7th International Building Physics Conference

IBPC2018

Proceedings

SYRACUSE, NY, USA

September 23 - 26, 2018

Healthy, Intelligent and Resilient Buildings and Urban Environments

ibpc2018.org | #ibpc2018

Study on multivariate regression model of indoor and outdoor particulate pollution in severe cold area of China

Yang Lv[1,*], Haifeng Wang[1], Yuwei Zhou[1], Tongke Zhao[1], Bin Chen[1], Lei Zhang[2], Qi Zhao[2]

[1]Dalian University of Technology, Dalian
[2]Northeast Petroleum University, Daqing

*Corresponding email: lvyang@dlut.edu.cn

ABSTRACT
At present, the widespread existence of haze phenomenon has a serious impact on indoor air quality. Indoor particulate pollution has been paid more and more attention by the society. However, the correlation and diffusion mechanism of indoor and outdoor particulate matter are still controversial. In order to explore the correlation between indoor and outdoor particulate matter of different building types in heating season and non-heating season, the indoor and outdoor particulate concentrations and meteorological parameters of 110 stations in severe cold area of China were monitored by experiments. The analysis shows that indoor and outdoor temperature, humidity, air velocity, wind direction and atmospheric pressure are the main factors affecting indoor and outdoor particulate concentration. And based on these factors, it can model the indoor predicted particulate concentrations by multivariate regression. It also shows a significant difference in the relationship between the concentration of particulate matter and factors of indoor and outdoor particulate matter. Therefore, this study provides a good premise for exploring the health risks and control measures of particulate matter.

KEYWORDS
Particulate Matter, Severe Cold Area, Correlation Analysis, Regression Model

INTRODUCTION

Human inhalable particulate matter (PM_{10}, $PM_{2.5}$) is the primary pollutant in most cities in China, especially for fine particulate matter ($PM_{2.5}$). More and more epidemiologists show that there is a positive correlation between human morbidity, mortality and mass concentration of particulate matter (Dockery et al. 1993; Zhao et al. 2015; Tseng et al. 2015; Cohen et al. 2017; Klemm R J et al. 1996). People spend 90% of their time indoors, so indoor air quality plays an important role in human health. The study showed that the outdoor pollution components can enter into the indoor air through natural ventilation, mechanical ventilation and infiltration ventilation. There is a significant correlation between the indoor and outdoor particles. The proportion of indoor $PM_{2.5}$ from outdoor is 30~75% (Dockery and Spengler, 1981; Koutrakis et al., 1992; Ozkaynak et al. 1995; Xiong et al. 2004).

The situation of indoor and outdoor particulate matter in severe cold area of China is different from that in other regions because of geographical location, climatic condition and building type. This paper will via the long-term monitoring of indoor and outdoor particulate concentrations to analyze the correlation between indoor particles concentrations and indoor and outdoor influencing factors of different building types in Daqing, which will provide parameters and basis for indoor particulate matter exposure assessment.

METHODS

Selection of measured objects and sampling points

The sampling points were located in five districts of Daqing which were total of 110 sampling points, including 30 classrooms, 30 offices and 50 residences, among which were divided into urban and rural residences. To get the seasonal variations of particulate pollution, we collected data of summer and winter from November 2016 to April 2017 and June to August 2016, respectively. A sampling point was sited in and out of each room, measured simultaneously. Each point was collected seven days of valid data. To analyze more accurately, the hourly average value of each sampling point was calculated. The measuring instrument includes QT50 particulate online monitor ($\pm 1 \mu g/m^3$). The measurements contents include indoor and outdoor temperature, relative humidity and $PM_{2.5}$ mass concentration. The monitoring time was set to start every 15 minutes, and the data were collected for 5 minutes each time.

Data analysis and processing method

In this paper, the analysis data is mainly based on the hourly average of each parameter. With the SPSS software, the statistical analysis of sampling data was completed, and finished the multiple analysis such as bivariate correlation analysis and multiple linear regression.

Bivariate correlation analysis is an important method to evaluate the relationship between two of variables. Pearson coefficient (r) could measure the extent of correlations, which could be expressed as:

$$r = \frac{\sum (X - \bar{X})(Y - \bar{Y})}{\sqrt{(X - \bar{X})^2 (Y - \bar{Y})^2}} \quad (1)$$

Where X and Y - variables, $\bar{X}$ and $\bar{Y}$- the averages of variables. In this part, Pearson coefficient (r) is used to express the relationship between indoor particulate concentration and indoor and outdoor influencing factors (such as outdoor particulate concentration, temperature and humidity, etc.).

Multivariate linear regression is a regression with one dependent variable and two or more independent variables, each of which is a single term. It is one of the most commonly used statistical methods in the measurement of microenvironment or exposure of particulate matter. The mathematical model of multivariate linear regression can be expressed as follows:

$$y = a + b_1 x_1 + b_2 x_2 + \cdots + b_n x_n \quad (2)$$

Where, y is a dependent variable; x_1, x_2 ... x_n are independent variables; b_1, b_2 ... b_n are coefficients of independence variables; a is a constant term. To ensure the accuracy of the regression model, the errors between the observed and predicted values of five performance indexes are used in this paper: the normalized absolute error (NAE), the root mean square error (RMSE), the prediction accuracy (PA), the determinant coefficient (R^2) and the index of agreement (IA).

RESULTS AND DISCUSSIONS

Correlation Analysis between Indoor $PM_{2.5}$ concentration and factors

There are many factors affecting indoor particulate concentration, such as indoor particulate source strength, settling rate, air exchanges rate, penetration coefficient, building types and meteorological conditions. The results show that there is a certain relationship between indoor particulate concentration and them above, while little attention has been paid to this aspect in the present researches. By the method of bivariate analysis, Pearson correlation coefficient (*r*) was used to express the extent of the linear correlation between the indoor particulate concentration and these factors. The probability distribution value ($P < 0.05$) was considered to be statistically significant.

Table 1. The correlation analysis between indoor $PM_{2.5}$ concentration and affecting factors

Building type			T_{in}	RH_{in}	$PM_{2.5}$(Out)	T_{out}	RH_{out}	P	WD	WS
Office	Summer	*r*	0.027	0.289	0.734	0.030	0.126	-0.048	-0.183	0.092
		P	0.392	0.000	0.000	0.341	0.000	0.126	0.000	0.003
		N	1041	1041	1041	1041	1041	1041	1041	1032
	Winter	*r*	0.001	0.007	0.903	0.206	-0.063	-0.296	-0.036	0.198
		P	0.978	0.846	0.000	0.000	0.075	0.000	0.305	0.000
		N	806	806	806	806	806	806	806	806
Classroom	Summer	*r*	0.199	0.245	0.745	0.143	0.118	-0.083	-0.176	0.052
		P	0.000	0.000	0.000	0.000	0.000	0.008	0.000	0.038
		N	1032	1032	1032	1032	1032	1032	1032	1032
	Winter	*r*	-0.006	0.349	0.809	0.227	0.071	-0.076	-0.101	0.000
		P	0.862	0.000	0.000	0.000	0.034	0.024	0.003	0.990
		N	886	886	886	886	886	886	886	886
Urban residence	Summer	*r*	0.071	0.497	0.837	0.21	0.067	-0.235	0.038	0.112
		P	0.006	0.000	0.000	0.000	0.009	0.000	0.146	0.000
		N	1485	1485	1485	1485	1485	1485	1485	1485
	Winter	*r*	0.008	0.045	0.604	-0.111	0.144	0.027	-0.04	0.033
		P	0.797	0.152	0.000	0.000	0.000	0.392	0.195	0.291
		N	1028	1028	1028	1028	1028	1028	1028	1028
Rural residence	Summer	*r*	0.159	0.336	0.848	0.205	0.02	-0.217	0.129	0.13
		P	0.010	0.000	0.000	0.001	0.750	0.000	0.037	0.035
		N	263	263	263	263	263	263	263	263
	Winter	*r*	0.284	0.123	0.862	0.105	-0.034	-0.02	-0.089	0.092
		P	0.000	0.123	0.000	0.190	0.669	0.802	0.267	0.248
		N	158	158	158	158	158	158	158	158

Note: *r*-Correlation , *P*- Significance, *N*- Sampling times, T - Temperature, RH - Relative humidity, P - Atmospheric pressure, WD- Wind direction, WS- Wind speed

Table 1 compares the correlation analysis between indoor $PM_{2.5}$ concentration and indoor and outdoor factors of four building types in summer and winter, statistics of which have significant correlation has been bold. On the whole, it can be found that outdoor $PM_{2.5}$ concentration has the strongest correlation with indoor $PM_{2.5}$ whether it is summer or winter. Indoor $PM_{2.5}$ concentration is positively correlated with indoor and outdoor temperature, humidity and outdoor wind speed, negatively correlated with outdoor atmospheric pressure, and both positively and negatively correlated with outdoor wind direction. However, all these factors are significantly related to indoor particulate concentration for summer. For winter, the main factor of indoor $PM_{2.5}$ concentration is outdoor particulate concentration, the influence of other factors is relatively little. The primary explanation is that the main ventilation mode of the measured buildings is natural ventilation during summer, outdoor particulate matter can enter into the indoor environment which is vulnerable to wind pressure and hot pressure. Outdoor wind speed and atmospheric pressure will affect the indoor and outdoor air exchange rate and the concentration of indoor particulate matter. For the season of winter, the indoor temperature is constant due to the use of heating equipments, which has little effect on the diffusion and transfer of particulate matter. At the meanwhile, doors and windows are always closed and the impermeability of enclosure structure is well, the change of outdoor atmospheric pressure has little effect on the permeation process of indoor particulate matter.

Multivariate regression model fitting

According to the correlation analysis above, it found that all the factors mentioned could affect indoor particulate concentration. Based on these factors, a multivariate linear regression model of indoor $PM_{2.5}$ concentration prediction was proposed. In order to ensure the reliability of the results, the data is divided into two groups, one composed of 70% of the original data for regression, and the other 30% is used to verify the regression model. Table 2 shows the regression model between indoor particulate concentration and factors of four building types in summer and winter. The simulation results are standardized to ensure the comparability among the model parameters, which can be found that the range of determining coefficient R^2 obtained by calculation of four building types in summer and winter is 0.57 ~ 0.85, indicating that the model has a strong fitting.

Table 2. The regression model between indoor particulate concentrations

	Season	R^2	model
Office	Summer	0.57	y=0.15x_1+0.28x_2+0.62x_3+0.09x_4+0.04x_5-0.05x_6-0.10x_7-0.05x_8
	Winter	0.85	y=0.02x_1-0.03x_2+0.94x_3-0.01x_4+0.01x_5+0.06x_6+0.06x_7+0.04x_8
Classroom	Summer	0.67	y=0.09x_1+0.14x_2+0.77x_3+0.02x_4+0.04x_5+0.14x_6-0.03x_7+0.02x_8
	Winter	0.69	y=-0.03x_1+0.15x_2+0.80x_3-0.10x_4-0.05x_5+0.01x_6+0.02x_7-0.03x_8
Urban residence	Summer	0.67	y=0.09x_1+0.23x_2+0.76x_3-0.02x_4-0.02x_5+0.02x_6-0.03x_7-0.05 x_8
	Winter	0.69	y=0.08x_1+0.07x_2+0.65x_3-0.04x_4+0.02x_5+0.01x_6+0.01x_7+0.02x_8
Rural residence	Summer	0.67	y=0.22x_1+0.14x_2+0.73x_3-0.06x_4-0.11x_5-0.05x_6+0.05x_7+0.03x_8
	Winter	0.69	y=0.20x_1+0.06x_2+0.83x_3-0.03x_4-0.09x_5+0.11x_6-0.05x_7-0.02x_8

Note: x_1- Indoor temperature; x_2-Indoor humidity; x_3-Outdoor $PM_{2.5}$ concentration; x_4-Outdoor temperature ; x_5-Outdoor humidity; x_6-Outdoor atmospheric pressure; x_7-Outdoor Wind Direction; x_8-Outdoor air velocity

Table 3. The results analysis of performance indicators

	Office		Classroom		Urban residence		Rural residence	
	Summer	Winter	Summer	Winter	Summer	Winter	Summer	Winter
NAE	0.290	0.344	0.341	0.306	0.271	0.385	0.269	0.527
RMSE	10.250	20.540	8.420	9.710	10.730	14.510	16.890	32.230
R^2	0.740	0.868	0.684	0.664	0.773	0.743	0.876	0.756
PA	0.537	0.834	0.820	0.630	0.122	0.661	0.405	0.628
AI	0.719	0.680	0.724	0.682	0.699	0.997	0.700	0.556

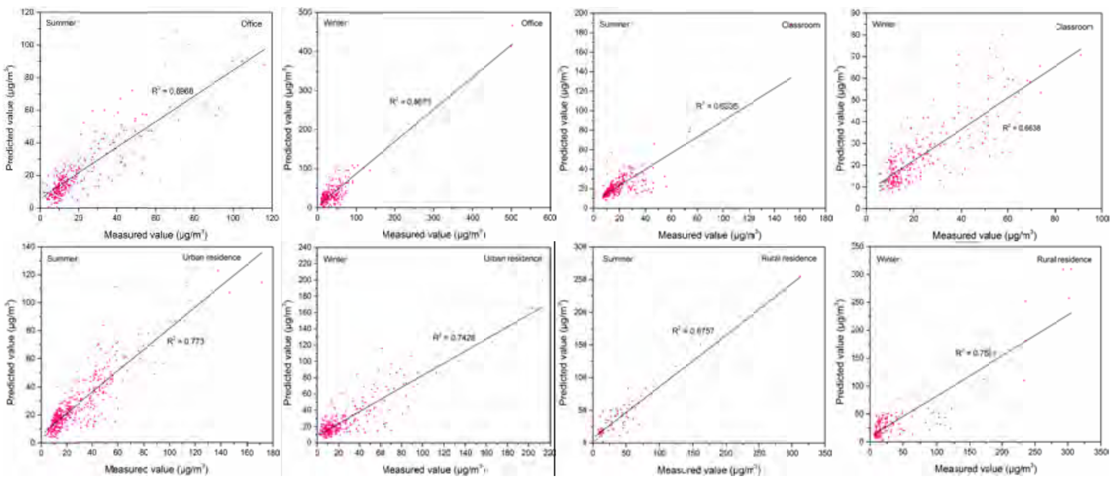

Figure 1. The comparison between predicted and measured values of regression model

According to the results of regression model in Table 2, the predicted value of the model is compared with the measured by 30% test data. Figure 1 is the result of comparison of each model. It can be found that there is a strong correlation between the measured and predicted values of indoor $PM_{2.5}$ concentration. Table 3 shows the analysis results of the errors between the measured values and the predicted values of five performance indexes, the values of NAE, RMSE are relatively small and R^2, PA and IA are close to 1, which indicates that the prediction model and the method are practical to some extent.

Comparative analysis of model parameters

Table 2 shows the regression results of the model parameters standardization. According to the comparative analysis , it can be found that the outdoor $PM_{2.5}$ concentration plays a leading role in the model, whether it is summer or winter, air velocity and wind direction are relatively weak, while the other parameters have their diversities respectively. For heating season, the influence of indoor temperature and humidity account for a large proportion of the model, while small in the non-heating season. Therefore, the correlation between indoor particulate concentration and all the parameters has significant seasonal differences, while no for different building types.

CONCLUSIONS

In this paper, the indoor and outdoor particulate concentrations of four building types were monitored to analyze the correlation and factors of indoor and outdoor particulate concentrations and establish the prediction model of indoor particulate concentration. Specific conclusions are as followed:

(1) Outdoor temperature and humidity, air velocity, wind direction and atmospheric pressure are the main factors affecting indoor particulate concentration. According to the results, indoor $PM_{2.5}$ concentration is positively correlated with indoor and outdoor temperature and humidity as well as outdoor air velocity, negatively with outdoor atmospheric pressure and both positively and negatively with outdoor wind direction.

(2) Based on these factors, the study uses multivariate linear regression to establish the indoor $PM_{2.5}$ concentration prediction model. No matter in summer and winter season, outdoor $PM_{2.5}$ concentration plays a leading role in the model, the outdoor air velocity and wind directions are relatively weak, while the other parameters have their diversities respectively.

ACKNOWLEDGEMENT

This research is supported by the China Environmental Protection Project (201509063), the National Nature Science Foundation of China (91743102, 51578103).

REFERENCES

Dockery D W, Pope C A, Xu X P, et al. 1993. An association between air pollution and morality in six United-States cities. *New England Journal of Medicine*, 329(24), 753-759.

Zhao J, Bo L, Gong C, et al.2015. Preliminary study to explore gene-PM2. 5 interactive effects on respiratory system in traffic policemen. *International journal of occupational medicine and environmental health*, 28(6), 971-983.

Tseng E, Ho W C, Lin M H, et al. 2015.Chronic exposure to particulate matter and risk of cardiovascular mortality: cohort study from Taiwan. *BMC public health*, 15(1), 936.

Cohen A J, Brauer M, Burnett R, et al. 2017.Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015. *The Lance*.

Klemm R J, Mason R M, Heilig C M, et al. 1996.Is daily mortality associated specifically with fine particles? *Journel of the Air&Waste Management Association*, 46(10), 927-965.

Dockery D W, Spengler J D. 1981. Indoor-outdoor relationships of respirable sulfates and particles. *Atmospheric Environment (1967)*, 15(3), 335-343.

Koutrakis P, Briggs S L K, Leaderer B P. 1992. Source apportionment of indoor aerosols in Suffolk and Onondaga Counties, New York. *Environmental Science & Technology*, 26(3), 521-527.

Ozkaynak H, Xue J, Spengler J, et al.1995. Personal exposure to airborne particles and metals: results from the Particle TEAM study in Riverside, California. *Journal of Exposure Analysis and Environmental Epidemiology*, 6(1), 57-78.

Xiong ZM, Zhang GQ, Peng JQ, et al. 2004. The research status of indoor inhalable particulate matter pollution. *HVAC*, 34 (4), 32-36.