Analysis and Evaluation of Urban Areas in Dresden*. IÖR.

Behling, R., Bochow, M., Foerster, S., Roessner, S., and Kaufmann, H. (2015). Automated GIS-Based Derivation of Urban Ecological Indicators Using Hyperspectral Remote Sensing and Height Information. *Ecol. indicators* 48, 218–234. doi:10.1016/j.ecolind.2014.08.003

Bolund, P., and Hunhammar, S. (1999). Ecosystem Services in Urban Areas. *Ecol. Econ.* 29, 293–301. doi:10.1016/s0921-8009(99)00013-0

Botequilha Leitão, A., and Ahern, J. (2002). Applying Landscape Ecological Concepts and Metrics in Sustainable Landscape Planning. *Landscape Urban Plann.* 59, 65–93. doi:10.1016/s0169-2046(02)00005-1

Bureau of Statistics of Haidian District (2020). *National Economic and Social Development Statistical Bulletin of Haidian District in 2019*. Beijing, China: Haidian District Government.

Cadenasso, M. L., Pickett, S. T. A., and Schwarz, K. (2007). Spatial Heterogeneity in Urban Ecosystems: Reconceptualizing Land Cover and a Framework for Classification. *Front. Ecol. Environ.* 5, 80–88. doi:10.1890/1540-9295(2007)5[80:shiuer]2.0.co;2

Carne, R. J. (1994). "Urban Vegetation: Ecological and Social Value," in Proceedings of the 1994 National Greening Australia Conference, 4–6.

Chinese Academy of Social Sciences (2019). *The 17th Report on the Competitiveness of Chinese Cities*.

Du, H., Zhou, F., Li, C., Cai, W., Jiang, H., and Cai, Y. (2020). Analysis of the Impact of Land Use on Spatiotemporal Patterns of Surface Urban Heat Island in Rapid Urbanization, a Case Study of Shanghai, China. *Sustainability* 12 (3), 1–17. doi:10.3390/su12031171

Eliasson, I., and Svensson, M. K. (2003). Spatial Air Temperature Variations and Urban Land Use - a Statistical Approach. *Met. Apps* 10, 135–149. doi:10.1017/s1350482703002056

Elmqvist, T. (2013). *Urbanization, Biodiversity and Ecosystem Services : Challenges and Opportunities : A Global Assessment : A Part of the Cities and Biodiversity Outlook Project*. Springer.

Estoque, R. C., and Murayama, Y. (2013). Landscape Pattern and Ecosystem Service Value Changes: Implications for Environmental Sustainability Planning for the Rapidly Urbanizing Summer Capital of the Philippines. *Landscape Urban Plann.* 116, 60–72. doi:10.1016/j.landurbplan.2013.04.008

Fennessy, M. S., Gernes, M., Mack, J. J., and Wardrop, D. H. (2002). *Methods for Evaluating Wetland Condition: Using Vegetation to Assess Environmental Conditions in Wetlands*. Washington, DC: Environmental Protection Agency, Office of Water. EPA-822-R-02-020 U.S. 2002.

Forman, R. T. T. (1995). Some General Principles of Landscape and Regional Ecology. *Landscape Ecol.* 10, 133–142. doi:10.1007/bf00133027

Fuller, R. A., Irvine, K. N., Devine-Wright, P., Warren, P. H., and Gaston, K. J. (2007). Psychological Benefits of Greenspace Increase with Biodiversity. *Biol. Lett.* 3, 390–394. doi:10.1098/rsbl.2007.0149

Giménez, M. G., de Jong, R., Della Peruta, R., Keller, A., and Schaepman, M. E. (2017). Determination of Grassland Use Intensity Based on Multi-Temporal Remote Sensing Data and Ecological Indicators. *Remote Sensing Environ.* 198, 126–139.

Giménez, M. G., Della Peruta, R., de Jong, R., Keller, A., and Schaepman, M. E. (2016). Spatial Differentiation of Arable Land and Permanent Grassland to Improve a Land Management Model for Nutrient Balancing. *IEEE J. Selected Top. Appl. Earth Observations Remote Sensing* 9, 5655–5665.

Gómez-Baggethun, E., and Barton, D. N. (2013). Classifying and Valuing Ecosystem Services for Urban Planning. *Ecol. Econ.* 86, 235–245.

Gonzalez, M., Ladet, S., Deconchat, M., Cabanettes, A., Alard, D., and Balent, G. (2010). Relative Contribution of Edge and Interior Zones to Patch Size Effect on Species Richness: An Example for Woody Plants. *For. Ecol. Manage.* 259, 266–274. doi:10.1016/j.foreco.2009.10.010

Harper, K. A., Macdonald, S. E., Burton, P. J., Chen, J., Brosofske, K. D., Saunders, S. C., et al. (2005). Edge Influence on forest Structure and Composition in Fragmented Landscapes. *Conservation Biol.* 19, 768–782. [CrossRef]. doi:10.1111/j.1523-1739.2005.00045.x

Henry, J. A., and Dicks, S. E. (1987). Association of Urban Temperatures with Land Use and Surface Materials. *Landscape Urban Plann.* 14, 21–29. doi:10.1016/0169-2046(87)90003-x

Huete, A., Justice, C., and Leeuwen, W. (1999). MODIS Vegetation Index (MOD 13) Algorithm Theoretical Basis. *Document Version* 3.

Huete, A., Liu, H. Q., Batchily, K., and van Leeuwen, W. (1997). A Comparison of Vegetation Indices Over a Global Set of TM Images for EOS-MODIS. *Remote Sensing Environ.* 59, 440–451. doi:10.1016/s0034-4257(96)00112-5

Liu, H. Q., and Huete, A. (1995). A Feedback Based Modification of the NDVI to Minimize Canopy Background and Atmospheric Noise. *IEEE Trans. Geosci. Remote Sensing* 33, 457–465. doi:10.1109/36.377946

Jim, C. Y., and Chen, S. S. (2003). Comprehensive Greenspace Planning Based on Landscape Ecology Principles in Compact Nanjing City, China. *Landscape Urban Plann.* 65, 95–116. doi:10.1016/s0169-2046(02)00244-x

Kaufmann, R. K., Seto, K. C., Schneider, A., Liu, Z., Zhou, L., and Wang, W. (2007). Climate Response to Rapid Urban Growth: Evidence of a Human-Induced Precipitation Deficit. *J. Clim.* 20, 2299–2306. doi:10.1175/jcli4109.1

Kong, F., Li, X., and Bai, Y. (2016). The Estimation of Aboveground Vegetation Carbon Storage Based on GF-1 WFV Satellite Images in Fengning County. *Resour. Sci.* 38, 1054–1064.

Kong, F., Yin, H., and Nakagoshi, N. (2007). Using GIS and Landscape Metrics in the Hedonic Price Modeling of the Amenity Value of Urban Green Space: A Case Study in Jinan City, China. *Landscape Urban Plann.* 79, 240–252. doi:10.1016/j.landurbplan.2006.02.013

Kuipers, H. J. G., Netten, J. J. C., and Hendriks, A. J. (2016). Explaining Ecological Quality by Using Variable Vegetation Densities in Hydrological Modelling. *Aquat. Bot.* 133, 38–44. [CrossRef]. doi:10.1016/j.aquabot.2016.05.008

Lakes, T., and Kim, H.-O. (2012). The Urban Environmental Indicator "Biotope Area Ratio"-An Enhanced Approach to Assess and Manage the Urban Ecosystem Services Using High Resolution Remote-Sensing. *Ecol. Indicators* 13, 93–103. doi:10.1016/j.ecolind.2011.05.016

Li, F., Wang, W., Zhang, X., Wang, R., and Zhao, D. (2009). Urban Ecological Land and its Optimization of Service Function. China Population. *Resour. Environ.* 19, 343–347.

Li, Y., Zhu, X., Sun, X., and Wang, F. (2010). Landscape Effects of Environmental Impact on Bay-Area Wetlands Under Rapid Urban Expansion and Development Policy: A Case Study of Lianyungang, China. *Landscape Urban Plann.* 94, 218–227. doi:10.1016/j.landurbplan.2009.10.006

Listopad, C. M. C. S., Masters, R. E., Drake, J., Weishampel, J., and Branquinho, C. (2015). Structural Diversity Indices Based on Airborne LiDAR as Ecological Indicators for Managing Highly Dynamic Landscapes. *Ecol. Indicators* 57, 268–279. doi:10.1016/j.ecolind.2015.04.017

Matsushita, B., Yang, W., Chen, J., Onda, Y., and Qiu, G. (2007). Sensitivity of the Enhanced Vegetation index (EVI) and Normalized Difference Vegetation index (NDVI) to Topographic Effects: A Case Study in High-Density Cypress Forest. *Sensors* 7, 2636–2651. doi:10.3390/s7112636

McGarigal, K. (2015). Spatial Pattern Analysis Program for Categorical and Continuous Maps. *Fragstats Help* 4, 2. [CrossRef].

NASA (2017). *Measuring Vegetation (NDVI & EVI): Normalized Difference Vegetation Index*. NDVI.

National Bureau of Statistics of China (2021). *Press Conference on the Operation of the National Economy in 2021*. Chinese Government 2022.

O'Neill, R. V., Gardner, R. H., Milne, B. T., Turner, M. G., and Jackson, B. (1991). *Heterogeneity and Spatial Hierarchies, Ecological Heterogeneity*. Springer, 85–96.

Owen, D. (2010). *Urbanization, Water Quality, and the Regulated Landscape*. Social ence Electronic Publishing.

Peng, J., Xie, P., Liu, Y., and Ma, J. (2016). Urban Thermal Environment Dynamics and Associated Landscape Pattern Factors: A Case Study in the Beijing Metropolitan Region. *Remote Sensing Environ.* 173, 145–155. doi:10.1016/j.rse.2015.11.027

Peng, J., Zhao, M., Guo, X., Pan, Y., and Liu, Y. (2017). Spatial-Temporal Dynamics and Associated Driving Forces of Urban Ecological Land: A Case Study in Shenzhen City, China. *Habitat Int.* 60, 81–90. doi:10.1016/j.habitatint.2016.12.005

Pettorelli, N., Laurance, W. F., O'Brien, T. G., Wegmann, M., Nagendra, H., and Turner, W. (2014). Satellite Remote Sensing for Applied Ecologists: Opportunities and Challenges. *J. Appl. Ecol.* 51, 839–848. doi:10.1111/1365-2664.12261

Pu, R., Gong, P., Tian, Y., Miao, X., Carruthers, R. I., and Anderson, G. L. (2008). Using Classification and NDVI Differencing Methods for Monitoring Sparse Vegetation Coverage: A Case Study of Saltcedar in Nevada, USA. *Int. J. Remote Sensing* 29, 3987–4011. doi:10.1080/01431160801908095

Rajabov, T. (2009). "Ecological Assessment of Spatio-Temporal Changes of Vegetation in Response to Biosphere Effects in Semi-Arid Rangelands of Uzbekistan," in *Land Restoration Training Programme* (Reykjavik, Iceland). http://www.unulrt.is/static/fellows/document/rajabov_t.pdf.

Ramakrishnan, P. (2014). An Approach to Classify the Object from the Satellite Image Using Image Analysis Tool. *IJIRST* 1 (4), 2349–6010.

Rose, R. A., Byler, D., Eastman, J. R., Fleishman, E., Geller, G., Goetz, S., et al. (2015). Ten Ways Remote Sensing Can Contribute to Conservation. *Conservation Biol.* 29, 350–359. doi:10.1111/cobi.12397

Sarrat, C., Lemonsu, A., Masson, V., and Guedalia, D. (2006). Impact of Urban Heat Island on Regional Atmospheric Pollution. *Atmos. Environ.* 40, 1743–1758. doi:10.1016/j.atmosenv.2005.11.037

Su, S., Xiao, R., Jiang, Z., and Zhang, Y. (2012). Characterizing Landscape Pattern and Ecosystem Service Value Changes for Urbanization Impacts at an Eco-Regional Scale. *Appl. Geogr.* 34, 295–305. doi:10.1016/j.apgeog.2011.12.001

Turner, M. G., Gardner, R. H., O'neill, R. V., Gardner, R. H., and O'Neill, R. V. (2001). *Landscape Ecology in Theory and Practice*. Springer. [CrossRef].

Uy, P. D., and Nakagoshi, N. (2008). Application of Land Suitability Analysis and Landscape Ecology to Urban Greenspace Planning in Hanoi, Vietnam. *Urban For. Urban Green.* 7, 25–40. doi:10.1016/j.ufug.2007.09.002

Wang, Z., Li, J., Yang, Y., Li, H., Wu, M., Wang, K., et al. (2017). *Quantitative Assess the Dynamic Change of Urban Ecological Environment Based on Remote Sensing—A Case Study in Yixing City, Jiangsu Province*. Journal of Ningxia University.

Xu, H. (2013). A Remote Sensing Urban Ecological Index and its Application. *Acta Ecologica Sinica* 33, 7853–7862.

Zhang, C. L., Chen, F., Miao, S. G., Li, Q. C., Xia, X. A., and Xuan, C. Y. (2009). Impacts of Urban Expansion and Future Green Planting on Summer Precipitation in the Beijing Metropolitan Area. *J. Geophys. Res.* 114, D02116. doi:10.1029/2008jd010328

Zhang, F., Wang, Y., Ma, X., Wang, Y., Yang, G., and Zhu, L. (2019). Evaluation of Resources and Environmental Carrying Capacity of 36 Large Cities in China

Based on a Support-Pressure Coupling Mechanism. *Sci. Total Environ.* 688, 838–854. doi:10.1016/j.scitotenv.2019.06.247

Zhang, H. (2007). Applications of RS Technology in Forest Biomass Research. *World For. Res.* 20, 30–34.

Zhou, T., Wu, J., and Peng, S. (2012). Assessing the Effects of Landscape Pattern on River Water Quality at Multiple Scales: A Case Study of the Dongjiang River Watershed, China. *Ecol. Indicators* 23, 166–175. doi:10.1016/j.ecolind.2012.03.013

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.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership