Journal of Environmental Management 342 (2023) 118145

Contents lists available at ScienceDirect

Journal of Environmental Management

journal homepage: www.elsevier.com/locate/jenvman

Research article

Spatiotemporally continuous estimates of daily 1-km $PM_{2.5}$ concentrations and their long-term exposure in China from 2000 to 2020

Qingqing He [a,*], Tong Ye [a], Weihang Wang [a], Ming Luo [b], Yimeng Song [c], Ming Zhang [a]

[a] School of Resource and Environmental Engineering, Wuhan University of Technology, Wuhan, 430070, China
[b] School of Geography and Planning, Sun Yat-sen University, Guangzhou, 510006, China
[c] School of the Environment, Yale University, New Haven, CT, 06511, USA

ARTICLE INFO

Keywords:
Fine particulate matter ($PM_{2.5}$)
Aerosol optical depth (AOD)
Ensemble learning
Full coverage
Long-term variations

ABSTRACT

Monitoring long-term variations in fine particulate matter ($PM_{2.5}$) is essential for environmental management and epidemiological studies. While satellite-based statistical/machine-learning methods can be used for estimating high-resolution ground-level $PM_{2.5}$ concentration data, their applications have been hindered by limited accuracy in daily estimates during years without $PM_{2.5}$ measurements and massive missing values due to satellite retrieval data. To address these issues, we developed a new spatiotemporal high-resolution $PM_{2.5}$ hindcast modeling framework to generate the full-coverage, daily, 1-km $PM_{2.5}$ data for China for the period 2000–2020 with improved accuracy. Our modeling framework incorporated information on changes in observation variables between periods with and without monitoring data and filled gaps in $PM_{2.5}$ estimates induced by satellite data using imputed high-resolution aerosol data. Compared to previous hindcast studies, our method achieved superior overall cross-validation (CV) R^2 and root-mean-square error (RMSE) of 0.90 and 12.94 μg/m^3 and significantly improved the model performance in years without $PM_{2.5}$ measurements, raising the leave-one-year-out CV R^2 [RMSE] to 0.83 [12.10 μg/m^3] at a monthly scale (0.65 [23.29 μg/m^3] at a daily scale). Our long-term $PM_{2.5}$ estimates show a sharp decline in $PM_{2.5}$ exposure in recent years, but the national exposure level in 2020 still exceeded the first annual interim target of the 2021 World Health Organization air quality guidelines. The proposed hindcast framework represents a new strategy to improve air quality hindcast modeling and can be applied to other regions with limited air quality monitoring periods. These high-quality estimates can support both long- and short-term scientific research and environmental management of $PM_{2.5}$ in China.

Author statement

Q. He: Conceptualization; Funding acquisition; Methodology; Supervision; Validation; Writing – original draft; Writing – review & editing. T. Ye: Data curation; Software; Visualization; Writing – review & editing. W. Wang: Visualization; Writing – review & editing. M. Luo: Data collocation; Writing – review & editing. Y. Song: Writing – review & editing. M. Zhang: Writing – review & editing.

1. Introduction

$PM_{2.5}$, as fine inhalable particulate matter with a diameter generally 2.5 μm and smaller, has been widely acknowledged as a leading health risk factor (Pope and Dockery, 2006). Both short-term and long-term exposures to ambient $PM_{2.5}$ are associated with adverse health impacts, including increased cardiovascular (Cao et al., 2012; Dominici et al., 2006; Liu et al., 2022) and respiratory morbidities (Habre et al., 2014; Peng et al., 2009), and even premature mortality (Burnett et al., 2018). With China's increasing industrialization and economic development, $PM_{2.5}$-related atmospheric pollution has substantially changed over the past decades (van Donkelaar et al., 2021). However, as the national air quality monitoring network was not established until the end of 2012, ground $PM_{2.5}$ measurements for China have been limited from 2013 to present. Moreover, even in the post-2013 period, the monitoring network consists of only a limited number of monitoring stations sparsely distributed in urban cities, leaving a considerable portion of China unmonitored. Therefore, there is still a need to develop methods to estimate an accurate nationwide high-resolution $PM_{2.5}$ dataset that covers the long-term historical periods, including those years without $PM_{2.5}$ observations.

* Corresponding author.
E-mail address: qqhe@whut.edu.cn (Q. He).

https://doi.org/10.1016/j.jenvman.2023.118145
Received 19 December 2022; Received in revised form 17 April 2023; Accepted 9 May 2023
Available online 19 May 2023

The derivation of high-resolution particulate concentrations from satellite data has garnered increasing attention (He et al., 2021a; Lin et al., 2016; Ma et al., 2022; van Donkelaar et al., 2021). Most relevant studies have adopted statistical/machine learning methods to derive ground $PM_{2.5}$ concentration data for pre-2013 periods for China from satellite-retrieved aerosol optical depth (AOD) (Bai et al., 2022; Geng et al., 2021; Huang et al., 2018; Liang et al., 2018; Ma et al., 2016; Meng et al., 2021; Wei et al., 2021; Xiao et al., 2017; Xue et al., 2019). Among those, a few of studies have provided daily high-resolution $PM_{2.5}$ estimates that cover the entire country and span a long period (Bai et al., 2022; Liang et al., 2020; Wei et al., 2021). Modeling long-term $PM_{2.5}$ over China, including the pre-2013 period without $PM_{2.5}$ observations, is much more complicated than what most of the literature focus on, which is to predict $PM_{2.5}$ based on statistical/machine-learning methods developed for periods with $PM_{2.5}$ observations (i.e., post-2013). Owing to the absence of pre-2013 measured $PM_{2.5}$ records, the relationship between $PM_{2.5}$ and its predictors cannot be directly determined for those years without $PM_{2.5}$ observations. To successfully generate $PM_{2.5}$ estimates prior to 2013, previous studies have assumed that the $PM_{2.5}$–predictors relationship in years without $PM_{2.5}$ observations is concordant with those developed for the observed period and then directly applied the post-2013 model to hindcast pre-2013 $PM_{2.5}$ (Ma et al., 2022). This assumption may not comply with the real-world scenario since $PM_{2.5}$ emission sources change over time due to changes in industrial activity, transportation patterns, energy consumption, and other factors, resulting in extant statistical/machine-learning models developed for the long-term $PM_{2.5}$ hindcasting usually exhibit larger modeling errors in years without $PM_{2.5}$ observations, particularly at daily scale with validated R^2 values of $\leq$0.55 (Meng et al., 2021; Xiao et al., 2017; Xue et al., 2019), compared to those for periods with $PM_{2.5}$ observations (Chen et al. 2022a, 2022b; Hough et al., 2021; Jiang et al., 2021; Song et al., 2022; Stafoggia et al., 2019). Therefore, it is necessary to design an advanced modeling strategy that can construct the $PM_{2.5}$–predictors relationship specific to the period without $PM_{2.5}$ observations to improve the accuracy of daily estimates for the long-term period.

Satellite-based $PM_{2.5}$ estimation methods are challenged by non-random missingness in satellite aerosol retrievals. Aerosol retrieval algorithms usually fail to derive AOD values on cloudy days, over highly reflective surfaces such as snow and desert areas, and during heavy particulate pollution episodes (Levy et al., 2013; Lyapustin et al., 2018), which resulted in discontinuous downstream estimates of $PM_{2.5}$ in space and time. As a result, most previous $PM_{2.5}$ datasets derived from satellite AOD data contain missing values (Fang et al., 2016; Li et al., 2017b; Wei et al., 2021; Yang et al., 2022), which can hamper the ability of the model to explore the relationship between $PM_{2.5}$ and AOD (Meng et al., 2021) and may further cause bias in research on $PM_{2.5}$ exposure and $PM_{2.5}$-related health effects (He et al., 2021c; Lv et al., 2017). To the date, a few studies have filled daily gaps in long-term and large-scale $PM_{2.5}$ estimates (Bai et al., 2022; Geng et al., 2021; Xue et al., 2019) but the impact of AOD on $PM_{2.5}$ modeling during years without $PM_{2.5}$ measurements has not been thoroughly examined.

In this work, we developed a spatiotemporal high-resolution hindcast modeling method that utilizes a machine learning algorithm to generate full-coverage $PM_{2.5}$ estimates in China at a daily scale in a long period of 2000–2020, at a spatial resolution of 0.01°. The proposed hindcast modeling strategy is designed to construct $PM_{2.5}$-predictors relationships that vary each year during the pre-2013 period, by incorporating information on the similarity of the predictors between periods with and without $PM_{2.5}$ measurements. In addition, the $PM_{2.5}$ gaps induced by satellite aerosol retrievals were addressed by full-coverage AOD imputations estimated from the high-resolution Multi-Angle Implementation of Atmospheric Correction (MAIAC) AOD product, and the impact of full-coverage AOD on $PM_{2.5}$ estimation in the period without $PM_{2.5}$ measurements was explored. Finally, based on the long-term, full-coverage $PM_{2.5}$ estimates, we examined the spatiotemporal changes in $PM_{2.5}$ exposure in China and discussed its possible drivers for further air quality management.

2. Data and methods

2.1. Data

2.1.1. Ground-level $PM_{2.5}$ measurements

Ground-level hourly $PM_{2.5}$ measurements of the national air quality monitoring network for mainland China (i.e., excluding Hong Kong, Taiwan, and Macau) from 2013 to 2020 were collected from the China National Environmental Monitoring Center (http://www.cnemc.cn/). The national air quality monitoring network was established in 2013, and since then the number of in situ monitors has increased; there were ~1700 monitoring sites from 2013 to 2020 (Fig. 1). The daily mean $PM_{2.5}$ concentrations were computed from at least 18 valid hourly measurements per day for each in situ monitoring station. In addition, we collected daily $PM_{2.5}$ measurements from the local environmental protection authorities in Hong Kong (18 sites in 2005–2020), Macau (6 sites in 2012–2020), and Taiwan (84 sites in 2013–2020). In situ hourly data were also acquired from five U.S. embassies: those in Beijing (2008–2017), Guangzhou (2011–2017), Shanghai (2011–2017), Chengdu (2012–2017), and Shenyang (2013–2017). Due to the lack of pre-2013 ground $PM_{2.5}$ measurements for China's expanse, only in situ measurements for 2013–2020 were used for model development, and the small number of pre-2013 measurements were employed only for validating the $PM_{2.5}$ estimates.

2.1.2. MAIAC AOD and gap filling

As high-resolution satellite AOD data have been identified as a crucial predictor for resolving local variabilities in ground $PM_{2.5}$ estimation (Chudnovsky et al., 2013; He et al., 2021c), daily 1-km MAIAC AOD data (Lyapustin et al., 2018) were used in this study. In previous validation studies (Liu et al., 2019; Martins et al., 2017), MAIAC AOD retrievals achieved an accuracy comparable to those of widely used aerosol products obtained using the Dark Target method and Deep Blue algorithms. High-resolution AOD data with quality assurance flags (MCD19A2) were collected from NASA Earth Data (https://earthdata.nasa.gov/). We calculated the spatial coverage of daily MAIAC AOD from Terra and Aqua satellites for the entire study period and each season (Fig. S1, Supplementary Information). The overall missing rate of MAIAC AOD observations is 64.61%. For each grid cell, 22.44%–100% of days over the past 21 years are missing.

To fill the gaps in $PM_{2.5}$ estimates induced by MAIAC retrievals, we employed daily full-coverage aerosol data with a 1-km spatial resolution (hereafter referred to as MAIAC-imputed AOD) for 2000 to 2020, which were estimated by random forest-based imputation models considering spatial and temporal terms with MAIAC AOD used as the target variable. The detailed imputation modeling process was published in He et al. (2023). The validation results (Fig. S2) demonstrate that on a daily scale, our daily 1-km AOD imputations agree well with the original MAIAC retrievals (correlation coefficient $r = 0.988$ on average in the hold-out validation) and ground-level measurements (overall $R^2 = 0.75$, and root mean square error [RMSE] = 0.24). The spatial distribution of the multi-year mean MAIAC-imputed AOD values is shown in Fig. S3.

2.1.3. Other predictors

Auxiliary variables that may affect the accuracy of ground $PM_{2.5}$ concentration estimations from satellite data, particularly in cases in which nationwide historical $PM_{2.5}$ levels are undocumented and thus have to be inferred, were also used. In this study, meteorological parameters, $PM_{2.5}$ simulations, geographical condition data, and population distribution data were used in modeling and estimating concentrations of ground $PM_{2.5}$ in China.

Meteorological conditions considerably affect $PM_{2.5}$ variation (Guo et al., 2017; He et al., 2021b). We acquired weather station data,

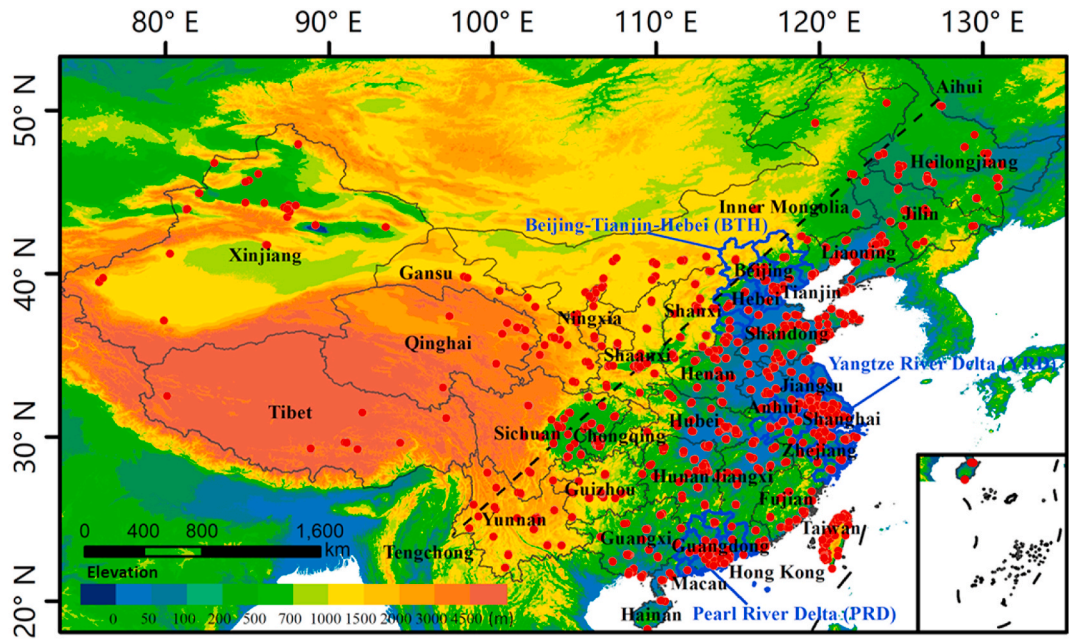

Fig. 1. Spatial map of ground-level $PM_{2.5}$ monitoring sites (shown as red dots), terrain, and three key megacity regions. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

including daily relative humidity (RH, %), wind speed (WS, m/s), and air temperature (T, °C), from the China Meteorological Data Sharing Service System (http://data.cma.cn/). We also collected 3-h planetary boundary layer height (PBLH, m) data from the European Centre for Medium-Range Weather Forecasts reanalysis dataset (http://www.ecmwf.int/), with a spatial resolution of 0.25° × 0.25°.

Major $PM_{2.5}$ constituents, i.e., sulfate, black carbon, organic carbon, dust, and sea salt, were collected from the MERRA-2 (Modern-Era Retrospective analysis for Research and Applications, Version 2) global reanalysis dataset (Randles et al., 2017), which has complete coverage at a spatial resolution of 0.5° × 0.625°. The MERRA-2 simulated $PM_{2.5}$ data have relatively high accuracy (Buchard et al., 2017) and provide additional daily $PM_{2.5}$ gradients on a coarse scale. The MERRA-2 data have been increasingly applied in aerosol- and particle-related studies (Liang et al., 2020; Meng et al., 2021). We calculated the daily $PM_{2.5}$ levels by weighting the five simulated species as Provençal et al. (2017).

Surface condition data, including normalized difference vegetation index (NDVI) at a 1-km spatial resolution and elevation data at a 30-m resolution, were respectively obtained from the MODIS Vegetation Indices Monthly L3 Global 1 km Product (https://search.earthdata.nasa.gov) and the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) global digital elevation model (DEM). We also obtained annual demographic distribution data (POP) at a 1-km spatial resolution from the LandScan global population database (Rose et al., 2021).

2.1.4. Data processing and integration

To produce high-resolution $PM_{2.5}$, a 1-km grid corresponding to the grid size of the MAIAC-imputed AOD data was created for the study area. The $PM_{2.5}$ monitoring sites were spatially matched to the 1-km grid cell according to site geolocation, and those monitoring sites falling in the same 1-km grid cell were merged. We aggregated predictor data with different spatial resolutions by averaging values within each grid cell (e.g., DEM) or through inverse distance weighted spatial interpolation (e.g., station-based meteorological parameters) and integration into 1-km grid cells. The data preprocessing and integration methods are detailed in Table S1.

2.2. Methodology

Basically, the commonly developed model for elucidating the relationship between $PM_{2.5}$ and predictors can be expressed as follows:

$$PM_{2.5} = f(obs\_sim\ predictors) \quad (1)$$

Where f denotes the nonlinear relationship between the response variable and explanatory variables learned by the machine-learning model; *obs_sim predictors* are the variables with observation or simulation data (i.e., the nine variables that are detailed in Section 2.1). This relationship can only be constructed for the period for which both $PM_{2.5}$ measurements and *obs_sim predictors* have valid data. For pre-2013 years without $PM_{2.5}$ monitoring data, the post-2013 relationship was used to make $PM_{2.5}$ predictions before 2013. However, although long-term $PM_{2.5}$ monitoring data is unavailable, the long-term observation and simulation data are available for *obs_sim predictors*, which could provide information on changes of each predictor in "historical" years (i.e., years without $PM_{2.5}$ measurements) relative to modeling years (i.e., years with $PM_{2.5}$ measurements). Previous studies have utilized similarity analysis (e.g., Euclidean Distance) to measure the changes in meteorological components between different years (Chen et al., 2021). Here, to facilitate the explanation of $PM_{2.5}$-predictors relationships in years without $PM_{2.5}$ measurements, we proposed a new $PM_{2.5}$ hindcast modeling framework that incorporates such change (or similarity) information, referred to as "historical similarity" in this study, into the hindcast modeling process for the first time. In addition, the spatial and temporal features were also integrated, as some studies found that including spatial and/or temporal autocorrelation information can enhance the predictive power of $PM_{2.5}$ (He et al., 2021a; Wei et al., 2021; Yang et al., 2022). Thus, the general model structure can be restated as follows:

$$PM_{2.5} = f(obs\_sim\ predictors, h, s, t) \quad (2)$$

where h, s, and t represent the additional dummy features: historical similarity features, spatial features, and temporal features, respectively, which are detailed in Sections 2.2.1–2.2.3.

2.2.1. Extraction of historical similarity features

This study conducted similarity analysis and used historical simi-

larity values as auxiliary features in the model development. The historical similarity of a predictor for a given historical year was determined by calculating the Manhattan distance between the data of the predictive historical year and the data of each model-training year (Eq. (3)).

$$h_m(a,b) = \sum_i \left| Var_{a,i} - Var_{b,i} \right| \quad (3)$$

where $h_m(a,b)$ represents the historical similarity of predictor m between the historical year a and the modeling-training year b, and Var_i represents the annual mean value of predictor m at grid cell i. To better represent the change in conditions between years, data from the full-coverage distribution of predictor m was used for Var other than that from the sample dataset. Other similarity methods, such as Pearson correlation analysis and Euclidean distance, were also attempted, but a sensitivity analysis showed that including historical similarity features based on Manhattan distance in the pre-2013 modeling resulted in better model performance. In this study, we calculated the similarities between each predictive year in the historical period of 2000–2012 and each model-training year in 2013–2020 for predictors with yearly-varying data.

The historical similarity values contain information about the variations in *obs_sim* predictors between the pre-2013 and post-2013 periods. By incorporating these between-year similarity values as features in the modeling, the $PM_{2.5}$–predictor relationship established for the modeling years can be correspondingly adjusted for the given predictive year. In other words, the developed $PM_{2.5}$-predictors relationship is tailored to each specific predictive year. Thus, our long-term $PM_{2.5}$ modeling framework is substantially different from previous prediction strategies that assume the relationships established for the model-training years remain constant throughout the predictive historical period.

2.2.2. Extraction of spatial features

Previous studies have demonstrated that $PM_{2.5}$ levels exhibit strong spatial patterns, and incorporating spatial information into the modeling can help resolve the spatial variability in $PM_{2.5}$ (He et al., 2021b; Wei et al., 2021; Yang et al., 2022; Zhang et al., 2018). Thus, we integrated spatial features in the modeling, including the widely-used coordinates (i.e., longitude and latitude), as well as additional spatial features by following the approach used in Wei et al. (2021), i.e., Haversine distances between the predictive grid cell and the four corners or the center of the circumscribed rectangle of the study area, to reflect the spatial pattern in the $PM_{2.5}$-predictors relationship. In the preliminary testing, we also attempted to use a batch of spatial convolutional layers, similar to those used by Hu et al. (2017). Sensitivity tests showed that the two methods of characterizing spatial information (i.e., Haversine distance and spatial convolutional layer) similarly contribute to the model performance, but the model structure based on Haversine distances was refined, taking less computational cost, which was finally applied in this study.

2.2.3. Extraction of temporal features

Like previous methods (He et al., 2021a; Yang et al., 2022), this study also includes temporal features to consider the seasonal variation in $PM_{2.5}$ and make a smooth connection between the last day of the current year and the first day of the following year. In addition to the Julian day of the year (DOY), its cosine form was also included as a temporal feature in the hindcast modeling. Additionally, similar to the calculation method of Haversine spatial features used in Section 2.2.2, we encode the position of one predictive grid cell in time by measuring the time circular distances according to the first, middle, and last days of the predictive year (Eq. (3)) to minimize the impact of seasonality.

$$t'(DOY) = \cos\left(2\pi \frac{DOY - t_0}{T}\right) \quad (4)$$

where t_0 represents the first, middle, or last days of the predictive year, and T = 365 for regular years, or 366 for leap years.

2.2.4. Model development and implementation

The random forest (RF) algorithm (Breiman, 2001) was used to construct the models, as it can effectively handle a large number of features and allows for the construction of flexible relationships between response and explanatory variables; these properties are essential for producing large-scale data with a high spatiotemporal resolution. The algorithm has gained tremendous popularity in recent air pollution-related studies (Bi et al., 2019; Brokamp et al., 2018; Hu et al., 2017; Jiang et al., 2021). RF is an advanced ensemble learning approach that operates by building multiple decision trees through bootstrap sampling with replacement. A subset of variables is randomly selected for each decision split in the learning process to fit each tree. For a regression application, predictions from all decision trees are averaged to obtain the model output. The number of decision trees grown (N_{tree}) and the number of variables for each decision split (N_{var}) are two core hyperparameters that were tuned via cross-validation and Bayesian optimization in this study. In addition, the variable importance, estimated by permutation of out-of-bag predictor observations for the regression-based random forest, was used to examine the relative importance of the variables to $PM_{2.5}$ estimation.

As only 2013–2020 nationwide observational data of $PM_{2.5}$ are available, we separately developed RF models for two periods to achieve high accuracy for the entire study period of 2000–2020: one directly uses the calendar year as an indicator to yield estimates for the period with nationwide $PM_{2.5}$ measurements, and the other incorporates historical similarities as features to hindcast pre-2013 concentrations of ground $PM_{2.5}$ across China, with both high spatial (i.e., 1 km) and temporal (daily) resolutions. A single model combining the two models can be considered a binary tree model, in which the dichotomous predictor is defined by the availability/non-availability of $PM_{2.5}$ measurements for the prediction year.

We used the collocated samples of 2013–2020 to develop the models for the two periods. The framework of the proposed hindcast modeling is presented in Fig. 2. For post-2013 years, we used the year identifier to estimate daily $PM_{2.5}$. For pre-2013 years, eight historical similarity features between the predictive year and modeling years were separately calculated for AOD, MERRA-2 $PM_{2.5}$, RH, T, WS, PBLH, NDVI, and POP, and integrated into the modeling as a surrogate for year identifiers. The other features for models built for both periods included the MAIAC-imputed AOD, meteorological fields (i.e., RH, T, WS, and PBLH), MERRA-2 $PM_{2.5}$ simulations, surface condition variables (i.e., NDVI, DEM, and POP), spatial locations (x–y coordinates of grid cells, Haversine distances between grid cells and upper-left/lower-left/upper-right/lower-right/central points of the study region), and temporal positions (dummy variables for the Julian day, cosine day, and cosine distances between the estimation day and the first/middle/last day of a year).

Through the proposed method combined with multisource data from ground-level measurements, satellite remote sensing, and numerical simulations, we predicted the long-term $PM_{2.5}$ data with full coverage and high spatiotemporal resolution (daily, 1 km) for China. However, similar to previous studies (Ma et al., 2022), including spatial features (e.g., longitude and Haversine distances in Section 2.2.2) in the modeling could lead to abnormal spatial variations in daily predicted maps in western China (Fig. S4a), due to air quality monitoring stations are sparsely distributed over this region. However, completely excluding spatial features from the modeling would degrade overall model performance (Table S2). Thus, in the final prediction process, we combined the outputs from the model with full features over eastern China and those without spatial features in Tibet, Qinghai, and Xinjiang provinces. In other words, we excluded spatial features from the models in the final prediction process to predict $PM_{2.5}$ concentrations in western China (i.e., Tibet, Qinghai, Xinjiang, and western Sichuan provinces) since we could

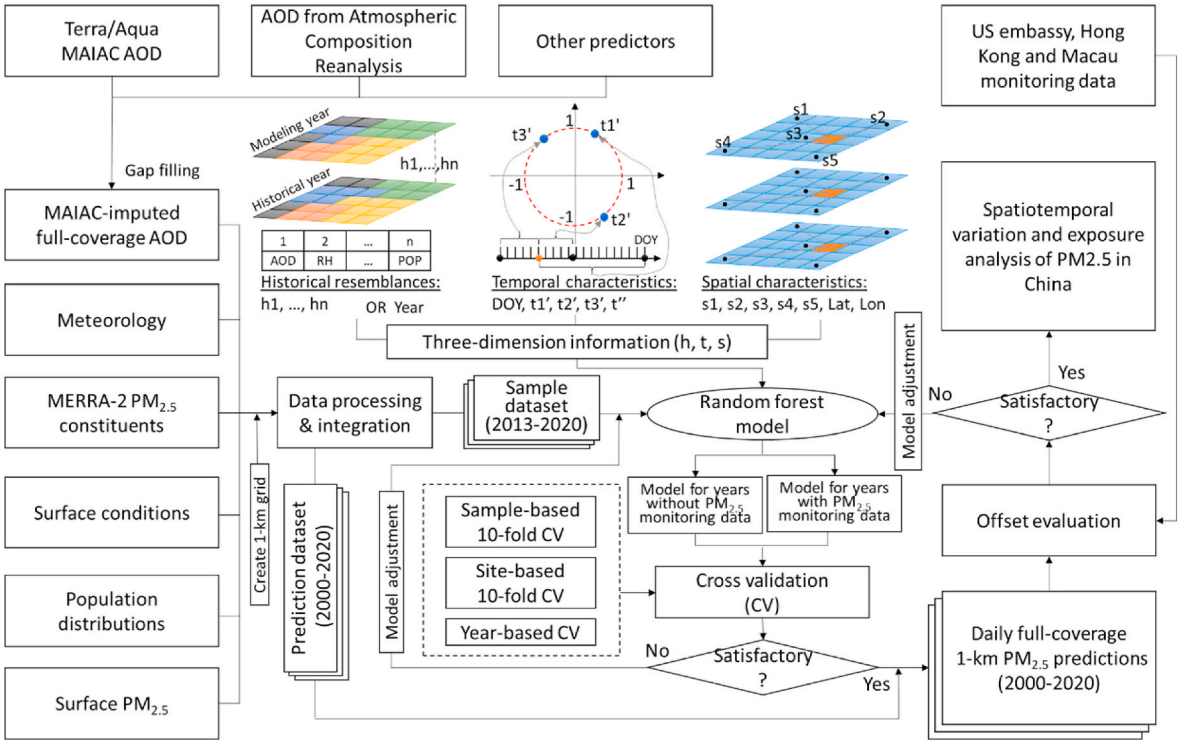

Fig. 2. Flowchart of long-term $PM_{2.5}$ prediction modeling with additional historical similarity, spatial, and temporal features.

not obtain smooth spatial surfaces of $PM_{2.5}$ predictions for those areas when including the spatial terms. The final daily 1-km full-coverage $PM_{2.5}$ dataset generated by this study is freely available via http://doi.org/10.5281/zenodo.4569557.

2.2.5. Model validation

For the post-2013 period, we evaluated the overall performance of the model using the popular sample-based 10-fold cross-validation (CV) approach. The sample dataset was randomly organized into 10 groups, of which one group (10% of samples) was reserved for testing, and the remaining nine groups (90% of samples) were used for training. The training process was conducted 10 times until all of the groups were tested. Furthermore, to explore the spatial predictive power of the model in unmonitored areas, a station-based 10-fold CV was performed, in which 10% of stations were randomly reserved for the validation procedure at each stage. The model performance in terms of the agreement between the predictions and observations was evaluated using statistical metrics, including the linearly regressed R^2, RMSE, and mean absolute error (MAE).

Since pre-2013 nationwide $PM_{2.5}$ observations were unavailable, we could not directly assess our model's performance across China before 2013. Instead, we conducted a leave-one-year-out CV to examine the reliability and fidelity of the $PM_{2.5}$ estimates in years without $PM_{2.5}$ measurements (2000–2012). Many previous studies (Wei et al., 2021; Xiao et al., 2017) adopted a simpler validation approach to indicate the accuracy of pre-2013 estimates, i.e., using one-year samples for modeling and another-year samples for testing. Here, we employed a stricter leave-one-year-out CV method, where one-year samples were reserved for model evaluation and seven-year samples were used for model development, and the validation process was conducted eight times until all of the samples in the modeling years were validated.

To further evaluate the pre-2013 model's hindcast ability, we conducted individual validation using $PM_{2.5}$ measurements from several in situ stations (i.e., U.S. embassy sites in mainland China and sites in Hong Kong and Macau). Although the limited number of pre-2013 measurements provided by these monitoring sites could not fully represent the nationwide $PM_{2.5}$ levels, this validation is the only direct method for testing the model's performance in the 2000–2012 period.

3. Results

3.1. Model performance

3.1.1. Modeling accuracy for 2013–2020

Based on the MAIAC-imputed AOD data, the final sample dataset used for estimating ground $PM_{2.5}$ contained 3,680,131 samples, three times the samples in the dataset based on the original MAIAC AOD retrievals (1,121,177). The sample data shows that the nationwide multi-year mean $PM_{2.5}$ concentration is 44.18 μg/m^3, with a standard deviation (std) of 39.40 μg/m^3, and the overall average of MAIAC-imputed AOD is 0.56 (std = 0.39).

Fig. 3 shows the density plots of the overall (sample-based) and spatial (station-based) 10-fold CV results over China, including statistical metrics of the agreement between the estimated and measured $PM_{2.5}$ concentrations: R^2, RMSE, and MAE. In general, the daily $PM_{2.5}$ estimates show good concordance with ground-level observations over China, with a high R^2 value of 0.89 and low RMSE (13.28 μg/m^3) and MAE (7.94 μg/m^3) in Fig. 3 (a). The monthly and yearly averages of sample-based CV results (Fig. 3b and c) indicate that the CV R^2 values increase by ~0.04 and ~0.05 to 0.93 and 0.94, respectively. Fig. 3d shows that the spatial CV R^2 and RMSE also reach relatively good values of 0.87 and 14.27 μg/m^3, respectively. The spatial CV presents only a slight decrease in model performance compared with the overall CV results, indicating that the developed model well captures the spatial variability in data and interprets the uncertainty in unobserved areas.

Table S3 summarizes the sample-based CV results of all estimates for each year from 2013 to 2020. Our model performs well and stably for each modeling year, with R^2 (RMSE) ranging between 0.86 (19.96 μg/m^3) and 0.90 (9.85 μg/m^3). The significant decrease in RMSE between 2013 and 2020 (10.11 μg/m^3) is mainly due to the considerable reduction in China's ambient $PM_{2.5}$ levels due to the air quality control policies implemented in recent years (Ma et al., 2016; Wei et al., 2021).

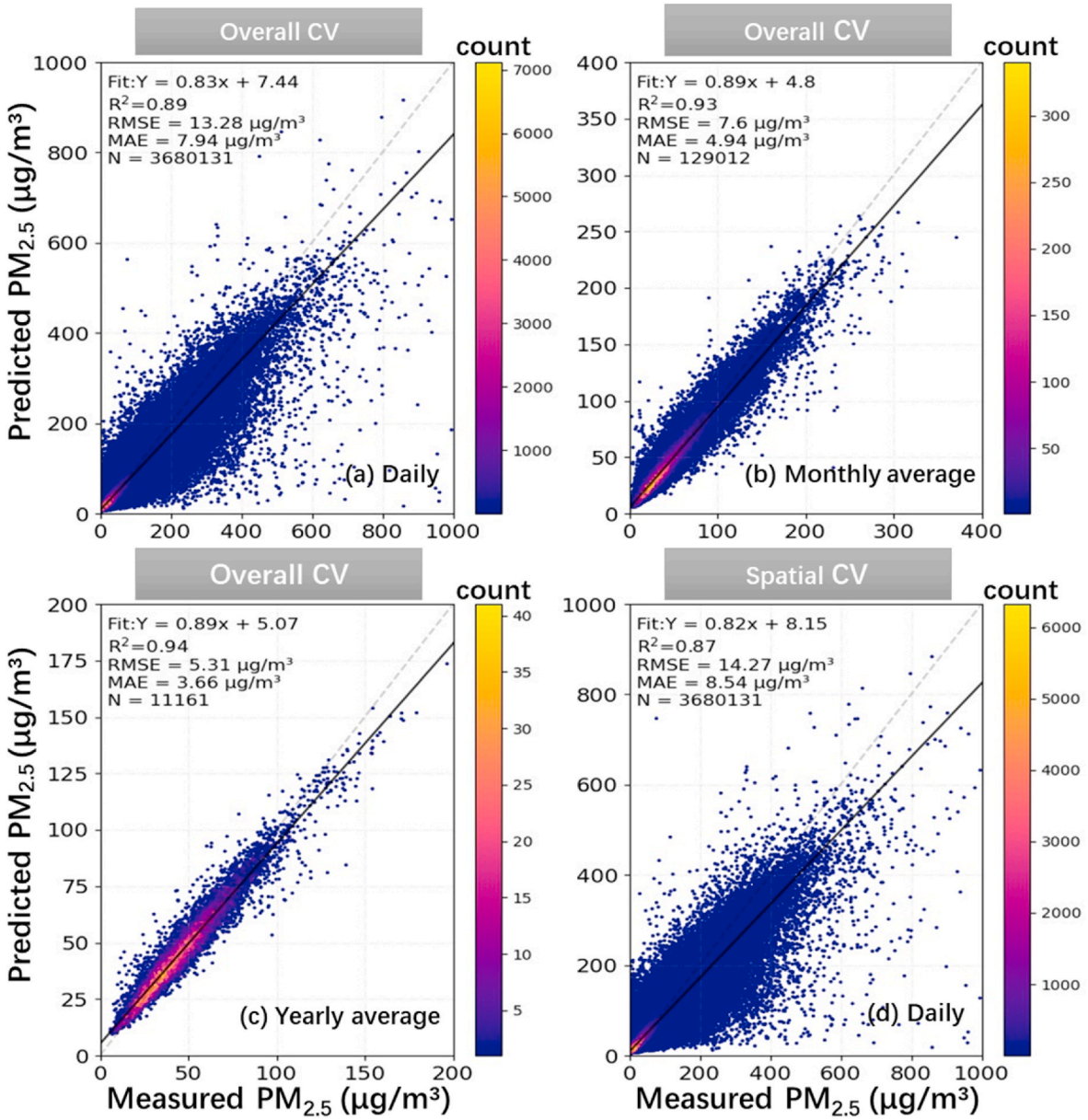

Fig. 3. Scatterplots of (a–c) overall and (d) spatial 10-fold CV between observed and estimated $PM_{2.5}$ concentrations. The black lines are the linear regression lines between $PM_{2.5}$ observations and estimates. The blue dotted lines are the one-to-one lines. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

The slopes (0.80–0.85) of the linear fitting lines of observed vs. estimated points are close to those of the one-to-one lines, and the intercepts (5.69–10.15) are small.

Figure S5 presents the spatiotemporal error statistics of $PM_{2.5}$ estimates based on station-based 10-fold CV results. Minor estimation errors exist for each monitoring site (i.e., the average CV R^2 of the sites is 0.83, and ~74% of sites show an R^2 of >0.80) and throughout the eight modeling years (i.e., the average CV R^2 of modeling days is 0.77, and ~65% of days show an R^2 of >0.75). Those above-mentioned validation results substantiate that our model can effectively resolve daily $PM_{2.5}$ variations at a site scale for the period with $PM_{2.5}$ monitoring data.

3.1.2. *Hindcasting performance in 2000–2012*

Fig. 4a–c presents the predictive power outside the model-training years. In the leave-one-year-out CV, the daily $PM_{2.5}$ estimates agree relatively well with the corresponding ground-level observations, with R^2, RMSE, and MAE values of 0.65, 23.29 μg/m^3, and 13.87 μg/m^3, respectively (Fig. 4a), indicating that our proposed hindcast model outperformed previous models, with leave-one-year-out CV R^2 values less than 0.59 (He et al., 2021a; Ma et al., 2016; Meng et al., 2021). Due to the input data completeness, the hindcast model performs much better for monthly and yearly averages than the daily level. The monthly and yearly averages derived from the leave-one-year-out CV results agree well with in situ measurements, with R^2 (RMSE, MAE) values of 0.83 (12.10 μg/m^3, 7.79 μg/m^3) for monthly results and 0.85 (8.05 μg/m^3, 5.59 μg/m^3) for yearly results (Fig. 4b and c).

To verify the accuracy of the hindcast estimates, we conducted an individual evaluation by comparing the hindcast $PM_{2.5}$ estimates against 2005–2012 surface observations (Fig. 5). The monitoring $PM_{2.5}$ data enveloped in the individual evaluation were collected from five U.S. embassy sites in mainland China and official monitoring sites in Hong Kong and Macau, and were not used in model development and prediction. The overall individual validation R^2 (RMSE) is 0.67 (18.16 μg/m^3) at a daily scale and 0.81 (9.52 μg/m^3) at a monthly scale (Fig. 5),

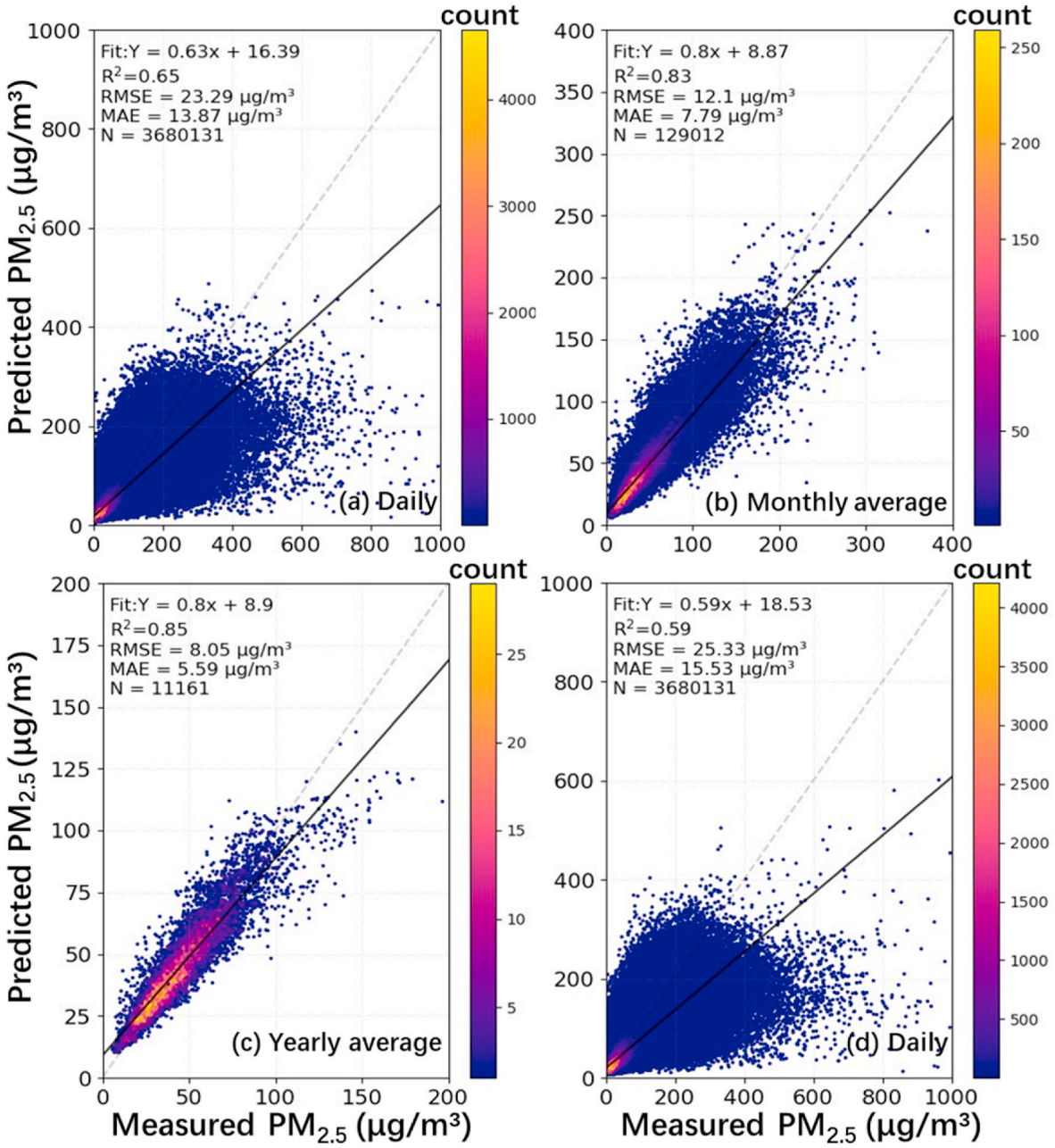

Fig. 4. Density plots of leave-one-year-out CV results of (a–c) our proposed model and (d) the simplified RF model.

suggesting that the leave-one-year-out CV results (Fig. 4) are comparable to the individual comparison. The spatial map of individual evaluation results (Fig. S6) shows the spatial variability in the pre-2013 modeling: R^2 values are 0.42–0.72 for grid cells with observations and the site RMSE value is 17.27 μg/m^3 on average. The R^2 value is generally associated with the spatially varying relationship between observed $PM_{2.5}$ and its predictors (He et al., 2021b). Generally, the daily full-coverage predictions outside the model-training years are reliable and thus can be used to analyze spatiotemporal $PM_{2.5}$ patterns and promote downstream applications for environmental management and epidemiological studies related to $PM_{2.5}$.

3.2. Improvements in $PM_{2.5}$ estimates prior to 2013

This section focuses on the added values of the proposed method to estimate $PM_{2.5}$ concentrations in years without $PM_{2.5}$ measurements. First, we developed a regular RF model with a simplified structure, i.e., without incorporating historical similarity, spatial, and temporal features, to estimate daily $PM_{2.5}$ concentrations (Fig. 4d). The input samples were the same as the model with full additional features to ensure comparability between the models. The comparison results demonstrate that our proposed hindcast model outperforms the simplified model, increasing the leave-one-year-out CV R^2 by 0.06 (0.65 vs. 0.59, Fig. 4a and d) at a daily scale. Then, to explore how the proposed hindcast method, especially the historical similarity features, improves the model's predictive power, we developed a batch of RF models with various structures and evaluated their performance through the leave-one-year-out CV. The validation results in Table 1 show that the model corrected by historical similarity features performed better, with leave-one-year-out CV R^2 and RMSE values of 0.65 and 23.24 μg/m^3, than the model with spatial/temporal features (0.60–0.64, 23.78–24.94 μg/m^3). In addition, the variable importance in Fig. S7 demonstrates that the historical similarity feature group is critical for predicting $PM_{2.5}$ outside the model-training period, having an importance degree of

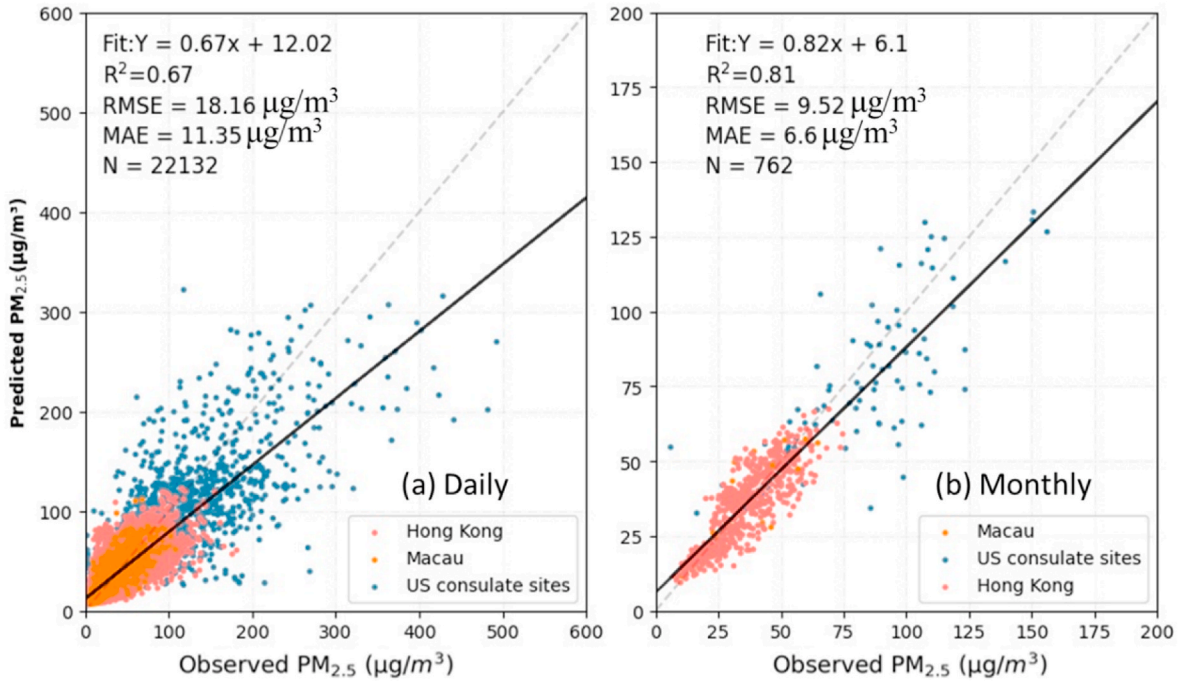

Fig. 5. Individual evaluation for the hindcasting performance with 2005–2012 surface $PM_{2.5}$ measurements.

Table 1
Leave-one-year-out CV results at a daily scale for random forest models with various model structures using samples in the Yangtze River Delta. (Sample size = 742,534).

Models	R^2	RMSE (μg/m^3)	Slope	Intercept
f(obs_sim variables)	0.59	25.34	0.61	21.88
f(obs_sim variables, Tfea)	0.60	24.94	0.63	20.98
f(obs_sim variables, Sfea)	0.64	23.78	0.64	19.72
f(obs_sim variables, Hfes)	0.65	23.24	0.65	18.86
Full model	0.66	22.88	0.65	18.71
Full model but no AOD	0.63	24.07	0.62	20.50

Obs_sim variables represent the observation or simulation predictors described in Section 2.1; Tfea, Sfea, and Hfes represent temporal, spatial, and historical similarity features mentioned in Section 2.2.1-2.2.3. Full model refers to models with the same variables and additional features as the model developed in the present study.

~20%. Thus, these comparison and validation results indicate that embedding the historical similarity information into the hindcast modeling enhances the ability of the model to explain the relationship between $PM_{2.5}$ and its predictors in the period without $PM_{2.5}$ observations.

Next, we examined the effect of imputed full-coverage AOD on $PM_{2.5}$ modeling performance for the pre-2013 period. Although the high-resolution AOD did not rank top in variable importance (Fig. S7), it should not be overlooked. Table 1 demonstrates that the pre-2013 modeling accuracy considerably reduces without AOD (the leave-one-year-out CV R^2 of 0.66 vs. 0.63), suggesting that the uncertainty in daily hindcast estimates was significantly alleviated by MAIAC-imputed AOD used in this study. More importantly, the imputed gap-free AOD helps us obtain $PM_{2.5}$ estimates with complete spatial coverage. Fig. 6 is an example showing gap-free maps of daily $PM_{2.5}$ estimates in a typical $PM_{2.5}$ episode, whereas the Terra and Aqua combined MAIAC retrievals can only provide 30%–48% of the data for China. Compared with the MAIAC-based predictions in the lower panel, our predictions (upper panel) demonstrate a complete picture of the evolution of the pollution event, which can help clarify the formation and mechanism of such an event. Moreover, as the only input with simultaneously spatial and temporal high resolution, AOD provides the unique opportunity to describe fine-scale gradients of daily particulate variations that other variables cannot replace. Thus, the MAIAC-imputed AOD enhanced the model performance in the years without $PM_{2.5}$ measurements as well as improved the spatiotemporal representativeness of $PM_{2.5}$ estimates.

3.3. *Spatiotemporal variations in long-term $PM_{2.5}$ exposure*

Using the full-coverage, high-resolution predictions, we investigated spatiotemporal variations in $PM_{2.5}$ concentrations and exposure in China, both nationwide and in selected regions, from 2000 to 2020. The obtained general spatial pattern and overall trend (Fig. S8) of $PM_{2.5}$ are consistent with previously reported results (van Donkelaar et al., 2021; Xue et al., 2019). Fig. 7 elucidates the temporal variations in population exposure to ambient $PM_{2.5}$ for China as a whole and for selected regions from 2000 to 2020 by relating the predicted $PM_{2.5}$ maps to yearly demographic distribution data. Fig. 7a demonstrates that on ~36% of the total days from January 2000 to December 2020, over 80% of the Chinese population was exposed to severe $PM_{2.5}$ levels of >37.5 μg/m^3 (the third interim target of the last air quality guideline by the World Health Organization (WHO)). The number of such days rapidly increased in the first six years over the study period and remained high at ~190 per year between 2006 and 2012, after which it continuously declined to 13 days in 2020. The severest years are 2008 and 2011, with such days exceeding 200. The national population-weighted mean concentration is 56 μg/m^3 on average, with the highest occurring in 2011 (67.14 μg/m^3) and the lowest in 2020 (35.40 μg/m^3; Fig. 7b).

In addition to investigating the annual variation in $PM_{2.5}$, we also explored seasonal variations in $PM_{2.5}$ exposure on both national and regional scales (Fig. 7c). Generally, the population-weighted mean concentrations are maximum in the cold months of each year and minimum in the warm months. The seasonal exposure trajectories also show that the peaks in $PM_{2.5}$ exposures in the summer months occur two to three years prior to those in the winter months, which is attributable to the changes in sulfate and nitrate emissions (Li et al., 2017a; Wang et al., 2019). In addition, significant regional discrepancies were observed in the seasonal trends. Fig. 7c demonstrates that although the BTH region remains among the most severely polluted in the country,

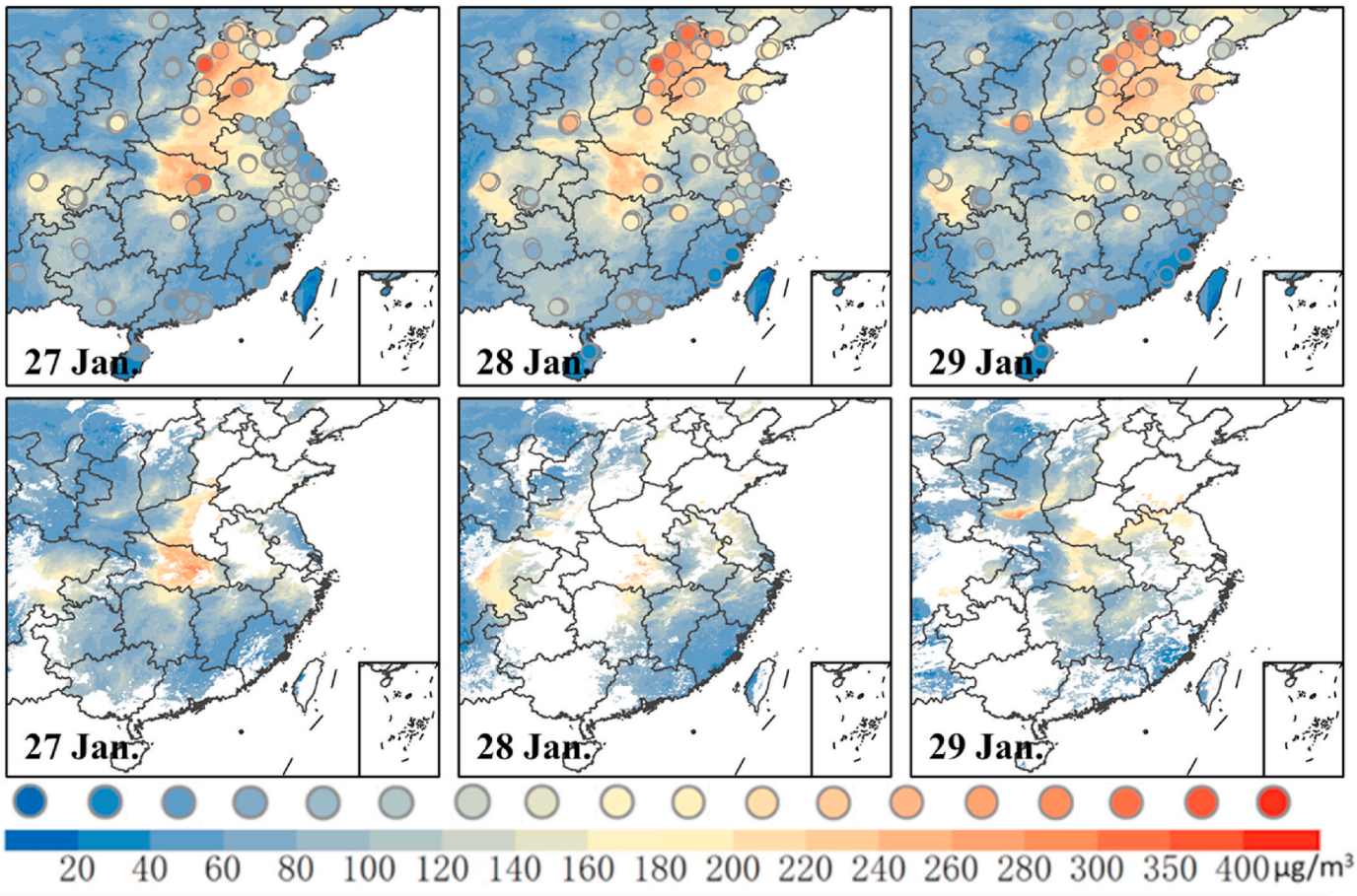

Fig. 6. A typical example of a severe $PM_{2.5}$ pollution episode during January 27–29, 2013, in eastern China. Upper panel: our full-coverage estimates based on MAIAC-imputed AOD. Lower panel: $PM_{2.5}$ estimates based on MAIAC AOD retrievals. The colored dots show that good accordance exists between the predicted and measured $PM_{2.5}$ levels.

2020 summertime exposures are below 37.5 μg/m^3, substantially less than the mean summer concentrations in 2002–2016, which are all greater than 50 μg/m^3. From 2015 to 2020, population exposures to $PM_{2.5}$ in the Pearl River Delta agglomeration were consistently lower than 25 μg/m^3 in summer but remained relatively high (>30 μg/m^3) in winter. In addition, we found that eastern $PM_{2.5}$ exposure increased by 3.72 μg/m^3 in the 2019 winter compared with the same period in 2018, especially in the BTH region, where the 2019 wintertime had an increase by 14.61 μg/m^3, offsetting 50% of the sharp decrease achieved in the previous winter.

Due to the long-term coverage of the dataset, we can quantitatively analyze $PM_{2.5}$-related issues in both historical and recent years. Fig. 8 exemplifies $PM_{2.5}$ variations in eastern China before (1–22 January), during (23 January to 7 April), and after (8 April to 31 May) the COVID-19 lockdown in Wuhan, providing insights into the $PM_{2.5}$ changes associated with COVID-19. Considering the various changing patterns in $PM_{2.5}$ in the past decades, we compared only the relative differences in $PM_{2.5}$ between 2019 and 2020. It can be clearly seen from Fig. 8 that the concentrations of $PM_{2.5}$ during the lockdown were significantly lower than during the same period in 2019, while the differences before and after the lockdown were smaller. This trend is somewhat reflected by the changes in the overall population-weighted mean values over eastern China. We observed a significant decrease in population-weighted mean concentrations during the COVID-19 outbreak (−8.72 μg/m^3), while only slight decreases were observed before (the difference in population-weighted mean concentration is −1.42 μg/m^3) and after (−1.27 μg/m^3) the lockdown period.

4. Discussion

It is difficult to reconstruct long-term $PM_{2.5}$ distributions for China using statistical/machine learning methods. Due to the lack of pre-2013 $PM_{2.5}$ observation data, it is impossible to directly calibrate the $PM_{2.5}$–predictor relationship before 2013. An alternative approach is to construct the $PM_{2.5}$–predictors relationship using post-2013 observation data, with the assumption that the pre-2013 and post-2013 relationships are identical, and then directly apply the post-2013 relationship to estimate the pre-2013 $PM_{2.5}$ levels. This assumption was the cornerstone for most previous works on hindcasting historical $PM_{2.5}$ in developing countries like China, and the methods achieved good model performance at longer (monthly, seasonal, and yearly) timescales (validated $R^2 > 0.7$) (He et al., 2021a; Huang et al., 2018; Ma et al., 2016; Meng et al., 2021; Wei et al., 2021; Xiao et al., 2017). However, at a daily scale, the performance of those previous hindcast models was poor, with validated R^2 values of 0.41–0.59 (the validation methods and results are detailed in Table S4), although advanced statistical models (e.g., spatiotemporal geographically and temporally weighted regression model) or data-driven machine learning methods (e.g., RF models) were used to resolve the complex relationship between $PM_{2.5}$ and its predictors (He et al., 2021a; Meng et al., 2021; Xiao et al., 2021). According to the validation results in Fig. 4, our proposed method achieves a leave-one-year-out CV R^2 of 0.65/0.83 at a daily/monthly scale for pre-2013 modeling, which is over most previous studies for China that also adopted leave-one-year-out CV (validated R^2 of 0.50–0.59 at a daily scale in Table S4). In addition, our model outputs were better in terms of spatial (1 km/full coverage) and temporal (daily/2000–2020)

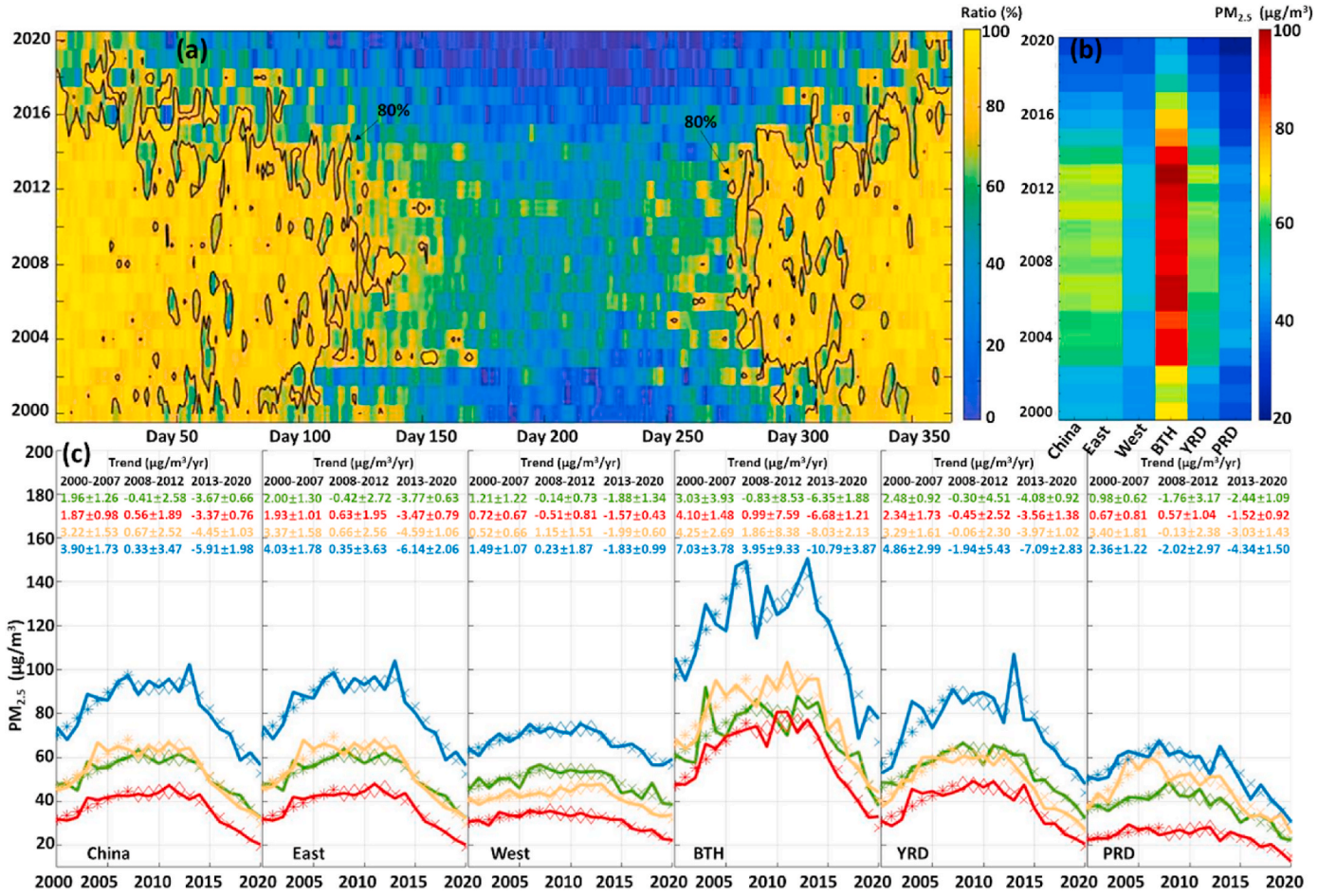

Fig. 7. Temporal variations in population exposure to $PM_{2.5}$ over China and selected regions: (a) Daily variations in the ratio of the population exposed to $PM_{2.5}$ concentration exceeding 37.5 μg/m^3 from 2000 to 2020; (b) annual mean population-weighted $PM_{2.5}$ concentrations over China and the five regions; (c) population-weighted mean $PM_{2.5}$ concentrations and linear trends for spring (in green), summer (in red), autumn (in yellow), and winter (in blue), where asterisks (*), diamonds (◊), and crosses (x) correspond to the 2001–2007, 2008–2012, and 2013–2020 trends, with 95% confidence intervals and colors matched to seasons. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

resolution/coverage than previous $PM_{2.5}$ datasets (He et al., 2021a; Wei et al., 2021; Xue et al., 2019). Thus, our proposed method achieves better performance (e.g., accuracy and data completeness) than, or at least comparable to, the previous long-term $PM_{2.5}$ models.

One reason for the improved pre-2013 accuracy was the incorporation of additional features described in Section 2.2.1-2.2.3 in the modeling, especially historical similarity features (Fig. 4, Table 1, and Fig. S7), which carried information regarding changes in the predictors between years with and without $PM_{2.5}$ monitoring data. This allowed the construction of $PM_{2.5}$–predictor relationships that were specific to each year in the period without $PM_{2.5}$ monitoring data. Another reason for the better performance of our method is the integration of *obs_sim predictors* in the long-term modeling. In addition to commonly used meteorological variables (e.g., RH and PBLH), MERRA-2 $PM_{2.5}$ was also incorporated in the modeling, ranking second among the nine observation predictors (Fig. S7). Moreover, the inclusion of imputed AOD in the modeling improved the accuracy of the pre-2013 estimates (Table 1) and helped to represent the complete picture of the fine-scale gradient of $PM_{2.5}$ (Fig. 6). It is worth noting that our hindcast validation was performed on the full-coverage samples. If only samples over the MAIAC-available areas were taken into account, our proposed method could gain higher hindcast performance since previous studies have shown that the hindcast model built only for areas with MAIAC AOD retrievals outperformed the full-coverage model (Meng et al., 2021; Xue et al., 2019).

A sharp countrywide decrease in $PM_{2.5}$ exposure was observed in China in recent years (Fig. 7 and S8), especially in BTH wintertime since 2013 (Fig. 7c). This was primarily attributed to the emission control policies implemented in China (Xiao et al., 2021; Zhong et al., 2021). To address the severe particulate pollution problem associated with rapid economic development, the Chinese government issued a series of strict air pollution control regulations and intervention measures from 2013 onwards, leading to significant reductions in nitroxide and sulfur oxide emissions as well as $PM_{2.5}$ levels (Ma et al., 2016). Zhong et al. (2021) and Xiao et al. (2021) decomposed the contributions of meteorology and emission to the $PM_{2.5}$ variabilities and found that although meteorological conditions were not conducive to $PM_{2.5}$ mitigation in the North China Plain and central China, $PM_{2.5}$ levels have decreased significantly in recent years. The recent decreasing trend indicates the reliability and feasibility of the issuance of strict air pollution control policies. However, previous studies found the unexpected slowness in wintertime $PM_{2.5}$ reduction in Henan province (Chen et al., 2023) and rebounded before the COVID-19 outbreak over northern cities (Dong et al., 2022). Likewise, we also found an increase in $PM_{2.5}$ exposure in 2019 compared to 2018, especially during eastern China (Fig. 7c), and observed only a slight decrease in $PM_{2.5}$ concentrations before and after the COVID-19 lockdown period of 2019–2020 (Fig. 8). According to Dong et al. (2022), anthropogenic factors could account for ~50% of the $PM_{2.5}$ rebound in northern China. Moreover, despite significant reductions in recent years, the annual population exposure level to $PM_{2.5}$, as indicated

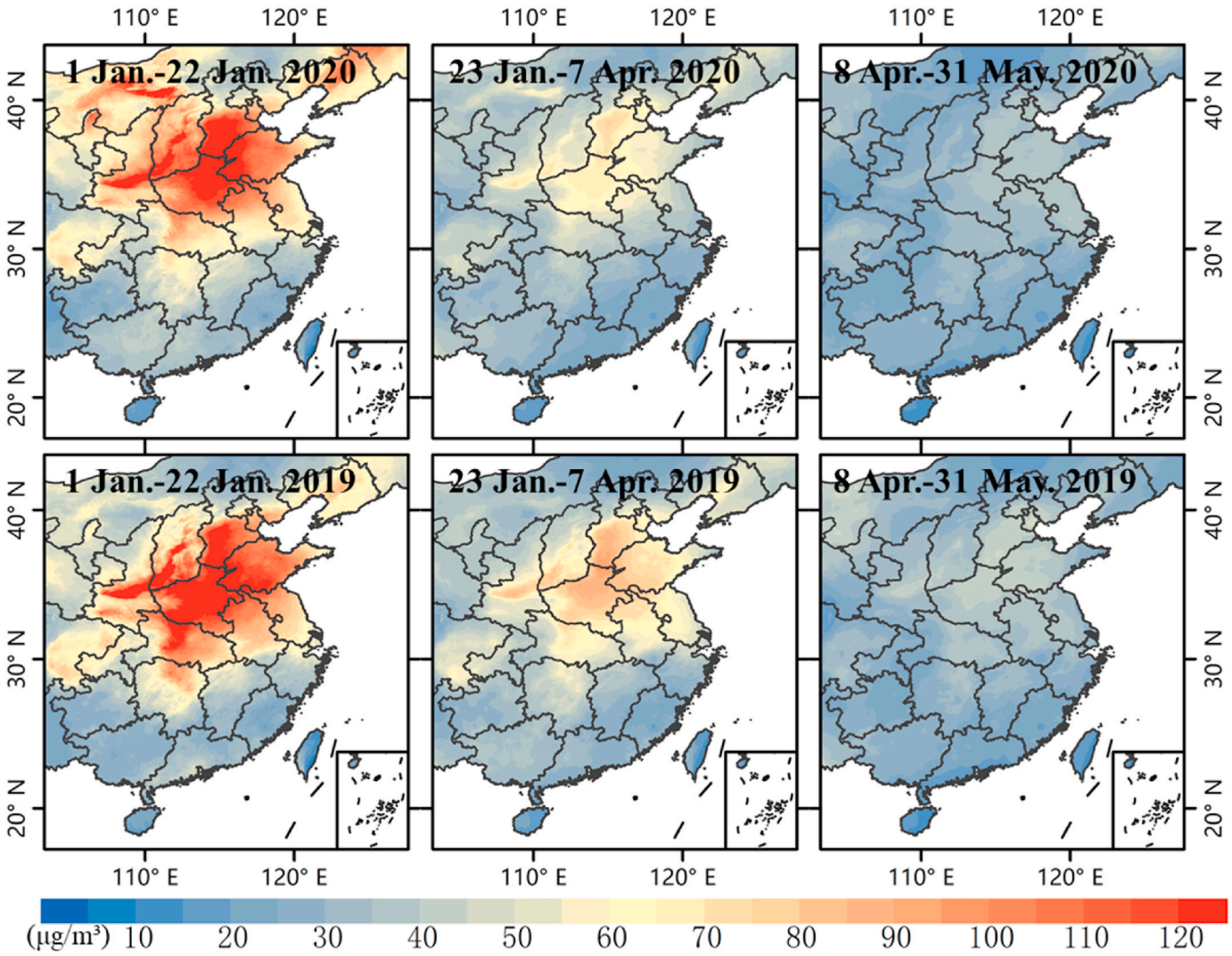

Fig. 8. Spatial maps of mean $PM_{2.5}$ concentrations before, during, and after the Wuhan lockdown due to the COVID-19 outbreak of 2020 and the same periods in 2019 for comparison.

by the population-weighted mean $PM_{2.5}$ concentration of 35.40 μg/m^3 in Fig. 7b, still falls short of meeting the first annual interim target of the 2021 WHO air quality guidelines, which is set at 35 μg/m^3. Thus, it appears that more action is needed to sustain the significant decreasing trend and eventually meet the latest WHO air quality guideline.

While $PM_{2.5}$ hindcasts have improved over the long historical period, this study has a few limitations. First, we only introduced annual pre-2013 information on predictors into the model development. More historical information, such as monthly similarities, could be further considered to enhance the model's performance in years without $PM_{2.5}$ observations. Second, while MERRA-2 $PM_{2.5}$ is important for modeling, it lacks the nitrate components, which may increase modeling bias over areas dominated by nitrate components. However, we also incorporated AOD, which reflects the total extinction coefficient of aerosols, into the modeling, helping to resolve this issue to some extent. Third, the modeling framework could be further refined, e.g., by using a simpler temporal feature to replace the current cosine forms in Section 2.2.3. In addition, due to the removal of the effect of spatial features in the western region, our predicted maps do not have obvious abnormal spatial variations (Fig. S4 and S8) while achieving higher overall accuracies (Figs. 3–4) than the national model without spatial features (Table S2). However, future users of our long-term estimates should be cautious with the relatively large uncertainty in the three western provinces.

5. Conclusion

We have developed a machine-learning hindcasting framework that can predict full-coverage, high-resolution $PM_{2.5}$ concentrations in China over a long period. In addition to incorporating spatial and temporal features, our method utilizes historical similarity features to capture variations in the observation/simulation between the period with and without $PM_{2.5}$ observations. We also addressed gaps in $PM_{2.5}$ estimates induced by satellite AOD by incorporating MAIAC-imputed AOD and discussed the effect of imputed AOD on full-coverage $PM_{2.5}$ modeling over the period without $PM_{2.5}$ observations. Our proposed method generates reliable daily $PM_{2.5}$ concentrations for China spanning 21 years at a 0.01° spatial resolution with a few modeling errors, including the 13 years before the establishment of the national air quality monitoring network. Our long-term estimates reveal that $PM_{2.5}$ levels across the country have significantly decreased since 2013. However, the national exposure level in 2020 still exceeded the first annual interim target of the 2021 WHO air quality guidelines, and wintertime $PM_{2.5}$ exposure rebounded slightly in 2019 in eastern China compared with 2018. These results highlight the need for more sustainable air pollution control policies. Our proposed hindcasting method can advance long-term modeling to supplement air quality monitoring data, and the resultant dataset can benefit environmental management and $PM_{2.5}$-related health studies.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The final daily 1-km full-coverage $PM_{2.5}$ dataset generated by this study is freely available via http://doi.org/10.5281/zenodo.4569557.

Acknowledgment

This study is supported by the National Natural Science Foundation of China (Grant NO.41901324).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jenvman.2023.118145.

References

Bai, K., Li, K., Ma, M., Li, K., Li, Z., Guo, J., Chang, N.-B., Tan, Z., Han, D., 2022. LGHAP: the Long-term Gap-free High-resolution Air Pollutant concentration dataset, derived via tensor-flow-based multimodal data fusion. Earth Syst. Sci. Data 14, 907–927.
Bi, J., Belle, J.H., Wang, Y., Lyapustin, A.I., Wildani, A., Liu, Y., 2019. Impacts of snow and cloud covers on satellite-derived PM2.5 levels. Rem. Sens. Environ. 221, 665–674.
Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32.
Brokamp, C., Jandarov, R., Hossain, M., Ryan, P., 2018. Predicting daily urban fine particulate matter concentrations using a random forest model. Environ. Sci. Technol. 52, 4173–4179.
Buchard, V., Randles, C., Da Silva, A., Darmenov, A., Colarco, P., Govindaraju, R., Ferrare, R., Hair, J., Beyersdorf, A., Ziemba, L., 2017. The MERRA-2 aerosol reanalysis, 1980 onward. Part II: evaluation and case studies. J. Clim. 30, 6851–6872.
Burnett, R., Chen, H., Szyszkowicz, M., Fann, N., Hubbell, B., Pope, C.A., Apte, J.S., Brauer, M., Cohen, A., Weichenthal, S., 2018. Global estimates of mortality associated with long-term exposure to outdoor fine particulate matter. Proc. Natl. Acad. Sci. USA 115, 9592–9597.
Cao, J., Xu, H., Xu, Q., Chen, B., Kan, H., 2012. Fine particulate matter constituents and cardiopulmonary mortality in a heavily polluted Chinese city. Environ. Health Perspect. 120, 373–378.
Chen, B., Huang, Y., Huang, J., Dong, L., Guan, X., Ge, J., Hu, Z., 2021. Using lidar and historical similar meteorological fields to evaluate the impact of anthropogenic control on dust weather during COVID-19. Front. Environ. Sci. 9.
Chen, B., Song, Z., Huang, J., Zhang, P., Hu, X., Zhang, X., Guan, X., Ge, J., Zhou, X., 2022a. Estimation of atmospheric PM10 concentration in China using an interpretable deep learning model and top-of-the-atmosphere reflectance data from China's new generation geostationary meteorological satellite, FY-4A. J. Geophys. Res. Atmos. 127, e2021JD036393.
Chen, B., Song, Z., Pan, F., Huang, Y., 2022b. Obtaining vertical distribution of PM2.5 from CALIOP data and machine learning algorithms. Sci. Total Environ. 805, 150338.
Chen, X., Li, X., Liang, J., Li, X., Li, S., Chen, G., Chen, Z., Dai, S., Bin, J., Tang, Y., 2023. Causes of the unexpected slowness in reducing winter PM2.5 for 2014–2018 in Henan Province. Environ. Pollut. 319, 120928.
Chudnovsky, A.A., Kostinski, A., Lyapustin, A., Koutrakis, P., 2013. Spatial scales of pollution from variable resolution satellite imaging. Environ. Pollut. 172, 131–138.
Dominici, F., Peng, R.D., Bell, M.L., Pham, L., McDermott, A., Zeger, S.L., Samet, J.M., 2006. Fine particulate air pollution and hospital admission for cardiovascular and respiratory diseases. JAMA 295, 1127–1134.
Dong, C., Li, J., Qi, Y., 2022. Decomposing PM2.5 air pollution rebounds in Northern China before COVID-19. Environ. Sci. Pollut. Control Ser. 29, 28688–28699.
Fang, X., Zou, B., Liu, X., Sternberg, T., Zhai, L., 2016. Satellite-based ground PM2.5 estimation using timely structure adaptive modeling. Rem. Sens. Environ. 186, 152–163.
Geng, G., Xiao, Q., Liu, S., Liu, X., Cheng, J., Zheng, Y., Xue, T., Tong, D., Zheng, B., Peng, Y., 2021. Tracking air pollution in China: near real-time PM2.5 retrievals from multisource data fusion. Environ. Sci. Technol. 55, 12106–12115.
Guo, J., Xia, F., Zhang, Y., Liu, H., Li, J., Lou, M., He, J., Yan, Y., Wang, F., Min, M., 2017. Impact of diurnal variability and meteorological factors on the PM2.5-AOD relationship: implications for PM2.5 remote sensing. Environ. Pollut. 221, 94–104.
Habre, R., Moshier, E., Castro, W., Nath, A., Grunin, A., Rohr, A., Godbold, J., Schachter, N., Kattan, M., Coull, B., 2014. The effects of PM2.5 and its components from indoor and outdoor sources on cough and wheeze symptoms in asthmatic children. J. Expo. Sci. Environ. Epidemiol. 24, 380–387.
He, Q., Gao, K., Zhang, L., Song, Y., Zhang, M., 2021a. Satellite-derived 1-km estimates and long-term trends of PM2.5 concentrations in China from 2000 to 2018. Environ. Int. 156, 106726.
He, Q., Wang, M., Yim, S.H.L., 2021b. The spatiotemporal relationship between PM2.5 and aerosol optical depth in China: influencing factors and implications for satellite PM2.5 estimations using MAIAC aerosol optical depth. Atmos. Chem. Phys. 21, 18375–18391.
He, Q., Wang, W., Song, Y., Zhang, M., Huang, B., 2023. Spatiotemporal high-resolution imputation modeling of aerosol optical depth for investigating its full-coverage variation in China from 2003 to 2020. Atmos. Res. 281, 106481.
He, Q., Zhang, M., Song, Y., Huang, B., 2021c. Spatiotemporal assessment of PM2.5 concentrations and exposure in China from 2013 to 2017 using satellite-derived data. J. Clean. Prod. 286, 124965.
Hough, I., Sarafian, R., Shtein, A., Zhou, B., Lepeule, J., Kloog, I., 2021. Gaussian Markov random fields improve ensemble predictions of daily 1 km PM2.5 and PM10 across France. Atmos. Environ. 264, 118693.
Hu, X., Belle, J.H., Meng, X., Wildani, A., Waller, L.A., Strickland, M.J., Liu, Y., 2017. Estimating PM2.5 concentrations in the conterminous United States using the random forest approach. Environ. Sci. Technol. 51, 6936–6944.
Huang, K., Xiao, Q., Meng, X., Geng, G., Wang, Y., Lyapustin, A., Gu, D., Liu, Y., 2018. Predicting monthly high-resolution PM2.5 concentrations with random forest model in the North China Plain. Environ. Pollut. 242, 675–683.
Jiang, T., Chen, B., Nie, Z., Ren, Z., Xu, B., Tang, S., 2021. Estimation of hourly full-coverage PM2.5 concentrations at 1-km resolution in China using a two-stage random forest model. Atmos. Res. 248, 105146.
Levy, R., Mattoo, S., Munchak, L., Remer, L., Sayer, A., Hsu, N., 2013. The Collection 6 MODIS aerosol products over land and ocean. Atmos. Meas. Tech. Discuss 6, 159–259.
Li, C., McLinden, C., Fioletov, V., Krotkov, N., Carn, S., Joiner, J., Streets, D., He, H., Ren, X., Li, Z., 2017a. India is overtaking China as the world's largest emitter of anthropogenic sulfur dioxide. Sci. Rep. 7, 1–7.
Li, T., Shen, H., Zeng, C., Yuan, Q., Zhang, L., 2017b. Point-surface fusion of station measurements and satellite observations for mapping PM2.5 distribution in China: methods and assessment. Atmos. Environ. 152, 477–489.
Liang, F., Xiao, Q., Huang, K., Yang, X., Liu, F., Li, J., Lu, X., Liu, Y., Gu, D., 2020. The 17-y spatiotemporal trend of PM2.5 and its mortality burden in China. Proc. Natl. Acad. Sci. USA 117, 25601–25608.
Liang, F., Xiao, Q., Wang, Y., Lyapustin, A., Li, G., Gu, D., Pan, X., Liu, Y., 2018. MAIAC-based long-term spatiotemporal trends of PM2.5 in Beijing, China. Sci. Total Environ. 616–617, 1589–1598.
Lin, C., Li, Y., Lau, A.K., Deng, X., Tim, K., Fung, J.C., Li, C., Li, Z., Lu, X., Zhang, X., 2016. Estimation of long-term population exposure to PM2.5 for dense urban areas using 1-km MODIS data. Rem. Sens. Environ. 179, 13–22.
Liu, C., Chen, R., Meng, X., Wang, W., Lei, J., Zhu, Y., Zhou, L., Kan, H., Xuan, J., 2022. Criteria air pollutants and hospitalizations of a wide spectrum of cardiovascular diseases: a nationwide case-crossover study in China. Eco-Environ. Health 1, 204–211.
Liu, N., Zou, B., Feng, H., Wang, W., Tang, Y., Liang, Y., 2019. Evaluation and comparison of multiangle implementation of the atmospheric correction algorithm, Dark Target, and Deep Blue aerosol products over China. Atmos. Chem. Phys. 19, 8243–8268.
Lv, B., Hu, Y., Chang, H.H., Russell, A.G., Cai, J., Xu, B., Bai, Y., 2017. Daily estimation of ground-level PM2.5 concentrations at 4km resolution over Beijing-Tianjin-Hebei by fusing MODIS AOD and ground observations. Sci. Total Environ. 580, 235–244.
Lyapustin, A., Wang, Y., Korkin, S., Huang, D., 2018. MODIS Collection 6 MAIAC algorithm. Atmos. Meas. Tech. 11.
Ma, Z., Dey, S., Christopher, S., Liu, R., Bi, J., Balyan, P., Liu, Y., 2022. A review of statistical methods used for developing large-scale and long-term PM2.5 models from satellite data. Rem. Sens. Environ. 269, 112827.
Ma, Z., Hu, X., Sayer, A.M., Levy, R., Zhang, Q., Xue, Y., Tong, S., Bi, J., Huang, L., Liu, Y., 2016. Satellite-based spatiotemporal trends in PM2.5 concentrations: China, 2004–2013. Environ. Health Perspect. 124, 184–192.
Martins, V., Lyapustin, A., Carvalho, L., Barbosa, C., Novo, E., 2017. Validation of high-resolution MAIAC aerosol product over South America. J. Geophys. Res. Atmos.
Meng, X., Liu, C., Zhang, L., Wang, W., Stowell, J., Kan, H., Liu, Y., 2021. Estimating PM2.5 concentrations in Northeastern China with full spatiotemporal coverage, 2005-2016. Rem. Sens. Environ. 253, 112203.
Peng, R.D., Bell, M.L., Geyh, A.S., McDermott, A., Zeger, S.L., Samet, J.M., Dominici, F., 2009. Emergency admissions for cardiovascular and respiratory diseases and the chemical composition of fine particle air pollution. Environ. Health Perspect. 117, 957–963.
Pope III, A., Dockery, D.W., 2006. Health effects of fine particulate air pollution: lines that connect. J. Air Waste Manag. Assoc. 56, 709–742.
Provençal, S., Buchard, V., da Silva, A.M., Leduc, R., Barrette, N., Elhacham, E., Wang, S.-H., 2017. Evaluation of PM2.5 surface concentration simulated by version 1 of the NASA's MERRA aerosol reanalysis over Israel and taiwan. Aerosol Air Qual. Res. 17, 253.
Randles, C., Da Silva, A., Buchard, V., Colarco, P., Darmenov, A., Govindaraju, R., Smirnov, A., Holben, B., Ferrare, R., Hair, J., 2017. The MERRA-2 aerosol reanalysis, 1980 onward. Part I: System description and data assimilation evaluation. J. Clim. 30, 6823–6850.
Rose, A., McKee, J., Sims, K., Bright, E., Reith, A., Urban, M., 2021. In: Oak Ridge, T.N. (Ed.), LandScan Global 2020. Oak Ridge National Laboratory.
Song, Z., Chen, B., Huang, J., 2022. Combining Himawari-8 AOD and deep forest model to obtain city-level distribution of PM2.5 in China. Environ. Pollut. 297, 118826.

Stafoggia, M., Bellander, T., Bucci, S., Davoli, M., de Hoogh, K., de' Donato, F., Gariazzo, C., Lyapustin, A., Michelozzi, P., Renzi, M., Scortichini, M., Shtein, A., Viegi, G., Kloog, I., Schwartz, J., 2019. Estimation of daily PM10 and PM2.5 concentrations in Italy, 2013–2015, using a spatiotemporal land-use random-forest model. Environ. Int. 124, 170–179.

van Donkelaar, A., Hammer, M.S., Bindle, L., Brauer, M., Brook, J.R., Garay, M.J., Hsu, N.C., Kalashnikova, O.V., Kahn, R.A., Lee, C., 2021. Monthly global estimates of fine particulate matter and their uncertainty. Environ. Sci. Technol. 55, 15287–15300.

Wang, N., Lyu, X., Deng, X., Huang, X., Jiang, F., Ding, A., 2019. Aggravating O3 pollution due to NOx emission control in eastern China. Sci. Total Environ. 677, 732–744.

Wei, J., Li, Z., Lyapustin, A., Sun, L., Peng, Y., Xue, W., Su, T., Cribb, M., 2021. Reconstructing 1-km-resolution high-quality PM2.5 data records from 2000 to 2018 in China: spatiotemporal variations and policy implications. Rem. Sens. Environ. 252, 112136.

Xiao, Q., Wang, Y., Chang, H.H., Meng, X., Geng, G., Lyapustin, A., Liu, Y., 2017. Full-coverage high-resolution daily PM2.5 estimation using MAIAC AOD in the Yangtze River Delta of China. Rem. Sens. Environ. 199, 437–446.

Xiao, Q., Zheng, Y., Geng, G., Chen, C., Huang, X., Che, H., Zhang, X., He, K., Zhang, Q., 2021. Separating emission and meteorological contributions to long-term PM2.5 trends over eastern China during 2000–2018. Atmos. Chem. Phys. 21, 9475–9496.

Xue, T., Zheng, Y., Tong, D., Zheng, B., Li, X., Zhu, T., Zhang, Q., 2019. Spatiotemporal continuous estimates of PM2.5 concentrations in China, 2000-2016: a machine learning method with inputs from satellites, chemical transport model, and ground observations. Environ. Int. 123, 345–357.

Yang, N., Shi, H., Tang, H., Yang, X., 2022. Geographical and temporal encoding for improving the estimation of PM2.5 concentrations in China using end-to-end gradient boosting. Rem. Sens. Environ. 269, 112828.

Zhong, Q., Tao, S., Ma, J., Liu, J., Shen, H., Shen, G., Guan, D., Yun, X., Meng, W., Yu, X., Cheng, H., Zhu, D., Wan, Y., Hu, J., 2021. PM2.5 reductions in Chinese cities from 2013 to 2019 remain significant despite the inflating effects of meteorological conditions. One Earth 4, 448–458.

Zhang, W., He, Q., Wang, H., Cao, K., He, S., 2018. Factor analysis for aerosol optical depth and its prediction from the perspective of land-use change. Ecol. Indicat. 93, 458–469.