Association Between Consecutive Ambient Air Pollution and Chronic Obstructive Pulmonary Disease Hospitalization: Time Series Study During 2015-2017 in Chengdu China

Yi Zhang[1,2†], Ziyue Wang[1,2†], Jingjing Xu[3], Yiqi Liu[4], Bin Zhou[5], Nan Zhang[5], Mingjie He[5], Jipeng Fan[5], Xianbo Liu[5], Jian Zhao[5], Qin Yang[1,2], Lifu Zhang[1,2], Yu Cao[1,2], and Steven Su[6]

Abstract— **This paper investigates the association between consecutive ambient air pollution and Chronic Obstructive Pulmonary Disease (COPD) hospitalization in Chengdu China. The three-year (2015-2017) time series data for both ambient air pollutant concentrations and COPD hospitalizations in Chengdu are approved for the study. The big data statistic analysis shows that Air Quality Index (AQI) exceeded the lighted air polluted level in Chengdu region are mainly attributed to particulate matters (i.e., $PM_{2.5}$ and PM_{10}). The time series study for consecutive ambient air pollutant concentrations reveal that AQI, $PM_{2.5}$, and PM_{10} are significantly positive correlated, especially when the number of consecutive polluted days is greater than nine days. The daily COPD hospitalizations for every 10 $\mu g/m^3$ increase in $PM_{2.5}$ and PM_{10} indicate that consecutive ambient air pollution can lead to an appearance of an elevation of COPD admissions, and also present that dynamic responses before and after the peak admission are different. Support Vector Regression (SVR) is then used to describe the dynamics of COPD hospitalizations to consecutive ambient air pollution. These findings will be further developed for region specific, hospital early notifications of COPD in responses to consecutive ambient air pollution.**

I. INTRODUCTION

Exacerbation of ambient air pollution has been becoming the critical focus research for health informatics, environment regulation, and big data analysis. It was noted by World Health Organization (WHO) Air Quality Standards that in China only five of the largest 500 cities attained complete air quality standards [1]. Recently Chronic Obstructive Pulmonary Disease (COPD) was progressed as the major cause of death in China [2]. The prevalence of COPD has been responsible for approximate one million deaths in China 2013, accounting for 31.1% of the total COPD deaths worldwide [3].

Medical and biological studies found associations between ambient air pollutants and morbidity of COPD. The ambient air pollutants with a diameter of 0.25-0.28 μm had a significant correlation with COPD symptoms [4]. Exposure to ambient air pollution generated the adverse risks to pneumonia [5] [6]. It also significantly correlated to children's emergency COPD hospital admissions and gave rise to oxidative stress [7]. Treatment for lung epithelial cells revealed that high concentrations of Polycyclic Aromatic Hydrocarbon (PAH) in polluted air significantly enhanced the definite diagnosis of COPD [8].

Daily counts of ambient air quality, characteristics of climatology, and hospital admissions were took into consideration for estimates of specific region based COPD hospitalizations. Days of being air polluted at low temperature and low relative humidity were more likely to cause the emergent COPD hospitalization [9] [10]. A multiple-region cooperative study among China, Ghana, India, Mexico, Russia, and South Africa, showed that per 10 $\mu g/m^3$ increase in $PM_{2.5}$, enhanced the number of COPD patients aged over 50 years old by 1.21%. In Shanghai, every 10 $\mu g/m^3$ increment in the concentration of $PM_{2.5}$ increased the relative risk of COPD by 9.0% [11].

Chengdu recently has been targeted as the study sample of ambient air pollution since both the urbanization progress and the special geography (i.e., steep topography, mountainous landform, cloudy and misty climate) conjointly give rise to consecutive ambient air pollution in various days. Studies based on 2015-2017 (731 days) time series data found that a positive correlation between air quality and mental illness hospitalization (e.g., every 10 $\mu g/m^3$ increase in $PM_{2.5}$ and PM_{10} concentrations responded the respective increases of 2.89% and 1.91% for mental illness hospitalizations) [12]. $PM_{2.5}$ and PM_{10} were significantly correlated to the number of hospitalized patients with respiratory and internal circulatory diseases (e.g., every 10 $\mu g/m^3$ increase in $PM_{2.5}$ and PM_{10} concentrations was responsible for hospitalization increases by 1.03% and 0.65%, respectively) [13]. Those

*This work was supported by the National Natural Science Foundation of China (61801094, 71661167005), the Fundamental Research Funds for the Central Universities China (ZYGX2019J086), the China Postdoctoral Science Foundation funded project (2017M612950), the International Scientific and Technological Innovation Cooperation of Science and Technology Department of Sichuan Province (2020YFH0093).

[1]School of Aeronautics and Astronautics, University of Electronic Science and Technology of China, Chengdu 611731, China

[2]Aircraft Swarm Intelpligent Sensing Cooperative Control, Key Laboratory of Sichuan Province, University of Electronic Science and Technology of China, Chengdu 611731, China yi.zhang@uestc.edu.cn; uestc_wzy@163.com; yangqinue@gmail.com; lifu@std.uestc.edu.cn; caoy.csu@gmail.com

[3]University of Electronic Science and Technology of China Hospital, Chengdu 610000 China xujj@uestc.edu.cn

[4]Xichang People's Hospital, Xichang 615000, China 898345905@qq.com

[5]Chengdu Gold Disk UESTC Multimedia Technology Co. Ltd, Chengdu 610000, China zhoubin@gdpacs.com; zhangnan@gdpacs.com; hemingjie@gdpacs.com; fanjipeng@gdpacs.com; lxb@gdpacs.com; zhaojian@gdpacs.com

[6]Biomedical Engineering School , University of Technology Sydney, Ultimo, NSW 2007, Australia steven.su@uts.edu.au

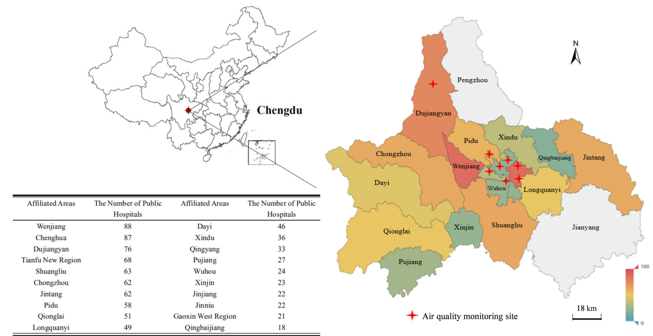
†Both of authors equally contribute to this work.

data in 2015-2016 (365 days) also revealed the positive correlation between air pollution and COPD admission rates in the low temperature [14].

Based on three-year (2015-2017) time series data in Chengdu, literatures provided adequate outcomes for the correlated link of hospital admissions with key variables such as ambient air pollutants, weather conditions, and population composition. In order to develop an accurate predictive and prescriptive big data approach for the early notification of COPD hospitalization, the association between consecutive ambient air pollution and COPD hospitalization is further investigated. The 2015-2017 consecutive polluted time series for both ambient air pollutant concentration data and COPD hospitalization data are handled. The dynamics of increase percentages in daily COPD hospitalizations for every 10-unit increase in air pollution concentration are statistically depicted, and then modelled by using robust Support Vector Regression (SVR). The study implies the region specific, early notification solutions for hospitalizations under consecutive ambient air pollution.

II. Data Description

The Institutions Ethical Review Board approved all experimental procedures involving human subjects. The Health Commission of Sichuan Province approved the University of Electronic Science and Technology of China (UESTC) and Goldisc Multimedia Technology Co. Ltd to utilize Electronic Medical Records (EMRs) of Chengdu hospitals during 2015-2017 for this study. Daily COPD hospital admissions were collected across 1073 local hospitals (Figure 1) in Chengdu and its neighboring townships (including Dujiangyan, Qionglai, and Chongzhou). The AQI criteria ranked over the lighted polluted level ($AQI \geq 100$) was applied to filter the air pollutant concentration and COPD hospitalization time series from January 1, 2015 to December 31, 2017. Then, the consecutive ambient air pollution time series can be identified, the statistical characteristics of which were showed in Table 1. The patient diagnoses cataloged to J44 related to International Classification of Diseases, 10^{th} Revision (ICD-10) were confirmed as COPD hospitalizations. There are 111,740 COPD hospital admissions recorded, where J44.1 (COPD with Acute Exacerbation) occupies accounted for 64.9%, followed by J44.101 (Chronic Obstructive Emphysema Bronchitis with Acute Exacerbation), J44.801 & J44.803 (Chronic Bronchitis with Emphysema) and J44.9 (Unspecified COPD), accounting for 16.0%, 8.0% and 4.7%, respectively. The applied data for ambient air pollutant concentrations included $PM_{2.5}$, PM_{10}, Sulfur Dioxide (SO_2), Nitrogen Dioxide (NO_2), Carbon Monoxide (CO), Ozone (O_3), and AQI. Since the limit settings of Ambient Air Quality Standard of China (GB3095-2012) are more stringent compared to WHO Air Quality Guidelines 2005, the former particularly in Class II Concentration Index was used as the benchmark of the study.



Affiliated Areas	The Number of Public Hospitals	Affiliated Areas	The Number of Public Hospitals
Wenjiang	88	Dayi	46
Chenghua	87	Xindu	36
Dujiangyan	76	Qingyang	33
Tianfu New Region	68	Pujiang	27
Shuangliu	63	Wuhou	24
Chongzhou	62	Xinjin	23
Jintang	62	Jinjiang	22
Pidu	58	Jinniu	22
Qionglai	51	Gaoxin West Region	21
Longquanyi	49	Qingbaijiang	18

Fig. 1. Hospitals and air monitoring stations in Chengdu. The red star indicates eight real-time monitoring stations for meteorological observatories, i.e., Sanwa Kiln, Shilidian, Junpinpg Street, Liangjia Avenue, Shahepu, Lingyan Temple, Caotang Temple, and Jinquan Lianghe. The color bar presents the number of public and private hospitals in Chengdu.

III. Methods

The increase percentage of daily COPD hospitalizations for every 10-unit increase in air pollution concentrations (AQI, $PM_{2.5}$, PM_{10}, SO_2, NO_2, CO, and O_3) can be described as

$$Q_i = \sum_{j=1}^{J} \frac{10\left(x_{i,j} - x\right)}{x\left(z_{i,j} - z\right)} / j \times 100\% \tag{1}$$

where i denotes the number of consecutive ambient air pollution days, $i \in [2, I]$, j presents the number of data segments for the i-day consecutive ambient air pollution data, $j \in [1, J]$, $x_{i,j}$ and $z_{i,j}$ respectively indicate the number of COPD hospitalizations and the air pollutant concentration in the j^{th} day at the data segment of the i-day consecutive ambient air pollution, x and z represent the baselines of the number of COPD hospitalizations and the air pollutant concentration.

Support Vector Regression (SVR) is applied to describe the dynamics of the daily increase percentage of COPD hospitalization for per 10-unit increase in air pollutant concentration. Assume that the mean of Q_i compose the input vector and its high dimensional feature is transformed by Gaussian Radial Basis Function $K(x_i, x) = \exp\left(-\frac{\|x_i - x\|}{2\sigma^2}\right)$. Then the approximation of SVR for the input vector can be described by

$$\mathrm{y_i} = w^T \cdot K(x_i, x) + b \tag{2}$$

The quadratic optimization problem can be formed by

$$\begin{aligned} & Minimize\ \frac{1}{2}\|w\|^2 + C\sum_{i=2}^{I}\left(\xi_i + \xi_i^*\right) \\ & subject\ \ to \\ & y_i - wx_i - b \leq \varepsilon + \xi_i \\ & wx_i + b - y_i \leq \varepsilon + \xi_i^* \\ & \xi_i, \xi_i^* \geq 0 \end{aligned} \tag{3}$$

where C, ε, ξ_i, and ξ_i^* are the free parameters.

TABLE I

STATISTIC RESULTS FOR AMBIENT AIR POLLUTION CONCENTRATIONS IN CONSECUTIVE POLLUTED DAYS WHEN THERE WERE OVER THE *AQI* CRITERIA

	Mean±SD	Class II Concentration Index**	Minimum	$P_{25^{th}}$***	$P_{50^{th}}$***	$P_{75^{th}}$***	Maximum	IQR***
Air Pollutants								
NO_2 (μg/m^3)	66.7±14	80	30.9	57.6	64.8	74.3	118	16.7
O_3 (μg/m^3)	108.1±60.1	N/A	23.2	58.8	94.5	143.2	289.8	84.5
$PM_{2.5}$ (μg/m^3)*	108.2±39.6	75	40.2	79.1	98.1	130.9	296.6	51.7
PM_{10} (μg/m^3)*	172.1±57.6	150	85.4	131	159.6	198.3	450	67.3
SO_2 (μg/m^3)	18.7±5.4	150	9.2	14.7	18.1	22.5	37.1	7.9
CO (μg/m^3)	0.9±0.36	10	0	0.755	0.88	1.05	2.18	0.295

* $PM_{2.5}$ (PM_{10}) indicates the averaged concentration value of the particulate matter in the air with aerodynamic diameter is less than or equal to 2.5 μm (10 μm) across twenty four hours.
** Class II Concentration Index refers to Ambient Air Quality Standard of China (GB3095-2012), over which the air quality was classified as a polluted level.
*** $P_{25^{th}}$, $P_{50^{th}}$, $P_{75^{th}}$, and IQR represents lower quartile, median, upper quartile, and inter-quartile range, respectively.

TABLE II

STATISTIC RESULTS FOR COPD HOSPITALIZATIONS IN CONSECUTIVE POLLUTED DAYS WHEN THERE WERE OVER THE *AQI* CRITERIA

	N*	Mean±SD	Minimum	$P_{25^{th}}$	$P_{50^{th}}$	$P_{75^{th}}$	Maximum	IQR	COPD Code**
COPD									
COPD with acute lower respiratory tract infection	4019	13.3± 8.4	1	7	11	17	44	10	J44.0
COPD with acute exacerbation	72609	239.6±151.6	29	132	206	293.5	760	162	J44.1
Chronic obstructive emphysema bronchitis with acute exacerbation	17924	59.2 ± 31.3	7	32.5	60	80	183	47.5	J44.101
Other specified COPD	182	0.6±0.9	0	0	0	1	6	1	J44.8
Chronic bronchitis with emphysema	5204	17.2±3.7	0	9	14	19.5	65	10.5	J44.801 & J44.803
Chronic asthmatic bronchitis	2431	8.0±4.6	0	4	8	11	22	7	J44.802 & J44.804
Chronic bronchiolitis	179	0.6±0.9	0	0	0	1	4	1	J44.805
Chronic obstructive bronchitis	273	0.9±1.2	0	0	0	1	1	5	J44.806
Unspecified COPD	8919	29.4±18.7	1	17	26	37	110	20	J44.9

* The total number of COPD hospitalization samples.
** COPD code referred to the code of COPD in ICD-10 codes released in 2013.

IV. RESULTS

Table I indicated that the three-year air pollutant concentrations had the right-skewed distributions. SO_2, NO_2, CO, and O_3 concentrations did not exceed Class II Concentration Index. However, the mean values of $PM_{2.5}$ and PM_{10} concentrations reached 108.2 $\mu g/m^3$ and 172.1 $\mu g/m^3$ respectively, while the limit sets according to Class II Concentration Index were referred to 75 $\mu g/m^3$ and 105 $\mu g/m^3$. In Table II, J44.1, J44.101, J44.801 & J44.803, J44.802 & J44.804, and J44.9 showed the right-skewed or normal distributions.

Based on statistic results in Tables I and II, the spearman correlation test was applied for the segmented consecutive polluted time series. The maximal consecutive polluted days of AQI, $PM_{2.5}$, and PM_{10} were 24, 18, and 17, respectively. The correlation tests for the segmented consecutive polluted time series in each duration was included in Table III. Tests between AQI vs. $PM_{2.5}$, AQI vs. PM_{10}, and $PM_{2.5}$ vs. PM_{10} showed a positive correlation from 2-17 days. Significant collections ($p < 0.05$) can be found in 9-17 days in AQI (vs. $PM_{2.5}$), 9-14 days in AQI (vs. PM_{10}), and 3-17 days in $PM_{2.5}$ (vs. PM_{10}). The outcomes indicated that the particulate matter $PM_{2.5}$ and PM_{10} had a significant influence on Chengdu ambient air quality when consecutive air pollution occurred.

The mean percentages of increase in daily COPD hospitalizations (95% confidence intervals) for every 10-unit increase in AQI, $PM_{2.5}$, and PM_{10} were statistically calculated based on the IQR values of daily COPD hospitalizations and ambient air pollution concentration (Figure 2). It can be seen that the peak increases were found at the 2^{nd}, 6^{th}, and 9^{th} consecutive days for AQI and $PM_{2.5}$, and the 3^{rd} and 9^{th} consecutive days for PM_{10}. After the 9^{st} consecutive day, a drop-off followed by a high-level holding with slight increases was introduced. Therefore, the different dynamics before and after the most peak increased day were separately modeled by using SVR, which has been seen in Figure 2.

The limitations of this study were only for regional (Chengdu China) big data analytics about specific COPD hospitalizations. The motivation of this work was originally from the increased hospital burden for COPD in Chengdu hospital systems. In the future work, the early notification approach of COPD hospitalization burden in response to the situation of consecutive ambient air pollution will be further explored.

V. CONCLUSION

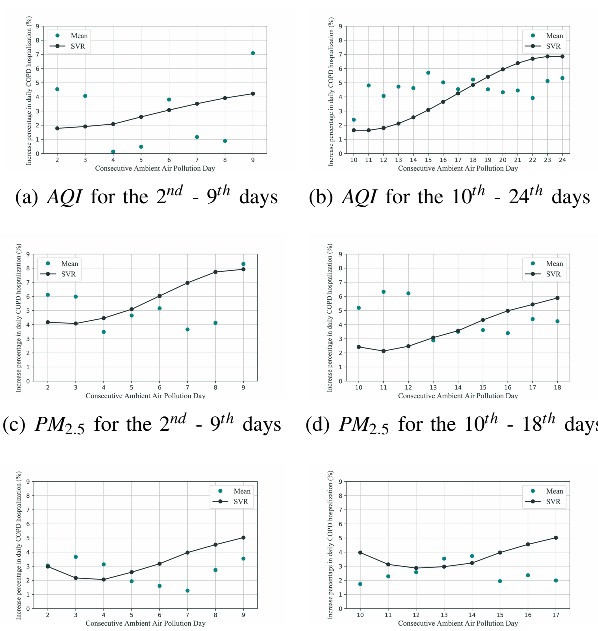
This study investigated the specific Chengdu region, daily COPD hospitalization increases to ambient air pollutions.

TABLE III

SPEARMAN CORRELATIONS AMONG *AQI*, $PM_{2.5}$, AND PM_{10} BASED ON CONSECUTIVE AMBIENT AIR POLLUTION IN 2015-2017 CHENGDU

Variables	Consecutive Ambient Air Pollution Day																
	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
$PM_{2.5}$ vs. PM_{10}	0.762	0.815**	0.893**	0.921**	0.932**	0.889**	0.907**	0.924**	0.898**	0.949**	0.908**	0.793	0.960**	0.967**	0.950**	0.863**	-
$PM_{2.5}$ vs. *AQI*	0.783	0.717**	0.637	0.536	0.516	0.656	0.741	0.784**	0.808**	0.811**	0.832**	0.841**	0.841**	0.829**	0.815**	0.815**	0.756
PM_{10} vs. *AQI*	0.762	0.687	0.667	0.645	0.743**	0.763**	0.705	0.808**	0.841**	0.871**	0.884**	0.839**	0.839**	0.803	0.725	0.716	-

** $p < 0.05$



(a) *AQI* for the 2^{nd} - 9^{th} days (b) *AQI* for the 10^{th} - 24^{th} days

(c) $PM_{2.5}$ for the 2^{nd} - 9^{th} days (d) $PM_{2.5}$ for the 10^{th} - 18^{th} days

(e) PM_{10} for the 2^{nd} - 9^{th} days (f) PM_{10} for the 10^{th} - 17^{th} days

Fig. 2. Increase percentages in daily COPD hospitalizations for every 10-unit increase in specific air pollution concentrations. The blue round dot indicated the mean value in consecutive ambient air pollution days, and the solid dot curve denoted SVR fitting dynamics of daily COPD hospitalizations with respect to consecutive ambient air pollution.

The day-to-day ambient air quality data and COPD hospitalization data from January 1, 2015 to December 31, 2017 were used for this big data analysis research. The consecutive ambient air pollution days were marked according to Air Quality Standard of China. The air pollution concentrations and COPD hospitalizations time series in consecutive ambient air pollution days were then segmented. Statistical analyses found particulate matters $PM_{2.5}$ and PM_{10} were the key variables for elevated COPD hospitalizations. Spearman correlation tests indicated a positive correlation among *AQI* and particulate matters. The dynamics of daily COPD hospitalization increases were then modeled by SVR, where the peak increase and drop-off afterwards can be well observed and depicted. In the future work, the early notification for Chengdu COPD hospitalization relative to ambient air pollution will be further explored.

REFERENCES

[1] Chen Z, Wang J N, Ma G X, et al. China tackles the health effects of air pollution[J]. Lancet (London, England), 2013, 382(9909): 1959.

[2] Fang L, Gao P, Bao H, et al. Chronic obstructive pulmonary disease in China: a nationwide prevalence study[J]. The Lancet Respiratory Medicine, 2018, 6(6): 421-430.

[3] Yin P, Wang H, Vos T, et al. A subnational analysis of mortality and prevalence of COPD in China from 1990 to 2013: findings from the global burden of disease study 2013[J]. Chest, 2016, 150(6): 1269-1280.

[4] Peng L, Xiao S, Gao W, et al. Short-term associations between size-fractionated particulate air pollution and COPD mortality in Shanghai, China[J]. Environmental Pollution, 2019: 113483.

[5] Ho S C, Chuang K J, Lee K Y, et al. Chronic obstructive pulmonary disease patients have a higher risk of occurrence of pneumonia by air pollution[J]. Science of the Total Environment, 2019, 677: 524-529.

[6] Halonen J I, Lanki T, Yli-Tuomi T, et al. Urban air pollution, and asthma and COPD hospital emergency room visits[J]. Thorax, 2008, 63(7): 635-641.

[7] MacNee W, Donaldson K. Mechanism of lung injury caused by PM10 and ultrafine particles with special reference to COPD[J]. European Respiratory Journal, 2003, 21(40 suppl): 47s-51s.

[8] Yang L, Wang W C, Lung S C C, et al. Polycyclic aromatic hydrocarbons are associated with increased risk of chronic obstructive pulmonary disease during haze events in China[J]. Science of the Total Environment, 2017, 574: 1649-1658.

[9] Sun X W, Chen P L, Ren L, et al. The cumulative effect of air pollutants on the acute exacerbation of COPD in Shanghai, China[J]. Science of The Total Environment, 2018, 622: 875-881.

[10] Qiu H, Yu I T S, Wang X, et al. Season and humidity dependence of the effects of air pollution on COPD hospitalizations in Hong Kong[J]. Atmospheric Environment, 2013, 76: 74-80.

[11] Lin H, Qian Z M, Guo Y, et al. The attributable risk of chronic obstructive pulmonary disease due to ambient fine particulate pollution among older adults[J]. Environment international, 2018, 113: 143-148.

[12] Qiu H, Zhu X, Wang L, et al. Attributable risk of hospital admissions for overall and specific mental disorders due to particulate matter pollution: A time-series study in Chengdu, China[J]. Environmental research, 2019, 170: 230-237.

[13] Zhu X, Qiu H, Wang L, et al. Risks of hospital admissions from a spectrum of causes associated with particulate matter pollution[J]. Science of The Total Environment, 2019, 656: 90-100.

[14] Qiu H, Tan K, Long F, et al. The burden of COPD morbidity attributable to the interaction between ambient air pollution and temperature in Chengdu, China[J]. International journal of environmental research and public health, 2018, 15(3): 492.