Environmental Pollution 233 (2018) 494–500

Contents lists available at ScienceDirect

Environmental Pollution

journal homepage: www.elsevier.com/locate/envpol

Estimation of residential fine particulate matter infiltration in Shanghai, China☆

Xiaodan Zhou [a, b, 1], Jing Cai [a, c, 1], Yan zhao [d, 1], Renjie Chen [a], Cuicui Wang [a], Ang Zhao [a, e], Changyuan Yang [a], Huichu Li [a], Suixin Liu [f], Junji Cao [f], Haidong Kan [a, *], Huihui Xu [b, **]

[a] School of Public Health, Key Laboratory of Public Health Safety of the Ministry of Education, Key Laboratory of Health Technology Assessment of the Ministry of Health, Fudan University, Shanghai, China
[b] Environmental Health Department, Shanghai Municipal Center for Disease Control and Prevention, Shanghai, China
[c] Shanghai Key Laboratory of Meteorology and Health, Shanghai, China
[d] Shanghai First Maternity and Infant Hospital, Tongji University School of Medicine, Shanghai, China
[e] Environmental & Occupational Health Evaluation Department, Shanghai Municipal Center for Disease Control & Prevention, Shanghai, China
[f] Institute of Earth Environment, Chinese Academy of Sciences, Xian, China

A R T I C L E I N F O

Article history:
Received 2 May 2017
Received in revised form
11 October 2017
Accepted 13 October 2017

Keywords:
$PM_{2.5}$ exposure
Infiltration factor
Model prediction
Seasonal variation

A B S T R A C T

Ambient concentrations of fine particulate matter ($PM_{2.5}$) concentration is often used as an exposure surrogate to estimate $PM_{2.5}$ health effects in epidemiological studies. Ignoring the potential variations in the amount of outdoor $PM_{2.5}$ infiltrating into indoor environments will cause exposure misclassification, especially when people spend most of their time indoors. As it is not feasible to measure the $PM_{2.5}$ infiltration factor (F_{inf}) for each individual residence, we aimed to build models for residential $PM_{2.5}$ F_{inf} prediction and to evaluate seasonal F_{inf} variations among residences. We repeated collected paired indoor and outdoor $PM_{2.5}$ filter samples for 7 continuous days in each of the three seasons (hot, cold and transitional seasons) from 48 typical homes of Shanghai, China. $PM_{2.5}$-bound sulfur on the filters was measured by X-ray fluorescence for $PM_{2.5}$ F_{inf} calculation. We then used stepwise-multiple linear regression to construct season-specific models with climatic variables and questionnaire-based predictors. All models were evaluated by the coefficient of determination (R^2) and root mean square error (RMSE) from a leave-one-out-cross-validation (LOOCV). The 7-day mean (±SD) of $PM_{2.5}$ F_{inf} across all observations was 0.83 (±0.18). F_{inf} was found higher and more varied in transitional season (12–25 °C) than hot (>25 °C) and cold (<12 °C) seasons. Air conditioning use and meteorological factors were the most important predictors during hot and cold seasons; Floor of residence and building age were the best transitional season predictors. The models predicted 60.0%–68.4% of the variance in 7-day averages of F_{inf}. The LOOCV analysis showed an R^2 of 0.52 and an RMSE of 0.11. Our finding of large variation in residential $PM_{2.5}$ F_{inf} between seasons and across residences within season indicated the important source of outdoor-generated $PM_{2.5}$ exposure heterogeneity in epidemiologic studies. Our models based on readily available data may potentially improve the accuracy of estimates of the health effects of $PM_{2.5}$ exposure.

1. Introduction

Epidemiologic studies have consistently suggested fine particulate matter ($PM_{2.5}$) as a risk factor for adverse health effects (Pope and Dockery, 2006). However, interpretation of findings from these studies have been hampered by uncertainties in exposures, because outdoor concentrations were universally used as an exposure proxy, even though most individuals spend more than 80% of their time indoors (Leech et al., 1996; Klepeis et al., 2001; EPA, 2013). Although it has been found indoor $PM_{2.5}$ commonly correlated with

☆ This paper has been recommended for acceptance by Dr. Hageman Kimberly Jill.

* Corresponding author. Mailbox 249, 130 Dong-An Road, Shanghai 200032, China.
** Corresponding author. Room 1527, 1380 West Zhongshan Road, Shanghai 200032, China.
E-mail addresses: haidongkan@gmail.com (H. Kan), xuhuihui@scdc.sh.cn (H. Xu).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.envpol.2017.10.054
0269-7491/

ambient concentrations as outdoor $PM_{2.5}$ can enter the indoor spaces (Chen and Zhao, 2011), spatial and temporal variations of $PM_{2.5}$ outdoor-to-indoor transport haven't been fully understood, which is needed for exposure assessment methods improvement.

$PM_{2.5}$ infiltration factor (F_{inf}), defined as the equilibrium proportion of outdoor fine particles that penetrates indoors and remains suspended (Chen and Zhao, 2011), was useful for quantifying the fraction of the total indoor particles with outdoor origin. Studies conducted in North America and Europe found substantially spatial and temporal variation of $PM_{2.5}$ F_{inf} (Chen and Zhao, 2011), which indicates that ignoring potential variations in the outdoor-indoor $PM_{2.5}$ infiltration would result in exposure misclassification (Allen et al., 2007; Meng et al., 2005; Long et al., 2001) that could further bias health effect estimates.

Particle-bound sulfate or sulfur has been commonly used to estimate $PM_{2.5}$ F_{inf} for residential homes (Wallace and Williams, 2005; Dockery and Spengler, 1967; Sarnat et al., 2002), because it is abundant in ambient particles, especially in the submicron particle size range (Hänninen et al., 2004) and with few indoor sources (Sarnat et al., 2002). This method requires both indoor and outdoor pollution measurements. However, it is extremely challenging to measure $PM_{2.5}$ F_{inf} for all individual residences in large population studies and establishing F_{inf} prediction models with available data on housing, environment and activities factors (Clark et al., 2010) could be a feasible solution.

Previous studies showed that $PM_{2.5}$ F_{inf} were differently influenced by residential factors between regions. Hystad et al (2009) used temperature, building value, and heating approaches to predict 54% of infiltration among detached residences from the U.S and Canada (Hystad et al., 2009). Chan et al (2005) found year of construction, size of dwelling and category of dwelling energy efficiency were important predictors of PM infiltration across the U.S. (Chan et al., 2005).

According to these previous studies, predictors of F_{inf} varied with regions and climates and the accessibility of some variables may differ from regions as well. F_{inf} models therefore may not be transferable to other locations. In China, only a limited number of studies have evaluated the variation of residential $PM_{2.5}$ F_{inf} between residences and within residence across seasons (Shi et al., 2015). In addition, even fewer studies explained the variation of $PM_{2.5}$ F_{inf} through modeling methods. In this study, we aim to estimate the infiltration of $PM_{2.5}$ in typical homes of Shanghai, China and to investigate key factors of residential $PM_{2.5}$ F_{inf} by establishing prediction models.

2. Methods

2.1. Study design

In this study, residences from the downtown area of Shanghai were recruited through flyers. To rule out the possibility of significant indoor PM or sulfur sources, such as smoking, frying, grilling and candle burning (Gorjinezhad et al., 2017; Amouei Torkmahalleh et al., 2017), we excluded residences with the following residences: 1) residences with smoking family members; 2) those using coal or wood as cooking fuels; 3) those with open kitchens; 4) those having habits of candle burning.

Among residences that met our criteria, apartment and Shikumen (a traditional Shanghainese architectural style characterized by brick-wood structure houses with shared stone gates and patios with lanes and alleys) were selected since they were typical building types in Shanghai and comprise more than 50% of the total housing stock of the city according to the Shanghai Yellow Pages.

2.2. Data collection

A total of 48 recruited residences were eventually monitored between June 2013 and January 2014. Indoor and outdoor sampling were conducted at participants' homes. For indoor sampling, equipment was set in the middle of the main activity room away from kitchens, air conditioners and ventilation. Outdoor sampling equipment was placed in the back yard, away from all structures; whereas for high-rise apartments the outdoor samplers were extended approximately 1-m out of an available window, with any cracks being sealed to prevent air exchange. At each residential site, measurements were conducted for three 7-day periods representing hot, cold and transitional season. All 48 homes had indoor and outdoor sampling equipment running simultaneously in both transitional and cold seasons. For the hot season, indoor $PM_{2.5}$ of 48 homes were monitored, however, only 19 homes had outdoor sampler due to the limited equipment availability.

We used samplers with a 2 L/min pump (PCXR8, SKC Inc., PA, USA) and a $PM_{2.5}$ impactor for indoor and outdoor sampling. To prevent overloading, effectively 72-h samples were collected on pre-weighed 37 mm Teflon filters (225–8303, SKC Inc., PA, USA) using a programmed schedule for each sampling event. All filters were pre-conditioned for 48 h prior to weighing at a constant air temperature of 20 °C ± 1 °C and constant relative humidity (RH) of 50% ± 5%. Field blanks comprised 10% of the total number of collected samples, and blank-corrected $PM_{2.5}$ mass concentrations were determined following gravimetric analysis. PM-bound sulfur in the filters was analyzed using energy dispersive X-ray fluorescence (Cooper Environmental Services, Portland, OR, USA). Real-time indoor and outdoor temperatures and relative humidity during each sampling period were recorded using data loggers (HOBO U10-003).

Information of resident behaviors related to F_{inf} and residence characteristics were gathered through a main questionnaire at recruitment, including building type and year constructed, family members, presence of air conditioning (AC) and heating facilities, presence of air filters/cleaners and sources of indoor particles (cooking fuel type and habits). For behaviors that vary seasonally or typical activities that may occur occasionally, participants were asked to record them with detailed information in a structured questionnaire during each sampling period, including activities related to ventilation, use of AC/heat, time and frequency of cooking and cleaning, and guest smoking.

2.3. F_{inf} calculation

F_{inf} of $PM_{2.5}$ was calculated based on sulfur infiltration factor. First, we calculated the sulfur infiltration factor using Eq. (1) for each residence based on the assumption that there are typically no indoor sources of sulfur.

$$F_{inf}{}^{S}{}_{i} = C^{S}_{Ii}/C^{S}_{Oi} \tag{1}$$

where i means individual-specific; $F_{inf}{}^{S}{}_{i}$ is PM-sulfur infiltration factor for residence i; C^{S}_{Ii} and C^{S}_{Oi} are the indoor (I) and outdoor (O) concentrations of sulfur for residence i, respectively.

Previous studies have reported that $F_{inf}{}^{PM2.5}$ (F_{inf} of $PM_{2.5}$) may differ from that of PM-bound sulfur possibly due to the change of sulfur proportion on $PM_{2.5}$ during the infiltration (Hänninen et al., 2004). Thus, for each season, the observed difference in $PM_{2.5}$ and sulfur infiltration factors was then corrected using the ratio of the corresponding regression coefficients according to Eq. (2).

$$F_{inf}{}^{PM2.5}{}_{i} = (\beta^{PM2.5}/\beta^{S})_{s} \times F_{inf}{}^{S}{}_{i} \text{ (Hänninen et al., 2004)} \tag{2}$$

where i and s means individual-specific and season-specific; $F_{inf}{}^{PM2.5}{}_i$ and $F_{inf}{}^{S}{}_i$ are $PM_{2.5}$ and PM-sulfur infiltration factor for residence i, respectively; $\beta^{PM2.5}$ and β^{S}, respectively, are the slopes of $PM_{2.5}$ and sulfur indoor–outdoor regression for each season (Table 3).

For data quality control, we excluded sampling data if (1) the sampling pump achieved less than 75% of the programmed schedule (i.e. effective sampling time less than 54 h); (2) the flow rate failed to maintain at 2 L/min (±0.2 L/min) from the start to the end of sampling. Eventually, 31 out of 144 total monitoring events (21.5%) were removed from analysis, leaving 113 monitoring events suitable for calculating F_{inf}.

2.4. Model building

Stepwise-multiple linear regression was applied to construct $F_{inf}{}^{PM2.5}$ models. We included temperature and relative humility, residence type, building age of residence, floor, daily cooking frequency, cooking fuel type, use of kitchen ventilators when cooking, area of the monitored room, window opening behavior, AC/heat use behavior as potential predictors because some of these variables were identified to relate with F_{inf} in previous studies (Chen and Zhao, 2011; Clark et al., 2010; Allen et al., 2012) and most of them were relatively easy to collected.

A three-stage approach was used for model establishment. First step was evaluating univariate regressions of the $F_{inf}{}^{PM2.5}$ with all available potential predictors. The predictor giving the highest adjusted R^2 was selected as the starting model for inclusion if the direction of effect was as defined a priori [e.g. windows opening would increase the $F_{inf}{}^{PM2.5}$ (positive direction) because of the increased ventilation]. Second, a manually supervised forward regression analysis was conducted to evaluate which of the remaining variables further improved the model adjusted R^2. Subsequent variables stayed in the model if they met the following criteria: 1) the adjusted R^2 of the model increased by at least 0.01 (1%) to that of the previous model; 2) the coefficient of the variable and the existing variables in the model accorded with the right direction of effect. These criteria ensured that models involving counterintuitive associations be avoided, even if they give a stronger basis for prediction (as indicated by adjusted R^2 value and RMSE). This process continued until no more variables met the criteria. As a final step, we added indicator variables of "day of the week" and "day" of sampling in the models to control any unknown temporal trends in $PM_{2.5}$ concentrations.

Season-specific models were constructed under the assumption that the $PM_{2.5}$ F_{inf} predictors and their model coefficients may vary with seasons. We categorized each 7-day period into a "hot", "transitional" or "cold" season based on the average outdoor temperature (>25 °C, 12–25 °C and ≤12 °C, respectively). We used 25 °C and 12 °C as cutoffs because they were supported by the data (Fig. 1). $F_{inf}{}^{PM2.5}{}_i$ greater than 1.50 were removed because it may indicate a strong indoor sulfur source; $F_{inf}^{PM2.5}{}_i$ between 1.0 and 1.50 were included to account for imprecision in the sulfur measurements. For hot season, only 19 homes had indoor and outdoor sampling equipment running simultaneously so that only these data were used for model construction.

Leave-one-out-cross-validation (LOOCV) was applied for model performance evaluation. Each season-specific model was fitted to N-1 residences with the variables unchanged, and the predicted concentration was estimated using the fitted model at the left-out residence. The overall fit R^2 and root mean squared error (RMSE) between the predicted and estimated concentrations in three seasons and for all residences were calculated to represent the model performance.

The statistical tests were two-sided, and values of $P < 0.05$ were considered statistically significant. All data analysis were performed using R software (Version 2.15.3) with "stats" package.

3. Results

3.1. Residence characteristics

As summarized in Table 1, of the 48 monitored residences, 65% were apartments, 56% were on the 1st-3rd floors and 58% were built after 1990. Forty-four families cooked at least once per day using natural gas as the only cooking fuel and 41 of them used ventilators during cooking. All residences installed AC and 5 of them had air filters/cleaners, but none of them used air filters/cleaners during the sampling periods.

3.2. Seasonal variability of air pollution and residential activities

Table 2 presents description of residential activities related to F_{inf} and their seasonal variation. Median (5^{th}-95th percentile) daily window opening time was 11.68 (3.33, 21.67), 9.56 (1.33, 22.85), 8.89 (0.57, 24.00) hours in hot, transitional and cold season, respectively. AC use was much more prevalent in the hot season (6.50 (0, 46.24) hours) than the cold season (0 (0, 13.32) hours). Information from the questionnaire indicated none of the 48 homes used AC in transitional seasons.

Seasonal variations of outdoor $PM_{2.5}$ showed as expected that cold season recorded the highest median concentration at 77.05 (42.37, 120.10) μg/m^3, whereas the lowest in hot season at 31.64 (17.35, 58.80) μg/m^3. Overall, both indoor levels of $PM_{2.5}$ were lower than but comparable to their corresponding outdoor levels. In addition, the indoor and outdoor $PM_{2.5}$ concentrations were significantly correlated and the correlation coefficient was 0.66 ($p < 0.01$) across all the observations (Table 3). Similar and stronger associations between indoor and outdoor levels were observed for $PM_{2.5}$-bound sulfur. Their corresponding values were showed in Tables 2 and 3.

3.3. Infiltration factors distribution

Table 3 provided slopes produced by indoor versus outdoor regression analyses. Consistent with previous studies (Hänninen et al., 2004), β^{S} is slightly larger than $\beta^{PM2.5}$ among different seasons, which varied between 0.84 (cold season) and 0.94 (transition season) for sulfur and between 0.62 (hot season) and 0.77 (cold season) for $PM_{2.5}$. Accordingly, the ratio of the corresponding regression coefficients differed from seasons and they were 0.67, 0.81 and 0.92 in hot, transitional and cold seasons, respectively. Then slope ratios ($\beta^{PM2.5}/\beta^{S}$) were used to correct the difference between $PM_{2.5}$ and sulfur infiltration factors.

We also found the small and statistically nonsignificant intercepts of the indoor–outdoor PM-sulfur regressions. The intercepts were −0.03 μg/m^3 (Standard error (SE) = 0.44, p = 0.95), 1.24 μg/m^3 (SE = 0.85, p = 0.16) and 0.51 μg/m^3 (SE = 0.29, p = 0.09) for hot, transitional and cold season, respectively. The finding supports our assumption of the absence of indoor PM-sulfur sources.

Table 4 compared the $PM_{2.5}$ infiltration factors under different factors, including seasons, AC use, house types and daily kitchen use frequency. The 7-day mean (±SD) of $PM_{2.5}$ F_{inf} across all observations was 0.83 ± 0.18. Transitional season had larger and more varied F_{inf} (0.92 ± 0.23) compared to the other two seasons ($p < 0.05$), while no significant difference was observed between hot and cold season.

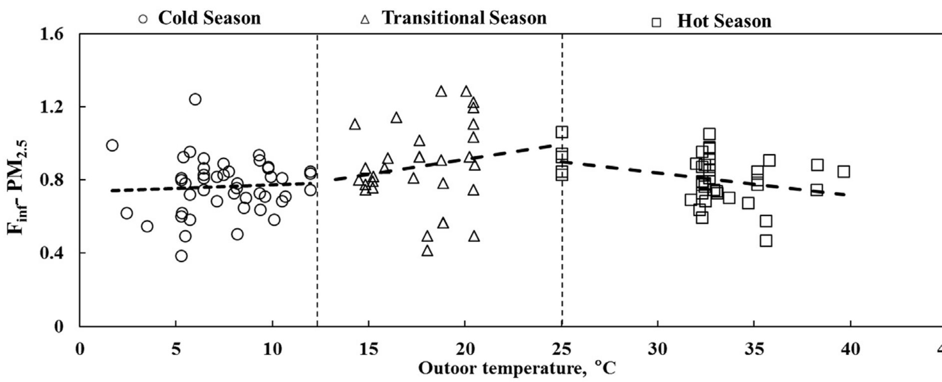

Fig. 1. Infiltration factor of $PM_{2.5}$ vs. average outdoor temperature during the 1-week sampling period in cold, transitional and hot seasons.

Table 1
Description of the monitored residences.

	Residences
Total monitored, *n*	48
Residence Types, *n*	
Apartment	31
Shikumen	17
Built Year, *n*	
<1950	14
1950–1989	6
1990–2000	17
>2000	11
Floor, *the n th*	
1^{st}-3rd	27
4^{th}-10th	11
>10th	10
Daily cooking frequency, *n*	
0	4
1	14
2	15
3	15
Cooking fuel type, *n*	
No cooking or cooking outside of residence	4
Natural gas	44
Use of kitchen ventilators when cooking, *n*	41
Area of the monitored room, Mean±SD[a], m^2	18.58 ± 7.46

[a] SD = Standard Deviation.

3.4. Infiltration factors model

We found that income, floors, daily kitchen use frequency, building age, building type, window open time, use time of AC, outdoor temperatures and humidity were associated with F_{inf} ($p < 0.05$) in the univariate regression.

Table 5 showed results of predictor selection for season-specific models. The model R^2 were 0.68, 0.68 and 0.60 for hot, transitional and cold season, respectively. The most consistent F_{inf} predictors during the hot season and cold season were variables related to use of AC and climate. Less total time of AC use ($p < 0.01$) were associated with higher $PM_{2.5}$ F_{inf} in both hot and cold seasons; while outdoor temperature ($p = 0.015$) only influence F_{inf} in cold season. Outdoor RH appeared to have different direction of effect in the two seasons; however it has a marginal significance ($p = 0.098$) in the cold-season model. For transitional season, the most important predictors were floor of residence ($p < 0.01$) and building age ($p < 0.01$). Residences on the higher floors ($p < 0.01$) and in the older buildings ($p < 0.01$) had smaller F_{inf}.

The comparisons between estimated and predicted F_{inf} was shown on Fig. 2. The overall model CV R^2 and RMSE were 0.515 and 0.11, respectively. We found all data points distributed evenly at two sides of 1:1 line further indicates reasonable agreement between estimations and prediction. For each season, model fit R^2 and RMSE was 0.34 and 0.10 for hot season, 0.48 and 0.15 for transitional season, and 0.38 and 0.09 for cold season (Table 5). Much lower R^2 observed in hot and cold seasons is probably due to the small variation of F_{inf}-$PM_{2.5}$ within season. As shown in Table 4, the standard deviation of F_{inf}-$PM_{2.5}$ was 0.12 and 0.14 in these two seasons compared to 0.23 for transitional season.

4. Discussion

In this study, we measured indoor and outdoor $PM_{2.5}$ concentrations in 48 typical households in Shanghai, China for three

Table 2
Distribution (P5, P50, P95)[a] of residential indoor and outdoor levels of $PM_{2.5}$ and sulfur, I/O ratio, climate and activities by seasons.

	Hot season	Transitional season	Cold season	Overall
Monitored residences, *n*	42	31	47	48
Outdoor temperature (°C)	(32.11, 32.68, 38.26)	(14.45, 18.75, 24.69)	(3.92, 7.29, 10.63)	(5.25, 17.62, 35.18)
Outdoor relative humidity (%)	(48.80, 61.19, 64.67)	(53.00, 64.83, 85.74)	(36.89, 54.06, 63.78)	(37.84, 60.71, 78.85)
Window opening time per day (h/d)	(3.33, 11.68, 21.67)	(1.33, 9.56, 22.85)	(0.57, 8.89, 24.00)	(0.98, 10.63, 24.00)
Total AC use time during sampling period (h)	(0, 6.50, 46.24)	0	(0, 0, 13.32)	(0, 0, 37.76)
Outdoor $PM_{2.5}$ ($\mu g/m^3$)	(17.35, 31.64, 58.80)	(31.10, 39.79, 113.82)	(42.37, 77.05, 120.10)	(18.86, 51.12, 118.23)
Indoor $PM_{2.5}$ ($\mu g/m^3$)	(11.10, 32.30, 60.59)	(31.90, 46.08, 97.27)	(24.81, 69.88, 102.62)	(13.93, 47.54, 101.16)
I/O ratio of $PM_{2.5}$	(0.24, 0.88, 1.27)	(0.64, 1.03, 1.37)	(0.46, 0.83, 1.09)	(0.44, 0.87, 1.32)
Outdoor sulfur ($\mu g/m^3$)	(2.53, 4.46, 9.55)	(2.88, 6.06, 12.11)	(2.43, 4.04, 9.30)	(2.53, 4.52, 10.57)
Indoor sulfur ($\mu g/m^3$)	(2.15, 4.02, 8.73)	(3.39, 5.80, 9.58)	(2.02, 3.98, 7.54)	(2.15, 4.32, 8.92)
I/O ratio of sulfur	(0.55, 0.89, 1.14)	(0.65, 0.89, 1.09)	(0.55, 0.90, 1.21)	(0.57, 0.87, 1.06)

Definition of abbreviations: $PM_{2.5}$-particulate matter less than 2.5 μm in aerodynamic diameter; I/O ratio –the ratio of indoor concentration versus outdoor concentrations.
[a] P5, the 5th centile of the distribution, P(50): the 50^{th} centile of the distribution; P(95): the 95th centile of the distribution.

Table 3
Regression analysis of the sulfur and $PM_{2.5}$ indoor–outdoor relationships.

	sulfur					$PM_{2.5}$					Slope ratio
	β_S	SE	R^2	N	p-value	$\beta_{PM2.5}$	SE	R^2	N	p-value	$\beta_{PM2.5}/\beta_S$
Hot season[a]	0.92	0.02	0.99	19	<0.01	0.62	0.13	0.50	19	<0.01	0.67
Transitional season[b]	0.94	0.04	0.96	29	<0.01	0.76	0.07	0.79	31	<0.01	0.81
Cold season[c]	0.84	0.02	0.98	44	<0.01	0.77	0.08	0.66	45	<0.01	0.92
Overall	0.87	0.02	0.96	92	<0.01	0.73	0.05	0.66	95	<0.01	0.84

Definition of abbreviations: $PM_{2.5}$-particulate matter less than 2.5 μm in aerodynamic diameter.
[a] Hot season: the average outdoor temperature is higher than 25 °C.
[b] Transitional season: the average outdoor temperature is between 12 and 25 °C.
[c] Cold season: the average outdoor temperature is lower than 12 °C.

Table 4
The distribution of F_{inf} of $PM_{2.5}$ categorized by different factors.

F_{inf}	Mean ± SD	Min	P25[a]	P75[b]	Max
Observations	0.83 ± 0.18	0.50	0.73	0.91	1.50
Season*					
Hot season	0.79 ± 0.12	0.58	0.70	0.88	1.06
Transitional season	0.92 ± 0.23	0.50	0.78	1.04	1.50
Cold season	0.79 ± 0.14	0.55	0.71	0.86	1.24
AC use					
Yes	0.80 ± 0.23	0.55	0.63	0.91	1.50
No	0.84 ± 0.17	0.50	0.74	0.91	1.29
Daily kitchen use frequency					
0	0.76 ± 0.14	0.59	0.61	0.86	0.95
1	0.83 ± 0.22	0.50	0.70	0.92	1.50
2	0.86 ± 0.20	0.58	0.70	0.93	1.29
3	0.83 ± 0.13	0.57	0.74	0.91	1.24
House type					
Apartment	0.83 ± 0.18	0.50	0.71	0.92	1.29
Shikumen	0.83 ± 0.19	0.58	0.75	0.90	1.50

*: Hot season: the average outdoor temperature is higher than 25 °C; transitional season: the average outdoor temperature is between 12 °C and 25 °C and cold season: the average outdoor temperature is lower than 12 °C.
[a] P25, the 25th centile of the distribution.
[b] P75: the 75th centile of the distribution.

Table 5
Final stepwise regression for season-specific F_{inf}.

Predictor	β	SE	*p-value*	Model		LOOCV	
				R^2	adjusted R^2	R^2	RMSE
Hot season[a] (n = 19)				0.684	0.594	0.34	0.10
Intercept	0.182	0.317	0.58				
Time of AC use per day (h)	−0.013	0.005	<0.01				
Relative humidity (%)	0.010	0.004	0.03				
Transitional season[b] (n = 29)				0.681	0.563	0.48	0.15
Intercept	2.373	0.752	<0.01				
Floor	−0.020	0.004	<0.01				
Age (year)	−0.005	0.002	<0.01				
Cold season[c] (n = 44)				0.600	0.520	0.38	0.09
Intercept	0.613	0.130	<0.01				
Time of AC use per day (h)	−0.010	0.003	<0.01				
Temperature (°C)	0.020	0.008	0.015				
Relative humidity (%)	−0.002	0.001	0.098				

Definition of abbreviations: AC, air conditioning.
[a] Hot season: the average outdoor temperature is higher than 25 °C.
[b] Transitional season: the average outdoor temperature is between 12 and 25 °C.
[c] Cold season: the average outdoor temperature is lower than 12 °C.

different seasons. Our results showed two-to three-fold difference of F_{inf} across households, ranging from 0.50 to 1.50. This result was consistent with two large-scale studies in multiple communities, the Multi-Ethnic Study of Atherosclerosis and Air Pollution (MESA Air) (Allen et al., 2012) and the Relationships of Indoor, Outdoor, and Personal Air (RIOPA) study (Meng et al., 1994) found similar ranges of F_{inf} across western residences, which are 0.2–1.0 and 0.1–1.3, respectively. Considerable spatial variation of F_{inf} between households supports that ignoring potential variations in outdoor-to-indoor $PM_{2.5}$ may increase exposure classification.

Temporal variability of F_{inf} found in this study was also significant. $PM_{2.5}$ F_{inf} reached its maximum in the transitional season,

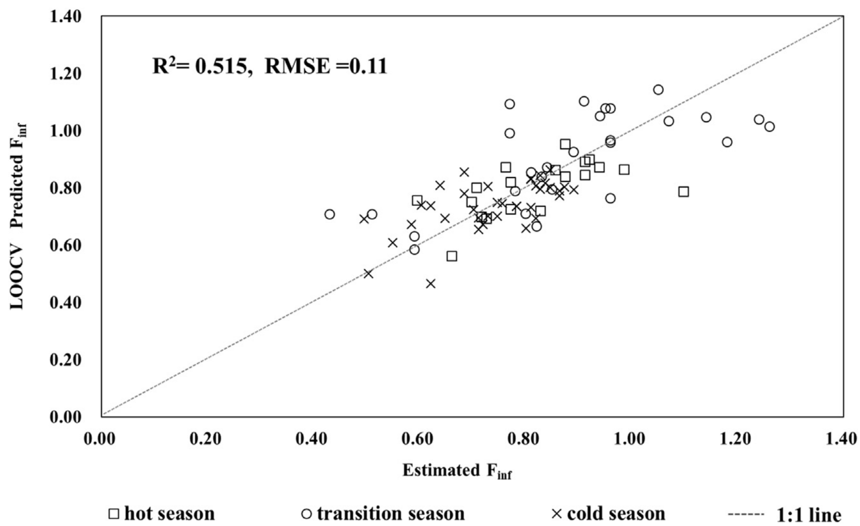

Fig. 2. Comparisons of estimated F_{inf} (x-axis) with values predicted from a LOOCV (y-axis) fit for the multiple regression models shown in Table 5. Dash Line represent 1:1.

which was consistent with the previous studies (Allen et al., 2012; Meng et al., 1994). It is plausible that temperature may influence PM infiltration by affecting residential behaviors. For example, Shanghai experiences relatively hot summers with average daily highs of 32.2 °C in July and cold winters with average daily lows of 2.1 °C in January (data retrieved from Dataset of daily climate data from Chinese surface stations for global exchange, 2013–2014, China Meteorological Administration). Therefore, both hot and cold weather necessitate that residences are sealed from outdoor air, and typical homes in Shanghai use mechanical heating or cooling for inhabitant comfort. Whereas in transitional season, less use of AC and more window opening, could significantly enhance air exchange and thereby increased F_{inf} (Clark et al., 2010; Howard-Reed et al., 2002; Wallace et al., 2002).

Our model predictors provided further explanation for the observed seasonal variation of F_{inf}. In the hot season and cold season, significant variation in F_{inf} was explained by meteorological conditions and inhabitants' activities. The most consistent predictor is AC use, which has been reported by studies conducted in Toronto, Canada (Clark et al., 2010) and six U.S. communities (Allen et al., 2012). These studies also found lower F_{inf} in homes with more often central AC use, which may influence F_{inf} by discouraging window opening and/or by increasing PM deposition on filters or in air ducts (Howard-Reed et al., 2003). The other important predictors were humidity for hot season and temperature for cold season. These variables probably contributed additional information on AC use beyond those captured by questionnaire.

Built year and floor predicted a large portion of variation of transition-season F_{inf}. Residence age has previously been reported to affect infiltration, although inconsistent across different studies. This is not unexpected as tightness of the residences mainly depend on the building methods differing across regions and over time. Our multivariate regression model found a significant association between higher F_{inf} with older residences and lower floors. A trend of lower infiltration in older residence was also observed by Allen et al. (2012), while Hystad et al. (2009) found the opposite results. No effect of residence age was observed by Meng et al. (1994).

This study has several strengths. First, we used sulfur as an outdoor $PM_{2.5}$ tracer to estimate F_{inf}. We carefully corrected the bias from the change of sulfur proportion on $PM_{2.5}$ during infiltration by season. Second, the repeated measurements in three seasons allows us to detect the temporal variation of $PM_{2.5}$ F_{inf} as well as to explore the influence of occupants' behavior between seasons. Third, our models established on readily accessible data explained a major portion of the variance in 1-week average F_{inf}. In addition, LOOCV analysis implied the stable performance of prediction models.

Our study has limitations. First, 32%–40% of the variability remains unexplained by our models and there are some other influential factors needed to be explored in the future. Second, the sample size in this study was relatively small, and we only investigated two typical architecture types within a small geographic area. One should be cautious when making extrapolations of these results to other residences or other microenvironments (e.g., office). In addition, the applicability of our model results in other parts of the world should be carefully investigated, especially for those areas with different climate and residence characteristics.

This work may have an applicability in air pollution epidemiologic studies of China or even other developing country. China, as the largest developing country, has suffered the most severe problem of $PM_{2.5}$ pollution. Given the importance of accounting for the variation of ambient PM infiltration and the challenge of F_{inf} -$PM_{2.5}$ measurement in individual home, our predictive models provide a feasible solution for large-scale population studies. The models built in this study based upon accessible variables potentially allows for more accurate and precise estimates of the health risks of $PM_{2.5}$ exposure in urban settings of China. Furthermore, the methodology of F_{inf} $PM_{2.5}$ modeling developed in this study could be transferable to other parts of the world. Specifically, careful correction of the difference between $PM_{2.5}$ F_{inf} and PM-sulfur F_{inf}, and selection of candidate predictors were two critical procedures of F_{inf} prediction. Lastly, this work may have implication on the development of exposure mitigation strategies. The factors identified in our predictive models (such as AC use) influenced the variability of F_{inf} PM and may serve as an intervention target for air pollution health challenges in China.

5. Conclusion

In this study, the finding of wide spatial and temporal variations in residential $PM_{2.5}$ F_{inf} highlighted the heterogeneity in exposure to $PM_{2.5}$ of outdoor origin. Ignoring potential variations in outdoor-to-indoor $PM_{2.5}$ may increase exposure classification. Based on relatively easily collected predictors, our predictive models of $PM_{2.5}$

F_{inf} could explain major portion of this variation, suggesting modeling approach can be a feasible solution for $PM_{2.5}$ F_{inf} estimation.

Acknowledgements

We thank the participants of this study. The work was supported by the Public Welfare Research Program of National Health and Family Planning Commission of China (201502003), National Natural Science Foundation of China (91543114 and 81222036), Key Discipline of Public Health in Shanghai (15GWZX0201) and Shanghai Health and Family Planning Commission (20124377).

The authors declare they have no competing financial interests.

References

Allen, R., et al., 2007. Evaluation of the recursive model approach for estimating particulate matter infiltration efficiencies using continuous light scattering data. J. Expo. Sci. Environ. Epidemiol. 17 (5), 468–477.
Allen, R.W.A., Avol, S.D., Cohen M, E., et al., 2012. Modeling the residential infiltration of outdoor PM($_{2.5}$) in the multi-ethnic study of Atherosclerosis and air pollution (MESA air). Environ. Health Perspect. 120 (6), 824–830.
Amouei Torkmahalleh, M., et al., 2017. Review of factors impacting emission/concentration of cooking generated particulate matter. Sci. Total Environ. 586, 1046–1056.
Chan, W.Y.R., et al., 2005. Analyzing a database of residential air leakage in the United States. Atmos. Environ. 39 (19), 3445–3455.
Chen, C., Zhao, B., 2011. Review of relationship between indoor and outdoor particles: I/O ratio, infiltration factor and penetration factor. Atmos. Environ. 45 (2), 275–288.
Clark, N.A., et al., 2010. Exploring variation and predictors of residential fine particulate matter infiltration. Int. J. Environ. Res. Public Health 7 (8), 3211–3224.
Dockery, D.W., Spengler, J.D., 1967. Indoor-outdoor relationships of respirable sulfates and particles. Atmos. Environ. 15 (3), 335–343, 1981.
EPA, C., 2013. Exposure Factors Handbook of Chinese Population (Adults). China Environmental Science Press, p. 849.
Gorjinezhad, S., et al., 2017. Quantifying trace elements in the emitted particulate matter during cooking and health risk assessment. Environ. Sci. Pollut. Res. Int. 24 (10), 9515–9529.
Hänninen, O.O., et al., 2004. Infiltration of ambient PM2.5 and levels of indoor generated non-ETS PM2.5 in residences of four European cities. Atmos. Environ. 38 (37), 6411–6423.
Howard-Reed, C., Wallace, L.A., Ott, W.R., 2002. The effect of opening windows on air change rates in two homes. J. Air & Waste Manag. Assoc. 52 (2), 147–159.
Howard-Reed, C., Wallace, L.A., Emmerich, S.J., 2003. Effect of ventilation systems and air filters on decay rates of particles produced by indoor sources in an occupied townhouse. Atmos. Environ. 37 (38), 5295–5306.
Hystad, P.U., et al., 2009. Modeling residential fine particulate matter infiltration for exposure assessment. J. Expo. Sci. Environ. Epidemiol. 19 (6), 570–579.
Klepeis, N.E., et al., 2001. The National Human Activity Pattern Survey (NHAPS): a resource for assessing exposure to environmental pollutants. J. Expo. Anal. Environ. Epidemiol. 11 (3), 231–252.
Leech, J.A., et al., 1996. The canadian human activity pattern survey: report of methods and population surveyed. Chronic Dis. Can. 17 (3–4), 118–123.
Long, C.M., et al., 2001. Using time- and size-resolved particulate data to quantify indoor penetration and deposition behavior. Environ. Sci. Technol. 35 (10), 2089–2099.
Meng, Q.Y., et al., 1994. Determinants of indoor and personal exposure to PM(2.5) of indoor and outdoor origin during the RIOPA study. Atmos. Environ. 43 (36), 5750–5758, 2009.
Meng, Q.Y., et al., 2005. *$PM_{2.5}$ of ambient origin: estimates and exposure errors relevant to PM epidemiology*. Environ. Sci. Technol. 39 (14), 5105–5112.
Pope 3rd, C.A., Dockery, D.W., 2006. Health effects of fine particulate air pollution: lines that connect. J. Air Waste Manag. Assoc. 56 (6), 709–742.
Sarnat, J.A., et al., 2002. Using sulfur as a tracer of outdoor fine particulate matter. Environ. Sci. Technol. 36 (24), 5305–5314.
Shi, S.S., Chen, C., Zhao, B., 2015. Air infiltration rate distributions of residences in Beijing. Build. Environ. 92, 528–537.
Wallace, L., Williams, R., 2005. Use of personal-indoor-outdoor sulfur concentrations to estimate the infiltration factor and outdoor exposure factor for individual homes and persons. Environ. Sci. Technol. 39 (6), 1707–1714.
Wallace, L.A., Emmerich, S.J., Howard-Reed, C., 2002. Continuous measurements of air change rates in an occupied house for 1 year: the effect of temperature, wind, fans, and windows. J. Expo. Analysis Environ. Epidemiol. 12 (4), 296–306.