

An Enhanced Geographically and Temporally Weighted Neural Network for Remote Sensing Estimation of Surface Ozone

Tongwen Li^{ID}, Jingan Wu^{ID}, Jiajia Chen, and Huanfeng Shen^{ID}, *Senior Member, IEEE*

Abstract—Surface ozone (O₃) pollution is a severe environmental problem that endangers human health. It is necessary to obtain high spatiotemporal resolution O₃ data to provide support for pollution monitoring and prevention. For this purpose, this study makes comprehensive use of remote sensing data, reanalysis data, and ground station observations and develops an enhanced geographically and temporally weighted neural network (EGTWNN) model to acquire high spatial and temporal resolutions of O₃ data. The EGTWNN model is nested by two neural networks (NNs). The first NN automatically learns the spatiotemporal proximity relationship to obtain spatiotemporal weights for the samples, and the spatiotemporal weights are then inputted into the second NN to conduct weighted modeling of the relationship between O₃ and influencing variables. The contribution of the proposed model is that the first NN replaces the traditional empirical weighting method and represents the spatiotemporal proximity relationship more accurately to improve estimation accuracy. Results indicate that the cross-validation R^2 and the root mean square error (RMSE) of EGTWNN are 0.81 and 21.24 $\mu\text{g}/\text{m}^3$, respectively, which are increased by 0.02 and decreased by $\sim 1 \mu\text{g}/\text{m}^3$ relative to those of the traditional empirical weighting method-based geographically and temporally weighted NN model. The results also show that, compared with the geographically and temporally weighted regression model, the proposed model achieves superior performance. In addition, the spatiotemporal weights obtained by the first NN of EGTWNN are highly consistent with those obtained by the traditional empirical weighting method, indicating that the results of NNs are highly interpretable.

Index Terms—Geographically and temporally weighted neural network (NN), nested NN modeling, ozone remote sensing.

I. INTRODUCTION

SURFACE ozone (O₃) pollution is a common environmental problem faced by mankind [1]. Transient or

Manuscript received 30 March 2022; revised 30 May 2022; accepted 20 June 2022. Date of publication 29 June 2022; date of current version 15 July 2022. This work was supported in part by the Guangdong Basic and Applied Basic Research Foundation under Grant 2022A1515010492 and in part by a grant from the State Key Laboratory of Resources and Environmental Information System. (*Corresponding author: Huanfeng Shen.*)

Tongwen Li is with the School of Geospatial Engineering and Science, Sun Yat-sen University, Zhuhai 519082, China, and also with the Key Laboratory of Natural Resources Monitoring in Tropical and Subtropical Area of South China, Ministry of Natural Resources, Guangzhou 510000, China (e-mail: litw8@mail.sysu.edu.cn).

Jingan Wu is with the School of Geospatial Engineering and Science, Sun Yat-sen University, Zhuhai 519082, China (e-mail: wujg5@mail.sysu.edu.cn).

Jiajia Chen is with the School of Resource and Environmental Science, Wuhan University, Wuhan 430072, China (e-mail: evechen@whu.edu.cn).

Huanfeng Shen is with the School of Resource and Environmental Science and the Collaborative Innovation Center of Geospatial Technology, Wuhan University, Wuhan 430072, China (e-mail: shenhf@whu.edu.cn).

Digital Object Identifier 10.1109/TGRS.2022.3187095

prolonged exposure to high levels of O₃ can cause premature death, respiratory diseases (e.g., asthma and respiratory infections), and cardiovascular diseases (e.g., stroke and arrhythmias) in humans [2]. In addition, O₃ is an important greenhouse gas because tropospheric O₃ can absorb and release 8–10 μm of Earth's infrared radiation [3], which can exert an important influence on climate change. As a result, O₃ pollution has a considerable impact on human health, climate change, and the ecosystem [4]–[6], and effective measures must be implemented for the monitoring of O₃ pollution.

Compared with ground monitoring stations, satellite remote sensing is advantageous due to its broad and long-term observations, and it has been increasingly used in the monitoring of surface O₃ pollution [7]–[9]. Researchers usually utilize statistical models and machine learning to construct the relationships between satellite-observed products (e.g., total O₃ products, O₃ profile products, and O₃ precursor products) and surface O₃ concentration, so as to obtain large-scale O₃ data. These models mainly include geographically weighted regression (GWR) [10], land-use regression (LUR) [11], random forest (RF) [12], extreme gradient boosting (XGBoost) [13], [14], the RF-generalized additive model (RF-GAM) [15], the RF-based data-fusion model [16], the light gradient boosting machine (Light-GBM) [17], [18], extremely randomized trees (ETs) [19], the neural network (NN) [20], [21], the deep forest (DF) [22], the self-adaptive geospatially local scheme based on categorical boosting (SGLBoost) [23], and so on. These models have important applications in the remote sensing estimation of surface O₃.

The aforementioned models can generally be divided into two categories. The first category is the global model, which establishes the relationship between variables by using constant coefficients but cannot consider the spatiotemporal heterogeneity of the relationship. The second category is the spatiotemporal model that can consider the spatiotemporal heterogeneity of the relationship between variables [24], such as GWR and SGLBoost (the others are global models). Both the two categories of models can achieve fruitful applications, but the second one may produce relatively better results on the local scale. Thereinto, the GWR model and its extension model, namely, geographically and temporally weighted regression (GTWR), are the most representative spatiotemporal models, which have been widely applied in the remote sensing estimation of atmospheric parameters, such as fine particulate matter (PM_{2.5}) [25]–[27].

GWR and GTWR models assume that samples with closer distances contribute more to the modeling [28], [29], and the weights of the samples are determined based on the spatiotemporal weighting of spatial and temporal distances. However, the spatiotemporal weighting method is often empirical, and two problems still exist. First, temporal and spatial distances have many expressions (e.g., the Euclidean distance and the Manhattan distance), so incorporating all expressions of distances into the empirical spatiotemporal weighting method is difficult, that is, establishing a unified expression of temporal and spatial distances is impossible. Second, it is still worth further discussing whether the weights obtained are optimal when the empirical formula (e.g., the Gaussian function) is used for spatiotemporal weighting.

Fortunately, NNs can incorporate all forms of distances for a unified expression and can automatically learn the best expression between spatial distances, temporal distances, and spatiotemporal weights. Meanwhile, NNs have strong nonlinear expression ability [30], [31] and can effectively mine the internal association between spatial distances, temporal distances, and spatiotemporal weights, possibly resulting in better results than those of the empirical weighting method. Therefore, it may be a promising idea to replace the empirical spatiotemporal weighting method in GWR and GTWR models with an NN.

In addition, aiming at the problem that GWR and GTWR models cannot address the nonlinear relationships between variables due to local linear modeling, our previous study [32] developed a geographically and temporally weighted NN (GTWNN) model, which is essentially an NN that integrates spatiotemporal weighting into NN modeling. Similarly, the spatiotemporal weighting method is empirical, and it is difficult to achieve the optimal expression between spatiotemporal weights, and spatial and temporal distances.

Therefore, this study focuses on the development of the enhanced GTWNN (EGTWNN) model, in which an NN is used to replace the empirical spatiotemporal weighting method of the previous GTWNN model. At the same time, this study also develops an enhanced GTWR (EGTWR) model, in which the empirical spatiotemporal weighting method of the original GTWR model is replaced by an NN. The proposed EGTWNN and EGTWR models can automatically learn the optimal expression of spatiotemporal weighting, which helps the model represent the spatiotemporal proximity relationship and improves the estimation accuracy. On the basis of the EGTWNN and EGTWR models, large-scale daily O_3 data can be obtained, which provides a data basis for O_3 pollution monitoring and control. Validation and mapping results indicate that the proposed models report a promising application prospect.

II. STUDY REGION AND DATA

A. Study Region

The study domain covers the Guangdong–Hong Kong–Macao Greater Bay Area (GBA) region in China, which includes nine cities in the Pearl River Delta of Guangdong (i.e., Guangzhou, Shenzhen, Zhuhai, Foshan,

Huizhou, Dongguan, Zhongshan, Jiangmen, and Zhaoqing), as well as Hong Kong and Macao (as shown in Fig. 1). The GBA region is one of the most open and economically dynamic regions in China, and it plays an outstanding driving role and occupies a pivotal strategic position in China's economic and social development. The GBA region is benchmarking other bay areas in the world, intending to build a world-class megacity region [33]. However, with the rapid urbanization and industrialization, the GBA region faces serious O_3 pollution [34]–[36], which has become an important obstacle to economic and social development.

B. Data

1) *Ground Station Measurements of O_3* : Hourly O_3 concentrations measured at ground monitoring stations in 2019 were obtained from the China National Environmental Monitoring Center (CNEMC, <https://air.cnemc.cn:18007/>), the Geophysical and Meteorological Bureau of Macao (<https://www.smg.gov.mo/zh/subpage/64/realtime-iqa-report>), and the Environmental Protection Department of Hong Kong (<https://cd.epic.epd.gov.hk/EPICDI/air/station/>). In total, 108 ground monitoring stations (84 stations in the Pearl River Delta and its surroundings, 18 stations in Hong Kong, and six stations in Macao) were obtained. In this article, a daily maximum 8-h average of O_3 (MDA8_ O_3) was adopted to investigate the O_3 pollution level in the GBA region. The study area was divided into a 0.05° grid for the modeling and mapping of surface O_3 , and the data of multiple monitors in one grid cell were averaged. As shown in Fig. 1, 10% of the grid cells that contain monitors are randomly selected as the model test dataset, and 90% of the grid cells are used as the modeling dataset.

2) *Satellite Observations*: The Tropospheric Monitoring Instrument (TROPOMI) sensor onboard the Sentinel-5 Precursor (S5P) satellite can observe O_3 precursor information, namely, nitrogen dioxide (NO_2) and formaldehyde (HCHO), providing a possible technical approach for remote sensing estimation of surface O_3 . In this study, TROPOMI NO_2 and HCHO products are obtained from the Google Earth Engine (<https://developers.google.com/earth-engine/datasets/catalog/sentinel>), with fields of “tropospheric_ NO_2 _column_number_density” and “tropospheric_HCHO_column_number_density,” respectively. The temporal resolution of the satellite observations is one day, and the spatial resolution is resampled to $0.05^\circ \times 0.05^\circ$. The TROPOMI NO_2 products were validated to show a high consistency with multi-axis differential optical absorption spectroscopy (DOAS) sites, and the correlation coefficient was 0.84 [37]. Also, the TROPOMI HCHO products have achieved a good performance compared to ground observations [38]. The satellite-observed NO_2 and HCHO products are used as primary predictors for remote sensing estimation modeling of surface O_3 .

3) *Reanalysis Data*: ERA5 reanalysis data [39] were included to provide meteorological variables, O_3 -related simulated variables, and radiation-related variables, which has been widely used in O_3 analysis and modeling [40], [41]. The meteorological variables include 10-m U-wind component, 10-m V-wind component, 2-m temperature, boundary

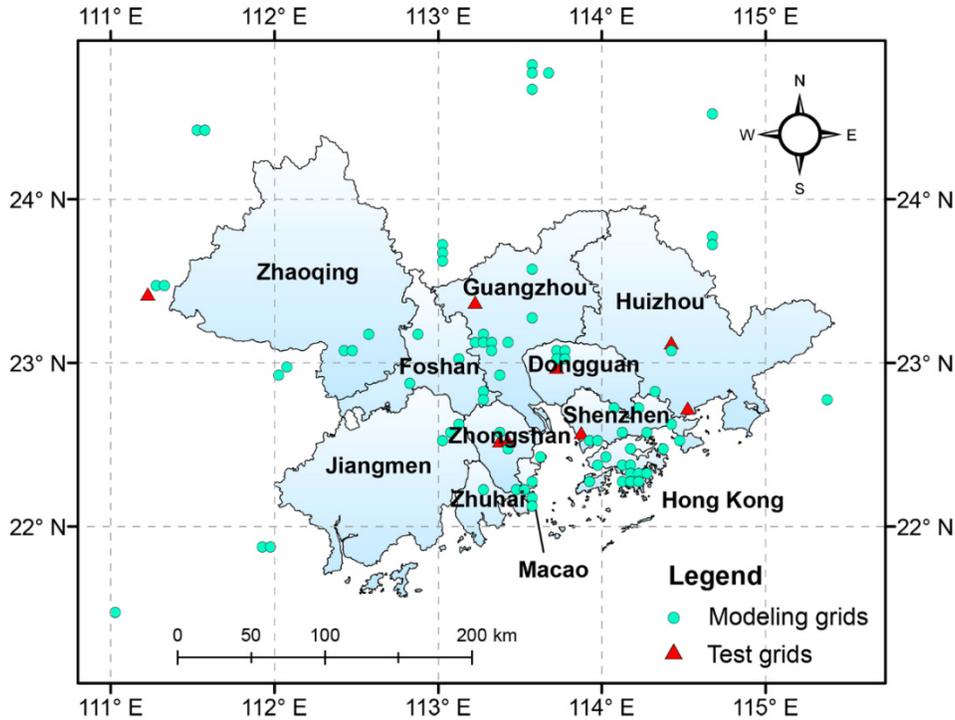


Fig. 1. Study region and distribution of monitoring stations.

layer height, relative humidity, evaporation, surface pressure, and total precipitation. The O₃-related simulated variables are composed of the O₃ mass mixing ratio and total column O₃. The radiation-related variables contain surface net solar radiation, downward UV radiation at the surface, and surface net thermal radiation. These ERA5 reanalysis data are downloaded from the Copernicus Climate Change Service (C3S) Climate Data Store (<https://cds.climate.copernicus.eu/#!/search?text=ERA5&type=dataset>).

C. Data Preprocessing

A grid with a spatial resolution of 0.05° was created for the remote sensing modeling and mapping of O₃ concentration in this study. All data were reprojected to the same coordinate system and resampled to the same resolution of 0.05°, and the modeling and mapping were based on this spatiotemporally consistent dataset.

For the GTWR and EGTWR models, the collinearity of the input variables will lead to instability. Therefore, we conducted a collinearity diagnostic for the input variables based on the variance inflation factor (VIF) and the GWR tool in ArcGIS, and high collinearity is found among some variables (VIF > 10). To address this issue, these variables are excluded from the modeling, and the following variables are selected for use in this study: NO₂, HCHO, boundary layer height, O₃ mass mixing ratio, and surface net thermal radiation.

III. METHOD

As mentioned previously, the spatiotemporal weighting method of GTWR and GTWNN models is empirical. It fails to

achieve a unified expression of commonly used temporal and spatial distances, and whether such an empirical weighting method is optimal still needs further discussion. Thus, the main purpose of this study is to use an NN to replace the empirical spatiotemporal weighting method of GTWR and GTWNN models; the proposed models are called EGTWR and EGTWNN, respectively. The relationships among the four models are shown in Fig. 2.

A. GTWNN

The GTWNN model was developed in our previous study [32], and it was inspired by the GTWR model. For ease of understanding, we first briefly describe the GTWR model for O₃ estimation, which can be depicted as $O_{3i} = \beta_0(u_i, v_i, t_i) + \sum_{k=1}^K \beta_k(u_i, v_i, t_i) \cdot x_{ik}$, where O_{3i} is the surface O₃ concentration at location (u_i, v_i) on day t_i ; β_0 stands for the intercept for location (u_i, v_i) on day t_i ; x denotes the influencing variables, such as NO₂ and HOCHO; and β_k denotes the location–time–specific slopes for the influencing variables. Meanwhile, K is the number of influencing variables. Let $\boldsymbol{\beta} = [\beta_0, \beta_1, \dots, \beta_K]^T$, and the model parameters (intercept and slopes) for location (u_i, v_i) on day t_i can be expressed as $\boldsymbol{\beta}(u_i, v_i, t_i) = (\mathbf{X}^T \mathbf{W}(u_i, v_i, t_i) \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W}(u_i, v_i, t_i) \mathbf{Y}$, where \mathbf{X} is the input matrix of influencing variables, \mathbf{Y} is the vector of O₃, and $\mathbf{W}(u_i, v_i, t_i) = \text{diag}(w_{i1}, w_{i2}, \dots, w_{ij}, \dots, w_{in})$, in which w_{ij} denotes the spatiotemporal weight of the j th sample, and n is the number of samples. An empirical Gaussian weighting function is often used to calculate the spatiotemporal weight [25], [29].

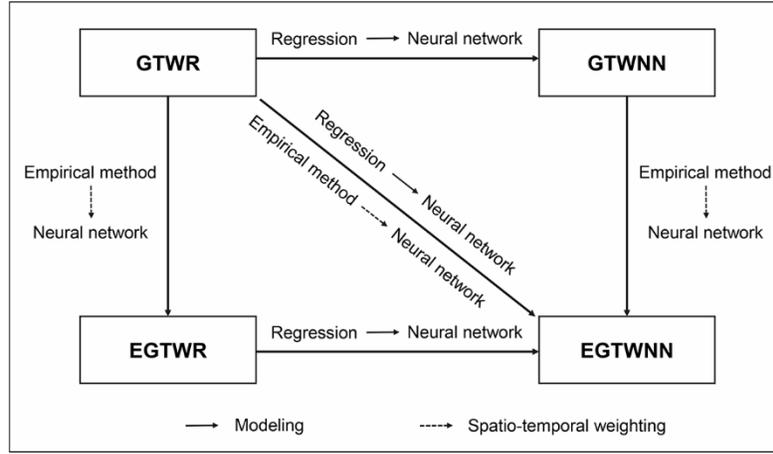


Fig. 2. Internal relationships among the models.

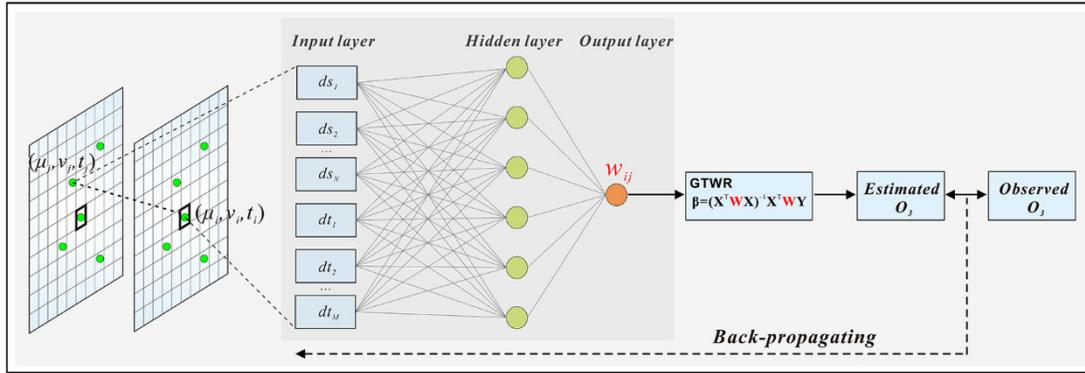


Fig. 3. Schematic of the EGTWR model.

Given that the GTWR model is locally linear, we developed a GTWNN model, which uses an NN to replace the regression process. Specifically, the GTWNN model can be described as

$$O_{3i} = f_{(u_i, v_i, t_i)}(x_{1i}, x_{2i}, \dots, x_{Ki}) \quad (1)$$

where $f_{(u_i, v_i, t_i)}$ denotes the location–time–specific NN, that is, the NN model varies with location and time. Specifically, a generalized regression NN (GRNN) with four layers is used (denoted as GTW-GRNN). The first layer is the input layer. The second layer is the pattern layer, which has as many neurons as the number of samples. The outputs of the pattern layer are passed on to the summation units in the third layer (summation layer). The last layer is the output layer. The details of the GRNN model can be referred to Specht [42].

The essence of constructing the GTW-GRNN model is to incorporate spatiotemporal weighting into the GRNN model, and the process is rough as follows. Assume that n samples are included to construct the GTW-GRNN model for location (u_i, v_i) on day t_i . Then, the number of neurons in the pattern layer is n . For the j th neuron of the pattern layer, its output can be expressed as $a_j = \exp(-\|\mathbf{x}_i - \mathbf{x}_j\|)$, where \mathbf{x}_i stands for the input vector of location (u_i, v_i) on day t_i , \mathbf{x}_j means the input vector of sample j , $\|\cdot\|$ is the Euclidean distance, $b = 0.8326/\text{spread}$ is the bias term, and spread is a parameter

to control the smoothness of the fitting function. The outputs of the pattern layer are passed to the summation layer, and the predicted O_3 concentration can be expressed as

$$O_{3i} = \frac{\sum_{j=1}^n a_j w_{ij} y_j}{\sum_{j=1}^n a_j w_{ij}} \quad (2)$$

where w_{ij} is the spatiotemporal weight of sample j for location (u_i, v_i) on day t_i calculated by the empirical weighting method and y_j denotes the O_3 concentration of sample j . According to (2), if the influencing variables and O_3 observations of the training samples are obtained, the O_3 concentration for location (u_i, v_i) on day t_i can then be acquired, by integrating the spatiotemporal weights to carry out joint weighting. The details of the GTW-GRNN model can be found in [32].

B. EGTWR

To improve the spatiotemporal weighting of the GTWR model, an EGTWR model is developed in this study, which replaces the empirical spatiotemporal weighting function with an NN. The basic idea of the EGTWR model is similar to that of [43], but the modeling process is different (the details are shown below). The structure of the EGTWR model is presented in Fig. 3. In general, the structure of the EGTWR

model is similar to that of GTWR, and the only difference lies in the method of spatiotemporal weighting.

In the EGTWR model, the fully connected NN (FNN) is used to establish the relationship between spatial and temporal distances, and spatiotemporal weights. As shown in Fig. 3, FNN has a three-layer structure. The first layer is the input layer, in which temporal and spatial distances are incorporated, and the second layer is the hidden layer, which carries out the nonlinear transformation for the input signals. The last layer is the output layer, which contains only one neuron, namely, the spatiotemporal weight. It should be noted that the outputted spatiotemporal weights for the samples will be normalized by dividing by the maximum. In this study, two kinds of spatial distances and two kinds of temporal distances are adopted ($M = 2$ and $N = 2$ in Fig. 3), and their specific expressions are given as follows:

$$ds_1 = R \cdot \arccos\left(\cos\left(\frac{v_i}{180} \cdot \pi\right) \cdot \cos\left(\frac{v_j}{180} \cdot \pi\right) \cdot \cos\left(\frac{u_i}{180} \cdot \pi - \frac{u_j}{180} \cdot \pi\right) + \sin\left(\frac{v_i}{180} \cdot \pi\right) \cdot \sin\left(\frac{v_j}{180} \cdot \pi\right)\right) \quad (3)$$

$$ds_2 = \sqrt{(u_i - u_j)^2 + (v_i - v_j)^2} \quad (4)$$

$$dt_1 = t_i - t_j (t_i \geq t_j) \quad (5)$$

$$dt_2 = 1 - \cos\left((t_i - t_j) \frac{2\pi}{T}\right) (T = 365) \quad (6)$$

where u and v denote longitude and latitude, respectively; R means Earth's radius; ds_1 is Earth's surface distance from location (u_i, v_i) to location (u_j, v_j) ; ds_2 stands for the latitude and longitude distance; dt_1 is the temporal distance; $t_i \geq t_j$ indicates that only the samples before day t_i are included for modeling; and dt_2 is the cosine distance that considers the seasonality of O₃ and the explanatory variables [26].

Through the above process, the EGTWR model can be constructed. However, the outputs of FNN are the spatiotemporal weights; we have no truth values for NN training. How to train and optimize the FNN model is the key problem for the EGTWR model. According to the solution of the GTWR model (see Section III-A), once the spatiotemporal weight matrix is determined, the model coefficient can be solved. In other words, the EGTWR model coefficients can be solved once the output values of FNN are substituted. Therefore, we use the spatiotemporal weights outputted by FNN to obtain the O₃ values. Subsequently, station O₃ observations are adopted as the label data, and the mean square error (mse) loss function is established. It is then backpropagated to optimize the weights of FNN. Notably, an FNN can be constructed for each location independently in theory, but, to reduce the computation and improve model stability, the weights of FNN are shared, that is, all FNNs constructed within the study region are the same.

As mentioned above, the EGTWR model proposed in this study is consistent with the basic idea of the model in [43], but some differences exist in the specific implementation. The equation for solving the coefficients of our model is $\beta = (\mathbf{X}^T \mathbf{W} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W} \mathbf{Y}$, whereas the solution method in [43] is $\beta = \mathbf{W} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}$. It can be seen that our

model completely follows the solution approach of the original GTWR model. Compared to their study, we further developed the EGTWNN model, as shown in Section III-C.

C. EGTWNN

Our previous study proposed the GTWNN model to simultaneously address the spatiotemporal heterogeneity and nonlinear relationships [32]. However, the spatiotemporal weighting function can still be improved. As a result, an EGTWNN model is proposed in this study. The idea of the EGTWNN model is similar to that of EGTWR, but the difference lies in the use of an NN rather than a linear regression model to build the relationships between variables. The model structure of EGTWNN, which is nested by two NNs, is shown in Fig. 4. The first NN constructs the relationship between spatial and temporal distances and spatiotemporal weights, and the results of the first NN are inputted into the second NN for joint weighting to obtain surface O₃ concentration. Among them, the first NN adopts the same structure as FNN in the EGTWR model, and the second is the GRNN model. Similar to the EGTWR model, all the FNNs in the EGTWNN share the same weights.

As indicated in (2), as long as the spatiotemporal weights are determined, the O₃ concentration can be obtained (other variables can be calculated from the input data). Therefore, we input the spatiotemporal weights obtained by FNN into GRNN [i.e., substituting the spatiotemporal weights to (2)], and O₃ estimates of the EGTWNN model can then be obtained. The mse loss function between the model estimates and the ground observations of O₃ can be established, and the weights of the NN can thus be updated through the backpropagation mechanism [44]. The modeling process of EGTWNN is given as follows.

Step 1 (Spatiotemporal Weighting): With temporal and spatial distances as inputs (using the same four kinds of distances as the EGTWR model), the FNN model was constructed to obtain the spatiotemporal weights. Given the local modeling strategy of EGTWNN, we cannot collect a large number of samples for training, so the FNN model only uses one hidden layer. The outputs of the FNN model are the spatiotemporal weights of the samples, and the truth values cannot be obtained. Therefore, the FNN model is not trained separately but adopts the idea of nested modeling and joint learning (see Steps 2 and 3).

Step 2 (Nested NN Modeling): The spatiotemporal weights obtained in Step 1 are inputted into the GRNN model, the joint weighting is implemented by combining with the GRNN weights, and the nested NN model (i.e., EGTWNN) is constructed. With this model, surface O₃ concentration can be obtained by inputting the influencing variables.

Step 3 (Error Backpropagation): On the basis of monitoring station observations of O₃ and the model estimates obtained in Step 2, the error can be calculated to construct the loss function. Afterward, the error is backpropagated, and the weights of the NNs can be updated. Notably, the weights of the GRNN model are mainly determined by the input data, so the weights of the FNN model are updated in this process.

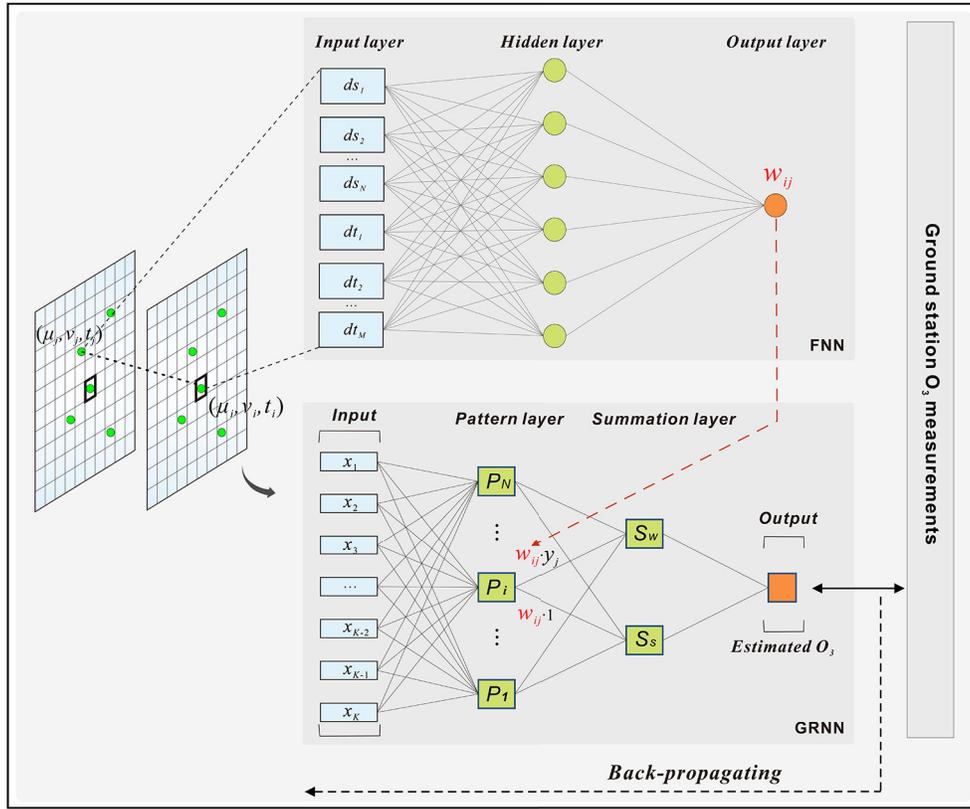


Fig. 4. Schematic of the EGTWNN model.

For the proposed EGTWNN model, four main categories of parameters need to be determined. The first category encompasses the spatial and temporal parameters for modeling and the six surrounding grid cells, and the samples in the five days before the target day are used to construct the model in this study. The second category encompasses the FNN model parameters, the input layer, the hidden layer, and the output layer containing four, six, and one neurons, respectively. The third category is the GRNN model parameter, and spread is set to 0.1 in this work. The last category encompasses the nested NN modeling parameters, the learning rate is 0.1, the number of iterations is set to 40, and the optimization method is Adam [45]. These parameters are selected through the cross-validation (CV) and test of the model, and Section III-D provides the specific process.

D. Model Evaluation

The model CV [46] and the model external test are adopted to determine the model parameters and evaluate model performance in this study, and the details are shown in Fig. 5. All grid cells containing monitoring stations are randomly partitioned into two parts, that is, the modeling dataset (90%) and the test dataset (10%), which can be seen in Fig. 1. The modeling dataset is used to establish the EGTWNN model, and the spatial-based CV is adopted to determine the model parameters during modeling. All the grid cells containing monitoring stations in the modeling dataset are randomly and equally partitioned into tenfold, and ninefold is used for

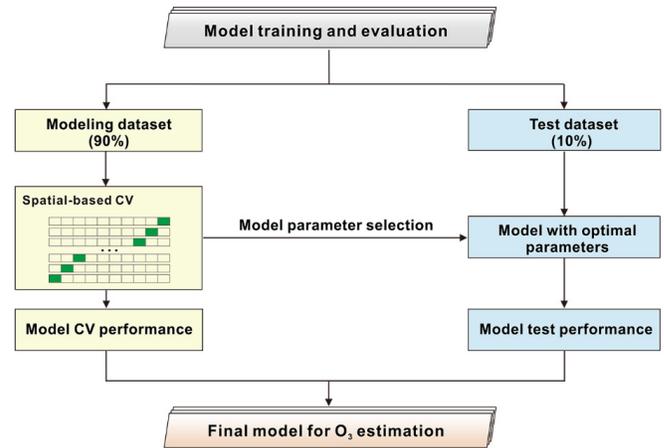


Fig. 5. Procedure of the model evaluation process.

the model establishment and the remaining one for model validation. The process would be repeated ten times until each fold has been used for model validation. Through spatial-based CV, the optimal model parameters can be determined, and the test dataset is then used to evaluate the model performance, so the model test performance can be achieved. In this study, both model CV performance and model test performance can be obtained, and they reflect the model performance and generalization ability together.

Several statistical indicators are used to quantitatively evaluate model performance, including linear regression between

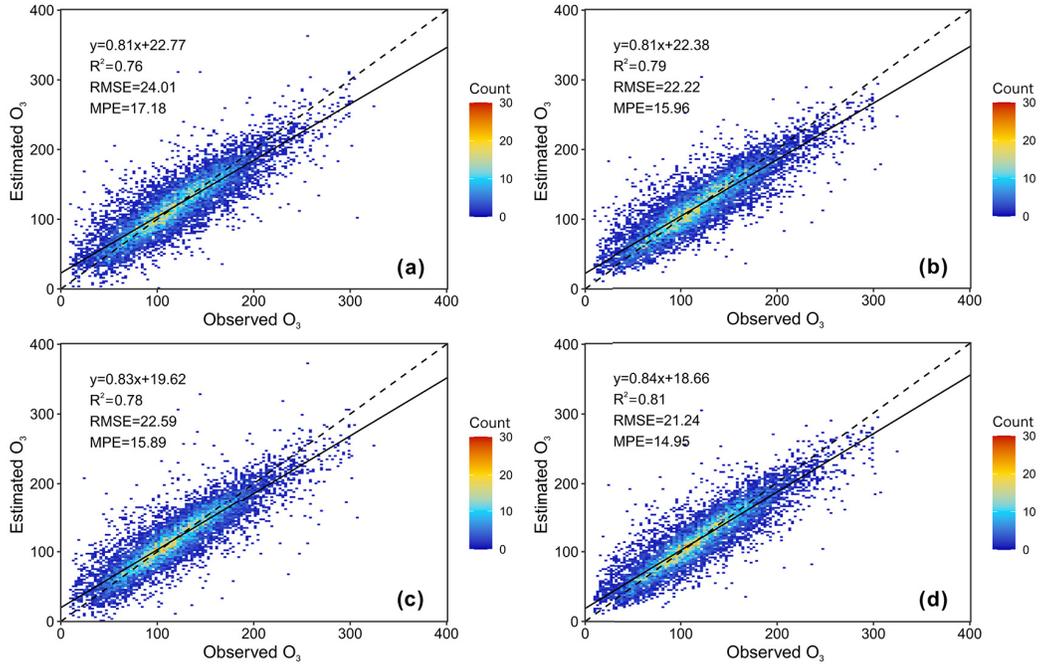


Fig. 6. CV performance of (a) GTWR, (b) GTWNN, (c) EGTWR, and (d) EGTWNN models.

model estimates and station O₃ observations (slope and intercept), coefficient of determination (R^2 , unitless) for the linear regression, root mse (RMSE, $\mu\text{g}/\text{m}^3$), and mean absolute predictive error (MPE, $\mu\text{g}/\text{m}^3$).

IV. RESULTS AND ANALYSIS

A. CV and Test Performance of the Models

The model CV is adopted to determine the model parameters and reflect the predictive ability, and the results are shown in Fig. 6. Due to the consideration of spatiotemporal heterogeneity, the GTWR model has obtained relatively satisfactory results, with R^2 , RMSE, and MPE values of 0.76, 24.01 $\mu\text{g}/\text{m}^3$, and 17.18 $\mu\text{g}/\text{m}^3$, respectively. To further consider the nonlinear relationships, the GTWNN model was proposed, and it obtained much better performance; the R^2 , RMSE, and MPE values are 0.79, 22.22 $\mu\text{g}/\text{m}^3$, and 15.96 $\mu\text{g}/\text{m}^3$, respectively. In addition, the EGTWR model is developed in this study to improve the empirical spatiotemporal weighting method of GTWR with an NN, and superior performance is reported compared with the GTWR model, with the R^2 value increasing by 0.02 (from 0.76 to 0.78), the RMSE value decreasing by 1.42 $\mu\text{g}/\text{m}^3$ (from 24.01 to 22.59 $\mu\text{g}/\text{m}^3$), and the MPE value decreasing by 1.29 $\mu\text{g}/\text{m}^3$ (from 17.18 to 15.89 $\mu\text{g}/\text{m}^3$), respectively. Finally, when comparing the EGTWNN model to the EGTWR model, the GRNN model rather than the regression model is used to address the nonlinear relationships between variables; when the EGTWNN model is compared with the GTWNN model, the empirical spatiotemporal weighting function is replaced with the FNN model to automatically learn the optimal spatiotemporal weights. As a result, the proposed EGTWNN model has achieved the best performance, with R^2 , RMSE,

and MPE values of 0.81, 21.24 $\mu\text{g}/\text{m}^3$, and 14.95 $\mu\text{g}/\text{m}^3$, respectively.

The models are also evaluated using the external dataset, and the model test performance is shown in Fig. 7. Overall, the model test performance presents a similar pattern as the model CV performance. Among the four models, the GTWR model achieves the poorest performance, and the R^2 , RMSE, and MPE values are 0.80, 22.37 $\mu\text{g}/\text{m}^3$, and 16.55 $\mu\text{g}/\text{m}^3$, respectively. When the nonlinear GRNN model is introduced to establish the GTWNN model, the performance shows some improvement, with R^2 , RMSE, and MPE values of 0.84, 20.35 $\mu\text{g}/\text{m}^3$, and 14.78 $\mu\text{g}/\text{m}^3$, respectively. In addition, when the empirical spatiotemporal weighting is improved by the EGTWR model, the R^2 , RMSE, and MPE values exhibit great advantages, being 0.82, 21.22 $\mu\text{g}/\text{m}^3$, and 15.90 $\mu\text{g}/\text{m}^3$, respectively. The proposed EGTWNN model demonstrates the best performance, with R^2 , RMSE, and MPE values of 0.86, 18.76 $\mu\text{g}/\text{m}^3$, and 13.68 $\mu\text{g}/\text{m}^3$, respectively. Notably, the model test performance is better than the model CV performance for the models, which can probably be attributed to that the test grid cells are generally located in the middle of the study region (see Fig. 1); thus, the test monitoring stations have sufficient modeling stations in their surroundings.

B. Temporal Evaluation of Model Performance

For the four models, the performance patterns exhibited in all seasons (spring: March, April, and May; summer: June, July, and August; autumn: September, October, and November; and winter: January, February, and December) are basically similar to the overall model performance in Section IV-A, that is, the EGTWNN model reports the best performance, the EGTWR and GTWNN models come in second, and the GTWR model performs the worst (see Table I). With autumn

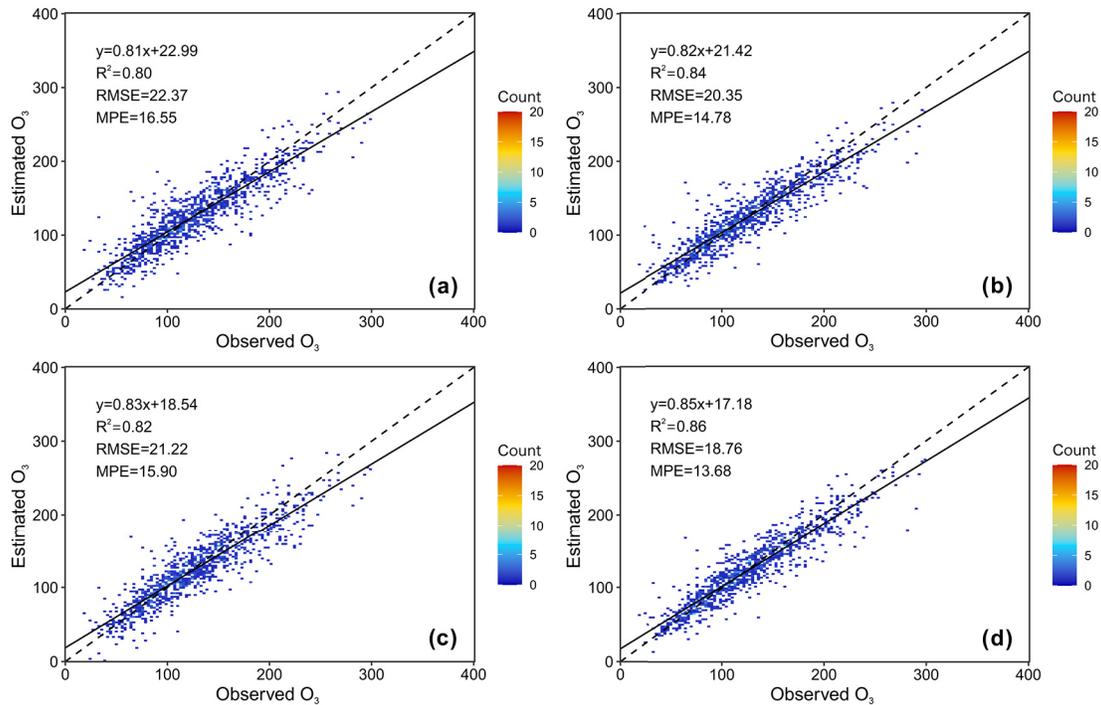


Fig. 7. Test performance of (a) GTWR, (b) GTWNN, (c) EGTWR, and (d) EGTWNN models.

TABLE I
SEASONAL PERFORMANCE OF THE MODELS

			Spring	Summer	Autumn	Winter
GTWR	Model CV	R^2	0.66	0.69	0.72	0.58
		RMSE	23.73	29.16	23.99	19.47
	Model test	R^2	0.69	0.78	0.74	0.66
		RMSE	22.40	25.04	23.73	16.66
GTWNN	Model CV	R^2	0.72	0.75	0.74	0.62
		RMSE	21.18	25.64	22.75	18.55
	Model test	R^2	0.74	0.83	0.78	0.74
		RMSE	20.82	22.48	21.83	14.63
EGTWR	Model CV	R^2	0.67	0.73	0.75	0.62
		RMSE	23.68	27.43	22.30	18.30
	Model test	R^2	0.70	0.80	0.78	0.67
		RMSE	22.51	24.08	21.89	16.39
EGTWNN	Model CV	R^2	0.73	0.79	0.77	0.63
		RMSE	20.82	23.87	21.67	18.25
	Model test	R^2	0.74	0.85	0.83	0.73
		RMSE	20.21	20.65	19.49	14.71

as an example, the values of CV R^2 are 0.77, 0.75, 0.74, and 0.72, respectively, and the values of test R^2 are 0.83, 0.78, 0.78, and 0.74, respectively. It is worth noting that the EGTWNN model has a similar performance as GTWNN in winter and reports no advantages.

Seasonally, the four models all perform the best in summer and autumn, and the worst in winter. Taking the EGTWNN

model as an example, the CV R^2 values of the four seasons are 0.73, 0.79, 0.77, and 0.63, respectively, and the test R^2 values are 0.74, 0.85, 0.83, and 0.73, respectively. A possible reason for this result is that summer and autumn have relatively high levels of O₃ concentration, which are easier for the models to predict, and the relationships between variables are more stable in these two seasons.

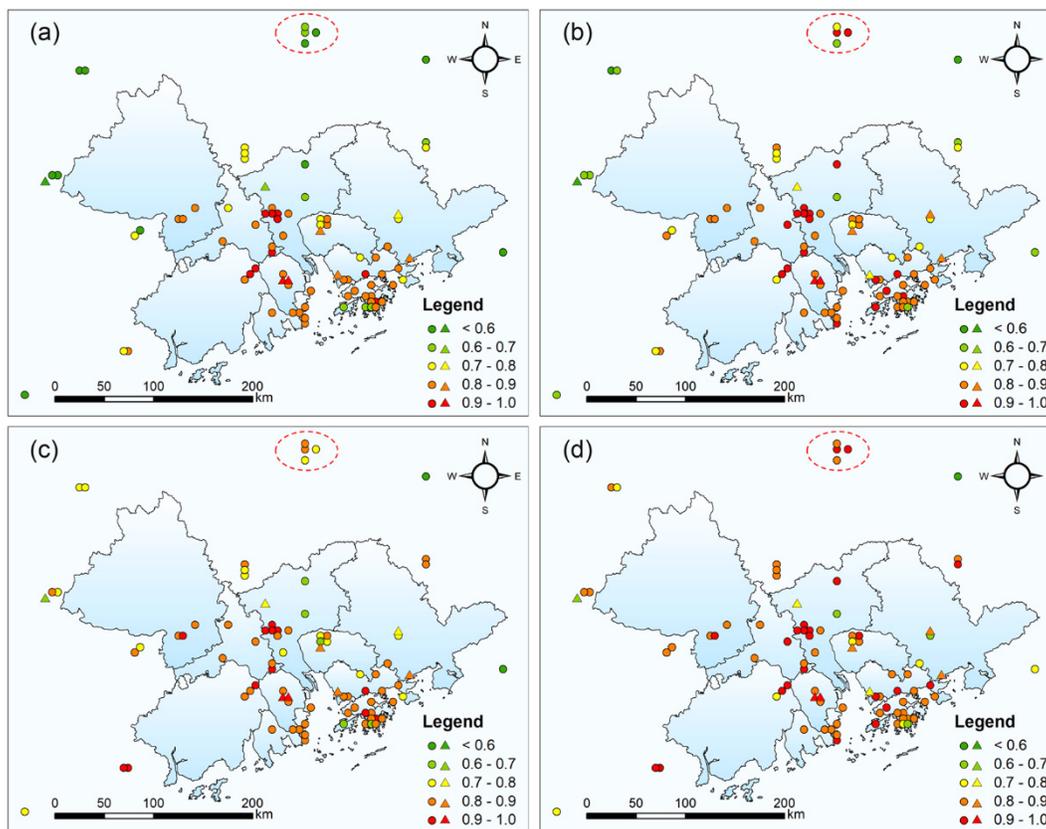


Fig. 8. Distribution of R^2 values for grid cells with monitoring stations. The circle denotes CV performance, and the triangle stands for test performance. (a) GTWR. (b) GTWNN. (c) EGTWR. (d) EGTWNN.

C. Spatial Evaluation of Model Performance

For each grid cell, the R^2 and RMSE values were calculated between the model CV/test results and station observations of O₃ to evaluate the spatial performance of the models. The spatial distribution of the R^2 values of each model is shown in Fig. 8. In general, the EGTWNN model performs the best, with mean CV R^2 and RMSE values of 0.86 and 19.54 $\mu\text{g}/\text{m}^3$, respectively. The EGTWR and GTWNN models come next, with mean CV R^2 (RMSE) values of 0.82 (21.15 $\mu\text{g}/\text{m}^3$) and 0.82 (20.83 $\mu\text{g}/\text{m}^3$), respectively. The GTWR model performs the worst, with mean CV R^2 and RMSE values of 0.77 and 22.77 $\mu\text{g}/\text{m}^3$, respectively. At the same time, the mean test R^2 values are 0.84, 0.80, 0.81, and 0.78 for the EGTWNN, EGTWR, GTWNN, and GTWR models, respectively. In addition, the proportions of R^2 values higher than 0.80 for these models are 86%, 71%, 71%, and 61%, respectively, indicating that the EGTWNN model has a great advantage in spatial performance.

In terms of the spatial distribution of R^2 values, the EGTWR and EGTWNN models are superior to the GTWR and GTWNN models on the whole, respectively, due to the use of an NN to replace the empirical spatiotemporal weighting method. The most obvious ones are the outer grid cells, such as those circled in Fig. 8. The R^2 values of the GTWR model are concentrated within 0.6–0.7, whereas the R^2 value of the EGTWR model increases to 0.7–0.9. Similarly, the R^2 values

of the EGTWNN model are improved compared with those of the GTWNN model, which are all higher than 0.80.

V. DISCUSSION

A. Spatial and Temporal Mappings of Surface O₃

Daily O₃ concentration data can be obtained based on the proposed EGTWNN model. This study selects November 21–24, 2019, as an example to show the O₃ remote sensing estimation results, as shown in Fig. 9. It indicates that the spatial distribution of O₃ estimated by satellite remote sensing is highly consistent with the O₃ observations of ground stations, but they have more spatial information than the ground station observations, indicating that the EGTWNN model proposed in this study has good application potential in the spatiotemporal fine-scale monitoring of surface O₃.

A pollution incident can be monitored and analyzed from the four-day O₃ mapping results. Jiangmen appears in a small high-value range of O₃ on November 21. It slowly spreads to Zhongshan, Dongguan, and Guangzhou on November 22 then to Zhaoqing and Zhuhai on November 23. Then, the wide-range O₃ pollution is formed in the mid-west of the GBA region on November 24. This example shows that satellite remote sensing has great potential to monitor and analyze pollution incidents and processes on a fine scale.

Furthermore, the annual mean surface O₃ estimates in the GBA region are mapped, which are shown in Fig. 10. The

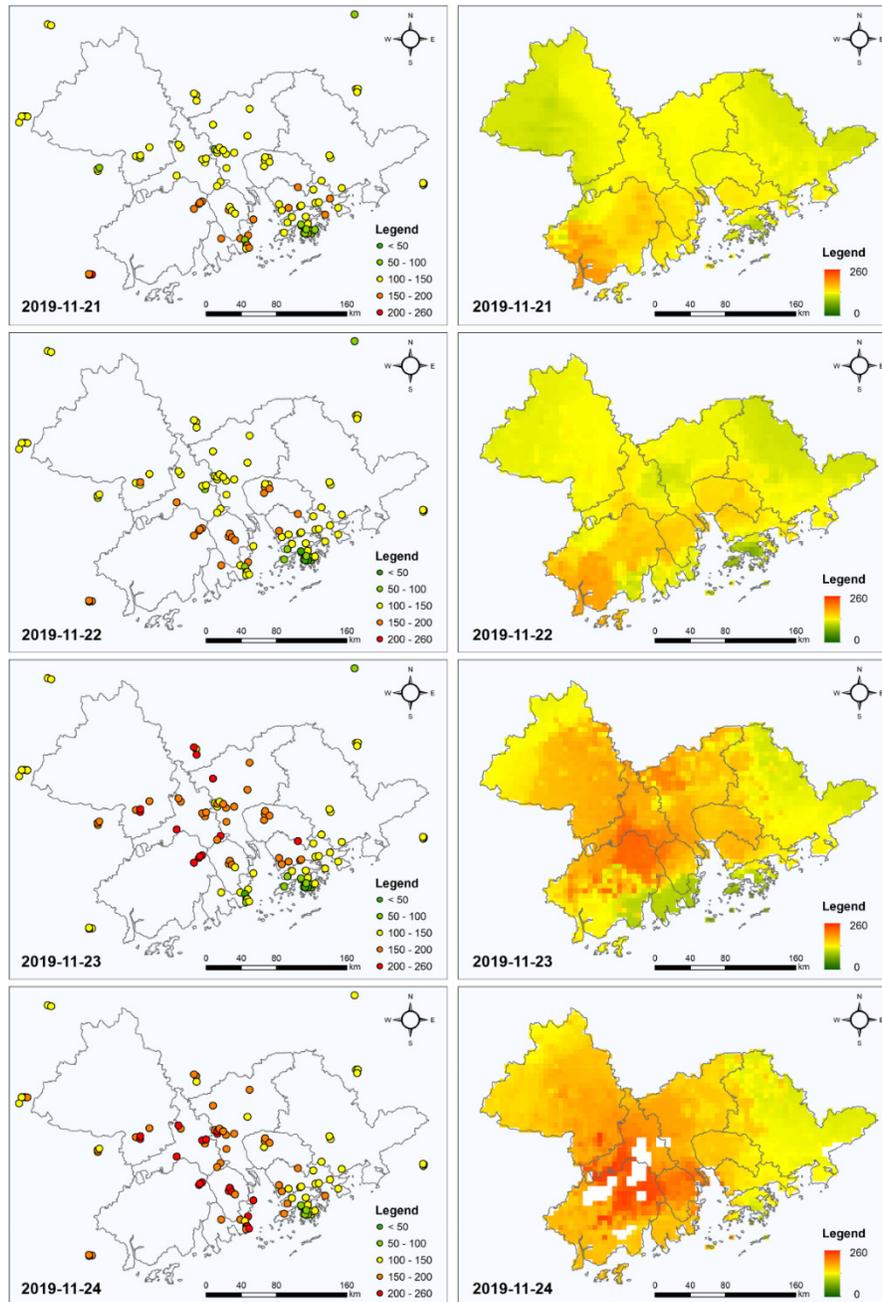


Fig. 9. Spatiotemporal mapping of surface O_3 concentrations. (Left) Station measurements of O_3 . (Right) Remote sensing retrievals of O_3 . The white regions indicate missing data.

O_3 pollution hotspots are located in Guangzhou, Dongguan, and Zhongshan cities, whereas Zhaoqing, Zhuhai, Macao, and Hongkong report relatively low levels of O_3 pollution. The mapping results indicate that the proposed model can achieve high-resolution O_3 data for pollution monitoring and analysis.

B. Interpretability of FNN for Spatiotemporal Weighting

The interpretability of NNs plays a critical role in accuracy improvement, causal inference, and so on. In this study, both the model CV and model test results show that the EGTWNN model can effectively establish the relationship

between O_3 and influencing variables. Is it reasonable for FNN in the EGTWNN model to obtain spatiotemporal weights? In other words, are the spatiotemporal weights obtained by the FNN model interpretable? To this end, we use the traditional empirical Gaussian weighting function to calculate the spatiotemporal weights of the samples in the process of model CV and compare the empirical weights with those obtained by the FNN model to calculate the Spearman correlation coefficient (r) and investigate the consistency of the two results.

Fig. 11(a) shows the statistical results of the r distribution. About 87% of the r values are greater than or equal to 0.90, and the minimum r value is greater than 0.40. These

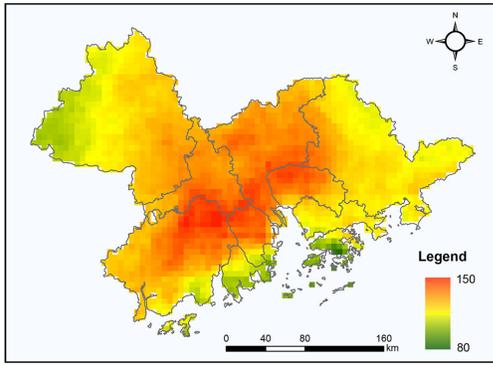


Fig. 10. Annual mean surface O₃ estimates in the GBA region.

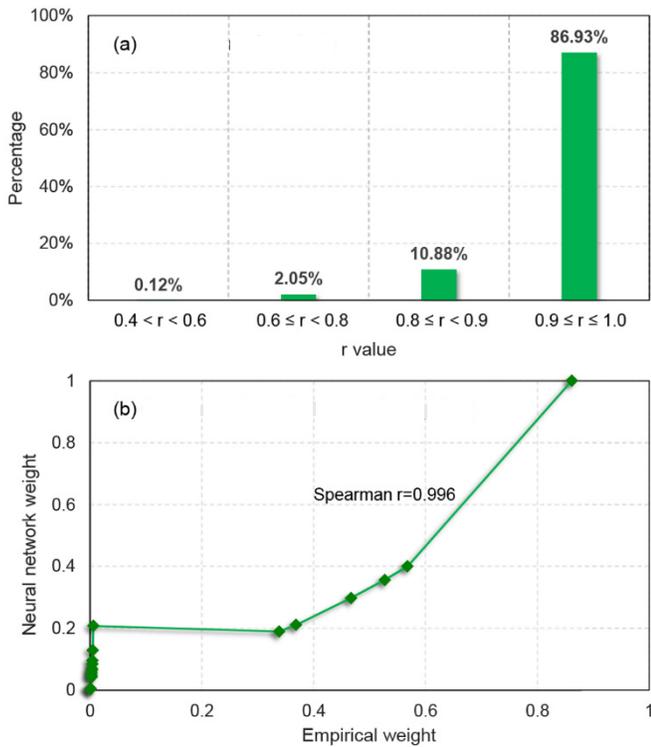


Fig. 11. Comparison of FNN-based spatiotemporal weights and empirical weighting function-based spatiotemporal weights. (a) Distribution of r values. (b) One case of spatiotemporal weight.

results show that the spatiotemporal weights automatically learned by FNN are highly consistent with those obtained by the empirical spatiotemporal weighting function, indicating that the spatiotemporal weights learned by FNN are highly interpretable. Fig. 11(b) presents a case with 16 samples participating in the modeling. The spatiotemporal weights obtained by FNN are highly consistent with the empirical results ($r = 0.996$) but are not equal numerically. Through NN learning and optimization, the spatiotemporal weights are mined more accurately to describe the spatiotemporal proximity relationship of the samples, thus resulting in improved model estimation accuracy.

C. Inspiration for Incorporating Neural Networks Into Mechanism Models

Mechanism models use physical and chemical mechanisms to construct retrieval equations for surface parameters, while NNs train relationships between variables based on a large amount of data. In recent years, scholars have extensively discussed the choice of mechanism model and NN, among which a critical view is that the coupling of the two is an important tendency [47]–[49]. This study provides a reference for the coupling of mechanism models and NNs, that is, using an NN to replace the partial molecular process of the mechanism model. The GTWR model is regarded as a mechanism model (although it is also an empirical statistical model), the FNN model is used in this study to replace the spatiotemporal weighting subprocess of the model, and the results demonstrate that higher estimation accuracy is obtained. This example indicates that it is possible to improve the accuracy and obtain more ideal results by using an NN to replace a part of the uncertain subprocess of the mechanism model.

D. Limitations and Uncertainties

Although the EGTWR and EGTWNN models proposed in this study have achieved good performance, limitations still exist in the following aspects. The first one is variable selection. For the variable selection of the EGTWR model, to avoid the instability caused by multicollinearity, it is necessary to use some statistical indicators (e.g., VIF) to screen variables, but this procedure is troublesome. Nevertheless, the proposed EGTWNN model has little restriction on variable screening. To maintain consistency and comparability, the same input variables as the EGTWR model are adopted for the EGTWNN model. Second, to reduce calculation, we only include a certain number of surrounding stations for modeling. If too few samples are collected, the EGTWR model is prone to instability. In this study, when the number of samples is less than nine, the EGTWR model tends to be unstable. Prosperously, the EGTWNN model is not subject to this problem, and the model can work with no less than two samples. The third aspect is the common problem faced by the EGTWR and EGTWNN models. That is, compared with GTWR and GTWNN models, the amount of modeling computation increases a lot, which is mainly spent on the joint training and optimization of NNs.

VI. CONCLUSION

Based on the GTWNN model, the EGTWNN model is proposed in this study, and its contribution lies in that an NN is used to replace the empirical spatiotemporal weighting method. The results show that the proposed EGTWNN model can achieve improved performance. This study can be summarized from three aspects as follows.

- 1) Owing to the automatic learning of the optimal weighting function and powerful nonlinear modeling ability of the NN, the proposed EGTWNN model exhibits advantageous performance compared with the GTWNN model, the CV R^2 value increases by 0.02 (from 0.79 to

0.81), and the RMSE value decreases by $\sim 1 \mu\text{g}/\text{m}^3$ (from 22.22 to 21.24 $\mu\text{g}/\text{m}^3$). Also, model test R^2 increases by 0.02 (from 0.84 to 0.86), and RMSE decreases by 1.59 $\mu\text{g}/\text{m}^3$ (from 20.35 to 18.76 $\mu\text{g}/\text{m}^3$).

- 2) The spatiotemporal weights automatically learned by the FNN model are highly consistent with the results obtained by the traditional empirical weighting method. About 87% of the Spearman correlation coefficient values are greater than or equal to 0.90, indicating that the proposed model has strong interpretability. However, the spatiotemporal weights obtained by the FNN model are not completely equal to those of the traditional empirical weighting method, indicating that FNN can describe the contribution of the samples more accurately.
- 3) Daily surface O_3 data can be obtained based on the proposed EGTWNN, and the results reveal the strong potential of the EGTWNN model for monitoring and analyzing O_3 pollution events.

ACKNOWLEDGMENT

The authors acknowledge the data providers for freely releasing the abovementioned data in Section II, among which the China National Environmental Monitoring Center (CNEMC) station data were acquired from the PM25.in website (<http://pm25.in/>) and the Beijing Air platform (<https://quotsoft.net/air/>). They would also like to thank the anonymous editors and reviewers.

REFERENCES

- [1] J. Zhang, Y. Wei, and Z. Fang, "Ozone pollution: A major health hazard worldwide," *Frontiers Immunology*, vol. 10, p. 2518, Oct. 2019.
- [2] *Health Risk and Exposure Assessment for Ozone*, U.S. EPA, Washington, DC, USA, 2014.
- [3] H. M. Worden, K. W. Bowman, J. R. Worden, A. Eldering, and R. Beer, "Satellite measurements of the clear-sky greenhouse effect from tropospheric ozone," *Nature Geosci.*, vol. 1, no. 5, pp. 305–308, May 2008.
- [4] T. Zhao, I. Markevych, M. Romanos, D. Nowak, and J. Heinrich, "Ambient ozone exposure and mental health: A systematic review of epidemiological studies," *Environ. Res.*, vol. 165, pp. 459–472, Aug. 2018.
- [5] H. Liu *et al.*, "Ground-level ozone pollution and its health impacts in China," *Atmos. Environ.*, vol. 173, pp. 223–230, Jan. 2018.
- [6] A. P. K. Tai, M. V. Martin, and C. L. Heald, "Threat to future global food security from climate change and ozone air pollution," *Nature Climate Change*, vol. 4, no. 9, pp. 817–821, Sep. 2014.
- [7] Z. Wang *et al.*, "Systematics of atmospheric environment monitoring in China via satellite remote sensing," *Air Qual., Atmos. Health*, vol. 14, no. 2, pp. 157–169, Feb. 2021.
- [8] N. Colombi, K. Miyazaki, K. W. Bowman, J. L. Neu, and D. J. Jacob, "A new methodology for inferring surface ozone from multispectral satellite measurements," *Environ. Res. Lett.*, vol. 16, no. 10, Oct. 2021, Art. no. 105005.
- [9] L. Shen *et al.*, "An evaluation of the ability of the ozone monitoring instrument (OMI) to observe boundary layer ozone pollution across China: Application to 2005–2017 ozone trends," *Atmos. Chem. Phys.*, vol. 19, no. 9, pp. 6551–6560, 2019.
- [10] X. Y. Zhang, L. M. Zhao, M. M. Cheng, and D. M. Chen, "Estimating ground-level ozone concentrations in eastern China using satellite-based precursors," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 7, pp. 4754–4763, Jul. 2020.
- [11] L. Chen *et al.*, "A hybrid approach to estimating long-term and short-term exposure levels of ozone at the national scale in China using land use regression and Bayesian maximum entropy," *Sci. Total Environ.*, vol. 752, Jan. 2021, Art. no. 141780.
- [12] Y. Zhan, Y. Luo, X. Deng, M. L. Grieneisen, M. Zhang, and B. Di, "Spatiotemporal prediction of daily ambient ozone levels across China using random forest for human exposure assessment," *Environ. Pollut.*, vol. 233, pp. 464–473, Feb. 2018.
- [13] R. Liu, Z. Ma, Y. Liu, Y. Shao, W. Zhao, and J. Bi, "Spatiotemporal distributions of surface ozone levels in China from 2005 to 2017: A machine learning approach," *Environ. Int.*, vol. 142, Sep. 2020, Art. no. 105823.
- [14] R. Li, L. Cui, F. Hongbo, J. Li, Y. Zhao, and J. Chen, "Satellite-based estimation of full-coverage ozone (O_3) concentration and health effect assessment across Hainan island," *J. Cleaner Prod.*, vol. 244, Jan. 2020, Art. no. 118773.
- [15] R. Li, Y. Zhao, W. Zhou, Y. Meng, Z. Zhang, and H. Fu, "Developing a novel hybrid model for the estimation of surface 8 h ozone (O_3) across the remote Tibetan Plateau during 2005–2018," *Atmos. Chem. Phys.*, vol. 20, no. 10, pp. 6159–6175, 2020.
- [16] T. Xue *et al.*, "Estimating spatiotemporal variation in ambient ozone exposure during 2013–2017 using a data-fusion model," *Environ. Sci. Technol.*, vol. 54, no. 23, pp. 14877–14888, 2020.
- [17] Y. Kang *et al.*, "Estimation of surface-level NO_2 and O_3 concentrations using TROPOMI data and machine learning over East Asia," *Environ. Pollut.*, vol. 288, Nov. 2021, Art. no. 117711.
- [18] Y. Wang, Q. Yuan, T. Li, L. Zhu, and L. Zhang, "Estimating daily full-coverage near surface O_3 , CO, and NO_2 concentrations at a high spatial resolution over China based on S5P-TROPOMI and GEOS-FP," *ISPRS J. Photogramm. Remote Sens.*, vol. 175, pp. 311–325, May 2021.
- [19] J. Wei *et al.*, "Full-coverage mapping and spatiotemporal variations of ground-level ozone (O_3) pollution from 2013 to 2020 across China," *Remote Sens. Environ.*, vol. 270, Mar. 2021, Art. no. 112775.
- [20] Q. Di, S. Rowland, P. Koutrakis, and J. Schwartz, "A hybrid model for spatially and temporally resolved ozone exposures in the continental United States," *J. Air Waste Manage. Assoc.*, vol. 67, no. 1, pp. 39–52, Jan. 2017.
- [21] T. Li and X. Cheng, "Estimating daily full-coverage surface ozone concentration using satellite observations and a spatiotemporally embedded deep learning approach," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 101, Sep. 2021, Art. no. 102356.
- [22] S. Zhu, H. Zhu, J. Xu, Q. Zeng, D. Zhang, and X. Liu, "Satellite remote sensing of daily surface ozone in a mountainous area," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022.
- [23] Y. Wang, Q. Yuan, L. Zhu, and L. Zhang, "Spatiotemporal estimation of hourly 2-km ground-level ozone over China based on Himawari-8 using a self-adaptive geospatially local model," *Geosci. Frontiers*, vol. 13, no. 1, Jan. 2022, Art. no. 101286.
- [24] T. Li, H. Shen, Q. Yuan, and L. Zhang, "A locally weighted neural network constrained by global training for remote sensing estimation of $\text{PM}_{2.5}$," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2021, Art. no. 4102513.
- [25] Q. He and B. Huang, "Satellite-based mapping of daily high-resolution ground $\text{PM}_{2.5}$ in China via space-time regression modeling," *Remote Sens. Environ.*, vol. 206, pp. 72–83, Mar. 2018.
- [26] Q. He and B. Huang, "Satellite-based high-resolution $\text{PM}_{2.5}$ estimation over the Beijing-Tianjin-Hebei region of China using an improved geographically and temporally weighted regression model," *Environ. Pollut.*, vol. 236, pp. 1027–1037, May 2018.
- [27] Y. Bai, L. Wu, K. Qin, Y. Zhang, Y. Shen, and Y. Zhou, "A geographically and temporally weighted regression model for ground-level $\text{PM}_{2.5}$ estimation from satellite-derived 500 m resolution AOD," *Remote Sens.*, vol. 8, no. 3, p. 262, Mar. 2016.
- [28] A. S. Fotheringham, M. E. Charlton, and C. Brunsdon, "Geographically weighted regression: A natural evolution of the expansion method for spatial data analysis," *Environ. Planning A*, vol. 30, no. 11, pp. 1905–1927, 1998.
- [29] B. Huang, B. Wu, and M. Barry, "Geographically and temporally weighted regression for modeling spatio-temporal variation in house prices," *Int. J. Geograph. Inf. Sci.*, vol. 24, no. 3, pp. 383–401, Mar. 2010.
- [30] Y. Bengio, Y. LeCun, and G. E. Hinton, "Deep learning for AI," *Commun. ACM*, vol. 64, pp. 58–65, Jun. 2021.
- [31] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, p. 436, 2015.
- [32] T. Li, H. Shen, Q. Yuan, and L. Zhang, "Geographically and temporally weighted neural networks for satellite-based mapping of ground-level $\text{PM}_{2.5}$," *ISPRS J. Photogramm. Remote Sens.*, vol. 167, pp. 178–188, Sep. 2020.

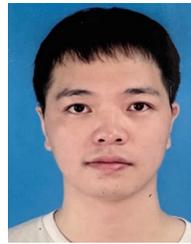
- [33] E. C. M. Hui, X. Li, T. Chen, and W. Lang, "Deciphering the spatial structure of China's megacity region: A new bay area—The Guangdong-Hong Kong-Macao Greater Bay area in the making," *ities*, vol. 105, Oct. 2020, Art. no. 102168.
- [34] *Blue Book of China's Atmospheric Ozone Pollution Prevention and Control*, PCOPC, Nanjing, China, 2020.
- [35] G. Yang, Y. Liu, and X. Li, "Spatiotemporal distribution of ground-level ozone in China at a city level," *Sci. Rep.*, vol. 10, no. 1, p. 7229, Dec. 2020.
- [36] K. Li, D. J. Jacob, L. Shen, X. Lu, I. De Smedt, and H. Liao, "Increases in surface ozone pollution in China from 2013 to 2019: Anthropogenic and meteorological influences," *Atmos. Chem. Phys.*, vol. 20, no. 19, pp. 11423–11433, Oct. 2020.
- [37] H. J. Eskes and K. U. Eichmann, "S5P MPC product readme nitrogen dioxide," Tech. Rep., 2020. [Online]. Available: <https://sentinel.esa.int/documents/247904/3541451/Sentinel-5P-Nitrogen-Dioxide-Level-2-Product-Readme-File>
- [38] J.-C. Lambert *et al.*, "Quarterly validation report of the Copernicus Sentinel-5 precursor operational data products #08: April 2018–August 2020," Tech. Rep., 2020. [Online]. Available: http://www.tropomi.eu/sites/default/files/files/publicS5P-MPC-IASB-ROCVR-02.0.2-20190411_FINAL.pdf
- [39] H. Hersbach *et al.*, "The ERA5 global reanalysis," *Quart. J. Roy. Meteorolog. Soc.*, vol. 146, no. 730, pp. 1999–2049, 2020.
- [40] Y. Xia, Y. Hu, Y. Huang, C. Zhao, F. Xie, and Y. Yang, "Significant contribution of severe ozone loss to the Siberian-Arctic surface warming in Spring 2020," *Geophys. Res. Lett.*, vol. 48, no. 8, 2021, Art. no. e2021GL092509.
- [41] Y. Xia *et al.*, "Concurrent hot extremes and high ultraviolet radiation in summer over the Yangtze plain and their possible impact on surface ozone," *Environ. Res. Lett.*, vol. 17, no. 6, Jun. 2022, Art. no. 064001.
- [42] D. F. Specht, "A general regression neural network," *IEEE Trans. Neural Netw.*, vol. 2, no. 6, pp. 568–576, Nov. 1991.
- [43] S. Wu, Z. Wang, Z. Du, B. Huang, F. Zhang, and R. Liu, "Geographically and temporally neural network weighted regression for modeling spatiotemporal non-stationary relationships," *Int. J. Geographical Inf. Sci.*, vol. 35, no. 3, pp. 582–608, Mar. 2021.
- [44] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, pp. 533–536, Oct. 1986.
- [45] D. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.
- [46] T. Li, H. Shen, C. Zeng, and Q. Yuan, "A validation approach considering the uneven distribution of ground stations for satellite-based PM_{2.5} estimation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 1312–1321, 2020.
- [47] Q. Yang, C. Jin, T. Li, Q. Yuan, H. Shen, and L. Zhang, "Research progress and challenges of data-driven quantitative remote sensing," *Nat. Remote Sens. Bull.*, vol. 26, no. 2, pp. 268–285, 2022.
- [48] Q. Yuan *et al.*, "Deep learning in environmental remote sensing: Achievements and challenges," *Remote Sens. Environ.*, vol. 241, May 2020, Art. no. 111716.
- [49] M. Reichstein, G. Camps-Valls, B. Stevens, M. Jung, J. Denzler, N. Carvalhais, and M. Prabhath, "Deep learning and process understanding for data-driven Earth system science," *Nature*, vol. 566, pp. 195–204, Feb. 2019.



Tongwen Li received the B.S. degree in geographic information systems and the Ph.D. degree in cartography and geographic information engineering from Wuhan University, Wuhan, China, in 2015 and 2020, respectively.

He is currently an Assistant Professor with the School of Geospatial Engineering and Science, Sun Yat-sen University, Zhuhai, China. He has published more than 30 papers in peer-reviewed international journals. His research interests include atmospheric remote sensing (including fine particulate matter,

ozone, carbon dioxide, and so on) and machine learning in environmental remote sensing.



Jingan Wu received the B.S. degree in geographic information systems from Anhui University, Hefei, China, in 2015, and the Ph.D. degree in cartography and geographic information engineering from Wuhan University, Wuhan, China, in 2021.

He is currently a Post-Doctoral Research Assistant with the School of Geospatial Engineering and Science, Sun Yat-sen University, Zhuhai, China. His research interests include spatiotemporal data fusion and missing information reconstruction for remote sensing images.



Jiajia Chen received the M.E. degree in geomatics engineering from Wuhan University, Wuhan, China, in 2018, where she is currently pursuing the Ph.D. degree in cartography and geographic information engineering.

Her research interests mainly include atmospheric remote sensing, multisource data fusion, and ozone air pollution.



Huanfeng Shen (Senior Member, IEEE) received the B.S. degree in surveying and mapping engineering and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2002 and 2007, respectively.

He is currently a Distinguished Professor with Wuhan University, where he also serves as the Dean of the School of Resource and Environmental Sciences. He was or is a Principal Investigator (PI) of two projects supported by the National Key Research and Development Program of China and six projects supported by the National Natural Science Foundation of China. He has authored or coauthored more than 150 peer-reviewed international journal articles, where over 60 appeared in IEEE journals, and published four books as a chief editor. His research interests include remote-sensing image processing, multisource data fusion, and intelligent environmental sensing.

Dr. Shen is also a fellow of the Institution of Engineering and Technology (IET), the Education Committee Member of the Chinese Society for Geodesy Photogrammetry and Cartography, and the Theory Committee Member of the Chinese Association for the Geospatial Information Society. He was a recipient of the First Prize in Natural Science Award of Hubei Province in 2011, the First Prize in Nature Scientific Award of China's Ministry of Education in 2015, and the First Prize in Scientific and Technological Progress Award of the Chinese Society for Geodesy Photogrammetry and Cartography in 2017. He is also a Senior Regional Editor of the *Journal of Applied Remote Sensing* and an Associate Editor of *Geography and Geo-Information Science* and *Journal of Remote Sensing*.