

Spatiotemporal impacts of purpose-specific human mobility on air pollution: Evidence from taxi trajectories and interpretable machine learning

Wenrui XU^a, Xinyue GU^{b,*}

^a School of Architecture, Tsinghua University, Beijing, China

^b Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hong Kong, China

ARTICLE INFO

Keywords:

Human mobility
Air pollution
Taxi trajectory data
Trip purpose inference
XGBoost
SHAP

ABSTRACT

Human mobility exerts significant influences on urban air pollution. Regrettably, most existing studies treated mobility as a homogeneous entity, neglecting that its effects may vary by travel purposes due to distinct spatiotemporal patterns. To address this gap, this study utilizes a trip purpose inference algorithm to classify mobility based on Beijing's three-month taxi trajectory data and examines its impact on air pollution using interpretable XGBoost-SHAP models. The correlational analysis indicates the substantial contribution of wind, temperature, and precipitation to air pollution. Human mobility's contribution is less significant than the abovementioned natural environments but greater than built environments, such as building density and height. In the long term, the negative correlation between work- and home-purpose mobility and pollution challenges the assumption that more mobility always increases pollution. Based on the case study in Beijing, this research eventually proposed possible practical implications and suggestions for sustainable urban planning and management, including promoting mixed-use development and work-residence integration, creating urban wind corridors and open green spaces, and adopting low-emission transportation while avoiding blanket traffic restrictions. This study uses interpretable machine learning models to clarify complex variable relationships, while future research could explore causality to better understand the underlying mechanisms.

1. Introduction

Air pollution has become one of the most severe environmental problems in urban areas worldwide, posing a serious threat to the residents' physical and mental health (Wu et al., 2020). Air pollutants such as PM_{2.5}, PM₁₀, NO₂, SO₂, CO, and O₃ can cause respiratory and cardiovascular diseases, severely affecting the expectancy of lifespan (Fuller et al., 2022; Landrigan et al., 2018). Over the past two decades, air pollution has incurred a significant increase in the number of deaths globally, especially in developing and less developed countries (Cakaj et al., 2023; Niu et al., 2024), with about 2.9 million and 4.5 million premature deaths in 2000 and 2019 respectively, which is the highest among the premature death number caused by all types of pollution (Landrigan et al., 2018).

The origin of urban mobility-environment interactions can be traced to the first industrial cities, where concentrated human activities began systematically impacting local air quality. Many studies have

demonstrated that human mobility has a significant influence on air pollution, particularly during the COVID-19 pandemic when widespread travel restrictions across numerous cities provided valuable opportunities for research (Bao & Zhang, 2020; Cai & Xie, 2007; Fu & Gu, 2017; Ghaffaripasand et al., 2024; Leroutier & Quirion, 2022; Rahman et al., 2021; Wu et al., 2021). For instance, panel data from 44 cities in northern China revealed that government-imposed travel restrictions during the pandemic led to a 69.85 % reduction in human mobility and corresponding decreases in SO₂, PM_{2.5}, PM₁₀, NO₂, and CO by 6.76 %, 5.93 %, 13.66 %, 24.67 %, and 4.58 %, respectively. This reduction in air pollution was strongly associated with a decrease in human mobility (Bao & Zhang, 2020).

In these studies, travel surveys, reports, or vehicle GPS data were primarily used to reflect human mobility (Ghaffaripasand et al., 2024). Among these, taxi trajectory data stands out for its high spatiotemporal resolution (Cai et al., 2014), allowing accurate tracking of daily individual trajectories and uncovering the heterogeneous characteristics of

* Corresponding author at: Block Z, The Hong Kong Polytechnic University, 11 Yuk Choi Rd, Hung Hom, Hong Kong, China.

E-mail address: xinyue.gu@connect.polyu.hk (X. GU).

<https://doi.org/10.1016/j.scs.2025.106411>

Received 16 August 2024; Received in revised form 23 March 2025; Accepted 26 April 2025

Available online 29 April 2025

2210-6707/© 2025 Elsevier Ltd. All rights reserved, including those for text and data mining, AI training, and similar technologies.

travels (Cai & Xu, 2013; Huang et al., 2017; Sui et al., 2019; Zhai et al., 2018). For example, Luo et al. (2017) analyzed travel patterns and emissions using trajectory data from approximately 13,600 taxis in Shanghai over a month, while Xia et al. (2023) leveraged large-scale taxi trajectory data to study the impact of traffic congestion on air pollution at varying times and locations. Regarding air pollution, existing studies mainly rely on ground monitoring data (Bao & Zhang, 2020; Ghaffar-pasand et al., 2024; Wu et al., 2021), remote sensing data, or model estimates derived from such data (Yu et al., 2020).

Contemporary challenges require understanding how differentiated mobility creates distinct environmental impacts because mobility encompasses various types, each exhibiting distinct spatiotemporal distribution patterns (Williams et al., 2012). Existing studies on the relationship between mobility and air pollution primarily focus on the impact of different transportation modes on air quality (Ercan et al., 2022), such as shared mobility (Huang et al., 2022) and electric vehicles (Ferrero et al., 2016). These studies are often conducted at multi-city scales for analysis and comparison (Sharma et al., 2023), with limited attention to finer spatiotemporal scales within cities. Another line of research examines the heterogeneous exposure to air pollution associated with mobility patterns among different population groups (Nyhan et al., 2019; Nyhan et al., 2016; Park & Kwan, 2017; Setton et al., 2011). For instance, Park & Kwan (2017) demonstrated that even individuals staying in an exact location may experience varying exposure levels at different times of the day due to spatiotemporal variations in environmental risk factors and human mobility. Similarly, individuals with differing daily movement patterns may be exposed to different ozone concentrations during the same period. While prior studies have established aggregate mobility-pollution relationships, the spatiotemporal specificity of purpose-driven impacts remains underexplored, particularly in megacities facing simultaneous economic growth and air quality mandates (Bell & Ward, 2000; Yan et al., 2013). Human mobility is inherently tied to its purpose, encompassing the destination and the activities undertaken, which shape unique temporal patterns and spatial associations with various urban functional zones (Schneider et al., 2013).

Based on existing studies on the impact of mobility on air pollution, several urban planning and traffic management policy recommendations have been proposed, including regulating urban travel activities (Wang & Liu, 2014), improving road infrastructure (Asamer et al., 2016), and promoting low-carbon transportation (Wang et al., 2015). For example, Bouscasse et al. (2022) demonstrated through cost-benefit analyses under different scenarios that reducing private vehicle use can significantly lower air pollution-related mortality. In practice, some regions have introduced regulations targeting car usage and urban travel activities. For instance, Europe has implemented "Sustainable Urban Mobility Plans" (Pisoni et al., 2019), while major cities in China, Argentina, and Chile have enforced license plate-based driving restrictions (Zhang et al., 2017) and license plate lotteries (Quan & Xie, 2022) for years. These policies include banning vehicles based on the last digit of their license plate on certain weekdays, limiting the total number of new license plates issued each month, and randomly allocating plates via lottery. However, the effectiveness of such policies is not guaranteed. For example, noncompliance is common, often leading to negligible changes in travel behavior (Wang et al., 2014). Furthermore, some commuters who previously did not drive may begin driving on unrestricted days due to reduced congestion from these policies (Jia et al., 2017). In terms of air pollution, the impact can even be counterproductive. Zhang et al. (2017) found that license plate-based restrictions can reduce NO levels while significantly increasing NO_x and O₃ concentrations. Moreover, blanket restrictions have caused significant inconveniences to citizens' daily lives (Jia et al., 2017).

In summary, the key gap in existing studies lies in their neglect of the heterogeneity of mobility in terms of travel purposes. To deepen our understanding of the intricate interplay between mobility and air pollution and promote sustainable cities, there is an urgent need to

categorize human mobility according to its purposes in order to explore the distinct spatiotemporal distribution characteristics and potential impacts of various types of mobility on air pollution.

To address the identified research gap, this study adopts and enhances a trip purpose inference algorithm to classify different types of human mobility based on long-term taxi trajectory data collected over three months in Beijing, provides a more comprehensive understanding of human mobility patterns, and investigates their spatiotemporally heterogeneous relationships with air pollution using interpretable XGBoost machine learning models. Beijing serves as an invaluable case study because it has been grappling with persistent air pollution for many years, and its urban planning features and the pollution control measures it once implemented share many similarities with those of other megacities globally. Ultimately, the study seeks to propose a refined urban traffic management policy and environmental planning strategy to mitigate urban air pollution, thus advancing the development of smart and sustainable cities (Asamer et al., 2016; Cai & Xu, 2013). The specific objectives of the study are threefold:

- To improve the trip purpose inference algorithm for a better understanding of the heterogeneity of diverse human mobility;
- To investigate the spatial and temporal patterns of human mobility for diverse purposes and compare them with the patterns of different air pollutants;
- To analyze the spatiotemporal impact of diverse human mobility on air pollution.

2. Materials

2.1. Research area

By the end of 2016, Beijing was a major metropolis with a population exceeding 21 million permanent residents, spread across 16 municipal districts encompassing an area of 16.4 thousand square kilometers. The research area is delineated by the 187.6-kilometer-long 6th Ring Road (Beijing Ring Expressway), covering approximately 2267 square kilometers and is home to 90% of the population. The research area includes six urban districts, namely Xicheng, Dongcheng, Haidian, Chaoyang, Fengtai, and Shijingshan, and connects with six suburban districts: Changping, Shunyi, Tongzhou, Daxing, Fangshan, and Mentougou.

The study utilizes traffic analysis zones (TAZ) as the primary unit of analysis, a spatial unit designed explicitly for traffic-related studies and widely adopted by transportation researchers and planning institutions (Feng et al., 2022; Xu et al., 2019). Leveraging road network data from OpenStreetMap (OSM), the research area is divided into 2,213 TAZs, with sizes ranging from 0.017 to 124.4 square kilometers and an average size of 1.025 square kilometers (Fig. 1).

In January 2017, Beijing experienced a significant smog event, marked by a heavy air pollution alert issued at midnight on Dec. 30th, 2016. This event set a historical record, with 25 consecutive days of highly severe air pollution (Gao et al., 2015, 2017; Wang et al., 2015). Therefore, the taxi trajectory data from this period is particularly valuable for this study.

2.2. Research data

The datasets used in this study comprise three parts. The independent variables are different types of human mobility, calculated based on taxi Origin-Destination (OD) and Points of Interest (POI) data. In addition, multi-source urban environmental data are adopted for control variables, including built and natural environmental factors. The dependent variable is the air pollution data.

2.2.1. Independent and control variables

Taxis constitute a significant portion of motorized road traffic activities and contribute substantially to carbon emissions (An et al., 2011;

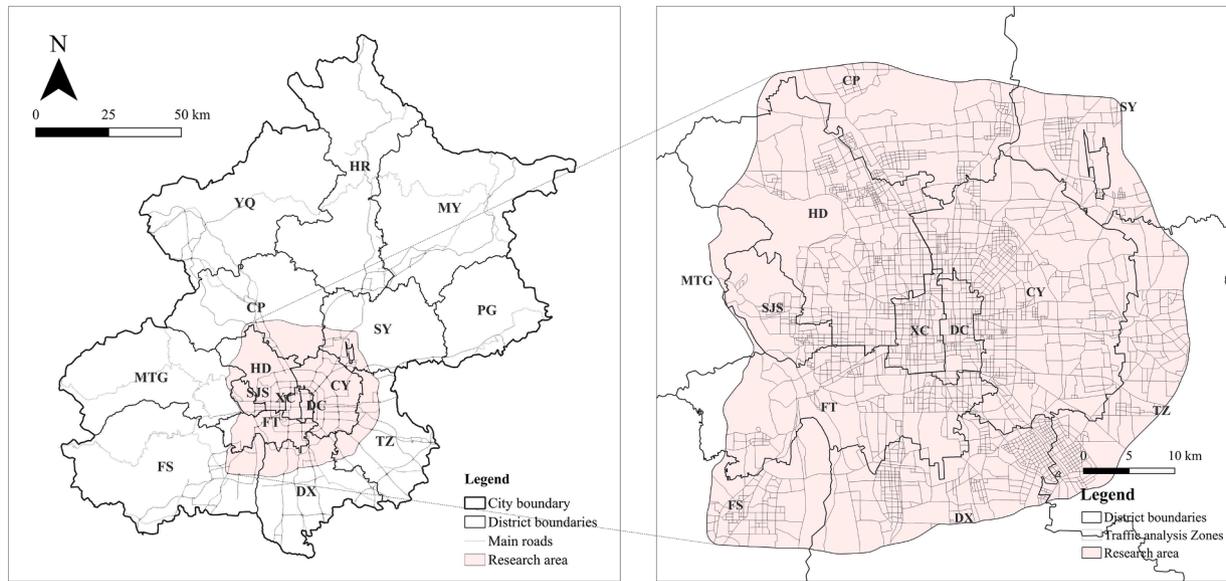


Fig. 1. Research area and traffic analysis zones (TAZs) in the study.

Luo et al., 2017; Wang et al., 2015). Therefore, this study employs taxi trajectory data from Beijing spanning January to March 2017 (a total of 90 days, including 61 workdays and 29 holidays) as an agency to reflect human mobility (Liu, 2017). This data includes GPS positions of approximately 30,000 taxis in Beijing, out of a total fleet of around 69,000 taxis as of late 2016 (Gong et al., 2016; Liu et al., 2017)). The dataset comprises 22,438,854 records, including latitude, longitude, UTC, direction, speed, and information on whether the vehicle is occupied or empty.

The OD data used in this study has been pre-processed by filtering the raw data based on vehicle occupancy to extract the UTC, latitude, and longitude of valid passenger pick-up and drop-off points within the research area. Additionally, 11,745 records that fall outside the study period were removed. The final dataset includes approximately 20.41 million records for the entire city, with about 18.78 million records within the Sixth Ring Road, the study area. Then, POI data was incorporated to develop the algorithm further for identifying trip purposes. This POI data was obtained from the open API of AMap in 2017, a leading map service provider in China known for its comprehensive and accurate mapping details (Xue & Li, 2020).

The urban environment plays a significant role in the distribution of air pollution (Cao et al., 2024; Han et al., 2022; O'Regan et al., 2022). Therefore, to investigate the spatial impact of human mobility on air pollution, additional multi-source urban environmental data were used as control variables, encompassing both built and natural environmental factors (Table 1). The built environment data include building density (BD), building height (BH), and floor area ratio (FAR), which are used to characterize development intensity and urban morphology and are calculated based on building vector data from map service providers (Gu et al., 2024; Li et al., 2021). The natural environment data include the Normalized Difference Vegetation Index (NDVI) and Digital Elevation Model (DEM) data, which characterize the greening level and topography of the city (Yang et al., 2021). The NDVI data were derived from 2017 Landsat remote sensing satellite data with a spatial resolution of 30 m (Yang et al., 2019). The DEM data were sourced from the Copernicus DEM released by the European Space Agency, with a spatial resolution of 30 m (European Space Agency, 2024; OpenTopography, 2021). These data were later aggregated into the TAZ units in GIS for spatial correlation analysis.

On the other hand, weather conditions play a crucial role in the formation and dispersion of air pollutants (Kallos et al., 1993; Yen et al., 2013). Consequently, meteorological data were incorporated as control

Table 1
Control variables adopted in the study.

| Variable | Abbreviation | Description | Unit |
|--|--------------|--|--------|
| Built environmental factors (n=3) | | | |
| Building density | BD | The proportion of the building footprint area relative to the total area of the study unit. | % |
| Building height | BH | The mean height of buildings within the study unit. | m |
| Floor area ratio | FAR | The ratio of the total floor area of all buildings to the total area of the study unit. | None |
| Natural environmental factors (n=6) | | | |
| Normalized Difference Vegetation Index | NDVI | The spatial distribution of green areas within the study unit, as determined by remote sensing data. | NDVI |
| Digital Elevation Model | DEM | The mean elevation of the study unit. | m |
| Daily lowest temperature | DLT | The daily lowest temperature within the study unit. | K |
| Daily precipitation | DP | The total daily precipitation within the study unit. | mm |
| Wind speed | WS | The daily average wind speed within the study unit. | m/s |
| Wind direction | WD | The daily average wind direction within the study unit. | degree |

variables (Han et al., 2022). These variables include daily lowest temperature (DLT), daily precipitation (DP), wind speed, and wind direction (Yen et al., 2013). Preliminary analysis revealed significant multicollinearity between daily highest temperature (DHT) and daily lowest temperature (DLT). As a result, only the DLT was retained in the final models (Table 4). The DLT and DP data were obtained from the High-Resolution (1 day, 1 km) and Long-Term (1961–2019) Gridded Dataset for Temperature and Precipitation across China (HRLT) dataset (Qin & Feng, 2022; Qin & Zhang, 2022). Wind data were obtained from the ERA5-Land post-processed daily statistics spanning from 1950 to the present (Copernicus Climate Change Service, Climate Data Store, 2024). The raw data includes daily mean wind speeds in the u- and v-directions. Through data processing, the daily average wind speed and wind direction were calculated, with the wind direction expressed as the angle measured clockwise from true north.

2.2.2. *Dependent variables*

Six pollutants (PM_{2.5}, PM₁₀, NO₂, SO₂, CO, and O₃) are commonly used to evaluate air quality (Xiao et al., 2018). In this study, the daily data of air pollution in Beijing during the research period is sourced from the High-resolution and High-quality Ambient Air Pollutants Dataset for China (CHAP). This dataset is developed using a combination of multiple big data sources, including ground-based measurements, satellite remote sensing products, atmospheric reanalysis, and model simulations to account for the spatiotemporal heterogeneity of air pollution (He et al., 2022; Wei et al., 2023; Wei et al., 2020, 2021; Wei, Li, et al., 2022; Wei, Liu, et al., 2022). The CHAP dataset includes major air pollutants such as PM_{2.5}, PM₁₀, NO₂, SO₂, CO, and O₃, with a daily temporal resolution and a spatial resolution of 1 km (10 km for NO₂, SO₂, and CO). It has been extensively applied and validated across various fields, including medical health, environmental science, atmospheric science, and remote sensing. Given the significant differences in the spatiotemporal patterns of these pollutants, this study does not combine them into a single comprehensive index but instead analyzes each pollutant individually.

3. **Methods**

The methodological framework of this study is illustrated in Fig. 2. First, POIs were reclassified into nine types according to travel purposes (Gong et al., 2016; Li et al., 2021; Liu et al., 2023; Zhao et al., 2017). Next, the trip purpose inference algorithm (Furletti et al., 2013; Li et al., 2021) was optimized and employed to categorize human mobility based on the POI data. Then, we analyzed the spatiotemporal distribution patterns of diverse human mobility and compared them with those of air pollution. Finally, we established a total of 18 XGBoost-SHAP models to assess the importance and correlation of eighteen variables, including human mobility, built environment, and natural environmental factors, with six air pollutants.

3.1. *Optimized trip purpose inference algorithm*

The essential idea of the trip purpose inference algorithm is to

associate each drop-off point with nearby POIs based on the premise that passengers typically alight from taxis near their destination and then walk to engage in activities there (Furletti et al., 2013; Gong et al., 2016; Qian & Ukkusuri, 2015; Tian et al., 2024; Xing et al., 2020). The POIs are pre-classified and mapped into nine categories based on activity types (Table 3): home, work, transport, shopping, dining, leisure, education, medical, and other daily activities, following previous studies in the field (Jiang et al., 2012; Li et al., 2021; Liu et al., 2023; Zhao et al., 2023). The algorithm for each drop-off point involves three main steps, as illustrated in Fig. 3: 1) filtering candidate POIs based on walking distance and operating hours; 2) calculating the visit probability for each candidate POI by considering distance, time, and the proportion of this specific POI type; 3) inferring the trip purpose by summing the probabilities of the corresponding POI points for each activity type.

Step 1: Filtering the candidate POIs

The first step involves filtering the candidate POIs based on spatial and temporal conditions (Zhao, 2017). The spatial condition is that the distance between the POI P_i and the drop-off point D should be within walkable distance WD . The temporal condition requires that the POI be open around the drop-off time t . WD represents the maximum distance within which most drop-off points should be able to find at least one POI, reflecting passengers' willingness to walk to the final destination (Li et al., 2021). In this study, WD is set as a constant value of 100 m based on preliminary results, as described in Fig. 4.

The opening hours H of each POI are determined based on the type of POI and whether the drop-off time falls on a workday or a holiday (Li et al., 2021; Zhao, 2017), as detailed in Appendix Table A.1. The filtering process is illustrated in Fig. 3. For example, if there is a drop-off point at 12 p.m. on a weekend, candidate POIs are represented as black dots, while non-candidate POIs are shown as gray dots.

$$CandidatePOI = \{P_i \mid d(P_i, D) \leq WD, t \in H\} \tag{1}$$

Given a list of POIs, those that meet both the spatial and temporal conditions are retained as candidate POIs for subsequent calculations. If no POI satisfies these conditions, the final probability for every activity

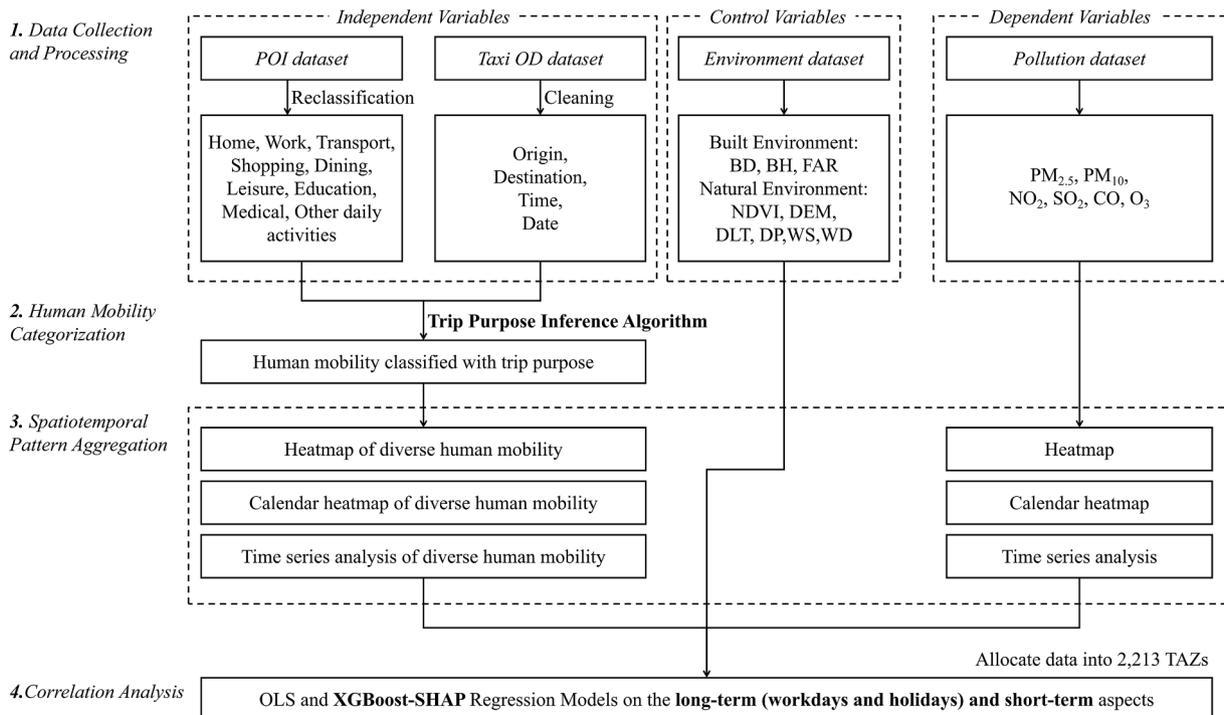


Fig. 2. The methodological framework of the study.

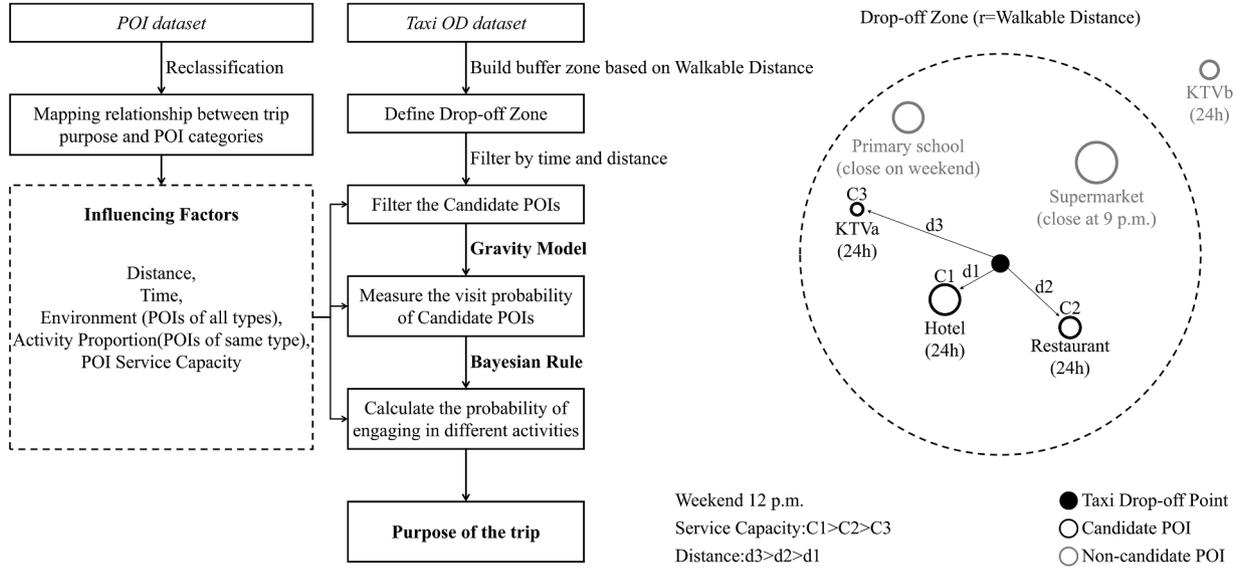


Fig. 3. Trip purpose inference algorithm framework.

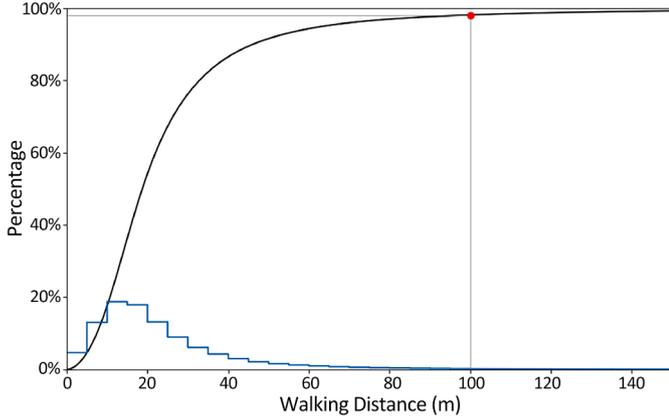


Fig. 4. Percentage of trip destinations that contain at least one POI within different walking distances.

associated with this OD record is set to zero.

Step 2: Measuring the visit probability of candidate POIs

The second step involves calculating the visit probability of each candidate POI, primarily based on the Gravity Model and Bayesian rules, taking into account factors such as distance, time, POI quantity, and capacity (Gong et al., 2016). Given a drop-off point D and the associated candidate POIs, the visit probability for each candidate POI P_i is defined as follows.

$$Pr(P_i|D, t) = \frac{Pr(D, t|P_i) \times Pr(P_i)}{Pr(D, t)} = \frac{Pr(D|P_i) \times Pr(t|P_i) \times Pr(P_i)}{Pr(D, t)} = Pr(D|P_i) \times Pr(P_i|t) \quad (2)$$

$Pr(P_i|D, t)$ denotes the probability that a passenger will visit P_i if dropped off at location D at time t . Given that the location D and time t can be considered conditionally independent, $Pr(D, t|P_i)$ can be expressed as the product of $Pr(D|P_i)$ and $Pr(t|P_i)$, and $Pr(D, t)$ equals to the product of $Pr(D)$ and $Pr(t)$. For a certain drop-off point, $Pr(D)$ is 1. $Pr(D|P_i)$ represents the probability that a passenger will drop off at location D if intend to visit POI P_i , which can be measured using the Gravity Model (Formula 3). The value of $Pr(P_i|t)$ is derived from prior studies on

the daily patterns of human activities in the city, with distinctions made between workdays and holidays according to the specific activity type (Jiang et al., 2012).

$$Pr(D|P_i) \propto G(D, P_i) = A_D A_i d(P_i, D)^{-\beta} = \rho_i C_i d(P_i, D)^{-1.5} \quad (3)$$

$$\rho_i = \frac{\varphi(P_i, activity)}{\sum_{j=1}^9 \varphi(activity_j)}, \quad j \in \{1, 2, \dots, 9\} \quad (4)$$

$$\varphi(activity_k) = \frac{\text{number}(candidatePOI_{activity=k})}{\text{number}(POI_{activity=k})} \quad (5)$$

In the Gravity Model, the distance decay coefficient β is set as 1.5, according to previous studies (Li et al., 2021; Zhao, 2017). A_D and A_i are the attractiveness of D and P_i , in which A_D is set as 1 to facilitate calculation, and A_i considers two factors: firstly, ρ_i is the distribution density of P_i 's type of activity (formula 4, 5); secondly, C_i is the service capacity of P_i . In this study, we adopted the service capacity index of different types of POI proposed by Li et al. (2021), as detailed in Appendix Table A.1. Therefore, given a certain drop-off point and time, the probability of visiting each candidate POI is:

$$Pr(P_i | D, t) = \frac{G(D, P_i) \times Pr(P_i|t)}{\sum_{j=1}^n G(D, P_i) \times Pr(P_j|t)} \quad (6)$$

Step 3: Inferring the purpose of the trips

Given a drop-off point and time, the probability of engaging in different activities can be calculated as formula 7. Finally, the trip purpose of each record of OD data has nine values representing the probability of engaging in different types of activities, and their sum is 1.

$$Pr(activity_i) = \sum_{j=1}^n Pr(P_j|D, t), \quad P_j \in \{candidatePOI_{activity=i}\} \quad (7)$$

Compared to existing research, this algorithm in this study has two key enhancements. Firstly, considering that different activities are likely to happen during different times within a day, the possibility of different daily activities through different times within a workday or holiday is taken into account (Huang et al., 2010; Jiang et al., 2012). Secondly, considering that the passengers are likely to visit multiple places after getting out of the taxi, the output is divided into nine values representing various types of activities rather than one (Li et al., 2021).

3.2. Interpretable machine learning models

To investigate the relationship between human mobility and air pollution, this study constructs regression models from both long-term and short-term perspectives.

From a long-term perspective, the study examines the distribution of daily average density of human mobility and pollution concentration across different areas over the entire research period, using TAZs as the unit of analysis. OD points are linked to TAZs through spatial joins, and the density of human mobility for diverse purposes within each TAZ is calculated and used as the independent variable in the models. Control variables are listed in Table 1. The dependent variables are the concentrations of six pollutants, PM_{2.5}, PM₁₀, NO₂, SO₂, CO, and O₃, resulting in six regression models to compare outcomes. Due to significant differences in human mobility patterns between workdays and holidays, separate analyses are conducted.

From a short-term perspective, the study focuses on the daily variations in human mobility density and pollution concentration within each TAZ unit, establishing regression models that use both TAZ and date as the units of analysis. The human mobility density within each TAZ unit on a given day is treated as the independent variable. The concentrations of the six pollutants within each TAZ unit on that day serve as the dependent variables.

3.2.1. OLS multiple linear regression model

Initially, the Ordinary Least Squares (OLS) model is employed to establish a multiple linear regression framework and perform multicollinearity testing to select independent variables. Multicollinearity is assessed using the Variance Inflation Factor (VIF). A VIF value less than 7.5 indicates that multicollinearity is not present, allowing the variables to be used in the regression model. The model's fit is evaluated using the adjusted R-squared (R²) statistic. In cases where the model demonstrates poor fit, nonlinear models are considered to better capture the intricate relationships among the variables.

3.2.2. XGBoost machine learning regression model

Among the various nonlinear models available, this study selected the eXtreme Gradient Boosting (XGBoost) machine learning regression model, as proposed by Chen and Guestrin (2016), after comparing it with several commonly used alternatives. XGBoost is an enhanced version of the Gradient Boosting Decision Tree (GBDT) and offers more efficient model training (Gu, et al., 2024). The dataset was divided into training and validation sets, with 75% allocated to training and 25% to validation. Grid search parameter tuning with cross-validation was then conducted to identify the optimal parameters for the model, which were subsequently used to establish the model. The model's performance was evaluated using adjusted R-squared and Root Mean Square Error (RMSE), with the results presented in Table 4.

3.2.3. Interpretation of nonlinear relationships with SHAP

While machine-learning models demonstrate strong performance and can partially indicate the relative importance of each feature, they often fall short in elucidating the precise impact of different independent variables on the prediction outcomes (Gu et al., 2024; Yang et al., 2024). To further clarify the contribution of each variable and its positive or negative effects within the model, this study introduces SHapley Additive exPlanations (SHAP) to interpret the models.

Drawing inspiration from cooperative game theory, SHAP constructs an additive explanation model, treating all features as "contributors" (Lundberg & Lee, 2017; Yang et al., 2024). For each predicted sample, the model generates a prediction value, and the SHAP value represents the contribution of each feature to that prediction. Given the sample i is x_i , the feature j of x_i is x_{ij} , the model's prediction for this sample is y_i , and the baseline prediction for the entire model (typically the mean of the target variable across all samples) is y_{base} , the SHAP value is determined by the following equation.

$$y_i = y_{base} + f(x_{i1}) + f(x_{i2}) + \dots + f(x_{ij}) \quad (8)$$

$f(x_{ij})$ is the SHAP value of x_{ij} , which represents the contribution of x_{ij} to the final predicted value y_i . When $f(x_{ij}) > 0$, it indicates that the feature increases the predicted value; otherwise, the feature reduces the predicted value.

4. Results

4.1. Spatial patterns of human mobility and air pollution

This section mainly examines the long-term spatial distribution of daily average human mobility density and air pollution concentration within three months. The mobility density is calculated by summing the total amount of activities within each TAZ unit and dividing by the unit's area. The results are grouped using the natural breaks classification method to visualize.

Figs. 5 and 7 illustrate the spatial distribution of human mobility density on workdays and holidays, showing a clear gradient from the city center to the suburbs, with most activity concentrated within the Fourth Ring Road. The Global Moran's I analysis found that human mobility and air pollution both showed clustered distributions at a very high confidence level ($p < 0.001$). The deep-colored areas on the maps represent the highest category of mobility density, and the descriptive statistics in the study area are summarized in Table 2.

Further determination of the locations of cold and hot spot clusters was made through the Getis-Ord Gi* analysis. On workdays, home-purpose mobility is predominantly concentrated in residential areas within the Fourth Ring Road, with hot spots extending to suburban residential areas such as Huilongguan, Tiantongyuan, and Wangjing. Work-purpose mobility is concentrated in major business districts and technology parks, including the CBD, Lize, Shangdi, and Zhongguancun. Transport-purpose mobility primarily aligns with subway lines and major transportation hubs, such as Beijing South Railway Station and the Capital Airport. Shopping- and dining-purpose mobility clusters in commercial areas like Wangfujing, Sanlitun, Xidan, and Guomao in the city center. Leisure-purpose mobility concentrates around large leisure facilities, such as the Olympic Park. The spatial distributions of education- and medical-purpose mobility correspond to the locations of universities and hospitals, with higher concentrations in the Haidian District, home to several prominent universities, and areas around Peking Union Medical College Hospital and Tiantan Hospital. These patterns suggest a strong correlation between human mobility and urban centrality, with the high-value areas of different mobility types closely aligning with the distribution of corresponding functional zones.

On holidays, the overall spatial distribution patterns of human mobility remain broadly consistent with those on workdays (Fig. 7), with significant changes observed primarily in terms of quantity (Table 2). The density of work-, transport-, and education-purpose mobility decreases notably during holidays, while dining- and leisure-purpose mobility increase significantly.

Figs. 6 and 8 depict the spatial patterns of six air pollutants, which are generally concentrated in the southeast and lower in the northwest. The average and maximum values and standard deviation within the study area are summarized in Table 2. On workdays, PM_{2.5} concentrations are elevated in the southern and eastern regions, while localized cold spots appear in the northwest areas within the Second and Third Ring Roads. Since PM₁₀ is primarily composed of PM_{2.5}, its distribution closely mirrors that of PM_{2.5}. A localized hot spot for PM₁₀ is observed near the northeastern Second Ring Road, while cold spots are observed in the central urban area within the Fourth Ring Road and the northern region of the Fifth Ring Road. NO₂ concentrations are higher in the southern areas and lower in the north, with notable hot spots near Zhongguancun and Qinghe, close to the North Sixth Ring Road. Both SO₂ and CO exhibit higher concentrations in the southern suburbs and

Workday

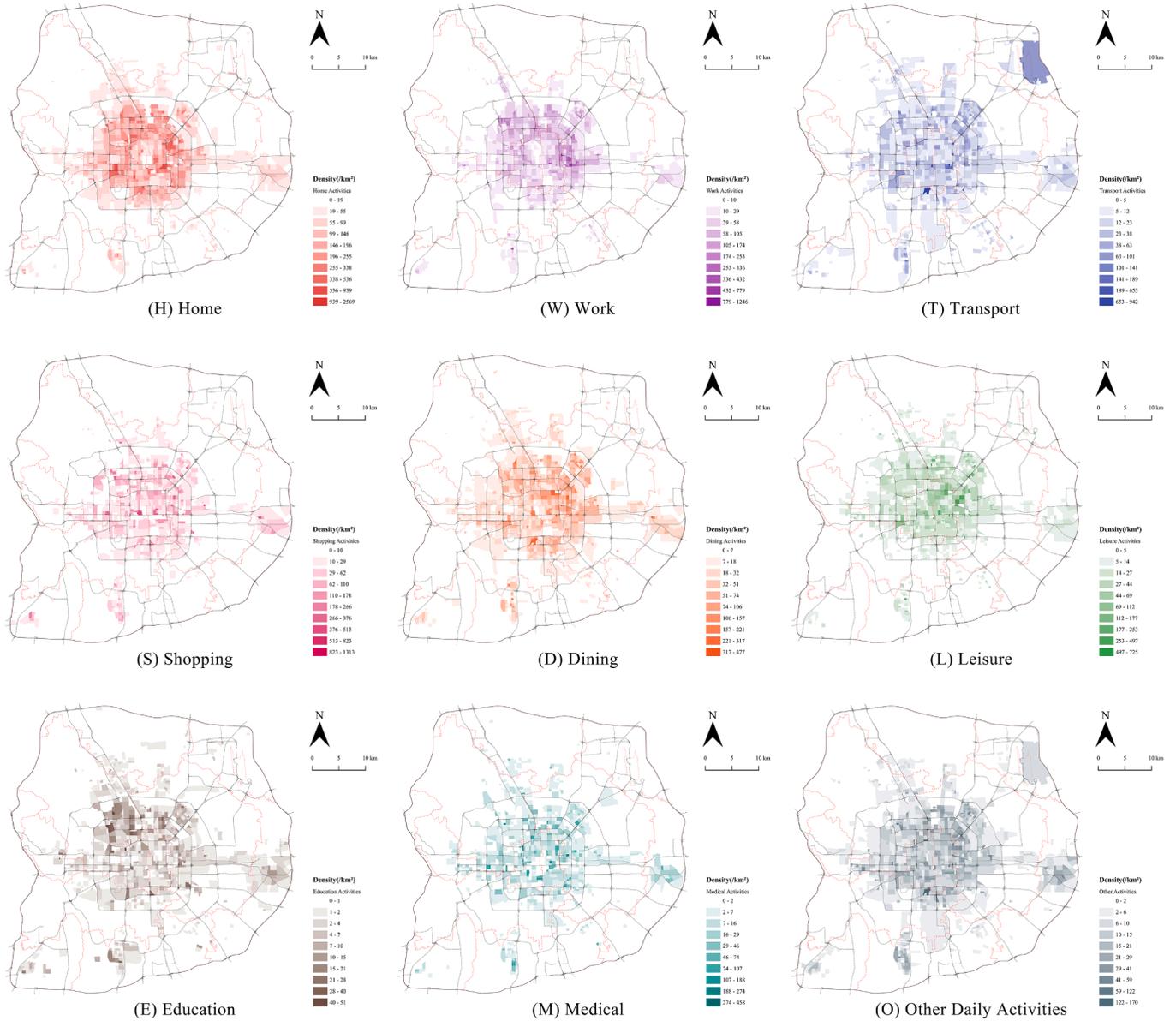


Fig. 5. Spatial distribution of daily average density of human mobility on workdays.

Table 2

Descriptive statistics of human mobility density and air pollutant concentration in the study area.

| | | Workday | | | Holiday | | |
|----------------|--|---------|----------|---------|---------|----------|---------|
| | | Ave. | Max. | S.D. | Ave. | Max. | S.D. |
| Human Mobility | Home (/km ²) | 74.203 | 3165.834 | 129.593 | 57.859 | 2569.461 | 104.449 |
| | Work (/km ²) | 104.744 | 3730.244 | 227.040 | 26.443 | 1246.34 | 60.799 |
| | Transport (/km ²) | 21.379 | 1611.088 | 61.130 | 12.917 | 941.822 | 36.819 |
| | Shopping (/km ²) | 21.528 | 1301.838 | 64.841 | 18.965 | 1313.371 | 58.274 |
| | Dining (/km ²) | 16.457 | 419.782 | 30.338 | 18.556 | 477.241 | 34.166 |
| | Leisure (/km ²) | 10.342 | 823.373 | 30.562 | 12.716 | 724.504 | 35.167 |
| | Education (/km ²) | 7.737 | 193.412 | 14.996 | 1.994 | 50.946 | 4.408 |
| | Medical (/km ²) | 8.468 | 743.789 | 33.014 | 5.191 | 457.767 | 18.023 |
| | Other (/km ²) | 2.967 | 90.490 | 5.351 | 6.869 | 170.347 | 11.301 |
| Air Pollutant | PM _{2.5} (µg/m ³) | 84.530 | 90.635 | 2.974 | 96.110 | 102.706 | 3.427 |
| | PM ₁₀ (µg/m ³) | 119.106 | 129.289 | 5.629 | 131.409 | 143.168 | 6.553 |
| | NO ₂ (µg/m ³) | 60.873 | 66.467 | 2.920 | 56.278 | 61.897 | 3.087 |
| | SO ₂ (µg/m ³) | 17.006 | 22.611 | 2.079 | 20.707 | 26.217 | 1.945 |
| | CO (mg/m ³) | 1.585 | 1.838 | 0.096 | 1.508 | 1.848 | 0.116 |
| | O ₃ (µg/m ³) | 64.057 | 74.893 | 3.607 | 68.414 | 84.553 | 3.114 |

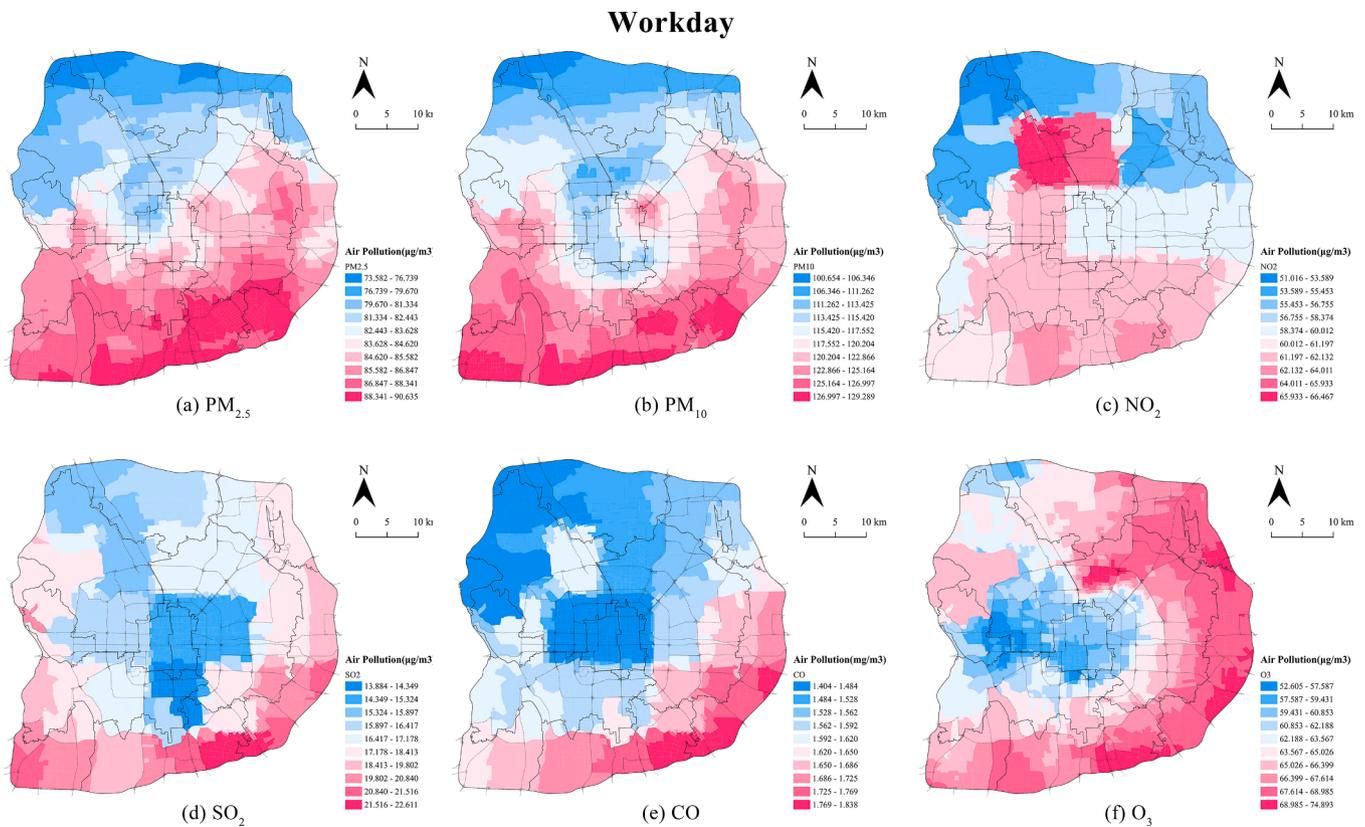


Fig. 6. Spatial distribution of daily average concentration of air pollution on workdays.

relatively lower levels in central urban areas. In contrast, O_3 concentrations are higher in the northeastern and southeastern regions, with a prominent hot spot around Capital International Airport, located near the northeast side of the Fifth Ring Road.

During holidays, the average values for $PM_{2.5}$, PM_{10} , SO_2 , and O_3 are higher than those observed on workdays, while NO_2 and CO are lower than on workdays. Notably, areas with elevated concentrations of PM_{10} and O_3 are more concentrated near the Capital International Airport on holidays. At the same time, the distribution patterns of other pollutants remain largely consistent with those observed on workdays.

Juxtaposing the spatial patterns of human mobility against those of air pollution reveals a lack of conspicuous positive correlation between the two. Human mobility generally decreases outward from the city center to the suburbs, while air pollution often manifests as a relative void in central urban areas. However, in particular business districts, science and technology parks, and large transportation hubs, there is a noticeable synchronization of local hot spots, where travel activities and air pollution concentrations are elevated.

In addition, this study does not provide exhaustive details on the short-term spatial distribution of each activity for each day due to our primary focus on long-term patterns, limitations in paper length, and the fact that distribution patterns observed on a single day or even within an hour generally exhibit strong consistency with long-term patterns, differing mainly in terms of quantity.

4.2. Temporal patterns of human mobility and air pollution

This section aggregates the total volume of human mobility across the entire research area to analyze the temporal patterns (Table 3). The calendar heat map (Fig. 9) visualizes the daily distribution of human mobility relating to various activities, while the time series plot (Fig. 11) presents the hourly distribution, revealing detailed variations throughout the day. For comparison, trends in air pollution are also visualized with these plots (Figs. 10, and 11). Although issues with the

raw taxi data affected the accuracy of hourly variations for five days, these data still accurately reflect the average daily levels of human mobility (Fig. 11).

The total volume of different types of human mobility (Table 3) aligns with existing studies and validates the accuracy of the trip purpose inference algorithm (Li et al., 2021; Zhao, 2017). On workdays, the proportions of home and work-purpose mobility are significantly higher than those of other mobility types. On holidays, the volume of work and education-purpose mobility decreases markedly while home, shopping, dining, leisure, and other daily activities increase.

Fig. 9 shows that work-, transport-, education-, and medical-purpose mobility exhibit a pronounced midweek peak, whereas dining, leisure, and other daily activities peak on holidays. The research period spans the Chinese Spring Festival, during which all types of mobility significantly decreased (January 27 - February 2).

The time-series graph (Fig. 11) delineates a consistent daily rhythm in all mobility types, characterized by pronounced peaks during the morning and evening commutes. However, there are discernible variations between the mobility patterns on workdays versus holidays and between standard weekends and the Chinese Spring Festival period.

The calendar heat map (Fig. 10) and time-series graph (Fig. 11) reveal that the concentration of six pollutants follows similar fluctuating patterns, cycling approximately every one to two weeks. $PM_{2.5}$, PM_{10} , NO_2 , and CO exhibit analogous distribution characteristics, with elevated levels at the beginning of the year persisting for over a week. Then, the concentrations of these pollutants, along with SO_2 , show a significant increase on Jan. 28th and after the Chinese Spring Festival (Jan. 28th - Feb. 2nd). In contrast, O_3 displays a gradual, long-term rise from January to March, with prominent peaks observed on Mar. 18th and 19th.

4.3. Impacts of human mobility on air pollution

This section develops two levels of regression models, resulting in a

Holiday

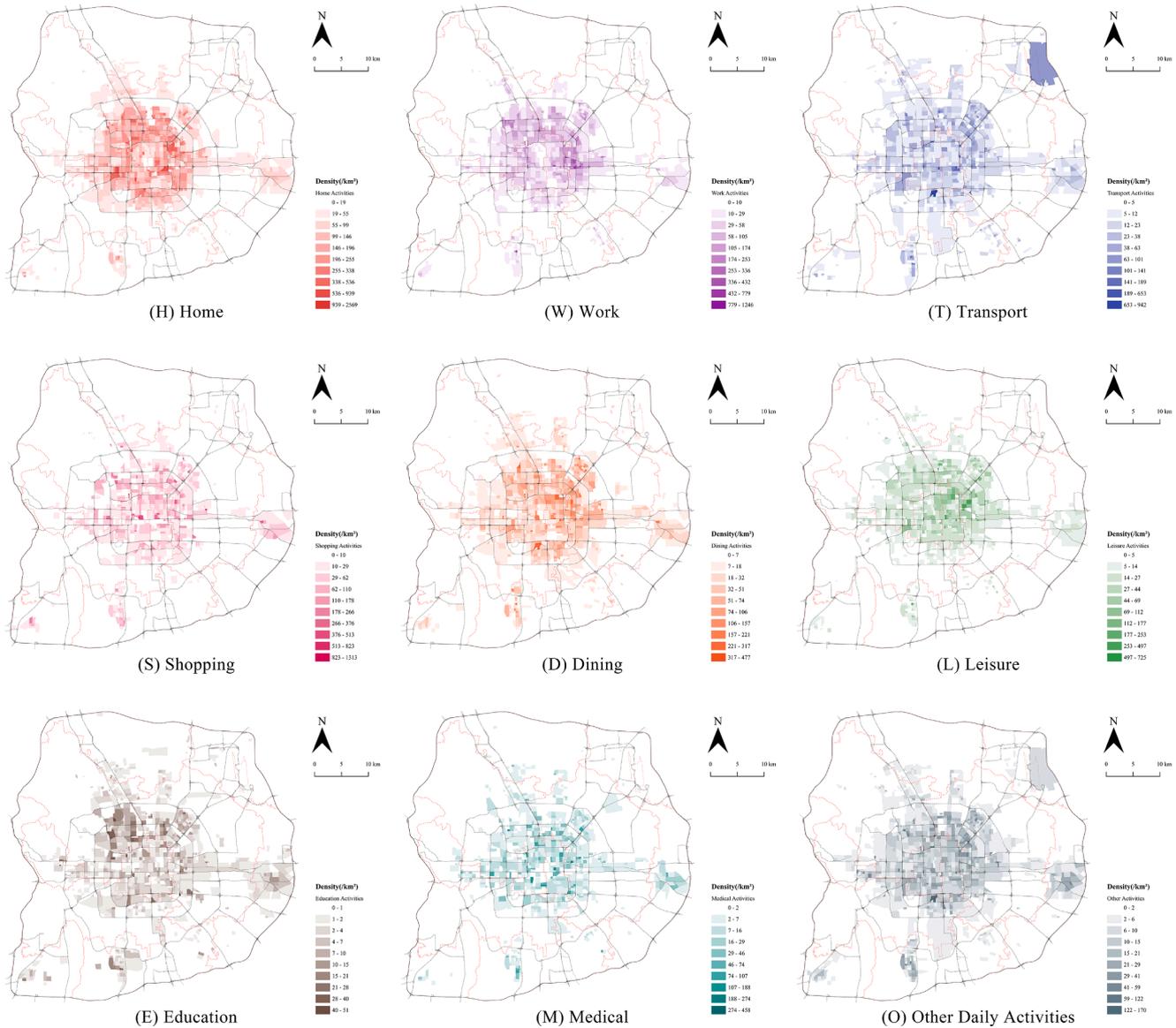


Fig. 7. Spatial distribution of daily average density of human mobility on holidays.

total of 18 models: 1) Long-term models that focus on the spatial distribution across different TAZs over the entire research period using aggregated daily data, with each model comprising 2,213 entries; 2) Short-term models that examine the spatiotemporal distribution for each day and each TAZ, with each model comprising 199,170 entries.

4.3.1. Comparison of the performances of regression models

First, the Ordinary Least Squares (OLS) model was employed for correlation analysis and multicollinearity testing. Based on the Pearson Correlation Coefficient and Variance Inflation Factor (VIF), all the variables are significantly correlated and free of multicollinearity.

However, as a classical linear regression model, OLS is limited to depicting linear relationships and struggles to capture non-linear interactions. The model's fit was assessed using the adjusted R-squared value, revealing that the results for models 1-1-c, 1-2-c, 1-2-d, and the first five short-term models were relatively low (Table 4). Consequently, XGBoost regression models were developed to enhance model performances. The regression evaluation metrics in Table 4 demonstrate that the XGBoost models achieved higher adjusted R-squared values and

lower Root Mean Square Errors (RMSE).

4.3.2. Relative importance of different variables

To interpret the XGBoost models, this study employed the Tree Explainer, a SHAP interpreter specifically designed for tree-based models, including the XGBoost models. The Tree Explainer was used to compute the Shapley values for each variable. Subsequently, the importance of each variable was ranked based on the average of the absolute Shapley values, and this ranking is visualized using bar plots (Fig. 12).

The contributions of human mobility to air pollution consistently rank fourth or lower across all 18 models (in both long-term and short-term and for all six pollutants). In contrast, natural environmental factors such as wind speed (WS), wind direction (WD), daily lowest temperature (DLT), daily precipitation (DP), DEM, and NDVI were consistently ranked as the most important variables in all models, with their mean absolute SHAP values significantly exceeding those of other variables. Built environmental factors, including building density (BD), building height (BH), and floor area ratio (FAR), typically rank just

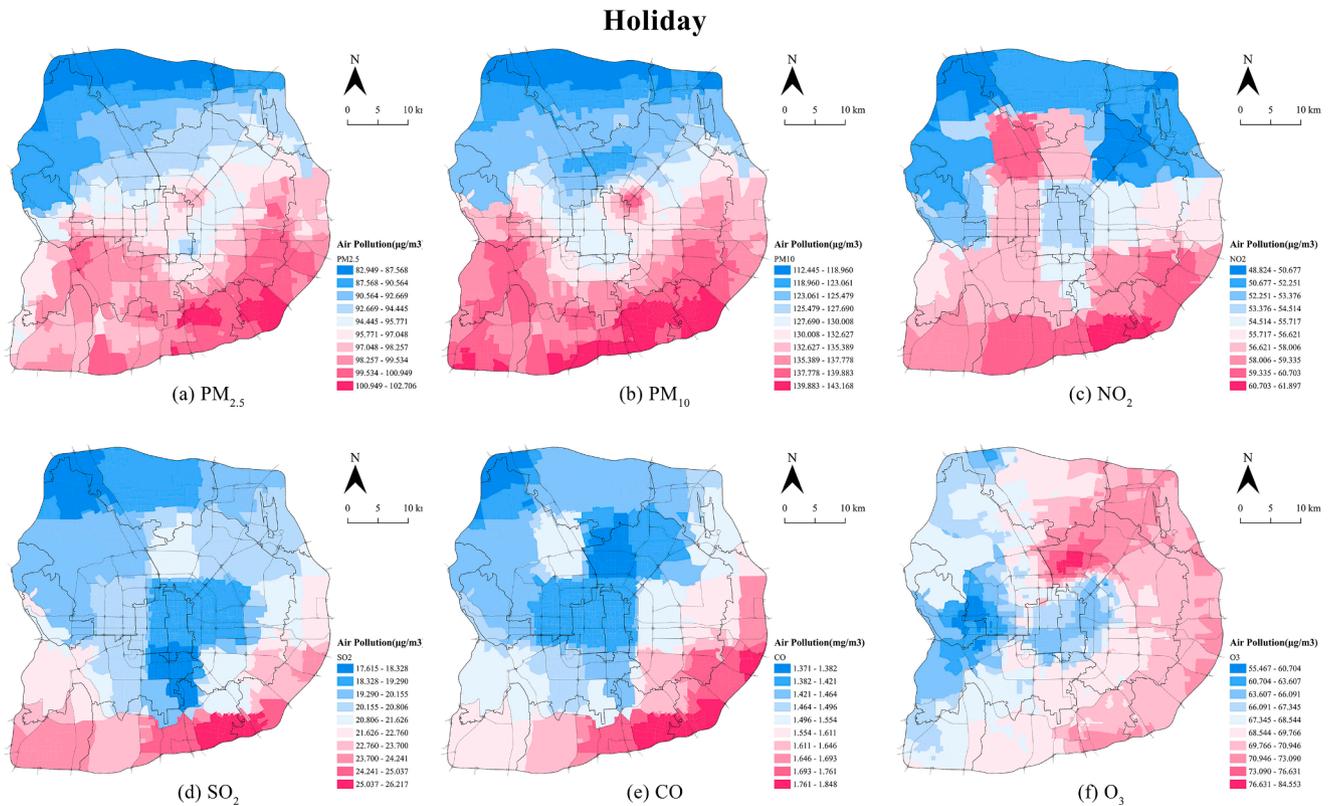


Fig. 8. Spatial distribution of daily average concentration of air pollution on holidays.

Table 3
Average daily volume and proportion of diverse human mobility on 61 workdays and 29 holidays.

| Activity | On Workdays | On Holidays |
|----------------------------|----------------|----------------|
| Home (H) | 75,479 (30.2%) | 59,260 (36.2%) |
| Work (W) | 79,986 (32.0%) | 22,082 (13.5%) |
| Transport (T) | 25,391 (10.1%) | 16,163 (9.9%) |
| Shopping (S) | 20,803 (8.3%) | 19,129 (11.7%) |
| Dining (D) | 17,768 (7.1%) | 19,370 (11.8%) |
| Leisure (L) | 9,536 (3.8%) | 11,594 (7.1%) |
| Education (E) | 8,907 (3.6%) | 2,542 (1.6%) |
| Medical (M) | 8,769 (3.5%) | 5,586 (3.4%) |
| Other daily activities (O) | 3,698 (1.5%) | 8,033 (4.9%) |
| Total | 250,337 (100%) | 163,759 (100%) |

below the relatively high-importance mobility variables, placing them in the mid-range among all 18 variables.

Although the mean absolute SHAP value of a single mobility variable was not as high as that of natural environmental factors, the combined mean SHAP values of the nine different types of mobility were still not negligible, as they exceeded those of built environmental factors. Furthermore, the mean SHAP values of mobility in long-term models were significantly higher than those in short-term models.

Across all models, home- and work-purpose mobility consistently ranked as the most important among the different types of mobility. Beyond these two variables, the importance of other mobility types varied depending on the model. In long-term workday models, transport-purpose mobility was relatively important in NO₂ and CO models, shopping-purpose mobility was prominent in SO₂ and O₃ models, leisure-purpose mobility was significant in PM_{2.5} models, and medical-purpose mobility showed higher importance in NO₂ models. In long-term holiday models, leisure-purpose mobility and shopping-purpose mobility were relatively more important. In short-term

models, the importance was primarily concentrated on home- and work-purpose mobility, with education-purpose mobility being notable in PM_{2.5}, PM₁₀, and SO₂ models.

4.3.3. The impacts of diverse human mobility on air pollution

This study further utilizes SHAP values to assess the impact of each variable on air pollution within the XGBoost models. As shown in Fig. 13, the x-axis position represents the SHAP values for each variable, indicating the variable's impact value on the model's output in each case. If the point is on the right side of the y-axis, it means a positive effect on the prediction. Conversely, it has a negative impact on the prediction if it is on the left. The color coding reflects the actual value of each variable. Vertically, the variables are ranked according to their relative importance, as illustrated in Fig. 12.

The local dependence plots (Fig. 14) illustrate the non-linear effects of human mobility with the highest relative importance on the concentration of six air pollutants. The three plots for each model are arranged from left to right according to the relative importance of the

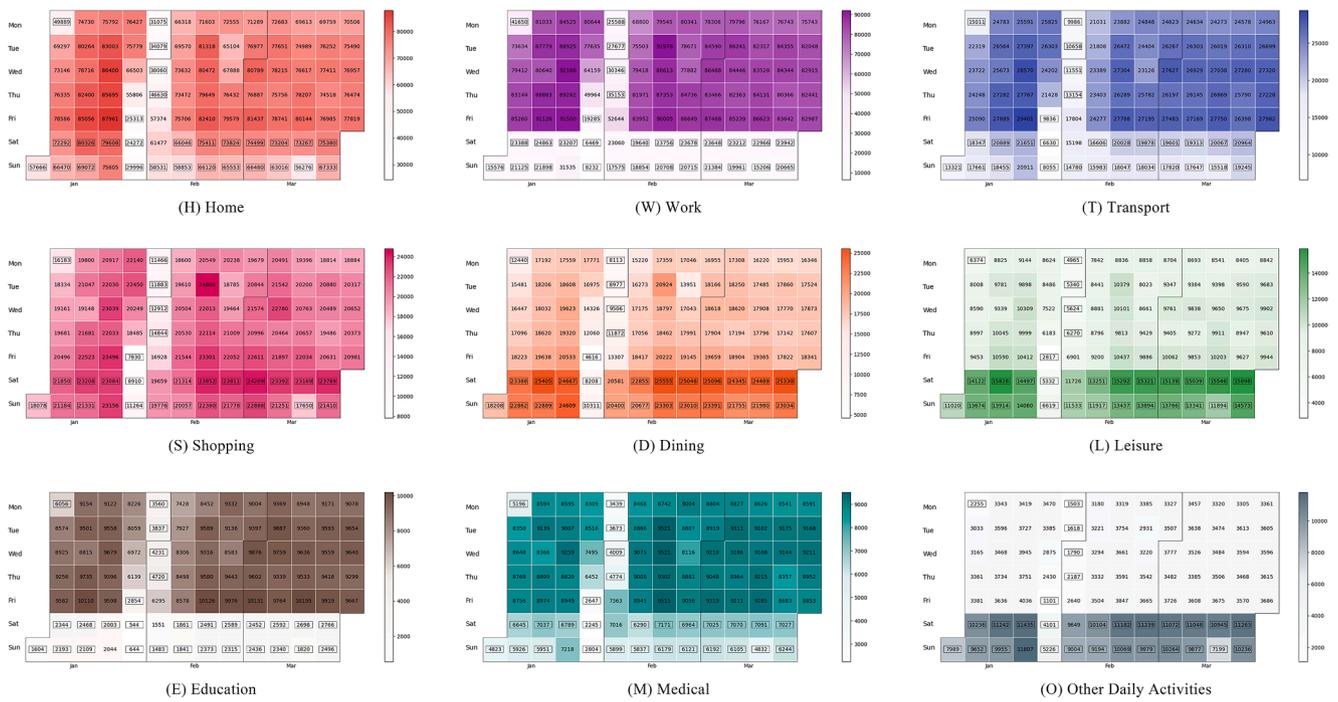


Fig. 9. Calendar heat map of human mobility (holidays highlighted with black frames).

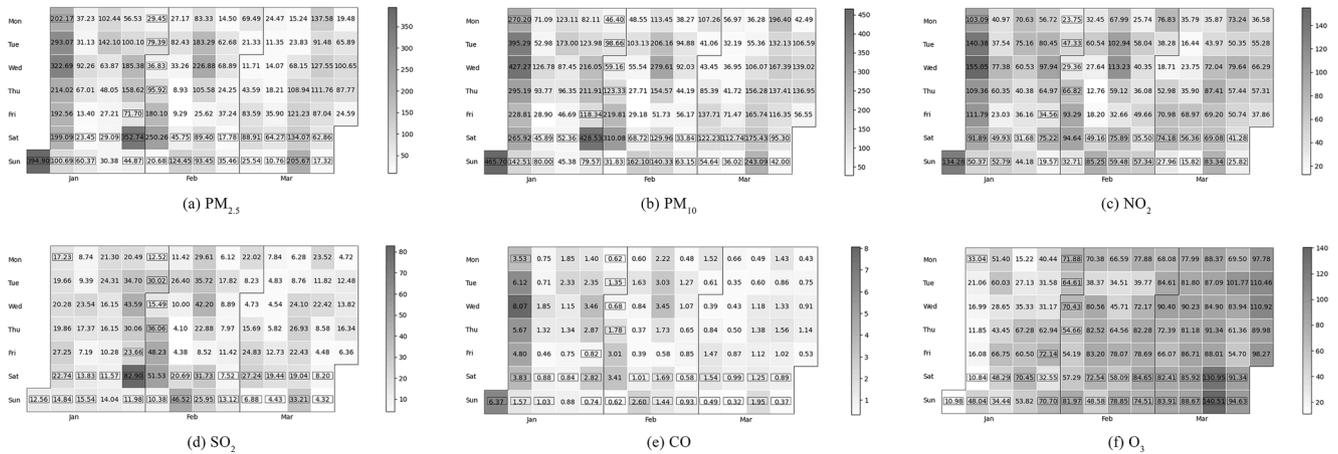


Fig. 10. Calendar heat map of air pollution.

corresponding mobility type. It is evident that, in the long-term models, the non-linear effects of mobility on air pollution are much clearer compared to the short-term models.

High values of home-purpose mobility (depicted in red) predominantly fall on the left side of the y-axis in most models where it ranks highly in relative importance, while its low values (shown in blue) are mainly positioned on the right side in the models (Fig. 13). This pattern suggests a clear negative correlation between home-purpose mobility and air pollution. The only exception occurs in the long-term holiday model, where, as home-purpose mobility increases, its relationship with O_3 is initially negative, but once a certain threshold is surpassed, the correlation becomes predominantly positive (Fig. 14).

Work-purpose mobility in long-term workday models primarily shows a negative correlation with NO_2 , CO , and O_3 and a positive correlation with PM_{10} . In holiday models, it shows a negative correlation with $PM_{2.5}$. In other models, work-purpose mobility does not exhibit a significant correlation with air pollutants.

Transport-purpose mobility in the long-term workday model shows a negative correlation with CO , and its relationship with NO_2 initially

becomes positive as mobility density increases, then shifts to negative, and eventually becomes less pronounced. In the holiday model, it shows a relationship with $PM_{2.5}$ that transitions from positive to negative.

Shopping-purpose mobility in the long-term workday model is primarily negatively correlated with O_3 , while in the holiday model, it shows a moderate positive correlation with PM_{10} . Leisure-purpose mobility initially indicates a negative correlation with $PM_{2.5}$, PM_{10} , and CO in the long-term holiday model, followed by a slightly positive correlation with O_3 . It primarily shows a transition from positive to negative correlation with O_3 .

As summarized in Section 4.3.2, the relative importance of other types of mobility in the models is relatively low. Dining-purpose mobility, as density increases, first shows a negative correlation with O_3 and then a slight positive correlation in the long-term holiday model. Education-purpose mobility in the long-term workday model initially shows a positive correlation with PM_{10} , which turns negative once a certain threshold is exceeded. Medical-purpose mobility in the long-term workday model shows a moderate negative correlation with NO_2 . Other daily activity-purpose mobility in the long-term holiday

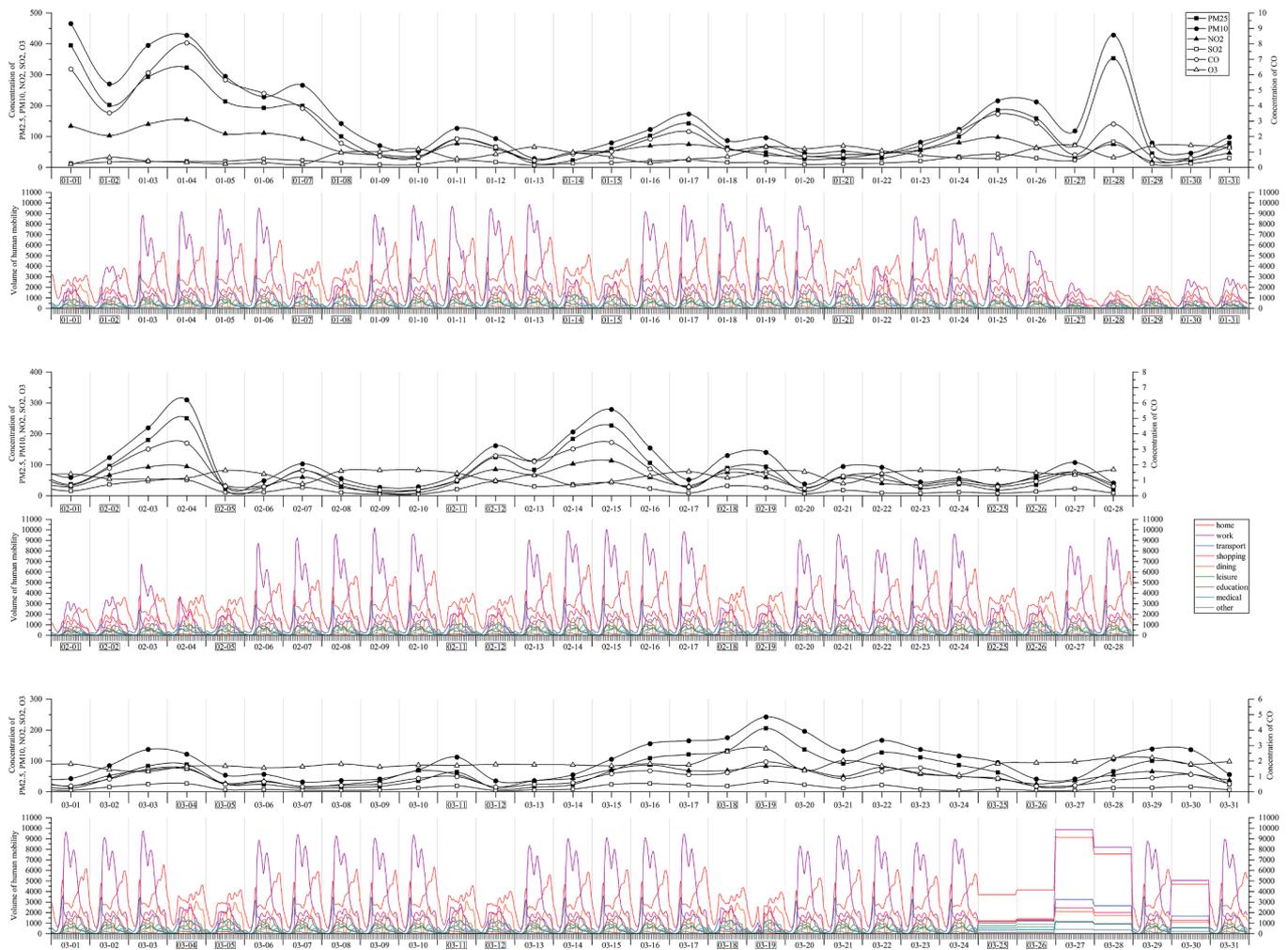


Fig. 11. Time-series graph illustrating the hourly volume of human mobility and air pollution levels.

Table 4
Model performance results of the OLS and XGBoost models.

| Regression Model | OLS Model | | XGBoost Model | |
|-----------------------------|---------------|--------|---------------|--------|
| | Adj. R-square | RMSE | Adj. R-square | RMSE |
| 1. Long-term models | | | | |
| 1-1. Workdays | | | | |
| 1-1-a: PM _{2.5} | 0.825 | 1.244 | 0.944 | 0.698 |
| 1-1-b: PM ₁₀ | 0.757 | 2.776 | 0.910 | 1.679 |
| 1-1-c: NO ₂ | 0.118 | 2.743 | 0.800 | 1.269 |
| 1-1-d: SO ₂ | 0.564 | 1.373 | 0.818 | 0.900 |
| 1-1-e: CO | 0.582 | 0.062 | 0.871 | 0.035 |
| 1-1-f: O ₃ | 0.692 | 2.002 | 0.926 | 0.960 |
| 1-2. Holidays | | | | |
| 1-2-a: PM _{2.5} | 0.677 | 1.849 | 0.921 | 0.946 |
| 1-2-b: PM ₁₀ | 0.605 | 3.947 | 0.915 | 1.897 |
| 1-2-c: NO ₂ | 0.188 | 2.717 | 0.870 | 1.109 |
| 1-2-d: SO ₂ | 0.329 | 1.539 | 0.810 | 0.860 |
| 1-2-e: CO | 0.463 | 0.083 | 0.897 | 0.037 |
| 1-2-f: O ₃ | 0.443 | 2.364 | 0.873 | 1.025 |
| 2. Short-term models | | | | |
| 2-a: PM _{2.5} | 0.243 | 72.830 | 0.951 | 18.551 |
| 2-b: PM ₁₀ | 0.210 | 88.114 | 0.946 | 23.055 |
| 2-c: NO ₂ | 0.372 | 24.001 | 0.948 | 6.882 |
| 2-d: SO ₂ | 0.279 | 11.539 | 0.934 | 3.483 |
| 2-e: CO | 0.249 | 1.275 | 0.958 | 0.300 |
| 2-f: O ₃ | 0.512 | 19.260 | 0.962 | 5.393 |

model shows a moderate negative correlation with SO₂.

For control variables, wind speed primarily shows a negative correlation with air pollution. The relationship between wind direction and air pollution is less clear. Both daily lowest temperature (DLT) and daily precipitation (DP) are generally positively correlated with air pollution in the long-term models, except for their negative correlation with O₃. In the short-term models, DLT and DP are primarily negatively correlated with air pollution, except for their positive correlation with O₃. DEM is mainly negatively correlated with air pollution in the long-term models, though it exhibits a more complex non-linear effect. In the short-term models, DEM is primarily positively correlated with air pollution, except for its negative correlation with O₃. Interestingly, in models where the relative importance of NDVI is higher, NDVI is positively correlated with air pollution, suggesting that simply increasing green space may not necessarily improve air quality.

5. Discussion

5.1. The correlation between diverse human mobility and air pollution

Numerous studies have highlighted the significant impact of human mobility on air pollution, yet they often overlook the heterogeneity of mobility based on travel purposes (Bell & Ward, 2000; Schneider et al., 2013; Yan et al., 2013). This study categorizes human mobility for diverse purposes to explore the distinct spatiotemporal distribution and its impact on air pollution. The results reveal a clear correlation between specific types of human mobility and air pollutants, with significant

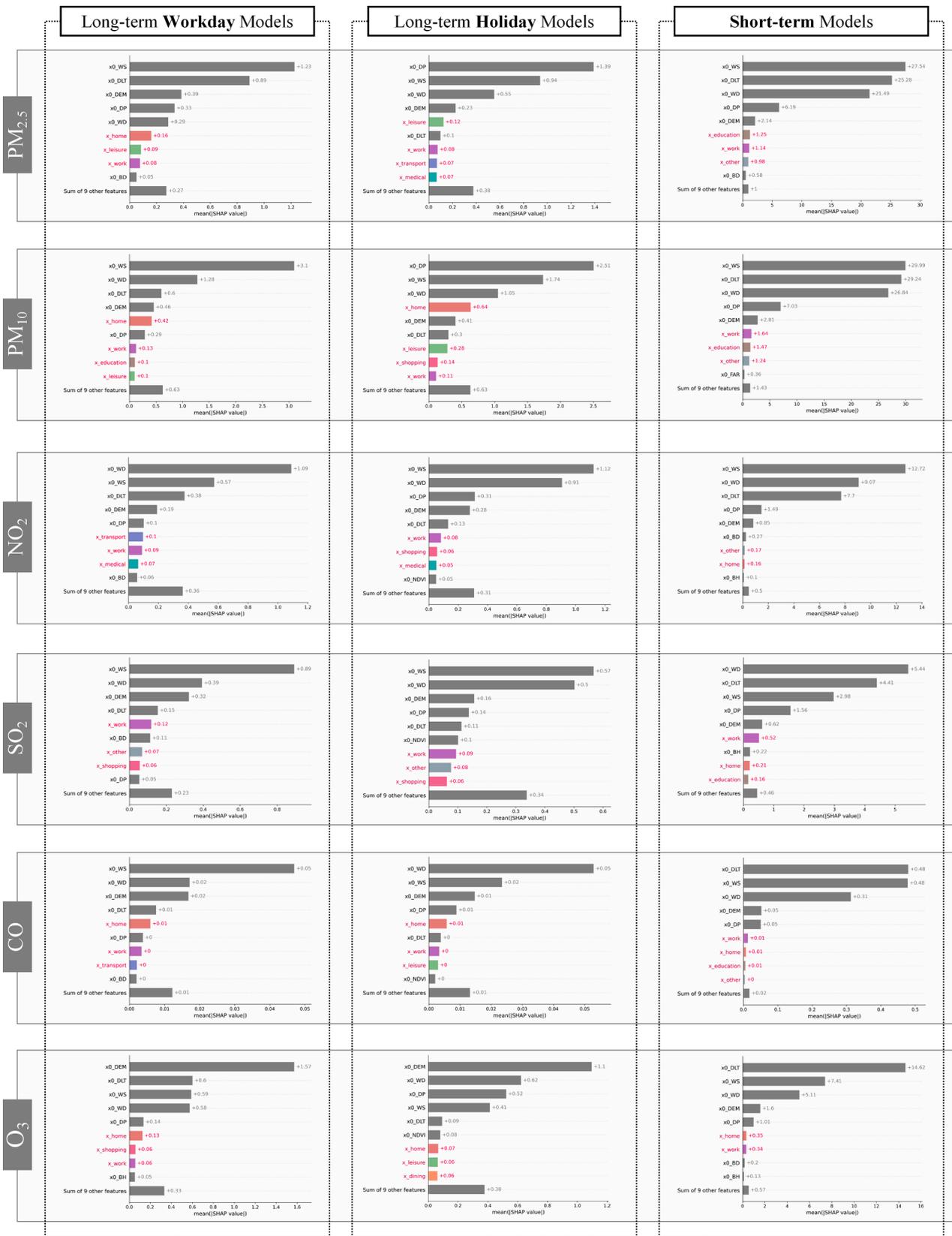


Fig. 12. Importance of different variables for six air pollutants in the long-term and short-term models.

differences observed across different variables and models.

In general, the impact of human mobility on air pollution is less significant than that of natural environments but more important than that of built environments. Moreover, when using average data over

three months for workdays or holidays, the relative importance of human mobility's impact on air pollution is higher than on daily data. The direction of the correlation between specific purpose mobility and air pollution is also clearer in the long term. This may be because, on a

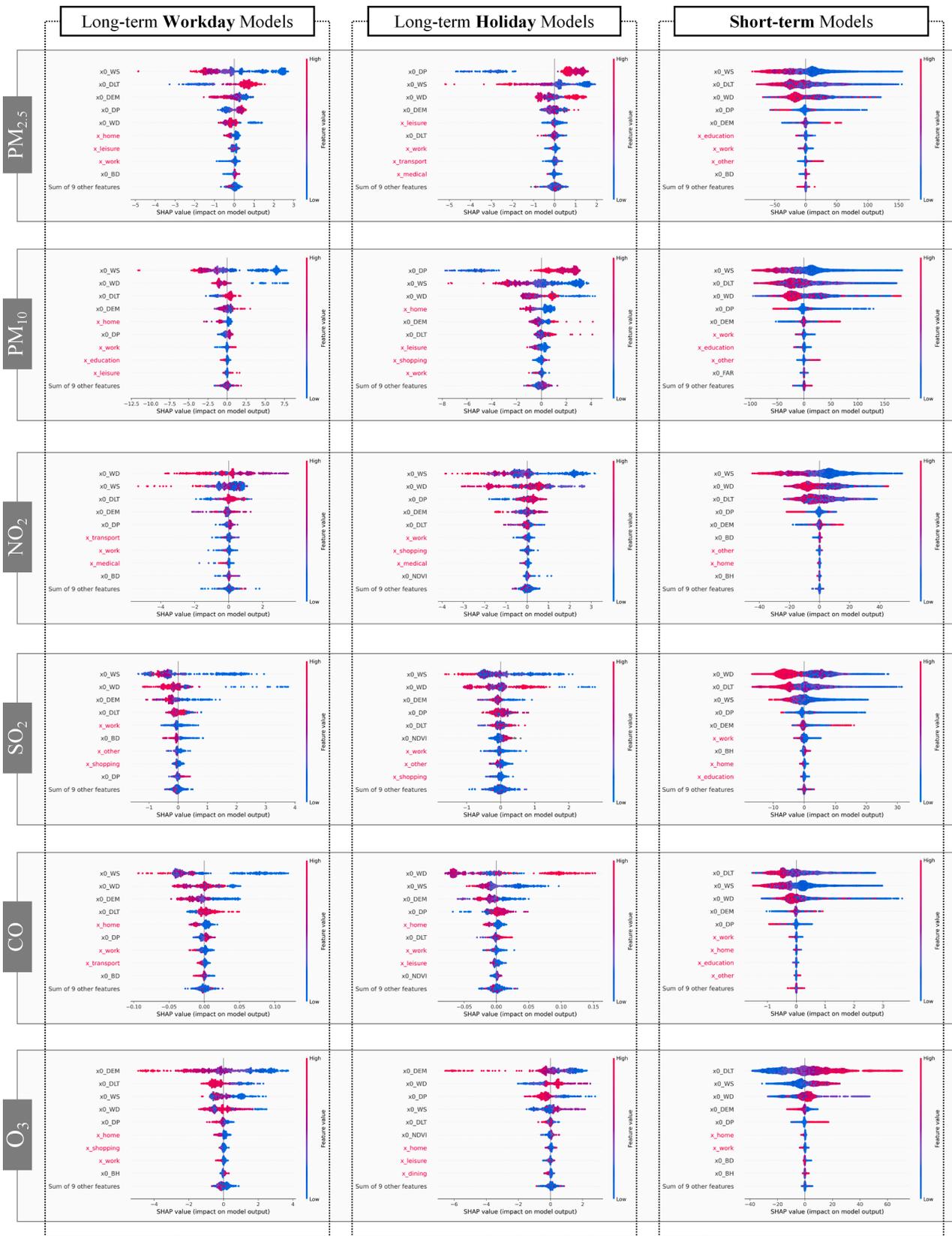


Fig. 13. SHAP values of independent variables for air pollution.

daily timescale, weather conditions exert a more direct and noticeable influence on the formation and dispersion of air pollutants, thereby masking the direct effect of human mobility (Archer et al., 2020). For example, higher wind speeds on a given day are more likely to facilitate

the reduction of pollution levels (Kallos et al., 1993; Yen et al., 2013).

Home- and work-purpose mobility have a more substantial impact on air pollution across all models, while leisure-purpose mobility shows a notable increase in relative importance in holiday models. These types

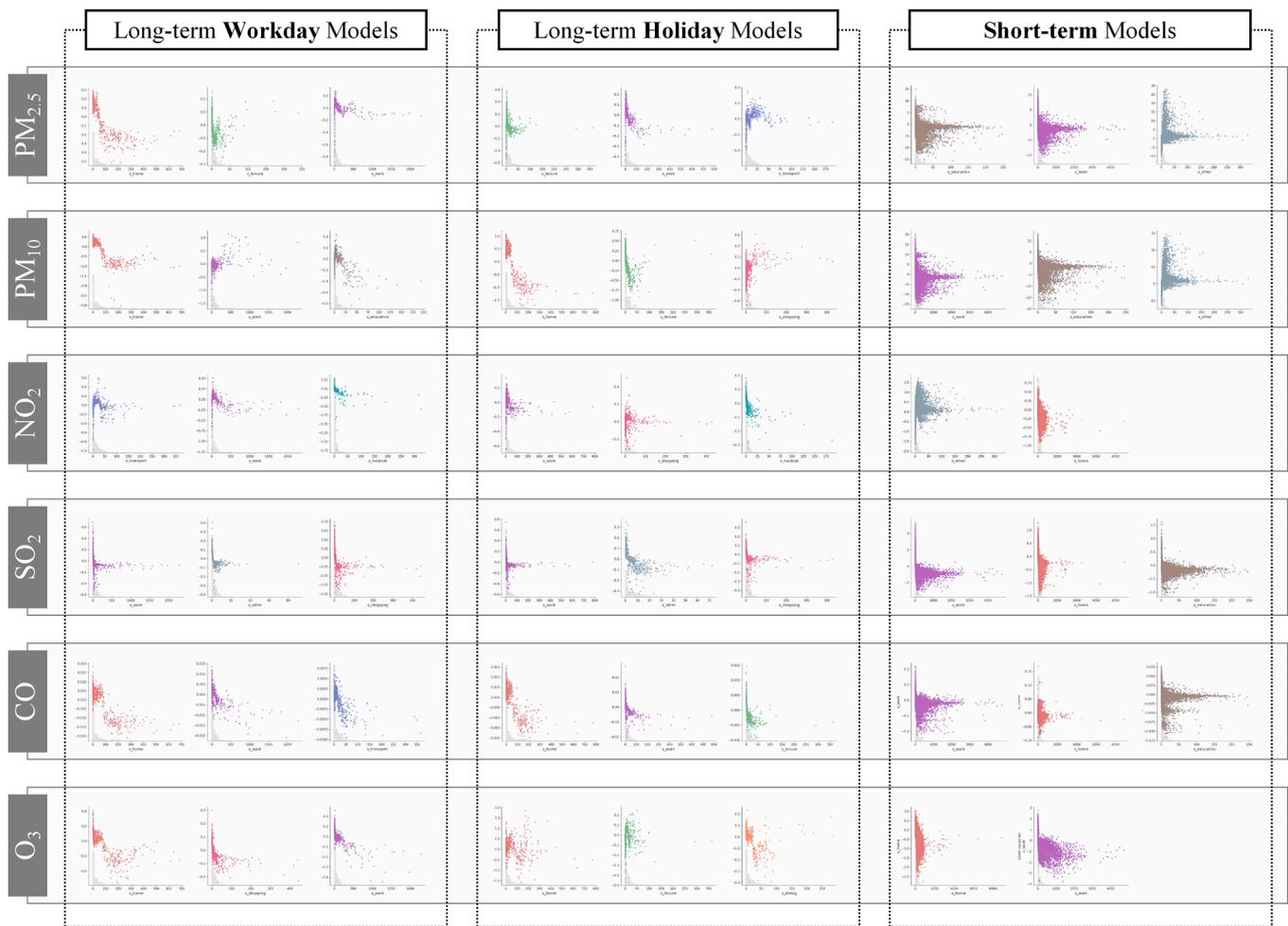


Fig. 14. Non-linear effects of human mobility with highest relative importance on six air pollutants.

of mobility also represent the largest share of volume among all travel purposes. However, for certain specific pollutants, other types of mobility with smaller shares, such as transport, dining, and education, also play an essential role.

The specific types of human mobility with high relative importance mostly show a negative or non-linear correlation with air pollution from a long-term perspective. This suggests that an increase in mobility is not always directly associated with a rise in pollution levels, which aligns with some previous studies (Archer et al., 2020; Munir et al., 2021) but may challenge the common intuition that all types of mobility should increase pollution. The correlational analysis results offer strong statistical support for this counterintuitive association. However, further research is needed to explore the underlying causal mechanisms: whether certain types of purpose-specific mobility reduce air pollution, whether lower pollution levels encourage increased mobility, or whether a more complex interplay of additional variables is involved. Therefore, Various explanations are possible: First, local air pollution may have a reverse effect on long-term human mobility patterns (Qiao et al., 2024), as people tend to engage in activities in lower-pollution areas or reduce travel during high-pollution periods (Xia, 2024). For example, certain residential and office areas may already have low pollution levels due to environmental factors, so even if mobility increases, the overall pollution level remains relatively low. Second, a higher proportion of mobility for specific purposes may result in less pollution compared to other types of mobility, or it could reduce the volume of other mobility types. This could be related to different travel behaviors, such as driving distance and time, as well as the activities at the destination. For example, returning home or transferring to public

transportation with lower per capita emissions could reduce pollution (Archer et al., 2020).

Moreover, the impact of human mobility on different air pollutants varies. For example, work-purpose mobility shows a negative correlation with $PM_{2.5}$ but a positive correlation with PM_{10} . Transport-purpose mobility on workdays exhibits a significant negative correlation only with CO, with no notable correlation with other pollutants. This may be due to the fact that different air pollutants have distinct sources and propagation characteristics (Elminir, 2005; Fernández et al., 2021; Xiao et al., 2018). Additionally, similar to other studies, O_3 in this research shows particularly distinct behavior compared to other pollutants (Marković et al., 2008; Venter et al., 2020; Xiao et al., 2018). For instance, leisure-purpose mobility during holidays generally shows a negative correlation followed by a slight positive correlation with most pollutants, while the relationship with O_3 is the opposite, starting with a positive correlation and then turning negative. Similarly, control variables, including daily lowest temperature, daily precipitation, and DEM, also display an inverse correlation with O_3 compared to most other pollutants.

5.2. Practical implications and suggestions

Based on the above findings from the Beijing case study and existing research (Asamer et al., 2016; Wang & Liu, 2014; Wang et al., 2015), this study proposes several recommendations for urban planning and traffic management policies to develop smarter and more sustainable cities.

The spatial distribution of mobility for different purposes largely aligns with the spatial distribution of urban functional zones, while the

temporal distribution follows the patterns of daily human activities (Schneider et al., 2013; Yan et al., 2013). Therefore, the spatiotemporal distribution characteristics of mobility for different purposes presented in this study can help urban planners and policymakers better understand when specific urban functional zones are more active (Liu et al., 2023; Wang et al., 2022; Wu et al., 2021; Yu et al., 2022). It is evident that mobility for specific purposes often peaks at particular times of the day (Fig. 11), potentially leading to congestion and pollution concentrated in specific urban areas, especially within single-functional urban zones. Although the goal of this study is not to establish an exact air pollution forecasting model, the spatiotemporal patterns observed over the long term can help policymakers more efficiently and intuitively predict areas in urban environments that may experience pollution issues (Baklanov & Zhang, 2020; Forehead & Huynh, 2018; Pantusheva et al., 2022; Zhong et al., 2016). For example, time-based dynamic traffic management could help reduce congestion and emissions in localized regions (Meng et al., 2020).

The negative correlation between work- and home-purpose mobility and air pollution challenges the traditional assumption that “more mobility equals more pollution,” suggesting that there is no need for “one-size-fits-all” restrictions on mobility, thereby minimizing the impact on residents’ travel convenience (Jia et al., 2017; Wang & Liu, 2014; Zhang et al., 2017). Instead, we can reduce urban air pollution by promoting electric taxis and other low-emission transport options, especially for high-frequency travel purposes such as commuting to work and home (Orset, 2019; Sun et al., 2021). Additionally, policies that promote the integration of residential and workplace areas can improve air quality (Babalik-Sutcliffe, 2013; Zhang & Zhang, 2023). By encouraging development in areas with shorter distances between home and work, such as creating mixed-use zones that combine residential and office spaces, long commutes can be reduced, leading to lower pollution exposure and better air quality.

Finally, natural environmental factors, such as wind speed and urban green spaces, have strong correlations with air pollution. Increased wind speed helps to disperse pollutants and reduce their concentration, suggesting the promotion of “urban wind corridors” (Huang et al., 2021; Yang et al., 2020). Additionally, merely increasing green space may not significantly reduce pollution. Different types of green spaces have varying effects on air quality, with open, well-connected green areas that facilitate air movement likely to be more beneficial in reducing pollution (Venter et al., 2024; Wu & Chen, 2023; Yang et al., 2020).

5.3. Limitations and future works

The primary methodological framework of this study focuses on correlations, which may limit the ability to infer explicit causal mechanisms, as discussed in Section 5.1. For example, improved air quality may encourage certain types of mobility (Cui et al., 2019; Park & Kwan, 2017; Tang et al., 2024; Zhang et al., 2019). Future research could benefit from more rigorous causal inference frameworks, such as structural causal models and the potential outcomes framework, to better understand the underlying mechanisms driving these correlations. Besides, the synergistic effects of different variables and potential confounding variables warrant further investigation. For instance, examining how built environmental factors and human mobility factors interact to influence air pollution could provide deeper insights into the varying performance of the same independent variable across different models (Yi et al., 2022; Zeng & He, 2023).

Taxis were employed as a proxy to represent the broader pollution externality of human mobility in this study. This includes not only direct emissions from various transportation modes but also the pollution arising from activities conducted at destinations. However, while taxi trajectory data offers advantages in spatiotemporal resolution, it only captures a portion of urban mobility. Future research could integrate data from other transportation modes, such as private cars, buses, and bicycles, to further refine the comprehensive analysis of how urban

mobility impacts air pollution (Guo et al., 2022; Zhao et al., 2023; Zhu et al., 2024). Additionally, constrained by data availability, this study examined human mobility over a three-month period, primarily covering winter. Although this duration is relatively long compared to existing studies and provides some insight into long-term trends (Fig. 11), it does not offer sufficient evidence for seasonal or cross-year variations. Therefore, future studies could encompass data spanning a year or even several years, allowing for the identification of patterns over a longer time scale (Benchrif et al., 2021; Shi et al., 2023). The development of such datasets is also urgently needed. Moreover, as the air pollution data in this study were limited to daily resolution, future research could explore the characteristics of air pollution at more refined temporal resolutions (Zhang & Li, 2024).

This study uses Beijing as a case study not only to address its specific challenges but also to provide insights and serve as a reference for cities worldwide, as discussed in Section 5.2. However, practical implications and policy recommendations should be adapted to the specific context of each city rather than applied universally. Future research could incorporate a broader range of cities across different regions and urban types to enhance the reliability and generalizability of the findings.

6. Conclusion

As air pollution becomes an increasingly severe environmental challenge in cities worldwide, there is a growing need to better understand the relationship between human mobility and air pollution. Compared to prior studies, this research provides a more comprehensive exploration of the spatiotemporal heterogeneity of purpose-specific human mobility, revealing the complex and sometimes counterintuitive relationship between urban mobility and air pollution.

Using fine-scale taxi trajectory data over three months in Beijing and interpretable machine-learning models, our correlational analysis reveals significant associations between purpose-specific mobility patterns and air pollution variations in Beijing. Key findings include the substantial contribution of wind, temperature, and precipitation to air pollution in the regression models. Human mobility’s contribution is less significant than that of natural environments but greater than built environments. Additionally, in the long term, the impact of human mobility is more pronounced, with clearer directions of correlation. Notably, the negative correlation between work- and home-purpose mobility and pollution challenges the assumption that more mobility always leads to more pollution. The study also highlights significant differences in the correlation between various pollutants and independent variables, particularly O_3 .

Based on these findings, the study suggests several recommendations for urban planning and management. These include promoting mixed-use development and supporting work-residence integration in functional zoning, encouraging the creation of urban wind corridors and open green spaces in environmental planning, and promoting low-emission transportation while avoiding blanket traffic restriction measures. However, these suggestions may be further validated through comparative studies in other regions and practical exploration.

In conclusion, this study in Beijing serves as a typical case for reevaluating and improving existing facility planning and traffic control policies in cities facing similar pollution problems, contributing to more effective data-driven management and the advancement of smart and sustainable cities.

CRedit authorship contribution statement

Wenrui XU: Writing – original draft, Visualization, Project administration, Methodology, Data curation, Conceptualization. **Xinyue GU:** Writing – review & editing.

Declaration of competing interest

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

Acknowledgment

The January-March 2017 Beijing taxi pick-up and drop-off point information dataset is sourced from the National Key R&D Program project “Geographic Big Data Mining and Spatiotemporal Pattern Discovery (2017YFB0503600) at China National Earth Observation Data Center”.

This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.scs.2025.106411](https://doi.org/10.1016/j.scs.2025.106411).

Data availability

The authors do not have permission to share data.

References

- An, S., Hu, X., & Wang, J. (2011). Urban taxis and air pollution: A case study in Harbin, China. *Journal of Transport Geography*, 19(4), 960–967. <https://doi.org/10.1016/j.jtrangeo.2010.12.005>
- Archer, C. L., Cervone, G., Golbazi, M., Al Fahel, N., & Hultquist, C. (2020). Changes in air quality and human mobility in the USA during the COVID-19 pandemic. *Bulletin of Atmospheric Science and Technology*, 1(3), 491–514. <https://doi.org/10.1007/s42865-020-00019-0>
- Asamer, J., Reinthaler, M., Ruthmair, M., Straub, M., & Puchinger, J. (2016). Optimizing charging station locations for urban taxi providers. *Transportation Research Part A: Policy and Practice*, 85, 233–246. <https://doi.org/10.1016/j.tra.2016.01.014>
- Babalik-Sutcliffe, E. (2013). Urban form and sustainable transport: Lessons from the Ankara case. *International Journal of Sustainable Transportation*, 7(5), 416–430. <https://doi.org/10.1080/15568318.2012.676152>
- Baklanov, A., & Zhang, Y. (2020). Advances in air quality modeling and forecasting. *Global Transitions*, 2, 261–270. <https://doi.org/10.1016/j.glt.2020.11.001>
- Bao, R., & Zhang, A. (2020). Does lockdown reduce air pollution? Evidence from 44 cities in northern China. *Science of the Total Environment*, 731, Article 139052. <https://doi.org/10.1016/j.scitotenv.2020.139052>
- Bell, M., & Ward, G. (2000). Comparing temporary mobility with permanent migration. *Tourism Geographies*, 2(1), 87–107. <https://doi.org/10.1080/146166800363466>
- Benchrif, A., Wheida, A., Tahri, M., Shubbar, R. M., & Biswas, B. (2021). Air quality during three covid-19 lockdown phases: AQI, PM2.5 and NO2 assessment in cities with more than 1 million inhabitants. *Sustainable Cities and Society*, 74, Article 103170. <https://doi.org/10.1016/j.scs.2021.103170>
- Bouscasse, H., Gabet, S., Kerneis, G., Provent, A., Rieux, C., Ben Salem, N., Dupont, H., Troude, F., Mathy, S., & Slama, R. (2022). Designing local air pollution policies focusing on mobility and heating to avoid a targeted number of pollution-related deaths: Forward and backward approaches combining air pollution modeling, health impact assessment and cost-benefit analysis. *Environment International*, 159, Article 107030. <https://doi.org/10.1016/j.envint.2021.107030>
- Cai, H., Jia, X., Chiu, A. S. F., Hu, X., & Xu, M. (2014). Siting public electric vehicle charging stations in Beijing using big-data informed travel patterns of the taxi fleet. *Transportation Research Part D: Transport and Environment*, 33, 39–46. <https://doi.org/10.1016/j.trd.2014.09.003>
- Cai, H., & Xie, S. (2007). Estimation of vehicular emission inventories in China from 1980 to 2005. *Atmospheric Environment*, 41(39), 8963–8979. <https://doi.org/10.1016/j.atmosenv.2007.08.019>
- Cai, H., & Xu, M. (2013). Greenhouse gas implications of fleet electrification based on big data-informed individual travel patterns. *Environmental Science and Technology*, 47(16), 9035–9043. <https://doi.org/10.1021/es401008f>. Scopus.
- Cakaj, A., Lisiak-Zielińska, M., Khaniabadi, Y. O., & Sicard, P. (2023). Premature deaths related to urban air pollution in Poland. *Atmospheric Environment*, 301, Article 119723. <https://doi.org/10.1016/j.atmosenv.2023.119723>
- Cao, W., Zhou, W., Yu, W., & Wu, T. (2024). Combined effects of urban forests on land surface temperature and PM2.5 pollution in the winter and summer. *Sustainable Cities and Society*, 104, Article 105309. <https://doi.org/10.1016/j.scs.2024.105309>
- Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 785–794). <https://doi.org/10.1145/2939672.2939785>
- Copernicus Climate Change Service, Climate Data Store. (2024). *ERA5-land post-processed daily-statistics from 1950 to present*. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). https://doi.org/10.24381/cds.e9c9c792_</Dataset>.
- Cui, C., Wang, Z., He, P., Yuan, S., Niu, B., Kang, P., & Kang, C. (2019). Escaping from pollution: The effect of air quality on inter-city population mobility in China. *Environmental Research Letters*, 14(12), Article 124025. <https://doi.org/10.1088/1748-9326/ab5039>
- Elminir, H. K. (2005). Dependence of urban air pollutants on meteorology. *Science of the Total Environment*, 350(1), 225–237. <https://doi.org/10.1016/j.scitotenv.2005.01.043>
- Ercan, T., Onat, N. C., Keya, N., Tatari, O., Eluru, N., & Kucukvar, M. (2022). Autonomous electric vehicles can reduce carbon emissions and air pollution in cities. *Transportation Research Part D: Transport and Environment*, 112, Article 103472. <https://doi.org/10.1016/j.trd.2022.103472>
- E.S. Agency. (2024). *Copernicus Global Digital Elevation Model* [Dataset]. Distributed by OpenTopography. <https://doi.org/10.5069/G9028PQB>.
- Feng, R., Feng, Q., Jing, Z., Zhang, M., & Yao, B. (2022). Association of the built environment with motor vehicle emissions in small cities. *Transportation Research Part D: Transport and Environment*, 107, Article 103313. <https://doi.org/10.1016/j.trd.2022.103313>
- Fernández, M. E., Gentili, J. O., & Campo, A. M. (2021). Air pollutants in an intermediate city: Variability and interactions with weather and anthropogenic elements in bahía blanca, argentina. *Environmental Processes*, 8(1), 349–375. <https://doi.org/10.1007/s40710-021-00502-6>
- Ferrero, E., Alessandrini, S., & Balanzino, A. (2016). Impact of the electric vehicles on the air pollution from a highway. *Applied Energy*, 169, 450–459. <https://doi.org/10.1016/j.apenergy.2016.01.098>
- Forehead, H., & Huynh, N. (2018). Review of modelling air pollution from traffic at street-level—The state of the science. *Environmental Pollution*, 241, 775–786. <https://doi.org/10.1016/j.envpol.2018.06.019>
- Fu, S., & Gu, Y. (2017). Highway toll and air pollution: Evidence from Chinese cities. *Journal of Environmental Economics and Management*, 83, 32–49. <https://doi.org/10.1016/j.jeem.2016.11.007>
- Fuller, R., Landrigan, P. J., Balakrishnan, K., Bathan, G., Bose-O'Reilly, S., Brauer, M., Caravanos, J., Chiles, T., Cohen, A., Corra, L., Cropper, M., Ferraro, G., Hanna, J., Hanrahan, D., Hu, H., Hunter, D., Janata, G., Kupka, R., Lanphear, B., ... Yan, C. (2022). Pollution and health: A progress update. *The Lancet Planetary Health*, 6(6), e535–e547. [https://doi.org/10.1016/S2542-5196\(22\)00090-0](https://doi.org/10.1016/S2542-5196(22)00090-0)
- Furletti, B., Cintia, P., Renso, C., & Spinsanti, L. (2013). Inferring human activities from GPS tracks. In *Proceedings of the 2nd ACM SIGKDD international workshop on urban computing* (pp. 1–8). <https://doi.org/10.1145/2505821.2505830>
- Gao, M., Guttikunda, S. K., Carmichael, G. R., Wang, Y., Liu, Z., Stanier, C. O., Saide, P. E., & Yu, M. (2015). Health impacts and economic losses assessment of the 2013 severe haze event in Beijing area. *Science of the Total Environment*, 511, 553–561. <https://doi.org/10.1016/j.scitotenv.2015.01.005>
- Gao, M., Saide, P. E., Xin, J., Wang, Y., Liu, Z., Wang, Y., Wang, Z., Pagowski, M., Guttikunda, S. K., & Carmichael, G. R. (2017). Estimates of health impacts and radiative forcing in winter haze in Eastern China through constraints of surface PM2.5 predictions. *Environmental Science & Technology*, 51(4), 2178–2185. <https://doi.org/10.1021/acs.est.6b03745>
- Ghaffaripasand, O., Okure, D., Green, P., Sayyahi, S., Adong, P., Sserunjogi, R., Bainomugisha, E., & Pope, F. D. (2024). The impact of urban mobility on air pollution in Kampala, an exemplar sub-Saharan African city. *Atmospheric Pollution Research*, 15(4), Article 102057. <https://doi.org/10.1016/j.apr.2024.102057>
- Gong, L., Liu, X., Wu, L., & Liu, Y. (2016). Inferring trip purposes and uncovering travel patterns from taxi trajectory data. *Cartography and Geographic Information Science*, 43(2), 103–114. <https://doi.org/10.1080/15230406.2015.1014424>
- Gu, X., Wu, Z., Liu, X., Qiao, R., & Jiang, Q. (2024). Exploring the Nonlinear interplay between urban morphology and nighttime thermal environment. *Sustainable Cities and Society*, 101, Article 105176. <https://doi.org/10.1016/j.scs.2024.105176>
- Guo, B., Yang, H., Zhou, H., Huang, Z., Zhang, F., Xiao, L., & Wang, P. (2022). Understanding individual and collective human mobility patterns in twelve crowding events occurred in Shenzhen. *Sustainable Cities and Society*, 81, Article 103856. <https://doi.org/10.1016/j.scs.2022.103856>
- Han, L., Zhao, J., Zhang, T., & Zhang, J. (2022). Urban ventilation corridors exacerbate air pollution in central urban areas: Evidence from a Chinese city. *Sustainable Cities and Society*, 87, Article 104129. <https://doi.org/10.1016/j.scs.2022.104129>
- He, L., Wei, J., Wang, Y., Shang, Q., Liu, J., Yin, Y., Frankenberg, C., Jiang, J. H., Li, Z., & Yung, Y. L. (2022). Marked impacts of pollution mitigation on crop yields in China. *Earth's Future*, 10(11). <https://doi.org/10.1029/2022EF002936>. e2022EF002936.
- Huang, G., Zhang, W., & Xu, D. (2022). How do technology-enabled bike-sharing services improve urban air pollution? Empirical evidence from China. *Journal of Cleaner Production*, 379, Article 134771. <https://doi.org/10.1016/j.jclepro.2022.134771>
- Huang, L., Li, Q., & Yue, Y. (2010). Activity identification from GPS trajectories using spatial temporal POIs' attractiveness. In *Proceedings of the 2nd ACM sigspatial international workshop on location based social networks* (pp. 27–30). <https://doi.org/10.1145/1867699.1867704>
- Huang, Y., Lei, C., Liu, C.-H., Perez, P., Forehead, H., Kong, S., & Zhou, J. L. (2021). A review of strategies for mitigating roadside air pollution in urban street canyons. *Environmental Pollution*, 280, Article 116971. <https://doi.org/10.1016/j.envpol.2021.116971>
- Huang, Z., Cao, F., Jin, C., Yu, Z., & Huang, R. (2017). Carbon emission flow from self-driving tours and its spatial relationship with scenic spots – A traffic-related big data method. *Journal of Cleaner Production*, 142, 946–955. <https://doi.org/10.1016/j.jclepro.2016.09.129>

- Jia, N., Zhang, Y., He, Z., & Li, G. (2017). Commuters' acceptance of and behavior reactions to license plate restriction policy: A case study of Tianjin, China. *Transportation Research Part D: Transport and Environment*, 52, 428–440. <https://doi.org/10.1016/j.trd.2016.10.035>
- Jiang, S., Ferreira, J., & González, M. C. (2012). Clustering daily patterns of human activities in the city. *Data Mining and Knowledge Discovery*, 25(3), 478–510. <https://doi.org/10.1007/s10618-012-0264-z>
- Kallos, G., Kassomenos, P., & Pielke, R. A. (1993). Synoptic and mesoscale weather conditions during air pollution episodes in Athens, Greece. *Boundary-Layer Meteorology*, 62(1), 163–184. <https://doi.org/10.1007/BF00705553>
- Landrigan, P. J., Fuller, R., Acosta, N. J. R., Adeyi, O., Arnold, R., Basu, N.(N), Baldé, A. B., Bertollini, R., Bose-O'Reilly, S., Boufford, J. I., Breysse, P. N., Chiles, T., Mahidol, C., Coll-Seck, A. M., Cropper, M. L., Fobil, J., Fuster, V., Greenstone, M., Haines, A., ... Zhong, M. (2018). The Lancet Commission on pollution and health. *The Lancet*, 391(10119), 462–512. [https://doi.org/10.1016/S0140-6736\(17\)32345-0](https://doi.org/10.1016/S0140-6736(17)32345-0)
- Leroutier, M., & Quirion, P. (2022). Air pollution and CO2 from daily mobility: Who emits and why? Evidence from Paris. *Energy Economics*, 109, Article 105941. <https://doi.org/10.1016/j.eneco.2022.105941>
- Li, H., Li, Y., Wang, T., Wang, Z., Gao, M., & Shen, H. (2021). Quantifying 3D building form effects on urban land surface temperature and modeling seasonal correlation patterns. *Building and Environment*, 204, Article 108132. <https://doi.org/10.1016/j.buildenv.2021.108132>
- Li, S., Zhuang, C., Tan, Z., Gao, F., Lai, Z., & Wu, Z. (2021). Inferring the trip purposes and uncovering spatio-temporal activity patterns from dockless shared bike dataset in Shenzhen, China. *Journal of Transport Geography*, 91. <https://doi.org/10.1016/j.jtrangeo.2021.102974>. Scopus.
- Liu, J., Meng, B., & Shi, C. (2023). A multi-activity view of intra-urban travel networks: A case study of Beijing. *Cities*, 143, Article 104634. <https://doi.org/10.1016/j.cities.2023.104634>
- Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 30. https://proceedings.neurips.cc/paper_files/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html
- Luo, X., Dong, L., Dou, Y., Zhang, N., Ren, J., Li, Y., Sun, L., & Yao, S. (2017). Analysis on spatial-temporal features of taxis' emissions from big data informed travel patterns: A case of Shanghai, China. *Journal of Cleaner Production*, 142, 926–935. <https://doi.org/10.1016/j.jclepro.2016.05.161>
- Marković, D. M., Marković, D. A., Jovanović, A., Lazić, L., & Mijić, Z. (2008). Determination of O₃, NO₂, SO₂, CO and PM₁₀ measured in Belgrade urban area. *Environmental Monitoring and Assessment*, 145(1), 349–359. <https://doi.org/10.1007/s10661-007-0044-1>
- Meng, X., Zhang, K., Pang, K., & Xiang, X. (2020). Characterization of spatio-temporal distribution of vehicle emissions using web-based real-time traffic data. *Science of the Total Environment*, 709, Article 136227. <https://doi.org/10.1016/j.scitotenv.2019.136227>
- Munir, S., Coskuner, G., Jassim, M. S., Aina, Y. A., Ali, A., & Mayfield, M. (2021). Changes in air quality associated with mobility trends and meteorological conditions during COVID-19 lockdown in northern England, UK. *Atmosphere*, 12(4), 4. <https://doi.org/10.3390/atmos12040504>
- Niu, L., Zhang, Z., Liang, Y., & Van Vliet, J. (2024). Spatiotemporal patterns and drivers of the urban air pollution island effect for 2273 cities in China. *Environment International*, 184, Article 108455. <https://doi.org/10.1016/j.envint.2024.108455>
- Nyhan, M., Grauwlin, S., Britter, R., Misstear, B., McNabola, A., Laden, F., Barrett, S. R. H., & Ratti, C. (2016). Exposure track—The impact of mobile-device-based mobility patterns on quantifying population exposure to air pollution. *Environmental Science & Technology*, 50(17), 9671–9681. <https://doi.org/10.1021/acs.est.6b02385>
- Nyhan, M. M., Kloog, I., Britter, R., Ratti, C., & Koutrakis, P. (2019). Quantifying population exposure to air pollution using individual mobility patterns inferred from mobile phone data. *Journal of Exposure Science & Environmental Epidemiology*, 29(2), 238–247. <https://doi.org/10.1038/s41370-018-0038-9>
- OpenTopography. (2021). Copernicus GLO-90 digital surface model. *OpenTopography*. <https://doi.org/10.5069/G9028PQB>
- O'Regan, A. C., Byrne, R., Hellebust, S., & Nyhan, M. M. (2022). Associations between Google Street View-derived urban greenspace metrics and air pollution measured using a distributed sensor network. *Sustainable Cities and Society*, 87, Article 104221. <https://doi.org/10.1016/j.scs.2022.104221>
- Orset, C. (2019). How do travellers respond to health and environmental policies to reduce air pollution? *Ecological Economics*, 156, 68–82. <https://doi.org/10.1016/j.ecolecon.2018.08.016>
- Pantushveva, M., Mitkov, R., Hristov, P. O., & Petrova-Antonova, D. (2022). Air pollution dispersion modelling in urban environment using CFD: A systematic review. *Atmosphere*, 13(10), 10. <https://doi.org/10.3390/atmos13101640>
- Park, Y. M., & Kwan, M.-P. (2017). Individual exposure estimates may be erroneous when spatiotemporal variability of air pollution and human mobility are ignored. *Health & Place*, 43, 85–94. <https://doi.org/10.1016/j.healthplace.2016.10.002>
- Pisoni, E., Christidis, P., Thunis, P., & Trombetti, M. (2019). Evaluating the impact of “sustainable urban mobility plans” on urban background air quality. *Journal of Environmental Management*, 231, 249–255. <https://doi.org/10.1016/j.jenvman.2018.10.039>
- Qian, X., & Ukkusuri, S. V. (2015). Spatial variation of the urban taxi ridership using GPS data. *Applied Geography*, 59, 31–42. <https://doi.org/10.1016/j.apgeog.2015.02.011>
- Qiao, R., Gao, S., Liu, X., Xia, L., Zhang, G., Meng, X., Liu, Z., Wang, M., Zhou, S., & Wu, Z. (2024). Understanding the global subnational migration patterns driven by hydrological intrusion exposure. *Nature Communications*, 15(1), 1. <https://doi.org/10.1038/s41467-024-49609-y>
- Qin, R., & Feng, Z. (2022). HRLT: A high-resolution (1 day, 1 km) and long-term (1961–2019) gridded dataset for temperature and precipitation across China [Dataset]. PANGAEA. <https://doi.org/10.1594/PANGAEA.940192>
- Qin, R., & Zhang, F. (2022). HRLT: A high-resolution (1 day, 1 km) and long-term (1961–2019) gridded dataset for temperature and precipitation across China [Dataset]. PANGAEA. <https://doi.org/10.1594/PANGAEA.941329>
- Quan, Y., & Xie, L. (2022). Serendipity of vehicle ownership restrictions: Beijing's license plate lottery cultivates non-driving behavior. *Transportation Research Part D: Transport and Environment*, 113, Article 103532. <https://doi.org/10.1016/j.trd.2022.103532>
- Rahman, Md. M., Paul, K. C., Hossain, Md. A., Ali, G. G. Md. N., Rahman, Md. S., & Thill, J.-C. (2021). Machine learning on the COVID-19 pandemic, human mobility and air quality: A review. *IEEE Access: Practical Innovations, Open Solutions*, 9, 72420–72450. <https://doi.org/10.1109/ACCESS.2021.3079121>. IEEE Access.
- Schneider, C. M., Belik, V., Couronné, T., Smoreda, Z., & González, M. C. (2013). Unravelling daily human mobility motifs. *Journal of the Royal Society, Interface*, 10(84), Article 20130246. <https://doi.org/10.1098/rsif.2013.0246>
- Setton, E., Marshall, J. D., Brauer, M., Lundquist, K. R., Hystad, P., Keller, P., & Cloutier-Fisher, D. (2011). The impact of daily mobility on exposure to traffic-related air pollution and health effect estimates. *Journal of Exposure Science & Environmental Epidemiology*, 21(1), 42–48. <https://doi.org/10.1038/jes.2010.14>
- Sharma, A., Shiwang, J., Lee, A., & Peng, W. (2023). Equity implications of electric vehicles: A systematic review on the spatial distribution of emissions, air pollution and health impacts. *Environmental Research Letters*, 18(5), Article 53001. <https://doi.org/10.1088/1748-9326/acc87c>
- Shi, X., Zhao, J., He, J., & Xu, H. (2023). Exploring year-to-year spatiotemporal changes in cycling patterns for bike-sharing system in the pre-, during and post-pandemic periods. *Sustainable Cities and Society*, 98, Article 104814. <https://doi.org/10.1016/j.scs.2023.104814>
- Sui, Y., Zhang, H., Song, X., Shao, F., Yu, X., Shibasaki, R., Sun, R., Yuan, M., Wang, C., Li, S., & Li, Y. (2019). GPS data in urban online ride-hailing: A comparative analysis on fuel consumption and emissions. *Journal of Cleaner Production*, 227, 495–505. <https://doi.org/10.1016/j.jclepro.2019.04.159>
- Sun, L., Zhang, T., Liu, S., Wang, K., Rogers, T., Yao, L., & Zhao, P. (2021). Reducing energy consumption and pollution in the urban transportation sector: A review of policies and regulations in Beijing. *Journal of Cleaner Production*, 285, Article 125339. <https://doi.org/10.1016/j.jclepro.2020.125339>
- Tang, J. H. C. G., Huang, Y., Zhu, Y., Yang, X., & Zhuge, C. (2024). The association between travel demand of docked bike-sharing and the built environment: Evidence from seven US cities. *Sustainable Cities and Society*, 106, Article 105325. <https://doi.org/10.1016/j.scs.2024.105325>
- Tian, D., Zhang, J., Li, B., Xia, C., Zhu, Y., Zhou, C., Wang, Y., Liu, X., & Yang, M. (2024). Spatial analysis of commuting carbon emissions in main urban area of Beijing: A GPS trajectory-based approach. *Ecological Indicators*, 159, Article 111610. <https://doi.org/10.1016/j.ecolind.2024.111610>
- Venter, Z. S., Aunan, K., Chowdhury, S., & Lelieveld, J. (2020). COVID-19 lockdowns cause global air pollution declines. *Proceedings of the National Academy of Sciences*, 117(32), 18984–18990. <https://doi.org/10.1073/pnas.2006853117>
- Venter, Z. S., Hassani, A., Stange, E., Schneider, P., & Castell, N. (2024). Reassessing the role of urban green space in air pollution control. *Proceedings of the National Academy of Sciences*, 121(6), Article e2306200121. <https://doi.org/10.1073/pnas.2306200121>
- Wang, J. J., & Liu, Q. (2014). Understanding the parking supply mechanism in China: A case study of Shenzhen. *Journal of Transport Geography*, 40, 77–88. <https://doi.org/10.1016/j.jtrangeo.2014.04.019>
- Wang, J., Wang, S., Voorhees, A. S., Zhao, B., Jang, C., Jiang, J., Fu, J. S., Ding, D., Zhu, Y., & Hao, J. (2015). Assessment of short-term PM_{2.5}-related mortality due to different emission sources in the Yangtze River Delta, China. *Atmospheric Environment*, 123, 440–448. <https://doi.org/10.1016/j.atmosenv.2015.05.060>
- Wang, L., Xu, J., & Qin, P. (2014). Will a driving restriction policy reduce car trips?—The case study of Beijing, China. *Transportation Research Part A: Policy and Practice*, 67, 279–290. <https://doi.org/10.1016/j.tra.2014.07.014>
- Wang, R., Zhang, X., & Li, N. (2022). Zooming into mobility to understand cities: A review of mobility-driven urban studies. *Cities*, 130, Article 103939. <https://doi.org/10.1016/j.cities.2022.103939>
- Wang, Z., Chen, F., & Fujiyama, T. (2015). Carbon emission from urban passenger transportation in Beijing. *Transportation Research Part D: Transport and Environment*, 41, 217–227. <https://doi.org/10.1016/j.trd.2015.10.001>
- 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 PM_{2.5} and sub-2.5 μm 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., Li, K., Dickerson, R. R., Pinker, R. T., Wang, J., Liu, X., Sun, L., Xue, W., & Cribb, M. (2022). Full-coverage mapping and spatiotemporal variations of ground-level ozone (O₃) pollution from 2013 to 2020 across China. *Remote Sensing of Environment*, 270, Article 112775. <https://doi.org/10.1016/j.rse.2021.112775>
- Wei, J., Li, Z., Lyapustin, A., Sun, L., Peng, Y., Xue, W., Su, T., & Cribb, M. (2021). Reconstructing 1-km-resolution high-quality PM_{2.5} data records from 2000 to 2018 in China: Spatiotemporal variations and policy implications. *Remote Sensing of Environment*, 252, 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 (NO₂, SO₂, and CO) in China: Daily seamless mapping and spatiotemporal

- variations. *Atmospheric Chemistry and Physics*, 23(2), 1511–1532. <https://doi.org/10.5194/acp-23-1511-2023>
- Williams, A. M., Foord, J., & Mooney, J. (2012). Human mobility in functional urban regions: Understanding the diversity of mobilities. *International Review of Sociology*, 22(2), 191–209. <https://doi.org/10.1080/03906701.2012.696961>
- Wu, J., Qian, Y., Wang, Y., & Wang, N. (2021). Analyzing the contribution of human mobility to changes in air pollutants: Insights from the COVID-19 lockdown in Wuhan. *ISPRS International Journal of Geo-Information*, 10(12), 12. <https://doi.org/10.3390/ijgi10120836>
- Wu, L., & Chen, C. (2023). Does pattern matter? Exploring the pathways and effects of urban green space on promoting life satisfaction through reducing air pollution. *Urban Forestry & Urban Greening*, 82, Article 127890. <https://doi.org/10.1016/j.ufug.2023.127890>
- Wu, R., Wang, J., Zhang, D., & Wang, S. (2021). Identifying different types of urban land use dynamics using Point-of-interest (POI) and Random Forest algorithm: The case of Huizhou, China. *Cities*, 114, Article 103202. <https://doi.org/10.1016/j.cities.2021.103202>
- Wu, Y., Wang, Y., Wang, L., Song, G., Gao, J., & Yu, L. (2020). Application of a taxi-based mobile atmospheric monitoring system in Cangzhou, China. *Transportation Research Part D: Transport and Environment*, 86, Article 102449. <https://doi.org/10.1016/j.trd.2020.102449>
- Xia, C. (2024). Escaping environmental hazards? Human mobility in response to air pollution and extreme cold events. *Annals of the American Association of Geographers*, 114(6), 1268–1290. <https://doi.org/10.1080/24694452.2024.2332651>
- Xia, F., Cheng, X., Lei, Z., Xu, J., Liu, Y., Zhang, Y., & Zhang, Q. (2023). Heterogeneous impacts of local traffic congestion on local air pollution within a city: Utilizing taxi trajectory data. *Journal of Environmental Economics and Management*, 122, Article 102896. <https://doi.org/10.1016/j.jeem.2023.102896>
- Xiao, K., Wang, Y., Wu, G., Fu, B., & Zhu, Y. (2018). Spatiotemporal characteristics of air pollutants (PM10, PM2.5, SO2, NO2, O3, and CO) in the inland basin city of Chengdu, southwest China. *Atmosphere*, 9(2), 2. <https://doi.org/10.3390/atmos9020074>
- Xing, Y., Wang, K., & Lu, J. J. (2020). Exploring travel patterns and trip purposes of dockless bike-sharing by analyzing massive bike-sharing data in Shanghai, China. *Journal of Transport Geography*, 87, Article 102787. <https://doi.org/10.1016/j.jtrangeo.2020.102787>
- Xu, C., Zhao, J., & Liu, P. (2019). A geographically weighted regression approach to investigate the effects of traffic conditions and road characteristics on air pollutant emissions. *Journal of Cleaner Production*, 239, Article 118084. <https://doi.org/10.1016/j.jclepro.2019.118084>
- Xue, Y., & Li, C. (2020). Extracting Chinese geographic data from Baidu Map API. *The Stata Journal: Promoting Communications on Statistics and Stata*, 20(4), 805–811. <https://doi.org/10.1177/1536867X20976313>
- Yan, X.-Y., Han, X.-P., Wang, B.-H., & Zhou, T. (2013). Diversity of individual mobility patterns and emergence of aggregated scaling laws. *Scientific Reports*, 3(1), 2678. <https://doi.org/10.1038/srep02678>
- Yang, J., Dong, J., Xiao, X., Dai, J., Wu, C., Xia, J., Zhao, G., Zhao, M., Li, Z., Zhang, Y., & Ge, Q. (2019). Divergent shifts in peak photosynthesis timing of temperate and alpine grasslands in China. *Remote Sensing of Environment*, 233, Article 111395. <https://doi.org/10.1016/j.rse.2019.111395>
- Yang, J., Shi, B., Shi, Y., Marvin, S., Zheng, Y., & Xia, G. (2020). Air pollution dispersal in high density urban areas: Research on the triadic relation of wind, air pollution, and urban form. *Sustainable Cities and Society*, 54, Article 101941. <https://doi.org/10.1016/j.scs.2019.101941>
- Yang, J., Yang, Y., Sun, D., Jin, C., & Xiao, X. (2021). Influence of urban morphological characteristics on thermal environment. *Sustainable Cities and Society*, 72, Article 103045. <https://doi.org/10.1016/j.scs.2021.103045>
- Yang, L., Yang, H., Yu, B., Lu, Y., Cui, J., & Lin, D. (2024). Exploring non-linear and synergistic effects of green spaces on active travel using crowdsourced data and interpretable machine learning. *Travel Behaviour and Society*, 34, Article 100673. <https://doi.org/10.1016/j.tbs.2023.100673>
- Yen, M.-C., Peng, C.-M., Chen, T.-C., Chen, C.-S., Lin, N.-H., Tzeng, R.-Y., Lee, Y.-A., & Lin, C.-C. (2013). Climate and weather characteristics in association with the active fires in northern southeast Asia and spring air pollution in taiwan during 2010 7-SEAS/dongsha experiment. *Atmospheric Environment*, 78, 35–50. <https://doi.org/10.1016/j.atmosenv.2012.11.015>
- Yi, H., Zhao, L., Qian, Y., Zhou, L., & Yang, P. (2022). How to achieve synergy between carbon dioxide mitigation and air pollution control? Evidence from China. *Sustainable Cities and Society*, 78, Article 103609. <https://doi.org/10.1016/j.scs.2021.103609>
- Yu, H., Yang, J., Li, T., Jin, Y., & Sun, D. (2022). Morphological and functional polycentric structure assessment of megacity: An integrated approach with spatial distribution and interaction. *Sustainable Cities and Society*, 80, Article 103800. <https://doi.org/10.1016/j.scs.2022.103800>
- Liu, Yu (2017). *The January-March 2017 Beijing taxi pick-up and drop-off point information dataset*. Peking University. </Dataset>.
- Yu, X., Ivey, C., Huang, Z., Gurram, S., Sivaraman, V., Shen, H., Eluru, N., Hasan, S., Henneman, L., Shi, G., Zhang, H., Yu, H., & Zheng, J. (2020). Quantifying the impact of daily mobility on errors in air pollution exposure estimation using mobile phone location data. *Environment International*, 141, Article 105772. <https://doi.org/10.1016/j.envint.2020.105772>
- Zeng, Q.-H., & He, L.-Y. (2023). Study on the synergistic effect of air pollution prevention and carbon emission reduction in the context of “dual carbon”: Evidence from China’s transport sector. *Energy Policy*, 173, Article 113370. <https://doi.org/10.1016/j.enpol.2022.113370>
- Zhai, Z., Song, G., & Yu, L. (2018). How much vehicle activity data is needed to develop robust vehicle specific power distributions for emission estimates? A case study in Beijing. *Transportation Research Part D: Transport and Environment*, 65, 540–550. <https://doi.org/10.1016/j.trd.2018.09.004>
- Zhang, F., Li, Z., Li, N., & Fang, D. (2019). Assessment of urban human mobility perturbation under extreme weather events: A case study in Nanjing, China. *Sustainable Cities and Society*, 50, Article 101671. <https://doi.org/10.1016/j.scs.2019.101671>
- Zhang, W., Lin Lawell, C.-Y. C., & Umanskaya, V. I. (2017). The effects of license plate-based driving restrictions on air quality: Theory and empirical evidence. *Journal of Environmental Economics and Management*, 82, 181–220. <https://doi.org/10.1016/j.jeem.2016.12.002>
- Zhang, X., & Li, N. (2024). An activity space-based gravity model for intracity human mobility flows. *Sustainable Cities and Society*, 101, Article 105073. <https://doi.org/10.1016/j.scs.2023.105073>
- Zhang, X., & Zhang, D. (2023). Urban carbon emission scenario prediction and multi-objective land use optimization strategy under carbon emission constraints. *Journal of Cleaner Production*, 430, Article 139684. <https://doi.org/10.1016/j.jclepro.2023.139684>
- Zhao, B., Deng, M., & Shi, Y. (2023). Inferring nonwork travel semantics and revealing the nonlinear relationships with the community built environment. *Sustainable Cities and Society*, 99, Article 104889. <https://doi.org/10.1016/j.scs.2023.104889>
- Zhao, P., Kwan, M.-P., & Qin, K. (2017). Uncovering the spatiotemporal patterns of CO2 emissions by taxis based on Individuals’ daily travel. *Journal of Transport Geography*, 62, 122–135. <https://doi.org/10.1016/j.jtrangeo.2017.05.001>
- Zhong, J., Cai, X.-M., & Bloss, W. J. (2016). Coupling dynamics and chemistry in the air pollution modelling of street canyons: A review. *Environmental Pollution*, 214, 690–704. <https://doi.org/10.1016/j.envpol.2016.04.052>
- Zhu, B., Hu, S., Kaparias, I., Zhou, W., Ochieng, W., & Lee, D.-H. (2024). Revealing the driving factors and mobility patterns of bike-sharing commuting demands for integrated public transport systems. *Sustainable Cities and Society*, 104, Article 105323. <https://doi.org/10.1016/j.scs.2024.105323>