Asian Growth Research Institute

調査報告書 25-05

What Drives Air Pollution in China? Evidence from Interpretable Machine Learning and Spatial Analysis of $PM_{2.5}$

令和 8（2026）年 3 月

公益財団法人　アジア成長研究所

What Drives Air Pollution in China? Evidence from Interpretable Machine Learning and Spatial Analysis of $PM_{2.5}$

Alvaro Domínguez[*]

[*]Asian Growth Research Institute, Kitakyushu, Japan.

February 13, 2026

Abstract

This paper investigates the determinants of fine particulate matter ($PM_{2.5}$) across 31 Chinese provinces during 2004–2020, a period of rapid economic transformation accompanied by rising environmental pressures. Using a provincial panel dataset and a flexible empirical framework, we identify the economic, demographic, and energy-related factors most closely associated with regional variation in air quality. Population pressure and the composition of the regional energy mix emerge as the most influential predictors, while higher levels of solar and wind power capacity are associated with lower predicted pollution. The analysis also reveals substantial geographic heterogeneity, indicating that the drivers of air quality differ markedly across regions. These findings underscore the importance of geographically targeted environmental policies that simultaneously expand renewable energy capacity and address urban emission sources, particularly in densely populated provinces.

Keywords: China, $PM_{2.5}$, Machine Learning, Spatial analysis.

1 Introduction

Air pollution has emerged as one of the most pressing environmental challenges accompanying rapid economic growth, particularly in developing economies where industrialization, urban expansion, and rising energy demand have intensified environmental pressures. China provides a particularly relevant context for studying these dynamics, as its sustained economic transformation has generated substantial improvements in income and infrastructure while simultaneously contributing to higher concentrations of fine particulate matter ($PM_{2.5}$). Elevated levels of $PM_{2.5}$ pose significant risks to public health, making it essential to understand the structural forces that drive regional differences in air quality. Identifying these forces is critical for designing policies capable of balancing continued economic development with environmental sustainability.

In this study, we analyze three central research questions: (a) which economic, demographic, and energy factors are most strongly associated with provincial variation in $PM_{2.5}$ levels in China; (b) the extent to which these relationships exhibit nonlinearities or interaction effects; and (c) whether meaningful spatial heterogeneity exists in the determinants of air pollution that would justify geographically differentiated environmental policies. By addressing these questions, the analysis aims to provide a comprehensive and empirically grounded assessment of the structural drivers of air quality across Chinese provinces.

To address these questions, we combine a rich provincial panel dataset with a flexible empirical framework based on machine learning (ML) methods designed to capture complex relationships while maintaining interpretability. This approach enables us to identify the factors most closely associated with pollution outcomes and to examine how their contributions vary across space, thereby providing insights that are directly relevant for policy design.

The use of ML is particularly suitable in this context because it allows the data to flexibly capture complex and potentially nonlinear relationships without requiring researchers to pre-specify functional forms. Among these approaches, ensemble methods such as gradient boosting have proven especially effective in high-dimensional settings, making them well suited for analyzing the structural drivers of air pollution.

This paper contributes to the literature in three ways. First, it provides a comprehensive assessment of the structural drivers of $PM_{2.5}$ across Chinese provinces using a flexible empirical framework capable of capturing nonlinear relationships and interaction effects. Second, it integrates interpretable machine learning techniques with spatial analysis to uncover geographic heterogeneity in the determinants of air pollution. Third, it generates policy-relevant insights by identifying the regional conditions under which renewable energy expansion and demographic pressures most strongly influence air quality outcomes.

The remainder of the paper is organized as follows. Section 2 reviews the relevant literature, Section 3 describes the data, Section 4 presents the results, and Section 5 concludes.

2 Literature Review

This study contributes to several strands of literature examining the determinants of air pollution, the relationship between economic activity and environmental quality, and the increasing use of machine learning methods in applied economic research.

A large body of work has analyzed the connection between economic growth and environmental degradation, often framed through the Environmental Kuznets Curve (EKC) hypothesis. Early contributions argue that pollution initially rises with income but eventually declines once economies reach higher levels of development (Grossman and Krueger, 1995; Selden and Song, 1994). Later research questions the universality of this inverted-U relationship, emphasizing instead the importance of policy interventions, technological progress, and structural transformation in shaping environmental outcomes (Stern, 2004; Dinda, 2004). In the Chinese context, rapid industrialization and export-led growth have been closely associated with rising emissions, although more recent investments in cleaner technologies and stronger environmental regulation have begun to moderate these patterns (Zhang et al., 2016, 2018).

Another related strand focuses specifically on the determinants and consequences of fine particulate matter. Exposure to $PM_{2.5}$ has been shown to significantly increase mortality risk and reduce life expectancy (Pope Iii et al., 2002; Chen et al., 2013). Beyond health impacts, air pollution can negatively affect labor productivity, cognitive performance, and broader economic activity (Zivin and Neidell, 2012; Chang et al., 2019). Empirical work on China identifies industrial activity, fossil-fuel-based energy consumption, particularly coal, and rapid urbanization as central drivers of particulate pollution (Lin and Zhu, 2018; Han et al., 2014). These findings highlight the importance of understanding how economic structure, demographic pressures, and energy systems jointly influence regional air quality.

The spatial dimension of pollution has also received increasing attention. Air pollution is inherently geographic, reflecting the interaction between local emissions, regional economic structures, and atmospheric transport. Studies document substantial spatial heterogeneity in pollution exposure across Chinese provinces, suggesting that uniform national policies may generate uneven environmental outcomes (Zhou et al., 2018; Wang et al., 2018). This evidence has motivated calls for geographically differentiated policy approaches that account for regional differences in industrial composition, population density, and energy infrastructure.

Methodologically, this paper relates to a growing literature integrating machine learning tools into empirical economics. Traditional econometric models often impose restrictive functional forms that may fail to capture nonlinearities and interaction effects present in complex environmental systems. ML methods offer a flexible alternative by allowing the data to reveal underlying patterns while maintaining strong predictive performance (Mullainathan and Spiess, 2017; Athey and Imbens, 2019). Ensemble techniques such as gradient boosting are particularly effective in high-dimensional settings where relationships among variables are potentially nonlinear (Friedman, 2001; Hastie et al., 2009; Chen and Guestrin, 2016).

Recent advances also emphasize the importance of interpretability when applying these methods to policy-relevant questions. Tools such as SHAP (SHapley Additive exPlanations) provide a theoretically grounded framework for decomposing model predictions into variable-level contributions, thereby improving transparency and facilitating economic interpretation (Lundberg and Lee, 2017). By combining predictive accuracy with interpretability, these approaches allow researchers to generate insights that are directly relevant for policy design.

3 Data

To analyze the determinants of $PM_{2.5}$ pollution across 31 provinces in China from 2004 to 2020, we use data obtained from the National Bureau of Statistics of China. This period captures a stage of sustained economic expansion, rapid industrialization, and urban growth, accompanied by increasing environmental pressures. As such, it provides an appropriate context for examining the interaction between economic development and air quality. Understanding this relationship is essential for evaluating the challenges policymakers face in promoting economic growth while maintaining environmental sustainability.

Table 1 reports the definitions of all variables used in the analysis. We group the variables into seven broad categories: air pollution, economic activity, demographic characteristics, human capital and innovation, energy production, transportation, and administrative identifiers. Air quality is measured using the annual average concentration of $PM_{2.5}$, complemented by its logarithmic transformation and alternative population-weighted indicators to ensure robustness to measurement choice.

The remaining variables capture the structural conditions under which pollution is generated. Economic activity is reflected in income levels, total output, sectoral composition, and foreign investment, while demographic measures control for population pressures. Education outcomes and patent activity serve as proxies for human capital and innovative capacity. Given the central role of energy use in driving emissions, the analysis incorporates detailed measures of electricity generation by source, fossil fuel consumption, and gas infrastructure. Transportation indicators, including freight activity, passenger flows, infrastructure mileage, and vehicle stocks, capture mobility patterns that may further influence pollution levels. Together, these variables provide a comprehensive framework for identifying the structural determinants of air quality differences across Chinese provinces.

Table 1: Variable Definitions

Category / Variable	**Description**
Air Pollution	
$PM_{2.5}$ Mean; log $PM_{2.5}$ Mean; alternative $PM_{2.5}$ measures	Annual average concentration of fine particulate matter, its logarithm, and cumulative or population-weighted indicators used to capture exposure to air pollution.
Economic Activity	
GDP per capita; Total GDP; Value added (primary, secondary); Foreign investment	Indicators reflecting income levels, the scale of economic activity, industrial composition, and external capital inflows.
Demographic Characteristics	
Population measures	Total population, density, and year-end population levels capturing demographic pressures.
Human Capital and Innovation	
Education indicators; Patent grants	Secondary and tertiary enrollment and graduation counts, along with domestic patents measuring innovative capacity.
	Continued on next page

Category / Variable	Description
Energy Production	
Electricity generation; Energy mix (thermal, hydro, nuclear, wind, solar); Gas infrastructure; LPG supply; Fuel consumption; New-energy buses	Measures describing electricity production, reliance on alternative energy sources, gas availability and coverage, fossil fuel consumption, and adoption of cleaner public transportation technologies.
Transportation	
Road and rail transport; Highway and railway mileage; Vehicle fleet composition	Freight and passenger flows, transport infrastructure, and the size and composition of civil and private vehicle stocks.
Administrative Identifiers	
Province, year, regional codes	Variables defining the panel structure of the dataset.

4 Results

Machine learning (ML) is a set of methods primarily designed to identify patterns in data and generate accurate predictions, often without requiring researchers to specify a predetermined functional form. While these methods are typically prediction-oriented, they can also complement causal frameworks when used appropriately. Instead of assuming linear relationships or specifying interactions ex ante, ML algorithms allow the data to reveal how explanatory variables jointly influence the outcome of interest. This flexibility is particularly valuable in the study of air pollution, where environmental outcomes are shaped by complex and potentially nonlinear relationships involving economic activity, population dynamics, and the energy mix, among others. By leveraging these methods, the analysis can capture hidden structures in the data that may remain undetected under traditional econometric approaches.

In this study, we use a machine learning method called Extreme Gradient Boosting (XGBoost), which improves predictions by combining many prediction trees rather than relying on a single equation. Specifically, we implement XGBoost as a regression model to predict provincial variation in log $PM_{2.5}$. Hyperparameters were selected using cross-validation to ensure robust predictive performance and to limit overfitting. The dependent variable is the logarithm of average $PM_{2.5}$ concentrations, which serves as a measure of particulate pollution across province-year observations. Formally, the model can be expressed as:

$$\log(\mathrm{PM}_{2.5})_{pt} = f(X_{pt}) + \alpha_p + \delta_t + \varepsilon_{pt}, \tag{1}$$

where p indexes provinces and t indexes years, X_{pt} denotes the vector of observed covariates, α_p represents province fixed effects, and δ_t captures year fixed effects. The function $f(\cdot)$ is unknown and is approximated using boosted prediction trees, allowing the conditional expectation of pollution to depend flexibly on the covariates without imposing a predetermined functional form.

The model relates this outcome to a broad set of covariates capturing demographic, economic, energy, and transportation conditions. These include population measures, GDP indicators, sectoral activity, proxies for education and innovation such as patents and graduates, multiple sources of electricity generation including wind and solar power, energy supply variables, and transportation volumes. A detailed description of all variables is provided in Table 1. Province and year fixed effects are included to control for unobserved location-specific and time-specific factors that may influence pollution levels.

The XGBoost model demonstrates strong predictive performance in explaining variation in log $\mathrm{PM}_{2.5}$ across Chinese provinces. Figure 1 shows the in-sample relationship between observed and predicted values, where the points closely follow the 45-degree line, indicating that the model captures much of the systematic variation in the data. The residuals are tightly centered around zero and display no systematic pattern, suggesting that little explanatory variation remains unaccounted for. To ensure that this strong fit is not driven by overfitting, Figure 2 reports the out-of-fold (OOF) results obtained through 10-fold cross-validation. Although slightly more dispersed, as expected when predicting unseen observations, the predictions remain highly accurate, with an OOF RMSE of 0.072. The approximately normal distribution of residuals further supports the model's ability to generalize beyond the training sample.

The high predictive accuracy is consistent with the strengths of XGBoost, which builds sequential decision trees to progressively reduce prediction errors while controlling model complexity. This approach enables the model to capture nonlinear relationships and interaction effects that are common in environmental and economic systems, thereby improving its ability to generalize to new data.

Figure 1: Observed vs Predicted Diagnosis

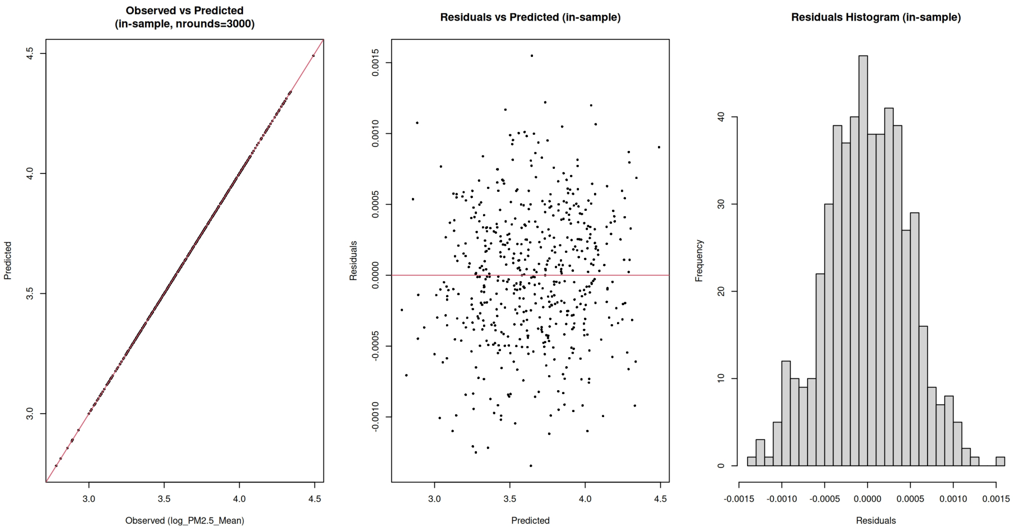
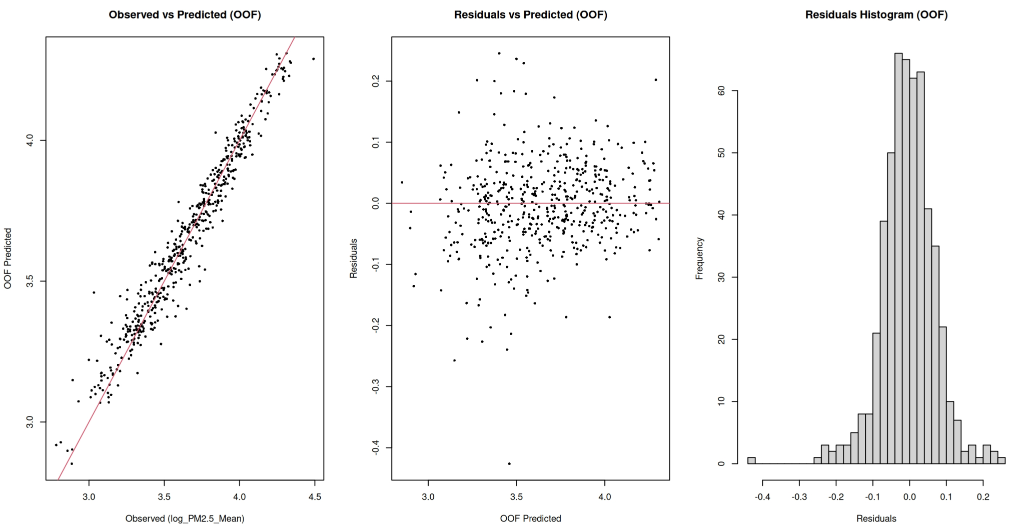


Figure 2: Observed vs Predicted Diagnosis (OOF)



Due to the "black box" nature of these models, it is difficult to interpret the results obtained from them. To address this issue we rely on interpretability tools. Figures 3 and 4 below present the global SHAP values, which quantify the contribution of each predictor to the model's predictions. The near-identical ranking between the in-sample and OOF SHAP results provides further evidence of model stability. Across both figures, *Solar_Power* is the most influential predictor by a substantial margin, followed by *POP_Mean* and *Wind_Power*. The prominence of these variables highlights the central role of the regional energy mix and demographic pressure in shaping particulate pollution levels. Provinces with larger populations typically experience greater transportation demand, industrial activity, and energy consumption, all of which are associated with higher emissions. Meanwhile, the importance of renewable energy variables suggests that differences in energy transition paths across provinces are strongly linked to air quality outcomes.

Figure 3: Global SHAPs

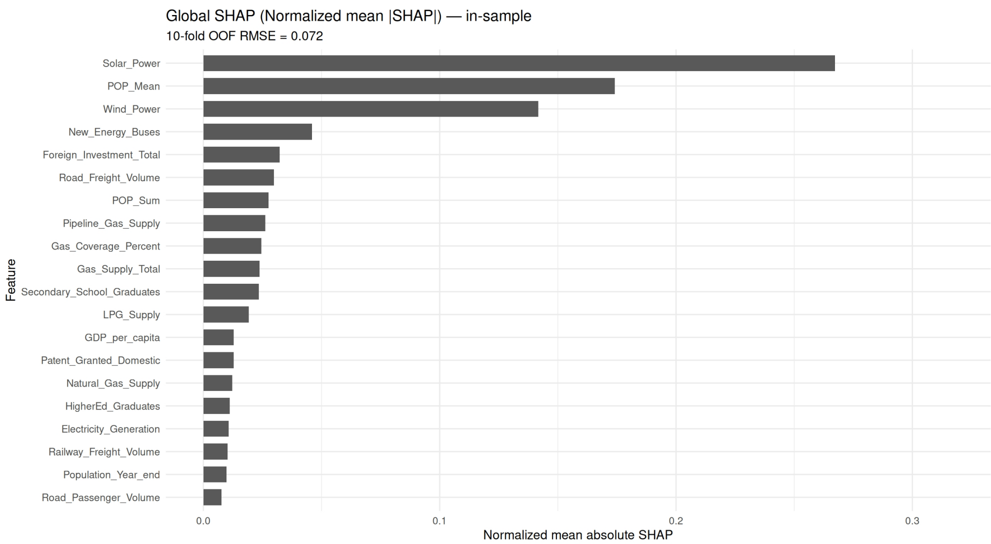
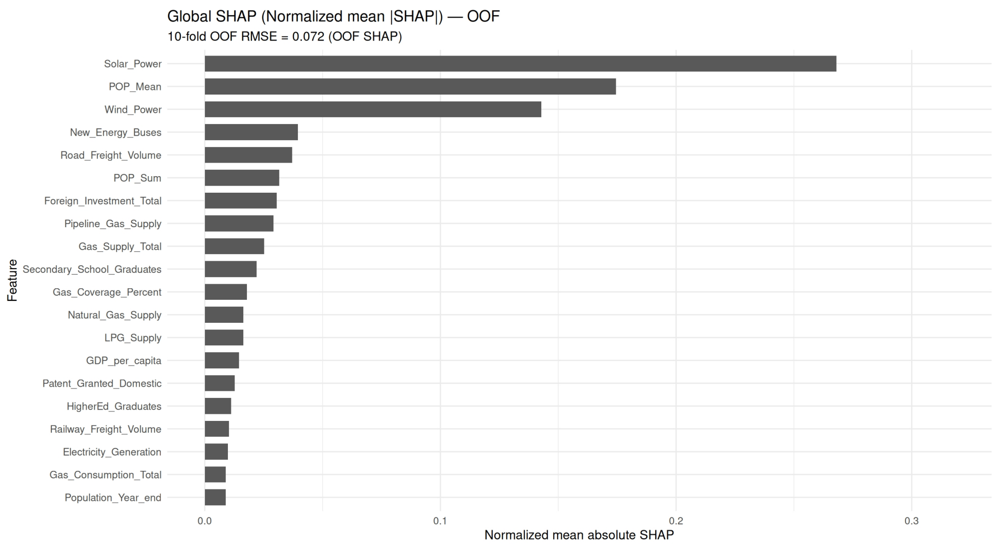


Figure 4: Global SHAPs (OOF)



To further examine how the most important predictors are associated with pollution levels, we analyze SHAP dependence plots for population density, solar power, and wind power, the three variables identified as most important in the global SHAP analysis (Figures 5, 6, and 7). These plots display how each variable contributes to the predicted value of $\log(PM_{2.5})$ across observations while accounting for the joint influence of the remaining predictors.

In these plots, observations are color-coded by geographic region (North, East, South, and West), allowing for a visual assessment of spatial heterogeneity in the relationships between the predictors and pollution levels. Unlike traditional regression coefficients, which impose a constant marginal effect, SHAP dependence plots allow the relationship between a predictor and the outcome to vary across the data. This flexibility makes them particularly well suited for detecting nonlinear patterns and heterogeneous effects, both of which are common in environmental systems. Such interpretability is especially valuable in environmental policy contexts, where identifying the factors most strongly associated with pollution is essential for designing targeted interventions.

Figure 5 suggests that greater solar power capacity is associated with negative contributions to predicted pollution, consistent with the view that renewable energy adoption is associated with lower particulate emissions. The color distribution indicates that this pattern is broadly observed across regions, although provinces in the western region appear more frequently at higher capacity levels, reflecting differences in renewable energy deployment.

A similar pattern emerges in Figure 6, where higher levels of wind power are linked to reductions in predicted pollution. The regional color coding reveals some geographic dispersion in the magnitude of these effects, suggesting that the environmental benefits of wind energy may vary depending on local economic structures and energy systems.

In contrast, Figure 7 shows that higher population density is generally associated with positive SHAP values, indicating that more densely populated provinces tend to contribute to higher predicted pollution levels. This positive association is visible across all regions but appears particularly pronounced in eastern provinces, where economic activity and urban concentration are highest. This finding aligns with the concentration of transportation demand, industrial production, and energy consumption typically observed in densely populated areas. From these results, we see the dual importance of demographic pressures and energy transition in explaining provincial variation in air quality, while also pointing to meaningful regional differences in these relationships.

Figure 5: SHAP Dependence for Solar Power (OOF)

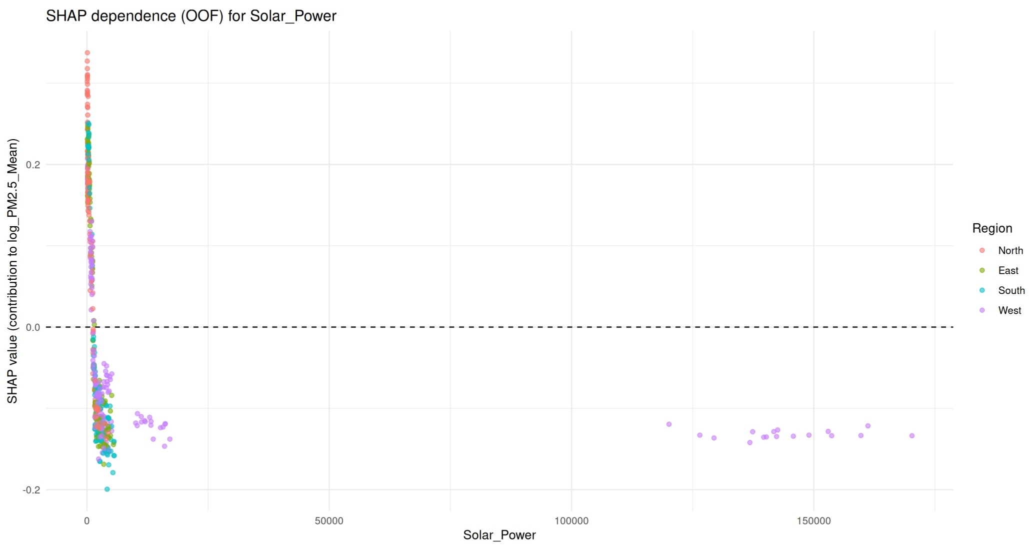


Figure 6: SHAP Dependence for Wind_Power (OOF)

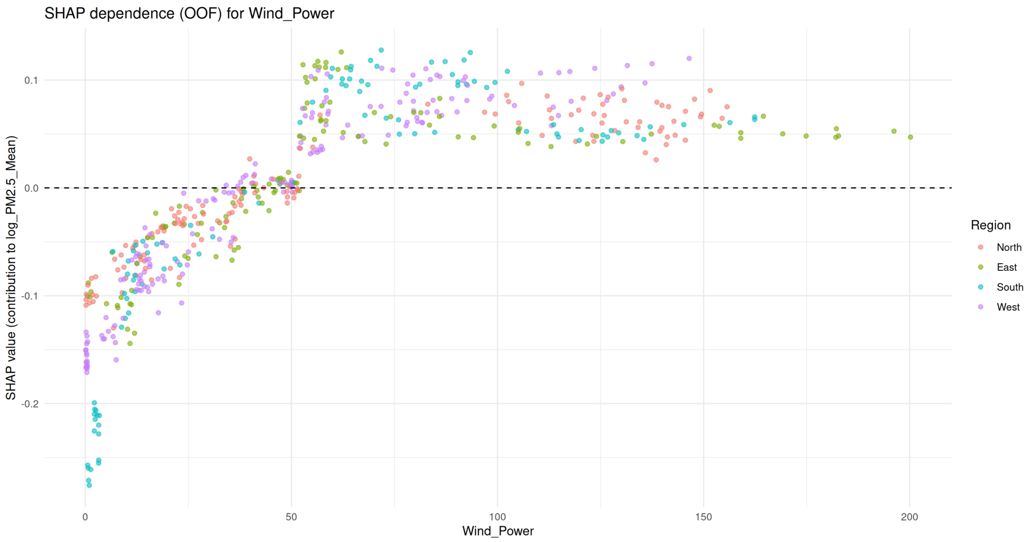
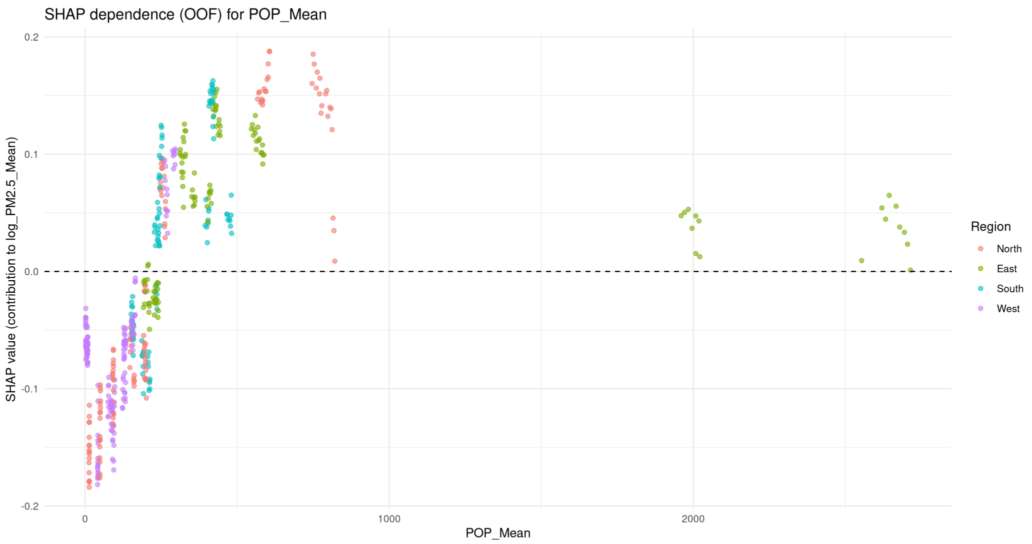


Figure 7: SHAP Dependence for Pop_MEAN (OOF)



We next present province-level maps to provide a spatial interpretation of the model's predictions. While the dependence plots describe how each variable is associated with pollution across observations, mapping mean SHAP values helps visualize where these relationships are strongest. In these maps, negative SHAP values indicate that a predictor lowers the model's predicted pollution level relative to the baseline, while positive values indicate the opposite. This spatial perspective is particularly valuable in environmental and economic research because pollution is inherently spatial, often shaped by regional economic structures, energy systems, and demographic conditions. Translating model outputs into maps facilitates the identification of clusters, regional disparities, and areas where policy interventions may yield the greatest environmental benefits.

Figure 8 presents the spatial distribution of the mean SHAP values for solar power capacity. Provinces with negative SHAP values, concentrated primarily in western and several central areas, indicate that solar deployment is associated with reductions in predicted particulate pollution. This pattern is consistent with the large-scale expansion of solar infrastructure in regions characterized by abundant land availability and favorable solar conditions. In contrast, several eastern provinces display positive SHAP contributions, suggesting that the existing level of solar capacity in these areas is not yet sufficient to offset the pollution generated by dense economic activity and energy demand.

Figure 8: SHAP Map for Solar_Power

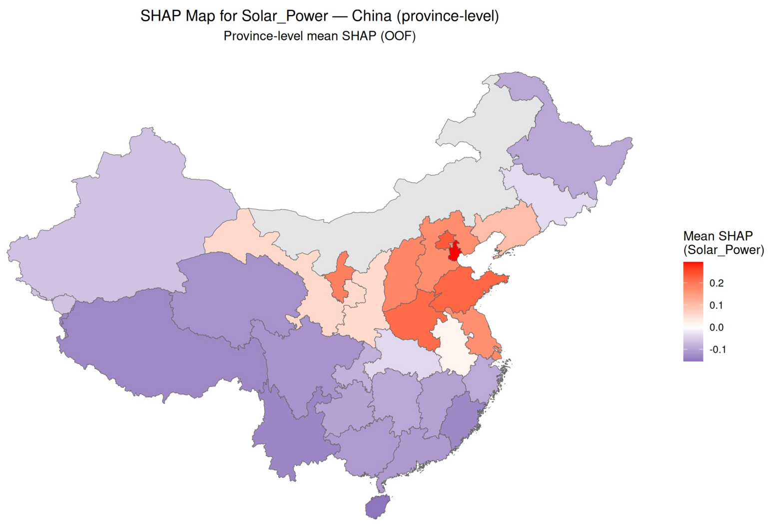
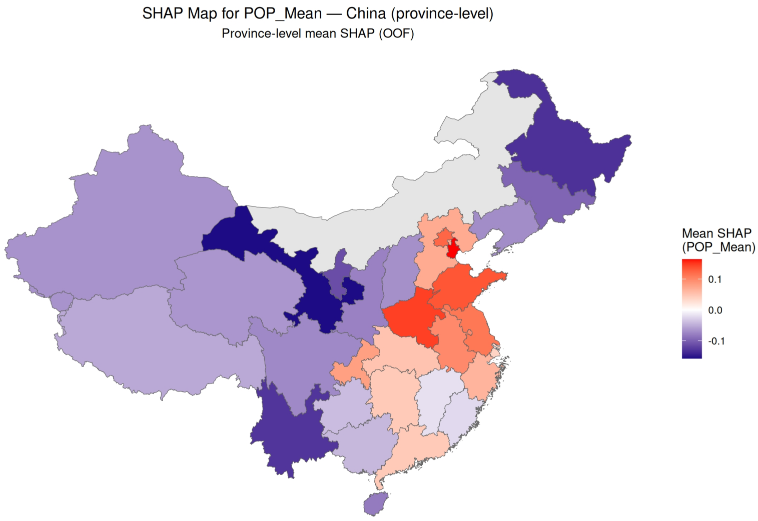


Figure 9 illustrates the geographic contribution of population to predicted pollution levels. Positive SHAP values are heavily concentrated in eastern and coastal provinces, indicating that demographic pressure is a major driver of higher predicted $PM_{2.5}$ concentrations in these areas. These provinces typically combine high urbanization rates with intense transportation flows and industrial activity, all of which contribute to elevated emissions. Conversely, many western provinces exhibit negative SHAP values, reflecting lower

Figure 9: SHAP Map for POP_Mean

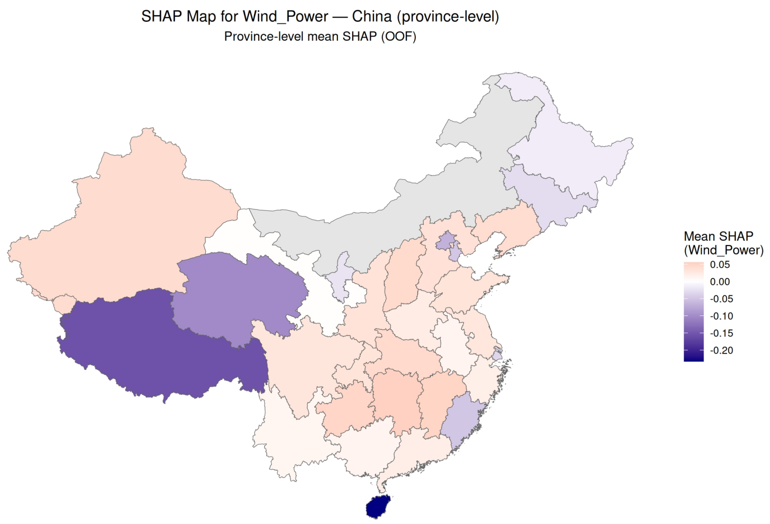


population densities and reduced pollution pressure from human activity.

Finally, Figure 10 exhibits the spatial pattern for wind power. Strong negative SHAP values appear in several western provinces, suggesting that wind generation plays a meaningful role in lowering predicted pollution where it has been extensively deployed. Many central and eastern provinces show values closer to zero or mildly positive, indicating either limited wind capacity or smaller marginal environmental effects. All these maps reveal substantial spatial heterogeneity in the determinants of air quality, with renewable energy appearing more strongly associated with pollution reductions in less densely populated regions, while demographic pressures remain a dominant factor in highly urbanized provinces.

From a policy perspective, these findings support the adoption of geographically targeted environmental strategies rather than uniform national policies. Expanding renewable energy capacity in high-emission eastern provinces could generate substantial marginal improvements in air quality, while continued investment in solar and wind infrastructure in western regions can reinforce their role as sources of cleaner energy. At the same time, densely populated provinces may benefit from policies aimed at reducing urban emissions, including stricter vehicle standards, expanded public transportation systems, congestion management, and incentives for energy-efficient buildings. Strengthening interprovincial electricity transmission could further enhance environmental outcomes by allowing renewable-rich regions to supply cleaner energy to more pollution-intensive areas. Collectively, a spatially differenti-

Figure 10: SHAP Map for Wind_Power



ated approach can improve policy effectiveness by directing resources toward the locations where they are most likely to produce meaningful reductions in particulate pollution.

5 Conclusions

In this study we examine the determinants of particulate matter pollution across Chinese provinces using Extreme Gradient Boosting (XGBoost), a flexible regression approach capable of capturing nonlinear relationships and interaction effects among economic, demographic, energy, and transportation variables. The model demonstrates strong predictive performance, and the consistency between in-sample and out-of-fold results suggests that the findings are not driven by overfitting. Global SHAP values identify solar power, population density, and wind power as the most influential predictors of provincial variation in $PM_{2.5}$. Dependence plots further reveal that renewable energy is generally associated with lower predicted pollution, while demographic pressure tends to increase it. By combining predictive accuracy with interpretability, the analysis provides a structured view of the factors most closely linked to air quality outcomes.

The spatial analysis reinforces these conclusions by showing that the contributions of key predictors vary considerably across provinces. Renewable energy appears to play a particularly important role in mitigating pollution in several western regions, whereas densely populated eastern provinces continue to face stronger upward pressure on $PM_{2.5}$. These results highlight the importance of geographically differentiated environmental policies. Expanding renewable energy in high-emission provinces, strengthening electricity transmission from cleaner regions, and implementing measures to reduce urban emissions could substantially improve air quality. More broadly, the findings illustrate the value of integrating machine learning with spatial analysis to better understand environmental challenges and to support the design of targeted, evidence-based policy interventions.

Several limitations should be considered when interpreting these results. The analysis identifies predictive associations rather than causal relationships, and unobserved factors such as regulatory differences or pollution spillovers may also influence air quality. In addition, the provincial level of aggregation may mask important within-region variation. Future research could extend this framework using finer spatial data and causal identification strategies.

References

Athey, S. and Imbens, G. W. (2019). Machine learning methods that economists should know about. *Annual Review of Economics*, 11(1):685–725.

Chang, T. Y., Graff Zivin, J., Gross, T., and Neidell, M. (2019). The effect of pollution on worker productivity: evidence from call center workers in china. *American Economic Journal: Applied Economics*, 11(1):151–172.

Chen, T. and Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 785–794. ACM.

Chen, Y., Ebenstein, A., Greenstone, M., and Li, H. (2013). Evidence on the impact of sustained exposure to air pollution on life expectancy from china's huai river policy. *Proceedings of the National Academy of Sciences*, 110(32):12936–12941.

Dinda, S. (2004). Environmental kuznets curve hypothesis: a survey. *Ecological economics*, 49(4):431–455.

Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of statistics*, pages 1189–1232.

Grossman, G. M. and Krueger, A. B. (1995). Economic growth and the environment. *The quarterly journal of economics*, 110(2):353–377.

Han, L., Zhou, W., Li, W., and Li, L. (2014). Impact of urbanization level on urban air quality: A case of fine particles (pm2. 5) in chinese cities. *Environmental Pollution*, 194:163–170.

Hastie, T., Tibshirani, R., and Friedman, J. (2009). The elements of statistical learning: data mining, inference, and prediction.

Lin, B. and Zhu, J. (2018). Changes in urban air quality during urbanization in china. *Journal of Cleaner Production*, 188:312–321.

Lundberg, S. M. and Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in neural information processing systems*, 30.

Mullainathan, S. and Spiess, J. (2017). Machine learning: an applied econometric approach. *Journal of Economic Perspectives*, 31(2):87–106.

Pope Iii, C. A., Burnett, R. T., Thun, M. J., Calle, E. E., Krewski, D., Ito, K., and Thurston, G. D. (2002). Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *Jama*, 287(9):1132–1141.

Selden, T. M. and Song, D. (1994). Environmental quality and development: is there a kuznets curve for air pollution emissions? *Journal of Environmental Economics and management*, 27(2):147–162.

Stern, D. I. (2004). The rise and fall of the environmental kuznets curve. *World development*, 32(8):1419–1439.

Wang, X., Tian, G., Yang, D., Zhang, W., Lu, D., and Liu, Z. (2018). Responses of pm2.5 pollution to urbanization in china. *Energy Policy*, 123:602–610.

Zhang, H., Wang, S., Hao, J., Wang, X., Wang, S., Chai, F., and Li, M. (2016). Air pollution and control action in beijing. *Journal of Cleaner Production*, 112:1519–1527.

Zhang, X., Chen, X., and Zhang, X. (2018). The impact of exposure to air pollution on cognitive performance. *Proceedings of the National Academy of Sciences*, 115(37):9193–9197.

Zhou, C., Chen, J., and Wang, S. (2018). Examining the effects of socioeconomic development on fine particulate matter (pm2. 5) in china's cities using spatial regression and the geographical detector technique. *Science of the Total Environment*, 619:436–445.

Zivin, J. G. and Neidell, M. (2012). The impact of pollution on worker productivity. *American Economic Review*, 102(7):3652–3673.

What Drives Air Pollution in China? Evidence from Interpretable Machine Learning and Spatial Analysis of $PM_{2.5}$

令和 8（2026）年 3 月発行

発行所　公益財団法人アジア成長研究所
〒803-0814　北九州市小倉北区大手町 11 番 4 号
Tel：093-583-6202／Fax：093-583-6576
URL：https://www.agi.or.jp
E-mail：office@agi.or.jp