International Journal of Biology and Life Sciences
ISSN: 2957-9511 | Vol. 8, No. 2, 2024

Bridge between Precision PM2.5 Exposure Study and Healthy Life of Citizens: A Mini-review

Yazhu Wang [1, *], Chang Gao [2], Junrui Zhao [2]

[1] The Co-innovation Center for Social Governance of Urban and Rural Communities in Hubei Province, Zhongnan University of Economics and Law, Wuhan Hubei, 430073, China
[2] Research Center for Environment and Health, Zhongnan University of Economics and Law, Wuhan Hubei, 430073, China

*** Corresponding author:** Yazhu Wang

Abstract: Air pollution poses a significant health risk to residents. To better address the health and well-being needs of the population, more in-depth research on air pollution is essential. This article focuses on key issues in refined PM2.5 exposure research, including the development of high-precision PM2.5 distribution maps, the estimation of indoor PM2.5 exposure, and the variations in health risks among different populations exposed to PM2.5. Refined PM2.5 exposure research primarily involves three factors: the creation of accurate PM2.5 maps, the assessment of indoor PM2.5 exposure, and the categorization of different population groups. By thoroughly analyzing existing research, this article proposes future research directions and optimization strategies, with the goal of providing valuable insights for furthering refined PM2.5 exposure studies.

Keywords: PM2.5; Precise Exposure; Indoor; Group Differences.

1. Introduction

PM2.5 refers to a type of air pollutant consisting of inhalable particles with an aerodynamic diameter of 2.5 micrometers or less. Due to their small size, these particles can penetrate deep into the lungs, reaching the alveoli, and potentially causing significant harm to human health. Numerous studies have established a strong correlation between prolonged exposure to high levels of PM2.5 and the development of lung diseases, cardiovascular conditions, nervous system disorders, and even premature mortality [1-2]. Over the past decade or so, due to the rapid progress of urbanization and industrialization in China, the adverse health effect has been appeared. One study showed that in 2015, the number of premature deaths related to PM2.5 in 161 cities in China was 652,000 [2]. After more than a decade of efforts to control air pollution, air quality has significantly improved in most areas of China [3]. However, for various reasons, there is a noticeable trend of PM2.5 pollution rebounding in many regions [4]. While scholars are focusing on understanding the causes of this rebound and exploring governance strategies, it is equally important to consider how residents can reduce their exposure to pollution and protect their health and well-being in this context. Estimating PM2.5 exposure risk is a key task in this regard. Large-scale PM2.5 exposure assessments typically use cities or urban clusters as units, relying on medium- and long-term data from air pollution monitoring stations along with population statistics to evaluate the overall severity of the exposure risk [3].

Indeed, the methods mentioned above are valuable from an academic standpoint and can help elucidate regional variations in the impacts of air pollution, offering a macro-level perspective that can assist policymakers in implementing effective management strategies [3]. However, it should be noted that their practical significance for guiding citizens toward healthier living still requires enhancement. There are several reasons for this. For instance, residents' daily routines and activities are generally stable and not easily modified, and city-level pollution warnings often fail to effectively prompt individuals to take self-protection measures. Moreover, short-term exposure to PM2.5 does not necessarily equate to a lower health risk. Some studies have shown that the contribution of short-term PM2.5 exposure to the mortality burden is approximately one-seventh of that from long-term exposure [5]. Additionally, the harmful effects of PM2.5 vary across different populations [6-8], and the exposure risk, as well as the necessary protective measures during pollution events, differ accordingly. Furthermore, many PM2.5 exposure studies have primarily focused on outdoor air quality data, while people spend about 85% of their time indoors [9], making it essential to separately address the differences between indoor and outdoor air pollution. These issues are often overlooked in large-scale studies. It is therefore believed that more refined research on PM2.5 exposure estimation can provide more accurate guidance for residents' self-protection and management strategies during periods of air pollution. This article focuses on several key aspects of refined PM2.5 exposure research: the creation of higher-precision PM2.5 pollution maps, investigation of indoor PM2.5 levels, and the exploration of population-specific variations in the harmful effects of PM2.5.

2. High-precision PM2.5 Distribution Map

2.1. Traditional Methods

Due to cost and management constraints, the distribution of air pollution monitoring stations in cities is typically sparse and uneven, and personal sampling data is both unstable and limited. Consequently, inferring and constructing a comprehensive PM2.5 distribution map from the data of a few fixed monitoring stations remains a significant challenge in air pollution research. Currently, the three most commonly used methods for this task are physical and chemical models, interpolation methods, and regression models. Physical and chemical models include diffusion models, chemical transport models, and other deduction models based on physical and chemical principles. These models are theoretically reliable but require large datasets and complex computations, typically only achievable on large-scale servers or supercomputers. They represent more traditional simulation methods. Interpolation methods, commonly employed in geographical research, aim to fill in gaps in data. The Inverse Distance Weighting (IDW) and Kriging methods are among the most frequently used in this context. These methods are simple to compute, well-established, and widely used in PM2.5 exposure research [10-11]. However, their accuracy can be compromised when the number of monitoring stations is limited or unevenly distributed [12]. Regression models, which combine monitoring station data with other geographical, remote sensing, and meteorological information, attempt to mathematically derive a regression equation based on the correlation with the target variable. Geographic Weighted Regression (GWR) and Land Use Regression (LUR) are two common examples. In practice, regression models often yield more accurate results than interpolation methods [13-14]. An additional advantage of regression models is that, with the development of machine learning, more complex nonlinear models can be integrated into the analysis, significantly improving both accuracy and computational efficiency. As a result, combining regression models with machine learning techniques represents a promising direction for future research.

2.2. Methods Combining with Machine Learning

Machine learning algorithms can be integrated with interpolation methods to enhance predictive accuracy. Ma (2019) [15] identified the issue of sparse monitoring stations and proposed a hybrid algorithm combining bidirectional long short-term memory networks (Bi-LSTM) and inverse distance weighted interpolation (IDW). This approach outperformed IDW in predicting the spatial distribution of PM2.5 concentration in Guangdong Province. Machine learning can also be paired with regression models. He (2019) [16] experimentally demonstrated that artificial neural networks (ANN) and Bayesian maximum entropy (BME) produced more accurate models than the simpler land use regression (LUR) model. Xiao (2020) [17] further combined LUR with BME, and after conducting multiple experiments with combined test data, the accuracy of PM2.5 concentration estimation in the Yangtze River Delta region exceeded that of both the simple LUR model (0.427) and the Kriging interpolation method (0.487) . Dai (2022) [13] integrated LUR with the LightGBM algorithm, yielding annual average data results following data source reorganization. Although this time-scale analysis is not ideal for fine-scale PM2.5 exposure studies, it provided a comparative assessment of several machine learning algorithms, including BP neural networks (BPNN), deep neural networks (DNN), random forests (RF), XGBoost, and LightBGM, highlighting their respective advantages and limitations. He (2022) [18] applied a gradient-boosted decision tree (GBDT) algorithm using aerosol data from the MAIAC system to develop a spatiotemporal inversion model for nationwide 1 km resolution PM2.5 concentration estimation . This study underscores the importance of incorporating diverse data sources—such as geographic, meteorological, and remote sensing data—to enhance the predictive power of machine learning models. Huang (2021) [19] also utilized aerosol data and created a prediction framework combining geographic and meteorological data. This framework successfully reduced the prediction resolution of daily PM2.5 concentration to 100m, aligning with data from air quality monitoring stations, thus providing valuable insights for more refined research.

3. Indoor PM2.5 Estimation

3.1. The Importance of Indoor Air Pollution Controlling

Indoor PM2.5 is the second major focus of this article. While most air pollution studies based on monitoring stations primarily focus on outdoor data, indoor air pollution has long been a significant threat to public health [20]. Globally, people spend approximately 90% of their time in various indoor environments [9], and indoor air pollution is responsible for millions of premature deaths each year [21]. The sources of indoor PM2.5 pollution can be categorized into three main types: particles that are produced outdoors and enter indoor spaces, particles generated by indoor human activities, and particles emitted by the building itself [9]. In China, the indoor PM2.5 situation varies significantly by region. In rural areas, the use of biomass and coal fuels for cooking and heating is the primary source of household air pollution, contributing substantially to the overall health burden in these populations [22]. In contrast, urban areas face different challenges. Some studies suggest that indoor PM2.5 in Chinese cities is mainly influenced by the transport of outdoor PM2.5 into indoor spaces [23]. These variations highlight two important points: First, they underscore the necessity of region-specific research on indoor PM2.5, as levels can differ significantly across different areas and cannot be adequately assessed using a single indicator. Second, they emphasize the complexity of conducting indoor PM2.5 research, as it remains challenging to identify models and theories that are universally applicable across diverse environments and conditions.

3.2. Measurement of Indoor PM2.5

Unlike outdoor air pollution, there are no national standard monitoring stations for indoor air quality operating year-round, which makes the measurement of indoor PM2.5 levels inherently problematic. While household air quality monitors are a potential solution for families with the means to invest in such equipment [24], most households lack these devices, making it difficult to collect and integrate these data into comprehensive research. As a result, the more commonly used research approach involves sampling and measuring air quality in a limited number of buildings, followed by analysis, modeling, and investigation of the sources and variations of particulate matter. In a review by Diapouli (2013) [25], several studies attempted to estimate the relative contributions of indoor and outdoor sources to indoor particle concentrations using statistical analysis or mass balance equations applied to time series of indoor and outdoor concentrations. One of the key challenges identified in these studies was accurately calculating the infiltration efficiency (P) and the deposition rate (k) separately. Gemenetzis (2006) [26] conducted a study at Aristotle University in Greece, measuring PM2.5 levels in 40 rooms, and found that rooms occupied by smokers had significantly higher PM2.5 concentrations—up to 70% higher—compared to non-smokers' rooms. Another key finding from this study was that PM2.5 concentrations tended to decrease slightly with increasing floor number, marking an early example of indoor

PM2.5 estimation. Chen (2011) [27] provides an updated review of both experimental and modeling approaches regarding the relationship between indoor and outdoor particles. The use of three key parameters—the indoor/outdoor (I/O) ratio, infiltration factor, and penetration factor—was examined to assess this relationship. The experimental data for these parameters, measured in both real homes and laboratory settings, were summarized and analyzed. The I/O ratios varied significantly due to factors such as size-dependent indoor particle emission rates, the geometry of cracks in building envelopes, and air exchange rates. As a result, drawing uniform conclusions is challenging, as detailed information remains sparse, making the I/O ratio less useful for understanding the indoor/outdoor relationship. Chatoutsidou (2015) [28] conducted measurements of real-time I/O particulate matter (PM10) mass concentration and particle count concentration in a modern office environment with mechanical ventilation. The results revealed that the presence of people during working hours significantly influenced both the quantity and quality of particles in the office, compared to non-working hours. In the absence of major indoor sources of new particles (such as printing equipment), the indoor environment was primarily impacted by human presence, with suspended activities being the leading source of particles larger than 1 μm. Furthermore, outdoor pollution sources had a considerable impact on the quantity and quality of indoor particle concentrations. Wallace (2022) [29] applied the Random Component Superposition (RCS) regression model to estimate the infiltration factors for 790,000 paired indoor and outdoor sites. These infiltration factors were used to estimate indoor-generated PM2.5 and outdoor-permeated PM2.5 concentrations. On average, indoor-generated particles accounted for 46%–52% of the total indoor PM2.5 concentration. However, the contribution of indoor pollution sources to the overall indoor PM2.5 varied widely, ranging from nearly zero to almost 100%. The influence of indoor-generated particles on potential exposure was markedly different from that of outdoor concentration variations. Ji (2015) [30] conducted a study based on actual measurements from 90 residential rooms in Beijing and applied a mass balance model to estimate the contribution of various sources to indoor PM2.5 levels. The results indicated that outdoor particles made the largest contribution to indoor PM2.5 concentrations, accounting for 54%–63% of the total when all windows were closed, and more than 92% when windows were open. The two most significant factors influencing indoor PM2.5 levels were outdoor PM2.5 concentration and the rate of particle settlement. Air exchange rate, cooking activities, and smoking intensity also had a notable impact on indoor PM2.5 concentrations, while the effect of the PM2.5 penetration coefficient was minimal. This study serves as an important reference for understanding the influence of outdoor PM2.5 on indoor air quality in Chinese cities.

4. Group Differences in PM2.5 Exposure Study

4.1. Group Differences in PM2.5 Exposure Study

The third issue addressed in this article concerns the differences in PM2.5 exposure estimates across various populations. Early studies often treated humans as static receptors, even when considering spatial distribution, and rarely accounted for the dynamic movement of individuals. In contrast, large-scale studies typically viewed populations as homogeneous, overlooking the fact that different groups face varying levels of risk from PM2.5 exposure. For instance, the elderly and children tend to have higher susceptibility and greater health risks.

4.2. Distribution of Population

Some studies have focused on the uneven spatial distribution of the population, adjusting air pollution exposure estimates through population weighting. Wang (2019) [31] observed that the spatial distribution of PM2.5 concentrations might not align with the spatial distribution of the population. The study found that population exposure to PM2.5 exhibited clear urban-suburban-rural differences: while it was consistent with population distribution, it did not match the distribution of PM2.5 concentrations. Song (2019) [32] examined the impact of dynamic changes in population distribution along roadways when estimating PM2.5 exposure in 13 cities within the Beijing-Tianjin-Hebei region of China. The study quantified population-weighted exposure risk using static population distribution data and PM2.5 observation data, but also highlighted that the actual dynamic population movement significantly influenced exposure assessment results, a factor that future researchers should carefully consider. Li (2021) [33] further addressed this by incorporating high-resolution PM2.5 data and heat maps based on LBS big data to simulate hourly population changes. The study revealed that population-weighted exposure was underestimated by 45.83% in suburban areas and overestimated by more than 34.87% in city centers. Additionally, PM2.5 exposure on weekdays was worse than previously assessed, with about 12.41% of individuals experiencing PM2.5 levels exceeding the census-based estimates, double the expected number.

4.3. Health Risks among Different Groups

Several studies have thoroughly examined the variations in health effects of PM2.5 exposure across different populations. Zhang (2022) [8] conducted a study to investigate gender differences in the relationship between long-term PM2.5 exposure and two cardiovascular diseases: ischemic heart disease (IHD) and stroke. The study found that women exposed to PM2.5 over an extended period had a higher risk of IHD, while no significant gender differences were observed in the risk of stroke. The findings suggested that gender differences should be considered when addressing cardiovascular disease risks potentially linked to PM2.5 exposure. Hu (2022) [6] explored the impact of PM2.5 exposure on three multi-morbidity patterns (respiratory, musculoskeletal, and cardiometabolic) as well as overall health outcomes across different age groups. The study highlighted that PM2.5 exposure posed greater harm to the health of older adults. Lin (2021) [7] investigated health disparities related to PM2.5 exposure among individuals with different educational backgrounds. The study found that increased PM2.5 exposure was positively correlated with higher rates of diabetes, hypertension, and overweight among individuals with lower educational levels, but this correlation was not observed in those with higher levels of education. This disparity may be linked to differences in job types and self-protection behaviors associated with varying education levels.

5. Conclusion

This article primarily addresses three key issues in fine-scale PM2.5 exposure estimation: the construction of high-precision PM2.5 maps, the estimation of indoor PM2.5 exposure, and the segmentation of exposed populations. The literature analysis leads to the following conclusions: (1) Machine learning methods have yielded promising results in constructing high-precision PM2.5 maps. Future research should focus on collecting more comprehensive data and selecting more effective machine learning models to achieve finer prediction scales and improve accuracy. (2) Indoor PM2.5 levels are influenced by various factors; however, in urban buildings without indoor emission sources, the primary source of indoor PM2.5 is outdoor air. The use of transmission models or mass balance models is a common approach for estimating indoor exposure. (3) The health risks associated with PM2.5 exposure vary across different populations, with factors such as age, gender, and occupation playing a significant role. These factors should be considered when studying fine-scale exposure. Additionally, several challenges and potential areas for optimization in the existing research are identified: (1) Indoor PM2.5 estimation is affected by the heterogeneity of different building types. Different models should be applied to different building categories, but few studies have incorporated building conditions as variables to minimize model variability and provide a more standardized approach to indoor PM2.5 estimation. (2) Population data, particularly dynamic distribution data, presents privacy concerns and is often difficult to obtain. Future research should aim to integrate more detailed population data into fine-scale PM2.5 exposure assessments while safeguarding privacy and avoiding other related issues.

References

[1] P.F. Peng, X.B. Guo, F.M.H. Cheung, et al. The association between PM2.5 exposure and neurological disorders: A systematic review and meta-analysis. Science of the Total Environment, 2019, 655: 1240-1248.

[2] K.J. Maji, A.K. Dikshit, M. Arora, et al. Estimating premature mortality attributable to PM2.5 exposure and benefit of air pollution control policies in China for 2020. Science of the Total Environment, 2018, 612: 683-693.

[3] J. Song, C.L. Li, M. Liu, et al. Spatiotemporal Distribution Patterns and Exposure Risks of PM2.5 Pollution in China. Remote Sensing, 2022, 14(13): 3173.

[4] C.G. Dong, J.Y. Li, Y. Qi. Decomposing PM2.5 air pollution rebounds in Northern China before COVID-19. Environmental Science and Pollution Research International, 2022, 29: 28688 - 28699.

[5] T.T. Li, Y.M. Guo, Y. Liu, et al. Estimating mortality burden attributable to short-term PM2.5 exposure: A national observational study in China. Environment International, 2019, 215: 245 - 251.

[6] K. Hu, K. Keenan, J.M. Hale, et al. A longitudinal analysis of PM2.5 exposure and multimorbidity clusters and accumulation among adults aged 45-85 in China. PLOS Global Public Health, 2022, 2(6): e0000520.

[7] J.F. Lin, H. Zheng, P. Xia, et al. Long-term ambient PM2.5 exposure associated with cardiovascular risk factors in Chinese less educated population. BMC Public Health, 2021, 21(1): 2241.

[8] J. Zhang, X.Y. Wang, M.F. Yan, et al. Sex Differences in Cardiovascular Risk Associated with Long-Term PM2.5 Exposure: A Systematic Review and Meta-Analysis of Cohort Studies. Frontiers in Public Health, 2022, 10: 802167.

[9] M. Mannan, S.G. Al-Ghamdi. Indoor Air Quality in Buildings: A Comprehensive Review on the Factors Influencing Air Pollution in Residential and Commercial Structure. International Journal of Environmental Research and Public Health, 2021, 18(6): 3276.

[10] N. Cressie. The origins of kriging. Mathematical Geology, 1990, 22: 239-252.

[11] G.Y. Lu, D.W. Wong. An adaptive inverse-distance weighting spatial interpolation technique. Computers & Geosciences, 2008, 34(9): 1044-105.

[12] M.A. Oliver, R. Webster. Kriging: a method of interpolation for geographical information systems. International Journal of Geographical Information Systems, 1990, 4(3): 313-332.

[13] H.B. Dai, G.Q. Huang, J.J. Wang, et al. Spatio-Temporal Characteristics of PM2.5 Concentrations in China Based on Multiple Sources of Data and LUR-GBM during 2016-2021. International Journal of Environmental Research and Public Health, 2022, 19(10): 6292.

[14] K.Y. Gu, Y. Zhou, H. Sun, et al. Spatial distribution and determinants of PM2.5 in China's cities: fresh evidence from IDW and GWR. Environmental Monitoring and Assessment, 2020, 193(1): 15.

[15] J. Ma, Y.X. Ding, V.J.L. Gan, et al. Spatiotemporal Prediction of PM2.5 Concentrations at Different Time Granularities Using IDW-BLSTM. IEEE Access, 2019, 7: 107897-107907.

[16] J.Y. He, G. Christakos, P. Jankowski. Comparative Performance of the LUR, ANN, and BME Techniques in the Multiscale Spatiotemporal Mapping of PM2.5 Concentrations in North China. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2019, 12(6): 1734-1747.

[17] L. Xiao, G. Christakos, J.Y. He, et al. Space-Time Ground-Level PM2.5 Distribution at the Yangtze River Delta: A Comparison of Kriging, LUR, and Combined BME-LUR Techniques. Journal of Environmental Informatics, 2020, 36(1): 33-42.

[18] W.H. He, H. Meng, J. Han, et al. Spatiotemporal PM2.5 estimations in China from 2015 to 2020 using an improved gradient boosting decision tree. Chemosphere, 2022, 296: 134003.

[19] C.H. Huang, J.L. Hu, T. Xue, et al. High-Resolution Spatiotemporal Modeling for Ambient PM2.5 Exposure Assessment in China from 2013 to 2019. Environmental Science & Technology, 2021, 55(3): 2152-2162.

[20] N. Bruce, R. Perez-Padilla, R. Albalak. Indoor air pollution in developing countries: a major environmental and public health challenge. Bulletin of the World Health Organization, 2020, 78(9): 1078-1092.

[21] V.V. Tran, D. Park, Y.C. Lee. Indoor Air Pollution, Related Human Diseases, and Recent Trends in the Control and Improvement of Indoor Air Quality. International Journal of Environmental Research and Public Health, 2020, 17(8): 2927.

[22] W. Hu, G.S. Downward, B. Reiss, et al. Personal and Indoor PM2.5 Exposure from Burning Solid Fuels in Vented and Unvented Stoves in a Rural Region of China with a High Incidence of Lung Cancer. Environmental Science & Technology, 2014, 48(15): 8456-8464.

[23] Z.J. Shao, X.J. Yin, J. Bi, et al. Spatiotemporal Variations of Indoor PM2.5 Concentrations in Nanjing, China. International Journal of Environmental Research and Public Health, 2019, 16(1): 144.

[24] Z.Q. Wang, W.W. Delp, B.C. Singer. Performance of low-cost indoor air quality monitors for PM2.5 and PM10 from residential sources. Building and Environment, 2020, 171: 106654.

[25] E. Diapouli, A. Chaloulakou, P. Koutrakis. Estimating the concentration of indoor particles of outdoor origin: A review. Journal of the Air & Waste Management Association, 2013, 63(10): 1113-1129.

[26] P. Gemenetzis, P. Moussas, A. Arditsoglou, et al. Mass concentration and elemental composition of indoor PM2.5 and PM10 in University rooms in Thessaloniki, northern Greece. Atmospheric Environment, 2006, 40 (17): 3195-3206.

[27] C. Chen, B. Zhao. Review of relationship between indoor and outdoor particles: I/O ratio, infiltration factor and penetration factor. Atmospheric Environment, 2011, 45(2): 275-288.

[28] S.E. Chatoutsidou, J. Ondracek, O. Tesar, et al. Indoor/outdoor particulate matter number and mass concentration in modern offices. Building and Environment, 2015, 92: 462-474.

[29] L.A. Wallace, T.K. Zhao, N.E. Klepeis. Indoor contribution to PM2.5 exposure using all PurpleAir sites in Washington, Oregon, and California. Indoor Air, 2022, 32(9): e13105.

[30] W.J. Ji, B. Zhao. Contribution of outdoor-originating particles, indoor-emitted particles and indoor secondary organic aerosol (SOA) to residential indoor PM2.5 concentration: A model-based estimation. Building and Environment, 2015, 90: 196-205.

[31] H. Wang, J.W. Li, Z.Q. Gao, et al. High-Spatial-Resolution Population Exposure to PM2.5 Pollution Based on Multi-Satellite Retrievals: A Case Study of Seasonal Variation in the Yangtze River Delta, China in 2013. Remote Sensing, 2019, 11(23): 2724.

[32] Y.M. Song, B. Huang, Q.Q. He, et al. Dynamic assessment of PM2.5 exposure and health risk using remote sensing and geo-spatial big data. Environmental Pollution, 2019, 253: 288-296.

[33] X. Li, T. Yang, Z.T. Zeng, et al. Underestimated or overestimated? Dynamic assessment of hourly PM2.5 exposure in the metropolitan area based on heatmap and micro-air monitoring stations. Science of the Total Environment, 2021, 779: 146283.