CHANGES OF SPATIAL DISPARITIES IN SATELLITE-DERIVED PM$_{2.5}$ EXPOSURES OVER 2010-2019 IN CHINA

Ming Liu$^{1,2}$, Gaoxiang Zhou$^1$, Ziyan Liu$^1$, Ling Han$^{1,2}$, Jonathan Li$^4$

$^1$ School of Land Engineering, Chang’an University, Xi'an, Shaanxi, 710064, China - (mingliu.2018901475, hanling) @chd.edu.cn
$^2$ Xi'an Key Laboratory of Territorial Spatial Information, Chang’an University, Xi'an, Shaanxi, 710064, China
$^3$ School of Human Settlements and Civil Engineering, Xi'an Jiaotong University, Xi'an, 710049, China - zhougaoxiang@xjtu.edu.cn
$^4$ Department of Geography and Environmental Management, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada - junli@uwwaterloo.ca

Commission III, WG III/8

KEY WORDS: PM$_{2.5}$, spatiotemporal disparity, satellite remote sensing, exposure, environmental justice, air pollution, China.

ABSTRACT:

Air pollution, especially fine particulate matter (PM$_{2.5}$), has attracted extensive attention due to its adverse impacts on public health. Although PM$_{2.5}$ pollution was significantly reduced in China over time, while little is known how the spatial disparity of PM$_{2.5}$ exposure has evolved, especially from both absolute and relative perspectives. Here, we estimate the long-term PM$_{2.5}$ exposures in China based on satellite observations and convolutional neural network, and characterize the spatial disparity of PM$_{2.5}$ exposure using Theil index and rank-rank relationship. The result shows that both PM$_{2.5}$ exposure and absolute spatial disparity were substantially reduced between 2010 and 2019. The nation-wide concentrations (Theil index) declined from 48.0μg/m$^3$ (0.13) to 35.5μg/m$^3$ (0.054). The inter-provincial disparities dominate the overall disparity in 2010, while the intra-provincial disparity contributed the most in 2019. However, while absolute disparities have diminished, relative disparities persist. PM$_{2.5}$ exposures in the least 20% percentile polluted cities have increased over time, while exposures in other regions declined. On average, the more (less) polluted cities in 2010 were still the more (less) polluted cities in 2019 (except for the very most 2 percentile polluted cities), indicating that the population in more polluted cities still experiences more air pollution than others. Spatial pattern of relative disparity changes was also observed. Overall, understanding not only absolute spatial disparity but also relative disparity is required to help formulate targeted policies for an equitable environment, leaving nobody behind.

1. INTRODUCTION

Suspended particulate matter with aerodynamic diameters less than 2.5 μm (PM$_{2.5}$) exposure can increase the risk of death from respiratory disease, cardiopulmonary disease, and lung cancer, as these small particles can penetrate the respiratory tract, alveoli and even blood stream. The Global Burden of Disease (GBD) study reported that PM$_{2.5}$ is the fourth leading risk factor for premature death in China, contributing to 11.1% of deaths in 2016 (Cohen et al., 2017; Naghavi et al., 2017).

Satellite remote sensing offers an effective approach to estimate large-scale ambient PM$_{2.5}$ concentrations. Various empirical models have been established based on satellite observations, including regression models (such as land use regression, geographical weighted regression and generalized additive models) and machine learning models (such as random forest and deep belief network) (Chen et al., 2018; Knibbs et al., 2018; Li et al., 2017; Liu et al., 2019b; Park et al., 2020; Xiao et al., 2018; Xue et al., 2020; Yan et al., 2021). These satellite-based models performed well in estimating PM$_{2.5}$ concentrations, while “concentration” is not the best indicator to characterize the level of threat to public health. For example, although PM$_{2.5}$ concentrations in the Taklamakan Desert are much higher than in other regions, the corresponding population exposure are still at a low level due to the low population density. Therefore, given the vast size and uneven population distribution of China, quantifying and analysing large-scale PM$_{2.5}$ exposure (rather than "concentration") is fulfilling for applying remote sensing to public health.

PM$_{2.5}$ pollution is known to be heterogeneous because of disparities in climate, geographical conditions, population density, industrial and economic activities, posing environmental justice challenges (Han et al., 2021; He and Huang, 2018a; Liu et al., 2021). Although related scholarship provided long-term and large-scale PM$_{2.5}$ concentrations (He et al., 2021; Liu et al., 2019a; Xue et al., 2020; Zheng et al., 2015), we have limited information on how magnitude of spatial variation of PM$_{2.5}$ exposure has evolved in China (especially over recent decades that saw tremendous change in economic output and air pollution). As equity and justice are inherently comparative, not only absolute changes of disparity, but also relative changes deserve to explore to advance the environmental justice, leaving nobody behind (Colmer et al., 2020).

Chinese government released the Ambient Air Quality Standards (GB3095-2012) in 2012, which updated the concentration limits of various air pollutants and included PM$_{2.5}$ in the standards for the first time. The Action Plan for the Prevention and Control of Air Pollution was then issued in 2013, which aims to reduce air pollution by optimizing industrial and energy structures, expanding air quality ground monitoring network, and establishing targeted emission reduction and regional collaborative control strategies. Studies documented that air quality in China was significant improved since 2013 (Li et al., 2020; Ma et al., 2019; Wei et al., 2021). Despite this, the 14th

* Corresponding author: Liu Ming (Email address: mingliu@chd.edu.cn)
Five-Year Plan in 2021 still involves PM$_{2.5}$ concentration as an important binding indicator to evaluate the effectiveness of air pollution policies, demonstrating the country has been attaching great importance to the prevention and control of air pollution.

Therefore, the decade of 2010-2019 were adopted as study period to verify the effectiveness of air pollution control policies by evaluating the changes of spatial disparity in PM$_{2.5}$ exposure, seeking to help answer the following questions: (1) whether spatial disparity of exposure have diminished over time or persist? (2) which spatial dimension contributes the most to the spatial disparities? (3) Have the most polluted regions changed in China change over time? To do this, we first estimated satellite-derived PM$_{2.5}$ exposures using convolutional neural network, and then quantified the spatial disparities of PM$_{2.5}$ exposures between 2010 and 2019 using the Theil decomposable method and rank-rank relationship from both absolute and relative perspectives.

2. DATA COLLECTION

2.1 Ground-level PM$_{2.5}$ observations

The study domain in this study is defined as the Mainland China. The measured PM$_{2.5}$ surface concentration data obtained from the National Environmental Monitoring Centre (http://www.cnemc.cn/), with more than 1,450 monitoring stations available. The daily concentration values were averaged from hourly observations for each station.

2.2 Explanatory variables

We adopted Moderate Resolution Imaging Spectroradiometer (MODIS) Collection 6 aerosol optical depth (AOD) from Terra and Aqua as part of our predictors for PM$_{2.5}$ estimation. Products from both “Dark Target (DT)” and the “Deep Blue (DB)” were used to fill the missing AOD values using inverse variance weighting method, which were also calibrated by the AERONET level 2.0 AOD measurements.

Additional variables involve meteorological data, surface information and categorical variables. Specifically, meteorological data include relative humidity, air temperature, wind speed, precipitation, visibility, surface pressure, and planetary boundary layer height, downloaded from NCEP GDAS/FNL 0.25 Degree Global Tropospheric Analyses and Forecast Grids (https://rda.ucar.edu/datasets/ds083.3/) and NCEP ADP Global Surface Observational Weather Data (http://rda.ucar.edu/datasets/ds461.0/); surface information includes plant cover quantified by normalized difference vegetation index (NDVI), and digital earth model (DEM) from Resources and Environmental Science Data Centre; categorical variables include coordinates and day of year reflecting spatial and temporal variations. All explanatory variables were unified with spatial resolution of 1 km and assigned to each PM$_{2.5}$ monitoring stations using nearest neighbour algorithm for modelling.

2.3 Population data

The national population gridded data were obtained from LandScan Global Vital Statistics database (https://landscan.orl.gov), with a resolution of 1 km (Bright et al., 2011; Rose et al., 2020). This dataset integrates census data and remote sensing information (such land cover and nighttime lights) to grid the population data, which has been widely used in various fields of research. We also calibrate the gridded population data based sixth and seventh census data using the least squares method.

3. METHODS

3.1 Exposure estimation

Considering the spatial autocorrelation of the PM$_{2.5}$ and its predictors, the convolutional neural networks (CNN) was adopted in this study to estimate PM$_{2.5}$ exposures, which shows good performances on spatial feature extraction nonlinear relationship characterization. As shown in Figure 1, PM$_{2.5}$ concentration at a given pixel is estimated by all the predictors of adjacent N*N grids centred around this pixel. For each given N*N input, CNN conducted by stacking a convolution layer and a fully-connected (FC) layer. A convolution layer consists of multiple convolution filters and an activation function. The convolutional filters run over the sample and calculate inner products, allowing adjacent information to be extracted into features. Each filter extracts distinct features from the input. We adopted 48 filters with dimension of $3^2$. The activation function (ReLU in this study) allows the model to learn complex nonlinear relationship between the inputs and outputs. The zero-padding method was adopted to remain the dimension unchanged. The FC layer with 128 nodes was embedded after convolutional layers to estimate PM$_{2.5}$ at the final stage. The dropout was used in our study to prevent overfitting. The standard mean-squared error (MSE) between the observed and estimated PM$_{2.5}$ concentrations were employed as the loss function to test the model accuracy.

![Figure 1. The structure of a CNN model for PM$_{2.5}$ estimation](image)

The hyperparameters in our model, such as the number of filters, layers and nodes of FC layers, were determined by the experiment according to the model accuracy. The estimation model was then established after the convolutional filters and weights of FC layers were learned by the standard gradient descent algorithm based on training dataset. Considering the limited number of measurements before 2013, two models (i.e. annual and daily model) were employed to estimate annual concentrations in 2010 and 2019. The annual concentrations after 2013 were averaged by daily PM$_{2.5}$ estimates using the model trained by daily-mean inputs in the corresponding year. The concentrations before 2013 (including 2013) were estimated using the model trained by annual-mean inputs. The model performance was validated by 10-fold cross-validation (CV) and test data using determination coefficient ($R^2$), root mean square error (RMSE) and mean predictive error (MPE).
Integrating ambient estimated PM$_{2.5}$ concentrations with population, the PM$_{2.5}$ exposures can be calculated by the following equations:

$$\text{Expo} = \frac{\sum_{i=1}^{n_i} (P_{Mi} \times Pop_i)}{\sum_{i=1}^{n_i} \text{Pop}_i} \quad (2)$$

where Expo refers to PM$_{2.5}$ exposure; $P_{Mi}$ refers to ambient PM$_{2.5}$ concentrations at grid $i$; Pop$_i$ refers to the number of grids within the city corresponding to grid $i$.

### 3.2 Absolute disparity assessment

Theil Index was adopted to quantify the spatial disparity, which is a statistical measure of inequality proposed by econometrician Henry Theil (Theil, 1972) based on the concept of information entropy in information theory. It has been used to measure regional and social differences (Azimi et al., 2018; Kang et al., 2019; Sitepu et al., 2018). Theil index can be decomposed to measure the contribution of intra-group and inter-group differences to the total difference. Therefore, we divided the whole country into seven regions, i.e. East China, Central China, South China, North China, Northwest China, Southwest China and Northeast China, to calculate the overall differences of PM$_{2.5}$ exposure among regions, among provinces and within provinces. The greater the Theil index, the greater the regional variation. The calculation formula is as follows:

$$T = \frac{1}{n} \sum_{i=1}^{n} E_i \ln \left( \frac{E_i}{E} \right) = T_R + T_P + T_C \quad (3)$$

$$T_R = \frac{1}{n} \sum_{k=1}^{7} \left( \frac{E_k}{E} \right) \ln \left( \frac{E_k}{E} \right) \quad (4)$$

$$T_P = \frac{1}{n} \sum_{k=1}^{7} \sum_{j=1}^{n_k} \left( \frac{E_{ij}}{E_k} \right) \ln \left( \frac{E_{ij}}{E_k} \right) \quad (5)$$

$$T_C = \frac{1}{n} \sum_{k=1}^{7} \sum_{j=1}^{n_k} \left( \frac{E_{ij}}{E} \right) \ln \left( \frac{E_{ij}}{E} \right) \quad (6)$$

where $T$ refers to the national Theil index, quantifying the spatial disparity across China; $T_R$ refers to the Theil index among regions; $T_P$ refers to the Theil index among provinces; $T_C$ refers to the Theil index within provinces (namely among cities); $E_k$ refers to the average exposure in region $k$; $E_{ij}$ refers to the average exposure in province $j$ within region $k$; $E_{ij}$ refers to the average exposure in city $i$ within province $j$; $n_k$ refers to the number of cities in the corresponding region; $n_{ij}$ refers to the number of cities in the corresponding province.

### 3.3 Relative disparity assessment

We used rank-rank relationship to characterize the relative changes of spatial disparity (Colmer et al., 2020). We build on this concept to explore whether the population living in the most/least polluted regions still exposed to the most/least severe PM$_{2.5}$ pollution over 2010-2019? To do this, the percentile rank (PR) of PM$_{2.5}$ exposure in each city were first calculated in 2010 and 2019. The higher the rank (one is the highest rank), the lower the PM$_{2.5}$ exposure, the better the air quality. The rank-rank relationship was then plotted with x-axis of rank in 2010 and y-axis of rank in 2019. The 1:1 line indicates there was no change in PM$_{2.5}$ rank, i.e. cities at the 10$^{th}$ percentile of PM$_{2.5}$ distribution in 2010 are also at the 10$^{th}$ percentile of PM$_{2.5}$ distribution in 2019. The larger the deviations from the 1:1 line, the bigger the average changes in PM$_{2.5}$ rank between 2010 and 2019. Points above the line represent the rank in 2019 lower than 2010, indicating that air quality decreased relatively compared with other cities. Points below the line represent rank increased from 2010 to 2019, indicating the improvements in relative air quality.

### 4. RESULTS AND DISCUSSION

#### 4.1 Patterns of PM$_{2.5}$ exposure

<table>
<thead>
<tr>
<th></th>
<th>$R^2$</th>
<th>RMSE ($\mu g/m^3$)</th>
<th>MPE ($\mu g/m^3$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Daily</td>
<td>0.82</td>
<td>13.8</td>
<td>10.0</td>
</tr>
<tr>
<td>Model</td>
<td>0.84</td>
<td>12.2</td>
<td>9.1</td>
</tr>
<tr>
<td>Test accuracy</td>
<td>0.62</td>
<td>20.0</td>
<td>16.4</td>
</tr>
<tr>
<td>Annual</td>
<td>0.83</td>
<td>12.5</td>
<td>9.8</td>
</tr>
<tr>
<td>Model</td>
<td>0.85</td>
<td>10.7</td>
<td>8.4</td>
</tr>
<tr>
<td>Test accuracy</td>
<td>0.65</td>
<td>17.9</td>
<td>14.0</td>
</tr>
</tbody>
</table>

Table 1. Performance of PM$_{2.5}$ estimation models

The PM$_{2.5}$ estimates were evaluated using 10-fold CV method. Table 1 shows the average accuracy of the CNN-based estimation models in various years. The annual models performed better than daily models with CV $R^2$ more than 0.8. We also validated daily and annual model using test data (which is independent with training dataset) to reveal the predictive capability of the models, with $R^2 > 0.6$ and RMSE $< 20 \mu g/m^3$. The result demonstrates a comparable accuracy with other studies(He and Huang, 2018b; Ma et al., 2016; You et al., 2016), suggesting the effectiveness of our estimation method.

Figure 2. PM$_{2.5}$ concentration and population exposure distribution in China in 2010 and 2019

The spatial distribution of PM$_{2.5}$ concentration and population exposure in China in 2010 and 2019 are shown in Figure 2. As can be seen in 2010, PM$_{2.5}$ concentration and population exposure in the east were higher than values in the west, and the dividing line was consistent with Heihe-Tengchong line. The PM$_{2.5}$ hotspots in eastern China are distributed in North China Plain, Sichuan Basin and Fenwei Plain. The high values were prevalent in the regions with intense population and anthropogenic activities (such as in the North China Plain, Hunan and Hubei region), fossil fuel combustion associated with rapid industrialization (such as in the Fenwei Plain where possess abundant coal-fired facilities and coal industry), and unfavourable meteorological and topographic conditions (such as
Sichuan Basin where surrounded by mountains with low altitude). Different from eastern China, the high PM$_{2.5}$ concentration in Tarim Basin in the northwest is mainly affected by natural factors such as sand and dust. After the State Council issued the “Action Plan of Air Pollution Prevention and Control” measures, air quality has been significantly improved across the country, especially in eastern China where industrial emissions and human activities dominate. PM$_{2.5}$ exposures reduced from 47.99μg/m$^3$ to 35.52μg/m$^3$ by 2019, with a decrease of 38.2% in central China and 31.5% in southwest China.

4.2 Changes of absolute spatial disparity

![Figure 3. Changes in PM$_{2.5}$ exposures from 2010 to 2019](image)

We compute the average, 10$^{th}$ and 90$^{th}$ percentiles of PM$_{2.5}$ provincial exposures using demographic data to accurately reflect the variations of exposures. As shown in Figure 3, not only the substantial reduction in PM$_{2.5}$ exposures, but also the gap between 10$^{th}$ and 90$^{th}$ percentiles have reduced. The difference between 10$^{th}$ and 90$^{th}$ percentiles was 22.6 μg/m$^3$ in 2019, decreasing from 51.4 μg/m$^3$ in 2010. The narrowed gap in PM$_{2.5}$ exposures indicate the absolute disparity in China has been diminished between 2010 and 2019, which was owing to the general improvement of air quality.

![Figure 4. Contributions of spatial disparities in PM$_{2.5}$ exposure from three dimensions](image)

Despite this, we still need more information on where the spatial disparity is from to help formulate targeted policies for an equitable environment. Therefore, the absolute disparity of PM$_{2.5}$ population exposure was quantified and decomposed using the Theil index, which allows the overall disparity to be decomposed into inter-regional differences, inter-provincial and intra-provincial differences in China. The results in Figure 4 show that the spatial disparity of PM$_{2.5}$ exposure was significantly decreased between 2010 and 2019, with the Theil index decreasing from 0.132 in 2010 to 0.054 in 2019. We found that the regional differences in PM$_{2.5}$ exposure in 2010 were dominated by inter-provincial differences with a contribution rate of 38.9%. The contribution rates of inter-regional and intra-provincial differences were similar, with respective value of 31.1% and 30.0%. Unlike patterns in 2010, the spatial disparity in 2019 were mainly contributed by intra-provincial differences, followed by inter-provincial and inter-regional differences. The results demonstrate that Chinese population not only were exposed to lower PM$_{2.5}$ concentrations, but also share a more equitable atmospheric environment after the implementation of air pollution control policies.

4.3 Changes of relative spatial disparity

![Figure 5. Rank-rank relationship between PM$_{2.5}$ percentile rank in 2010 and in 2019](image)

Though absolute spatial disparity in China was diminished from 2010 to 2019, as illustrated in Figures 3-4, how magnitude the relative disparity of exposure has changed over time remain obscure. Hence, to explore whether population in the most or least polluted regions was still exposed to the most or least severe pollution over time, we assigned each city to a percentile following the exposure distribution in 2010 and computed the corresponding mean PR in 2019. The correlation coefficient between PR in 2010 and 2019 is 0.96, demonstrating PR in 2010 can explain 92.4 % of PR in 2019. As shown in Figure 5, PM$_{2.5}$ ranks in 2010 were similar with the ranks in 2019 as 80% of the percentile pairs were distributed within the ±10% range of the 1:1 line, indicating that the most/least polluted cities in 2010 were still the most/least polluted cities on average in 2019. For the points that fall outside the ±10% range, we observed that the points below the range dominate and are distributed at the lower ranks, indicating relative air quality improvements in these cities, especially in the very most 2 percentile polluted cities (99$^{th}$ and 100$^{th}$).

![Figure 6. Changes in PM$_{2.5}$ exposure between 2010 and 2019 for each percentile bin in 2010](image)
Therefore, it is plausible to suppose that the persistence in relative spatial disparity is that PM2.5 reductions were proportional across most cities. The changes of PM2.5 exposure was plotted in Figure 6. We observed that PM2.5 in less polluted cities (below 20th percentile of PM2.5 distribution) in 2010 increased, while air quality in more polluted regions were getting better. While the baseline values in these regions are higher, the blue line in Figure 6 illustrates that PM2.5 exposure has fallen more in the more polluted regions in 2010. Despite this, the trend of purple line illustrates that the percentage changes of PM2.5 exposure remain almost stable in the above-median regions, demonstrating the most polluted areas have not experienced disproportional declines between 2010 and 2019.

The spatial distribution of changes in PM2.5 PR in Figure 7 shows that PR in most cities remain unchanged (coloured by yellow). Nevertheless, the local variation presents a certain pattern. Sichuan Basin, southern Shaanxi and eastern Gansu became relatively less polluted, while air quality in the south-eastern coast and northeastern cities were relatively declined. The possible reasons may include regional environmental policies and industrial structure.

![Figure 7. Spatial distribution of changes in PM2.5 PR between 2010 and 2019](image)

### 5. CONCLUSIONS

This study investigates the changes of spatial disparity in PM2.5 exposures between 2010 and 2019 using satellite data in China from both absolute and relative perspectives. The PM2.5 exposures were first estimated by CNN-based models, with average CV R² of 0.82 and RMSE (MPE) of 13.1 (9.9) µg/m³. Substantial declines of PM2.5 exposures were observed from 47.99µg/m³ in 2010 to 35.52µg/m³ in 2019. The absolute disparity was then quantified and decomposed by Theil method. We found that the absolute difference in PM2.5 exposure across China are shrinking over time, with Theil index decreased from 0.132 to 0.054. The spatial disparity in 2019 were mainly contributed by intra-provincial differences, which differs from disparity in 2010 that inter-provincial disparity dominates. Though absolute spatial disparities have diminished, relative disparities persist. PM2.5 exposures in the least 20th percentile polluted cities have increased over time, while air quality in other regions were getting better. On average, the more/less polluted cities in 2010 were still the more/less polluted cities in 2019, except for the very most 2 percentile polluted cities. The result indicates the population in more polluted cities still experiences more air pollution than others. Spatially, compared with other regions, Sichuan Basin, southern Shaanxi and eastern Gansu became relatively less polluted, while south-eastern coast and northeast cities became "relatively more polluted. Therefore, the study suggests that more targeted intervention should be warranted to address these exposure spatial disparities.

**ACKNOWLEDGEMENTS**

This work is supported by the Fundamental Research Funds for the Central Universities (Grant No.300102351301) and Natural Science Basic Research Program of Shaanxi (Program No. 2022JQ-247). We would like to thank the NASA MODIS, AERONET and UCAR for their publicly available data.

**REFERENCES**


This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.

https://doi.org/10.5194/isprs-annals-V-3-2022-297-2022 | © Author(s) 2022. CC BY 4.0 License.