

## Evaluating and optimizing PM<sub>2.5</sub> stations in Yangtze River Delta from a spatial representativeness perspective

Heming Bai<sup>a,\*</sup>, Wenkang Gao<sup>b</sup>, Myeongsu Seong<sup>a</sup>, Rusha Yan<sup>c</sup>, Jing Wei<sup>d</sup>, Chong Liu<sup>e,f</sup>

<sup>a</sup> Research Center for Intelligent Information Technology, Nantong University, Nantong, China

<sup>b</sup> State Key Laboratory of Atmospheric Boundary Layer Physics and Atmospheric Chemistry, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China

<sup>c</sup> State Environmental Protection Key Laboratory of Formation and Prevention of Urban Air Pollution Complex, Shanghai Academy of Environmental Sciences, Shanghai, China

<sup>d</sup> Earth System Science Interdisciplinary Center, Department of Atmospheric and Oceanic Science, University of Maryland, College Park, MD, USA

<sup>e</sup> CMA Earth System Modeling and Prediction Centre, China Meteorological Administration, Beijing, China

<sup>f</sup> State Key Laboratory of Severe Weather, Chinese Academy of Meteorological Sciences, Beijing, China

### ARTICLE INFO

Handling Editor: J Peng

### ABSTRACT

The spatial representativeness (SR) of air quality monitoring stations is an important parameter when using site observations for air quality evaluation and health assessment. In this study, by using daily 1-km-resolution PM<sub>2.5</sub> concentrations from China High Air Pollutants dataset from 2016 to 2020, we adopted a Concentration Similarity Frequency method to estimate SR of the current PM<sub>2.5</sub> stations in 25 cities over Yangtze River Delta (YRD) in Eastern China. These stations were further adjusted based on our proposed optimization scheme. For the current stations, SR areas cover 68.53% of urban area and 79.63% of urban population in YRD, but only cover 25.82% of rural area and 40.50% of rural population. Additionally, annual population-weighted mean (PWM) PM<sub>2.5</sub> based on SR is more accurate for urban regions than rural regions. Compared to full coverage PWM PM<sub>2.5</sub>, the attributable deaths using SR-based PWM PM<sub>2.5</sub> for urban and rural regions of YRD were overestimated by 1.04% and 4.09%. These overestimations were only 0.10% and 2.26% when using the optimized stations. Applying the optimization scheme also led to a 25.71% reduction in the number of stations. Our findings would provide a valuable reference for deploying new stations in YRD, especially in rural regions.

### 1. Introduction

Spatial representativeness (SR) is an important parameter when interpreting measured data from air quality monitoring stations (Righini et al., 2014; Shi et al., 2018). Regarding health assessment, SR is useful to quantify population exposure to the atmospheric pollution measured in a given station (Piersanti et al., 2015). Besides, optimizing air quality monitoring network also should take SR into account to cover large spatial area with the minimum number of sites (Hao & Xie, 2018; Ikeda et al., 1981). Moreover, SR can provide useful information on the point-grid matching tasks, such as data assimilation in air quality models (Elbern et al., 2007) and validating satellite-derived air pollution concentration (Piersanti et al., 2015). Overall, SR is one of the most important factors to evaluate the layout of air quality monitoring stations.

As the nationwide monitoring network in China has become increasingly dense since 2013 (Zhai et al., 2019), it is valuable to estimate and evaluate SR performance of these stations. Such analyses are particularly urgent from an urban-rural difference perspective given a severe imbalance of spatial distribution of stations in China, with most sites being located in urban areas (Gao et al., 2020; Wei et al., 2023). To date, several studies have investigated SR of air quality monitoring stations in China (Costabile et al., 2006; Hao & Xie, 2018; Hohenberger et al., 2021; Ma et al., 2019; Shi et al., 2018; Yu et al., 2018). Most of these previous studies were limited to single city, and based on the specific measurement campaigns with dense monitoring networks or air quality models with high spatiotemporal resolution. However, it is hard to apply these methods to investigate a large research area since they are expensive or require more computational cost. By contrast, using satellite-based air pollution data is a cost-effective method to examine

\* Corresponding author.

E-mail addresses: [hemingbai@ntu.edu.cn](mailto:hemingbai@ntu.edu.cn) (H. Bai), [gaowenkang@mail.iap.ac.cn](mailto:gaowenkang@mail.iap.ac.cn) (W. Gao), [mysms0318@gmail.com](mailto:mysms0318@gmail.com) (M. Seong), [yanrs@saes.sh.cn](mailto:yanrs@saes.sh.cn) (R. Yan), [weijing\\_rs@163.com](mailto:weijing_rs@163.com) (J. Wei), [liuchong@cma.gov.cn](mailto:liuchong@cma.gov.cn) (C. Liu).

**Table 1**  
Summary of the data sets in this study.

Dataset	Content	Spatial resolution	Temporal resolution	Period for this study
ChinaHighPM <sub>2.5</sub>	PM <sub>2.5</sub>	1 km × 1 km	Daily	2016–2020
GPWv4	Population	1 km × 1 km	Annual	2020
Seventh China Census GBD 2019	Population	City-level	–	2020
GUB_Global_2018	Mortality rate Urban boundary	National-level 30 m × 30 m	Annual –	2019 2018

SR over a large study area. For instance, Yu et al. (2018) combined remote-sensing data with the stratified sampling approach to assess SR of monitoring stations in Beijing-Tianjin-Hebei. Yet, their SR estimates may suffer uncertainty due to a coarse spatial resolution of satellite product ( $0.125^\circ \times 0.125^\circ$ ). Here, we use satellite-derived 1 km resolution air pollutant data to estimate SR of stations in YRD, and further investigate their differences between urban and rural regions.

SR estimation can be used to improve health assessment from site observation. Specifically, for estimating city-level air pollution exposure, SR can provide information on the population weights of different stations in a given city. Numerous studies have used site-based observation for health assessment and generally directly averaged air pollution data from each site within a city to derive the exposure concentrations (e.g., Bai, Gao, et al., 2022; Chen et al., 2017; Guan et al., 2021; Shi et al., 2021; Song et al., 2017). This strategy assumes the equal population weight of different stations, which would likely result in uncertainty in health assessment. Additionally, for rural regions without deploying air quality monitoring stations, some rural population may be covered by SR areas of urban sites; consequently, in these cases, SR estimation can be also used to assess rural air pollution exposure. To date, such studies are few. Thus, one important objective of this study is to estimate health assessment by using SR, and further examine their uncertainties.

Air quality monitoring stations in China are primarily deployed in urban regions (Gao et al., 2020), and tend to be clustered in areas with

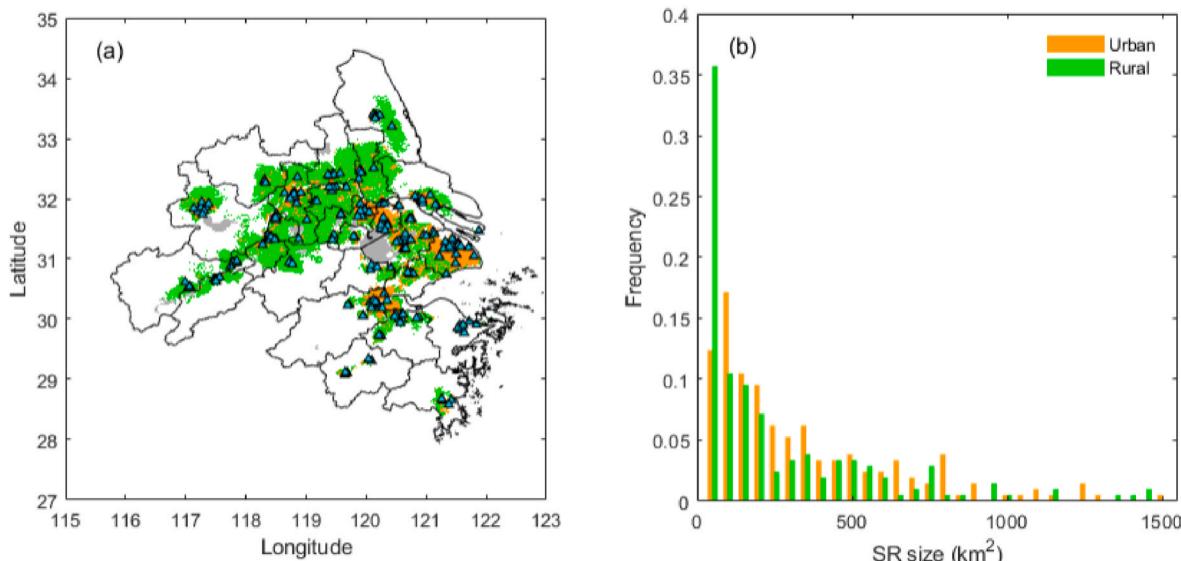
poor air quality, such as in street canyons and city centers (Yu et al., 2018). Consequently, the stations may have poor SR, especially in rural areas. Additionally, according to ‘technical regulation for selection of ambient air quality monitoring stations’ (on trial) (HJ 664–2013) released in China in 2013 (MEPC, 2013), the principles and requirements for determining station’s location are specified. However, the method for optimizing the network stations is not described in detail in HJ 664–2013 (Bai et al., 2017). Therefore, optimizing air quality monitoring network in China is important for efficient monitoring of air quality. Several studies used SR estimations to adjust existing monitoring networks in China (Costabile et al., 2006; Hao & Xie, 2018; Yu et al., 2018). Yet, these studies did not further investigate health assessment from optimized stations. To this end, this study not only optimizes the current stations, but also systematically evaluates the optimized stations with respect to SR performance and health assessment.

This study aims to investigate SR of the current air quality monitoring stations in Yangtze River Delta (YRD) using multi-year satellite-derived air pollutant data with 1 km spatial resolution. Given that particulate matter with an aerodynamic diameter of 2.5  $\mu\text{m}$  (PM<sub>2.5</sub>) or less is the primary pollutant in most cities in China (Lv et al., 2015), we only focus on SR of PM<sub>2.5</sub> stations. Additionally, we analyze the difference of SR performance between urban and rural regions. Furthermore, PM<sub>2.5</sub> exposure concentrations and deaths attributable to PM<sub>2.5</sub> exposure are estimated by using SR, and their uncertainties are also examined. Finally, we propose an optimization approach to adjust the current network stations, and further evaluate SR performance and health assessment based on the optimized stations.

The remainder of this paper was organized as follows. Section 2 introduced data and methods used to analyze SR performance and optimize the layout of stations. The results were presented in Section 3. Section 4 discussed the limitation and provided some policy recommendations. Finally, the conclusions were made in Section 5.

## 2. Data and methods

The main data sets used in this study were briefly listed in Table 1. Note that the different data sets should be restricted to the same year. However, using single-year PM<sub>2.5</sub> data may cause a large uncertainty in SR estimates. This study does not aim to provide an accurate estimate of



**Fig. 1.** (a) The spatial representativeness (SR) areas of existing PM<sub>2.5</sub> monitoring stations (triangles) in Yangtze River Delta, and orange (green) areas stand for urban (rural) SR areas. Panel (b) presents the frequency distribution of urban (orange) and rural (green) SR sizes of stations.

**Table 2**The performance of PM<sub>2.5</sub> monitoring stations for 25 cities in Yangtze River Delta.

City	Results of existing stations (N = 210)				Results of optimized stations (N = 156)			
	Area ratio	Population ratio	Redundancy	I_SR	Area ratio	Population ratio	Redundancy	I_SR
Anqing	0.09	0.20	0.10	0.14	0.79	0.91	0.09	0.81
Hefei	0.16	0.48	0.06	0.31	0.89	0.93	0.05	0.88
Chizhou	0.16	0.36	0.06	0.25	0.91	0.91	0.04	0.89
Chuzhou	0.16	0.24	0.09	0.19	0.91	0.93	0.08	0.88
Ma'anshan	0.48	0.73	0.27	0.52	0.79	0.91	0.34	0.71
Tongling	0.60	0.84	0.37	0.59	0.97	0.96	0.10	0.92
Wuhu	0.72	0.89	0.13	0.75	0.93	0.96	0.13	0.88
Xuancheng	0.12	0.23	0.07	0.17	0.87	0.90	0.04	0.87
Changzhou	0.81	0.92	0.09	0.83	0.82	0.93	0.14	0.81
Nanjing	0.88	0.94	0.07	0.87	0.91	0.95	0.10	0.89
Nantong	0.20	0.35	0.04	0.27	0.83	0.91	0.03	0.85
Suzhou	0.44	0.74	0.06	0.57	0.67	0.91	0.03	0.78
Taizhou*	0.60	0.68	0.06	0.62	0.86	0.91	0.08	0.85
Wuxi	0.47	0.75	0.07	0.59	0.76	0.91	0.05	0.81
Yancheng	0.19	0.24	0.02	0.21	0.83	0.92	0.08	0.84
Yangzhou	0.49	0.67	0.11	0.55	0.84	0.90	0.12	0.82
Zhenjiang	0.52	0.63	0.07	0.56	0.93	0.93	0.13	0.87
Shanghai	0.62	0.77	0.07	0.67	0.77	0.91	0.04	0.82
Hangzhou	0.15	0.69	0.11	0.40	0.78	0.91	0.03	0.83
Huzhou	0.18	0.36	0.01	0.27	0.86	0.91	0.05	0.86
Jiaxing	0.36	0.45	0.07	0.39	0.91	0.93	0.11	0.87
Jinhua	0.05	0.21	0.16	0.12	0.69	0.91	0.05	0.78
Ningbo	0.04	0.22	0.08	0.13	0.83	0.90	0.03	0.85
Shaoxing	0.28	0.62	0.13	0.42	0.83	0.91	0.02	0.86
Taizhou**	0.11	0.31	0.05	0.21	0.89	0.91	0.04	0.88

N: the number of stations; I\_SR: the index for evaluating SR performance of stations (see Section 2.2.3 for details); Taizhou\* from Jiangsu province; Taizhou\*\* from Zhejiang province.

attributable deaths for a specific year. One of the objectives of this work is to give robust SR estimates using multi-year PM<sub>2.5</sub> data, and further apply the SR estimates to calculate attributable deaths. The SR-based estimates of attributable deaths are compared to those attributable death estimates using the multi-year average of PWM PM<sub>2.5</sub>. Thus, the above calculation scheme is not designed to provide health assessments for a specific year, and it can be treated as a sensitivity experiment to explore the uncertainty of SR-based health assessments.

## 2.1. Data sets

This study focused on the YRD region, which is located in Eastern China and contains a total of 26 cities. This region represents the strongest economy power in China, but is also one of the regions with the heaviest air pollution in China (Bai et al., 2021; Yang et al., 2019). Among those cities in YRD, Zhoushan has the smallest land area, but the city's population density is not the lowest. However, we excluded this city because the assessment of SR of PM<sub>2.5</sub> monitoring stations in the city may suffer from a large uncertainty, given that Zhoushan is surrounded by sea and PM<sub>2.5</sub> data used in the study are unavailable in water area. Finally, we estimated SR of 210 p.m.<sub>2.5</sub> monitoring stations in 25 cities in YRD. The spatial distribution of these current stations is shown in Fig. 1a, and the specific cities are listed in Table 2. The details of the data used in this study are introduced as follows.

(1) Gridded daily PM<sub>2.5</sub> data. We estimated SR by using the 1-km-resolution daily PM<sub>2.5</sub> concentrations (i.e., ChinaHighPM<sub>2.5</sub>) collected from the CHAP data set (<https://weijing-rs.github.io/product.html>). This dataset was generated from the combined MODIS/Terra + Aqua MAIAC AOD products together with other auxiliary data using a newly developed space-time extremely randomized trees model (Wei et al., 2020, 2021). The daily PM<sub>2.5</sub> estimates have high accuracy with an average cross-validation

coefficient of determination of 0.92. Here, we employed the ChinaHighPM<sub>2.5</sub> data for the period from 2016 to 2020 to ensure that more samples could be collected.

- (2) Urban boundary data. The urban boundary data in 2018 were applied to identify whether each grid (1 km × 1 km) in YRD corresponds to urban grid or rural grid (Li et al., 2020). The data were generated by using the global artificial impervious area data (Gong et al., 2020), and can be freely accessed from <http://data.ess.tsinghua.edu.cn/gub.html>. Urban areas in YRD can be seen in Fig. S1.
- (3) Population and mortality data. Population data with 1 km spatial resolution for 2020 were from the Gridded Population of the World, Version 4 (Doxsey-Whitfield et al., 2015), and were obtained at <https://sedac.ciesin.columbia.edu/data/collection/gpw-v4>. City-level population data for 2020 from the Seventh China Census were also used to adjust the above gridded population data: for a given city, we scaled the gridded population counts by the ratio of the city's total population from census data to gridded data. Data on the age structure at national level for 2019, as well as the age specific and disease-specific mortality were obtained from Global Burden of Disease Study 2019 (GBD 2019) dataset (<https://vizhub.healthdata.org/gbd-compare/>). Note that we estimated the attributable deaths based on the multi-year average of PM<sub>2.5</sub>. This is because that SR-based PM<sub>2.5</sub> exposure and corresponding attributable deaths were calculated by using SR estimates, which are mean values based on a large sample size of pairs of PM<sub>2.5</sub> data from 2016 to 2020.

## 2.2. Methods

### 2.2.1. Spatial representativeness

We estimated SR of PM<sub>2.5</sub> monitoring stations by adopting a method of Concentration Similarity Frequency (CSF) defined as the following

(Piersanti et al., 2015):

$$F_{site}(X_{grid}, Y_{grid}) = \frac{\sum_{k=1}^{N_t} Flag}{N_t}, \text{ where } Flag = \begin{cases} 1, & \frac{|PM2.5(X_{site}, Y_{site}, t_k) - PM2.5(X_{grid}, Y_{grid}, t_k)|}{PM2.5(X_{site}, Y_{site}, t_k)} < 0.2 \\ 0, & \frac{|PM2.5(X_{site}, Y_{site}, t_k) - PM2.5(X_{grid}, Y_{grid}, t_k)|}{PM2.5(X_{site}, Y_{site}, t_k)} > 0.2 \end{cases} \quad (1)$$

where  $F_{site}(X_{grid}, Y_{grid})$  is a frequency function to determine whether a grid point ( $X_{grid}$ ,  $Y_{grid}$ ) is included in SR area of a site ( $X_{site}$ ,  $Y_{site}$ ).  $PM2.5(X, Y, t)$  represents surface concentration field from  $PM_{2.5}$  data at location (X, Y) at time t. Flag is the concentration similarity at time  $t_k$  by comparing  $\Delta PM_{2.5}/PM_{2.5}$  with a threshold of 20% (Piersanti et al., 2015). We used gridded daily  $PM_{2.5}$  data from 2016 to 2020 to match with the  $PM_{2.5}$  site;  $N_t$  is the number of successfully matched  $PM_{2.5}$  data pairs for the site. To robustly calculate the frequency function  $F_{site}(X_{grid}, Y_{grid})$ , we set the minimum value of  $N_t$  as 180 (about 10% of  $PM_{2.5}$  image count from 2016 to 2020). We assumed the maximum SR area is a box of  $100 \text{ km} \times 100 \text{ km}$  centered on the site.  $N_t$  for each  $100 \text{ km} \times 100 \text{ km}$  box for each  $PM_{2.5}$  site in YRD can be seen in Fig. S2, which shows that nearly all grids within each box meet the threshold of 180 except for grids in water region. After calculating  $F_{site}(X_{grid}, Y_{grid})$  for each grid point in the box, SR area of the site was assessed as the area where the condition  $F_{site}(X_{grid}, Y_{grid}) > 0.9$  is met on multi-year basis (Piersanti et al., 2015).

### 2.2.2. Optimization approach for adjusting current stations

We proposed an optimization approach to adjust the layout of the current  $PM_{2.5}$  monitoring stations for more population covered by SR. This approach includes three major steps. First, we estimated SR area for each grid ( $1 \text{ km} \times 1 \text{ km}$ ) in a given city, and subsequently calculated populations corresponding to these SR areas. Second, the grids were sorted by populations, and the first grid point was selected as one of the optimized stations. Third, for all grids except the selected one, the above-mentioned two steps were repeated until population covered by the overall SR area of all selected grids reached a threshold (90% of population of the city). Note that if SR area of a grid overlaps SR areas of the selected grids in the first step, then the overlapping area was cut out from SR area for this grid.

In the first step of the optimization approach, we did not use the data of the current  $PM_{2.5}$  stations; we treated each grid point in the study area as a potential  $PM_{2.5}$  station and evaluated SR performance for these grid points. The optimization approach in the present study only considered the SR and population. This approach aims to provide an accurate health assessment using site-based  $PM_{2.5}$  observations. The actual site selection should consider several other aspects besides SR, e.g., the category of monitoring sites (industrial, traffic, and port stations) (MEPC, 2013). However, it is hard to obtain data on such categories of land use in the entire YRD region. Information on different categories of current stations is also unavailable. To improve the optimization approach, we plan to perform city-level analysis in the future, considering the actual site selection requirements.

### 2.2.3. Evaluation of SR performance

We developed an index ( $I_{SR}$ ) to comprehensively evaluate SR performance of stations for a given city  $i$ , which is defined as the following:

$$I_{SR} = \frac{r\_area_i + r\_pop_i}{e^{R_i} + 1} \quad (2)$$

where  $r\_area_i$  is the ratio of the overall SR area of all stations in city  $i$  to city area;  $r\_pop_i$  is the ratio of the population covered by the overall SR

areas to city population;  $R_i$  represents the redundancy of stations defined

as the following formula (Bai, Yan, et al., 2022):

$$R_i = \frac{\sum_{j=2}^n (j-1) \times S_{ij}}{(n-1) \times S_i} \quad (3)$$

where  $n$  is the number of stations in city  $i$ , and  $S_i$  is the overall SR area of  $n$  stations, and  $S_{ij}$  is the overlap SR area. equation (3) implies that a weight of overlap SR area ( $S_{ij}$ ) is positively correlated with the number of stations (i.e.,  $j$ ) that contribute to this overlap area. The weight is assigned as  $j-1$  ( $j \geq 2$ ) to make the maximum  $R_i$  be 1.  $R_i$  ranges from 0 to 1, 0 indicating no overlap area and thus no redundancy in SR, and 1 indicating that the SR areas among  $n$  stations are exactly the same.

$I_{SR}$  comprehensively takes into account the spatial coverage of SR, population coverage of SR and redundancy of stations. This index ranges from 0 to 1, with larger values meaning better SR performance.

### 2.2.4. Estimating deaths attributable to $PM_{2.5}$ pollution

Deaths attributable to exposure to annual mean  $PM_{2.5}$  were estimated with equation (4), following a method from the GBD 2019 project (GBD 2019 Risk Factors Collaborators):

$$M_i(z) = \sum_a \sum_d \left( POP_{i,a} \times MB_{i,a,d} \times \frac{RR_{a,d}(z) - 1}{RR_{a,d}(z)} \right) \quad (4)$$

where  $i$ ,  $a$ , and  $d$  represent city, age group and cause of death, respectively;  $M$  denotes the attributable deaths at exposure level  $z$ ;  $POP$  stands for the population;  $MB$  is the baseline mortality rate.  $RR(z)$  is the relative risk associated with exposure at level  $z$ . Note that we used national-level age group and  $MB$  from GBD 2019 given that these city-level data are unavailable.

We estimated five mortality endpoints associated with  $PM_{2.5}$  pollution, including stroke, ischemic heart disease (IHD), chronic obstructive pulmonary disease (COPD), lung cancer (LC) and lower respiratory infection (LRI). Additionally, we applied an updated RR from a recent study (McDuffie et al., 2021), which was based on a Meta Regression-Bayesian, Regularized, Trimmed (MR-BRT) spline from the GBD 2019.

RR is based on annual population-weighted mean (PWM)  $PM_{2.5}$ . Here, calculating SR-based PWM  $PM_{2.5}$  is based on the following formula:

$$PWM_{SR,i} = \frac{\sum_j (PM_j \times POP_{SR,j})}{\sum_j POP_{SR,j}} \quad (5)$$

where  $PWM_{SR,i}$  is SR-based PWM  $PM_{2.5}$  for a given city  $i$ ;  $PM_j$  is annual  $PM_{2.5}$  concentration at station  $j$  and  $POP_{SR,j}$  is population covered by SR area of station  $j$ . To examine their reliability, SR-based PWM  $PM_{2.5}$  values were further compared to PWM  $PM_{2.5}$  calculated by using full coverage satellite-derived  $PM_{2.5}$  and gridded population data (referred to as full coverage PWM  $PM_{2.5}$ ).

A flow chart of this study is depicted in Fig. S3. Note that, for a given station in a given city, SR area of the station frequently covers rural areas in other cities. We consider this situation only for plotting the SR area for

the entire YRD region. For other analyses, such as frequency distribution of SR area and city level performance of SR, we focus on SR area within the city in which the station is located.

### 3. Results

#### 3.1. Spatial representativeness analysis for existing stations

For existing PM<sub>2.5</sub> monitoring stations in YRD, there is a large urban-rural difference in SR area. Urban (rural) SR areas are indicated by orange (green) areas in Fig. 1a. As can be seen, SR areas of the stations in YRD are mainly contributed by rural SR areas. The total area of rural SR (43487 km<sup>2</sup>) is about twice larger than that of urban area (20383 km<sup>2</sup>) (Fig. 1a). By contrast, the ratio of rural SR area to the entire rural area in YRD (25.82%) is much smaller than that for urban SR area (68.53%). This is likely due to the fact that almost all stations (93.33%) are located in urban areas (see triangles in Fig. 1a) in conjunction with a low ratio of urban area to entire YRD area (15.01%). Additionally, a difference in the frequency distribution of SR size between urban and rural areas is clearly observed, as shown in Fig. 1b. The frequency distribution is calculated by SR area of urban and rural regions represented by each station. Compared to urban results, the frequency distribution for rural areas shift to lower SR size (Fig. 1b), which suggests that small rural areas are frequently covered by SR areas.

The urban-rural difference in SR at city level is further examined. Blue (red) bars in Fig. 2a show ratios of urban (rural) SR area to total city area for each city in YRD. Blue (red) asterisks in Fig. 2a stand for urban (rural) area to total city area. As can be seen in Fig. 2a, rural SR area is larger than urban SR area for all cities except Suzhou, Wuxi, Shanghai, Hangzhou, Jinhua and Ningbo. These exceptional cities generally correspond to cities where urban area is close to rural area (see blue and red asterisks in Fig. 2a). On the other hand, for most cities, population covered by rural SR areas is smaller than that covered by urban SR areas, as shown in Fig. 2b. Note that population ratio is higher than area ratio for urban SR (blue bars in Fig. 2), and it is opposite for rural SR (red bars in Fig. 2). This can be partially attributed to a stronger population aggregation effect for urban region than rural region (Gaughan et al., 2016).

Fig. 2 shows ratios of SR area to total city area and ratios of population covered by SR area to total city population. For some cities in Fig. 2, small ratios do not indicate poor performance of SR. For instance,

ratio of urban SR area to total area of Chizhou city is only 0.92% (see blue bars in Fig. 2a). This low value is mainly due to the small size of the urban area in the city (1.4% of total city area). The ratio becomes 65.71% after dividing the urban SR area by urban area in Chizhou city. To address these issues, we further calculated ratios of urban (rural) SR areas of stations in a given city to the corresponding city's urban (rural) areas; these urban and rural area ratios for each city in YRD are shown in blue and red bars in Fig. 3, respectively. Additionally, blue (red) asterisks in Fig. 3 show ratios of urban (rural) population covered by SR areas of stations in each city to the corresponding city's urban (rural) population. For rural regions, SR performance is generally poor. Fig. 3b shows that there are only five cities where both rural area ratio and rural population ratio are higher 50%. Besides, there are nine cities with rural population ratios smaller than 20%, and these cities' population is dominated by rural population except Ningbo (rural population dominated cities: Anqing, Hefei, Chizhou, Chuzhou, Tongling, Wuhu, Xuan-cheng, Nantong, Taizhou, Yancheng, Yangzhou, Zhenjiang, Huzhou, Jinhua, Shaoxing and Taizhou). By contrast, SR performs well for urban regions in YRD. As shown in Fig. 3a, both urban area ratio and urban population ratio are higher than 50% for 16 cities.

#### 3.2. Health assessment based on SR

Combining the annual PM<sub>2.5</sub> concentration at each station with population covered by SR area of the corresponding station, we estimated SR-based PWM PM<sub>2.5</sub> for urban and rural areas in each city in YRD, and further investigated their uncertainties by comparing these estimates with the ones using gridded PM<sub>2.5</sub> and population datasets. An illustration of this comparison is given in Fig. 4, where the left panel indicates that SR-based PWM PM<sub>2.5</sub> for urban region is relatively accurate in most cities. The difference in urban PWM PM<sub>2.5</sub> between using SR and full coverage data is less than 5% for all cities except Xuancheng (8.45%) in Anhui province (Fig. 4a). On the other hand, SR-based PWM PM<sub>2.5</sub> for rural region is generally severely overestimated, even the one is overestimated by more than 10% for eight cities compared to full coverage PWM PM<sub>2.5</sub>, as shown in Fig. 4b. The reason for this overestimation is that calculating SR-based PWM PM<sub>2.5</sub> for rural region generally considers a minority of rural population living in highly polluted areas that covered by SR areas of stations; the calculation ignores most of rural population with low pollution exposure. By using SR-based PWM PM<sub>2.5</sub>, attributable deaths in the whole urban and rural

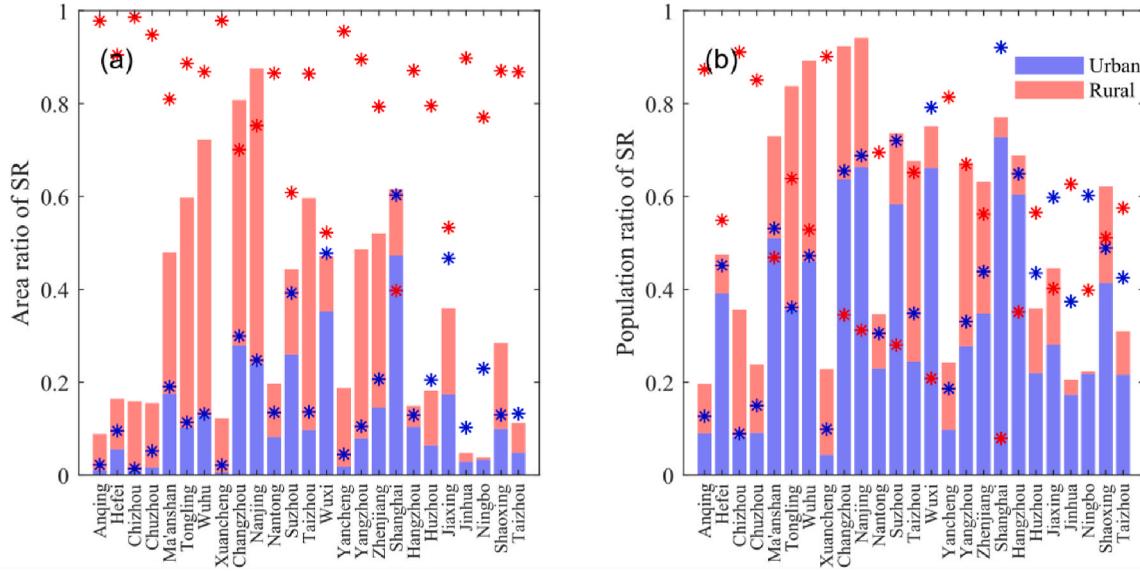


Fig. 2. Blue (red) bars in panel (a) present area ratio of urban (rural) SR to total city area for each city in YRD, and blue (red) asterisks in this panel stand for urban (rural) area to total city area. Panel (b) is the same as (a) but for population ratio.

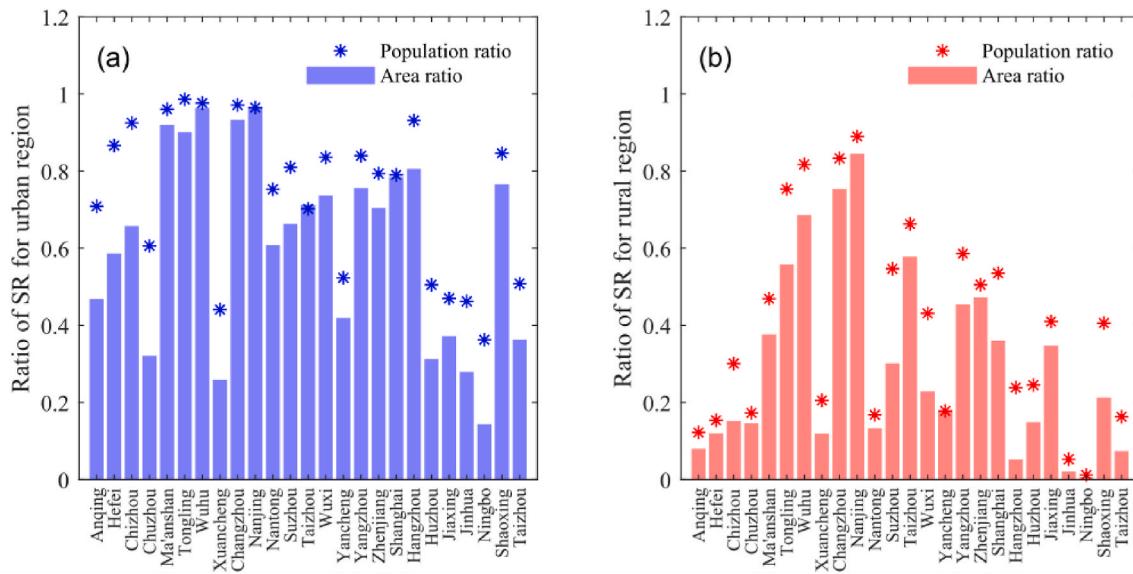


Fig. 3. Blue (red) asterisks denote ratios of urban (rural) population covered by SR areas of stations in each city to the corresponding city's urban (rural) population. Blue (red) bars stand for ratios of urban (rural) SR areas of stations in each city to the corresponding city's urban (rural) areas.

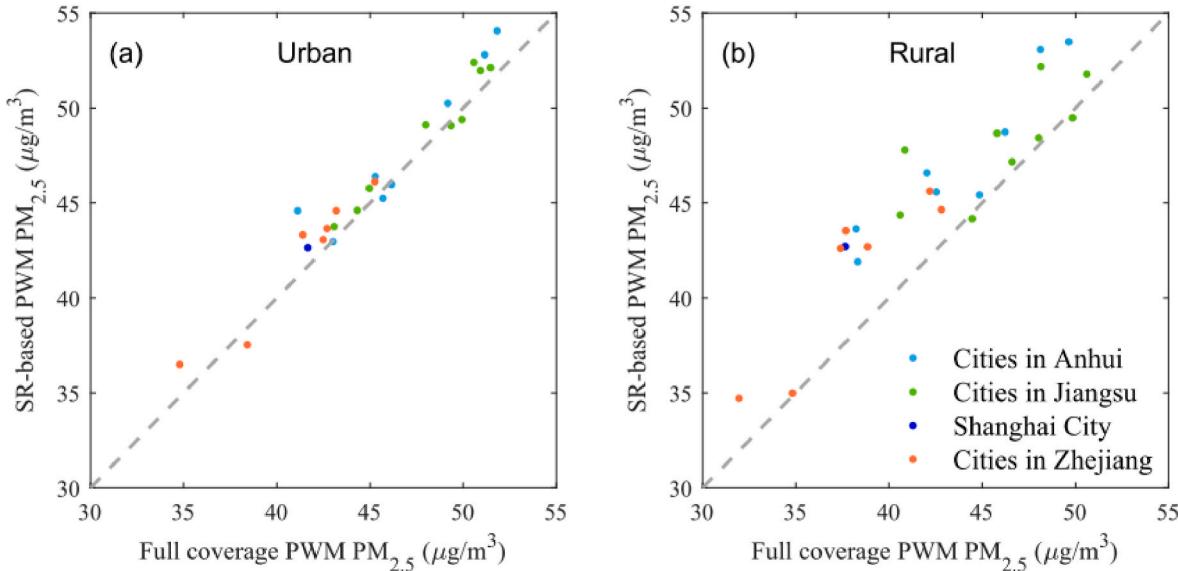


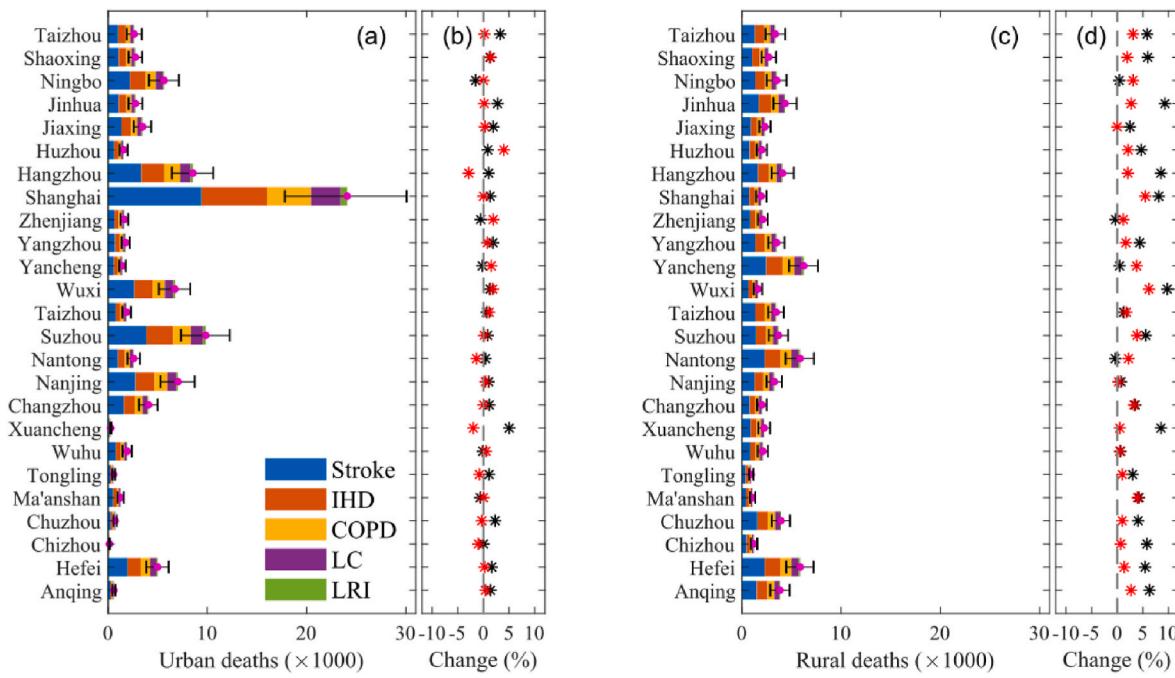
Fig. 4. Comparisons of annual population-weighted mean (PWM) between using SR estimates and using full coverage PM<sub>2.5</sub> data for (a) urban and (b) rural regions. Different color points correspond to cities in different provinces or municipality.

regions of YRD are 99.84 thousand (95% confidence interval (CI): 74.80–123.63) and 80.40 thousand (95% CI: 60.40–99.39) respectively, and overestimated by 1.04% and 4.09% compared to those by using full coverage PWM PM<sub>2.5</sub>. The differences of attributable deaths for each city are shown in Fig. 5. As can be seen, the difference for urban regions is relatively small, ranging from -1.53% in Ningbo to 5.03% in Xuancheng (see black asterisks in Fig. 5b). By contrast, the difference for rural regions varies greatly with cities, and is more than 5% in about half of the cities in YRD (black asterisks in Fig. 5d). Additionally, in rural population dominated cities except for Shaoxing, attributable deaths using full coverage PWM PM<sub>2.5</sub> in rural regions are higher than these in urban regions, with an increase ranging from 10.19% in Wuhu to 851.53% in Chizhou. This urban-rural difference expands when using SR-based PWM PM<sub>2.5</sub> and covers a range from 3.13% in Shaoxing to 909.63% in Chizhou. For attributable deaths based on SR-based PWM PM<sub>2.5</sub>, rural estimates are >100% higher than urban estimates for half of the cities

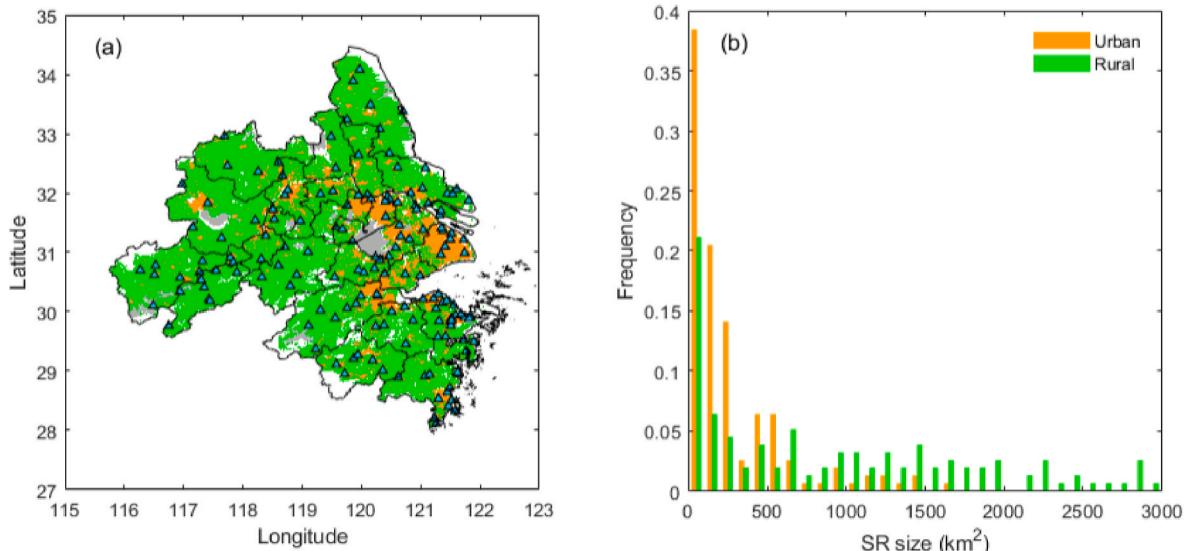
where population is dominated by rural population.

### 3.3. Analysis for optimized stations

The results above show that the layout of the current stations is somewhat unreasonable, especially for cities where population is dominated by rural population. In these cities, SR areas of the current stations cover a limited rural population, which leads to a high uncertainty in rural health assessment. Thus, we developed an optimization approach to adjust the current stations to account for more rural population. The distribution of the optimized stations is shown in Fig. 6a. As can be seen, comparing to the existing stations, the number of the optimized stations decreases by 54 (25.71%) while SR area of the optimized stations increases nearly twice for the entire YRD region. The results are partly because the optimized stations include more rural stations with larger SR size (Fig. 6b).



**Fig. 5.** Deaths attributable to PM<sub>2.5</sub> exposure by using full coverage PWM PM<sub>2.5</sub> for (a) urban and (c) rural regions in each city in YRD. The differences in urban and rural deaths between using full coverage PWM PM<sub>2.5</sub> and using SR-based PWM PM<sub>2.5</sub> are shown in (b) and (d) respectively, and black (red) asterisks stand for results based on existing (optimized) monitoring stations. Note that we calculated full coverage PWM PM<sub>2.5</sub>, SR-based PWM PM<sub>2.5</sub> and the attributable deaths by using the multi-year average of PM<sub>2.5</sub> data from 2016 to 2020.



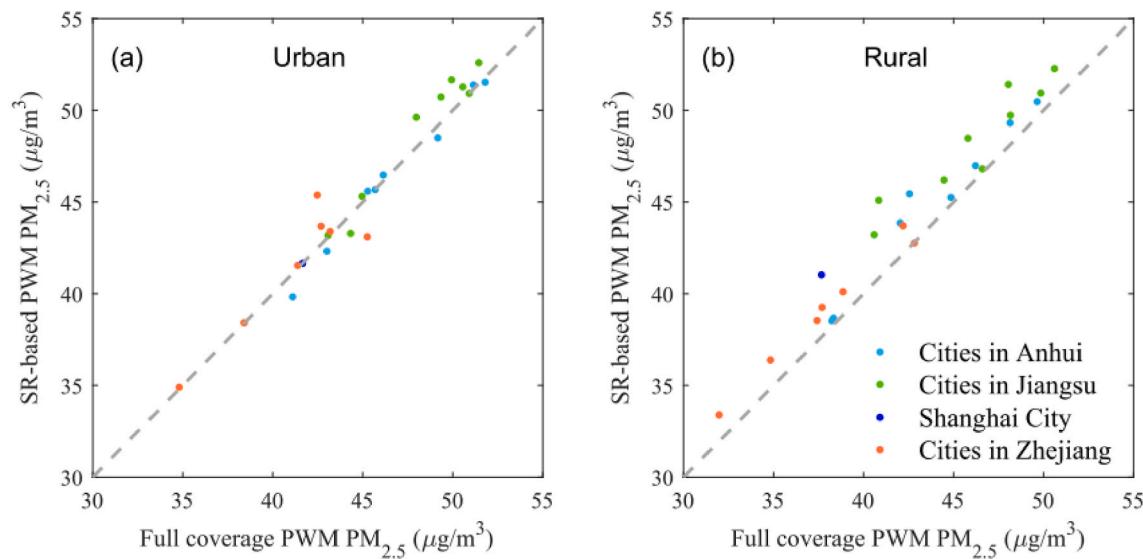
**Fig. 6.** (a) The spatial representativeness (SR) areas of optimized PM<sub>2.5</sub> monitoring stations (triangles) in Yangtze River Delta, and orange (green) areas stand for urban (rural) SR areas. Panel (b) presents the frequency distribution of urban (orange) and rural (green) SR sizes of optimized stations.

For the optimized stations in the entire YRD, there is a smaller urban-rural difference in SR area than that for the current stations. Specifically, the ratio of rural SR area to the entire rural area in YRD (84.49%) is comparable to that for urban SR area (93.93%). Additionally, both about 95% of urban and rural population in YRD are covered by SR areas of the optimized stations. These small urban-rural differences are also observed at city level (not shown).

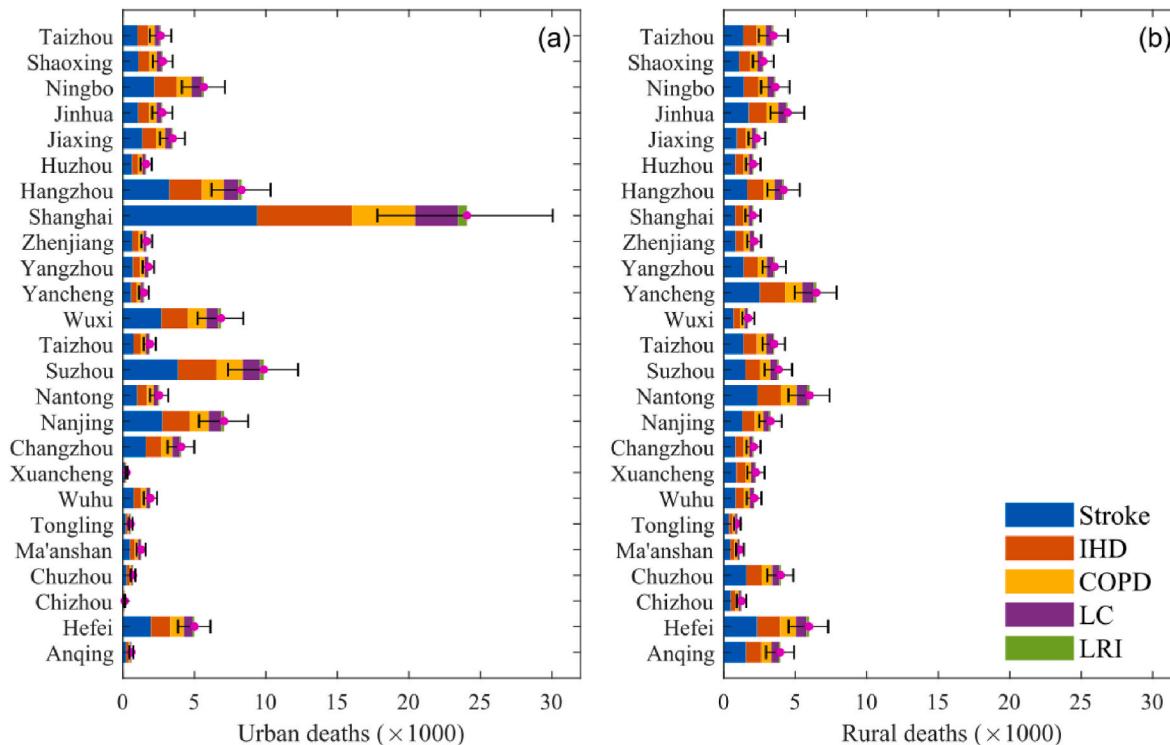
SR-based PWM PM<sub>2.5</sub> is more accurate for most cities after applying the optimized stations than the existing stations, and this improvement is particularly obvious in rural regions (Fig. 4 vs. 7). Based on the optimized stations, the difference in rural PWM PM<sub>2.5</sub> between using SR

estimates and full coverage data is less than 5% for 19 out of 25 cities (Fig. 7b), and the difference in urban PWM PM<sub>2.5</sub> is even less than 1% for about half of the cities (Fig. 7a). Note that both urban and rural PWM PM<sub>2.5</sub> are still generally overestimated even after applying optimized stations (Fig. 7). One possible reason for this is that the optimized stations tend to be located in densely populated areas coupled with heavy pollution emissions (Bi et al., 2019; van Donkelaar et al., 2006).

The improvement in attributable deaths assessment is also observed for most cities by applying the optimized stations (Figs. 5 and 8), especially for cities where population is dominated by rural population (red asterisks in Fig. 5). For instance, for Xuancheng City where about



**Fig. 7.** Comparisons of annual population-weighted mean (PWM) between using SR estimates and using full coverage  $\text{PM}_{2.5}$  data for (a) urban and (b) rural regions. This figure is based on the optimized stations.



**Fig. 8.** Deaths attributable to  $\text{PM}_{2.5}$  exposure by using SR-based PWM  $\text{PM}_{2.5}$  for (a) urban and (b) rural regions in each city in YRD. This figure is based on the optimized monitoring stations.

90% of population is located in rural regions, attributable deaths are overestimated by 5.03% and 8.53% for urban and rural regions respectively when using the existing stations. Correspondingly, after applying the optimized stations, the differences in attributable deaths between using SR-based PWM  $\text{PM}_{2.5}$  and full coverage PWM  $\text{PM}_{2.5}$  are only -1.98% and 0.50% for urban and rural regions, respectively. For the entire urban and rural areas in YRD, these differences are 0.10% and 2.26%.

To comprehensively evaluate the performance of the optimized stations, four factors regarding SR were estimated for each city in YRD, including area ratio, population ratio, redundancy and  $I_{SR}$  (see Section

2.2.3 for details). As shown in Table 2, for the optimized stations, both area ratio and population ratio substantially increase for all cities in YRD compared to that based on the existing stations. Meanwhile, the redundancy of stations decreases for 16 out of 25 cities. In addition, we can see a clear increase in  $I_{SR}$ , except two cities (Changzhou and Nanjing) where  $I_{SR}$  changes little. Overall, the above results suggest that the layout of the optimized stations proposed by this study is reasonable.

## 4. Discussion

### 4.1. Policy recommendations

Several reasons explain the unreasonable layout of the existing PM<sub>2.5</sub> stations in YRD. First, the location of existing stations is unsatisfactory. For about half of cities in YRD, SR areas of stations are not large enough to cover most of the residents. For example, based on the 9 existing stations in Nantong, the total SR area only covers 19.72% of the city area and 34.64% of the city population; after adjusting the locations of these stations, these values increase to 82.90% and 90.81%, respectively. Second, for most cities, there are too many urban stations, while rural stations are severely scarce. This situation may lead to the redundancy of stations in urban regions and the limited rural population covered by SR. Taking Hangzhou as an example, there are 15 urban stations and 2 rural stations. Correspondingly, the redundancy of stations is 11.19% and only 23.85% of the rural population is covered by SR areas.

Given the above-mentioned reasons regarding the unreasonable layout of the existing PM<sub>2.5</sub> stations in YRD, we proposed several policy recommendations. First, for cities with high redundancy of stations (e.g., Tongling), some urban stations with poor SR performance should be eliminated. Second, based on our results of the optimized stations, more rural stations should be deployed, especially for rural population dominated cities, e.g., Chizhou. Third, for some coastal cities (Taizhou and Ningbo), more stations should be deployed to capture large pollutant horizontal gradients influenced by a combination of land-sea breeze recirculation and emissions (Ding et al., 2004; Russo et al., 2016). Finally, it is necessary to evaluate layout of PM<sub>2.5</sub> stations from an SR perspective. This evaluation should systematically consider SR area, population coverage of SR and redundancy of stations. The study provided an alternative metric for policymakers to achieve these goals.

### 4.2. Limitations

A limitation of this study is the lack of comparison with previous results of SR. Our estimates of SR sizes vary greatly with stations, ranging from 2 to 3624 km<sup>2</sup> for the existing stations. Previous studies have rarely reported SR of PM<sub>2.5</sub> monitoring stations in YRD. Given that SR strongly depends on local meteorology and emissions (Righini et al., 2014; Santiago et al., 2013; Vardoulakis et al., 2005), it is not appropriate to directly compare SR studies across different regions. Nevertheless, previous studies can still provide some useful references for SR. For example, Piersanti et al. (2015) reported an SR size of 224 km<sup>2</sup> for a station located near coastline of Italy, and this estimate was much smaller than another one (800 km<sup>2</sup>) for an inland station. This coast-inland difference is also observed in our study (not shown). Additionally, Martin et al. (2014) found that the SR sizes ranging from 75 to 300 grid cells (0.1° × 0.1°) were more frequent for the Spanish PM<sub>10</sub> rural background stations. Their estimates are much higher than our results of the optimized stations, most of which located in rural regions. Different from the previous studies above using model dataset, Shi et al. (2018) used a high-resolution network observation over an urban city in North China, and found that SR sizes of PM<sub>2.5</sub> stations range from 0.25 to 16.25 km<sup>2</sup>. Their estimates, however, are likely to be underestimated due to the sparse distribution of stations. Further work is needed to confirm and validate our SR estimates, perhaps in combination with model simulations.

Another limitation of this work is that we only focus on PM<sub>2.5</sub> when optimizing the current stations. A monitoring station usually observes multiple pollutants, such as ozone and oxides of sulfur and nitrogen. Thus, optimizing the network stations should consider the discrepancy in SR among different pollutants, which is needed to further investigate in the future. Additionally, for urban-rural difference in health assessment, we did not consider the difference in age structure between urban and rural areas. We also did not consider PM<sub>2.5</sub> exposure due to different emission sources, such as rural population exposure to agricultural

chemicals. It would be useful to explore these limitations in the future.

Additionally, the optimized stations may fail to meet the site selection principles and distribution requirements according to HJ 664–2013 (MEPC, 2013). For example, one of the principles for the monitoring site layout is the stability of site location. But we relocated all the previous sites in the optimized scheme. Additionally, the site selection should consider the monitoring site category according to HJ 664–2013 (MEPC, 2013). However, data on specific categories of land use (e.g., industrial, traffic, and port categories) are hard to obtain for the entire YRD region. Future work is needed to improve the optimized scheme by considering the category of sites and the stability of site location.

## 5. Conclusions

In this study, based on daily 1-km-resolution PM<sub>2.5</sub> data, we estimated spatial representativeness (SR) of the current PM<sub>2.5</sub> monitoring stations in the Yangtze River Delta (YRD). We also proposed an optimization approach to adjust the current network stations and compared the SR performance between the current stations and the optimized stations. Our results show that 68.53% of urban area and 79.63% of urban population are covered by SR areas of the current stations for the entire YRD region; these ratios are only 25.82% and 40.50% for rural area and rural population, respectively. After optimizing the layout of the current stations, the number of stations decreases from 210 to 156 while SR area increases nearly twice for the entire YRD region. Moreover, SR areas of the optimized stations cover about 95% of urban and rural population in YRD. The optimization strategy of this study, however, only focused on PM<sub>2.5</sub>. Other pollutants, such as ozone, should be considered for air quality evaluation and management. Thus, future work on optimizing the network stations would need to consider SR estimates among different pollutants.

## Credit author statements

Heming Bai: Formal analysis, Writing – original draft, Writing – review & editing. Wenkang Gao: Resources, Data curation. Myeongsu Seong: Writing – review & editing. Rusha Yan: Investigation, Methodology. Jing Wei: Data curation. Chong Liu: Data curation.

## Acknowledgments

This work was supported by the National Natural Science Foundation of China (No. 42205180), the National Key Research and Development Program of China (No. 2022YFC3703003), the Natural Science Research Program for Higher Education in Jiangsu Province (No. 22KJB170019), and the Nantong Basic Science Research Program (No. JC12022074).

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.apgeog.2023.102949>.

## References

- Bai, H., Gao, W., Zhang, Y., & Wang, L. (2022). Assessment of health benefit of PM<sub>2.5</sub> reduction during COVID-19 lockdown in China and separating contributions from anthropogenic emissions and meteorology. *Journal of Environmental Sciences*, 115, 422–431. <https://doi.org/10.1016/j.jes.2021.01.022>
- Bai, Z., Han, J., & Azzi, M. (2017). Insights into measurements of ambient air PM 2.5 in China. *Trends in Environmental Analytical Chemistry*, 13, 1–9. <https://doi.org/10.1016/j.teac.2017.01.001>
- Bai, H., Yan, R., Gao, W., Wei, J., & Seong, M. (2022). Spatial representativeness of PM<sub>2.5</sub> monitoring stations and its implication for health assessment. *Air Quality, Atmosphere & Health*, 15(9), 1571–1581. <https://doi.org/10.1007/s11869-022-01202-2>
- Bai, H., Zheng, Z., Zhang, Y., Huang, H., & Wang, L. (2021). Comparison of satellite-based PM<sub>2.5</sub> estimation from aerosol optical depth and top-of-atmosphere reflectance. *Aerosol and Air Quality Research*, 21(2), Article 200257. <https://doi.org/10.4209/aaqr.2020.05.0257>

- 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. *Remote Sensing of Environment*, 221(October 2018), 665–674. <https://doi.org/10.1016/j.rse.2018.12.002>
- Chen, K., Zhou, L., Chen, X., Bi, J., & Kinney, P. L. (2017). Acute effect of ozone exposure on daily mortality in seven cities of Jiangsu province, China: No clear evidence for threshold. *Environmental Research*, 155(January), 235–241. <https://doi.org/10.1016/j.envres.2017.02.009>
- Costabile, F., Desantis, F., Hong, W., Liu, F., Salvatori, R., Wang, F., & Allegrini, I. (2006). Representativeness of urban highest polluted zones for siting traffic-oriented air monitoring stations in a Chinese city. *JSME International Journal Series B*, 49(1), 35–41. <https://doi.org/10.1299/jsmbe.49.35>
- Ding, A., Wang, T., Zhao, M., Wang, T., & Li, Z. (2004). Simulation of sea-land breezes and a discussion of their implications on the transport of air pollution during a multi-day ozone episode in the Pearl River Delta of China. *Atmospheric Environment*, 38 (39), 6737–6750. <https://doi.org/10.1016/j.atmosenv.2004.09.017>
- van Donkelaar, A., Martin, R. V., & Park, R. J. (2006). Estimating ground-level PM2.5 using aerosol optical depth determined from satellite remote sensing. *Journal of Geophysical Research: Atmospheres*, 111(21), Article D21201. <https://doi.org/10.1029/2005JD006996>
- Doxsey-Whitfield, E., MacManus, K., Adamo, S. B., Pistolesi, L., Squires, J., Borkovska, O., & Baptista, S. R. (2015). Taking advantage of the improved availability of census data: A first look at the gridded population of the World, version 4. *Papers in Applied Geography*, 1(3), 226–234. <https://doi.org/10.1080/23754931.2015.1014272>
- Elbern, H., Strunk, A., Schmidt, H., & Talagrand, O. (2007). Emission rate and chemical state estimation by 4-dimensional variational inversion. *Atmospheric Chemistry and Physics*, 7(14), 3749–3769. <https://doi.org/10.5194/acp-7-3749-2007>
- Gao, L., Yue, X., Meng, X., Du, L., Lei, Y., Tian, C., & Qiu, L. (2020). Comparison of ozone and PM2.5 concentrations over urban, suburban, and background sites in China. *Advances in Atmospheric Sciences*, 37(12), 1297–1309. <https://doi.org/10.1007/s00376-020-0054-2>
- Gaughan, A. E., Stevens, F. R., Huang, Z., Nieves, J. J., Sorichetta, A., Lai, S., Ye, X., Linard, C., Hornby, G. M., Hay, S. I., Yu, H., & Tatem, A. J. (2016). Spatiotemporal patterns of population in mainland China, 1990 to 2010. *Scientific Data*, 3(1), Article 160005. <https://doi.org/10.1038/sdata.2016.5>
- GBD 2019 Risk Factors Collaborators. (2020). Global burden of 87 risk factors in 204 countries and territories, 1990–2019: A systematic analysis for the global burden of disease study 2019. *The Lancet*, 396(10258), 1223–1249. [https://doi.org/10.1016/S0140-6736\(20\)30752-2](https://doi.org/10.1016/S0140-6736(20)30752-2)
- Gong, P., Li, X., Wang, J., Bai, Y., Chen, B., Hu, T., Liu, X., Xu, B., Yang, J., Zhang, W., & Zhou, Y. (2020). Annual maps of global artificial impervious area (GAIA) between 1985 and 2018. *Remote Sensing of Environment*, 236(October 2019), Article 111510. <https://doi.org/10.1016/j.rse.2019.111510>
- Guan, Y., Xiao, Y., Wang, Y., Zhang, N., & Chu, C. (2021). Assessing the health impacts attributable to PM2.5 and ozone pollution in 338 Chinese cities from 2015 to 2020. *Environmental Pollution*, 287(February), Article 117623. <https://doi.org/10.1016/j.envpol.2021.117623>
- Hao, Y., & Xie, S. (2018). Optimal redistribution of an urban air quality monitoring network using atmospheric dispersion model and genetic algorithm. *Atmospheric Environment*, 177(January), 222–233. <https://doi.org/10.1016/j.atmosenv.2018.01.011>
- Hohenberger, T. L., Che, W., Fung, J. C. H., & Lau, A. K. H. (2021). A proposed population-health based metric for evaluating representativeness of air quality monitoring in cities: Using Hong Kong as a demonstration. *PLoS One*, 16(5), Article e0252290. <https://doi.org/10.1371/journal.pone.0252290>
- Ikeda, S., Nakamori, Y., & Sawaragi, Y. (1981). Optimal sensor allocation in air monitoring network with spatial representativeness—design of no x monitoring system in kyoto city. *IFAC Proceedings Volumes*, 14(2), 3799–3804. [https://doi.org/10.1016/S1474-6670\(17\)64040-0](https://doi.org/10.1016/S1474-6670(17)64040-0)
- Li, X., Gong, P., Zhou, Y., Wang, J., Bai, Y., Chen, B., Hu, T., Xiao, Y., Xu, B., Yang, J., Liu, X., Cai, W., Huang, H., Wu, T., Wang, X., Lin, P., Li, X., Chen, J., He, C., ... Zhu, Z. (2020). Mapping global urban boundaries from the global artificial impervious area (GAIA) data. *Environmental Research Letters*, 15(9), Article 094044. <https://doi.org/10.1088/1748-9326/ab9be3>
- Lv, B., Liu, Y., Yu, P., Zhang, B., & Bai, Y. (2015). Characterizations of PM2.5 pollution pathways and sources analysis in four large cities in China. *Aerosol and Air Quality Research*, 15(5), 1836–1843. <https://doi.org/10.4209/aaqr.2015.04.0266>
- Ma, M., Chen, Y., Ding, F., Pu, Z., & Liang, X. (2019). The representativeness of air quality monitoring sites in the urban areas of a mountainous city. *Journal of Meteorological Research*, 33(2), 236–250. <https://doi.org/10.1007/s13351-019-8145-7>
- Martin, F., Fileni, L., Palomino, I., Vivanco, M. G., & Garrido, J. L. (2014). Analysis of the spatial representativeness of rural background monitoring stations in Spain. *Atmospheric Pollution Research*, 5(4), 779–788. <https://doi.org/10.5094/APR.2014.087>
- McDuffie, E. E., Martin, R. V., Spadaro, J. V., Burnett, R., Smith, S. J., O'Rourke, P., Hammer, M. S., van Donkelaar, A., Bindle, L., Shah, V., Jaegle, L., Luo, G., Yu, F., Adeniran, J. A., Lin, J., & Brauer, M. (2021). Source sector and fuel contributions to ambient PM2.5 and attributable mortality across multiple spatial scales. *Nature Communications*, 12(1), 3594. <https://doi.org/10.1038/s41467-021-23853-y>
- MEPC (Ministry of Environment Protection of China). (2013). *Technical regulation for selection of ambient air quality monitoring stations (on trial)*. HJ 664-2013.
- Piersanti, A., Vitali, L., Righini, G., Cremona, G., & Giancarella, L. (2015). Spatial representativeness of air quality monitoring stations: A grid model based approach. *Atmospheric Pollution Research*, 6(6), 953–960. <https://doi.org/10.1016/j.apr.2015.04.005>
- Righini, G., Cappelletti, A., Ciucci, A., Cremona, G., Piersanti, A., Vitali, L., & Giancarella, L. (2014). GIS based assessment of the spatial representativeness of air quality monitoring stations using pollutant emissions data. *Atmospheric Environment*, 97, 121–129. <https://doi.org/10.1016/j.atmosenv.2014.08.015>
- Russo, A., Gouveia, C., Levy, I., Dayan, U., Jerez, S., Mendes, M., & Trigo, R. (2016). Coastal recirculation potential affecting air pollutants in Portugal: The role of circulation weather types. *Atmospheric Environment*, 135, 9–19. <https://doi.org/10.1016/j.atmosenv.2016.03.039>
- Santiago, J. L., Martin, F., & Martilli, A. (2013). A computational fluid dynamic modelling approach to assess the representativeness of urban monitoring stations. *The Science of the Total Environment*, 454–455, 61–72. <https://doi.org/10.1016/j.scitotenv.2013.02.068>
- Shi, W., Bi, J., Liu, R., Liu, M., & Ma, Z. (2021). Decrease in the chronic health effects from PM2.5 during the 13th Five-Year Plan in China: Impacts of air pollution control policies. *Journal of Cleaner Production*, 317(June), Article 128433. <https://doi.org/10.1016/j.jclepro.2021.128433>
- Shi, X., Zhao, C., Jiang, J. H., Wang, C., Yang, X., & Yung, Y. L. (2018). Spatial representativeness of PM 2.5 concentrations obtained using observations from network stations. *Journal of Geophysical Research: Atmospheres*, 123(6), 3145–3158. <https://doi.org/10.1002/2017JD027913>
- Song, C., He, J., Wu, L., Jin, T., Chen, X., Li, R., Ren, P., Zhang, L., & Mao, H. (2017). Health burden attributable to ambient PM2.5 in China. *Environmental Pollution*, 223, 575–586. <https://doi.org/10.1016/j.envpol.2017.01.060>
- Vardoulakis, S., Gonzalez-Flesca, N., Fisher, B. E. A., & Pericleous, K. (2005). Spatial variability of air pollution in the vicinity of a permanent monitoring station in central Paris. *Atmospheric Environment*, 39(15), 2725–2736. <https://doi.org/10.1016/j.atmosenv.2004.05.067>
- Wei, J., Li, Z., Cribb, M., Huang, W., Xue, W., Sun, L., Guo, J., Peng, Y., Li, J., Lyapustin, A., Liu, L., Wu, H., & Song, Y. (2020). Improved 1 km resolution PM2.5 estimates across China using enhanced space-time extremely randomized trees. *Atmospheric Chemistry and Physics*, 20(6), 3273–3289. <https://doi.org/10.5194/acp-20-3273-2020>
- 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. *Remote Sensing of Environment*, 252(January 2020), Article 112136. <https://doi.org/10.1016/j.rse.2020.112136>
- Wei, J., Li, Z., Wang, J., Li, C., Gupta, P., & Cribb, M. (2023). Ground-level gaseous pollutants ( $\text{NO}_2$ ,  $\text{SO}_2$ , and  $\text{CO}$ ) in China: daily seamless mapping and spatiotemporal variations. *Atmospheric Chemistry and Physics*, 23, 1511–1532. <https://doi.org/10.5194/acp-23-1511-2023>
- Yang, W., Yuan, G., & Han, J. (2019). Is China's air pollution control policy effective? Evidence from Yangtze River Delta cities. *Journal of Cleaner Production*, 220, 110–133. <https://doi.org/10.1016/j.jclepro.2019.01.287>
- Yu, T., Wang, W., Ciren, P., & Sun, R. (2018). An assessment of air-quality monitoring station locations based on satellite observations. *International Journal of Remote Sensing*, 39(20), 6463–6478. <https://doi.org/10.1080/01431161.2018.1460505>
- Zhai, S., Jacob, D. J., Wang, X., Shen, L., Li, K., Zhang, Y., Gui, K., Zhao, T., & Liao, H. (2019). Fine particulate matter (PM2.5) trends in China, 2013–2018: Separating contributions from anthropogenic emissions and meteorology. *Atmospheric Chemistry and Physics*, 19(16), 11031–11041. <https://doi.org/10.5194/acp-19-11031-2019>