A citywide spatiotemporal perspective of particulate matter concentration on underground subway platforms

Check for updates

Tianchen Qin[1,2], Ran Tu[3] ✉, An Wang[4], Jinyitao Wang[1], Tianyuan Li[5], Suyang Xu[6], Xinke Li[1], Qiuzi Chen[6], Jihye Kim[7], Caiqing Yan[8], Liubing Huang[8], Jialiang Feng[1] & Shunyao Wang[1] ✉

Indoor air quality (IAQ) in semi-enclosed subway environments poses critical public health concerns. This study investigates IAQ in the Shanghai subway system through year-round monitoring of $PM_{2.5}$ and PM_{10} across representative stations. Concentrations of particulate matter (PM) were markedly higher indoors, peaking during winter weekday morning rush hours, with indoor/outdoor ratios exceeding unity. An interpretable machine learning model was developed to elucidate key factors affecting subway IAQ, identifying platform screen door design and train frequency as dominant influencing factors. By integrating the model with network-wide operational data and passenger boarding records, we quantified city-scale PM levels and commuter exposure. Further analysis incorporating point-of-interest data revealed that stations near residential zones exhibited the highest exposure, while commercial density intensified pollution. These findings provide a system-level assessment of PM exposure within urban rail transit and highlight the need for targeted ventilation and operational strategies to enhance IAQ and safeguard commuter health.

Subway is a crucial mode of public transportation in metropolitan areas. As of 2020, 193 cities have established subway systems worldwide, with a combined network length exceeding 17,000 km[1]. However, concerns have emerged regarding the indoor air quality (IAQ) in subways. Due to the semi-enclosed environment of subway stations, the concentration of air pollutants inside subway systems can be up to 2.2–10 times higher than the outdoor levels[2–9], posing significant public health concerns. This is particularly relevant for subway workers, regular commuters, as well as "green travelers" who prioritize eco-friendly commuting options over personal vehicles. Thus, there is a pressing need to evaluate IAQ and its influencing factors in subways to safeguard the health of passengers and transit employees.

Studies have categorized the origin of subway particulate matter (PM) into indoor sources, which are generated from mechanical wear, human activities[10–12], and outdoor sources, which enter stations via natural/mechanical ventilation from outdoors[13,14]. So far, studies have identified several key factors influencing indoor PM levels, including train arrival frequency[15,16], the effectiveness of Heating, Ventilation, and Air Conditioning (HVAC) systems[6,16–20], and platform design, such as the design of

platform facilities, the depth of the platform, and the height of platform screen doors (PSDs)[21–24].

While studies investigating subway IAQ are well-established, an accurate assessment of its influences on city-wide commuter exposure risks requires integrated spatiotemporal analysis on passenger traveling behavior and passenger load distribution across the city-wide network holistically. Suárez et al.[25] compared PM exposure levels among subway, car, bicycle, and bus commuters, demonstrating how the interaction between transport mode choice and spatiotemporal routes can significantly influence personal exposure. Similarly, biomonitoring data from Montreal revealed that peak-hour commuting behavior (regardless of transport mode) leads to elevated exposure risks due to increased passenger load and concentrated emissions[26]. Parallel findings in Shanghai and Beijing subway systems further confirmed that subway PM concentrations surge during rush hours, underscoring that those who travel during air pollution peaks may experience higher exposure risk to PM[27,28].

Despite these advancements, most of the studies were based on monitoring data, while comprehensive inference of the subway IAQ and its passenger exposure across multiple subway stations and operation scenarios

[1]School of Environmental and Chemical Engineering, Shanghai University, Shanghai, China. [2]School of Environmental Science and Technology, Sun Yat-sen University, Guangzhou, China. [3]North Central Texas Council of Governments, Arlington, TX, USA. [4]Department of Civil and Environmental Engineering, Hong Kong Polytechnic University, Hong Kong, China. [5]Department of Civil and Environmental Engineering, University of Waterloo, Waterloo, ON, Canada. [6]School of Transportation, Southeast University, Nanjing, China. [7]Department of Metallurgical and Materials Engineering, Colorado School of Mines, Golden, CO, USA. [8]Environmental Research Institute, Shandong University, Qingdao, China. ✉e-mail: rtu@nctcog.org; syw@shu.edu.cn
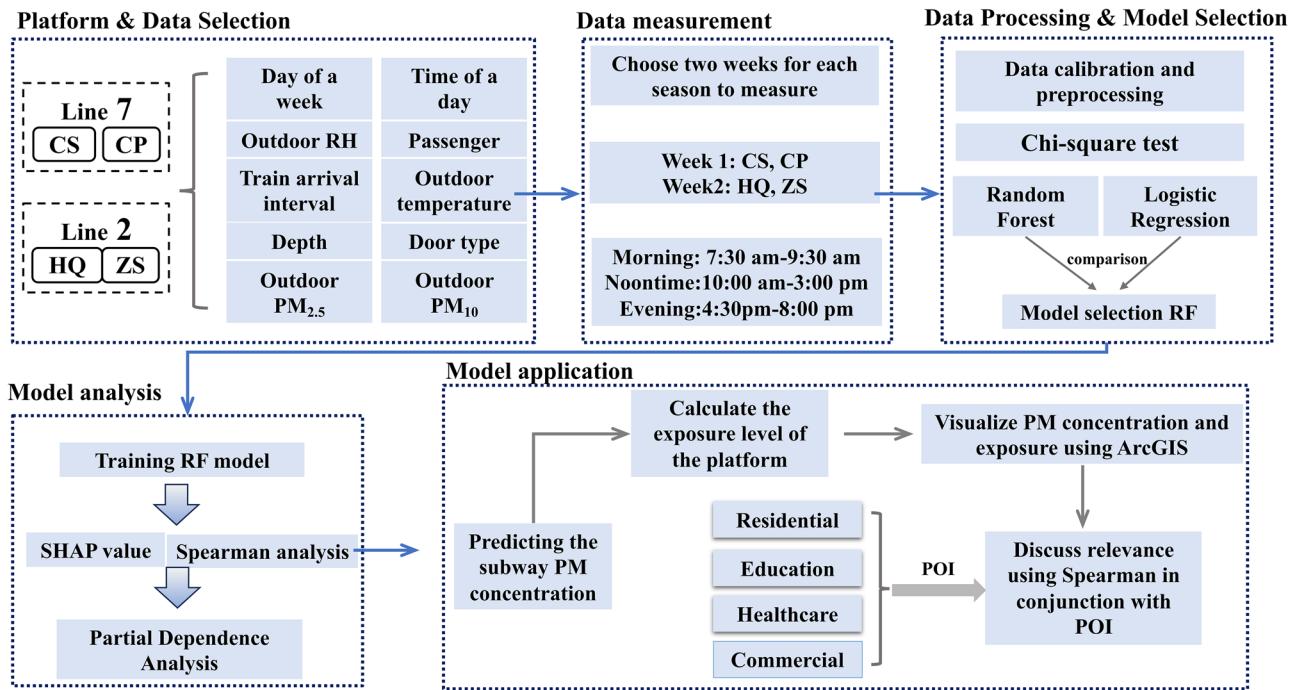
Fig. 1 | Overview of the methodology applied in this study. The data of this study were collected at four underground subway stations in Shanghai for AM, noon, and PM periods during weekdays and weekends for four seasons, followed by data pre-processing and model selection. We selected Random Forest with SHAP value for influencing factor analysis, applied the model to all the subway stations in the whole city, and analyzed the association between the PM concentration, exposure, and the point-of-interest distribution.

remains limited. The present study adopts the Metro system in Shanghai as a typical case (International Union of Public Transport 2022), with the objectives involving: (1) investigating the influencing factors for $PM_{2.5}$ and PM_{10} at subway platforms during rush/non-rush hours across different seasons; (2) analyzing the contributions from various factors to $PM_{2.5}$ and PM_{10} using interpretable machine learning methods; (3) estimating city-wide PM concentrations and exposure level across subway network, and interpreting it with point-of-interest (POI) to discuss potential mitigation strategies for subway IAQ as well as the health benefits for urban commuters.

Results
In this study, concurrent real-time monitoring of indoor/outdoor (I/O) $PM_{2.5}$ and PM_{10} concentrations was conducted at four representative subway stations with typical characteristics in the Shanghai Metro system. It should be noted that other gaseous pollutants (e.g., VOCs, SO_2, NO_2, and O_3) can also represent subway IAQ. In this study, we simply refer to the monitored concentrations of $PM_{2.5}$ and PM_{10} as subway IAQ, given their ubiquitous prevalence and wide investigations in subway-related studies. Measurements were taken during both rush (morning and evening) and non-rush hours (noontime), on weekdays and weekends, across four seasons, allowing us to capture the variations in subway PM levels and identify potential influencing factors. An interpretable machine learning method was applied to identify and deconvolve the main factors influencing the concentrations of $PM_{2.5}$ and PM_{10} in subway microenvironments, followed by a city-wide network estimation of subway PM concentration and passenger exposure. Figure 1 provides an overview of the methodology and its implementation in this study. Novelty of the present study lies in a holistic view of both city-wide POI and subway IAQ, which was predicted by an interpretable machine learning model based on the long-term characteristic data from online monitoring of indoor and outdoor PM at typical subway stations across Shanghai, providing strategic evidence for future sustainable urban planning and subway IAQ regulation.

Measured subway PM and key influencing features
The four stations selected for this study include Changshu Road Station (CS), Changping Road Station (CP), Hongqiao Railway Station (HQ), and Zhongshan Park Station (ZS). These four stations are located on Shanghai Subway Line 7 and Line 2, chosen for their strategic location, design characteristics, and passenger demand. Detailed information on these stations is provided in Fig. 2 and Table 1. CS and CP stations on Line 7 are located in the central downtown area, with service starting in 2009 with full-height platform screen doors (full PSDs). ZS station on Line 2 is also located in the city center but experiences much higher passenger demand and is in a more aged condition with a half-height platform screen door (half PSD) installed. The HQ station, located at the end of Line 2 since 2010, serves as a hub connecting to the inter-city high-speed railway and Hongqiao Airport. Limited site samples may introduce biased results. However, the selection of four sites can well represent the typical types of underground subway stations. In addition, in order to increase the amount of measurement data as much as possible, this study chose to compensate for the sample size issue with measurement duration (i.e., the experiment lasted for 1 year). Studies with similar or smaller sample sizes can be found in existing literature[6,29,30].

Figure 3 illustrates the average indoor and outdoor PM concentrations at the four studied stations (CS, CP, HQ, and ZS) across different seasons. Attention should be given that PM concentration at the ZS site was found to be significantly higher compared to the other sites, with 16-126 μg/m³ and 26-287 μg/m³ for $PM_{2.5}$ and PM_{10} (data refer to Table S1), respectively, which can be attributed to the half PSD that separates the rail track from the standing area while CS, CP, and HQ stations all have full PSD installed. This is consistent with previous studies showing the important role of PSD in controlling platform PM levels inside subway microenvironments[24,31,32]. On the other hand, the ZS station has been in operation for a significantly longer duration (over 20 years) compared to the other three stations (nearly 10 years), which may also elevate PM concentration with the aging facilities[33]. The average concentration of $PM_{2.5}$ and PM_{10} across the four selected sites in this study was 51.9 μg/m³ and 101 μg/m³, respectively.
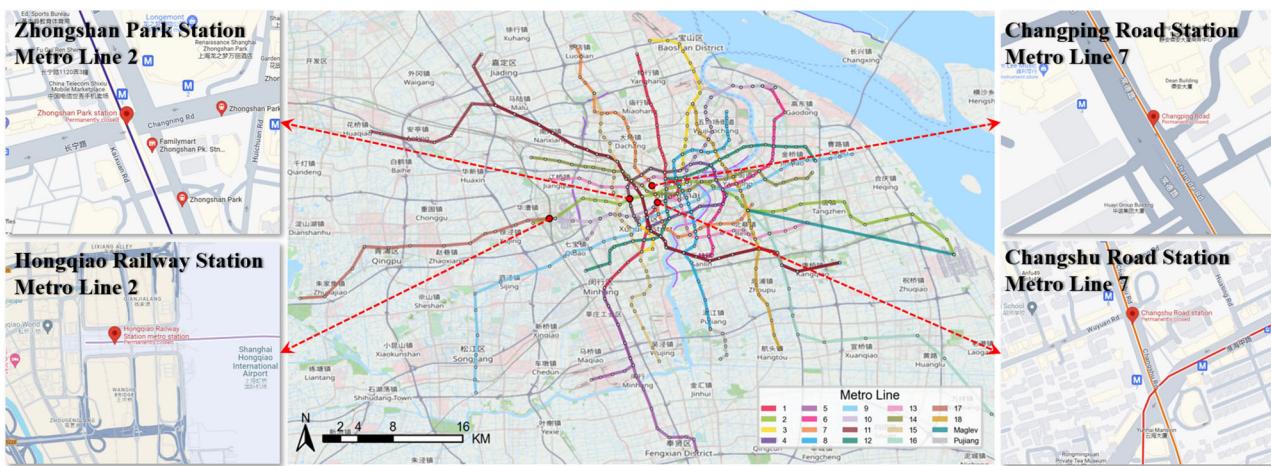
Fig. 2 | Sampling sites on Line 7 and Line 2 of the Shanghai subway system. Four stations were selected: CS (Changshu Road Station) and CP (Changping Road Station) are located on Line 7. HQ (Hongqiao Railway Station) and ZS (Zhongshan Park Station) are located on Line 2.

Table 1 | Characteristics of the stations chosen for measurement

Line #	Station	Platform type[a]	Service time	Track-platform separation	Transfer station
7	CS	Island	Dec, 2009	Full PSD	No
	CP	Side	Dec, 2009	Full PSD	No
2	HQ	Island	Jul, 2010	Full PSD	Yes
	ZS	Island	Jun, 2000	Half PSD	Yes

[a]Island refers to one platform at the center and the rail tracks on both sides; side platform refers to two platforms at each side with rail tracks in the center.

PM$_{2.5}$ and PM$_{10}$ concentrations during rush hour and non-rush hour, and their corresponding I/O ratios, are presented in Fig. 3. Concurrent outdoor concentrations of PM$_{2.5}$ and PM$_{10}$ during rush hour and non-rush hour are provided in Fig. S1 as a reference. Data collected at the CS, CP, and ZS stations showed that PM concentrations measured during rush hour were higher than those measured during the non-rush hour (average concentration of PM during rush and non-rush hour can be seen in Table S1). Such a characteristic is similar to what has been reported in previous studies. For instance, the rush-hour PM$_{2.5}$ concentration in Suzhou was measured to be 265 μg/m^3, being significantly higher than the non-rush-hour concentration[34]. A study in Shanghai also identified concentration peaks of PM$_{2.5}$ and PM$_{10}$ during rush hour, which appeared at 7:00–10:00 am and 4:00–7:00 pm[14]. It is worth noting that the diurnal variation of PM at HQ was distinct from the other three stations, which had the highest concentration during morning rush hour, followed by noon and evening. This can be attributed to the special function of HQ, which connects to inter-city high-speed rail and the international airport with an extended layout, and the daily maximum passenger flow is expected in the morning rush hour. The variation in I/O ratios of PM$_{2.5}$ and PM$_{10}$ during rush hour and non-rush hour is presented in Fig. 3b. The average values of the I/O ratio for PM$_{2.5}$ and PM$_{10}$ measured at four different stations were above 1, indicating a generally higher PM level inside the subway microenvironment compared to the outdoor environment. Comparing the I/O ratio at different times of the day, the I/O ratio in the morning rush hours was slightly higher than the I/O ratio at noontime and evening, which coincides with the variation of outdoor PM, and this might result from the enhanced particle resuspension from passenger movement and crowding during this rush hour.

Feature importance analysis for subway indoor PM
Multiple internal and external factors can influence PM inside the subway microenvironment[6,35–40]. According to existing literature and our

observation during the monitoring campaign, we considered the day of the week, time of day, door type, depth, train arrival interval, meteorology (obtained from the Bureau of Meteorology of Shanghai), passenger flow (obtained from Shanghai Government Data Portal), and outdoor PM$_{2.5}$ and PM$_{10}$ concentrations (obtained from Shanghai Air Quality Monitoring Station) as influencing factors for subway indoor PM concentration. Table 2 details the processing of the raw data. Data regarding the four subway stations were subjected to a chi-square test, which indicated significant differences ($p < 0.05$) in PM concentration among these factors (Table S2).

The Random Forest (RF) Regression and Shapley Additive Explanations (SHAP) value analysis were applied to quantify the relative contributions and feature importance of different factors to the PM concentration inside subway stations, as shown in Fig. 4. Door type, outdoor PM$_{10}$, and train arrival interval were identified as the highest importance, suggesting that indoor mechanical wear and the exchange of outdoor air have the highest association with subway indoor PM. Other notable factors, including outdoor temperature, depth, and passenger flow, were also found to play an important role, with different rankings for PM$_{2.5}$ and PM$_{10}$. Furthermore, Spearman's correlation analysis was performed on these feature values to determine the direction of their impacts on PM$_{2.5}$ and PM$_{10}$ (Table S3). The results reveal that there is a positive and significant correlation between indoor PM and door type (0.62, $p < 0.05$) and outdoor PM$_{10}$ (0.44, $p < 0.05$), with a negative correlation with train arrival interval (−0.43, $p < 0.05$), which aligns with the findings from the SHAP analysis.

The Partial Dependency Plot (PDP) was further adopted to investigate the marginal effect of the top influencing factors on PM$_{2.5}$ and PM$_{10}$ inside the subway (Fig. 5). An abrupt decline of PM$_{2.5}$ and PM$_{10}$ concentrations can be found along with the enhancement of the screen door height (Fig. 5a, d), showing that a lower subway PM level can be achieved with the isolation between the train track and passengers on the platform using PSD. Figure 5b, e illustrates the effects of outdoor PM$_{10}$ pollution on indoor PM$_{2.5}$ and PM$_{10}$. As outdoor PM$_{10}$ increases, the SHAP value for both indoor PM$_{2.5}$ and PM$_{10}$ exhibits a close to linear increase, indicating that higher outdoor PM concentrations are associated with higher indoor PM concentrations. In contrast, Fig. 5c, f demonstrates the negative marginal effects of train arrival interval on indoor PM$_{2.5}$ and PM$_{10}$. The results show that indoor PM concentration decreases notably as the train arrival interval increases, further supporting the finding that subway mechanical wear is one of the major influencing factors to the subway IAQ.

Machine learning-assisted estimation of city-wide subway PM and exposure
This section evaluates the city-wide spatial distribution of PM$_{2.5}$ and PM$_{10}$ across subway stations in Shanghai and investigates their associations with
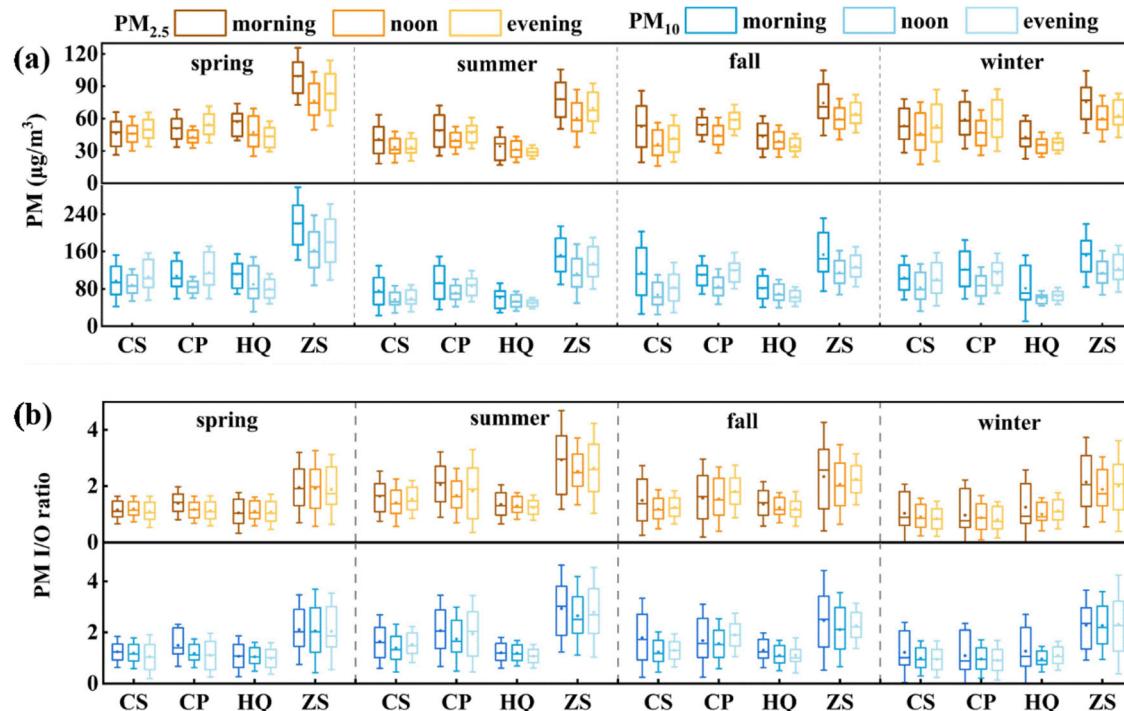
Fig. 3 | Indoor concentrations and PM I/O ratio. a $PM_{2.5}$ (yellow) and PM_{10} (blue) at four stations in the morning (rush hour), noon (non-rush hour), and evening (rush hour) across four seasons; **b** the I/O ratios of $PM_{2.5}$ (yellow) and PM_{10} (blue) at four stations in the morning, noon, and evening across four seasons.

Table 2 | Details of each variable before analysis

Feature factors	Data source and pre-processing
Passenger flow	The daily passenger flow across different stations from October 21 to October 27, 2023, was estimated based on the scaling coefficients derived from the available passenger flow data on January 15, 2018 (Shanghai Government Data Portal).
Day of the week	Weekdays and weekends.
Time of the day	Morning rush hour (7:30 am to 9:30 am), noontime non-rush hour (10:00 am to 3:00 pm), and evening rush hour (4:30 pm to 8:00 pm).
Door type	The automatic screen door type on each platform is sorted according to the height of the screen door, from half to full.
Depth	The excavation depth of the standard section was surveyed by personal conversation with subway staff. This study classified the depth (in meters) into three levels using the K-means clustering method.
Train arrival interval	The time interval between two consecutive trains arriving at the station was derived from the official website of Shanghai Subway (https://service.shmetro.com/hcskb/index.htm).
Outdoor $PM_{2.5}$	Concurrent outdoor $PM_{2.5}$ near each subway station. The data was derived from the Shanghai Air Quality Monitoring Station, with spatial connection assigned by ArcGIS.
Outdoor PM_{10}	Concurrent outdoor PM_{10} near each subway station. The data was derived from the Shanghai Air Quality Monitoring Station, with spatial connection assigned by ArcGIS.
Temperature	Temperature outside each subway station, which was derived from the Shanghai Meteorological Bureau, with spatial connection assigned by ArcGIS.
RH (Relative Humidity)	RH outside each subway station, which was derived from the Shanghai Meteorological Bureau, with spatial connection assigned by ArcGIS.

the surrounding urban living environment. By applying the fitted RF model established in Section 2.2, $PM_{2.5}$ and PM_{10} concentrations were estimated across all subway stations in Shanghai. Passenger exposure to $PM_{2.5}$ and PM_{10} was also estimated using the equation $E = N \times C \times IR \times ET$ where E is the total daily exposure of the group (μg); N represents the number of passengers; C is the calculated PM concentration inside the subway (μg/m³); IR is the inhalation rate, which is identified as 0.498 m³/h according to the China Population Exposure Manual[41]; ET is the exposure time at the subway platform, which took the passenger waiting time, as half of the subway headway (hour) by assuming that passenger arrival time is independent from *train arrival time*. We further transfer the absolute exposure to the relative exposure level based on the quantile of the absolute for a clear comparison across spatiotemporal spans.

A spatial visualization of subway PM and the exposure level integrated with multiple categories of POI (Fig. S2), including residential, healthcare, education, and commercial, was performed. As illustrated in Fig. 6, each subway station was represented using the hexagonal unit, and the color intensity represents the estimated $PM_{2.5}$ concentration (Fig. 6a–c) or passenger exposure level (Fig. 6d–f) during rush hour and non-rush hour on weekdays. Other distributions of $PM_{2.5}/PM_{10}$ concentration and exposure level on weekends and weekdays can be found in Figs. S3–S6.

The concentrations of $PM_{2.5}$ and PM_{10} were significantly higher during the morning and evening rush hours compared to non-peak hours at noon on weekdays across the entire city (Fig. 6a–f). However, it should be noted that passenger exposure to $PM_{2.5}$ and PM_{10} showed distinct patterns during the investigated time periods, which demonstrated a higher exposure
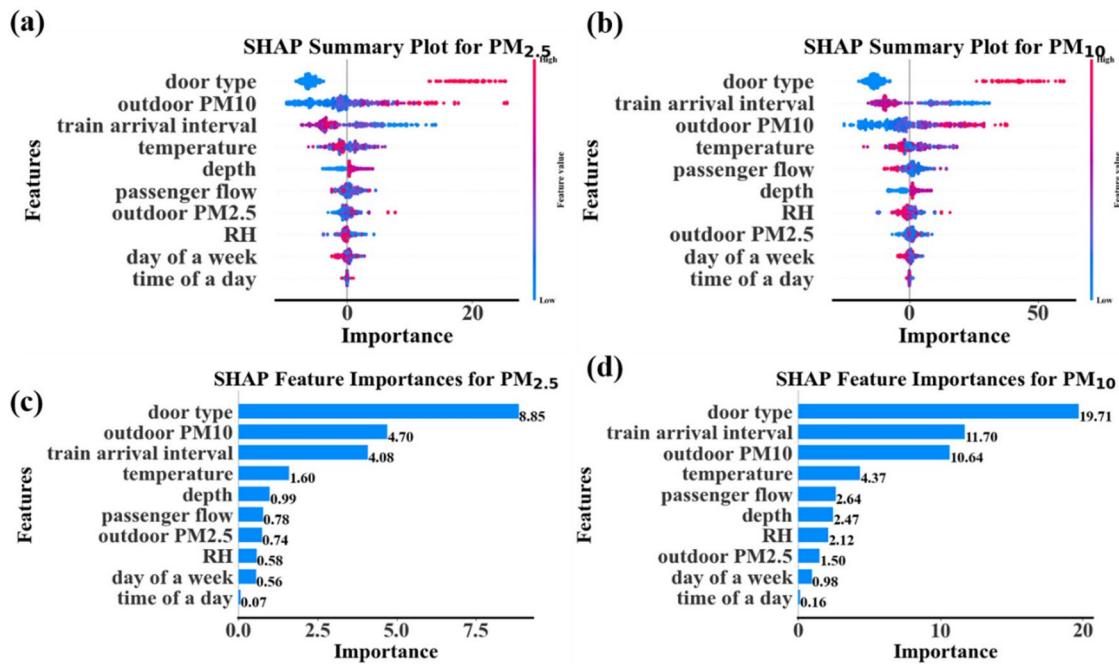
Fig. 4 | Feature importance analysis. SHAP value analysis for influencing factors on the subway platform **a** PM$_{2.5}$ and **b** PM$_{10}$. Feature importance analysis for **c** PM$_{2.5}$ and **d** PM$_{10}$ on the subway platform. Door type has the highest importance for both pollutants, followed by outdoor PM$_{10}$ and train arrival intervals.

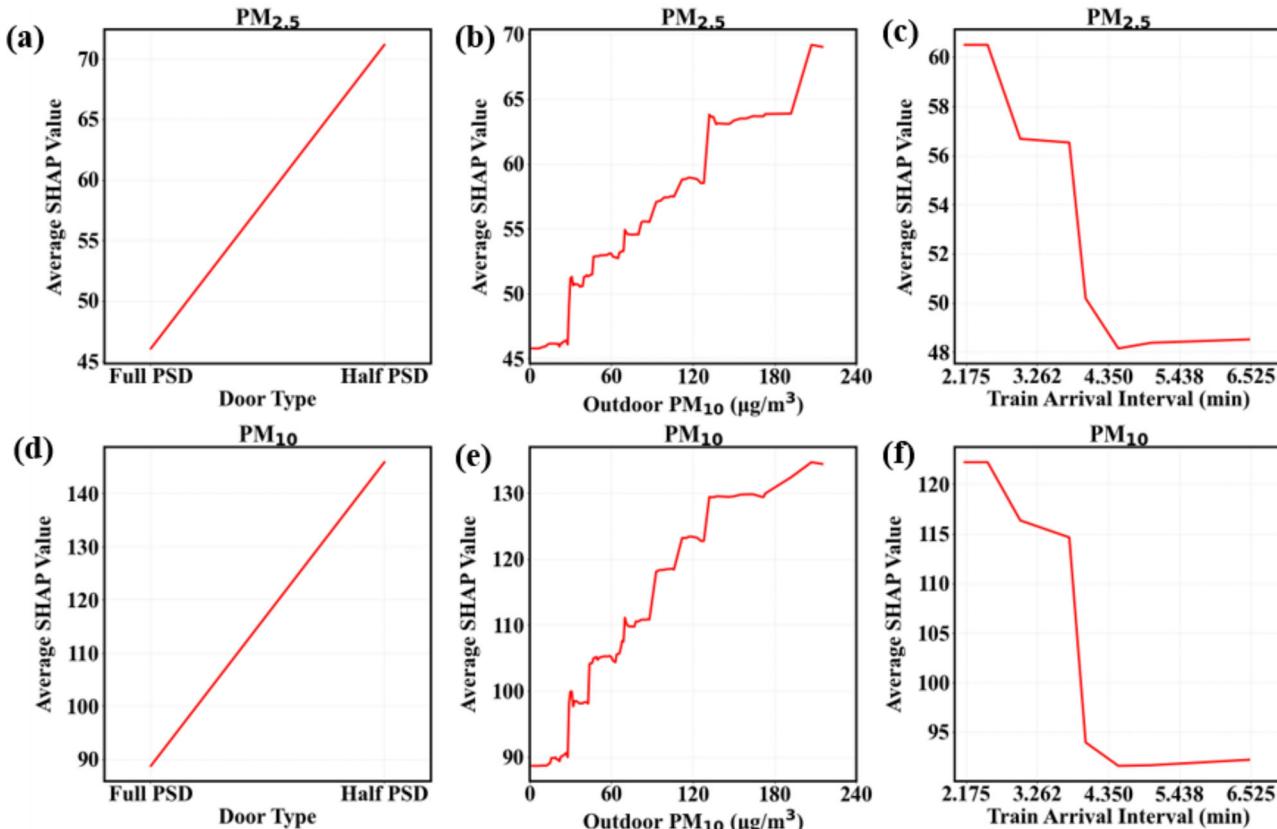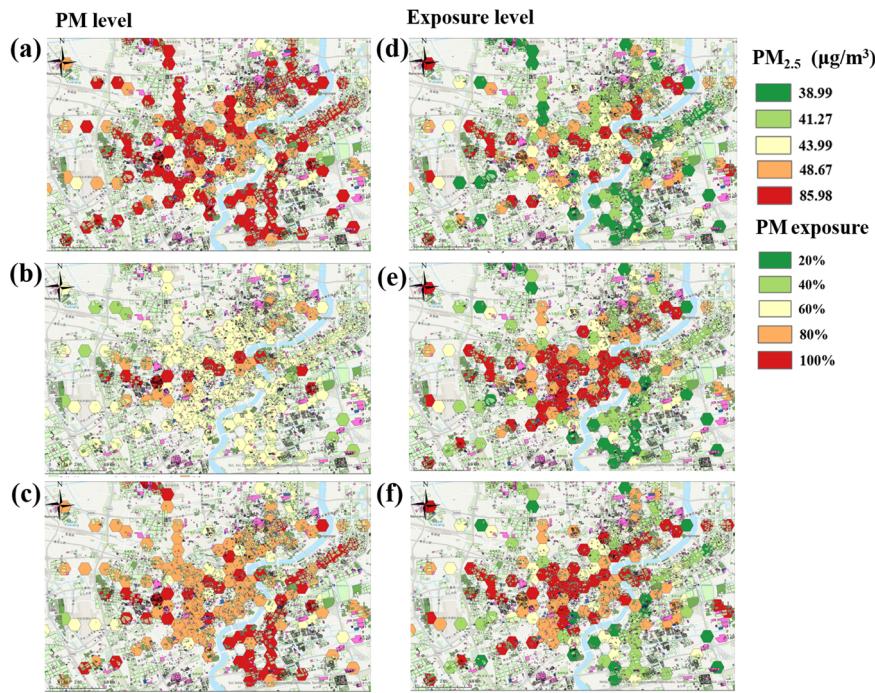**Fig. 5 | Partial Dependency Plot (PDP) analysis.** The figures present PDP analysis for marginal effects from platform screen door (PSD) on **a** PM$_{2.5}$ and **d** PM$_{10}$, outdoor PM$_{10}$ on **b** PM$_{2.5}$ and **e** PM$_{10}$, and train arrival interval on **c** PM$_{2.5}$ and **f** PM$_{10}$. Half PSD leads to higher PM concentrations compared to full PSD. The PM concentrations also increase with higher outdoor PM levels and higher train arrival frequency.

Fig. 6 | Platform PM$_{2.5}$ and the passenger exposure estimated in the Shanghai subway network. The figures illustrate PM$_{2.5}$ concentrations estimated during the **a** morning, **b** noon, and **c** evening episodes on weekdays, respectively, and passenger PM$_{2.5}$ exposure level during the **d** morning, **e** noon, and **f** evening, respectively. The exposure levels to PM were categorized into five percentile-based categories for visualization (20%, 40%, 60%, 80%, and 100%).

burden during the noontime and evening rush hours. Compared with rush hours, the decrease in train departure frequency during non-rush hours at noon significantly prolongs passenger waiting time. Although the instantaneous PM concentration could be relatively low during noontime, the prolonged platform occupancy time leads to an increase in exposure levels over noontime. Therefore, the PM exposure on subway platforms during non-rush hours at noon exceeded the value during rush hours in the morning, highlighting the need for excessive ventilation during the noon non-peak.

Based on the overall distribution of all four types of POIs across Shanghai (Fig. S2), the total area and number of the four different POIs are summarized in Fig. 7a, b, respectively. The associations between different types of POIs with PM concentrations and exposure levels were analyzed using Spearman's correlation (Fig. 7, Tables S4 and S5). During weekdays, the correlations between the number of residential POIs and the concentrations of PM$_{2.5}$ and PM$_{10}$ were relatively low and mostly insignificant. However, during the noon period on weekdays, the correlations for both PM$_{2.5}$ and PM$_{10}$ become significant ($p < 0.05$), indicating that a higher pollutant exposure is associated with an increased residential area around, which can be attributed to a higher passenger volume as illustrated by the linear regression analysis (Fig. S7) and Pearson correlation analysis (Fig. S8). The correlation pattern for healthcare and educational POIs was found to be similar, such that the association becomes higher at noontime compared to morning and evening ($p < 0.001$). Over the weekend, the overall correlation for healthcare and educational POIs increases, especially during noon and evening, indicating higher pollutant concentrations in subway areas around healthcare and educational districts during these periods. Pearson's correlation test (Table S6) shows that the distance to the nearest healthcare spot is negatively related to the exposure level at a 95% confidence interval, especially during the period of weekday noon. The finding highlights that a higher vulnerable group at healthcare locations is more likely to be exposed to the indoor PM at subway stations. Special attention should be given to those stations regarding their cleanness, retrofitting, and ventilation.

In addition, the present study included area and number-weighted average PM concentration and exposure, which quantifies the combined effects of both concentration intensity and spatial distribution, as illustrated in Fig. 7c, d (details can be found in Figs. S9–S12). After normalizing the data using residential areas, it was found that healthcare POIs showed the most

significant association with the area-weighted average PM concentration, followed by commercial areas, education areas, and residential areas. In terms of PM exposure assessment, the overall trend for healthcare and commercial zones aligned with concentration levels, significantly exceeding those of education and residential areas. The POI number-weighted analysis provided an additional perspective, which also indicated that commercial areas contributed the most regarding pollution concentration and exposure risk. Commercial areas, despite their limited spatial distribution, exhibited the strongest association with pollution, a trend particularly pronounced during weekends due to shopping activities, where exposure risks increased significantly.

Discussion

This study investigated the spatiotemporal variations of PM$_{2.5}$ and PM$_{10}$ across typical subway stations in Shanghai and identified key factors influencing indoor PM levels through interpretable machine learning approaches. The trained machine learning model was applied to visualize city-scale concentration and exposure level of subway PM with POIs. PSD type, outdoor PM$_{10}$, and train arrival frequency are identified as dominant influencing factors to subway indoor PM variability based on the machine learning analysis, highlighting the urgent need for infrastructure upgrades, such as retrofitting aging stations with full PSD systems, optimizing HVAC efficiency, and adopting cleaner rail materials to reduce mechanical wear. Additionally, tailored strategies for high-risk POI zones, such as commercial, healthcare, and educational areas near city centers, should be prioritized by enhanced real-time air quality monitoring, ventilation, and public awareness to realize the public health benefits of green travelers in urban areas.

The findings from this study indicate that subway PM can be mainly divided into indoor sources, which are mainly contributed by in situ wheel-rail wear, brake pad wear, and pantograph/contact grid wear[10–12,42,43], and outdoor sources, including indoor and outdoor air exchange via the ventilation system and entrance[6,44]. Higher levels of PM concentrations during rush hours with increased passenger demand, as well as in colder seasons when outdoor air pollution levels were observed in the current work. Compared to the previous work (Table S4), the PM$_{2.5}$ concentration in Shanghai subway stations was observed to be higher than those measured in Chengdu (7.33–82.69 μg/m^3)[45], Tianjin (35.5–43.6 μg/m^3)[16], and Beijing
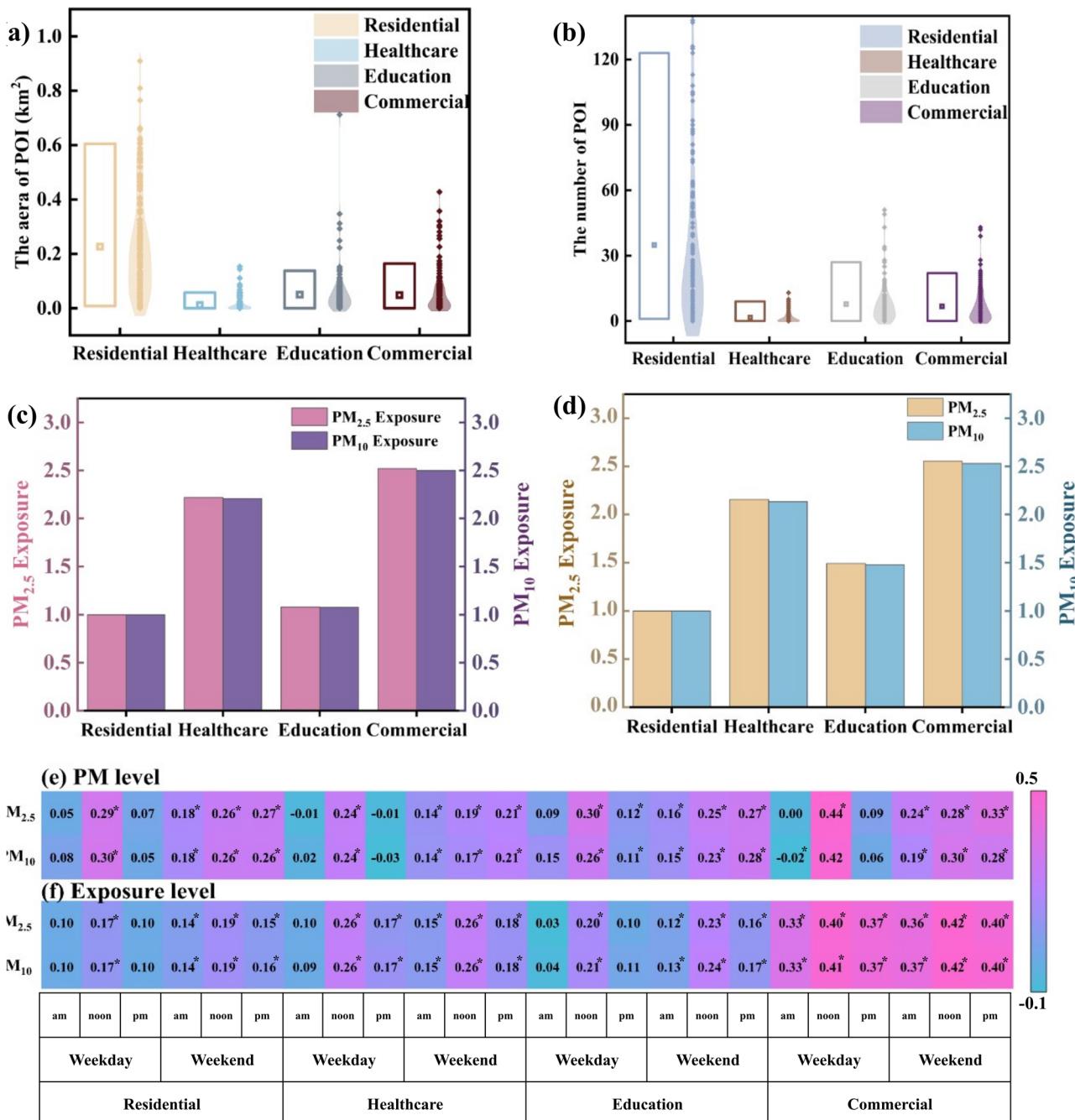
Fig. 7 | The association between POI and subway platform PM exposure. The figures present: **a** total aera of four different POIs in Shanghai; **b** the number of four types of POIs in Shanghai; **c** area-weighted average exposure level to PM$_{2.5}$ and PM$_{10}$ among four types of POI (normalized to the residential POI); **d** number-weighted average exposure level to PM$_{2.5}$ and PM$_{10}$ among four types of POI (normalized to the residential POI); Spearman correlation analysis of **e** PM concentration and **f** PM exposure level at different time periods with four types of POI.

(20–90 μg/m³)[29], but lower compared to New York (6.81–317.2 μg/m³)[46], Boston (7.2–638 μg/m³)[46], and Washington (5.0–720 μg/m³)[46]. Given their varied industrialization and economic development levels, this comparison provides a global view of the subway station indoor IAQ across the world. Overall, the PM level in Shanghai subway stations falls into the higher range, especially for stations like ZS where half PSD is installed.

By performing the influencing factor analysis, train arrival interval and PSD type, both associated with the dispersion of indoor mechanical wear, were ranked as the dominant PM sources on subway platforms. Physical separation (i.e., full PSD) between the subway train and the platform could be the most effective method to reduce passenger exposure to PM[16,47]. In

addition, PSDs also help isolate other types of airborne pollutants, such as bioaerosols. Hwang et al.[48] reported that the bacterial concentrations in Korean stations without PSDs were higher than those with PSDs. Based on findings from this study, we advocate for the renovation and installation of full PSDs, especially in aged subway stations, to improve the health benefits of passengers.

Another critical message delivered by the current study is that the influences come from the infiltration of outdoor PM$_{2.5}$ and PM$_{10}$ by passenger flow and the indoor-outdoor air exchange facilitated by mechanical ventilation systems in the semi-enclosed subway microenvironments[16,49,50]. Meanwhile, HVAC systems, which effectively filter indoor PM, can reduce

indoor PM concentrations by up to 30% when in operation[17]. Hence, upgrading or renovating HVAC systems can significantly reduce indoor PM levels and control the potential impacts from outdoor air pollution on subway indoor PM levels. Furthermore, regular cleaning activities, such as mopping and vacuuming, have also been shown to effectively reduce PM concentrations in subway stations[4,30,51].

Here, we also observed that a higher train arrival frequency is associated with a higher level of subway PM. When the interval between train arrivals is shorter than 4.35 min, the impact on PM becomes more pronounced. In such cases, a significant amount of PM generated by brake wear and air exchange between the tunnel and platform cannot be adequately filtered by the HVAC system in a timely manner. The piston wind effect caused by the coming train will resuspend the particles, resulting in persistently high PM concentrations at the platform. Moreover, while reducing train intervals may help mitigate PM accumulation, prolonged waiting durations on the platform can raise the overall exposure levels for commuters. Consequently, an optimization of train arrival frequency is highly needed in different subway systems around the world.

This study also highlights passenger exposure across a variety of POIs when addressing the improvements of subway IAQ. Residential areas tend to be the most vulnerable places in terms of subway PM exposure. Another finding from Beijing also depicts that residential quarters are one of the highest POIs that exhibit high individual exposures to NO_x and $PM_{2.5}$ for cyclists[36]. These findings underscore the significance of improving both indoor and outdoor air quality around residential POIs for the safeguarding of green commuters. While for areas such as commercial, educational, and healthcare zones, where higher PM concentrations are present due to the proximity of dense traffic and pollution sources, the exposure falls much lower than that of residential areas. Kan et al.[52] integrated traveling trajectory GPS records with personal exposure analysis, revealing the necessity of considering the spatiotemporal services of traveling facilities into the exposure, besides individual traveling diaries. Although the referred study focuses on road traffic, its result is consistent with our study, which suggests that improved train scheduling and the operation of the HVAC system inside the station should be further investigated with the combination of daily traveling distribution across different land use types.

In recent years, vigorous efforts have been made to promote sustainable urban travel. However, as one of the primary modes of urban transport, subways often have higher air pollution levels than outdoor environments, raising concerns about social equity and public health[8,53–55]. It is thus crucial to ensure that passengers, who play a key role in supporting green urban transport, are protected from excessive exposure. This study highlights the need for protecting subway passengers and workers from poor IAQ. Future improvements should be function-specific, considering the unique challenges posed by the characteristics of the stations and the impact of different types of POIs. A comprehensive approach, integrating technological upgrades like full PSD installations, HVAC improvements, and real-time air quality monitoring with urban planning and public health measures tailored to POI distributions, can ensure healthier environments for subway passengers and workers.

It should be acknowledged that this study has a limitation on the spatiotemporal scope of PM across subway lines and train cabins, and a more accurate prediction result is limited by the number of study sites and the data spatiotemporal resolution. In addition, estimated passenger flow data may not accurately represent actual ridership during the measurement periods (2023–2024), further contributing to potential modeling deviations. To minimize such errors in future research, direct on-site monitoring of these parameters using appropriate instrumentation is strongly recommended to enhance data accuracy and model reliability. Integrating composition and health risk assessments of PM toxicity will further elucidate the public health implications of prolonged exposure in subway microenvironments. It is worth noting that the mechanism by which POI affects the exposure level at indoor subway platforms remains unclear due to the lack of full-chain passenger activity data. Addressing these gaps will support the formulation of holistic policies to achieve equitable air quality

improvements, aligning with global efforts to promote sustainable and health-conscious urban transit systems.

Methods
Field measurement
The Shanghai Metro system is currently the longest metro network (808 km) with one of the highest annual passenger loadings in the world. From 2023 to 2024, the study was divided into four distinct phases to capture seasonal variations, with each phase consisting of 2 weeks real-time monitoring of indoor and outdoor subway air quality (March 20th–April 2nd, July 17th–July 30th, October 16th–October 29th in 2023, and January 8th–January 21st in 2024 representing spring, summer, autumn, and winter, respectively). During the first week, measurements were conducted from Monday to Sunday at CS and CP stations in Line 7. In the second week, measurements were shifted to HQ and ZS in Line 2. Line 2 is a conventional railway started from 2000 with mostly half PSD installed, while Line 7 is relatively new, started from 2009 with full PSadopted the RF model to estimate the PM concentration based on the model performanceD installed. On each monitoring day, real-time recording of both subway platform IAQ ($PM_{2.5}$, PM_{10}, RH, temperature) and nearby outdoor air quality ($PM_{2.5}$, PM_{10}, RH, temperature) was performed. The location of the monitoring devices was chosen according to undisturbed measurement from nearby traffic. Inside the subway, the platform monitoring site was set 3 m away from the PSD. The outdoor monitoring site was set at the nearest entrance of the target platform, both 3 m away from the main street and the subway entrance.

All tests took place during three distinct time periods, including morning rush hour (7:30 to 9:30 am), non-rush hours (10:00 am to 3:00 pm), and evening rush hour (4:30 to 8:00 pm). Two portable PM monitors (Dylos, DC1700-PM air quality monitors) were used for simultaneous measurements of I/O $PM_{2.5}$ and PM_{10} during each test. The PM inlets were placed at roughly 1.5 m above the ground level to represent normal passenger exposure height. The Dylos PM monitors were calibrated against a Tapered Element Oscillating Microbalance combined with a Filter Dynamic Measurement System (TEOM-FDMS, TEOM 1405-F, Thermo Fisher Scientific Inc., USA) across four seasons (Fig. 8).

Influencing factor data processing
Influencing factors that were included in this study cover different passenger demand, temporal variations, design of the platform, train operations, outdoor air quality, and meteorology. The processing of the data is listed in Table 2.

The passenger flow across all stations in Shanghai between October 16th and 22nd, 2023, when our PM data were monitored, was estimated by applying the scaling coefficients (Eq. 1) derived from available passenger data for Shanghai subway stations on January 15th, 2018. The resulting data were considered as representative of the passenger flow across station platforms in Shanghai from Monday to Sunday during the observation period in the present study.

$$q_{(x,2023)} = \frac{Q_{(x,2023)}}{Q_{(x,2018)}} q_{(x,2018)} \tag{1}$$

Where x is the day of the week; $q_{(x,2023)}$ is the passenger flow of a certain platform in 2023; $q_{(x,2018)}$ is the passenger flow of this platform in 2018; $Q_{(x,2023)}$ is the passenger flow of the corresponding line at this platform on a certain day in 2023; $Q_{(x,2018)}$ is the passenger flow of the corresponding line at this platform on a certain day in 2018.

In this study, outdoor $PM_{2.5}$ and PM_{10} concentration data were obtained from measurements conducted at various monitoring stations under the Shanghai Air Quality Monitoring Network, while outdoor temperature and relative humidity (RH) data were provided by meteorological observation stations operated by the Shanghai Meteorological Bureau. To establish an association model between metro stations and ambient environmental factors, we first integrated geographic information from air
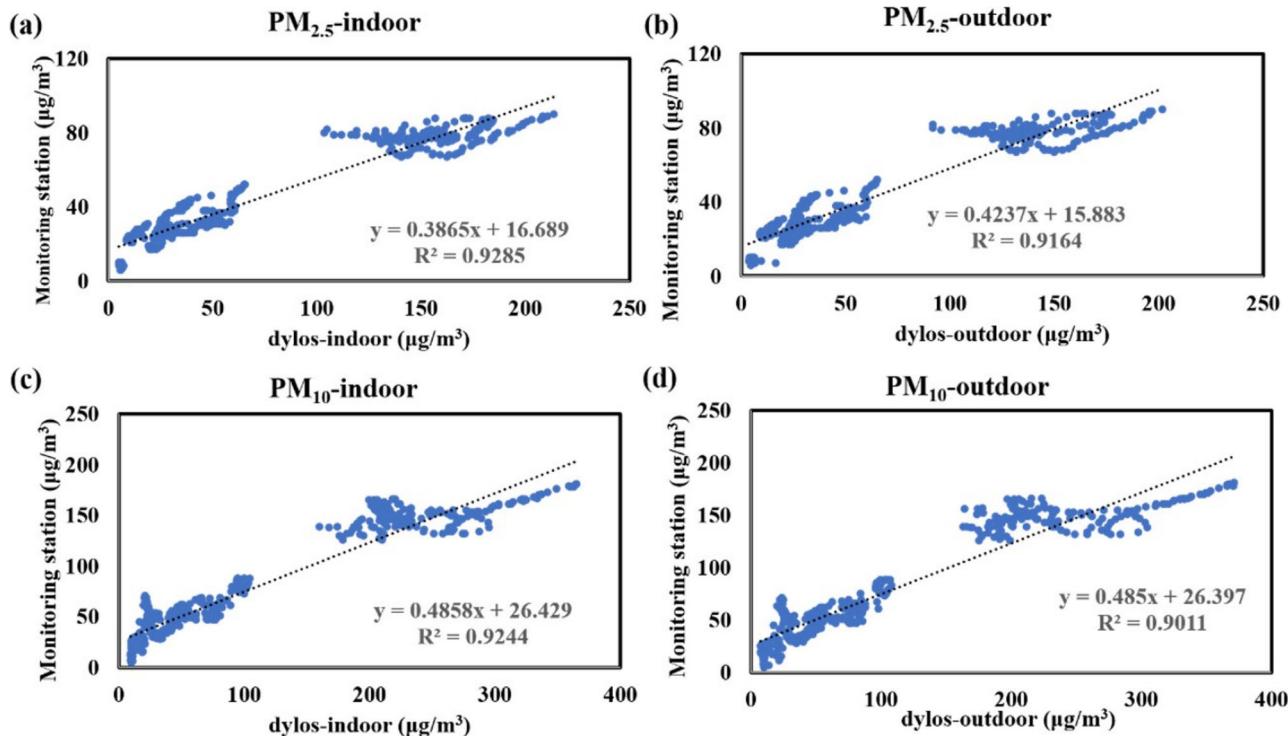
Fig. 8 | Measurement data calibration. Calibration between Dylos (DC1700-PM) and TEOM (TEOM-FDMS, TEOM 1405-F) collocated at the monitoring station during summer and winter for **a** indoor PM$_{2.5}$, **b** outdoor PM$_{2.5}$, **c** indoor PM$_{10}$, and **d** outdoor PM$_{10}$.

quality monitoring stations, meteorological observation stations, and Shanghai metro stations within the ArcGIS Pro platform. Based on spatial proximity principles, buffer analysis and nearest-neighbor matching algorithms were applied to spatially link each metro station with its corresponding nearest air quality monitoring station and meteorological observation site. This spatial matching process enabled the assignment of outdoor PM$_{2.5}$, PM$_{10}$ concentrations, temperature, and RH values to each metro station, thereby establishing a data foundation for subsequent exposure assessments.

Interpretable machine learning model
Both RF and Logistic Regression (LR) were used for PM prediction in this study. The RF model aggregates the prediction results of multiple decision trees to generate the final output. Each individual decision tree employs the mean square error (MSE) as the splitting criterion, and the model iteratively selects the optimal features and segmentation points by minimizing the MSE. The final prediction is obtained by averaging the outputs of all decision trees (Eq. 2):

$$\hat{y}_1 = \frac{1}{T}\sum_{t=1}^{T} h_{1t}(x), \hat{y}_2 = \frac{1}{T}\sum_{t=1}^{T} h_{2t}(x) \quad (2)$$

Where $\hat{y}_1$ and $\hat{y}_2$ are the predicted concentration of PM$_{2.5}$ and PM$_{10}$; x is the features; T is the number of decision trees in the RF, and in this article, the optimal number of decision trees is determined to be 200 through grid search; $h_{1t}(x)$ and $h_{2t}(x)$ are the predicted value of PM$_{2.5}$ and PM$_{10}$ for the tth tree.

For the LR model, input feature values are linearly combined, and the results are mapped to probability values using the Sigmoid function (Eq. 3):

$$P(y = 1|x) = \frac{1}{1 + e^{-(w_0 + w_1 x_1 + w_2 x_2 + \cdots + w_p x_p)}} \quad (3)$$

Where $P(y = 1/x)$ is the probability that the target y takes on a value of 1 given the feature x; w is the model weight; x is the feature. The model

optimizes weights by maximizing the logarithmic likelihood function (Eq. 4):

$$L(w) = \sum_{i=1}^{N} \left[y_i \log\log(P(y_i = 1|x_i)) + (1 - y_i)\log\log(1 - P(y_i = 1|x_i)) \right] \quad (4)$$

Due to the requirement for categorical variables in LR, PM$_{2.5}$ was categorized into 0–35 μg/m^3, 35–75 μg/m^3, and >75 μg/m^3. PM$_{10}$ was categorized into 0–75 μg/m^3, 75–150 μg/m^3, and >150 μg/m^3 (Ambient Air Quality Standards (GB 3095-2012)). We compared the performance of the RF model and the LR model, and eventually adopted the RF model to estimate the PM concentration based on the model performance (illustrated in Fig. 9).

The SHAP value analysis was applied as the explanation of the trained RF model[56]. The SHAP value not only indicates the contribution of each feature to the model's prediction, but also helps identify key influencing factors. When applying the SHAP value (Eq. 5).

$$\Phi_i = \sum_{S \subseteq F\{i\}} \frac{|S|!(|F| - |S| - 1)!}{|F|!} \cdot \left(f(S \cup \{i\}) - f(S) \right) \quad (5)$$

Where F is the set of all features; S is the feature subset, $S \subseteq F$; Φ_i is the Shapley value of the feature for the input x; $f(S)$ is the predicted value of the model under the feature subset S; $|S|$ is the size (number of features) of the feature subset S. $\frac{|S|!(|F|-|S|-1)!}{|F|!}$ is a weight term used to fairly allocate the contribution of feature i. In addition, in order to better evaluate the marginal effects of each feature, we also used PDPs to explore the marginal impact of influencing factors on the prediction results.

Model application for the network subway indoor PM estimation
To apply the trained model for PM concentration prediction in the Shanghai subway, relevant features were collected from various subway stations (in addition to our four studied stations) in Shanghai, including "Passenger flow," "Temperature," "Day of the week," "Time of day," "PSD," "Depth," "Train arrival interval," "Outdoor PM$_{2.5}$," "Outdoor PM$_{10}$," and
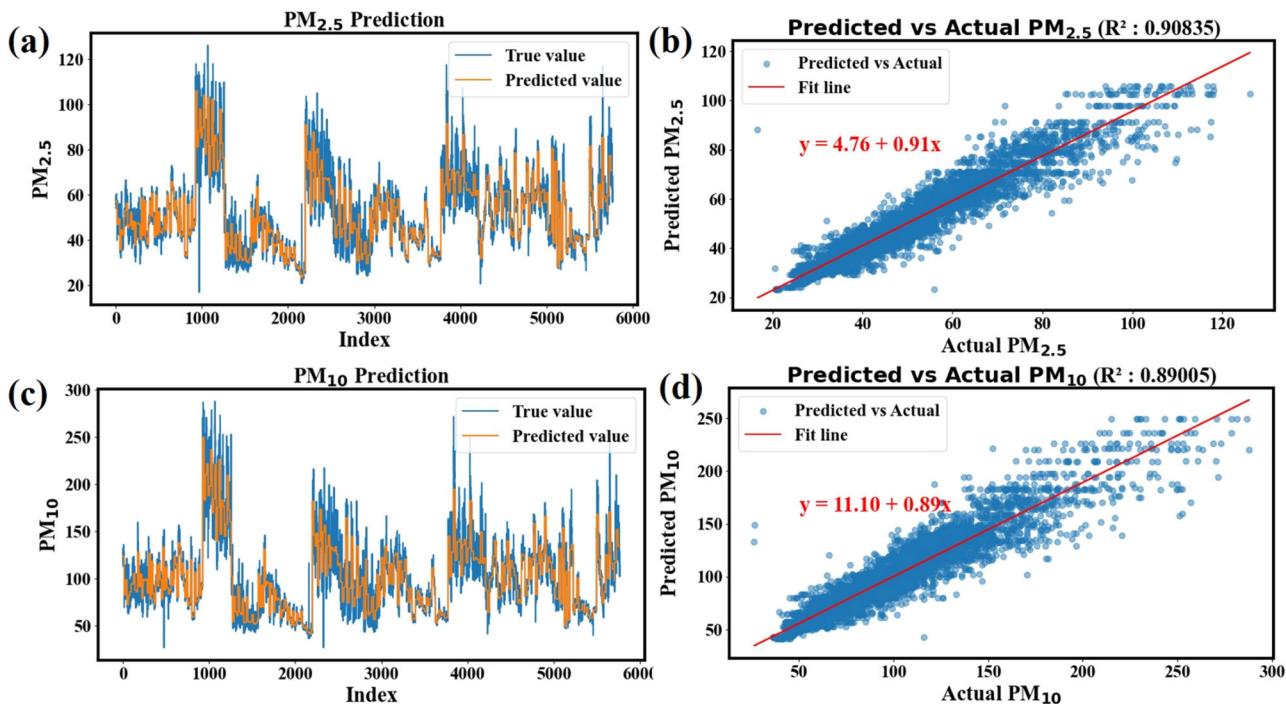
Fig. 9 | Random Forest model performance. Random Forest fitted time series of **a** PM$_{2.5}$ and **c** PM$_{10}$. **b** Predicted PM$_{2.5}$ value plotted against the measured PM$_{2.5}$ value. **d** Predicted PM$_{10}$ value plotted against measured PM$_{10}$ value.

"Relative Humidity." These data were then processed in the same way as described in Table 2. In addition to the estimated concentrations of PM$_{2.5}$ and PM$_{10}$, we calculated the PM exposure inside the subway using Eq.6 for the estimated PM concentration:

$$E = N \times C \times \text{IR} \times \text{ET} \qquad (6)$$

Where E is the total daily exposure of the group (μg); N represents the number of passenger; C is the calculated PM concentration inside the subway (μg/m³); IR is the inhalation rate, which is identified as 0.498 m³/h according to the China Population Exposure Manual[57]; ET is the exposure time in the subway (h). In this study, the exposure time is defined as the waiting time of passengers on the platform, which was approximated as half of the average train arrival interval during each observation period.

Spatial distribution of subway PM$_{2.5}$ and PM$_{10}$ and its association with urban point-of-interests

Following the successful development and validation of our predictive model, this study quantified the average PM concentrations across multiple subway stations in Shanghai across six distinct temporal segments: morning rush hours, midday off-peak periods, and evening rush hours for both weekdays and weekends. To effectively communicate the spatial and temporal variations in PM concentrations and associated exposure levels, we employed advanced geospatial visualization techniques using ArcGIS Pro (version 3.1.5). The map of Shanghai city was first segmented into regular hexagonal grids, each covering an area of 1 km². This grid system allows for an effective integration of the predicted PM data with specific spatial locations, enhancing the accuracy of spatial analysis. The data aligned with these hexagonal grids were utilized to highlight regions exhibiting either elevated or reduced PM concentrations and passenger exposure. Additionally, POI data obtained from OpenStreetMap, including residential, educational, healthcare, and commercial areas, were incorporated into the GIS mapping. This integrated approach not only assesses IAQ across the entire Shanghai subway network but also identifies functional areas that may correlate with the concentrations of PM$_{2.5}$ and PM$_{10}$. To further evaluate the impact of PM on the subway near different types of POI, including PM concentration and

PM exposure level, this study estimated the POI area and POI number-weighted PM concentration and PM exposure (Eqs. 7–10)

$$\text{PM}_{\text{POI}-\text{area weighted}} = \frac{\sum_{i=1}^{n} \left(\text{area}_{(\text{POI},i)} \times \text{PM}_i \right)}{\sum_{i=1}^{n} \text{area}_{(\text{POI},i)}} \qquad (7)$$

$$\text{PM Exposure}_{\text{POI}-\text{area weighted}} = \frac{\sum_{i=1}^{n} \left(\text{area}_{(\text{POI},i)} \times \text{PM Exposure}_i \right)}{\sum_{i=1}^{n} \text{area}_{(\text{POI},i)}} \qquad (8)$$

$$\text{PM}_{\text{POI}-\text{number weighted}} = \frac{\sum_{i=1}^{n} \left(\text{number}_{(\text{POI},i)} \times \text{PM}_i \right)}{\sum_{i=1}^{n} \text{number}_{(\text{POI},i)}} \qquad (9)$$

$$\text{PM Exposure}_{\text{POI}-\text{number weighted}} = \frac{\sum_{i=1}^{n} \left(\text{number}_{(\text{POI},i)} \times \text{PM Exposure}_i \right)}{\sum_{i=1}^{n} \text{number}_{(\text{POI},i)}} \qquad (10)$$

Where area$_{(\text{POI},i)}$ represents the total area of a certain type of POI within the i hexagon; PM$_i$ and PMExposure$_i$ the PM concentration or PM exposure of the i hexagon; number$_{(\text{POI},i)}$ represents the number of POIs of a certain type within the i hexagon.

Data availability
The datasets generated and/or analyzed during the current study are not publicly available due to regional security issues, but are available from the corresponding author on reasonable request and with the permission of the Shanghai Government Data Portal.

Code availability
The underlying code for this study is available to qualified researchers on request from the corresponding author.

References

1. Wang, S. et al. Indoor air quality in subway microenvironments: pollutant characteristics, adverse health impacts, and population inequity. *Environ. Int.* **190**, 108873 (2024).
2. Adams, H., Nieuwenhuijsen, M., Colvile, R., McMullen, M. & Khandelwal, P. Fine particle (PM$_{2.5}$) personal exposure levels in transport microenvironments, London, UK. *Sci. Total Environ.* **279**, 29–44 (2001).
3. Barmparesos, N., Assimakopoulos, V. D., Assimakopoulos, M. N. & Tsairidi, E. Particulate matter levels and comfort conditions in the trains and platforms of the Athens underground metro. *AIMS Environ. Sci.* **3**, 199–219 (2016).
4. Braniš, M. The contribution of ambient sources to particulate pollution in spaces and trains of the Prague underground transport system. *Atmos. Environ.* **40**, 348–356 (2006).
5. Johansson, C. & Johansson, P.-Å Particulate matter in the underground of Stockholm. *Atmos. Environ.* **37**, 3–9 (2003).
6. Martins, V. et al. Exposure to airborne particulate matter in the subway system. *Sci. Total Environ.* **511**, 711–722 (2015).
7. Passi, A., Nagendra, S. M. S. & Maiya, M. P. Characteristics of indoor air quality in underground metro stations: a critical review. *Build. Environ.* **198**, 107907 (2021).
8. Xu, B. & Hao, J. Air quality inside subway metro indoor environment worldwide: a review. *Environ. Int.* **107**, 33–46 (2017).
9. Martins, V. et al. Factors controlling air quality in different European subway systems. *Environ. Res.* **146**, 35–46 (2016).
10. Lee, Y., Lee, Y. C., Kim, T., Choi, J. S. & Park, D. Sources and characteristics of particulate matter in subway tunnels in Seoul, Korea. *Int. J. Environ. Res. Public Health* **15**, 2534 (2018).
11. Perrino, C., Marcovecchio, F., Tofful, L. & Canepari, S. Particulate matter concentration and chemical composition in the metro system of Rome, Italy. *Environ. Sci. Pollut. Res.* **22**, 9204–9214 (2015).
12. Yang, J., Fan, X., Zhang, H., Zheng, W. & Ye, T. A review on characteristics and mitigation strategies of indoor air quality in underground subway stations. *Sci. Total Environ.* **869**, 161781 (2023).
13. Aksenov, A. A. et al. Human Activities Shape Indoor Volatile Chemistry. *Environ. Sci. Technol. Lett.* **10**, 965–975 (2023).
14. Ma, H., Shen, H., Liang, Z., Zhang, L. & Xia, C. In *Proc. 8th International Symposium on Heating, Ventilation and Air Conditioning.* (Springer Berlin, Heidelberg 2014).
15. Huang, S. et al. Characteristics and source identification of fine particles in the Nanchang subway, China. *Build. Environ.* **199**, 107925 (2021).
16. Ren, J. et al. A field study of CO$_2$ and particulate matter characteristics during the transition season in the subway system in Tianjin, China. *Energy Build.* **254**, 111620 (2022).
17. Alves, C. et al. Chapter characterisation of airborne particulate matter in different European subway systems. In *Urban Transport Systems.* InTechOpen (2017).
18. Furuya, K. et al. Seasonal variation and their characterization of suspended particulate matter in the air of subway stations. *J. Trace Microprobe Tech.* **19**, 469–485 (2001).
19. Qiao, T. et al. Preliminary investigation of PM$_1$, PM$_{2.5}$, PM$_{10}$ and its metal elemental composition in tunnels at a subway station in Shanghai, China. *Transp. Res. Part D Transp. Environ.* **41**, 136–146 (2015).
20. Vilcassim, M. J., Thurston, G. D., Peltier, R. E. & Gordon, T. Black carbon and particulate matter (PM$_{2.5}$) concentrations in New York City's subway stations. *Environ. Sci. Technol.* **48**, 14738–14745 (2014).
21. Cha, Y., Tu, M., Elmgren, M., Silvergren, S. & Olofsson, U. Variation in airborne particulate levels at a newly opened underground railway station. *Aerosol Air Qual. Res.* **19**, 737–748 (2019).
22. Han, H., Jang, K.-J., Shin, C. & Lee, J.-Y. In *Proc. Healthy Building* 262–267 (2012).
23. Kim, T.-H., Choi, B.-H., Kang, M.-S. & Lee, H.-J. Removal of iron oxide from indoor air at a subway station using a vegetation biofilter: a case study of Seoul, Korea. *Atmosphere* **12**, 1463 (2021).
24. Querol, X. et al. Variability of levels and composition of PM$_{10}$ and PM$_{2.5}$ in the Barcelona metro system. *Atmos. Chem. Phys.* **12**, 5055–5076 (2012).
25. Suárez, L. et al. Personal exposure to particulate matter in commuters using different transport modes (bus, bicycle, car and subway) in an assigned route in downtown Santiago, Chile. *Environ. Sci. Process. Impacts* **16**, 1309–1317 (2014).
26. Miao, Q., Bouchard, M., Chen, D., Rosenberg, M. W. & Aronson, K. J. Commuting behaviors and exposure to air pollution in Montreal, Canada. *Sci. Total Environ.* **508**, 193–198 (2015).
27. Li, T.-T., Bai, Y.-H., Liu, Z.-R. & Li, J.-L. In-train air quality assessment of the railway transit system in Beijing: a note. *Transp. Res. Part D Transp. Environ.* **12**, 64–67 (2007).
28. Zhang, Y. et al. Rush-hour aromatic and chlorinated hydrocarbons in selected subway stations of Shanghai, China. *J. Environ. Sci.* **24**, 131–141 (2012).
29. Zhang, M. et al. Atmospheric iron particles in PM$_{2.5}$ from a subway station, Beijing, China. *Atmos. Environ.* **283**, 119175 (2022).
30. Mugica-Álvarez, V., Figueroa-Lara, J., Romero-Romo, M., Sepúlveda-Sánchez, J. & López-Moreno, T. Concentrations and properties of airborne particles in the Mexico City subway system. *Atmos. Environ.* **49**, 284–293 (2012).
31. Kim, M. J. et al. Statistical evaluation of indoor air quality changes after installation of the PSD system in Seoul's metro. *Indoor Built Environ.* **20**, 187–197 (2011).
32. Targino, A. C. et al. Spatio-temporal variability of airborne particulate matter in the São Paulo subway. *Build. Environ.* **189**, 107526 (2021).
33. Moreno, T. et al. Subway platform air quality: Assessing the influences of tunnel ventilation, train piston effect and station design. *Atmos. Environ.* **92**, 461–468 (2014).
34. Cao, S. J. et al. An investigation of the PM$_{2.5}$ and NO$_2$ concentrations and their human health impacts in the metro subway system of Suzhou, China. *Environ. Sci. Process. Impacts* **19**, 666–675 (2017).
35. He, S. et al. Commuter health risk and the protective effect of three typical metro environmental control systems in Beijing, China. *Transp. Res. Part D Transp. Environ.* **62**, 633–645 (2018).
36. Huang, S. et al. Source apportionment of fine particulate matter at different underground sites in the Chengdu metro system in summer. *Build. Environ.* **248**, 111081 (2024).
37. Hwang, S., Kim, S.-Y., Choi, S., Lee, S. & Park, D.-U. Correlation between levels of airborne endotoxin and heavy metals in subway environments in South Korea. *Sci. Rep.* **11**, 17086 (2021).
38. Juraeva, M., Ryu, K. J., Jeong, S.-H. & Song, D. J. Influences of the train-wind and air-curtain to reduce the particle concentration inside a subway tunnel. *Tunn. Undergr. Space Technol.* **52**, 23–29 (2016).
39. Morawska, L., He, C., Hitchins, J., Gilbert, D. & Parappukkaran, S. The relationship between indoor and outdoor airborne particles in the residential environment. *Atmos. Environ.* **35**, 3463–3473 (2001).
40. Naddafi, K. et al. Concentrations of airborne particulate matters in underground and surface stations of Tehran subway system. *J. Air Pollut. Health* **1**, 7–14 (2016).
41. Duan, J. et al. Distinctions in source regions and formation mechanisms of secondary aerosol in Beijing from summer to winter. *Atmos. Chem. Phys.* **19**, 10319–10334 (2019).
42. Jung, H.-J. et al. Source identification of particulate matter collected at underground subway stations in Seoul, Korea using quantitative single-particle analysis. *Atmos. Environ.* **44**, 2287–2293 (2010).
43. Sundh, J., Olofsson, U., Olander, L. & Jansson, A. Wear rate testing in relation to airborne particles generated in a wheel-rail contact. *Lubr. Sci.* **21**, 135–150 (2009).

44. Nieuwenhuijsen, M., Gomez-Perales, J. & Colvile, R. Levels of particulate air pollution, its elemental composition, determinants and health effects in metro systems. *Atmos. Environ.* **41**, 7995–8006 (2007).

45. Huang, S. et al. Characteristics and health risk assessments of fine particulate matter at the overground and underground subway sites in Chengdu. *Build. Environ.* **242**, 110577 (2023).

46. Luglio, D. G., Huynh, T., Saporito, A. & Gordon, T. Investigation of a river-tunnel effect on $PM_{2.5}$ concentrations in New York City subway stations. *Transp. Res. Part D Transp. Environ.* **115**, 103579 (2023).

47. He, S. et al. Characteristics of subway air quality and favorable locations for passengers waiting and taking the subway. *Atmos. Pollut. Res.* **13**, 101482 (2022).

48. Hwang, S. H., Yoon, C. S., Ryu, K. N., Paik, S. Y. & Cho, J. H. Assessment of airborne environmental bacteria and related factors in 25 underground railway stations in Seoul, Korea. *Atmos. Environ.* **44**, 1658–1662 (2010).

49. Minguillon, M. C. et al. Aerosol sources in subway environments. *Environ. Res.* **167**, 314–328 (2018).

50. Reche, C. et al. Factors controlling particle number concentration and size at metro stations. *Atmos. Environ.* **156**, 169–181 (2017).

51. Salma, I., Weidinger, T. & Maenhaut, W. Time-resolved mass concentration, composition and sources of aerosol particles in a metropolitan underground railway station. *Atmos. Environ.* **41**, 8391–8405 (2007).

52. Kan, Z. et al. Assessing individual activity-related exposures to traffic congestion using GPS trajectory data. *J. Transp. Geogr.* **98**, 103240 (2022).

53. Cartenì, A., Cascetta, F. & Campana, S. Underground and ground-level particulate matter concentrations in an Italian metro system. *Atmos. Environ.* **101**, 328–337 (2015).

54. Van Ryswyk, K. et al. Metro commuter exposures to particulate air pollution and $PM_{2.5}$-associated elements in three Canadian cities: the urban transportation exposure study. *Environ. Sci. Technol.* **51**, 5713–5720 (2017).

55. Wang, B.-Q., Liu, J.-F., Ren, Z.-H. & Chen, R.-H. Concentrations, properties, and health risk of $PM_{2.5}$ in the Tianjin City subway system. *Environ. Sci. Pollut. Res.* **23**, 22647–22657 (2016).

56. Lundberg, S. M. & Lee, S.-I. A unified approach to interpreting model predictions. In *Proc. Advances in Neural Information Processing Systems* Vol. 30 (2017).

57. Duan, X., Zhao, X., Wang, B., Chen, Y. & Cao, S. *Highlights of the Chinese Exposure Factors Handbook* (Science Press, Beijing 2015).

Author contributions
T.Q.: data collection, methodology, illustration, analysis, paper drafting; R.T.: supervision, methodology, conceptualization, analysis, paper drafting and revision; A.W.: methodology, paper revision; J.W.: illustration, paper drafting; T.L.: paper revision; S.X.: illustration, methodology; X.L.: illustration; Q.C.: illustration; J.K.: paper revision; C.Y.: resource, paper revision; L.H.: funding source, paper revision; J.F.: supervision; S.W: supervision, methodology, conceptualization, analysis, funding source, resource, paper drafting and revision. All authors have read and approved the manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s44333-025-00070-4.

Correspondence and requests for materials should be addressed to Ran Tu or Shunyao Wang.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.