scientific reports

OPEN

Assessing the impact of meteorological factors and air pollution on respiratory disease mortality rates: a random forest model analysis (2017–2021)

Yousef Dowlatabadi[1], Shaghayegh abadi[2], Maryam Sarkhosh[3✉], Mitra Mohammadi[2✉] & Seyed Mohammad Mahdi Moezzi[4]

Air pollution poses a significant threat to the health of all living beings on our planet. It has been scientifically established as a crucial factor affecting mortality rates, respiratory illnesses, mental well-being, and overall health. This study aimed to investigate the impact of air pollution and meteorological factors on respiratory disease mortality rates in Mashhad in 2017–2021 using a Random Forest (RF) model. At first, the daily statistics of meteorological parameters (pressure, humidity, temperature, solar radiation) during 2017–2021 were collected. The information related to pollutants pollutants such as $PM_{2.5}$ (which is defined as particulate matter with less than 2.5 micrometer diameter and potentially harmful to humans), PM_{10} (Particles with a diameter of 10 micrometers or less that can negatively impact both human health and environmental conditions.), sulfur dioxide (SO_2), nitrogen dioxide (NO_2), and carbon monoxide (CO) was collected. the mortality statistics from respiratory diseases were collected from the Health system registaration (Sina). Then we used the RF regression model in Excel and Python software to investigate the interaction between the mentioned variables. The escalating trend of air pollution in Mashhad has led to an expected increase in respiratory-related hospitalizations. This necessitates urgent air pollution control measures. Concurrently, the study of pollutants and climatic elements, as substantiated by global epidemiological studies, is crucial. In Mashhad, the second most polluted city in Iran, climatic factors like humidity, sunshine duration, temperature, pressure, and sunlight intensity exacerbate atmospheric pollution levels, impacting human health and ecosystems. The R^2, RSME, and MAE of RF model are 0.73, 2.52, and 2 which indicate that the model has successfully identified the relationship between input variables and the target variable, and it will exhibit high accuracy in predictions. In this study, the most influential factor was identified when the Variance Inflation (VI) factor reached a value of 0.548. This indicates a strong correlation between this factor and the death rate of patients during the specified period. Furthermore, we analyzed by excluding the day and month plans from our model. The results showed that the factor with the highest coefficient in the executive model was related to pressure, with a VI value of 0.049. This suggests that pressure plays a significant role in our model and has a substantial effect on the death rate of patients. In the study of various pollutants, it was found that PM_{10} had the most significant impact on the mortality rate of patients with respiratory conditions, with a VI of 0.039. Following PM_{10}, the pollutants with the next highest coefficients of importance were NO_2 (VI = 0.034), SO_2 (VI = 0.033), PM2.5 (VI = 0.029), and CO (VI = 0.025), respectively. Furthermore, the study observed a notable increase in the mortality rate of respiratory patients over time. Specifically, for every five days, the death rate increased by 35.5%. Indeed, climate change and air pollution significantly contribute to the mortality rate from respiratory diseases. Therefore, it is crucial for individuals, particularly those with respiratory conditions, to heed meteorological warnings.

[1]School of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran. [2]Department of Environmental Science, Kheradgarayan Motahar Institute of Higher Education, Mashhad, Iran. [3]Department of Environmental

Health Engineering, School of Health, Mashhad University of Medical Sciences, Mashhad, Iran. ⁴Department of Civil Engineering, Sharif University of Technology, Tehran, Iran. ✉email: marya.sarkhosh@yahoo.com; mitramohammadi@gmail.com; m.mohammadi@motahar.ac.ir

Air pollution, a pressing issue in modern cities, has numerous harmful effects on humans and the environment. A key area of concern is air pollutants, which have been linked to respiratory diseases and allergies, including pneumonia and Chronic Obstructive Pulmonary Disease (COPD). With industrial development on the rise, human-generated pollutants are also increasing. The World Health Organization (WHO) has indeed recognized air pollution as a significant environmental health risk, leading to various illnesses and fatalities, including cancer, heart disease, and respiratory disorders[1]. Exposure to air pollutants can intensify respiratory disease episodes, such as asthma[2,3]. Urban living and proximity to main streets are risk factors for asthma[3].

To combat this, most countries have established centers to monitor air pollutants, including Particulate Matter (PM), Ozone (O3), Carbon Monoxide (CO), Sulfur Dioxide (SO2), Nitrogen Dioxide (NO2), and Lead (Pb). Weather conditions, which can fluctuate air quality across large city areas, also significantly affect human mortality and disease rates. Chinese research indicates a significant correlation between the risk of influenza and air pollutants, particularly O3, SO2, and NO2, predominantly in the cold season[4]. Research has indeed indicated that long-term exposure to polluted air can contribute to the development of diabetes mellitus, specifically type 2[5]. Exposure to air pollutants such as CO may exacerbate rheumatoid arthritis. Additionally, elevated NO2 levels can adversely impact the quality of life in patients with autoimmune disorders, including systemic lupus erythematosus[6,7]. The study in Tehran indicates an adult asthma prevalence of 11.73%, surpassing the national average of 8.9%. A notable association is observed between asthma and the presence of CO and PM10, potentially implying exacerbation of asthma due to pollution[8].

Respiratory diseases are among the leading causes of morbidity and mortality worldwide. Exposure to air pollutants exacerbates existing respiratory conditions and contributes to premature deaths. Research conducted in China indicated that brief exposure to PM2.5 could potentially elevate the risk of hospital admissions due to cardiorespiratory issues[9]. A 2021 meta-analysis revealed a statistically significant correlation between PM2.5 exposure levels and lung cancer incidence, underscoring the potential carcinogenicity of fine particulate matter[10]. Vulnerable groups, such as children, the elderly, and individuals with pre-existing health conditions, bear the brunt of air pollution's adverse effects. People with conditions like asthma, bronchitis, or cardiovascular diseases face heightened susceptibility to pollution-related health complications. According to the World Health Organization (WHO), over 90% of the global population resides in areas where air quality exceeds the recommended standards. This highlights the urgent need for strategies to improve air quality and reduce the associated health risks[11].

Mashhad, a metropolis grappling with air pollution and a high production of various pollutants, ranks second in air pollution in the country. The increasing trend of air pollution in Mashhad suggests a potential rise in mortality rates. However, with appropriate management programs and accurate environmental climate information, the city's air pollution can be significantly controlled, thereby safeguarding human health. Adherence to the annual average PM2.5 guideline in South Africa is projected to yield substantial public health benefits, potentially averting 28,000 fatalities annually. This reduction constitutes a significant decrease, accounting for over 5% of the nation's overall mortality rate[12]. The 2013 Mashhad study by Bonyadi and team applied the Air Q model to determine the health effects of $PM_{2.5}$, concluding that higher average levels of $PM_{2.5}$ correlated with increased deaths due to respiratory and heart conditions[13]. However, our study's advantage is its extensive four-year scope and the inclusion of various air pollutants.

In the intricate labyrinth of machine learning, one model stands out for its power and popularity - the Random Forest (RF) regression model. The RF model is one of the commonly used models in the field of machine learning for predicting continuous numerical values. This model is composed of several regression decision trees, each constructed using random subsets of the data and features. The final prediction of the model is made based on the average estimate of the trees. This model is well-known for its high efficiency and its ability to mitigate issues related to overfitting and handling large datasets. Additionally, this approach is useful for dealing with noisy data and showcasing the importance of variables[14].

In this study, a comprehensive dataset encompassing climatic variables (humidity, temperature, pressure, and radiation), pollutants (CO, NO2, SO2, PM10, and PM2.5), and mortality rates from respiratory diseases was compiled. The primary aim of this study was to investigate the potential correlation between levels of climatic variables, air pollution and the mortality rate attributed to respiratory complications in Mashhad City, northeastern Iran from 2017 to 2021 using RF model. This objective was pursued to contribute to the existing body of knowledge in the field of environmental health, specifically in understanding how air quality may impact human health outcomes. The findings of this study could have significant implications for public health policies and interventions aimed at mitigating the adverse health effects of poor air quality.

Methods

In this analytical-descriptive study, we aimed to explore the relationship between meteorological parameters, air pollutants, and respiratory disease mortality using RF regression model.

Data collection

Our dataset comprised information collected from various sources:

1) Meteorological Parameters:

We obtained daily statistics for meteorological parameters, including pressure, humidity, temperature, and solar radiation, during the statistical period from 2017 to 2021. These data were sourced from the Mashhad Meteorological Organization, located on the Shahid Kalantari highway.

2) Air Pollutants:

Information related to air pollutants, specifically PM2.5, PM10, SO2, NO2, and CO, was collected from the statistical period. These pollutants play a crucial role in environmental health and can impact respiratory well-being.

3) Respiratory Disease Mortality Statistics:

To understand the health implications, we gathered data on respiratory mortality from 2017 to 2021. These statistics were obtained from the Health system registaration (SINA).

Random forest model (RF)

Model implementation and validation

The implementation of the RF model involved several steps. Initially, the entire climate data (pressure, humidity, temperature, and solar radiation), as well as pollutant data (PM2.5, PM10, SO2, NO2, and CO), along with information related to the mortality rate due to respiratory diseases, were randomly divided into two subsets for training and testing the model. The dataset contains daily information on features and mortality rates spanning 5 years from March 21, 2017, to March 19, 2021. 25% of the total data was designated testing data and kept separate from model training to ensure precise model validation.

Subsequently, the model was fitted to the data, and its accuracy was measured using various metrics. For precise validation of the model, K-fold cross-validation was employed. In this method, the entire dataset used for model training is divided into K equal subsets, and the model is trained K times separately. In each iteration, the model utilizes K-1 subsets as training data and one subset for validation. The final accuracy values reported for the model are obtained by averaging its accuracy across all iterations. Also, the hyperparameter tuning of the model is conducted simultaneously with this approach. Table 1 showcases the ultimate values of the most significant hyperparameters of the RF model.

Performance evaluation metrics

The coefficient of determination, or R^2, is a metric used in regression analysis to evaluate model performance. It indicates the percentage of variation in the dependent variable explained by the model. In other words, R^2 shows how much of the variance in the dependent variable is explained by the independent variables. R^2 ranges between zero and one, where a value of zero signifies poor model fit to the data and a value of one signifies perfect model fit. In other words, the closer the R^2 value is to one, the better the model performs. If y represents the observed or actual data, f represents the predicted data, n represents the number of data points used for validation, and $\hat{y}$ represents the mean of all observed data points, the mathematical equation for R^2 is as Eq. 1.

$$R^2 = 1 - \frac{SS_{res}}{SS_t} \quad , \quad SS_{res} = \sum_i^n (y_i - f_i)^2 \quad , \quad SS_t = \sum_i^n (y_i - \bar{y})^2 \tag{1}$$

Additionally, the metrics of Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) were also utilized, with their equations defined as Eq. 2 (14).

$$MAE = \frac{1}{n} \left(\sum_i^n |y_i - f_i| \right) \quad , \quad RMSE = \sqrt{\frac{1}{n} \left(\sum_i^n (y_i - f_i)^2 \right)} \tag{2}$$

Permutation variable importance (PVI)

To determine the importance of each variable, the permutation importance method was utilized. Firstly, the model's accuracy, measured by the R^2 metric, was computed in its original state. Then, the data for each variable were permuted, meaning the values of different rows were shuffled. Subsequently, the model's accuracy in this permutation scenario was calculated. The importance of each variable was derived from the difference between the initial accuracy and the accuracy after permutation. The greater the importance of a variable, the larger the difference in model accuracy it creates, indicating its significance. To calculate variable interactions, multiple variable rows were simultaneously permuted to compute the interaction value[15]. Finally, the final importance or interaction values of variables were determined iteratively according to Eq. 3 performed M times for enhanced

Hyperparameter	n-estimators	Max-depth	Random state
Value	1000	100	50
Definition	The number of decision trees in the forest	The maximum depth of each decision tree	The seed for the random generation of numbers

Table 1. The hyperparameters of the final RF model.

	The Year of study			
Variable	2017	2018	2019	2020
Minimum	0	0	0	0
Maximum	6	8	10	6
Mean	1.44	2.05	2.48	1.37
Standatd deviation	1.41	1.49	1.73	1.21

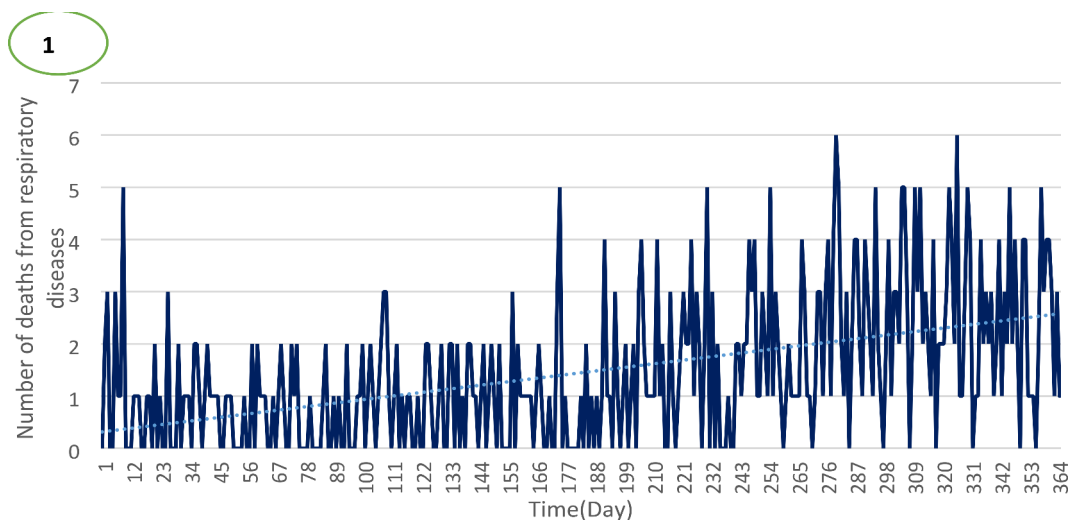**Table 2.** Respiratory mortality rate per day from 2017–2021 in Mashhad.

Fig. 1. Respiratory mortality rate in Mashhad City freom 2017–2020. (2017(1)- 2018(2)- 2019 (3)- 2020 (4)).

accuracy, and the average values obtained were utilized. In this equation, R^2_{base} represents the initial model accuracy, and R^2_{per} represents the model accuracy after permuting each variable.

$$PVI_i = \frac{1}{M} \left(\sum_{j=1}^{N} \left(R^2_{\text{base}} - R^2_{\text{per}} \right) \right) \tag{3}$$

Partial dependence
Partial Dependence, or the average change in the predicted value of the model, measures the relationship between a variable and the expected outcome. Equation 4 illustrates the mathematical formula for calculating partial dependence for variables $S\ (X_S)$

$$F_S = \frac{1}{n} \sum_{i=1}^{n} \widehat{f}\left(X_s, X_c \right) \tag{4}$$

In this equation, F_S represents the partial dependence function, $\widehat{f}$ denotes the machine learning model, and *n* is the number of rows of data used in model training. Essentially, Partial Dependence Functions (PDPs) are generated using this equation for each point of the variable(s) X_s,without altering the variables X_c. PDP plots, which utilize Eq. 4 to compute the values for each point, depict the relationship between the variable(s) of interest and the model's target variable[16]. For instance, if linear models like Linear Regression are employed, the relationship between all variables and the target variable is represented linearly. As mentioned, one of the advantages of RF model is their ability to represent non-linear relationships. One-way PDP plots illustrate the relationship between a variable and the target variable, while two-way PDPs depict the simultaneous effects of two variables on the target variable.

Results
Respiratory mortality rate
The average daily mortality of respiratory patients from 2017 to 2021 was 1.44, 2.05, 2.48, and 1.37 respectively (Table 2).

Figure 1 presents the daily mortality rates attributed to respiratory diseases from 2017 to 2020 in Mashhad. The data reveals variability in the number of fatalities across different days within this period. Notably, the latter half of each year consistently recorded a higher daily death toll compared to the first half. Furthermore, an ascending trend in the daily mortality rates is observed throughout the study period from 2017 to 2021. This trend underscores the escalating severity of respiratory diseases in the region.

4

Type of Pollutant		Standard Approved in 2013		Mean Daily Pollutant			
		µg/m³	PPm	2017	2018	2019	2020
CO (Carbon Monoxide) PPm	Maximum of 8 h	10,000	9	1.25	1.97	1.66	1.47
	A maximum of 1 h	40,000	35				
$PM_{2.5}$ (Particulate Matter) µg/m³	Annualy	10	-	30.46	25.50	26.96	27.74
	Maximum 24 h	25	-				
PM_{10}(Particulate Matter) µg/m³	Annualy	20	-	51.35	44.94	46.80	50.19
	Maximum 24 h	50	-				
NO_2(Nitrogen Oxide) PPb	Annualy	40	0.021	49.12	38.23	33.10	27.70
SO_2(Sulfur Dioxide) PPb	Annualy	20	0.007	7.61	7.18	8.65	8.95
	Maximum 24 h	100	0.037				

Table 3. Comparison of the concentration of pollutants in the years 2017–2021 with the Standard Concentration of Pollutants in 2010 in Mashhad City.

Validation metric	R^2	RMSE	MAE
Value	0.73	2.52	2.00

Table 4. The values of various Metrics for Model Validation under the Cross-validation Method.

Daily changes in meteorological parameters and air pollutants

According to the data presented in Table 3, the average daily concentration of the PM2.5 pollutant in 2017 was recorded as 30.46 µg/m³. This value exceeds the standard daily limit for this pollutant, which is set at 25 µg/m³, by 21.84%. In the subsequent years of 2018, 2019, and 2020, the concentration of this pollutant (PM2.5) continued to surpass the approved standard by 2%, 7.84%, and 10.96% respectively. Thus, it can be inferred that PM2.5 has been a significant contributor to the pollution levels in the city of Mashhad during the years under study.

For the PM10 pollutant, an increase of 2.7% and 0.38% was observed in 2017 and 2020 respectively, compared to the approved standard. However, a noticeable reduction in this pollutant was recorded in the years 2018 and 2019, with values of 10.4% and 10.12% below the standard.

It is noteworthy that the average daily concentration of the CO pollutant during the studied years (2017–2020) was significantly lower than the standard approved by the Supreme Council of Environmental Protection of Iran. This observation suggests that CO pollution was not a major concern during this period.

Random forest model (RF)

In this study, we aimed to explore the relationship between meteorological parameters and air pollutants with mortality rates from respiratory diseases. We employed the RF model as our primary analytical tool.

After implementing the k-fold cross-validation approach, the values for each validation metric were calculated. Table 4 illustrates the values corresponding to various validation metrics of the model. These values indicate that the model has successfully identified the relationship between input variables and the target variable or output, and it will exhibit high accuracy in predictions.

In addition, to quantify the importance of each parameter's impact on the mortality rate from respiratory diseases, we utilized the permutation feature importance method. This method assigns a numerical value to each variable based on its impact on the model's accuracy. A variable that has a greater impact on mortality from respiratory diseases will create a larger difference in the model's accuracy. Therefore, a larger number (closer to 1) indicates a higher importance of the variable in predicting mortality rates from respiratory diseases.

Table 5 shows the impact of different meteorological factors and pollutants on respiratory disease mortality rates in Mashhad from 2017 to 2020, as analyzed by the RF model. This offers a strong statistical assessment of each factor's significance.

We explored the impact of distinct meteorological factors and air pollutants on the mortality rates of patients suffering from respiratory ailments. This investigation was conducted independently, utilizing the RF model. To enhance the understanding of the relationships, partial dependence diagrams were employed.

Temporal and meteorological influences on respiratory disease mortality

Figure 2 illustrates the impact of cumulative changes in various parameters over 5 days on the mortality rates of patients with respiratory conditions. During the study period (2017–2020) we witnessed an escalating trend in respiratory disease mortality. The initial years (2017–2018) exhibited less variability in death rates (4–10 individuals), peaking at 11.8. However, the latter years (2019–2020) saw a surge, with the average death rate exceeding 16, marking a 35.5% increase compared to the initial years. This underscores the growing burden of respiratory diseases.

As illustrated in Fig. 3, we observed a significant seasonal variation in the mortality rate of respiratory patients over the course of a year. The death rate was found to be higher in the second half of the year, peaking in

Variable Name	Variable Importance
Month	0.112
Cumulative day (Sum of 5days)	0.548
Pmax (Maximum Pressure)	0.049
Mmax (Maximum Humidity)	0.043
Tmax (Maximum Temperature)	0.046
Radiation	0.037
CO (Carbon Monoxide)	0.025
NO_2 (Nitrogen Dioxide)	0.034
SO_2 (Sulfur Dioxide)	0.033
PM_{10} (Particulate matter with a diameter of 10 micrometres or less)	0.039
$PM_{2.5}$ (Particulate matter with a diameter of 2.5 micrometres or less)	0.029

Table 5. The impact of different Meteorological factors and pollutants on respiratory Disease Mortality Rates in Mashhad from 2017–2021 by Random Forest Model.
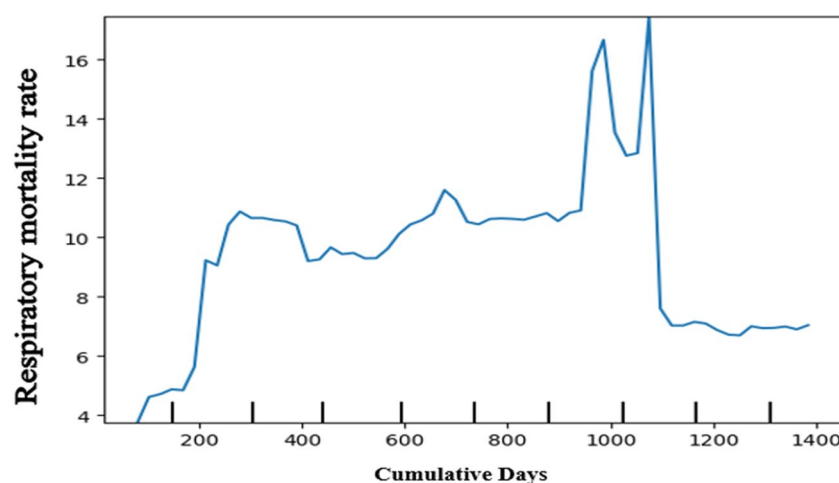
Fig. 2. The effects of 5-day cumulative changes on the mortality of respiratory patients in Mashhad From 2017–2021.

February with more than 10 deaths. Specifically, the cold season (temperatures below 30 °C) witnessed a higher death rate (more than 8 people) compared to the hot season (less than 8 people).

As illustrated in Fig. 4, during the statistical period of 2017–2021 in Mashhad, an increase in air pollutants was noted at high pressures. This escalation in air pollution was associated with a rise in the mortality rate of respiratory patients.

We also examined the impact of solar radiation on patient mortality during the statistical period. The findings suggest that sunlight had limited effects on the mortality of respiratory patients (Fig. 5).

A detailed analysis of temperature effects revealed that in low temperatures (0–30 °C), the average mortality rate of respiratory patients fluctuated significantly, reaching its highest value between 9.30 and 9.32 people (Fig. 6). However, at temperatures above 30 °C, the death rate decreased to 1.71.

We further investigated the relationship between humidity and mortality rate. As illustrated in Fig. 7, it was observed that lower humidities (30-70%) were associated with more deaths from respiratory diseases (9.5–9.3 people). Conversely, at higher humidities (70–100%), the death rate due to respiratory diseases decreased to less than 9 people.

Air pollutants and respiratory mortality

The pollutants examined in this study included SO_2, CO, PM10, PM2.5, and NO2. As illustrated in Fig. 8, an initial increase in SO_2 concentration (6–8 ppb) corresponded with a rise in the mortality rate of respiratory patients, with more than 9 additional deaths observed. However, at concentrations exceeding 8 ppb, the mortality rate decreased by 1.61%. The CO concentrations increased (2–25 ppm), the average mortality rate of respiratory patients also rose (ranging from 9.08 to 22). The mortality rate appeared to stabilize at higher concentrations, with no significant changes observed. As illustrated in Fig. 8 about PM10, at lower concentrations (15–44 µg/m³), an increase in PM10 was associated with a 1.09% decrease in the mortality rate. However, at concentrations between 44 and 100 µg/m³, the mortality rate increased by 2.76%.
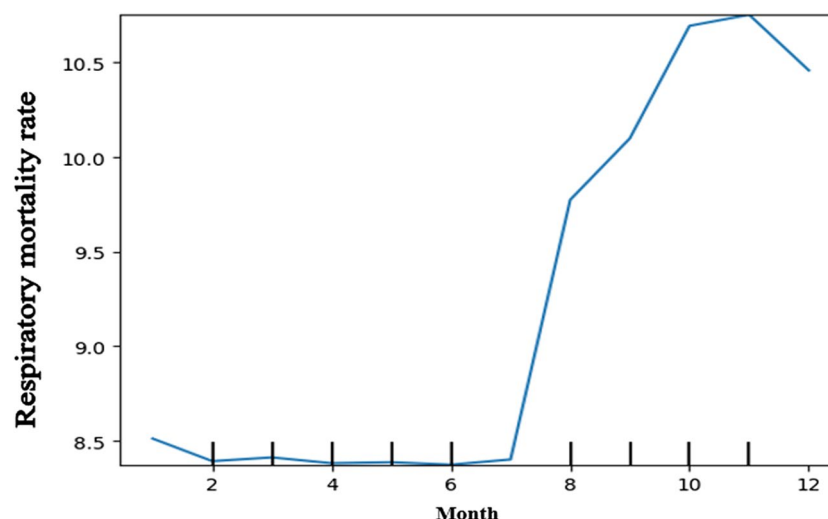
Fig. 3. The effects of changes in the months of the year on the mortality of respiratory patients in Mashhad during the yeasr 2017–2021.
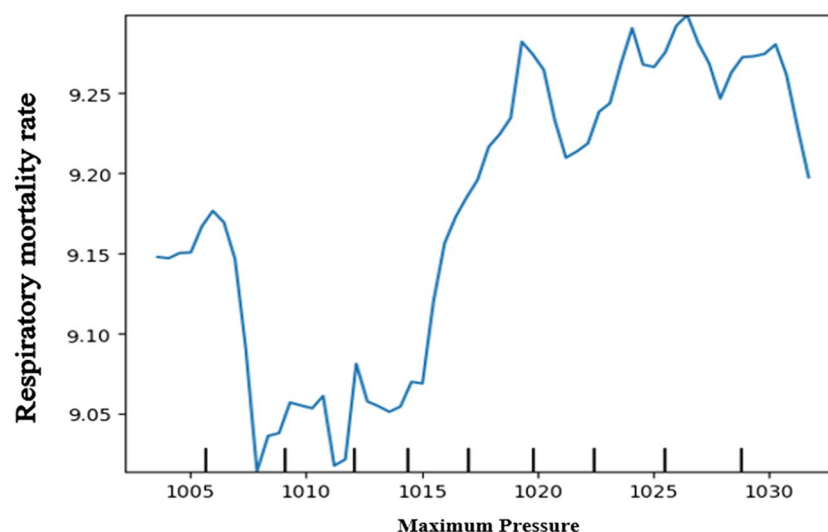

Fig. 4. The effects of changes in pressure on the mortality rate of respiratory patients in Mashhad during the years 2017–2021.

As illustrated in Fig. 8, as CO concentrations increased (2–25 ppm), the average mortality rate of respiratory patients also rose (ranging from 9.08 to 22). The mortality rate appeared to stabilize at higher concentrations, with no significant changes observed.

As illustrated in Fig. 8, at lower concentrations (10–25 μg/m^3), the average mortality rate was between 9.14 and 20.9. However, as PM2.5 concentrations increased (25–50 μg/m^3 and beyond), the average mortality rate also rose, ranging from 9.20 to 9.28.

As illustrated in Fig. 8, at concentrations of 20–30 ppb, the average mortality rate decreased by 2.68%. However, at concentrations between 30 and 60 ppb, the average mortality rate increased by 3.31%. Overall, the average mortality rate of respiratory patients related to NO2 increased by 2.19% during the study period.

Discussion

This study employs the RF statistical model to investigate the influence of meteorological parameters (pressure, temperature, humidity, solar radiation) and air pollutants (CO, NO2, SO2, PM2.5, PM10) on respiratory disease mortality rates in Mashhad from 2017 to 2021. The research aims to highlight the significance of these parameters in environmental health assessments and to estimate their impact on respiratory-related deaths. This underscores the intricate relationship between climatic factors, air pollution, and human health.

Rising air pollution in Mashhad is expected to increase respiratory hospitalizations, necessitating urgent attention and control measures from officials and experts. The time parameter emerged as the most influential factor (Variable
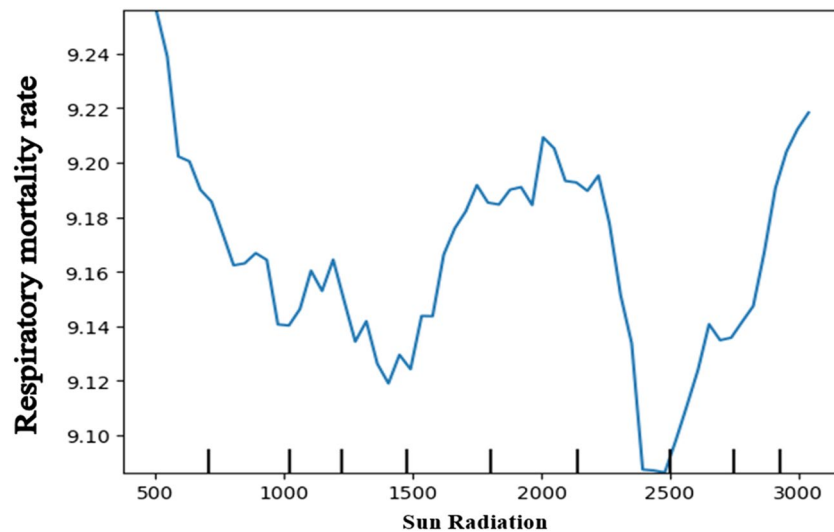
Fig. 5. The effects of changes solar radiation on the mortality rate of respiratory patients in Mashhad during the years 2017–2021.
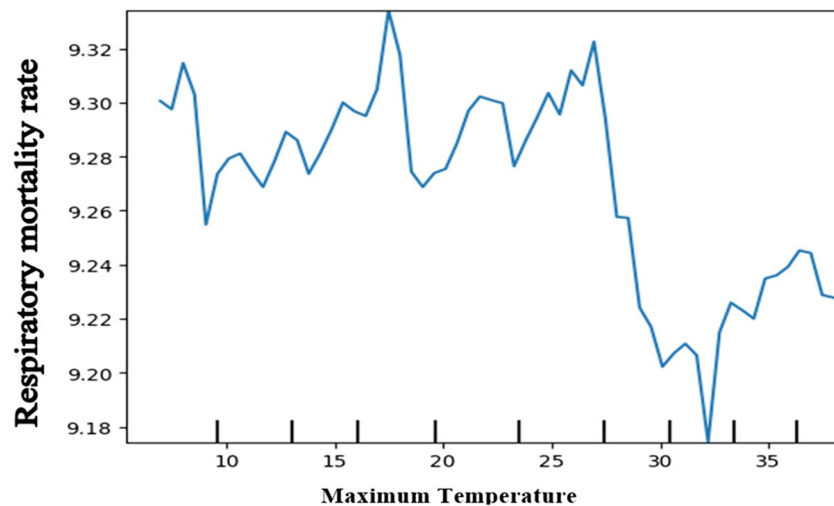

Fig. 6. The effect of changes in temperature on the mortality rate of respiratory patients in Mashhad during the years 2017–2021.

Importance = 0.548). When the day and month parameters were excluded, the pressure parameter had the highest importance coefficient (VI = 0.049). In a separate investigation, the impact of pollutants (PM2.5, PM10, O3, SO2, and NO2) on mortality rates was examined over cumulative time delays of 0–3, 0–4, and 0–5 days. The findings of this study indicated that the influence of these pollutants on mortality rates was most pronounced over a cumulative time delay of 0–5 days, compared to other time delays. The findings of this study indicated that the influence of these pollutants on mortality rates was most pronounced over a cumulative time delay of 0–5 days, compared to other time delays. Consequently, a cumulative time delay of 0–5 days was deemed optimal for obtaining more insightful results in this study[17]. Several studies have demonstrated an association between air pollutants and mortality rates. These studies utilized cumulative day lags of 0–3, 0–4, 0–5, 0–10, 0–20, and 0–30 days. Among these, the 0–5 day time lag was found to be the most effective in demonstrating the impact of air pollutants on mortality rates. Therefore, we have chosen to use a 5-day cumulative time lag for our study[18].

Pressure parameters revealed less fluctuation in respiratory disease-related deaths at the start of the statistical period, with a decreasing trend observed at pressures of 1005–1005 Hpa. However, an increase in the average death rate (2.64%) was noted at higher pressures of 1030 − 1015 Hpa towards the end of the period. In a parallel study conducted in the city of Shiraz, which aligns with the context of this research, it was observed that a correlation exists between meteorological conditions and mortality rates. Specifically, a decrease in temperature was often accompanied by an increase in air pressure. The study by Future and colleagues in 2011 demonstrated a notable inverse relationship between atmospheric pollutants and certain weather conditions, specifically wind speed and relative humidity[19].
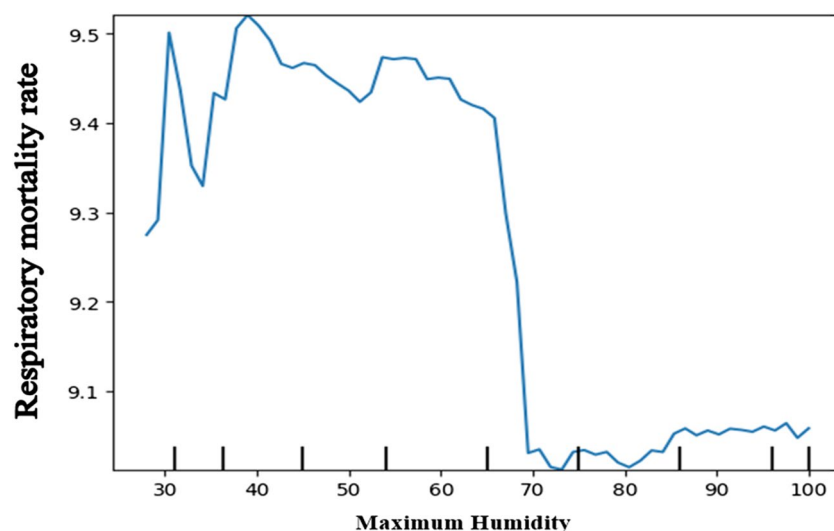
Fig. 7. The effects of changes in humidity on the mortality rate of respiratory patients in Mashhad during the years 2017–2021.
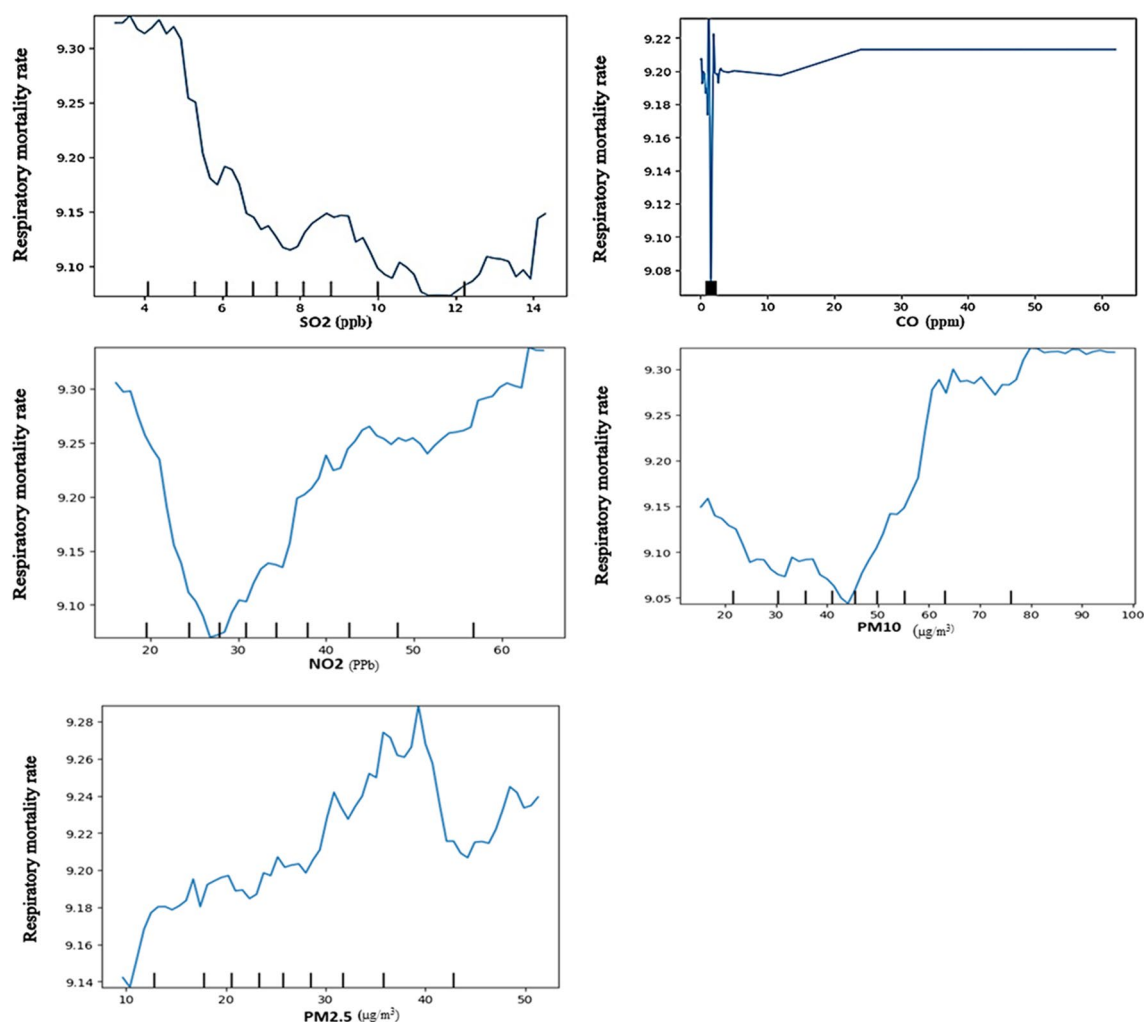

Fig. 8. The Effects of Changes in the Concentration of Pollutants on the Mortality Rate of Respiratory Patients in Mashhad During the Years 2017–2021.

This meteorological shift was associated with a rise in mortality rates, particularly among patients with respiratory conditions. The radiation parameter had the lowest importance coefficient (VI = 0.037) among these parameters. Radiation parameters showed higher average death rates during periods of low radiation (500–1500), typically in colder seasons. High radiation levels and warmer conditions may reduce virus replication and survival, while viruses may persist longer in moist and cold settings, potentially leading to higher infection rates and increased mortality during seasons with lower radiation and colder temperatures. An inverse relationship was observed with increasing radiation (1500–3000), indicating a decrease in average death rates. However, the radiation parameter fluctuated significantly during the statistical period of 2016–2019, suggesting no significant effect of solar radiation on respiratory patient mortality. However, In a recent epidemiological study conducted in Manaus, a significant linear correlation was observed between the mortality rate attributed to COVID-19 and the degree of exposure to solar radiation. The data suggested that individuals who experienced higher levels of solar radiation exposure weekly demonstrated lower mortality rates from COVID-19[20]. This finding underscores the potential role of environmental factors in disease outcomes and warrants further investigation.

The study findings suggest a temporal increase in mortality rates from respiratory diseases, with a 35.5% rise observed in 2019 and 2020 compared to 2017 and 2018. Seasonal variations were noted, with higher mortality rates in colder seasons compared to warmer ones. During the winter, individuals often use heaters and keep windows closed, leading to stagnant indoor air that can facilitate the transmission of viral infections. This situation increases susceptibility to illness, particularly among the elderly, raising the risk of mortality. Evidence from Jinan, China, indicates that PM2.5 levels peak in winter and are lowest in summer. Consequently, the elevated PM2.5 levels during winter likely exacerbate respiratory issues and increase mortality rates[21]. This seasonal trend could explain the higher mortality observed in our study during the winter months[22]. Contrary to our findings, a study by Dardir et al. in 2022 demonstrated a positive correlation between high temperatures during warm seasons and increased mortality rates and emergency room visits among elderly individuals with respiratory issues. This discrepancy suggests that air pollution, in conjunction with elevated temperatures, may exacerbate respiratory problems in this demographic[23]. Further research is needed to reconcile these divergent results and to fully understand the complex interplay between environmental factors and respiratory health.

Temperature parameters revealed the highest average mortality rates at lower temperatures (0–30 degrees Celsius), with a decrease in mortality rate (1.71%) observed at temperatures above 30 degrees Celsius. A Brazilian study found that cold temperatures elevated mortality risks for circulatory, and respiratory disease while warm temperatures increased cardiovascular disease mortality[24]. Humidity parameters showed higher mortality rates at lower humidity levels (30–70%), with a decrease in mortality rates observed at higher humidity levels (70–100%). A Chinese study identified both high and low absolute humidity as significant factors in respiratory disease mortality, with high humidity having a greater impact[25].

We also evaluated the importance of different pollutants. The PM10 pollutant had the highest importance coefficient (VI = 0.039), suggesting it had the greatest impact on the mortality rate caused by respiratory diseases among the studied pollutants. Particulate matter smaller than 10 micrometers in diameter can accumulate within the lungs, impairing alveolar macrophages. This can exacerbate respiratory issues and infections, ultimately increasing the risk of death in vulnerable patients Following PM10, the pollutants with the highest importance coefficients were NO2 (VI = 0.034), SO2 (VI = 0.033), PM2.5 (VI = 0.029), and CO (VI = 0.025). These coefficients highlight the varying degrees of impact each pollutant can have on respiratory disease mortality rates. A comprehensive systematic review corroborates our findings. It demonstrated that during the warmer seasons, an increase in the levels of PM10, PM2.5, CO, NO2, and O3 pollutants corresponded with a heightened rate of hospital admissions due to respiratory complications. This evidence further substantiates our study's conclusions, reinforcing the significant impact of these pollutants on respiratory health[26]. Some hidden factors should also be considered. For example, during the COVID-19 pandemic in 2020, $PM_{2.5}$ levels in Shanghai decreased significantly compared to 2019 due to lockdown measures. Such unexpected phenomena can influence air quality and subsequently impact respiratory health and mortality rates[27].

For PM10, an initial decrease in mortality rate (1.09%) was observed with increasing concentrations (15–44 µg/m^3), followed by a 2.76% increase at higher concentrations (44–100 µg/m^3). Similarly, for PM2.5, lower concentrations (10–25 µg/m^3) were associated with a lower average death rate (9.14–20.9 individuals), but an upward trend in mortality was observed with increasing concentrations (25–50 µg/m^3 and above), with an average mortality rate of 9.20–9.28 individuals. Particulate matter with a diameter of 0.25 to 1 micron, often referred to as PM2.5, exhibits prolonged retention in the respiratory system due to its unique motion characteristics. The composition of these particles is complex, encompassing organic and elemental carbon, mineral dust, trace elements, and water. The presence of these compounds has been linked to the onset of respiratory diseases, underscoring the potential health risks associated with exposure to such particulates. In a research study carried out in China, it was observed that the impact of PM10 on cardiovascular and respiratory mortality was more pronounced at elevated temperatures compared to lower temperatures. This suggests a potential interaction between temperature and PM10 levels in influencing health outcomes, particularly cardiovascular and respiratory mortality[28].

In the case of NO_2, a decrease in the average mortality rate (2.68%) was noted at concentrations of 20–30 ppb at the beginning of the statistical period. However, at higher concentrations (30–60 ppb), the average mortality rate increased by 3.31%. Overall, the study concludes that deaths caused by respiratory diseases related to NO_2 pollution increased by 2.19% during the study period. Additionally, new evidence confirms that $PM_{2.5}$ can significantly influence the formation of ozone (O_3) and further $PM_{2.5}$ particles[29]. Our findings indicate that a rise in SO_2 concentrations (6–8 ppb) was associated with higher mortality rates. Interestingly, we observed a 1.61% decrease in mortality when SO_2 levels surpassed 8 ppb. While it's conceivable that elevated SO_2 levels could contribute to increased respiratory mortality, it seems counterintuitive that lower SO_2 levels would lead to reduced respiratory mortality rates. In a parallel study conducted by Chong, K.C. et al. in 2021, no significant relationship was found between SO_2 concentrations and hospitalizations due to pediatric bronchitis[30]. Thus,

although we identified a significant inverse correlation between the prevalence of respiratory mortality rates and the concentration of SO_2. However, this appears to be coincidental, as there is no known biological mechanism to explain a protective effect of SO_2 against respiratory mortality.

For CO pollutants, an increase in mortality was observed with increasing levels (2–25 ppm), but no noticeable effect was observed during the statistical period. The Ahvaz study demonstrated a significant link between NO_2, CO, and O_3 exposure and increased mortality rates in patients with cardiovascular and respiratory conditions, with NO_2 being the most impactful[31].

Limitation of the study

It is suggested that future studies could explore additional determinants of respiratory patient mortality rates, including variables like age, profession, and educational background. Furthermore, it would be beneficial to extend the investigation of the correlation between meteorological parameters, air pollutants, and mortality rates from respiratory diseases over a more extended statistical period (e.g., five years or more). This study was conducted over four years, and a longer duration could potentially yield more precise results.

During this study, we encountered the COVID-19 pandemic, which may have influenced our findings. Additionally, the quarantine measures and travel-ban during that period likely altered air quality. Therefore, further studies are needed to confirm our findings.

Conclusion

The RF model successfully identified the relationship between meteorological factors, air pollution, and mortality rates, demonstrating high accuracy. The interplay of climate change and air pollution significantly impacts respiratory disease mortality, with PM2.5, PM10, and NO2 being particularly harmful. This underscores the need for comprehensive strategies to mitigate the health impacts of environmental pollutants. Policymakers should enforce stricter air quality regulations, and public health initiatives must raise awareness and provide protective measures. Further research is essential to explore long-term effects and develop effective interventions.

Data availability

The datasets generated and/or analyzed during the current study are not publicly available due to confidentiality but are available from the corresponding author on reasonable request.

References

1. Syuhada, G. et al. Impacts of air pollution on health and cost of illness in Jakarta, Indonesia. *Int. J. Environ. Res. Public Health*. **20** (4), 2916 (2023).
2. Sun, H-L., Chou, M-C. & Lue, K-H. The relationship of air pollution to ED visits for asthma differ between children and adults. *Am. J. Emerg. Med.* **24** (6), 709–713 (2006).
3. Idani, E. et al. Risk factors associated with asthma among adults in Khuzestan, Southwest Iran. *Clin. Epidemiol. Global Health*. **8** (2), 350–355 (2020).
4. Meng, Y., Lu, Y., Xiang, H. & Liu, S. Short-term effects of ambient air pollution on the incidence of influenza in Wuhan, China: a time-series analysis. *Environ. Res.* **192**, 110327 (2021). Epub 20201017.
5. Park, S. K. et al. Long-term exposure to air pollution and type 2 diabetes mellitus in a multiethnic cohort. *Am. J. Epidemiol.* **181** (5), 327–336 (2015). Epub 20150217.
6. Sahebari, M., Rajabi, E., Esmaili, H., Orooji, A. & Dowlatabadi, Y. Effects of air pollution on disease activity and health-related quality of life of systemic Lupus Erythematous patients: an Iranian Observational Longitudinal Study. *Curr. Rheumatol. Rev.* **19** (2), 222–229 (2023).
7. Sahebari, M. et al. Influence of air pollutants on the disease activity and quality of life in rheumatoid arthritis, an Iranian observational longitudinal study. *Med. J. Islam Repub. Iran.* **37**, 65 (2023). Epub 20230612.
8. Faraji, M. et al. Exposure to ambient air pollution and prevalence of asthma in adults. *Air Qual. Atmos. Health*. **14** (8), 1211–1219 (2021).
9. Ren, Z. et al. Effect of ambient fine particulates (PM(2.5)) on hospital admissions for respiratory and cardiovascular diseases in Wuhan, China. *Respir Res.* **22** (1), 128 (2021). Epub 20210428.
10. Brunekreef, B. et al. Mortality and morbidity effects of long-term exposure to low-level PM(2.5), BC, NO(2), and O(3): an analysis of European cohorts in the ELAPSE Project. *Res. Rep. Health Eff. Inst.* **2021** (208), 1–127 (2021).
11. World Health Organization WHO. WHO releases country estimates on air pollution exposure and health impact. *Whoint* (2016).
12. Altieri, K. E. & Keen, S. L. Public health benefits of reducing exposure to ambient fine particulate matter in South Africa. *Sci. Total Environ.* **684**, 610–620 (2019). Epub 20190525.
13. Bonyadi, Z., Ehrampoush, M. H., Ghaneian, M. T., Mokhtari, M. & Sadeghi, A. Cardiovascular, respiratory, and total mortality attributed to PM 2.5 in Mashhad, Iran. *Environ. Monit. Assess.* **188**, 1–6 (2016).
14. Moezzi, S. M. M., Mohammadi, M., Mohammadi, M., Saloglu, D. & Sheikholeslami, R. Machine learning insights into PM2. 5 changes during COVID-19 lockdown: LSTM and RF analysis in Mashhad. *Environ. Monit. Assess.* **196** (5), 453 (2024).
15. Kaneko, H. Cross-validated permutation feature importance considering correlation between features. *Anal. Sci. Adv.* **3** (9–10), 278–287 (2022).
16. Inglis, A., Parnell, A. & Hurley, C. B. Visualizing variable importance and variable interaction effects in machine learning models. *J. Comput. Graphical Stat.* **31** (3), 766–778 (2022).
17. San Tam, W. W., Wong, T. W. & Wong, A. H. Association between air pollution and daily mortality and hospital admission due to ischaemic heart diseases in Hong Kong. *Atmos. Environ.* **120**, 360–368 (2015).
18. Ravindra, K., Rattan, P., Mor, S. & Aggarwal, A. N. Generalized additive models: building evidence of air pollution, climate change and human health. *Environ. Int.* **132**, 104987 (2019). Epub 20190806.
19. Bonyadi, Z. et al. Impact of exposure to ambient air pollutants on the admission rate of hospitals for asthma disease in Shiraz, southern Iran. *Chemosphere*. **262**, 128091 (2021). Epub 20200823.
20. Barcellos, D. D. S., Fernandes, G. M. K. & de Souza, F. T. Data based model for predicting COVID-19 morbidity and mortality in metropolis. *Sci. Rep.* **11** (1), 24491 (2021). Epub 20211229.

21. Guo, Q., He, Z. & Wang, Z. Change in air quality during 2014–2021 in Jinan City in China and its influencing factors. *Toxics*. **11** (3), 210 (2023). Epub 20230224.
22. Kangas, T, Gadeyne, S., Lefebvre, W., Vanpoucke, C. & Rodriguez-Loureiro, L. Are air quality perception and PM(2.5) exposure differently associated with cardiovascular and respiratory disease mortality in Brussels? Findings from a census-based study. *Environ. Res.* **219**, 115180 (2023). Epub 20221228.
23. Dardir, M., Wilson, J. & Berardi, U. Air Quality and Heat-related Health Impacts during Heat Events. (2022).
24. Ferreira, L. C. M. et al. Ambient temperature and mortality due to acute myocardial infarction in Brazil: an ecological study of time-series analyses. *Sci. Rep.* **9** (1), 13790 (2019).
25. Chen, S. et al. The role of absolute humidity in respiratory mortality in Guangzhou, a hot and wet city of South China. *Environ. Health Prev. Med.* **26** (1), 109 (2021). Epub 20211117.
26. Grigorieva, E. & Lukyanets, A. Combined effect of hot weather and outdoor air pollution on respiratory health. *Literature Rev. Atmos.* **12** (6), 790 (2021).
27. Guo, Q., He, Z. & Wang, Z. Predicting of Daily PM(2.5) concentration employing wavelet artificial neural networks based on meteorological elements in Shanghai, China. *Toxics*. **11** (1), 51 (2023). Epub 20230103.
28. Tian, L. et al. The effects of interaction between particulate matter and temperature on mortality in Beijing, China. *Environ. Science: Processes Impacts*. **20** (2), 395–405 (2018).
29. Guo, Q., He, Z. & Wang, Z. The characteristics of air quality changes in Hohhot City in China and their relationship with meteorological and socio-economic factors. *Aerosol Air Qual. Res.* **24** (5), 230274 (2024).
30. Leung, S. Y. et al. Short-term association among meteorological variation, outdoor air pollution and acute bronchiolitis in children in a subtropical setting. *Thorax*. **76** (4), 360–369 (2021).
31. Sarizadeh, G., Jaafarzadeh, N., Moattar, F., Roozbehani, M. M. & Tahmasebi, Y. *Relationship between the urban air pollution and rate of cardiovascular and respiratory death in Ahvaz in the period of 10 years (2007–2017)* CABI Databases. (2021).

Acknowledgements

Author contributions

Y.D. and S.A. and M.S. conceived and designed the study, analyzed the data, and wrote the main manuscript text. M.M. performed statistical analysis. SMM.M. prepared figures. All authors reviewed the manuscript.

Funding

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-74440-2.

Correspondence and requests for materials should be addressed to M.S. or M.M.

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

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

Abbreviations

NO_2	Nitrogen Dioxide
CO	Carbon Monoxide
SO_2	Sulfur Dioxide
PM_{10}	Inhalable Particles
$PM_{2.5}$	Fine Particulate Matter
O_3	Ozone
COVID-19	coronavirus disease 2019