Environmental Science and Pollution Research (2023) 30:88431–88443
https://doi.org/10.1007/s11356-023-28682-8

RESEARCH ARTICLE

Application of machine learning to predict hospital visits for respiratory diseases using meteorological and air pollution factors in Linyi, China

Jing Yang[1] · Xin Xu[2] · Xiaotian Ma[3] · Zhaotong Wang[2] · Qian You[4] · Wanyue Shan[2] · Ying Yang[2] · Xin Bo[2,5] · Chuansheng Yin[1]

Received: 10 March 2023 / Accepted: 4 July 2023 / Published online: 12 July 2023

Abstract

Urbanization and industrial development have resulted in increased air pollution, which is concerning for public health. This study evaluates the effect of meteorological factors and air pollution on hospital visits for respiratory diseases (pneumonia, acute upper respiratory infections, and chronic lower respiratory diseases). The test dataset comprises meteorological parameters, air pollutant concentrations, and outpatient hospital visits for respiratory diseases in Linyi, China, from January 1, 2016 to August 20, 2022. We use support vector regression (SVR) to build models that enable analysis of the effect of meteorological factors and air pollutants on the number of outpatient visits for respiratory diseases. Spearman correlation analysis and SVR model results indicate that NO_2, $PM_{2.5}$, and PM_{10} are correlated with the occurrence of respiratory diseases, with the strongest correlation relating to pneumonia. An increase in the daily average temperature and daily relative humidity decreases the number of patients with pneumonia and chronic lower respiratory diseases but increases the number of patients with acute upper respiratory infections. The SVR modeling has the potential to predict the number of respiratory-related hospital visits. This work demonstrates that machine learning can be combined with meteorological and air pollution data for disease prediction, providing a useful tool whereby policymakers can take preventive measures.

Keywords Air pollution · Meteorological factors · Machine learning · Respiratory disease · Daily hospital visits · Correlation analysis

Introduction

Air pollution has been increasing at an alarming rate and is now considered to be an important risk factor for human morbidity and mortality (Al-Kindi et al. 2020). According to the World Health Organization (WHO), 99% of the world's population lives in areas where contaminant levels exceed the air quality threshold values stated in the most recent WHO guidelines; this exposure has resulted in millions of deaths from illnesses caused by air pollution (Requia et al.

Responsible Editor: Lotfi Aleya

Highlights

• Machine learning is used to examine correlations between meteorological and air pollution factors and outpatient visits.
• Pneumonia patient visits are strongly correlated with NO_2, $PM_{2.5}$, and PM_{10}.
• The SVR model shows the best performance for pneumonia.
• The link between meteorological and air pollution factors and respiratory diseases is examined.

✉ Chuansheng Yin
yinchuansheng@lysrmyy.cn

1 Intersection of Wohushan Road and Wuhan Road in Beicheng New Area, Linyi People's Hospital, Linyi 276000, Shandong, People's Republic of China

2 Department of Environmental Science and Engineering, Beijing University of Chemical Technology, Beijing 100029, People's Republic of China

3 School of Information and Control Engineering, Jilin Institute of Chemical Technology, Jilin City 132022, People's Republic of China

4 School of Management and Engineering, Capital University of Economics and Business, Beijing 100070, People's Republic of China

5 BUCT Institute for Carbon-Neutrality of Chinese Industries, Beijing 100029, People's Republic of China

2021; Yang et al. 2022a). Alarmingly, 7 million deaths were caused by air pollution in 2018 (WHO 2018). Excessive exposure to air pollution can lead to respiratory diseases and even cancer. Air pollution adversely affects public health and the image of a city, especially in cities with uncontrolled growth of industrial activity and high fossil fuel consumption (Almetwally et al. 2020; Margiana et al. 2022; Choi and Myong 2018; Grigorieva and Lukyanets 2021). Additionally, the health risks resulting from extreme weather events (such as heat waves and cold snaps) have attracted increasing attention (Costello et al. 2009; Kan et al. 2012; Semenza and Paz 2021). For example, Bodaghkhani et al. (2019) carried out systematic research on the effects of meteorological factors on hospitalizations in adult patients with asthma and found that the weather plays an important role in exacerbating asthma by affecting water loss and the evaporation of surface fluid in airways. Qi et al. (2021) demonstrated that multiple meteorological factors have strong associations with influenza activity in Chongqing, while Javorac et al. (2021) confirmed the apparent effects of temperature, atmospheric pressure, humidity, and heat on chronic obstructive pulmonary disease. Therefore, assessments of the relationships between meteorological and air pollution factors and the rate of hospital visits for respiratory diseases are vital in supporting intervention measures.

Predicting the effect of air pollution concentrations on the rate of hospital visits for respiratory diseases is challenging because this type of pollution is spatially and temporally dynamic and can be affected by complex chemical mechanisms, meteorological conditions, points of interest, and other factors (Zhang et al. 2022a). Predicting air pollutant concentrations involves nonlinear dynamics, as the pollutants diffuse in the atmosphere, and such predictions are expensive to carry out (Delavar et al. 2019). Various methods have been used to forecast air pollutant concentrations, including traditional statistical prediction methods, numerical prediction methods, artificial intelligence, and mixed-model theory and its applications (Bai et al. 2018). In recent years, machine learning and artificial intelligence have received considerable attention in the field of air quality prediction and atmospheric investigation and have been widely used in clinical decision-making for the diagnosis and management of air pollution-related diseases (Masmoudi et al. 2020; Masood and Ahmad 2021). Support vector regression (SVR) is a nonlinear model that has been widely used for disease diagnosis and prediction, as it is less likely to be affected by data overfitting than other classification algorithms (Otchere et al. 2021). SVR uses sparse training data and is highly efficient in identifying nonlinear relationships between input and output data (Li et al. 2021). For example, Razavi-Termeh et al. (2022) used an SVR model to analyze the effects of different seasons on the occurrence of asthma and provided a reference for establishing a respiratory disease prediction model. Ravindra et al. (2023) established an SVR model that explored the effect of air pollution and meteorological parameters on patient hospital visits for respiratory diseases and showed that this model had the potential to predict the likelihood of hospital visits for acute respiratory infections. Analogously, SVR models have exhibited prediction potential in daily pediatric inpatient hospital admissions relating to ambient air pollutant exposure (Zhou et al. 2019). These studies demonstrate that SVR models can capture the complex nonlinear relationships between air pollutant concentrations, meteorological factors, and respiratory diseases, enabling better interventions to mitigate air pollution and reduce public exposure risk and morbidity.

Air pollution is strongly related to the occurrence of respiratory diseases. Previous studies have shown that air pollutants (NO_2, SO_2, PM_{10}, and $PM_{2.5}$) can accumulate in lung tissue (Falcon-Rodriguez et al. 2016), and ultra-fine particles can be transferred to distal organs and tissues through the blood, causing respiratory infections, asthma, and even lung cancer (Darrow et al. 2014; Jacquemin et al. 2015). In addition, several cohort studies have suggested that PM_{10} and $PM_{2.5}$ are positively associated with respiratory diseases (Kloog et al. 2014; Lu et al. 2015). Pothirat et al. (2019) showed that PM_{10} and $PM_{2.5}$ were associated with the acute exacerbation of chronic obstructive pulmonary disease, while emergency visits for community-acquired pneumonia were associated with SO_2, CO, and O_3. Air pollution is a major challenge worldwide, especially in rapidly developing industrialized cities. Linyi is a typical industrialized city located in Shandong Province, China; it is the largest city in the province, with a total land area of 17,200 km^2 (Zhao et al. 2017). The population of Linyi is approximately 10,810,000—similar to that of several European countries (Scott et al. 2018; Blomgren and Virta 2020). The GDP of Linyi has steadily grown in recent years and was recently assessed as 480 billion yuan (Zhang et al. 2022b; Bureau 2021). With increasing economic development and energy consumption, air pollution has also increased, which has resulted in some economic losses for the city (Yi et al. 2016; Wu et al. 2021). To date, no study has used machine learning to explore the effect of meteorological and air pollution factors on patients with respiratory diseases in Linyi. The use of this method to identify relationships between meteorological and air pollution factors and human health in Linyi can provide a reference for decision-makers in Linyi to enact better air quality policies and take disease intervention measures.

This paper describes an SVR model for analyzing the complex links between air pollutants, weather parameters, and the number of patients with respiratory diseases. The model is applied to data recorded in Linyi from January 1, 2016 to August 20, 2022. We use several metrics (mean absolute error, mean squared error, and coefficient of determination) to evaluate the performance of the regression model. The meteorological and air pollution factors affecting respiratory diseases are identified based on the prediction accuracy. This study provides support and a reference for policymakers in industrialized cities to enact interventions that will reduce the incidence of respiratory diseases and increase the health and well-being of urban residents.

Material and methods

Study area

Linyi is located in the southeast of Shandong Province, China. Covering nine counties and three districts (Fig. 1), with a total land area of 17,200 km^2 and a population of 10,810,000, Linyi is the largest and most populous city in Shandong Province (Zhao et al. 2017).

In recent years, Linyi has experienced rapid development in heavy industries such as machinery, metallurgy, and building materials, which has driven considerable growth in energy demand. The environmental problems brought about by this economic development, especially air pollution, have caused widespread concern (Yang et al. 2022b; Jiang et al. 2022). The burning of fossil fuels related to industrial development releases large amounts of toxic compounds. The pollution is exacerbated by the region's geography—the mountains on three sides of the city limit the diffusion of pollutants, thus aggravating the pollution concentrations in the atmosphere above Linyi. In addition, the annual precipitation in this area is uneven and mainly falls in summer. The low precipitation levels in winter accentuate the air pollution problems and contribute to the formation of smog (Yin et al. 2021; Ji and Li 2016). In recent years, climate change has led to periods of extremely high temperatures and cold snaps in Linyi (Wu et al. 2022). The Ministry of Environmental Protection of the People's Republic of China has supervised the air quality in Linyi since 2015 (Zhang et al. 2018). Although the air quality has improved in recent years, the incidence of respiratory diseases has risen owing to the lagging effect of air pollution on public health (Li et al. 2020).

A previous study found that air pollutant emissions affect the prevalence of respiratory system diseases in Linyi (Wu et al. 2021). However, few studies have investigated how the combination of meteorological and air pollution factors relate to respiratory diseases. In this study, we used machine learning to assess the effects of weather and air pollutants on respiratory diseases in residents, providing data that support city managers in their efforts to enact intervention measures to prevent air pollution. In this way, the findings of this study will benefit public health and help to maintain the city's image.

Data sources and quality control

A flowchart of the data quality control procedure employed in this study is shown in Fig. 2. Meteorological data were obtained from the China Meteorological Administration

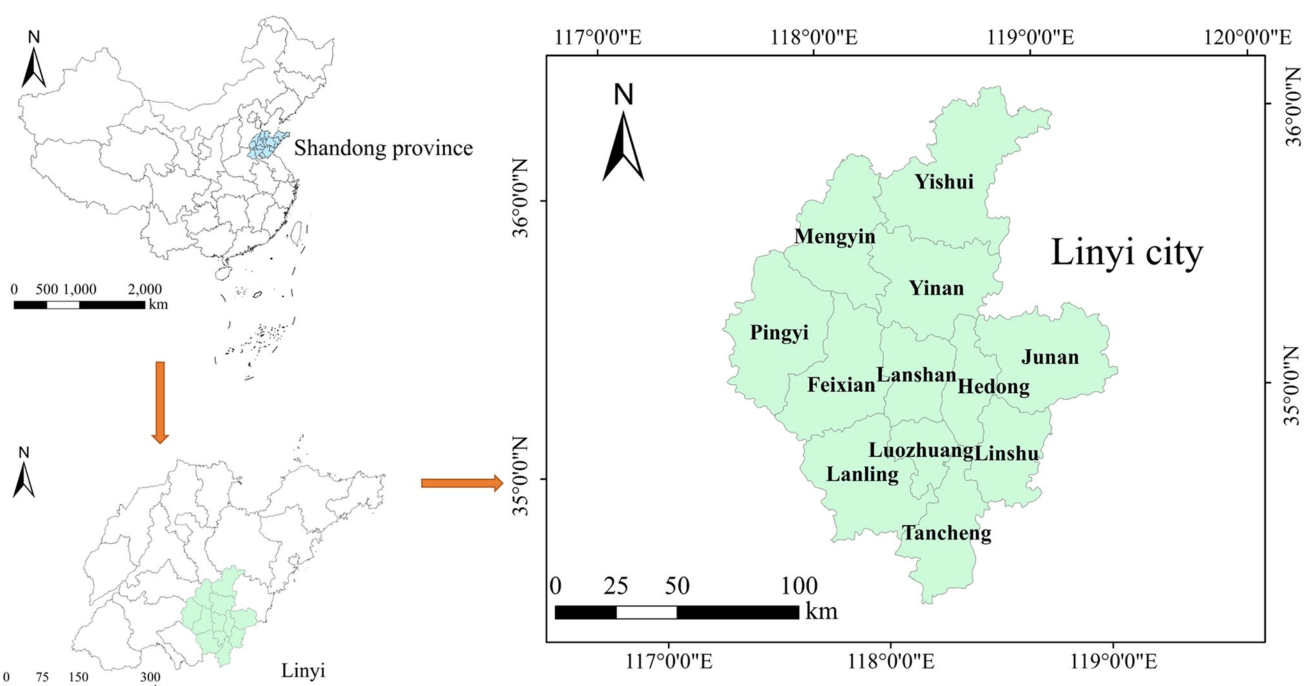


Fig. 1 Topographic information and location of the study area

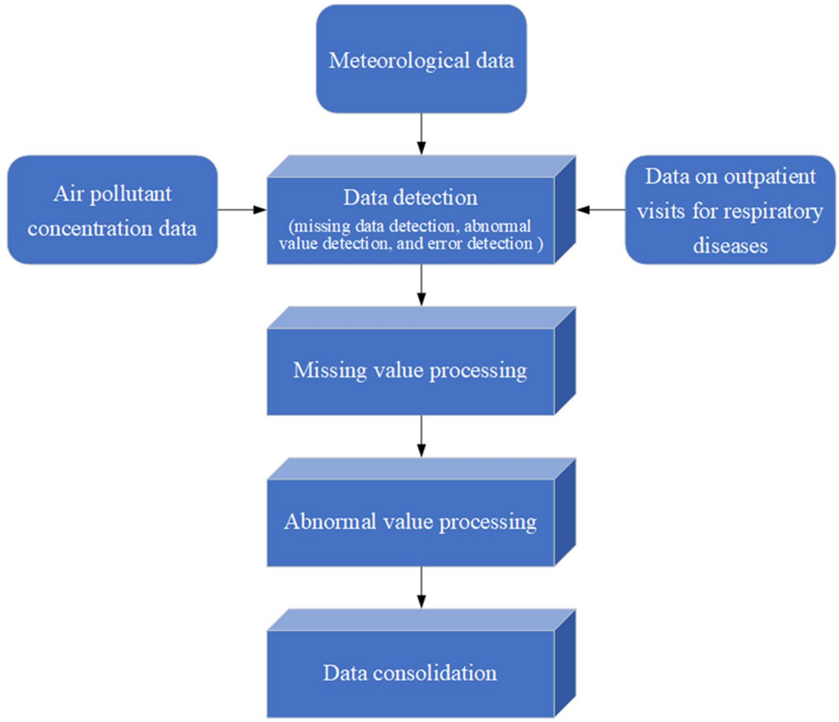


Fig. 2 Flowchart of data quality control in this study

observation station of Linyi (35.03°N, 118.24°E); these data included the average temperature, average humidity, and average wind speed, and all meteorological parameters were calculated from hourly observation data. All quality controls for these data were performed by the meteorological department.

The air pollutant concentrations were obtained from four state-controlled monitoring stations in Linyi, namely, Lunan Pharmaceutical Factory, Hedong Insurance Company, Yihe Community, and Xinguang Woolen Mill. The concentrations of air pollutants (NO_2, SO_2, CO, O_3, PM_{10}, and $PM_{2.5}$) were collected from January 1, 2016 to August 20, 2022. Quality control was applied to the original data, including missing data detection, error detection, and possible error correction, to support the applicability and availability of the data (Lippmann and Maynard 1999; Gómez-Carracedo et al. 2014; Appaia and Palraj 2023).

The dataset of daily patient hospital visits for respiratory diseases was obtained from Linyi People's Hospital, which is the only Grade 3, Class A general hospital in Linyi. This dataset covered the period from January 1, 2016 to August 20, 2022. The original outpatient data were checked by management personnel and information management personnel according to the national standards for tertiary hospitals. Cases with incomplete outpatient information, repeat outpatient visits, and unclear diagnoses were removed. In total, 99,894 cases of respiratory disease, according to the Tenth Edition of the International Classification of Diseases codes J00–J99 (ICD-10: J00–J99), were identified. Based on the characteristics of the hospital dataset, we selected three respiratory diseases for inclusion in this study: acute upper respiratory infections (ICD-10: J00–J06), pneumonia (ICD-10: J12–J18), and chronic lower respiratory diseases (ICD-10: J40–J47).

Spearman correlation coefficient

The Spearman rank correlation coefficient represents the degree of correlation between two sequence variables using the rank of a variable (Keshavarz and Mostofinejad 2019; Thirumalai et al. 2017). In this study, the different datasets were not normally distributed, so Spearman correlation analysis was used to explore the effect of air quality and meteorological factors on outpatient visits for respiratory diseases.

Support vector regression

SVR is a supervised learning model that is used to analyze data for classification and regression analysis; it mainly uses the inductive principle of structural risk minimization (Araste et al. 2022; Khan et al. 2022). The objective of SVR is to select characteristic subsets from training samples so that their classification will be equivalent to the division of the entire dataset and then to classify the data by determining the hyperplane that maximizes the boundary between classes in the training data (Zheng et al. 2019). This maximizes the margin between two vector classes and minimizes the error of the training dataset, thus reducing the influence of data overfitting (Mai et al. 2022). SVR has been successfully used for disease diagnosis and prediction (Hao et al. 2020; Choudhury and Gupta 2019; Nilashi et al. 2019).

Table 1 Patient and environmental characteristics

	2016	2017	2018	2019	2020	2021	2022	Total
Total ED visits, *n*	9295	16782	17693	20176	12946	14618	8384	99894
Age (*mean* ± *SD*)	32.73 ± 30.94	34.96 ± 30.85	35.33 ± 32.15	37.54 ± 32.69	47.29 ± 33.11	49.15 ± 30.79	51.99 ± 30.09	41.28 ± 7.32
0–15, *n* (%)	4771 (51.33%)	8210 (48.92%)	8572 (48.45%)	9079 (45.00%)	3895 (30.09%)	3955 (27.06%)	2002 (23.88%)	40484 (40.53%)
16–60, *n* (%)	1491 (16.04%)	2479 (14.77%)	2590 (14.64%)	3157 (15.65)%	2538 (19.60%)	3063 (20.95%)	1738 (20.73%)	17056 (17.07%)
> 60, *n* (%)	3033 (32.63%)	6093 (36.31%)	6531 (36.91%)	7940 (39.35%)	6513 (50.31%)	7600 (51.99%)	4644 (55.39%)	42354 (42.40%)
Male, *n* (%)	5831 (62.73%)	10285 (61.29%)	11052 (62.47%)	12651 (62.70%)	8358 (64.56%)	9445 (64.61%)	5446 (64.96%)	63068 (63.13%)
Female, *n* (%)	3464 (37.27%)	6497 (38.71%)	6641 (37.53%)	7525 (37.30%)	4588 (35.44%)	5173 (35.39%)	2938 (35.04%)	36826 (36.87%)
P (J12–J18)	2727 (29.34%)	5372 (32.01%)	6173 (34.89%)	7040 (34.89%)	3380 (26.11%)	3479 (23.8%)	1569 (18.71%)	29740 (29.77%)
AURI J00–J06	2442 (26.27%)	3814 (22.73%)	3682 (20.81%)	3950 (19.58%)	1860 (14.37%)	2385 (16.32%)	1781 (21.24%)	19914 (19.94%)
CLRD (J40–J47)	4126 (44.39%)	7596 (45.26%)	7838 (44.3%)	9186 (45.53%)	7706 (59.52%)	8754 (59.89%)	5034 (60.04%)	50240 (50.29%)
$PM_{2.5}$ (*mean* ± *SD*)	66.13 ± 4.76	57.07 ± 3.74	51.73 ± 2.20	58.06 ± 1.68	48.37 ± 1.55	44.87 ± 1.60	41.03 ± 26.99	52.47 ± 7.97
PM_{10} (*mean* ± *SD*)	124.79 ± 11.41	114.95 ± 7.86	107.77 ± 3.13	112.89 ± 5.22	92.00 ± 3.16	91.28 ± 2.80	71.23 ± 37.93	102.13 ± 16.91
SO_2 (*mean* ± *SD*)	28.82 ± 5.04	24.02 ± 3.74	19.06 ± 2.48	15.86 ± 2.49	11.59 ± 1.38	11.81 ± 0.83	11.2 ± 4.99	17.48 ± 6.36
NO_2 (*mean* ± *SD*)	42.05 ± 3.36	43.88 ± 1.41	40.29 ± 1.47	38.79 ± 2.08	35.82 ± 3.19	33.09 ± 1.47	27.83 ± 12.80	37.39 ± 5.16
CO (*mean* ± *SD*)	1.27 ± 0.11	1.14 ± 0.08	0.88 ± 0.07	0.86 ± 0.10	0.91 ± 0.11	0.80 ± 0.07	0.69 ± 0.24	0.94 ± 0.19
O_3 (*mean* ± *SD*)	116.95 ± 6.92	131.28 ± 5.68	121.41 ± 6.90	126.72 ± 4.84	126.14 ± 4.61	121.14 ± 4.57	121.16 ± 44.12	123.54 ± 4.41
AT (*mean* ± *SD*)	14.95 ± 10.16	15.00 ± 9.74	14.82 ± 10.39	14.88 ± 9.50	14.68 ± 9.20	15.03 ± 9.34	16.05 ± 10.49	15.00 ± 9.80
AH (*mean* ± *SD*)	67.40 ± 15.55	63.74 ± 18.00	64.55 ± 17.95	63.28 ± 17.00	66.36 ± 17.77	66.14 ± 18.52	60.78 ± 19.04	64.82 ± 17.73
AWS (*mean* ± *SD*)	2.05 ± 0.96	2.86 ± 1.05	3.00 ± 1.10	2.86 ± 1.09	2.75 ± 0.98	2.31 ± 1.13	2.50 ± 0.87	2.62 ± 1.09

P pneumonia (J12–J18), *AURI* acute upper respiratory infections (J00–J06), *CLRD* chronic lower respiratory diseases (J40–J47), *AT* average temperature, AH average humidity, *AWS* average wind speed

Model performance evaluation indicators

Cross-validation was carried out to extract more useful information from limited data, prevent the algorithm from becoming trapped around local minima, and avoid the overfitting problem (Singh and Agarwal 2023; Lahmiri and Bekiros 2019; Vakharia and Gujar 2019). To evaluate the performance and effect of the SVR model, ten-fold cross-validation was conducted to evaluate the SVR model, considering the data volume and computational cost. Furthermore, the mean absolute error (*MAE*), mean squared error (*MSE*), and coefficient of determination (R^2) were used to

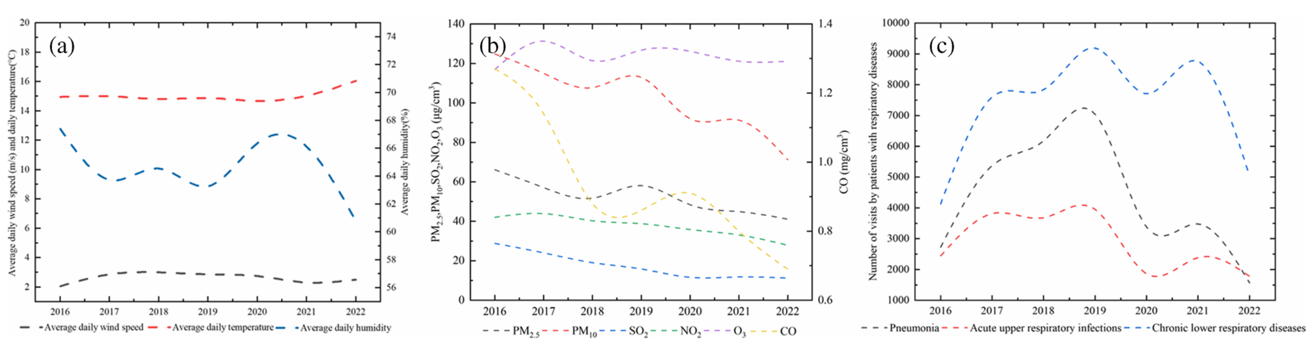


Fig. 3 Trends of **a** meteorological parameters, **b** air pollutants, and **c** number of respiratory-related hospital visits during the study period

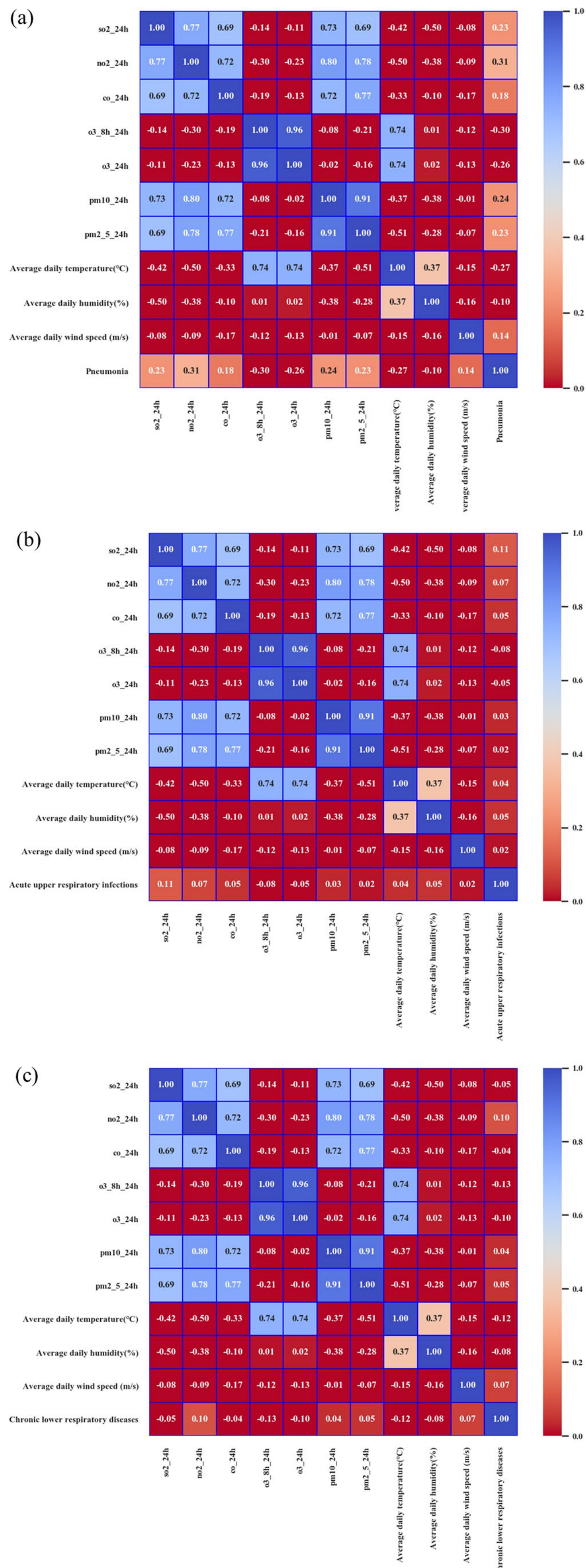


Fig. 4 Spearman correlation results: **a** pneumonia, **b** acute upper respiratory infections, and **c** chronic lower respiratory diseases

evaluate the prediction accuracy of the SVR model. These metrics were calculated as follows:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|\hat{y}_i - y_i\right|, \quad (1)$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(\hat{y}_i - y_i\right)^2, \quad (2)$$

$$R^2 = 1 - \frac{\sum_{i=0}^{n}\left(\hat{y}_i - y_i\right)^2}{\sum_{i=0}^{n}\left(y_i - \bar{y}_i\right)^2} \quad (3)$$

where i is the i th day; y_i represents the number of respiratory patients per day; $\bar{y}_i$ represents the mean value of the number of patients per day; $\hat{y}_i$ represents the predicted value, and n is the number of samples in the test set.

The *MSE* is an indicator of the model stability; the *MAE* measures the error arising from the predicted and observed values, and R^2 measures how well the model fits the data. Smaller *MSE* and *MAE* values denote higher model stability and accuracy, respectively, and an R^2 value close to 1 indicates accurate predictions (Zhang et al. 2022a; He et al. 2021).

Results

Data characteristics

In total, 99,894 patients diagnosed with a respiratory disease were admitted to Linyi People's Hospital from January 1, 2016 to August 20, 2022 (Table 1). Among these patients, 9295 (9.31%) visited the hospital in 2016; 16,782 (16.80%) visited in 2017; 17,693 (17.71%) visited in 2018; 20,176 (20.20%) visited in 2019; 12,946 (12.96%) visited in 2020; 14,618 (14.63%) visited in 2021; and 8384 (8.39%) visited in 2022. The largest proportion of hospital visits for respiratory diseases occurred in 2019. The three main respiratory diseases over the study period were chronic lower respiratory diseases (45.53%), pneumonia (34.89%), and acute upper respiratory infections (19.58%). The number of patients attending Linyi People's Hospital with one of these three respiratory diseases showed the same trend in each year. In 2019, the proportion of male patients (12,651, 62.70%) was higher than that of female patients, with an average male patient age of 37.54 ± 32.69 years. Most respiratory disease patients were aged from 0 to 15 years (45.00%), with those aged from 16 to 60 years, accounting for 15.65%, and those aged 61 years and above accounting for 39.35%.

Regarding the air pollutants, the concentrations of PM_{10}, $PM_{2.5}$, NO_2, and SO_2 showed decreasing trends throughout the study period. The interannual variation of the CO concentration was large, while that of O_3 was small. The interannual changes in outpatient visits for the three abovementioned respiratory diseases were similar to one another. The average daily temperature and average daily wind speed fluctuated around the mean level, with little interannual change, while the average daily humidity fluctuated considerably during the study period (Fig. 3).

Correlation analysis results

Figure 4 shows the results of Spearman correlation analysis between respiratory disease occurrence and meteorological and air pollution factors. Regarding the correlations between air quality factors and the number of hospital visits for respiratory diseases, the incidence of pneumonia is most strongly correlated with $PM_{2.5}$, (0.23) and PM_{10} (0.24). In addition, the $PM_{2.5}$ and PM_{10} concentrations are highly correlated (0.91). The SO_2 concentration is most strongly correlated with pneumonia (0.23) and acute upper respiratory infections (0.11). Similarly, the CO concentration is most strongly correlated with pneumonia (0.18) and acute upper respiratory infections (0.05). NO_2 is positively correlated with pneumonia (0.31), acute upper respiratory infections (0.07), and chronic lower respiratory diseases (0.10). Regarding the correlation between meteorological factors and respiratory diseases, the daily average temperature and daily average relative humidity are negatively correlated with pneumonia (−0.27 and −0.10, respectively). These two meteorological factors are also negatively correlated with chronic lower respiratory diseases. In contrast, the daily average temperature and daily average relative humidity are positively correlated with acute upper respiratory infections (0.04 and 0.05, respectively). The wind speed is positively correlated with the number of patients with either chronic lower respiratory disease, pneumonia, or acute upper respiratory infection, with correlation coefficients ranging from 0.02 to 0.14.

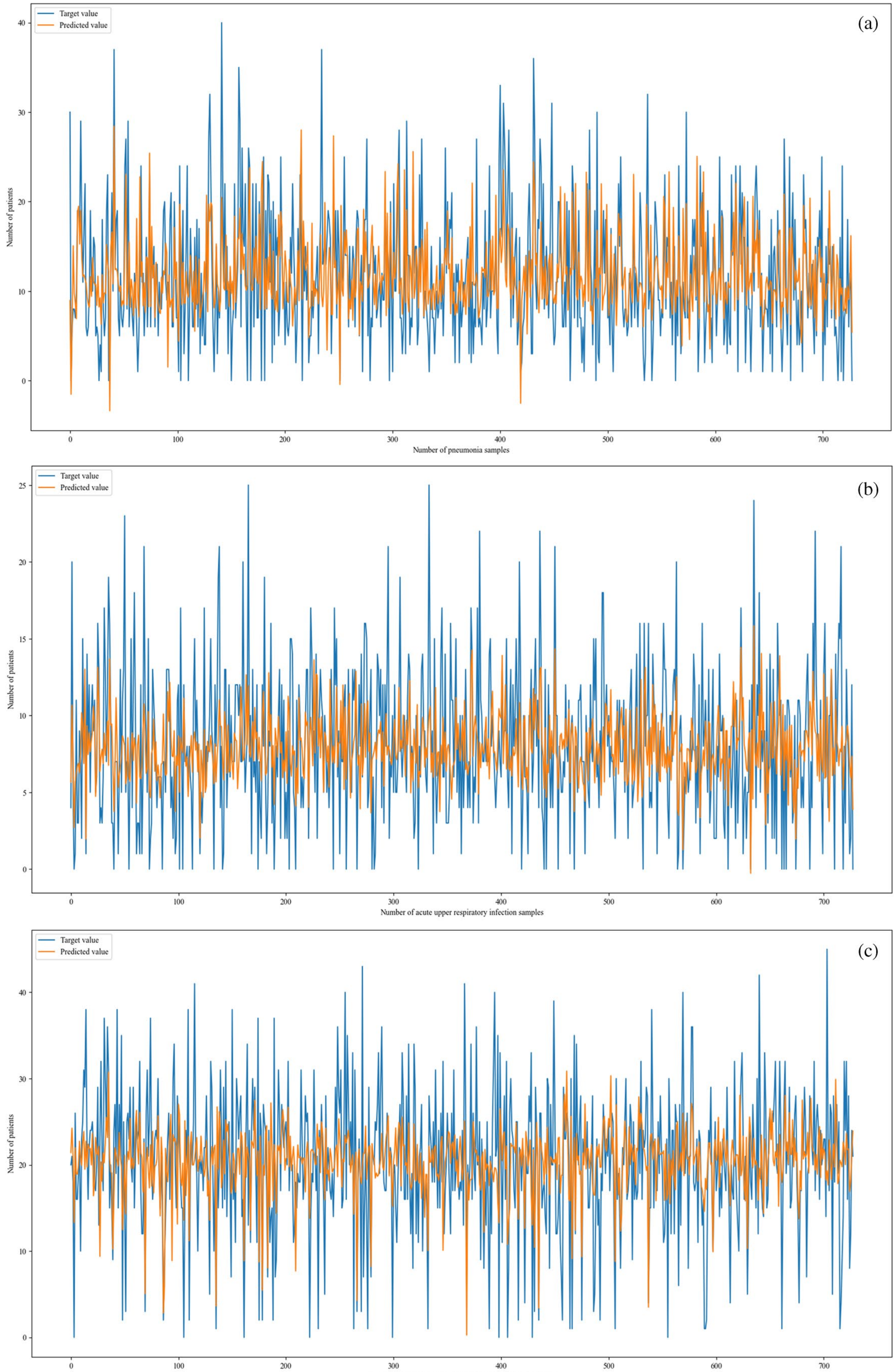
SVR model results and model evaluation

For this analysis, we used the number of patient hospital visits for respiratory disease as the regression target. Seventy percent of the data was used for training, and 30% was used for testing. Table 2 summarizes the outcomes of the SVR model

Table 2 SVR performance indicators for predicting respiratory-related hospital visits

Disease type	R^2 value	*MAE*	*MSE*
Pneumonia	0.308	0.655	0.740
Acute upper respiratory infections	0.155	0.720	0.828
Chronic lower respiratory diseases	0.185	0.711	0.832

MAE mean absolute error, *MSE* mean squared error, R^2 coefficient of determination


Target value
Predicted value
(a)
Number of patients
Number of pneumonia samples
(b)
Number of acute upper respiratory infection samples
(c)
Number of chronic lower respiratory tract infection samples

◀**Fig. 5** Scatterplots of the test values and predicted values: **a** pneumonia, **b** acute upper respiratory infections, and **c** chronic lower respiratory diseases

in predicting the number of visits for respiratory diseases. Three indicators were used to characterize the model performance: *MAE*, *MSE*, and R^2. The *MAE* ranges from 0.655 to 0.720. The *MSE* index values for pneumonia, acute upper respiratory infections, and chronic lower respiratory diseases are 0.740, 0.828, and 0.832, respectively. These results demonstrate that the SVR-based method has the potential to forecast the number of hospital visits for respiratory disease.

Figure 5 shows the number of hospital visits in the test sample estimated by the SVR model. The *x*-axis shows the sample number, and the *y*-axis shows the number of hospital visits for respiratory disease. The coincidence degree of the test value and predicted value reflects the model's performance. Among the three respiratory diseases, the SVR model produces the best prediction effect for pneumonia, which is consistent with the results in Table 2.

Discussion

Accelerated urbanization and rapid industrialization have led to air pollutants being emitted in increased quantities. As a result, the world's population is being exposed to higher concentrations of these pollutants (Cohen et al. 2017). This situation damages public health by exacerbating respiratory diseases. In Linyi, the incidence of respiratory system diseases has been increasing in recent years (Li et al. 2020). In this study, we used correlation analysis and SVR modeling to explore the relationships between hospital visits for respiratory diseases and meteorological and air pollution factors in Linyi and assessed the potential of the SVR model to predict the number of respiratory disease outpatient visits. Furthermore, we attempted to elucidate the effects of meteorological and air pollution factors on three types of respiratory diseases to provide a reference for relevant policymakers.

Regarding the effect of air pollutants on the occurrence of pneumonia, acute upper respiratory infections, and chronic lower respiratory diseases, Spearman correlation analysis showed that NO_2 had a more significant effect on pneumonia (0.31) than on acute upper respiratory infections (0.07) and chronic lower respiratory diseases (0.10). An important source of NO_2 emissions is the combustion process in vehicles and heating systems (Matthaios et al. 2019). The annual average concentrations of NO_2 ranged from 27.83 to 43.88 $\mu g \cdot m^{-3}$ (Table 1). The NO_2 concentrations trended downward from 2020 to 2022, possibly because of restrictions on transportation and industrial activities during the COVID-19 pandemic. Studies have shown that increased NO_2 concentrations in ambient air are associated with an increase in the number of hospital visits for respiratory diseases (Soleimani et al. 2019; Liu et al. 2016). Related to this, our results showed that the concentrations of $PM_{2.5}$ and PM_{10} were positively associated with the occurrence of pneumonia, acute upper respiratory infections, and chronic lower respiratory diseases. The strongest associations were for pneumonia, with correlation coefficients of 0.23 ($PM_{2.5}$) and 0.24 (PM_{10}). Accumulating evidence has demonstrated a close link between respiratory diseases and PM concentrations. Liu et al. (2019) found that the concentrations of ambient $PM_{2.5}$ and PM_{10} were strongly correlated with daily mortality from respiratory diseases in 652 cities worldwide. Croft et al. (2019) reported that the incidence of pneumonia and influenza was closely linked with the $PM_{2.5}$ concentration in the previous weeks. One study found that $PM_{2.5}$ has a greater adverse effect on human health than PM_{10}. This was mainly attributed to the fact that $PM_{2.5}$, along with other harmful substances, can enter the respiratory system faster than PM_{10}, and can thus cause respiratory diseases to appear sooner, especially pneumonia (Tao et al. 2014). In our study, the $PM_{2.5}$ and PM_{10} concentrations were strongly correlated (0.91) and had similar effects on the number of patients with the three main types of respiratory diseases. Therefore, we speculate that these two pollutants have similar effects on the mechanism of respiratory disease occurrence. We also found that hospital visits for pneumonia and acute upper respiratory infections increased with increasing SO_2 and CO concentrations. A previous study reported that increases in SO_2 and CO led to the occurrence or aggravation of respiratory diseases (Su et al. 2019). SO_2 is primarily released and enters the atmosphere, by the burning of fossil fuels (van der A et al. 2017), while CO mainly comes from residential and transportation sources (Lian et al. 2020). The inhalation of SO_2 and CO can damage lung tissue and cause acute neutrophilic pulmonary inflammation (Wigenstam et al. 2016), such as pneumonia and acute upper respiratory infections.

Meteorological factors such as the daily average temperature, humidity, and wind speed can affect respiratory system diseases by altering the respiratory defense system (Ciencewicki and Jaspers 2007). Our results showed that a higher daily average temperature and humidity reduced hospital visits for pneumonia and chronic lower respiratory diseases. Studies have shown that temperature and humidity can affect the survival and proliferation of pathogens that cause respiratory tract infections, thereby affecting the incidence of such infections (Shima et al. 2016). Lower temperatures facilitate respiratory viral infections, mainly because the lipid envelope of viruses resists degradation at low temperatures. This can stabilize viruses (such as respiratory syncytial virus) and allow them to be transmitted more easily in secretions, thus increasing the incidence of respiratory disease (Vandini et al. 2015). In addition, it has been reported that respiratory infections are more likely to spread in low-humidity environments. In a low-humidity environment, smaller pathogenic droplets expelled during coughing or sneezing remain airborne for longer, thus contributing to their spread (Stiti et al. 2022). Furthermore, the ability of respiratory epithelial cells to reject virus particles decreases with decreasing humidity, which increases the possibility of respiratory tract infection (Moriyama et al. 2020). The ability of climate factors to influence the occurrence of respiratory diseases varies by species. Research on the link between wind speed and respiratory disease is limited. We speculate that wind speed mainly affects respiratory diseases by influencing the concentration of air pollutants.

To evaluate the model, the R^2 value between the tested and predicted values, the *MAE*, and the *MSE* were examined. The SVR model presented the lowest MAE (0.655) and *MSE* (0.740) values and the highest R^2 (0.308) value for pneumonia outpatient visits. To better assess the potential for predicting daily hospital visits for respiratory diseases, our findings and different reported machine learning methods for patients with respiratory diseases were compared (Table S1). Other studies using machine learning methods have shown the potential to predict respiratory diseases using air quality factors (Ku et al. 2022; Alvarez-Mendoza et al. 2020). Similarly, the SVR method developed in our study showed the potential to forecast the number of respiratory disease-related hospital visits, especially for pneumonia.

Conclusion

Some air pollutants and meteorological conditions have been identified as key risk factors to public health. Revealing the relationships between these pollutants, meteorological conditions, and hospital visits for respiratory disease provides an important warning to the relevant authorities about the effects of air pollution. Using a combination of Spearman correlation analysis and SVR modeling, this study investigated the associations between meteorological factors, air pollution, and hospital visits for respiratory diseases, including pneumonia, acute upper respiratory infections, and chronic lower respiratory diseases. The SVR model considered the number of patients with respiratory diseases, air pollutant concentrations, and meteorological information from January 1, 2016 to August 20, 2022. Model evaluations were conducted using ten-fold cross-validation, and the data were randomly divided into training and testing sets at a 7:3 ratio. Moreover, we assessed the *MAE*, *MSE*, and R^2 to evaluate the performance of the regression models. The results showed that the NO_2, $PM_{2.5}$, and PM_{10} concentrations were positively correlated with hospital visits for the three investigated respiratory diseases, with the strongest correlation appearing for pneumonia. In addition, increases in the SO_2 and CO concentrations enhanced the occurrence of pneumonia and acute upper respiratory infections. Regarding meteorological factors, an increase in the daily average temperature and humidity decreased the occurrence of pneumonia and chronic lower respiratory diseases. Assessment of the SVR models showed that the best prediction effect was for pneumonia (*MAE* and *MSE* of 0.655 and 0.740, respectively). Here, we attempted to find associations between climate and air pollution factors and three main types of respiratory diseases. Future studies could include other variables that may contribute to the occurrence of these diseases, such as formaldehyde and nitrogen oxide concentrations. Additionally, the complex mechanisms by which meteorological factors and air pollutants influence respiratory diseases require further study.

Study limitations

This study had several limitations. First, the outpatient volume data were from Linyi Central Hospital. The total outpatient volume of this hospital accounted for more than 80% of the total number of outpatients throughout the city, and this sample may reasonably reflect the respiratory disease situation for the whole urban area of Linyi (Li et al. 2020); nevertheless, the true situation could have been different. Second, nine meteorological and air pollutant variables were analyzed (daily mean wind speed, temperature, humidity, CO, O_3, SO_2, NO_2, $PM_{2.5}$, and PM_{10}), as these variables are associated with worsening respiratory disease. However, other factors also contribute to the occurrence of respiratory diseases, such as formaldehyde and nitrogen oxide exposure (Neamtiu et al. 2019; Saki et al. 2020). Third, there may have been some error in the measured average concentrations of the air pollutants monitored by the four state-controlled monitoring stations in Linyi. Furthermore, while our study found associations between respiratory diseases and meteorological and air pollution factors, we could not accurately calculate the risk posed by these air pollutants, and the complex mechanisms by which meteorological and air pollution factors affect respiratory disease require further study. Additionally, other factors affecting the number of hospital visits for respiratory diseases, such as medical insurance policies and the reimbursement ratio, were not considered in this study and need further exploration in the future. From the perspective of the SVR model, the ratio of the training and testing data may not have been ideal and could have affected the results. Future research should consider different cutoff points in the dataset. In addition, the time lag effect between meteorological and air quality factors and the number of patients with respiratory diseases needs to be considered in future studies.

Despite these limitations, our results demonstrated clear relationships between meteorological and air pollution factors and respiratory diseases, providing support for government health agencies to enact prevention and control measures. In addition, our study demonstrated the potential of applying machine learning to predict the occurrence of disease and provides a methodological reference for future studies of the possible relationships between other air pollutant concentrations and epidemiological data.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11356-023-28682-8.

Author contribution Jing Yang: methodology, investigation, data curation, and writing—original draft. Xin Xu: methodology, investigation, data curation, and writing—original draft. Xiaotian Ma: investigation and methodology. Zhaotong Wang: data curation. Qian You: writing—review and editing. Wanyue Shan: investigation. Ying Yang: software. Xin Bo: conceptualization and writing–review and editing. Chuansheng Yin: conceptualization, methodology, and writing—review and editing. All the authors have read and agreed to the published version of the manuscript.

Funding This research was supported by the Fundamental Research Funds for the Central Universities (Grant No. buctrc202133), the National Natural Science Foundation of China (Grant No. 72174125), and the Environmental Health Risk Monitoring and Response Based on Big Data (Grant No. H2022006).

Data availability The authors do not have permission to share data.

Declarations

Ethics approval Not applicable.

Consent to participate All the authors declare that they consented to participate in this study.

Consent for publication All the authors declare that they consented to publish this study.

Competing interests The authors declare no competing interests.

References

Al-Kindi SG, Brook RD, Biswal S, Rajagopalan S (2020) Environmental determinants of cardiovascular disease: lessons learned from air pollution. Nat Rev Cardiolol 17(10):656–672. https://doi.org/10.1038/s41569-020-0371-2

Almetwally AA, Bin-Jumah M, Allam AA (2020) Ambient air pollution and its influence on human health and welfare: an overview. Environ Sci Pollut Res 27:24815–24830. https://doi.org/10.1007/s11356-020-09042-2

Alvarez-Mendoza CI, Teodoro A, Freitas A, Fonseca J (2020) Spatial estimation of chronic respiratory diseases based on machine learning procedures—an approach using remote sensing data and environmental variables in quito, Ucuador. Appl Geogr 123:102273. https://doi.org/10.1016/j.apgeog.2020.102273

Appaia L, Palraj S (2023) On replacement of outliers and missing values in time series. Int J Environ Qual 53:1–10. https://doi.org/10.6092/issn.2281-4485/16184

Araste Z, Sadighi A, Jamimoghaddam M (2022) Fault diagnosis of a centrifugal pump using electrical signature analysis and support vector machine. J Vib Eng Technol 2022:1–11. https://doi.org/10.1007/s42417-022-00687-6

Bai L, Wang J, Ma X, Lu H (2018) Air pollution forecasts: an overview. Int J Environ Res Pub Health 15(4):780. https://doi.org/10.3390/ijerph15040780

Blomgren J, Virta LJ (2020) Socioeconomic differences in use of public, occupational and private health care: a register-linkage study of a working-age population in Finland. PLoS One 15(4):e0231792. https://doi.org/10.1371/journal.pone

Bodaghkhani E, Mahdavian M, MacLellan C, Farrell A, Asghari S (2019) Effects of meteorological factors on hospitalizations in adult patients with asthma: a systematic review. Can Respir J 2019:3435103. https://doi.org/10.1155/2019/3435103

Bureau TARS (2021) National Bureau of Statistics Tibet Survey Corps. Tibet Statistical Yearbook. China Statistics Press, Beijing; 2020 (In Chinese)

Choi H, Myong JP (2018) Association between air pollution in the 2015 winter in South Korea and population size, car emissions, industrial activity, and fossil-fuel power plants: an ecological study. Ann Occup Environ Med 30(1):1–7. https://doi.org/10.1186/s40557-018-0273-5

Choudhury A, Gupta D (2019) A survey on medical diagnosis of diabetes using machine learning techniques, In Recent developments in machine learning and data analytics. Springer, pp 67–78. https://doi.org/10.1007/978-981-13-1280-9_6

Ciencewicki J, Jaspers I (2007) Air pollution and respiratory viral infection. Inhal Toxicol 19(14):1135–1146. https://doi.org/10.1080/08958370701665434

Cohen AJ, Brauer M, Burnett R, Anderson HR, Frostad J, Estep K, Balakrishnan K, Brunekreef B, Dandona L, Dandona R (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. Lancet 389(10082):1907–1918. https://doi.org/10.1016/S0140-6736(17)30505-6

Costello A, Abbas M, Allen A, Ball S, Bell S, Bellamy R, Friel S, Groce N, Johnson A, Kett M (2009) Managing the health effects of climate change: lancet and University College London Institute for Global Health Commission. Lancet 373(9676):1693–1733. https://doi.org/10.1016/S0140-6736(09)60935-1

Croft DP, Zhang W, Lin S, Thurston SW, Hopke PK, Masiol M, Squizzato S, van Wijngaarden E, Utell MJ, Rich DQ (2019) The association between respiratory infection and air pollution in the setting of air quality policy and economic change. Ann Am Thorac Soc 16(3):321–330. https://doi.org/10.1513/AnnalsATS.201810-691OC

Darrow LA, Klein M, Flanders WD, Mulholland JA, Tolbert PE, Strickland MJ (2014) Air pollution and acute respiratory infections among children 0–4 years of age: an 18-year time-series study. Am J Epidemiol 180(10):968–977. https://doi.org/10.1093/aje/kwu234

Delavar MR, Gholami A, Shiran GR, Rashidi Y, Nakhaeizadeh GR, Fedra K, Hatefi Afshar S (2019) A novel method for improving air pollution prediction based on machine learning approaches: a case study applied to the capital city of Tehran. ISPRS Int J Geo-Inf 8(2):99. https://doi.org/10.3390/ijgi8020099

Falcon-Rodriguez CI, Osornio-Vargas AR, Sada-Ovalle I, Segura-Medina P (2016) Aeroparticles, composition, and lung diseases. Front Immunol 7:3. https://doi.org/10.3389/fimmu.2016.00003

Gómez-Carracedo MP, Andrade JM, López-Mahía P, Muniategui S, Prada D (2014) A practical comparison of single and multiple imputation methods to handle complex missing data in air quality datasets. Chemometr Intell Lab 134:23–33. https://doi.org/10.1016/j.chemolab.2014.02.007

Grigorieva E, Lukyanets A (2021) Combined effect of hot weather and outdoor air pollution on respiratory health: literature review. Atmosphere 12(6):790. https://doi.org/10.3390/atmos12060790

Hao Y, Xu T, Hu H, Wang P, Bai Y (2020) Prediction and analysis of corona virus disease 2019. PloS One 15(10):e0239960. https://doi.org/10.1371/journal.pone.0239960

He Z, Yang M, Wang L, Bao E, Zhang H (2021) Concentrated photovoltaic thermoelectric hybrid system: an experimental and machine learning study. Eng Sci 15:47–56. https://doi.org/10.30919/es8d440

Jacquemin B, Siroux V, Sanchez M, Carsin AE, Schikowski T, Adam M, Bellisario V, Buschka A, Bono R, Brunekreef B, Cai Y, Cirach M, Clavel-Chapelon F, Declercq C, de Marco R, de Nazelle A, Ducret-Stich Regina E, Ferretti Virginia V, Gerbase Margaret W, Hardy R, Heinrich J, Janson C, Jarvis D, Al Kanaani Z, Keidel D, Kuh D, Le Moual N, Nieuwenhuijsen Mark J, Marcon A, Modig L, Pin I, Rochat T, Schindler C, Sugiri D, Stempfelet M, Temam S, Tsai MY, Varraso R, Vienneau D, Vierkötter A, Hansell Anna L, Krämer U, Probst-Hensch Nicole M, Sunyer J, Künzli N, Kauffmann F (2015) Ambient air pollution and adult asthma incidence in six European cohorts (ESCAPE). Environ Health Perspect 123(6):613–621. https://doi.org/10.1289/ehp.1408206

Javorac J, Jevtić M, Živanović D, Ilić M, Bijelović S, Dragić N (2021) What are the effects of meteorological factors on exacerbations of chronic obstructive pulmonary disease? Atmosphere 12(4):442. https://doi.org/10.3390/atmos12040442

Ji C, Li L (2016) Research on sponge city construction in Lin Yi city. In International conference on Education, Management, Computer and Society. Atlantis Press, pp 1818–1821. https://doi.org/10.2991/emcs-16.2016.457

Jiang S, Yu H, Li Z, Geng B, Li T (2022) Study on the evolution of the spatial-temporal pattern and the influencing mechanism of the green development level of the Shandong Peninsula urban agglomeration. Sustainability 14(15):9549. https://doi.org/10.3390/su14159549

Kan H, Chen R, Tong S (2012) Ambient air pollution, climate change, and population health in China. Environ Int 42:10–19. https://doi.org/10.1016/j.envint.2011.03.003

Keshavarz Z, Mostofinejad D (2019) Porcelain and red ceramic wastes used as replacements for coarse aggregate in concrete. Constr Build Mater 195:218–230. https://doi.org/10.1016/j.conbuildmat.2018.11.033

Khan MA, Abbas K, Su'ud MM, Salameh AA, Alam MM, Aman N, Mehreen M, Jan A, Hashim NAABN, Aziz RC (2022) Application of machine learning algorithms for sustainable business management based on macro-economic data: supervised learning techniques approach. Sustainability 14(16):9964. https://doi.org/10.3390/su14169964

Kloog I, Nordio F, Zanobetti A, Coull BA, Koutrakis P, Schwartz JD (2014) Short term effects of particle exposure on hospital admissions in the Mid-Atlantic states: a population estimate. PloS One 9(2):e88578. https://doi.org/10.1371/journal.pone.0088578

Ku Y, Kwon SB, Yoon JH, Mun SK, Chang M (2022) Machine learning models for predicting the occurrence of respiratory diseases using climatic and air-pollution factors. Clin Exp Otorhinolaryngol 15(2):168–176. https://doi.org/10.21053/ceo.2021.01536

Lahmiri S, Bekiros S (2019) Can machine learning approaches predict corporate bankruptcy? Evidence from a qualitative experimental design. Quant Financ 19(9):1569–1577. https://doi.org/10.1080/14697688.2019.1588468

Li H, Wu P, Bo X, Wu X, Yang J, Liu H, Wei M, Hao P (2020) Pollution characterization of major air pollutants and their impacts on resident health in Linyi City. Acta Sci Circumst 40(8):2919–2934. https://doi.org/10.13671/j.hjkxxb.2020.0091

Li L, Rong S, Wang R, Yu S (2021) Recent advances in artificial intelligence and machine learning for nonlinear relationship analysis and process control in drinking water treatment: a review. Chem Eng J 405:126673. https://doi.org/10.1016/j.cej.2020.126673

Lian X, Huang J, Huang R, Liu C, Wang L, Zhang T (2020) Impact of city lockdown on the air quality of COVID-19-hit of Wuhan city. Sci Total Environ 742:140556. https://doi.org/10.1016/j.scitotenv.2020.140556

Lippmann M, Maynard RL (1999) Air quality guidelines and standards 1. Air Pollut Health, pp 983–1017. https://doi.org/10.1016/B978-012352335-8/50118-6

Liu P, Wang X, Fan J, Xiao W, Wang Y (2016) Effects of air pollution on hospital emergency room visits for respiratory diseases: urban-suburban differences in eastern China. Int J Environ Res Pub Health 13(3):341. https://doi.org/10.3390/ijerph13030341

Liu C, Chen R, Sera F, Vicedo-Cabrera AM, Guo Y, Tong S, Coelho MS, Saldiva PH, Lavigne E, Matus PM (2019) Ambient particulate air pollution and daily mortality in 652 cities. New Engl J Med 381(8):705–715. https://doi.org/10.1056/NEJMoa1817364

Lu F, Xu D, Cheng Y, Dong S, Guo C, Jiang X, Zheng X (2015) Systematic review and meta-analysis of the adverse health effects of ambient PM2.5 and PM10 pollution in the Chinese population. Environ Res 136:196–204. https://doi.org/10.1016/j.envres.2014.06.029

Mai H, Le TC, Chen D, Winkler DA, Caruso RA (2022) Machine learning for electrocatalyst and photocatalyst design and discovery. Chem Rev 122(16):13478–13515. https://doi.org/10.1021/acs.chemrev.2c00061

Margiana RH, Yousefi A, Afra A, Agustinus WK, Abdelbasset M, Kuznetsova S, Mansourimoghadam EHA, Mohammadi MJ (2022) The effect of toxic air pollutants on fertility men and women, fetus and birth rate. Rev Environ Health. https://doi.org/10.1515/reveh-2022-0032

Masmoudi S, Elghazel H, Taieb D, Yazar O, Kallel A (2020) A machine-learning framework for predicting multiple air pollutants' concentrations via multi-target regression and feature selection. Sci Total Environ 715:136991. https://doi.org/10.1016/j.scitotenv.2020.136991

Masood A, Ahmad K (2021) A review on emerging artificial intelligence (AI) techniques for air pollution forecasting: fundamentals, application and performance. J Clean Prod 322:129072. https://doi.org/10.1016/j.jclepro.2021.129072

Matthaios VN, Kramer LJ, Sommariva R, Pope FD, Bloss WJ (2019) Investigation of vehicle cold start primary NO2 emissions inferred from ambient monitoring data in the UK and their implications for urban air quality. Atmos Environ 199:402–414. https://doi.org/10.1016/j.atmosenv.2018.11.031

Moriyama M, Hugentobler WJ, Iwasaki A (2020) Seasonality of respiratory viral infections. Annu Rev Virol 7(1), annurev-virology-012420–022445.https://doi.org/10.1146/annurev-virology-012420-022445

Neamtiu IA, Lin S, Chen M, Roba C, Csobod E, Gurzau ES (2019) Assessment of formaldehyde levels in relation to respiratory and allergic symptoms in children from Alba County schools, Romania. Environ Monit Assess 191(9):591. https://doi.org/10.1007/s10661-019-7768-6

Nilashi M, Ahmadi H, Shahmoradi L, Ibrahim O, Akbari E (2019) A predictive method for hepatitis disease diagnosis using ensembles of neuro-fuzzy technique. J Infect Public Health 12(1):13–20. https://doi.org/10.1016/j.jiph.2018.09.009

Otchere DA, Arbi Ganat TO, Gholami R, Ridha S (2021) Application of supervised machine learning paradigms in the prediction of petroleum reservoir properties: comparative analysis of ANN and SVM models. J Petrol Sci Eng 200:108182. https://doi.org/10.1016/j.petrol.2020.108182

Pothirat C, Chaiwong W, Liwsrisakun C, Bumroongkit C, Deesomchok A, Theerakittikul T, Limsukon A, Tajarernmuang P, Phetsuk N (2019) Acute effects of air pollutants on daily mortality and hospitalizations due to cardiovascular and respiratory diseases. J Thorac Dis 11(7):3070–3083. https://doi.org/10.21037/jtd.2019.07.37

Qi L, Liu T, Gao Y, Tian D, Tang W, Li Q, Feng L, Liu Q (2021) Effect of meteorological factors on the activity of influenza in Chongqing, China, 2012–2019. Plos One 16(2):e0246023. https://doi.org/10.1371/journal.pone.0246023

Ravindra K, Bahadur SS, Katoch V, Bhardwaj S, Kaur-Sidhu M, Gupta M, Mor S (2023) Application of machine learning approaches to predict the impact of ambient air pollution on outpatient visits for acute respiratory infections. Sci Total Environ 858:159509. https://doi.org/10.1016/j.scitotenv.2022.159509

Razavi-Termeh SV, Sadeghi-Niaraki A, Choi SM (2022) Spatio-temporal modelling of asthma-prone areas using a machine learning optimized with metaheuristic algorithms. Geocarto Int 37(25):9917–9942. https://doi.org/10.1080/10106049.2022.2028903

Requia WJ, Amini H, Mukherjee R, Gold DR, Schwartz JD (2021) Health impacts of wildfire-related air pollution in Brazil: a nationwide study of more than 2 million hospital admissions between 2008 and 2018. Nat Commun 12(1):1–9. https://doi.org/10.1038/s41467-021-26822-7

Saki H, Goudarzi G, Jalali S, Barzegar G, Farhadi M, Parseh I, Geravandi S, Salmanzadeh S, Yousefi F, Mohammadi MJ (2020) Study of relationship between nitrogen dioxide and chronic obstructive pulmonary disease in Bushehr, Iran. Clin Epidemiol Glob 8(2):446–449. https://doi.org/10.1016/j.cegh.2019.10.006

Scott N, Ólafsson S, Gottfreðsson M, Tyrfingsson T, Rúnarsdóttir V, Hansdottir I, Hernandez UB, Sigmundsdóttir G, Hellard M (2018) Modelling the elimination of hepatitis C as a public health threat in Iceland: a goal attainable by 2020. J Hepatol 68(5):932–939. https://doi.org/10.1016/j.jhep.2017.12.013

Semenza JC, Paz S (2021) Climate change and infectious disease in Europe: impact, projection and adaptation. Lancet Reg Health-Eu 9:100230. https://doi.org/10.1016/j.lanepe.2021.100230

Shima K, Coopmeiners J, Graspeuntner S, Dalhoff K, Rupp J (2016) Impact of micro-environmental changes on respiratory tract infections with intracellular bacteria. FEBS Lett 590(21):3887–3904. https://doi.org/10.1002/1873-3468.12353

Singh R, Agarwal BB (2023) An automated brain tumor classification in MR images using an enhanced convolutional neural network. Int J Inf Tecnol 15:665–674. https://doi.org/10.1007/s41870-022-01095-5

Soleimani Z, Boloorani AD, Khalifeh R, Teymouri P, Mesdaghinia A, Griffin DW (2019) Air pollution and respiratory hospital admissions in Shiraz, Iran, 2009 to 2015. Atmos Environ 209:233–239. https://doi.org/10.1016/j.atmosenv.2019.04.030

Stiti M, Castanet G, Corber A, Alden M, Berrocal E (2022) Transition from saliva droplets to solid aerosols in the context of COVID-19 spreading. Environ Res 204:112072. https://doi.org/10.1016/j.envres.2021.112072

Su W, Wu X, Geng X, Zhao X, Liu Q, Liu T (2019) The short-term effects of air pollutants on influenza-like illness in Jinan, China. BMC Public Health 19(1):1319. https://doi.org/10.1186/s12889-019-7607-2

Tao Y, Liu Y, Mi S, Guo Y (2014) Atmospheric pollution characteristics of fine particles and their effects on human health. Acta Sci Circumst 34(3):592–597. https://doi.org/10.13671/j.hjkxxb.2014.0107

Thirumalai C, Chandhini SA, Vaishnavi M (2017) Analysing the concrete compressive strength using Pearson and Spearman. 2017 International conference of electronics, communication and aerospace technology (iCECA) (Vol. 2, pp. 215–218). IEEE. https://doi.org/10.1109/ICECA.2017.8212799

Vakharia V, Gujar R (2019) Prediction of compressive strength and portland cement composition using cross-validation and feature ranking techniques. Constr Build Mater 225:292–301. https://doi.org/10.1016/j.conbuildmat.2019.07.224

van der A RJ, Mijling B, Ding J, Koukouli ME, Liu F, Li Q, Mao H, Theys N (2017) Cleaning up the air: effectiveness of air quality policy for SO2 and NOx emissions in China. Atmos Chem Phys Chem Glasses 17(3):1775–1789

Vandini S, Bottau P, Faldella G, Lanari M (2015) Immunological, viral, environmental, and individual factors modulating lung immune response to respiratory syncytial virus. BioMed Res Int 2015. https://doi.org/10.1155/2015/875723

WHO (World Health Organization) (2018) World Health Organization releases new global air pollution data. Geneva Switzerland, 2018. https://www.energy-community.org

Wigenstam E, Elfsmark L, Bucht A, Jonasson S (2016) Inhaled sulfur dioxide causes pulmonary and systemic inflammation leading to fibrotic respiratory disease in a rat model of chemical-induced lung injury. Toxicology 368–369:28–36. https://doi.org/10.1016/j.tox.2016.08.018

Wu X, Li D, Feng M, Liu H, Li H, Yang J, Wu P, Lei X, Wei M, Bo X (2021) Effects of air pollutant emission on the prevalence of respiratory and circulatory system diseases in Linyi, China. Environ Geochem Health 43(11):4475–4491. https://doi.org/10.1007/s10653-021-00931-0

Wu X, Wang L, An J, Wang Y, Song H, Wu Y, Liu Q (2022) Relationship between soil organic carbon, soil nutrients, and land use in Linyi City (East China). Sustainability 14(20):13585. https://doi.org/10.3390/su142013585

Yang J, Ma J, Sun Q, Han C, Guo Y, Li M (2022a) Health benefits by attaining the new WHO air quality guideline targets in China: a nationwide analysis. Environ Pollut 308:119694. https://doi.org/10.1016/j.envpol.2022.119694

Yang T, Wang H, Li H, Guo X, Wang D, Chen X, Wang F, Xin J, Sun Y, Wang Z (2022b) Quantitative attribution of wintertime haze in coastal east China to local emission and regional intrusion under a stagnant internal boundary layer. Atmos Environ 276:119006. https://doi.org/10.1016/j.atmosenv.2022.119006

Yi Z, Yu X, Li G, Chen P (2016) Analysis of the present situation and governance measures of haze in Linyi City. J Appl Sci Eng Innov 3(2):52–56

Yin Y, Chen H, Wang G, Xu W, Wang S, Yu W (2021) Characteristics of the precipitation concentration and their relationship with the precipitation structure: a case study in the Huai River basin, China. Atmos Res 253:105484. https://doi.org/10.1016/j.atmosres.2021.105484

Zhang S, Li Y, Hao Y, Zhang Y (2018) Does public opinion affect air quality? Evidence based on the monthly data of 109 prefecture-level cities in China. Energ Policy 116:299–311. https://doi.org/10.1016/j.enpol.2018.02.025

Zhang B, Rong Y, Yong R, Qin D, Li M, Zou G, Pan J (2022a) Deep learning for air pollutant concentration prediction: a review. Atmos Environ 119347. https://doi.org/10.1016/j.atmosenv.2022.119347

Zhang Y, Wen Z, Hu Y, Zhang T (2022b). Waste flow of wet wipes and decision-making mechanism for consumers' discarding behaviors. J Clean Prod 364: 132684.https://doi.org/10.1016/j.jclepro.2022.132684

Zhao J, Wang Y, Wang X (2017) Spatial autocorrelation of urban economic growth in Shandong province, China by using time-series data of Per Capita GDP. Geo-Spatial Knowledge and Intelligence: 5th International Conference, GSKI 2017, Chiang Mai, Thailand, December 8-10, 2017. Revised Selected Papers, Part I 5. Springer, Singapore, 2018, pp 23–31

Zheng Q, Tian X, Yang M, Su H (2019) The email author identification system based on support vector machine (SVM) and analytic hierarchy process (AHP). IAENG. Int J Comput Sci 46(2):178–191

Zhou H, Wang T, Zhou F, Liu Y, Zhao W, Wang X, Chen H, Cui Y (2019) Ambient air pollution and daily hospital admissions for respiratory disease in children in Guiyang, China. Front Pediatr 7:400. https://doi.org/10.3389/fped.2019.00400