



Article

Spatiotemporally Continuous Reconstruction of Retrieved PM_{2.5} Data Using an Autogeo-Stacking Model in the Beijing-Tianjin-Hebei Region, China

Wenhao Chu ¹, Chunxiao Zhang ^{1,2,*} , Yuwei Zhao ¹, Rongrong Li ³ and Pengda Wu ⁴

¹ School of Information Engineering, China University of Geosciences in Beijing, No. 29, Xueyuan Road, Haidian District, Beijing 100083, China

² Observation and Research Station of Beijing Fangshan Comprehensive Exploration, Ministry of Natural Resources, Beijing 100083, China

³ Institute of Space and Earth Information Science, The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong, China

⁴ Chinese Academy of Surveying and Mapping, Beijing 100830, China

* Correspondence: zcx@cugb.edu.cn; Tel.: +86-186-1821-9549



Citation: Chu, W.; Zhang, C.; Zhao, Y.; Li, R.; Wu, P. Spatiotemporally Continuous Reconstruction of Retrieved PM_{2.5} Data Using an Autogeo-Stacking Model in the Beijing-Tianjin-Hebei Region, China. *Remote Sens.* **2022**, *14*, 4432. <https://doi.org/10.3390/rs14184432>

Academic Editors: Yong Ge, Xiaomei Yang and Lianfa Li

Received: 4 August 2022

Accepted: 4 September 2022

Published: 6 September 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Aerosol optical depth (AOD) observations have been widely used to generate wide-coverage PM_{2.5} retrievals due to the adverse effects of long-term exposure to PM_{2.5} and the sparsity and unevenness of monitoring sites. However, due to non-random missing and nighttime gaps in AOD products, obtaining spatiotemporally continuous hourly data with high accuracy has been a great challenge. Therefore, this study developed an automatic geo-intelligent stacking (autogeo-stacking) model, which contained seven sub-models of machine learning and was stacked through a Catboost model. The autogeo-stacking model used the automated feature engineering (autofeat) method to identify spatiotemporal characteristics of multi-source datasets and generate extra features through automatic non-linear changes of multiple original features. The 10-fold cross-validation (CV) evaluation was employed to evaluate the 24-h and continuous ground-level PM_{2.5} estimations in the Beijing-Tianjin-Hebei (BTH) region during 2018. The results showed that the autogeo-stacking model performed well in the study area with the coefficient of determination (R²) of 0.88, the root mean squared error (RMSE) of 17.38 µg/m³, and the mean absolute error (MAE) of 10.71 µg/m³. The estimated PM_{2.5} concentrations had an excellent performance during the day (8:00–18:00, local time) and night (19:00–07:00) (the cross-validation coefficient of determination (CV-R²): 0.90, 0.88), and captured hourly PM_{2.5} variations well, even in the severe ambient air pollution event. On the seasonal scale, the R² values from high to low were winter, autumn, spring, and summer, respectively. Compared with the original stacking model, the improvement of R² with the autofeat and hyperparameter optimization approaches was up to 5.33%. In addition, the annual mean values indicated that the southern areas, such as Shijiazhuang, Xingtai, and Handan, suffered higher PM_{2.5} concentrations. The northern regions (e.g., Zhangjiakou and Chengde) experienced low PM_{2.5}. In summary, the proposed method in this paper performed well and could provide ideas for constructing geo-features and spatiotemporally continuous inversion products of PM_{2.5}.

Keywords: autogeo-stacking model; 24-h PM_{2.5} mapping; spatiotemporal continuity; automated feature engineering

1. Introduction

Fine particulate matter with a diameter of less than 2.5 µm (referred to as PM_{2.5}) is one of the primary air pollutants which causes public health concerns. Long-term exposure to PM_{2.5} causes harms to humans, such as impaired vascular function, adverse effects to the lungs, and even premature death [1–3]. Accurate PM_{2.5} observations of ground-based stations have been available in China since January 2013. However, the sparsity

and unevenness of observation sites limited the spatiotemporal analysis at a regional scale. Hence, satellite remote sensing technology attracted researchers' attention due to its capability for wide-range imaging [4–6]. For instance, Wang et al., analyzed the relationship between satellite-derived aerosol optical thickness (AOT) and PM_{2.5} in 2003 [7]. Limited by the technology of satellite remote sensing then, follow-up studies are more focused on daily scale products that are incomplete in space and time with traditional methods, such as the linear mixed-effect (LME) model, the generalized additive model (GAM), and the geographically weighted regression (GWR) model [8–10].

With the advancement of a new generation of geostationary satellites (e.g., GOCI, Himawari-8, FY-4A), the spatial and temporal resolutions have improved significantly. The spatial resolution can be up to 0.5 km, and the temporal resolution to 10 min [11,12]. At the same time, the retrieval methods have improved; new models were proposed, such as the random forest (RF) method and the geographically and temporally weighted regression (GTWR) method. With the high-resolution images and the improved retrieval methods, robust and hourly PM_{2.5} retrieval becomes feasible. Studies that used machine learning models to estimate hourly PM_{2.5} concentrations are frequent. The deep neural network (DNN) approach, the Light Gradient Boosting Machine (LightGBM) method, and other models are applied to derive hourly PM_{2.5} concentrations based on images of these geostationary satellites [13–16]. Compared to traditional means (e.g., LME and GAM), the accuracy of estimated hourly PM_{2.5} has improved significantly. However, there is still a problem of spatiotemporal discontinuity.

Due to complex weather conditions, high surface reflectivity, and others, aerosol optical depth (AOD) observations encounter 40–80% non-random missingness on average [17]. In addition, the satellite's orbit patterns and surface topography also cause the missing of AOD (i.e., missing AOD gaps during the daytime, and nighttime AOD gaps) [18,19]. Various research has made efforts to obtain full-coverage products, such as Jiang et al., who integrated multi-source hourly AOD products to obtain continuous products of AOD and PM_{2.5} in space and time at 1 km via the two-stage random forest model in 2018 [20]. Spatiotemporally continuous PM_{2.5} retrieval can not only be produced from seamless AOD products but also can be completed based on PM_{2.5} products which contain missing values. According to Xiao et al., [21], the difference between filling AOD data gaps and filling PM_{2.5} data gaps is small, as the performances of the two are similar in the Beijing-Tianjin-Hebei (BTH) and the Yangtze River Delta (YRD) region (daily, 1 km) during 2013. From the perspective of filling PM_{2.5} gaps, Zhan et al., coupled the gradient boosting machine (GBM) model with the GWR model to retrieve daily and full-coverage 1 km resolution PM_{2.5} over China in 2014 [22]. Brokamp et al., trained two different random forest (RF) models to deal with the situations of AOD existence and absence to generate daily PM_{2.5} at 1 km in the seven-county area surrounding Cincinnati, OH, area from 2000 to 2015 [23]. Moreover, the spatiotemporal fusion methods (e.g., the spatial and temporal non-local filter based fusion model) also can be integrated to fill nighttime PM_{2.5} gaps based on the hourly PM_{2.5} retrievals of the geo-intelligent deep belief network (Geoi-DBN) method at 0.05° in the Wuhan Urban Agglomeration (WUA) region, 2016 [24,25].

In the studies cited above, various models were used to fit the relationship between PM_{2.5} and potential factors (e.g., reanalysis products, AOD observations, land use/land cover data). Moreover, these studies also employed spatiotemporal features, such as simple features (e.g., day of year, longitude, latitude) and complex spatiotemporal features (e.g., constructed features used in Geoi-DBN, space-time random forest (STRF), space-time extra-trees (STET), and fast space-time LightGBM (STLG)); please refer to supplementary Equations (S1)–(S11) [9,15,22,26–29]. These simple geoi-features are inherent space and time properties common to the selected datasets. The complicated geoi-features are the spatiotemporal autocorrelation of ground-based PM_{2.5} observations. These previous studies indicate that spatiotemporal information has improved the performance of models to some extent. These geoi-features enhance a specific type of data in space and time (e.g., PM_{2.5} observations or AOD observations) but do not fully consider the spatiotemporal influences

between multi-source datasets (e.g., $PM_{2.5}$ observations and AOD observations). The spatiotemporal influences between multi-source datasets are usually learned by retrieval models or designed by professional numeric model developers. Moreover, the process of proposing and verifying these kinds of geoi-features (also called feature engineering in machine learning) is troublesome. It is complicated, time-consuming, and requires high expertise [30,31]. For example, after a series of experiments and analyses, it was determined that the effects between secondary aerosol and ambient meteorological factors were presented as non-linear relationships, and the non-linear relationships were worthy of attention (e.g., AOD^2) [32,33].

To deal with these issues, we have developed an automatic geo-intelligent stacking (autogeoi-stacking) model to achieve high precision and seamless $PM_{2.5}$ concentrations by incorporating the automated feature engineering method (autofeat) into the model ensemble method. From the perspective of filling in $PM_{2.5}$ gaps (i.e., missing $PM_{2.5}$ gaps during the daytime, nighttime $PM_{2.5}$ gaps), the autogeoi-stacking model was trained for the day (8:00–18:00, local time) and night (19:00–7:00). In order to improve the utilization of the spatiotemporal relationship, a technology called autofeat was introduced that can automatically generate additional spatiotemporal features from candidate features. The autogeoi-stacking model has been used to generate 24 h and continuous $PM_{2.5}$ concentrations in the BTH region during 2018. The structure of the rest of this paper is as follows: Section 2 introduces the study region and materials. Section 3 provides details of the development of the autogeoi-stacking model. Section 4 demonstrates the impact of autofeat and models' performances with the validation of 24 h $PM_{2.5}$ estimation in both spatial and temporal dimensions. The paper is finished with the discussion of results and conclusions.

2. Study Area and Data

2.1. Study Region

As one of China's most influential urban agglomerations, the Beijing-Tianjin-Hebei (BTH) region contains tens of millions of people. However, the BTH region is one of the smoggiest areas in China due to industrial pollution, coal burning, vehicle exhaust, and other adverse factors. It is significant to obtain refined 24 h data in this area, which can help assess $PM_{2.5}$ exposure and formulate air pollution policies. The location of the BTH region is shown in Figure 1. The research period is the whole year of 2018 (local time), the same as the China high air pollutants (CHAP) dataset [15].

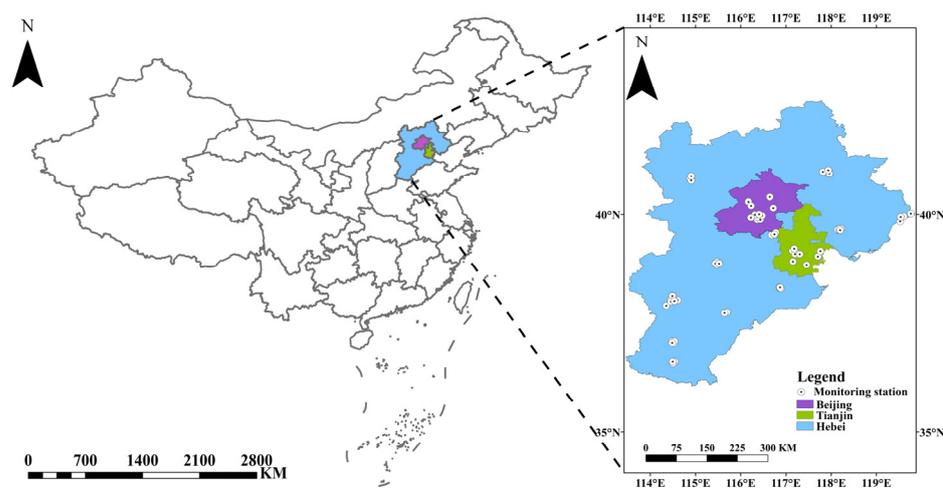


Figure 1. Study area and the spatial distributions of ground monitoring stations.

2.2. Data

This study collected 29 variables with potential effects on $PM_{2.5}$ concentrations, including station monitoring observations, satellite-derived $PM_{2.5}$ data, meteorological data, and topographic data. All selected products (Table 1) were preprocessed (e.g., data cleaning,

clip, define projection) and resampled to the exact spatial resolution ($0.05^\circ \times 0.05^\circ$) in the same coordinate system (WGS 84, EPSG:4326) through Python.

Table 1. Summary of datasets and sources used in this study.

| Categories | Abbreviation | Content | Unit | Spatial Resolution | Data Source |
|---------------------|-------------------|---------------------------|-------------------|----------------------------------|-------------|
| Ground Truth | PM _{2.5} | PM _{2.5} | μg/m ³ | Point | CNEMC |
| Satellite Retrieval | PM _{2.5} | PM _{2.5} | μg/m ³ | $0.05^\circ \times 0.05^\circ$ | CHAP |
| Meteorological | 10WU | 10 m u-component of wind | m/s | $0.25^\circ \times 0.25^\circ$ | ERA5 |
| | 10WV | 10 m v-component of wind | - | - | - |
| | 100WU | 100 m u-component of wind | - | - | - |
| | 100WV | 100 m v-component of wind | - | - | - |
| | T2M | 2 m temperature | K | - | - |
| | D2M | 2 m dewpoint temperature | - | - | - |
| | RH | Relative humidity | % | - | - |
| | SP | Surface pressure | Pa | - | - |
| | BLH | Boundary-layer height | m | - | - |
| | PRE | Total precipitation | - | - | - |
| | KX | K index | K | - | - |
| Aerosols | PM _{2.5} | PM _{2.5} | μg/m ³ | $0.5^\circ \times 0.625^\circ$ | MERRA2 |
| | BC | Black carbon aerosol | - | - | - |
| | OC | Organic carbon aerosol | - | - | - |
| | DUST | Dust aerosol | - | - | - |
| | SO ₄ | Sulfate aerosol | - | - | - |
| | SS | Sea salt aerosol | - | - | - |
| | CO | Carbon monoxide | - | $0.136^\circ \times 0.136^\circ$ | CAQRA |
| O ₃ | Ozone | - | - | - | |
| Topographic | SRTM | Surface elevation | m | 90 m | SRTM |

2.2.1. Ground-Level Measurements

Hourly data was collected from China National Environmental Monitoring Center (CNEMC) (<https://quotsoft.net/air>, accessed on 24 July 2022). To make full use of PM_{2.5} station measurements, all monitors within the range of 35.0°N to 43.0°N and 112.0°E to 120.0°E were included; 178 stations were included in this study. The distributions of stations in the BTH region are shown in Figure 1. As for the ground-based observations, data cleaning was made (e.g., eliminating NaN values and continuous repeated values).

2.2.2. China High Air Pollutants

Hourly data from the CHAP dataset was selected. The hourly data (08:00–18:00, local time) has been generated from Himawari-8 AOD products (missing gaps in space) together with other auxiliary data using the fast space-time LightGBM model (STLG) [15]. The STLG model consists of the LightGBM model and a constructed complex feature (please refer to supplementary Equations (S10) and (S11)). Hourly PM_{2.5} concentrations were evaluated using the monitoring data. Compared with ground-level observations, the results have shown the robustness of this hourly dataset (the cross-validation coefficient of determination, $CV\text{-}R^2 = 0.85$), with the root mean square error (RMSE) and the mean absolute error (MAE) of 13.62 and 8.49 μg/m³, respectively.

2.2.3. Reanalysis Information

The meteorological factors played vital roles in PM_{2.5} retrieval. Previous studies used a variety of reanalysis datasets to take advantage of these meteorological factors. Moreover, the European Centre for Medium-range Weather Forecasts (ECMWF) Reanalysis v5 (ERA5) data performs best in China according to the comparison of multiple reanalysis datasets (i.e., ERA5, the second Modern-Era Retrospective analysis for Research and Applications (MERRA-2), the Japanese 55-year Reanalysis (JRA55), and the NCEP/DOE Reanalysis 2 (NCEP-2) [34]. Due to the limitation of spatiotemporal resolution in this study, the ERA5 dataset from the ECMWF was the most suitable. Eleven ERA5 hourly meteorological parameters were used, such as surface temperature, relative humidity, surface pressure, boundary-layer height, and total precipitation.

At the same time, to improve the accuracy of the reconstruction, auxiliary data were also added [21]. A high-resolution air quality reanalysis dataset over China (CAQRA) produced by the Institute of Atmospheric Physics [35] was selected for this study. The CAQRA dataset was obtained from the assimilation of the ground-based observations from the CNEMC using the ensemble Kalman filter (EnKF) and Nested Air Quality Prediction Modeling System (NAQPMS). Two correlation factors were selected according to the spearman coefficient: CO and O₃.

As an influential reanalysis dataset, the MERRA-2 dataset was also added. This dataset is a comprehensive dataset resulting from multisource data, including ground-based observations, model simulations, and satellite observations [36]. It helps to provide information about trends in space and time [37]. Surface black carbon, organic carbon, dust, sulfate, and sea salt concentration data were obtained from this dataset. These five components are the significant chemical compositions of PM_{2.5}. The PM_{2.5} of the MERRA-2 dataset was calculated using Equation (S12) in the supplementary.

2.2.4. Ancillary Data

The Shuttle Radar Topography Mission (SRTM) products were employed in this study, which characterize the study area's digital elevation and surface deformation. In addition, the simple spatiotemporal characteristics of the data were included as input variables (i.e., latitude, longitude, day of year, day of week, week of year, month, day, hour, and season).

3. Research Methods

3.1. Machine Learning Models

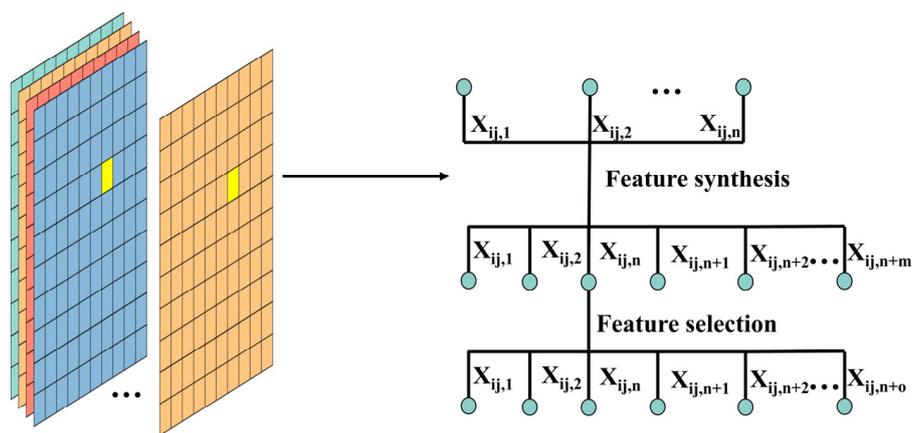
Here, seven models were adopted: random forest (RF), extremely randomized trees (ET), gradient boosting decision tree (GBDT), light gradient boosting machine (LightGBM), extreme gradient boosting (Xgboost), histogram-based gradient boosting machine (Hist-GBM), and Catboost. Each model was used by default for the regression task. For detailed comparisons between these models, please refer to Table A1 (Appendix A).

1. Random forest (RF): RF is an algorithm based on the bagging method and the classification and regression decision tree (CART) method, which was initially proposed by Breiman [38]. During the training process, RF will randomly sample many times and create a corresponding number of weak learners (i.e., various CART trees). At each node of the CART tree, the feature set (M) will be randomly selected from all the features (N , $N > M$), and the optimal feature set will be selected according to the size of the Gini coefficient. The randomness of sampling from the whole samples and the selection of feature set (M) improve the robustness of RF. The output of the regression model is the mean value of every weak learner.
2. Extremely randomized trees (ET): ET is a variant of RF that is more extreme and random than RF [39]. Unlike extensive random samplings from the data set during the training of RF, the ET method always trains on the entire data samples. In addition, when ET divides the feature set at each node, it does not select the optimal features according to the Gini coefficient but randomly divides the feature set. These differences enhance the generalization ability of ET to some extent.

3. Gradient boosting decision tree (GBDT): GBDT is a model based on the boosting method and CART regression tree, which differs from bagging in RF and ET [40]. During training, every iteration of GBDT is to fit the residual between the previous round of the model prediction result and the actual value due to the boosting method. In addition, GBDT can evaluate the importance of each feature (i.e., the selected independent variables) by calculating the frequency of every variable used.
4. Extreme gradient boosting (Xgboost): Xgboost is a new algorithm proposed under the framework of GBDT, and its performance is better than GBDT [41]. Xgboost extends the cost function by employing the second-order Taylor expansion, which further enhances the fitting ability of the model. Moreover, Xgboost utilizes L2 regularization to decrease the overfitting of the model. During training, it selects the optimal split node to handle with NaN values by comparing the split coefficient of the left and right nodes. Xgboost will sort all features before traversing them for segmentation (i.e., find the optimal node) and then generate the child node; this process will lead to additional memory occupation and time consumption, which is one of the reasons why LightGBM is proposed.
5. Light gradient boosting machine (LightGBM): LightGBM is an algorithm proposed after Xgboost; it is more optimized for the defects of GBDT and Xgboost [42]. LightGBM adopts the histogram algorithm. During training, by discretizing the continuous values of features into K integers, the histogram of the corresponding width is constructed (1/K represents a box). Then the box statistics are carried out according to the discretized value of the data. In this case, LightGBM can discover the optimal segmentation point in each box instead of traversing the entire set of discrete values, significantly decreasing memory usage and calculation costs and improving the algorithm's efficiency. Due to the lower memory usage and computation costs, LightGBM can handle large-scale data.
6. Histogram-based gradient boosting (HistGBM): HistGBM is a new algorithm based on GBDT, inspired by LightGBM [43]. As a GBDT method, HistGBM can also evaluate the importance of each feature. Moreover, it also adopts the histogram algorithm to decrease memory usage and calculation costs, thereby increasing the model's robustness, speed, and stability. The efficiency of HistGBM on large data sets is preferable to the original GBDT. It has native support for NaN values. Based on the potential coefficient, the HistGBM tree grows at each split point to determine whether samples with NaN values should go left or right during training.
7. Catboost: Catboost is a boosting algorithm proposed by Yandex under the GBDT framework, which is efficient, robust, and natively supports GPU acceleration [44]. It adopts an entirely symmetric tree, which is the same in the division of left and right nodes, to reduce the possibility of model over-fitting. In addition, Catboost adopts a novel algorithm called ordered boosting, which overcomes the prediction shift and gradient bias problems of the original GBDT algorithm. The model is compatible with processing categorical features and will automatically combine categorical features to generate more useful information (i.e., newly constructed variables) [45].

3.2. Automated Feature Engineering

Automated feature engineering (autofeat) is a way of training by automatically creating features from a candidate dataset and selecting a number of optimal ones from them. The autofeat method usually includes two parts: automatic feature synthesis and automatic feature selection [46]. The main procedure of the autofeat method is shown in Figure 2. Relevant studies have demonstrated its effects on time series forecasting, reinforcement learning, classification, and regression for tabular data [47–50]. Compared to manual feature engineering, autofeat is an automated process with the advantages of reducing time costs and not requiring expert knowledge. It also helps to discover and develop generic features while preserving the interpretability of the model. This study used a python library named autofeat (version 2.0.10) to employ the automated feature engineering method.



(a) All independent variables: X_1, X_2, \dots, X_n (b) Automated feature engineering

Figure 2. Schematic diagram of the automated engineering method (autofeat): (a) they represent all the selected potential variables from left to right, and the yellow area is the point X_{ij} (b) feature synthesis and feature selection in the process of the autofeat method ($n + m > n + o > n$).

Feature synthesis (i.e., deep feature synthesis) means the non-linear changes of original features and the combination of original and engineered features. Figure 3 indicates details about the feature synthesis of the autofeat method. For a given pixel, assume that a set A ($\{X_1, X_2, X_3, \dots, X_n\}$) consists of features $X_1, X_2, X_3, \dots, X_n$ (i.e., the selected independent variables). Taking the assumed features as an example, the main procedure of feature synthesis is as follows:

1. The process of first depth is the non-linear changes of raw data (e.g., $\log X_1, \sqrt{X_2}, \frac{1}{X_3}, X_1^2, \cos^{-1} X_2, e^{X_n}$).
2. The combination of constructed and original features happens in the second depth (e.g., $\log X_1 + \sqrt{X_2}, \sqrt{X_2} * \frac{1}{X_3}, \sqrt{\frac{1}{X_3} + X_1^2}, X_1^2 / e^{X_n}$). The following steps are identical to this step.

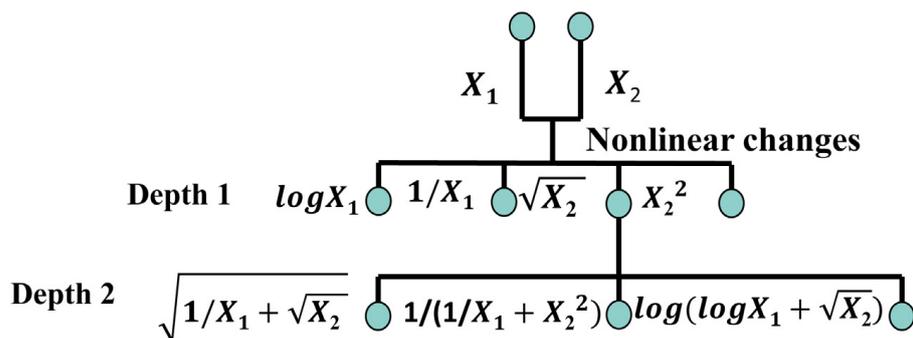


Figure 3. Schematic diagram of the feature synthesis part applies in the autofeat method for depths 1 and 2.

After feature synthesis, set A is combined with newly constructed features and turned into another set B, and then the feature selection works. Firstly, the constructed features highly related to the original features are excluded. Then, a multi-step feature selection approach starts based on the Lasso regression model and an L1-regularized logistic regression model. Algorithm 1 shows the pseudo-code of the feature selection. For more details, please refer to this article [46].

Algorithm 1. The pseudo-code of feature selection process

Process: Feature selection

Input: Set B, which contains m features

Output: Set D, which consists of l features ($l < m$)

1. Calculating the correlation of features, remove the features highly related to the original features, k features preserved ($k < m$);
2. Using the L1-regularized linear model to select the features to set C preliminarily, j features preserved ($j < k$);

for $i = 1$ to n **do**

The remaining features ($k - j$) are split into equal chunks (f , less than $(k - j)/2$);

Training models on the set C and each chunk (i.e., $j + f$);

Scoring models on the set C and each chunk;

end for

return set D

3.3. Model Development

Inspired by the model ensemble method, the two-stage model used in $PM_{2.5}$ concentrations retrieval, and the two targets used in soil moisture content retrieval [20,51], the autogeoi-stacking model was developed. The main procedure is shown in Figure 4. The proposed method contains stage 1 (the part before the model ensemble) and stage 2 (the model ensemble part). For each stage, a model is trained three times on the datasets in different steps (steps 0, 1, 2). In this study, the process of model training on the dataset (i.e., the raw training dataset) is called step 0. Step 1 refers to the model training on the dataset after the autofeat method of depth 1, and step 2 is the process of models employing the dataset, which contains the engineered features at depth 2.

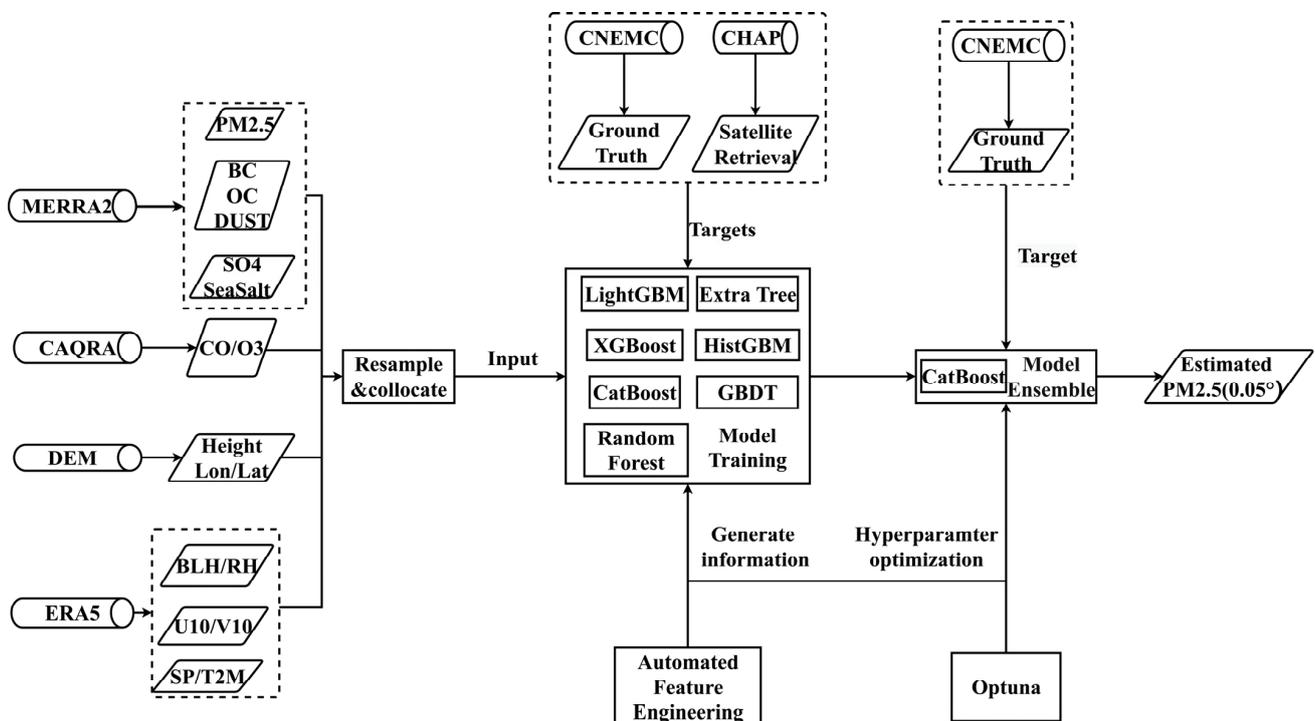


Figure 4. The main procedure of the proposed $PM_{2.5}$ reconstruction method developed in this study.

For each stage, there are four independent datasets: the training dataset, the validation dataset, the test set, and the out-of-station dataset. These four datasets were divided from the raw dataset. The raw dataset involved all available values of the selected variables according to the sites. NaN values in the targets were eliminated. The out-of-station dataset contained available values of 20 stations. The remaining part was divided into

three data sets according to the hold-out method [52]: 60% for model training (i.e., the training set), 20% for model tuning (i.e., the validation set), and 20% for model evaluation (i.e., the test set). Performances were validated based on the 10-fold cross-validation with the coefficient of determination (R^2), root mean squared error (RMSE), and mean absolute error (MAE) [53].

1. Stage 1 (the part before model ensemble): two targets were used: values from the CHAP dataset and ground truth (8:00–18:00, local time) to estimate $PM_{2.5}$ in this stage. The proposed method measured the spatial heterogeneity of in situ values and the temporal variation in ground-based observations toward the ground truth samples. When the target sample was the satellite retrieved data, the spatial heterogeneity of $PM_{2.5}$ distributions of the CHAP dataset was studied. The selected models were trained three times (step 0, 1, 2) with default parameters toward the two targets (total: $7 \times 3 \times 2 = 42$). The best one (at last: $7 \times 1 \times 2 = 14$) was chosen for the three steps. For the two targets, 14 models were retained at last. Another hyperparameter optimization process started via the optuna library (version 2.9.1). The numbers of iterations of hyperparameter optimization were different according to the training speed of the models.
2. Stage 2 (model ensemble part): Since $PM_{2.5}$ concentrations at night (19:00–07:00, local time) have not been inducted into the 14 models, a large prediction bias would be caused if the 24 h $PM_{2.5}$ distributions were generated directly. Thus, when training on the dataset, which consists of the estimated results of stage 1 (00:00–23:00, local time), the 24 h ground-based observations were used to limit the Catboost model in stage 2. The hyperparameters of the best model were tuned. For the Catboost method, the process of the hyperparameter optimization method and the optimal hyperparameters can be found in supplementary (Table S2, Figure S1).

4. Results

4.1. Autofeat Impacts

To reveal how the autofeat method influences the autogeoi-stacking model, an attempt at interpretable machine learning is made (Figure 5a–d); Shapley additive explanation (SHAP) as a classic post interpretation framework, was uniformly used for evaluation [54]. Moreover, the performances of cross-validation were evaluated based on the out-of-station dataset (Figure 5e,f).

The changes in the top ten features of the Catboost method in different steps are shown in Figure 5a–d. The feature X2 (the estimated values of the Catboost model to ground truth targets) is the most useful. The features X4, X8, X13, and X14 (the estimations of the ET, HistGBM models to the actual observations, and the estimations of the Xgboost model to the two targets) are always in the top ten features. The employed features in stage 2 are depicted in the supplementary. According to Figure 5e,f, the CV- R^2 results vary from 0.75 to 0.79 based on the out-of-station dataset, and the slopes of best-fit lines are between 0.74 and 0.78. From the changes of the top ten features and the improvement of accuracy, it can be seen that the autofeat method does work by constructing valuable features.

In the case of adding nine features constructed by the autofeat method, two of the top ten features were replaced by the newly added features (Figure 5b), with the performance almost unchanged (Figure 5f). With the newly constructed 66 features, Figure 5c,d displayed similar phenomena (i.e., two engineered features replaced two of the top 10 features) to Figure 5b. Taking the autofeat method into account, the CV- R^2 promotion of the Catboost model is approximately 1.33% (Figure 5g), and after optimizing the hyperparameters, the performance is improved by about 5.33% (Figure 5h). The improvement caused by the autofeat method is not particularly large, which may not be as effective as the hyperparameter optimization approach. However, taking the curse of dimensionality into consideration [55], the result is acceptable. In addition, the dimensional disaster caused by information redundancy requires an additional feature selection process [56], which needs

further studies. Consequently, it is recommended to integrate the autofeat method in step 2 and the hyperparameter optimization approach.

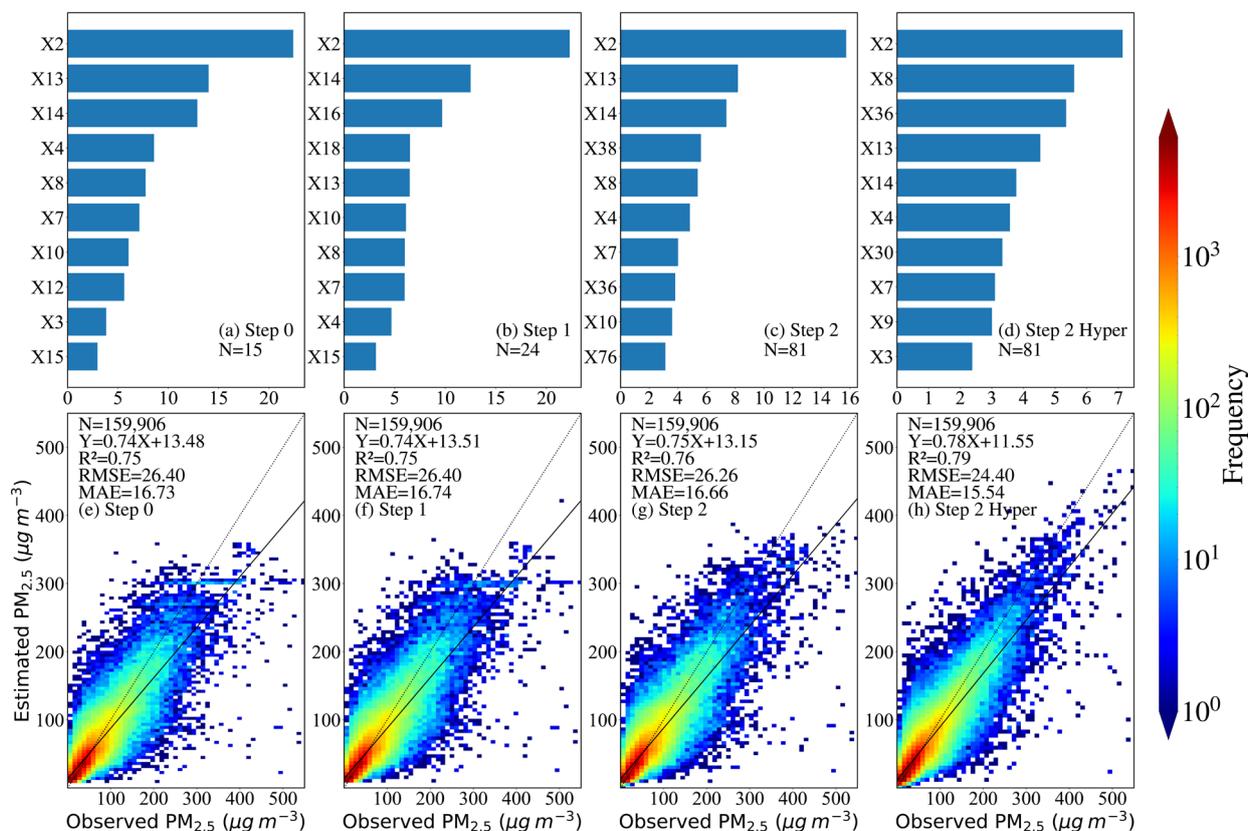


Figure 5. Sorted normalized SHAP value (%) of the top 10 features and the CV performance of the Catboost model in stage 2 based on the out-of-station dataset: (a–d) represent the sorted normalized SHAP value in different steps; (e–h) show the CV performances of models in different steps.

4.2. Spatial Performance

Spatial cross-validation performances of areas and stations are displayed in this part (Figures 6 and 7). Figure 6a reveals the robust ability of model prediction ($CV-R^2 = 0.88$) in the BTH region during 2018. In Beijing city, the proposed model shows a robust fitting result and a strong regression slope ($CV-R^2 = 0.89$, slope = 0.85) as before (Figure 6b). The proposed model performs best in Tianjin ($CV-R^2 = 0.89$, slope = 0.87) (Figure 6c). Performances of the proposed model in Hebei ($CV-R^2 = 0.88$, $CV-RMSE = 17.92 \mu g/m^3$) slightly dropped (Figure 6d), which may be due to worsening $PM_{2.5}$ pollution and more complicated geographical conditions (e.g., more mountains and terrain fluctuation). However, the high-value underestimation ($PM_{2.5}$ concentrations larger than $450 \mu g/m^3$) appears in Figure 6 as the large distance between the solid line and the dashed line. According to prior research, this situation may be solved through data augmentation in that part, such as the synthetic minority over-sampling technique (SMOTE) [57]. Overall, the results of each area in this study show that the model is robust ($CV-R^2$: 0.88–0.89, slope: 0.85–0.87), but the high-value underestimation deserves attention.

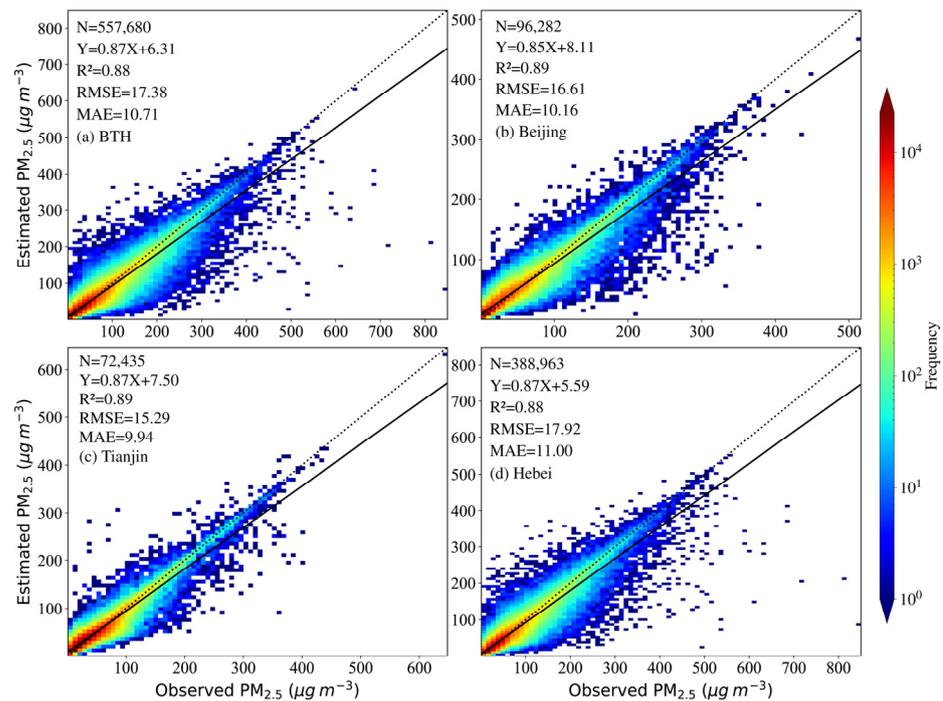


Figure 6. Density scatterplots of results of $PM_{2.5}$ estimates ($\mu\text{g}/\text{m}^3$) in different areas during 2018: (a) the BTH region (b) Beijing (c) Tianjin (d) Hebei. The dashed and solid lines denote 1:1 and best-fit lines from linear regression, respectively.

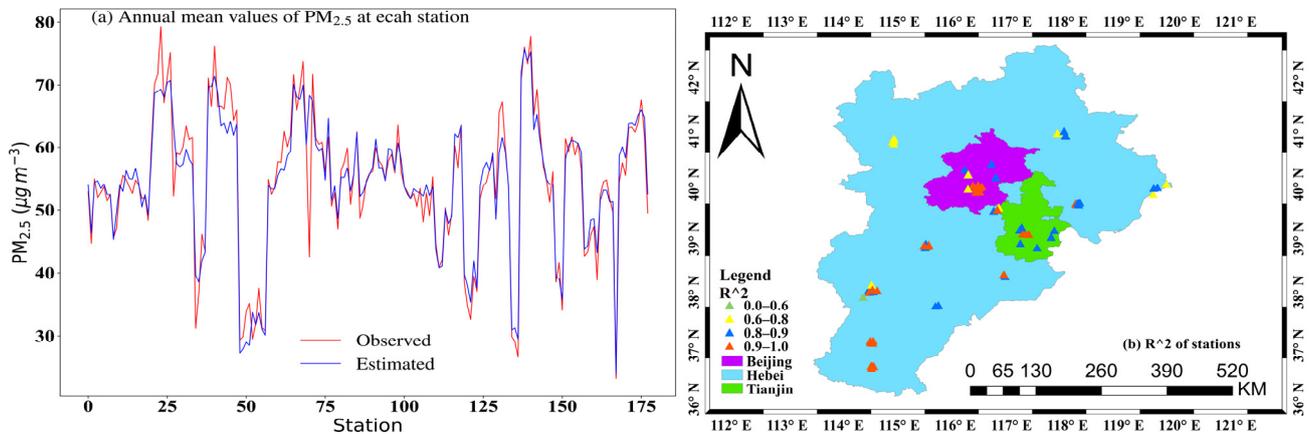


Figure 7. Station-based validation of accuracy in the BTH region during 2018: (a) annual mean values of $PM_{2.5}$ at each station; (b) $CV-R^2$ of stations.

Figure 7 exhibits the performance of each station in the BTH region during 2018. Figure 7a shows the annual mean values of $PM_{2.5}$ at each station used in this study. The observations and estimates at each station are generally similar. The maximum value of annual estimated $PM_{2.5}$ concentrations over all stations is $75.66 \mu\text{g}/\text{m}^3$ (the maximum ground-based $PM_{2.5}$ value: $79.32 \mu\text{g}/\text{m}^3$), and the minimum value is $23.79 \mu\text{g}/\text{m}^3$ (observed: $23.20 \mu\text{g}/\text{m}^3$). Figure 7b presents the $CV-R^2$ of stations in the BTH region, the performances of 13 stations have room for improvement ($CV-R^2 \leq 0.8$), and other stations perform well ($CV-R^2 > 0.8$). Figure S2 (refer to the supplementary) shows the locations and performances of these 13 stations. The $CV-R^2$ performances of ten stations are between 0.7 and 0.8, and the performances of the other three stations are below 0.7. The lower performances of these ten stations may be due to the generalization ability of the model, as the available data of the ten stations was selected as the out-of-station dataset. Two stations

(1065A and 1063A) endure lower $PM_{2.5}$ concentrations (annual mean values: $31.90 \mu\text{g}/\text{m}^3$, $31.27 \mu\text{g}/\text{m}^3$) with lower performances. The higher elevation and more complicated geographical conditions may be responsible for these at the two stations. Station 1035A has the lowest performance of all stations in the BTH region, which may be due to the irrational industrial structure that caused the higher $PM_{2.5}$ concentrations in Hebei province and the proposed model's lower ability in the high-value area (larger than $450 \mu\text{g}/\text{m}^3$). In general, the performances of spatial validation revealed that the autogeoi-stacking model well described the spatial heterogeneity of $PM_{2.5}$ distribution in the BTH region.

4.3. Temporal Performance

Applying the proposed approach, $PM_{2.5}$ concentrations in 2018 were estimated in space and time. To demonstrate the advantages of our method in 24 h $PM_{2.5}$ mapping and strengthen the understanding of severe ambient air pollution, we selected the values of 12 November to map the 24 h distributions in the BTH region (Figure 8) as a severe air pollution incident has occurred since that day and ended on November 15 [58]. Many meteorological conditions were unfavorable on that day, such as high relative humidity, low boundary layer height, high surface pressure, and low wind speed (Figures S3–S7). The worsening of ambient pollution from 0:00 (local time, mean $PM_{2.5}$ concentrations in the BTH region: $67.38 \mu\text{g}/\text{m}^3$) to 23:00 ($80.76 \mu\text{g}/\text{m}^3$) was shown in Figure 8 with the expansion of the red area (larger than $75 \mu\text{g}/\text{m}^3$). The most polluted area was southwest of the BTH region on 12 November 2018. With the spatiotemporal integrity of $PM_{2.5}$ concentrations, the temporal variation in $PM_{2.5}$ can easily be identified. From 0:00 to 4:00, the mean value of estimations has decreased from $67.38 \mu\text{g}/\text{m}^3$ to $59.76 \mu\text{g}/\text{m}^3$, which may be due to the less impact of human activity such as vehicle emissions. Then, the mean $PM_{2.5}$ concentration kept slowly increasing until 10:00 ($75.68 \mu\text{g}/\text{m}^3$), which may be related to vehicle emissions and adverse meteorological conditions such as high relative humidity and low boundary layer height. The drop happened until 16:00 ($68.62 \mu\text{g}/\text{m}^3$) due to the variations of boundary layer height. After 16 o'clock, because of the start of winter heating, the inversion of temperature, the vehicle emissions during the evening peak hours, the decrease in boundary layer height, and the increase in relative humidity, the level of $PM_{2.5}$ increased significantly until the end of the day and reached an early nighttime peak between 21:00 ($83.48 \mu\text{g}/\text{m}^3$) and 23:00 ($80.76 \mu\text{g}/\text{m}^3$). The diurnal variations were consistent with prior research [59–61]; for example, there are two peaks and the occurrence time of peaks is the same (i.e., 9:00–11:00, 20:00–23:00). Combined with Figure 9, the maximum average value of stations occurs at 21:00 (average estimated $PM_{2.5}$ in the BTH region: $83.48 \mu\text{g}/\text{m}^3$) in the estimated results (the mean value of estimated and observed $PM_{2.5}$ concentrations in situ: $107.66 \mu\text{g}/\text{m}^3$, $110.34 \mu\text{g}/\text{m}^3$). From the two subplots (Figure 9a,b), the trend of the estimated values is similar to that of the observations, such as the increase between 7:00 and 10:00 and between 14:00 and 21:00, the decrease from 10:00 to 14:00 and after 21:00. The range of average estimated $PM_{2.5}$ in stations (14.77 – $205.52 \mu\text{g}/\text{m}^3$) is similar to the mean observed concentrations (16.0 – $228.0 \mu\text{g}/\text{m}^3$). In general, the proposed autogeoi-stacking model well captured the 24 h variation even during a severe ambient pollution event.

Then, the time range is extended to day and night (Figure 10). The $CV\text{-}R^2$ result in the daytime (8:00–18:00, local time) is 0.90, which is higher than that at night (19:00–7:00) ($CV\text{-}R^2 = 0.88$), and other aspects such as the slope of best-fit lines and $CV\text{-}RMSE$, are also better ($CV\text{-}RMSE = 14.84 \mu\text{g}/\text{m}^3$). Figure 10b illustrates the robustness of the proposed model with a strong regression slope and high $CV\text{-}R^2$ at night, indicating that the constraints of $PM_{2.5}$ concentrations did work as mentioned in Section 3.3. As shown in Figure 10, the slopes of the best-fit lines (slopes = 0.89, 0.85) imply the low possibility of the low-value ($<450 \mu\text{g}/\text{m}^3$) overestimations. Moreover, the rising distance between the solid and dashed lines indicates that the proposed method underestimates the higher value (larger than $450 \mu\text{g}/\text{m}^3$) more seriously compared to the low-value area.

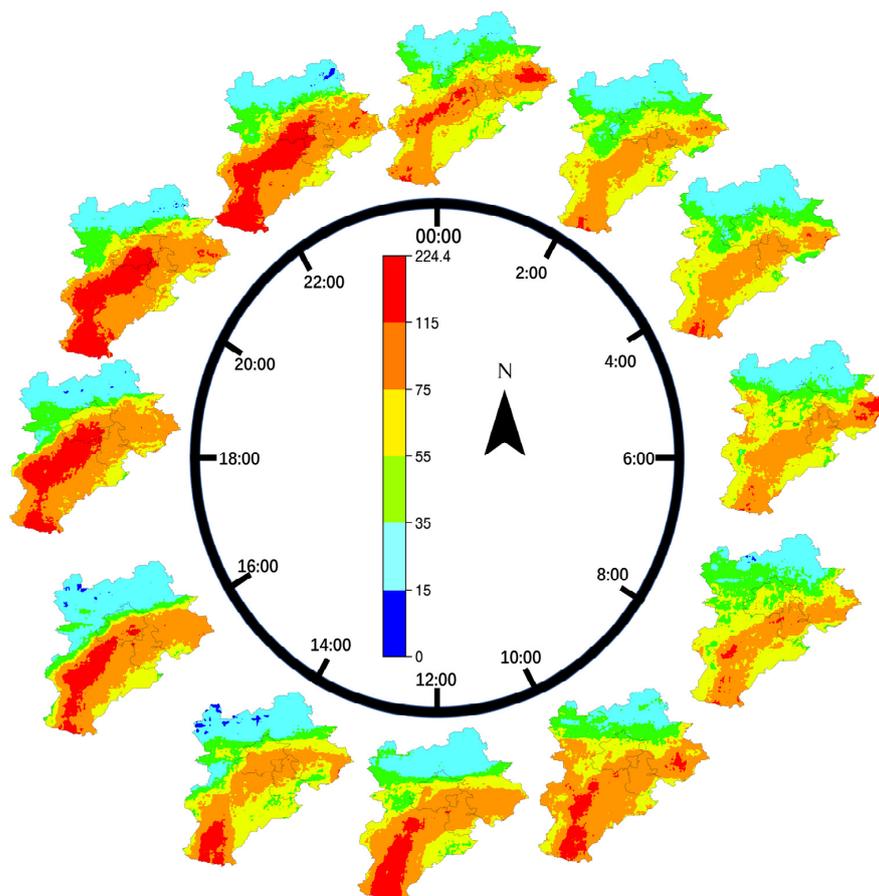


Figure 8. The hourly and continuous PM_{2.5} concentration in the BTH area was mapped on 12 November 2018 (local time).

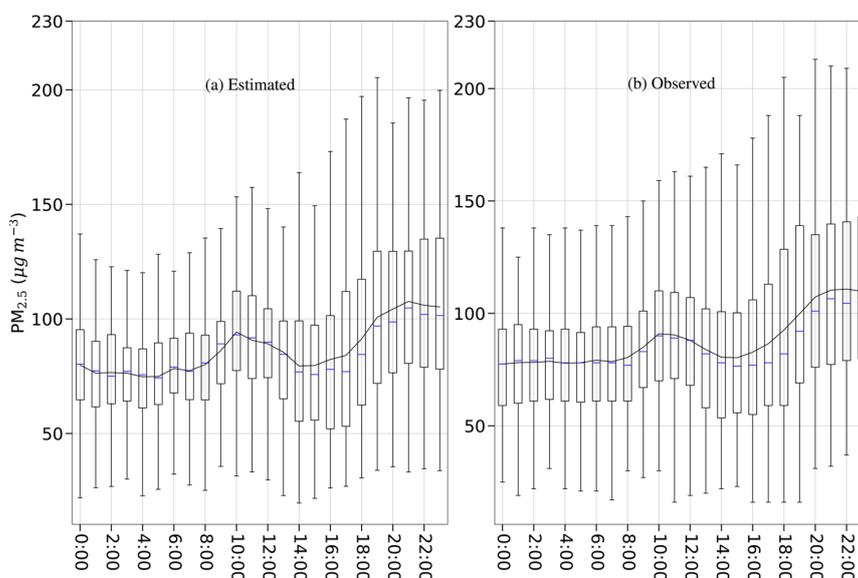


Figure 9. Boxplots of the temporal variance of PM_{2.5} concentrations on 11 November 2018: (a) mean estimated values of the proposed model at each station; (b) average ground truth of monitoring sites. In each box, the lowest and highest lines are the range of PM_{2.5} concentrations, and the three lines forming the box represent the 75th percentile, median, and 25th percentile values. The solid black lines in the two figures represent the average values for the whole 24 h.

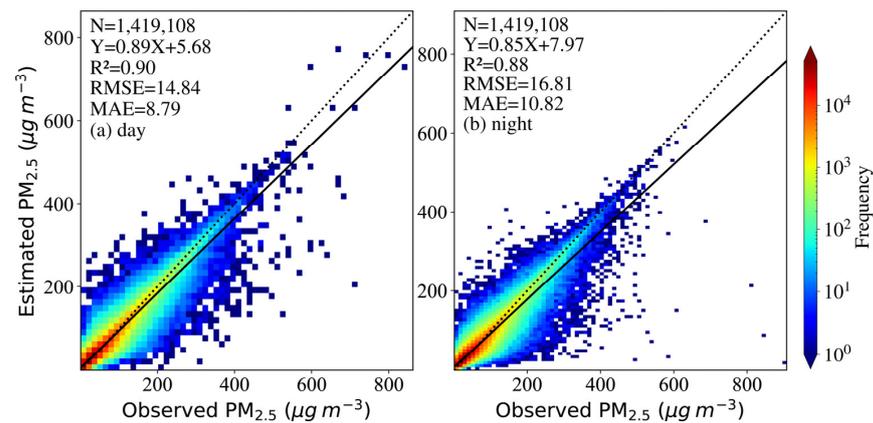


Figure 10. Density scatterplots of results of $PM_{2.5}$ estimates ($\mu\text{g}/\text{m}^3$) during daytime and night: (a) day and (b) night. The dashed and solid lines denote 1:1 and best-fit lines from linear regression.

Another validation in different seasons was also made (Figure 11), and the $CV-R^2$ results vary from 0.67 to 0.91. The performance of our proposed model is poor during summer ($CV-R^2 = 0.67$), even with the lowest $CV-RMSE$ and $CV-MAE$ over the four seasons ($CV-RMSE = 12.24 \mu\text{g}/\text{m}^3$, $CV-MAE = 8.47 \mu\text{g}/\text{m}^3$); this has also happened according to previous studies (Table 2) [12,13,17,22,23]. The lower R^2 and $RMSE$ can be conducted to the under-fitting of the proposed model and the lower variance of data in summer (details in supplementary Equations (S13)–(S15)). The under-fitting may be due to the frequent weather changes, the cloud cover in summer, the nighttime error, the additional uncertainty in the CHAP dataset, and the prediction errors of the proposed model in summer. The worsening performance in summer is due to the nighttime error and the uncertainty in the CHAP dataset. Moreover, the proposed model performs excellently over the other three seasons ($CV-R^2$: 0.86–0.91, $CV-RMSE$: 15.10–19.40 $\mu\text{g}/\text{m}^3$). Compared with previous research, the performances of our model are quite robust ($CV-R^2 = 0.86, 0.67, 0.89, 0.91$; $CV-RMSE = 16.69 \mu\text{g}/\text{m}^3, 12.24 \mu\text{g}/\text{m}^3, 15.10 \mu\text{g}/\text{m}^3, 19.40 \mu\text{g}/\text{m}^3$).

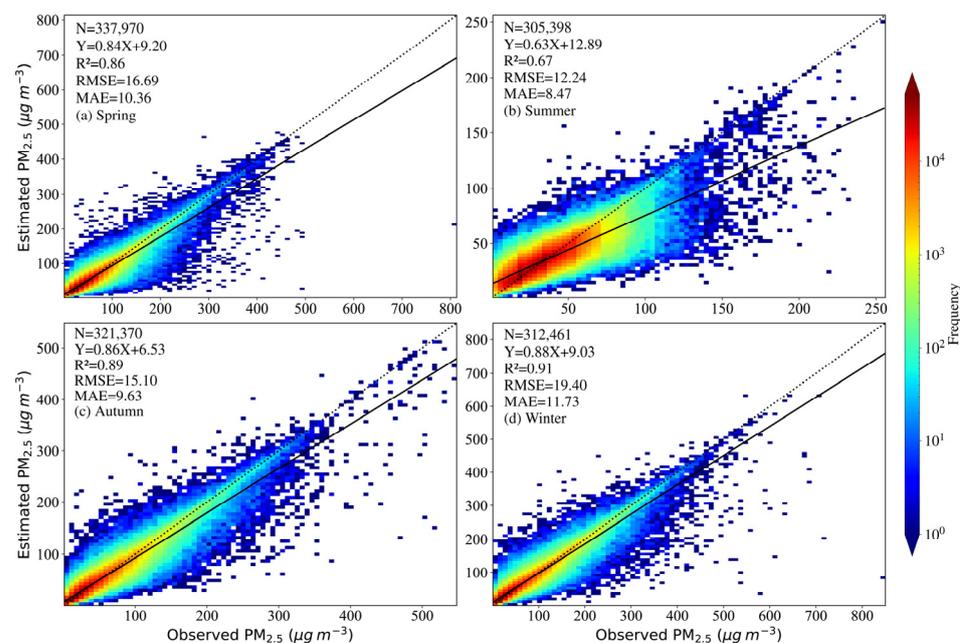


Figure 11. Density scatterplots of results of $PM_{2.5}$ estimates ($\mu\text{g}/\text{m}^3$) in different seasons, and the dashed and solid lines denote 1:1 and best-fit lines from linear regression: (a) spring (March to May), (b) summer (June to August), (c) autumn (September to November), and (d) winter (December to February).

Table 2. Comparison of the seasonal performances with previous studies, and the two lines (R^2 , RMSE) from left to right represent spring, summer, autumn, and winter.

| Study | Domain | Resolution | Gaps | Period | R^2 | RMSE ($\mu\text{g}/\text{m}^3$) |
|-------------------------|--------|---------------|------|-----------|------------------------|-----------------------------------|
| Wei et al., 2021 [12] | China | 0.05°, hourly | Yes | 2018 | 0.82, 0.71, 0.87, 0.86 | 14.55, 9.63, 11.83, 17.57 |
| Chen et al., 2019 [13] | China | 0.05°, hourly | Yes | 2016 | 0.82, 0.72, 0.86, 0.86 | 15.90, 11.00, 16.40, 21.40 |
| Jiang et al., 2021 [17] | China | 1 km, hourly | No | 2018–2019 | 0.85, 0.80, 0.85, 0.90 | 12.27, 7.78, 9.50, 14.83 |
| Wei et al., 2020 [24] | China | 1 km, daily | Yes | 2017–2018 | 0.88, 0.79, 0.90, 0.88 | 11.23, 7.23, 8.97, 12.84 |
| Wei et al., 2019 [25] | China | 1 km, daily | Yes | 2015–2016 | 0.81, 0.69, 0.84, 0.85 | 14.79, 9.62, 14.59, 20.06 |
| This study | BTH | 0.05°, hourly | No | 2018 | 0.86, 0.67, 0.89, 0.91 | 16.69, 12.24, 15.10, 19.40 |

Distributions of annual mean $\text{PM}_{2.5}$ concentrations are displayed in Figure 12. The differences between the CHAP dataset and the estimates of the autogeoi-stacking model are compared. The mean estimated value in the BTH region is $43.36 \mu\text{g}/\text{m}^3$. From the perspective of less than $35 \mu\text{g}/\text{m}^3$ and larger than $35 \mu\text{g}/\text{m}^3$, the spatial distributions of $\text{PM}_{2.5}$ is consistent (Figure 12a,b), which is also close to the distributions in previous studies [62,63]. The mean value of the estimated $\text{PM}_{2.5}$, which is less than $35 \mu\text{g}/\text{m}^3$ is $25.57 \mu\text{g}/\text{m}^3$ (CHAP: $30.19 \mu\text{g}/\text{m}^3$), and the mean value of the part which is greater than $35 \mu\text{g}/\text{m}^3$ is $52.94 \mu\text{g}/\text{m}^3$ (CHAP: $51.57 \mu\text{g}/\text{m}^3$). Due to the completeness of $\text{PM}_{2.5}$ in space and time, more details can be observed, such as the area less than $15 \mu\text{g}/\text{m}^3$ and the larger red area (greater than $55 \mu\text{g}/\text{m}^3$). The estimated $\text{PM}_{2.5}$ concentrations in the BTH region are consistent with those in the stations, such as the area larger than $55 \mu\text{g}/\text{m}^3$ in Handan (actual: $67.08 \mu\text{g}/\text{m}^3$, estimated: $69.32 \mu\text{g}/\text{m}^3$), Xingtai (actual: $68.61 \mu\text{g}/\text{m}^3$, estimated: $68.18 \mu\text{g}/\text{m}^3$), and Shijiazhuang City (actual: $68.85 \mu\text{g}/\text{m}^3$, estimated: $67.81 \mu\text{g}/\text{m}^3$), the yellow area ($35\text{--}55 \mu\text{g}/\text{m}^3$) in Beijing (actual: $48.07 \mu\text{g}/\text{m}^3$, estimated: $50.74 \mu\text{g}/\text{m}^3$), Tianjin City (actual: $51.88 \mu\text{g}/\text{m}^3$, estimated: $53.89 \mu\text{g}/\text{m}^3$), and the green ($35\text{--}55 \mu\text{g}/\text{m}^3$) area in Zhangjiakou (actual: $30.63 \mu\text{g}/\text{m}^3$, estimated: $27.12 \mu\text{g}/\text{m}^3$), Chengde City (actual: $31.71 \mu\text{g}/\text{m}^3$, estimated: $31.67 \mu\text{g}/\text{m}^3$). The $\text{PM}_{2.5}$ concentrations are lower in the north, such as Zhangjiakou ($27.12 \mu\text{g}/\text{m}^3$) and Chengde ($31.67 \mu\text{g}/\text{m}^3$), and are higher in the south: Shijiazhuang ($67.81 \mu\text{g}/\text{m}^3$), Xingtai ($68.18 \mu\text{g}/\text{m}^3$), and Handan ($69.32 \mu\text{g}/\text{m}^3$). The higher $\text{PM}_{2.5}$ concentrations might be caused by the development of heavy industries, adverse geographical conditions, and meteorological conditions (e.g., less wind, increasing relative humidity). The lower can be due to the developed tourism, favorable topographic conditions (e.g., dense mountains, more greening), and meteorological factors (e.g., frequent wind).

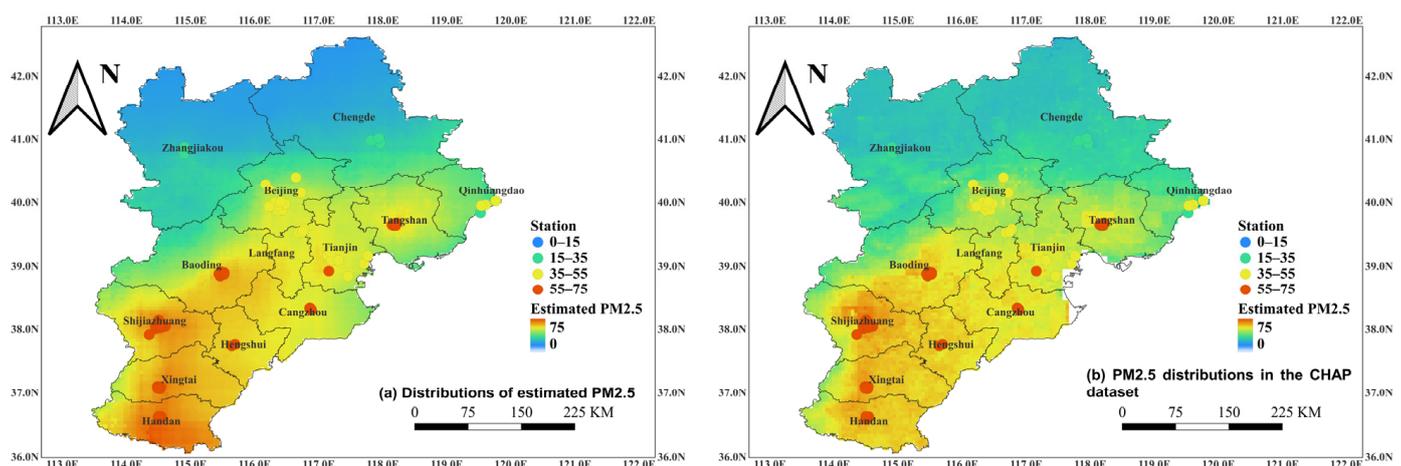


Figure 12. Distributions of annual mean values in the BTH area. (a) Estimated $\text{PM}_{2.5}$ based on the proposed method. (b) Estimated $\text{PM}_{2.5}$ of CHAP dataset.

5. Discussion

Towards the two issues: (a) spatiotemporal discontinuity of the hourly retrieved $\text{PM}_{2.5}$ concentrations and (b) the complexity of constructing and less utilizing spatiotemporal relationship, an attempt was made via the model ensemble approach and the autofeat

method in this study. 24 h $PM_{2.5}$ concentrations were continuously reconstructed in space and time based on the autogeoi-stacking model. The spatial and temporal performances of the autogeoi-stacking model were verified. Whether in out-of-station ($CV-R^2 = 0.75\text{--}0.79$, Figure 5) or regional validation ($CV-R^2 = 0.88\text{--}0.89$, Figure 6), the model performs very well, which serves to identify hot and cold spots of $PM_{2.5}$ in the BTH region. In temporal verification, the model behaves robustly ($CV-R^2 = 0.88\text{--}0.90$, Figure 10); it reveals the changes of $PM_{2.5}$ over 24 h. These results confirm the effectiveness of the autogeoi-stacking model in dealing with spatiotemporal discontinuities.

To address the second issue, the autofeat method was introduced. The impacts of the autofeat approach have been verified by the SHAP values and the performances (Figure 5). Compared to the original stacking model, the improvement brought by the autofeat method is about 1.33%. The improvement is up to 5.33% with the autofeat method and the process of hyperparameter optimization. Considering the high dimensions ($N = 81$) caused by the autofeat method, information redundancy may be responsible for the lower improvement (1.33%). Thus, an additional feature selection process could be useful in further studies. Moreover, autofeat does not conflict with hyperparameters optimization, and the two are suggested to do together to increase the upper limit of models. At the same time, the autogeoi-stacking model can also be considered to combine with the previously constructed features (e.g., features used in Geoi-DBN, STLG).

The proposed model works well through validation in spatial and temporal dimensions. However, there are still limitations: (a) the phenomena of underestimating the higher $PM_{2.5}$ concentrations ($\geq 450 \mu\text{g}/\text{m}^3$) and (b) the performances in summer. These two are typical problematic situations, according to previous research [15,16,25,26]. The former is due to the small number of samples in the high-value area with severe air pollution. Considering the environmental and health problems caused by the severe air pollution event [64], it is essential to enhance the performance of model in the high-value area. Moreover, the SMOTE method, which can be applied to generate synthetic samples utilizing information about the minority class in the training data [65], may be helpful. This method has been used to deal with this problem in previous studies [57,66,67]. The second limitation is much more complicated and is related to weather conditions such as relative humidity, temperature, and cloud cover [16,68]. It may be helpful to apply one model to fit the data in summer and another separate model to fit the data in other seasons. Moreover, the potential factors such as Normalized Difference Vegetation Index (NDVI), land use/land cover data, and population may also be helpful to enhance the performance in summer.

6. Conclusions

In this study, an autogeoi-stacking model was proposed to deal with the spatiotemporal discontinuity and the utilization of spatiotemporal characteristics. Based on the autofeat method and the model ensemble approach, the autogeoi-stacking model was developed, including the Catboost, ET, GBDT, HistGBM, LightGBM, RF, Xgboost approaches, and stacks through a Catboost model. A spatiotemporally continuous $PM_{2.5}$ retrieval data at the hourly level was obtained in the BTH region. The 10-fold cross-validation method was used for evaluation.

The proposed model works well at the regional scale ($CV-R^2 = 0.88\text{--}0.89$). The autogeoi-stacking model has excellent performance during the day and night ($CV-R^2 = 0.90$, $CV-R^2 = 0.88$) and can capture the hourly variations well even in the severe ambient air pollution event. On the seasonal scale, the R^2 values from high to low are winter, autumn, spring, and summer. Distributions of annual mean values show that the southern areas suffer higher $PM_{2.5}$ concentrations and the northern regions experience lower in the BTH region. In consideration of the autofeat method, it improved the $CV-R^2$ result to some extent (1.33%, 5.33%). In general, the proposed method can provide ideas for spatiotemporal continuity data and constructing geoi-features. Nevertheless, there is still room for improvement along this line of research. Further studies should concentrate on (a) the

high-value area, (b) the performance in summer, and (c) the combination of constructed features in previous studies.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/rs14184432/s1>, Figure S1: Two optimization objectives in the process of optimizing the hyperparameters of CBR model with the optuna library (version 2.9.1), RMSE on the left and log (MAE) on the right.; Figure S2: Distributions of the 13 stations that perform lower than other stations used in this study; Figure S3: Distributions of the boundary layer height in the BTH region on 12 November 2018; Figure S4: Distributions of the relative humidity in the BTH region on 12 November 2018; Figure S5: Distributions of the surface pressure in the BTH region on 12 November 2018. The 101,325 Pa means the standard atmospheric pressure; Figure S6: Distributions of the wind direction in the BTH region on 12 November 2018. 0 (360) means the north direction. 90 represents the east direction. 180 is the south direction and 270 indicates the west direction; Figure S7: Distributions of the wind speed in the BTH region on 12 November 2018; Table S1: All features used in the stage 2 step 0. The features are the output of the model at different steps; Table S2: Top 10 features of the Catboost method in stage 2, * means the useful information; Table S3: Hyperparameters of the Catboost method in Stage 2, Step 2.

Author Contributions: The authors' contributions are as follows: methodology, validation and writing—original draft, W.C.; writing—reviewing and editing, funding acquisition, C.Z.; visualization and editing the manuscript, Y.Z.; writing—reviewing and editing, R.L. and P.W. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Natural Science Foundation of China [grant numbers 61872325, 62172373]; the Fundamental Research Funds for the Central Universities, China [grant numbers 2652019028, 2652018082].

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: The CHAP dataset is provided by Jing Wei from the University of Maryland, thanks for him and this dataset. Moreover, we also wish to thank Zhaocheng Zeng from Peking University for his help in writing and editing.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

| | |
|-------------------|---|
| AOD | Aerosol Optical Depth |
| AOT | Aerosol optical thickness |
| Autofeat | Automated feature engineering |
| Autogeo-stacking | The automatic geo-intelligent stacking model |
| BTH | Beijing-Tianjin-Hebei |
| CAQRA | A high-resolution air quality reanalysis dataset over China |
| CART | Classification and regression decision tree |
| CNEMC | China National Environmental Monitoring Center |
| CV | Cross-validation |
| CV-R ² | Cross-validation coefficient of determination |
| DNN | The deep neural network method |
| ECMWF | European Centre for Medium-range Weather Forecasts |
| EnKF | Ensemble Kalman filter |
| ERA5 | ECMWF Reanalysis v5 |
| ET | Extremely randomized trees |
| GAM | Generalized additive model |
| GBDT | Gradient boosting decision tree |
| GBM | Gradient boosting machine |
| Geoi-DBN | Geo-intelligent deep belief network |
| GTWR | Geographically and temporally weighted regression model |
| GWR | Geographically weighted regression model |
| HistGBM | Histogram-based gradient boosting machine |

| | |
|-------------------|--|
| JRA55 | The Japanese 55-year Reanalysis |
| LightGBM | Light Gradient Boosting Machine |
| LME | Linear mixed-effect model |
| MAE | Mean absolute error |
| MERRA-2 | The second Modern-Era Retrospective analysis for Research and Applications |
| NAQPMS | Nested Air Quality Prediction Modeling System |
| NCEP-2 | The NCEP/DOE Reanalysis 2 |
| NDVI | Normalized Difference Vegetation Index |
| PM _{2.5} | Fine particulate matter with a diameter of less than 2.5 µm |
| R ² | Coefficient of determination |
| RF | Random forest |
| RMSE | Root-mean square error |
| SHAP | Shapley additive explanation |
| SMOTE | Synthetic minority over-sampling technique |
| SRTM | The Shuttle Radar Topography Mission |
| STET | The space-time extra-trees method |
| STLG | The fast space-time LightGBM model |
| STRF | The space-time random forest model |
| Xgboost | Extreme gradient boosting |
| YRD | Yangtze River Delta |

Appendix A

Table A1. Comparison of all machine learning models from the training device, RAM (mean value in stage 1), mean R², and time cost during inference aspects.

| Model | Training Device | RAM (M) | Mean R ² | Time Cost (Minute) |
|----------|-----------------|---------|---------------------|--------------------|
| Catboost | GPU | 0.94 | 0.78 | 2 h 6 |
| ET | CPU | 1588 | 0.87 | 1 h 52 |
| GBDT | CPU | 0.17 | 0.72 | 2 h 24 |
| HistGBM | CPU | 0.43 | 0.80 | 2 h 36 |
| LightGBM | CPU | 0.28 | 0.80 | 5 h 10 |
| RF | CPU | 992.64 | 0.85 | 1 h 58 |
| Xgboost | CPU | 0.66 | 0.82 | 1 h 46 |

References

1. Brook, R.D.; Urch, B.; Dvonch, J.T.; Bard, R.L.; Speck, M.; Keeler, G.; Morishita, M.; Marsik, F.J.; Kamal, A.S.; Kaciroti, N.; et al. Insights Into the Mechanisms and Mediators of the Effects of Air Pollution Exposure on Blood Pressure and Vascular Function in Healthy Humans. *Hypertension* **2009**, *54*, 659–667. [[CrossRef](#)] [[PubMed](#)]
2. Xing, Y.-F.; Xu, Y.-H.; Shi, M.-H.; Lian, Y.-X. The Impact of PM_{2.5} on the Human Respiratory System. *J. Thorac. Dis.* **2016**, *8*, 6.
3. Shi, Y.; Zhao, A.; Matsunaga, T.; Yamaguchi, Y.; Zang, S.; Li, Z.; Yu, T.; Gu, X. Underlying Causes of PM_{2.5}-Induced Premature Mortality and Potential Health Benefits of Air Pollution Control in South and Southeast Asia from 1999 to 2014. *Environ. Int.* **2018**, *121*, 814–823. [[CrossRef](#)] [[PubMed](#)]
4. Xu, Y.; Huang, Y.; Guo, Z. Influence of AOD Remotely Sensed Products, Meteorological Parameters, and AOD-PM_{2.5} Models on the PM_{2.5} Estimation. *Stoch. Environ. Res. Risk Assess.* **2021**, *35*, 893–908. [[CrossRef](#)]
5. Lin, C.; Labzovskii, L.D.; Mak, H.W.L.; Fung, J.C.H.; Lau, A.K.H.; Kenea, S.T.; Bilal, M.; Hey, J.D.V.; Lu, X.; Ma, J. Observation of PM_{2.5} Using a Combination of Satellite Remote Sensing and Low-Cost Sensor Network in Siberian Urban Areas with Limited Reference Monitoring. *Atmos. Environ.* **2020**, *227*, 117410. [[CrossRef](#)]
6. Li, J.; Zhang, H.; Chao, C.-Y.; Chien, C.-H.; Wu, C.-Y.; Luo, C.H.; Chen, L.-J.; Biswas, P. Integrating low-cost air quality sensor networks with fixed and satellite monitoring systems to study ground-level PM_{2.5}. *Atmos. Environ.* **2020**, *223*, 117293. [[CrossRef](#)]
7. Wang, J.; Christopher, S.A. Intercomparison between Satellite-Derived Aerosol Optical Thickness and PM_{2.5} Mass: Implications for Air Quality Studies. *Geophys. Res. Lett.* **2003**, *30*, 2095. [[CrossRef](#)]
8. Xie, Y.; Wang, Y.; Zhang, K.; Dong, W.; Lv, B.; Bai, Y. Daily Estimation of Ground-Level PM_{2.5} Concentrations over Beijing Using 3 Km Resolution MODIS AOD. *Environ. Sci. Technol.* **2015**, *49*, 12280–12288. [[CrossRef](#)]
9. Guo, Y.; Tang, Q.; Gong, D.-Y.; Zhang, Z. Estimating Ground-Level PM_{2.5} Concentrations in Beijing Using a Satellite-Based Geographically and Temporally Weighted Regression Model. *Remote Sens. Environ.* **2017**, *198*, 140–149. [[CrossRef](#)]
10. Ma, Z.; Hu, X.; Huang, L.; Bi, J.; Liu, Y. Estimating Ground-Level PM_{2.5} in China Using Satellite Remote Sensing. *Environ. Sci. Technol.* **2014**, *48*, 7436–7444. [[CrossRef](#)]

11. Ranjan, A.K.; Patra, A.K.; Gorai, A.K. A Review on Estimation of Particulate Matter from Satellite-Based Aerosol Optical Depth: Data, Methods, and Challenges. *Asia-Pac. J. Atmos. Sci.* **2021**, *57*, 679–699. [[CrossRef](#)]
12. Zhang, Y.; Li, Z.; Bai, K.; Wei, Y.; Xie, Y.; Zhang, Y.; Ou, Y.; Cohen, J.; Zhang, Y.; Peng, Z.; et al. Satellite remote sensing of atmospheric particulate matter mass concentration: Advances, challenges, and perspectives. *Fundam. Res.* **2021**, *1*, 240–258. [[CrossRef](#)]
13. Lee, C.; Lee, K.; Kim, S.; Yu, J.; Jeong, S.; Yeom, J. Hourly Ground-Level PM_{2.5} Estimation Using Geostationary Satellite and Reanalysis Data via Deep Learning. *Remote Sens.* **2021**, *13*, 2121. [[CrossRef](#)]
14. Lu, X.; Wang, J.; Yan, Y.; Zhou, L.; Ma, W. Estimating Hourly PM_{2.5} Concentrations Using Himawari-8 AOD and a DBSCAN-Modified Deep Learning Model over the YRDUA, China. *Atmos. Pollut. Res.* **2021**, *12*, 183–192. [[CrossRef](#)]
15. Wei, J.; Li, Z.; Pinker, R.T.; Wang, J.; Sun, L.; Xue, W.; Li, R.; Cribb, M. Himawari-8-Derived Diurnal Variations in Ground-Level PM_{2.5} Pollution across China Using the Fast Space-Time Light Gradient Boosting Machine (LightGBM). *Atmos. Chem. Phys.* **2021**, *21*, 7863–7880. [[CrossRef](#)]
16. Chen, J.; Yin, J.; Zang, L.; Zhang, T.; Zhao, M. Stacking Machine Learning Model for Estimating Hourly PM_{2.5} in China Based on Himawari 8 Aerosol Optical Depth Data. *Sci. Total Environ.* **2019**, *697*, 134021. [[CrossRef](#)]
17. Song, Z.; Fu, D.; Zhang, X.; Han, X.; Song, J.; Zhang, J.; Wang, J.; Xia, X. MODIS AOD Sampling Rate and Its Effect on PM_{2.5} Estimation in North China. *Atmos. Environ.* **2019**, *209*, 14–22. [[CrossRef](#)]
18. Shin, M.; Kang, Y.; Park, S.; Im, J.; Yoo, C.; Quackenbush, L.J. Estimating Ground-Level Particulate Matter Concentrations Using Satellite-Based Data: A Review. *GISci. Remote Sens.* **2020**, *57*, 174–189. [[CrossRef](#)]
19. Chen, Z.-Y.; Zhang, T.-H.; Zhang, R.; Zhu, Z.-M.; Yang, J.; Chen, P.-Y.; Ou, C.-Q.; Guo, Y. Extreme Gradient Boosting Model to Estimate PM_{2.5} Concentrations with Missing-Filled Satellite Data in China. *Atmos. Environ.* **2019**, *202*, 180–189. [[CrossRef](#)]
20. Jiang, T.; Chen, B.; Nie, Z.; Ren, Z.; Xu, B.; Tang, S. Estimation of Hourly Full-Coverage PM_{2.5} Concentrations at 1-Km Resolution in China Using a Two-Stage Random Forest Model. *Atmos. Res.* **2021**, *248*, 105146. [[CrossRef](#)]
21. Xiao, Q.; Geng, G.; Cheng, J.; Liang, F.; Li, R.; Meng, X.; Xue, T.; Huang, X.; Kan, H.; Zhang, Q.; et al. Evaluation of Gap-Filling Approaches in Satellite-Based Daily PM_{2.5} Prediction Models. *Atmos. Environ.* **2021**, *244*, 117921. [[CrossRef](#)]
22. Zhan, Y.; Luo, Y.; Deng, X.; Chen, H.; Grieneisen, M.L.; Shen, X.; Zhu, L.; Zhang, M. Spatiotemporal Prediction of Continuous Daily PM_{2.5} Concentrations across China Using a Spatially Explicit Machine Learning Algorithm. *Atmos. Environ.* **2017**, *155*, 129–139. [[CrossRef](#)]
23. Brokamp, C.; Jandarov, R.; Hossain, M.; Ryan, P. Predicting Daily Urban Fine Particulate Matter Concentrations Using a Random Forest Model. *Environ. Sci. Technol.* **2018**, *52*, 4173–4179. [[CrossRef](#)] [[PubMed](#)]
24. Li, T.; Zhang, C.; Shen, H.; Yuan, Q.; Zhang, L. Real-time and seamless monitoring of ground-level pm_{2.5} using satellite remote sensing. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2018**, *IV-3*, 143–147. [[CrossRef](#)]
25. Wu, J.; Li, T.; Zhang, C.; Cheng, Q.; Shen, H. Hourly PM_{2.5} Concentration Monitoring With Spatiotemporal Continuity by the Fusion of Satellite and Station Observations. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 8019–8032. [[CrossRef](#)]
26. Li, T.; Shen, H.; Yuan, Q.; Zhang, X.; Zhang, L. Estimating Ground-Level PM_{2.5} by Fusing Satellite and Station Observations: A Geo-Intelligent Deep Learning Approach: Deep Learning for PM_{2.5} Estimation. *Geophys. Res. Lett.* **2017**, *44*, 11985–11993. [[CrossRef](#)]
27. Wei, J.; Li, Z.; Cribb, M.; Huang, W.; Xue, W.; Sun, L.; Guo, J.; Peng, Y.; Li, J.; Lyapustin, A.; et al. Improved 1 Km Resolution PM_{2.5} Estimates across China Using Enhanced Space-Time Extremely Randomized Trees. *Atmos. Chem. Phys.* **2020**, *20*, 3273–3289. [[CrossRef](#)]
28. Wei, J.; Huang, W.; Li, Z.; Xue, W.; Peng, Y.; Sun, L.; Cribb, M. Estimating 1-Km-Resolution PM_{2.5} Concentrations across China Using the Space-Time Random Forest Approach. *Remote Sens. Environ.* **2019**, *231*, 111221. [[CrossRef](#)]
29. Li, H.; Yang, Y.; Wang, H.; Li, B.; Wang, P.; Li, J.; Liao, H. Constructing a Spatiotemporally Coherent Long-Term PM_{2.5} Concentration Dataset over China during 1980–2019 Using a Machine Learning Approach. *Sci. Total Environ.* **2021**, *765*, 144263. [[CrossRef](#)]
30. Zhang, J.; Fogelman-Soulie, F.; Largeton, C. Towards Automatic Complex Feature Engineering. In Proceedings of the International Conference on Web Information Systems Engineering, Dubai, United Arab Emirates, 12–15 November 2018; Springer: Berlin/Heidelberg, Germany, 2018; pp. 312–322.
31. Domingos, P. A Few Useful Things to Know about Machine Learning. *Commun. ACM* **2012**, *55*, 78–87. [[CrossRef](#)]
32. He, Q.; Gu, Y.; Zhang, M. Spatiotemporal Trends of PM_{2.5} Concentrations in Central China from 2003 to 2018 Based on MAIAC-Derived High-Resolution Data. *Environ. Int.* **2020**, *137*, 105536. [[CrossRef](#)]
33. He, Q.; Gao, K.; Zhang, L.; Song, Y.; Zhang, M. Satellite-Derived 1-Km Estimates and Long-Term Trends of PM_{2.5} Concentrations in China from 2000 to 2018. *Environ. Int.* **2021**, *156*, 106726. [[CrossRef](#)]
34. Ma, J.; Zhang, R.; Xu, J.; Yu, Z. MERRA-2 PM_{2.5} Mass Concentration Reconstruction in China Mainland Based on LightGBM Machine Learning. *Sci. Total Environ.* **2022**, *827*, 154363. [[CrossRef](#)]
35. Kong, L.; Tang, X.; Zhu, J.; Wang, Z.; Li, J.; Wu, H.; Wu, Q.; Chen, H.; Zhu, L.; Wang, W.; et al. A 6-Year-Long (2013–2018) High-Resolution Air Quality Reanalysis Dataset in China Based on the Assimilation of Surface Observations from CNEMC. *Earth Syst. Sci. Data* **2021**, *13*, 529–570. [[CrossRef](#)]
36. Zhao, Q.; Zhao, W.; Bi, J.; Ma, Z. Climatology and Calibration of MERRA-2 PM_{2.5} Components over China. *Atmos. Pollut. Res.* **2021**, *12*, 357–366. [[CrossRef](#)]
37. Ma, J.; Xu, J.; Qu, Y. Evaluation on the Surface PM_{2.5} Concentration over China Mainland from NASA's MERRA-2. *Atmos. Environ.* **2020**, *237*, 117666. [[CrossRef](#)]
38. Rigatti, S.J. Random Forest. *J. Insur. Med.* **2017**, *47*, 31–39. [[CrossRef](#)]

39. Wei, J.; Li, Z.; Lyapustin, A.; Sun, L.; Peng, Y.; Xue, W.; Su, T.; Cribb, M. Reconstructing 1-Km-Resolution High-Quality PM_{2.5} Data Records from 2000 to 2018 in China: Spatiotemporal Variations and Policy Implications. *Remote Sens. Environ.* **2021**, *252*, 112136. [[CrossRef](#)]
40. Zhan, Q.; Fan, Z.; Yan, S.; Yang, S.; Yang, C. New MAIAC AOD Product Based High Resolution PM_{2.5} Spatial-Temporal Distribution Change at Urban Scale—Case Study of Wuhan. In Proceedings of the 2019 10th International Workshop on the Analysis of Multitemporal Remote Sensing Images (MultiTemp), Shanghai, China, 5–7 August 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 1–4.
41. Gui, K.; Che, H.; Zeng, Z.; Wang, Y.; Zhai, S.; Wang, Z.; Luo, M.; Zhang, L.; Liao, T.; Zhao, H.; et al. Construction of a Virtual PM_{2.5} Observation Network in China Based on High-Density Surface Meteorological Observations Using the Extreme Gradient Boosting Model. *Environ. Int.* **2020**, *141*, 105801. [[CrossRef](#)]
42. Zhong, J.; Zhang, X.; Gui, K.; Wang, Y.; Che, H.; Shen, X.; Zhang, L.; Zhang, Y.; Sun, J.; Zhang, W. Robust Prediction of Hourly PM_{2.5} from Meteorological Data Using LightGBM. *Natl. Sci. Rev.* **2021**, *8*, nwaa307. [[CrossRef](#)]
43. Guryanov, A. Histogram-Based Algorithm for Building Gradient Boosting Ensembles of Piecewise Linear Decision Trees. In Proceedings of the International Conference on Analysis of Images, Social Networks and Texts, Kazan, Russia, 17–19 July 2019; Springer: Berlin/Heidelberg, Germany, 2019; pp. 39–50.
44. Prokhorenkova, L.; Gusev, G.; Vorobev, A.; Dorogush, A.V.; Gulin, A. CatBoost: Unbiased Boosting with Categorical Features. In Proceedings of the Advances in Neural Information Processing Systems, Montréal, QC, Canada, 3–8 December 2018; Volume 31, pp. 6638–6648.
45. Dorogush, A.V.; Ershov, V.; Gulin, A. CatBoost: Gradient Boosting with Categorical Features Support. *arXiv* **2018**, arXiv:1810.11363.
46. Horn, F.; Pack, R.; Rieger, M. The Autofeat Python Library for Automated Feature Engineering and Selection. In Proceedings of the Joint European Conference on Machine Learning and Knowledge Discovery in Databases, Würzburg, Germany, 16–20 September 2019; Springer: Berlin/Heidelberg, Germany, 2019; pp. 111–120.
47. Selvam, S.K.; Rajendran, C. Toffee-Tree: Automatic Feature Engineering Framework for Modeling Trend-Cycle in Time Series Forecasting. *Neural Comput. Appl.* **2021**, 1–20. [[CrossRef](#)]
48. Wang, M.; Ding, Z.; Pan, M. LbR: A New Regression Architecture for Automated Feature Engineering. In Proceedings of the 2020 International Conference on Data Mining Workshops (ICDMW), Sorrento, Italy, 17–20 November 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 432–439.
49. Shi, Q.; Zhang, Y.-L.; Li, L.; Yang, X.; Li, M.; Zhou, J. SAFE: Scalable Automatic Feature Engineering Framework for Industrial Tasks. In Proceedings of the 2020 IEEE 36th International Conference on Data Engineering (ICDE), Dallas, TX, USA, 20–24 April 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 1645–1656.
50. Khurana, U.; Samulowitz, H.; Turaga, D. Feature Engineering for Predictive Modeling Using Reinforcement Learning. In Proceedings of the AAAI Conference on Artificial Intelligence, New Orleans, LA, USA, 2–7 February 2018; Volume 32.
51. Zhang, Y.; Liang, S.; Zhu, Z.; Ma, H.; He, T. Soil Moisture Content Retrieval from Landsat 8 Data Using Ensemble Learning. *ISPRS J. Photogramm. Remote Sens.* **2022**, *185*, 32–47. [[CrossRef](#)]
52. Yadav, S.; Shukla, S. Analysis of K-Fold Cross-Validation over Hold-Out Validation on Colossal Datasets for Quality Classification. In Proceedings of the 2016 IEEE 6th International Conference on Advanced Computing (IACC), Bhimavaram, India, 27–28 February 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 78–83.
53. Rodriguez, J.D.; Perez, A.; Lozano, J.A. Sensitivity Analysis of K-Fold Cross Validation in Prediction Error Estimation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2010**, *32*, 569–575. [[CrossRef](#)]
54. Lundberg, S.M.; Lee, S.-I. A Unified Approach to Interpreting Model Predictions. In Proceedings of the Advances in Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017; Volume 30, pp. 4768–4777.
55. Altman, N.; Krzywinski, M. The Curse(s) of Dimensionality. *Nat. Methods* **2018**, *15*, 399–400. [[CrossRef](#)]
56. Dao, F.-Y.; Lv, H.; Wang, F.; Feng, C.-Q.; Ding, H.; Chen, W.; Lin, H. Identify Origin of Replication in *Saccharomyces Cerevisiae* Using Two-Step Feature Selection Technique. *Bioinformatics* **2019**, *35*, 2075–2083. [[CrossRef](#)]
57. Geng, G.; Xiao, Q.; Liu, S.; Liu, X.; Cheng, J.; Zheng, Y.; Xue, T.; Tong, D.; Zheng, B.; Peng, Y.; et al. Tracking Air Pollution in China: Near Real-Time PM_{2.5} Retrievals from Multisource Data Fusion. *Environ. Sci. Technol.* **2021**, *55*, 12106–12115. [[CrossRef](#)]
58. Xu, S.; Zhang, Z.; Du, X.; Li, Y.; Zhang, S.; Xu, P.; Zhang, B.; Meng, F. Impact of Residential Coal Combustion Control in Beijing-Tianjin-Hebei and Surrounding Region on PM_{2.5} in Beijing. *Res. Environ. Sci.* **2021**, *34*, 2876–2886. [[CrossRef](#)]
59. Zhang, L.; An, J.; Liu, M.; Li, Z.; Liu, Y.; Tao, L.; Liu, X.; Zhang, F.; Zheng, D.; Gao, Q.; et al. Spatiotemporal Variations and Influencing Factors of PM_{2.5} Concentrations in Beijing, China. *Environ. Pollut.* **2020**, *262*, 114276. [[CrossRef](#)]
60. Zhao, H.; Zheng, Y.; Li, C. Spatiotemporal Distribution of PM_{2.5} and O₃ and Their Interaction During the Summer and Winter Seasons in Beijing, China. *Sustainability* **2018**, *10*, 4519. [[CrossRef](#)]
61. Manning, M.I.; Martin, R.V.; Hasenkopf, C.; Flasher, J.; Li, C. Diurnal Patterns in Global Fine Particulate Matter Concentration. *Environ. Sci. Technol. Lett.* **2018**, *5*, 687–691. [[CrossRef](#)]
62. Wang, L.; Xiong, Q.; Wu, G.; Gautam, A.; Jiang, J.; Liu, S.; Zhao, W.; Guan, H. Spatio-Temporal Variation Characteristics of PM_{2.5} in the Beijing-Tianjin-Hebei Region, China, from 2013 to 2018. *Int. J. Environ. Res. Public Health* **2019**, *16*, 4276. [[CrossRef](#)]
63. Ding, Y.; Chen, Z.; Lu, W.; Wang, X. A CatBoost Approach with Wavelet Decomposition to Improve Satellite-Derived High-Resolution PM_{2.5} Estimates in Beijing-Tianjin-Hebei. *Atmos. Environ.* **2021**, *249*, 118212. [[CrossRef](#)]

64. Zheng, G.J.; Duan, F.K.; Su, H.; Ma, Y.L.; Cheng, Y.; Zheng, B.; Zhang, Q.; Huang, T.; Kimoto, T.; Chang, D.; et al. Exploring the Severe Winter Haze in Beijing: The Impact of Synoptic Weather, Regional Transport and Heterogeneous Reactions. *Atmos. Chem. Phys.* **2015**, *15*, 2969–2983. [[CrossRef](#)]
65. Blagus, R.; Lusa, L. SMOTE for High-Dimensional Class-Imbalanced Data. *BMC Bioinform.* **2013**, *14*, 106. [[CrossRef](#)]
66. Yu, Z.; Qu, Y.; Wang, Y.; Ma, J.; Cao, Y. Application of Machine-Learning-Based Fusion Model in Visibility Forecast: A Case Study of Shanghai, China. *Remote Sens.* **2021**, *13*, 2096. [[CrossRef](#)]
67. Vu, B.N.; Bi, J.; Wang, W.; Huff, A.; Kondragunta, S.; Liu, Y. Application of Geostationary Satellite and High-Resolution Meteorology Data in Estimating Hourly PM_{2.5} Levels during the Camp Fire Episode in California. *Remote Sens. Environ.* **2022**, *271*, 112890. [[CrossRef](#)]
68. Hu, H.; Hu, Z.; Zhong, K.; Xu, J.; Zhang, F.; Zhao, Y.; Wu, P. Satellite-Based High-Resolution Mapping of Ground-Level PM_{2.5} Concentrations over East China Using a Spatiotemporal Regression Kriging Model. *Sci. Total Environ.* **2019**, *672*, 479–490. [[CrossRef](#)]