

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/362174307>

Time Series Analysis and Forecasting of Air Pollutants Based on Prophet Forecasting Model in Jiangsu Province, China Citation

Article in *Frontiers in Environmental Science* · July 2022

DOI: 10.3389/fenvs.2022.945628

CITATIONS

4

READS

101

21 authors, including:



Ahmad Hasnain

Nanjing Normal University

17 PUBLICATIONS 177 CITATIONS

SEE PROFILE



Uzair Aslam Bhatti

Hainan University

95 PUBLICATIONS 666 CITATIONS

SEE PROFILE



Aamir Hussain

Muhammad Nawaz Shareef University of Agriculture

30 PUBLICATIONS 360 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



Fire simulation, Smoke propagation and Ventilation evaluation. [View project](#)



Remote sensing using Lidar Technology [View project](#)



Time Series Analysis and Forecasting of Air Pollutants Based on Prophet Forecasting Model in Jiangsu Province, China

Ahmad Hasnain^{1,2,3}, Yehua Sheng^{1,2,3*}, Muhammad Zaffar Hashmi⁴, Uzair Aslam Bhatti⁵, Aamir Hussain⁶, Mazhar Hameed⁷, Shah Marjan⁸, Sibghat Ullah Bazai⁹, Mohammad Amzad Hossain¹⁰, Md Sahabuddin¹¹, Raja Asif Wagan¹² and Yong Zha^{1,2,3*}

OPEN ACCESS

Edited by:

Dervis Kirikkaleli,
European University of Lefka, Turkey

Reviewed by:

Rita Yi Man Li,
Hong Kong Shue Yan University, Hong
Kong SAR, China
Minhaj Ali,
Islamia University of Bahawalpur,
Pakistan

*Correspondence:

Yehua Sheng
shengyehua@njnu.edu.cn
Yong Zha
yzha@njnu.edu.cn

Specialty section:

This article was submitted to
Environmental Economics and
Management,
a section of the journal
Frontiers in Environmental Science

Received: 16 May 2022

Accepted: 16 June 2022

Published: 22 July 2022

Citation:

Hasnain A, Sheng Y, Hashmi MZ, Bhatti UA, Hussain A, Hameed M, Marjan S, Bazai SU, Hossain MA, Sahabuddin M, Wagan RA and Zha Y (2022) Time Series Analysis and Forecasting of Air Pollutants Based on Prophet Forecasting Model in Jiangsu Province, China. *Front. Environ. Sci.* 10:945628. doi: 10.3389/fenvs.2022.945628

¹Key Laboratory of Virtual Geographic Environment, Ministry of Education, Nanjing Normal University, Nanjing, China, ²School of Geography, Nanjing Normal University, Nanjing, China, ³Jiangsu Center for Collaborative Innovation in Geographical Information, Resource Development and Application, Nanjing, China, ⁴Department of Chemistry, COMSATS University Islamabad, Islamabad, Pakistan, ⁵School of Information and Communication Engineering, Hainan University, Haikou, China, ⁶Department of Computer Science, Muhammad Nawaz Shareef University of Agriculture Multan, Multan, Pakistan, ⁷School of Information and Communication Engineering North University of China, Taiyuan, China, ⁸Department of Software Engineering, Balochistan University of Information Technology, Engineering and Management Sciences (BUITEMS), Quetta, Pakistan, ⁹Department of Computer Engineering, Balochistan University of Information Technology, Engineering and Management Sciences (BUITEMS), Quetta, Pakistan, ¹⁰School of Business and Economics (SOBE), United International University (UIU), Dhaka, Bangladesh, ¹¹College of Economics and Management, Nanjing University of Aeronautics and Astronautics (NUA), Nanjing, China, ¹²Department of Information Technology, Balochistan University of Information Technology, Engineering and Management Sciences (BUITEMS), Quetta, Pakistan

Due to recent developments in the global economy, transportation, and industrialization, air pollution is one of main environmental issues in the 21st century. The current study aimed to predict both short-term and long-term air pollution in Jiangsu Province, China, based on the Prophet forecasting model (PFM). We collected data from 72 air quality monitoring stations to forecast six air pollutants: PM₁₀, PM_{2.5}, SO₂, NO₂, CO, and O₃. To determine the accuracy of the model and to compare its results with predicted and actual values, we used the correlation coefficient (R), mean squared error (MSE), root mean squared error (RMSE), and mean absolute error (MAE). The results show that PFM predicted PM₁₀ and PM_{2.5} with R values of 0.40 and 0.52, RMSE values of 16.37 and 12.07 µg/m³, and MAE values of 11.74 and 8.22 µg/m³, respectively. Among other pollutants, PFM also predicted SO₂, NO₂, CO, and O₃ with R values are between 5 µg/m³ to 12 µg/m³; and MAE values between 2 µg/m³ to 11 µg/m³. PFM has extensive power to accurately predict the concentrations of air pollutants and can be used to forecast air pollution in other regions. The results of this research will be helpful for local authorities and policymakers to control air pollution and plan accordingly in upcoming years.

Keywords: prophet forecasting model, time series model, air pollution, machine learning, jiangsu province, China

INTRODUCTION

Due to developments in the global economy, transportation, and industrialization, air pollution is one of the main widespread environmental issues. The World Air Quality Report noted that many Asian countries have experienced high levels of air pollution, particularly in cities in China, Pakistan, India, and Bangladesh (AirVisual, 2018; AirVisual, 2019). China, with its substantial population, transportation, and industries, is the largest developing country in the world, and in the last 3 decades, many cities and regions of the country have faced serious air pollution (Zhao et al., 2020). In the last few years, due to strict restrictions on industrial emissions, transportation, and heating activities, China experienced a slight decline in air pollution, but efforts are still required to secure the environment at a significant level (Wu et al., 2020). Associated with adverse effects and an impact on climate change, air pollution has attracted widespread interest from scholars and administrations (Lee et al., 2020; Wang et al., 2022).

Pollutants that threaten human health include particulate matter ($PM_{2.5}$ and PM_{10}), sulfur dioxide (SO_2), nitrogen dioxide (NO_2), carbon monoxide (CO), and ozone (O_3), as studied by Li and Shi (2016). Particulate matter can be defined as a combination of strong particles and liquid droplets, which has a more significant impact on human health than other pollutants (WHO-World Health Organization, 2018; Guo et al., 2020). PM_{10} and $PM_{2.5}$ contain inducing materials such as lipopolysaccharide and polycyclic aromatic hydrocarbons and transition metals (zinc, copper, manganese) (Bilal et al., 2021; Yang et al., 2022). PM_{10} can infiltrate the lower airways in the form of thoracic particles with an aerodynamic diameter of fewer than 10 μm . PM_{10} causes various lung diseases, such as asthma, and cardiovascular diseases (Huang et al., 2016; Hasnain et al., 2021). $PM_{2.5}$, which is particulate matter with a diameter of 2.5 μm or less, has been proven to have adverse health impacts (Conibear et al., 2018; Zhang et al., 2020).

Two other important air pollutant factors are SO_2 and NO_2 , which pose a hazardous risk to human health and threaten the environment (Zhang et al., 2019). SO_2 is extremely reactive with other compounds, which can cause additional environmental pollution, such as particulate matter and acid rain (Tang et al., 2016; Liu and Sun, 2019). NO_2 severely degrades the human respiratory system, leading to respiratory symptoms and asthma (Liu and Sun, 2019). NO_2 can also cause additional environmental pollution with other compounds and cause severe damage to both human health and the environment (Su et al., 2017; Shairsingh et al., 2021). CO is another major primary air pollutant in the atmosphere that affects human health and causes air pollution (Choi et al., 2017; Petetin et al., 2018). CO is an odorless and colorless gas with a long-life cycle of roughly 2–3 months. In the atmosphere, CO plays a vital role that also involves O_3 production (Liu et al., 2018; Fan et al., 2020; Qayyum et al., 2021). Although O_3 is a necessity for life on earth and plays an important role, it also has harmful effects and is linked with numerous respiratory issues such as lung scarring and loss of lungs (Fajersztajn et al., 2013; Dehghani et al., 2017; Wang et al., 2019; Aamir et al., 2021; Zhao et al., 2020). Due to its

adverse effects and high phytotoxicity, O_3 is also responsible for the reduced production of agricultural yields worldwide (Salonen et al., 2019; Duan et al., 2021; Yang et al., 2021).

In early 2013, due to an unprecedented pollution event that caused inordinate risk to people's lives and property, there was need to implement preventive measures to improve the air quality of the country. In this scenario, the Chinese State Council issued the Action Plan on Prevention and Control of Atmospheric Pollution on 10 September 2013 (The State Council of China, 2013; Dai et al., 2018). After the plan was implemented, the air pollution status across the country was reduced compared with the previous year. No doubt due to control policies and strict restrictions, China experienced a slight decline in recent years, but many cities and regions of the country still suffer from haze pollution (Duan et al., 2021). In this scenario, there is a dire need to take preventive measures and accurately forecast pollution events as early as possible. For this purpose, this paper attempts to predict main air pollutant factors (PM_{10} , $PM_{2.5}$, SO_2 , NO_2 , CO and O_3) based on the Prophet forecasting model (PFM) in Jiangsu Province, China. The study provides useful information and credible outcomes for local authorities and policymakers to control air pollution and to plan accordingly in upcoming years.

The rest of the paper is organized as follows: Section 2 comprises a literature review. Section 3 presents the study area, data set, data processing, the model and statistical analysis to evaluate the model's performance. Section 4 presents the results and discussion. Section 5 discusses the conclusions, policy implications, limitations and future research directions.

LITERATURE REVIEW

The time series prediction method is widely used by scholars and researchers in many fields to predict air pollutant concentrations (Appel et al., 2017; Zhao et al., 2020; Qayyum et al., 2022). Chen and Li (2022) used the hedonic regression model, Google AutoML and Microsoft AutoML for forecasting the housing prices and found that the Google AutoML model is robust and performs better with R square (0.820) is higher. About other models, it has been noted that to make long-term air pollution predictions, current non-machine-learning models have major flaws and errors, such as the community multiscale air quality modeling system (CMAQ), which is severely disadvantaged due to its intricate and complex system and source list and its requirement for and regular updates (Xi et al., 2015; Deters et al., 2017). In air pollution prediction models, this has led to many studies concentrating on the premise of machine learning. Li et al. (2015) and Deters et al. (2017) presented models that were unable to make accurate predictions and relied deeply on meteorological parameters. As a result, these models are unable to effectively predict meteorological events amid the climate change crisis (Scher and Messori, 2019). Pasero and Mesin (2010) and Maleki et al. (2019) used artificial neural networks (ANNs) for air pollution prediction. They noted that the ANNs relied on meteorological factors as highlighted in previous studies and showed many deficiencies. Their results indicated that due to overfitting of data, the model had the poor capability to predict time series and still failed

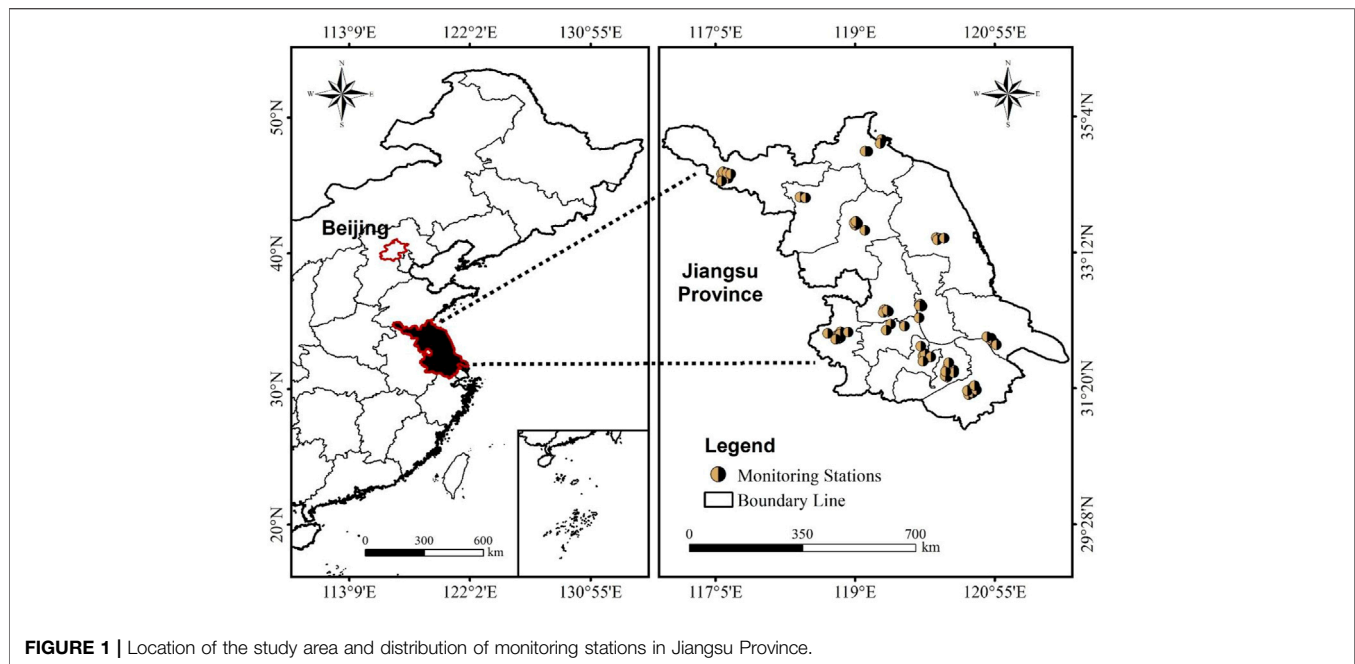


FIGURE 1 | Location of the study area and distribution of monitoring stations in Jiangsu Province.

to accurately forecast air quality. Meanwhile, for air pollution forecasting based on ANNs, Cabaneros et al. (2019) showed that to forecast long-term air pollutant factors with optimization approaches, feed-forward and hybrid ANN models were predominantly used. Yang et al. proposed a vector regression model to predict air pollutant concentrations by considering spatial assortment (Deng et al., 2018). In recent years, such data-driven methods with low accuracy in predicting air quality have been studied. Bhatti et al. (2021) used a time-series prediction model for forecasting air pollution using SARIMA and a factor analysis approach. This method using SARIMA approached the accuracy of 67% with better results than the ARIMA time series model. Used AutoML approach for forecasting the price index approach and providing useful predictions for future prices. Bhatti et al. (2019) used K-nearest neighbor method for recommendation and prediction using pattern recognition of datasets which provided an accuracy of more than 70% for the used datasets. Kamińska (2018) presented the same machine learning method with random forest regression (RFR); due to its nonlinear pattern, it is a popular approach based on the alteration of a few parameters. However, RFR has low accuracy of prediction because its highest and lowest values are bound to the training set data (Ivanov et al., 2018). Fuller et al. (2002) developed a model to predict PM_{10} concentration and its relationship with NO_x . Their results revealed that the model did not have good potential to predict other important air pollutants. Another study predicted the concentration of PM_{10} using generalized linear models (GLMs). The study focused on the relationship between meteorological variables and air pollutant concentrations (Garcia et al., 2016). The concentration of PM_{10} was considered a dependent variable in GLM, while meteorological variables and gaseous pollutants were considered independent variables. He et al. (2018) presented two methods to predict $PM_{2.5}$, the linear method and the non-linear method of support vector regression.

Against this background, we used the Prophet forecasting model (PFM), which was developed by Facebook, concerning time forecasts with an anticipated and desired variable. The PFM has an extensive capacity to predict accurately without an overabundance of multifaceted or complex sources, and the model is also able to activate successfully even if the data have missing values and numerous outliers (Taylor and Letham, 2017). In addition, in the prediction of air pollution, PFM has been successfully established over other acceptable models, such as autoregressive integrated moving average (ARIMA) and seasonal autoregressive integrated moving average (SARIMA), which are widely used in many fields. Compared with ARIMA, PFM takes approximately 10 times less time to train (Scher and Messori, 2019). As mentioned earlier, compared with other models, PFM has a unique ability to predict accurately and effectively without the use of other parameters, which makes it a preferred method over current approaches to predict air pollution, and only a few studies have applied it (Shen et al., 2020). Comparatively, China's air pollution is different from air pollution in other places in the world, and the current work aimed to examine and predict the air quality in the most polluted region of China. In this research, we aimed to predict main air pollutants, PM_{10} , $PM_{2.5}$, SO_2 , NO_3 , CO , and O_3 , based on the Prophet forecasting model (PFM) for accurate short-term and long-term forecasting in Jiangsu Province, China, which will help support appropriate measures to control air pollution and plan accordingly in upcoming years.

MATERIALS AND METHODS

Study Area

Jiangsu Province, located at the Yangtze River in eastern China (Figure 1), is one of the most developed areas of the country. The

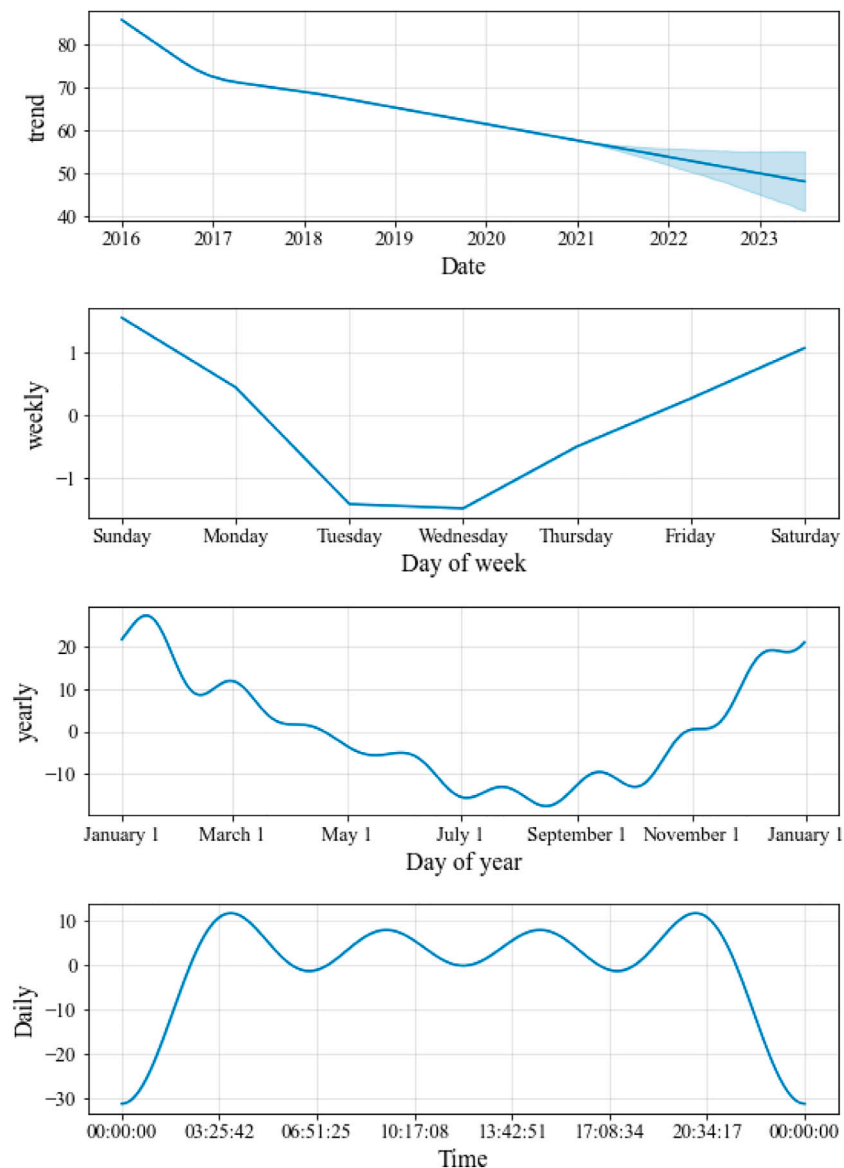


FIGURE 2 | Components of the Model ($PM_{2.5}$ trends graph, overall trend, yearly, weekly and daily).

province has a large population and well-developed industrial sectors. Occupying an area of around 107,200 km², it has 13 cities and, as of 2019, a population of approximately 80.5 million. In recent years, due to rapid economic development, Jiangsu Province has experienced the worst air quality (Zhang et al., 2020; Bhatti et al., 2022).

Data

Data on the daily average concentrations of air pollutant factors, including PM_{10} , $PM_{2.5}$, SO_2 , NO_2 , CO , and O_3 , were collected between 1 January 2016 and 31 December 2020. The website of historical data of air quality in China has allowed downloading of air pollution data since 13 May 2014 (Wang 2019). The data on the selected air pollutants were from the China Environmental Monitoring Station (CNEMC, 2019). In Jiangsu Province, 72

TABLE 1 | PM_{10} 2020 model performance statistics with different time period in Jiangsu Province.

Time	y	R	MSE	RMSE	MAE
15 Days	$0.31x+42.58$	0.31	835.63	28.91	24.25
1 Month	$0.36x+38.60$	0.41	570.36	23.88	19.27
3 Months	$0.32x+40.09$	0.36	433.39	20.82	17.46
6 Months	$0.39x+32.74$	0.36	287.41	16.95	13.33
1 Year	$0.43x+25.69$	0.40	268.03	16.37	11.74

monitoring stations that collect and record air quality data are scattered over 13 cities. The distribution of these monitoring stations is shown in **Figure 1**.

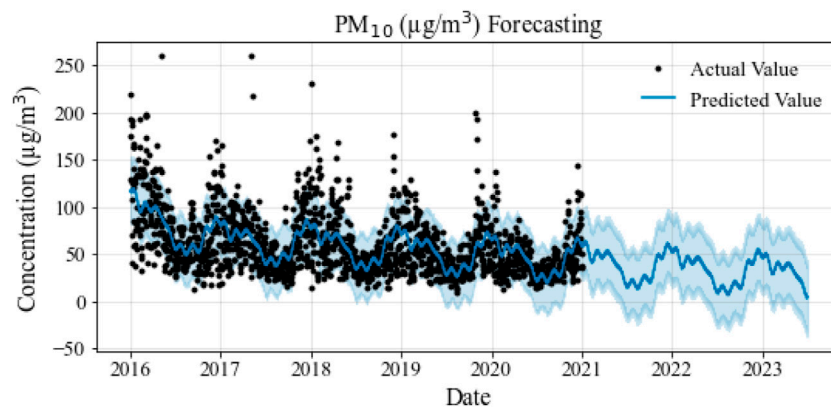


FIGURE 3 | PM₁₀ (µg/m³) forecasting in jiangsu province.

TABLE 2 | PM_{2.5} 2020 model performance statistics with different time period in Jiangsu Province.

Time	y	R	MSE	RMSE	MAE
15 Days	$0.27x+37.62$	0.39	632.79	25.16	21.30
1 Month	$0.34x+31.95$	0.45	428.40	20.70	16.44
3 Months	$0.40x+27.73$	0.52	273.27	16.53	13.18
6 Months	$0.50x+19.16$	0.51	162.03	12.73	9.50
1 Year	$0.52x+15.29$	0.52	145.76	12.07	8.22

Data Processing

To predict the concentrations of air pollutants, to meet the input requirements of the model, the values were manually processed in Excel for appropriate fitting and to satisfy the PFM formatting requirements, with timestamps in YYYY-MM-DD format. PFM has the broad ability to replace missing values with nearby points by fitting the model for the prediction, and any error was replaced with “NA” read from the devices denoted by “-1”.

Proposed Model

For analyzing time series and forecasting with trends, seasonality, and holidays, PFM is an influential and powerful tool for accurate

and effective prediction, and it takes only a few seconds to fit the model. The model uses the following formula:

$$y(t) = g(t) + s(t) + h(t) + \epsilon t. \quad (1)$$

To evaluate the performance of the PFM, **Eq 1** was employed, where $y(t)$ is the predicted value determined by a linear or logistic equation; $g(t)$ and $s(t)$ represent seasonality or time series based on yearly, monthly, daily or another period; $h(t)$ is the holiday outliers; and ϵt represents the unexpected error. For better understanding, the model has numerous parameters and the type of model can be assumed as linear or logistic. There is no maximum or minimum limit set in a linear model, while in a logistic model, the highest and lowest values are specified and used for saturated forecasts. Linear models were utilized to make sure the PFM reported and accounted for the typical outliers observed in air quality tendencies. To forecast and smooth time series data, the model accepts a Bayesian-based curve fitting technique, which is one of the model's most distinctive features compared to other forecasting models, such as ARIMA and the Holt-Winters method. PFM has more power to handle temporal patterns easily compared to traditional

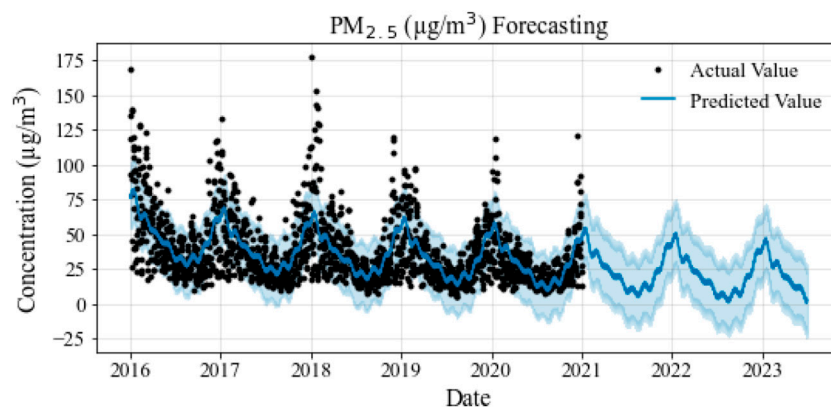


FIGURE 4 | PM_{2.5} (µg/m³) forecasting in jiangsu province.

TABLE 3 | SO₂ 2020 model performance statistics with different time period in Jiangsu Province.

Time	y	R	MSE	RMSE	MAE
15 Days	0.66x+4.21	0.16	7.59	2.76	2.32
1 Month	0.43x+5.11	0.10	6.77	2.60	2.22
3 Months	0.19x+5.86	0.04	5.93	2.44	2.15
6 Months	0.32x+5.18	0.08	3.63	1.91	1.52
1 Year	0.35x+4.10	0.17	2.46	1.57	1.16

exponential smoothing models and has requirements for frequent measurements (Taylor and Letham, 2017).

In PFM, change points are significant parameters and the explicit values of change points or the fitting scale can be specified; with higher change points, the model showed better performance when fitting the data. However, in predicting future trends, the model lost efficacy and effectiveness due to the use of fewer change points and was unable to be well-fitted. Initially, PFM plots a large value to determine the number of change points, and then the model uses L1 regularization to pick out a few points to use. To avoid overfitting due to a lack of change points, L1 regularization was used for this purpose, which picks out only the significant change points.

$$L(x, y) \equiv \sum_{i=1}^n (y_i - h_{\theta}(x_i))^2 + \lambda \sum_{i=1}^n |\theta_i| \quad (2)$$

Eq 2 represents L1 regularization, where x and y are the coordinates of the change points.

$\sum_{i=1}^n (y_i - h_{\theta}(x_i))^2$ represents the change between actual and forecasted value squared. The purpose of $\lambda \sum_{i=1}^n |\theta_i|$ is to maintain the balance in weights to avoid overfitting, where λ shows how much the weights are penalized; a high-value results in underfitting, while a low value means high bias. PFM has a value in determining the value of λ based on the number of estimators, or the average rate of change value can be specified.

Seasonality is another important parameter in predicting new values. PFM represents seasonality over daily, weekly, yearly, or

other periods (Figure 2). Figure 2 shows the seasonality of PM_{2.5} concentrations with daily, weekly, yearly, and overall time trends. PFM can also plot any custom seasonality or holiday component. Cross-validation was performed for the model to forecast the error given historic data. Data from the first 4 years (2016–2019) comprised the initial training data. PFM was used to forecast air quality for 2020 after training data from 2016 to 2019. To determine the model performance, these values were compared to the actual values. The model then predicted air quality for the upcoming 2.5 years.

Statistical Analysis

In this work, four metrics were employed to evaluate the model's performance: correlation coefficient (R), mean squared error (MSE), root mean squared error (RMSE), and mean absolute error (MAE). The correlation coefficient (R) was used to determine the degree of fit of the forecasted values to the overall actual values. MSE is the average change of actual and forecast values, and RMSE is the square root of MSE, while MAE is proposed as the average change between actual and predicted values. Compared with MAE, RMSE places higher importance on the variance between data outliers. These metrics are calculated with the following formulas:

$$R = \frac{n \sum_{i=1}^n (x_i \hat{x}_i) - (\sum_{i=1}^n x_i)(\sum_{i=1}^n \hat{x}_i)}{\sqrt{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2} \sqrt{n \sum_{i=1}^n \hat{x}_i^2 - (\sum_{i=1}^n \hat{x}_i)^2}} \quad (3)$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (x_i - \hat{x}_i)^2 \quad (4)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \hat{x}_i)^2} \quad (5)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |x_i - \hat{x}_i| \quad (6)$$

where x_i and \hat{x}_i are the actual and predicted values and n represents the number of samples.

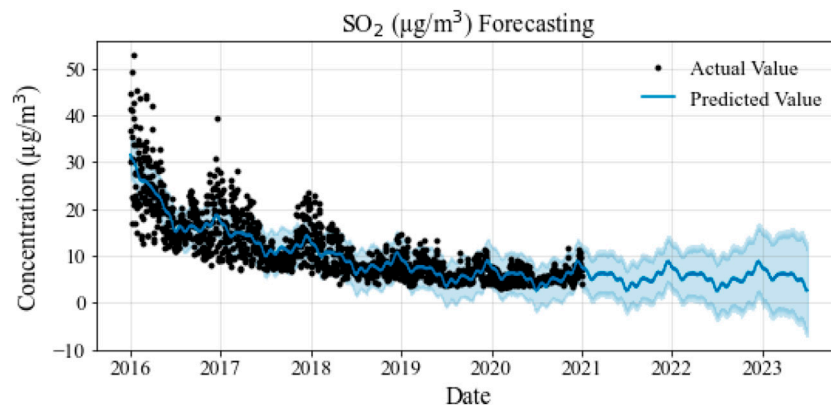
**FIGURE 5** | SO₂ (µg/m³) forecasting in jiangsu province.

TABLE 4 | NO₂ 2020 model performance statistics with different time period in Jiangsu Province.

Time	y	R	MSE	RMSE	MAE
15 Days	0.50 x +16.95	0.37	29.81	5.46	4.57
1 Month	0.52 x +16.60	0.68	58.58	7.65	6.61
3 Months	0.49 x +16.60	0.57	68.97	8.30	7.15
6 Months	0.49x+14.16	0.46	46.74	6.84	5.49
1 Year	0.53x+11.79	0.54	45.13	6.72	4.99

RESULTS AND DISCUSSION

PM₁₀ Prediction

A linear model was inputted to specify the features of PFM, and error and change points were determined using the LI regularization technique from PFM. The model accurately predicts PM₁₀ concentrations, and it is not overfitted during the entire year of 2020. We can see that all of the statistical indicators (R, MSE, RMSE, and MAE) show that the model predicts PM₁₀ concentration more accurately for the long-term prediction than short-term. In a 15 days prediction, the model predicted PM₁₀ with R, RMSE, and MAE values of 0.31, 28.91 µg/m³, and 24.25 µg/m³, respectively (Table 1). It can be seen that as the prediction time increased, the RMSE value decreased. The MAE value also showed a decreasing trend with the increase of time from 15 days to 1 year. In the 1 year prediction of PFM, the R, RMSE, and MAE values for PM₁₀ are 0.40, 16.37 µg/m³, and 11.74 µg/m³, respectively (Table 1). The decreasing error values show that the model predicts PM₁₀ with high accuracy as the time frame increases. Sayegh et al. (2014) developed five models to predict PM₁₀ concentration, while Ye (2019) utilized an ARIMA-PFM model to predict PM₁₀ concentrations in the time frame of 2 days. All indicators show that in the current work, PM₁₀ concentrations are forecast more effectively compared with those studies. The results of predicted PM₁₀ concentrations in Jiangsu Province are shown in Figure 3 and Table 1.

PM_{2.5} Prediction

To illustrate the performance of PFM, in a 15 days prediction the model predicted PM_{2.5} with R, RMSE, and MAE values of 0.39, 25.16 µg/m³, and 21.30 µg/m³, respectively (Table 2). The results demonstrate that as the time increased from 15 days to 1 year the model predicted PM_{2.5} more efficiently according to the statistical indicators shown in Table 2. In a 1-year prediction by PFM, the R, RMSE, and MAE values for PM_{2.5} were 0.52, 12.07 µg/m³, and 8.22 µg/m³, respectively (Table 2). The results indicate that the model appropriately forecast the concentration of PM_{2.5} in Jiangsu Province, as shown in Figure 4 and Table 2. Many studies have proposed methods for predicting PM_{2.5} concentrations (Li et al., 2015; Xi et al., 2015; Deters et al., 2017; Ye, 2019), and compared with these studies our study provides more accurate results for PM_{2.5} concentrations.

SO₂ Prediction

In the prediction of SO₂ concentrations by PFM, within the time frame of 15 days, the R, RMSE and MAE values for SO₂ are 0.16, 2.76 µg/m³, and 2.32 µg/m³, respectively (Table 3). Similarly particulate prediction, the performance of the model in predicting SO₂ concentrations improved with an increase of time as indicated by RMSE and MAE values. In the 1-year time frame, the model predicts SO₂ concentrations with R, RMSE, and MAE values of 0.17, 1.57 µg/m³, and 1.16 µg/m³, respectively (Table 3). In previous studies, Shaban et al. (2016) developed M5P model trees, artificial neural network (ANN), and support vector machine to predict short-term air pollution, and Ye (2019) proposed an ARIMA-PFM model to forecast air pollution. Compared with these studies, our results of RMSE and MAE values in the prediction of SO₂ concentrations are improved. The prediction results of PFM for SO₂ in Jiangsu Province are presented in Figure 5 and Table 3.

NO₂ Prediction

PFM provides better performance in predicting NO₂ concentrations in Jiangsu Province, with R, RMSE and MAE

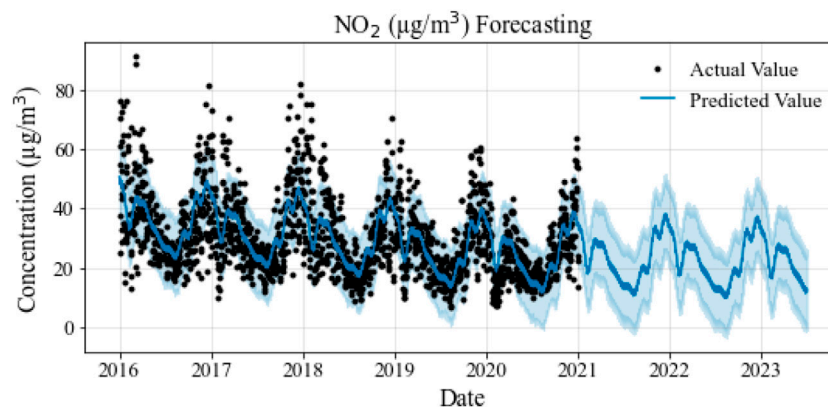
**FIGURE 6 |** NO₂ (µg/m³) forecasting in jiangsu province.

TABLE 5 | CO 2020 model performance statistics with different time period in Jiangsu Province.

Time	y	R	MSE	RMSE	MAE
15 Days	$0.22x+0.63$	0.26	0.04	0.20	0.17
1 Month	$0.32x+0.52$	0.39	0.03	0.17	0.14
3 Months	$0.39x+0.46$	0.46	0.03	0.17	0.15
6 Months	$0.43x+0.38$	0.34	0.02	0.14	0.11
1 Year	$0.49x+0.33$	0.38	0.01	0.12	0.09

values of 0.37, 5.46 $\mu\text{g}/\text{m}^3$, and 4.57 $\mu\text{g}/\text{m}^3$, respectively, for a period of 15 days (Table 4). PFM predicted NO_2 with an R of 0.68, which was the best performance in predicting NO_2 concentrations within the time frame of 1 month. The results indicate that in contrast to other air pollutants, the model provides better performance in short-term prediction with RMSE and MAE values for NO_2 . It can be seen that there were slight fluctuations in these values over an entire year (2020). In 1-year prediction by PFM, the R, RMSE, and MAE

values for NO_2 are 0.54, 6.72 $\mu\text{g}/\text{m}^3$, and 4.99 $\mu\text{g}/\text{m}^3$, respectively (Table 4). Comparing short- and long-term predictions, the MAE and RMSE values for NO_2 are similar in Jiangsu Province, with small differences (Figure 6 and Table 4).

CO Prediction

In predicting CO concentrations, PFM predicts with R, RMSE and MAE of 0.26, 0.20 $\mu\text{g}/\text{m}^3$, and 0.17 $\mu\text{g}/\text{m}^3$, respectively, in a period of 15 days (Table 5). The model predicts CO with R of 0.46 in 3-months prediction. Similar to other air pollutant parameters, the results show that the accuracy of the model improved with an increase of time for CO prediction. From 15 days to 1 year, the RMSE and MAE values gradually decreased from 0.20 to 0.12 $\mu\text{g}/\text{m}^3$ and 0.17 to 0.09 $\mu\text{g}/\text{m}^3$, respectively. With 1 year prediction, the model predicts with R = 0.38, RMSE = 0.12, and MAE = 0.09 in Jiangsu Province (Table 5). Overall, the results demonstrate that PFM provides significant results for CO concentrations over long-term prediction intervals in Jiangsu Province (Figure 7 and Table 5).

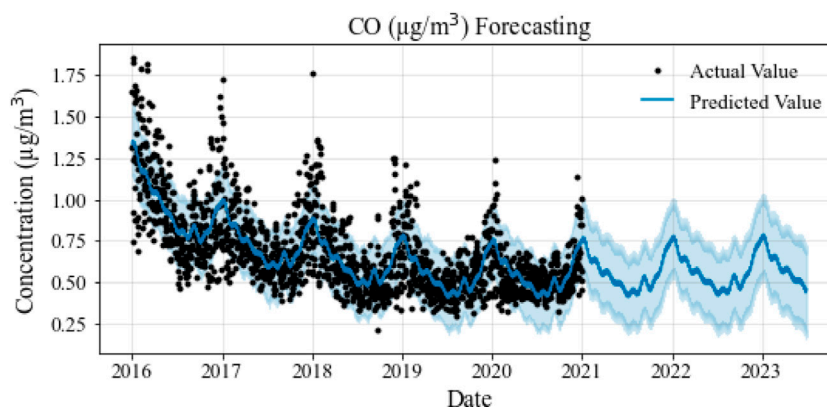
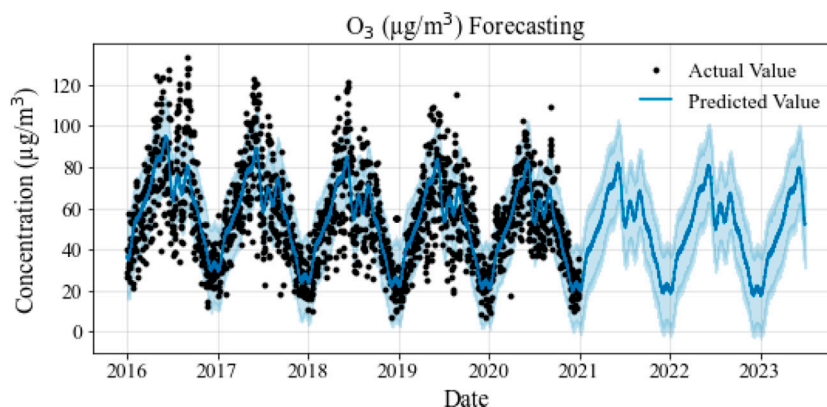
**FIGURE 7** | CO ($\mu\text{g}/\text{m}^3$) forecasting in jiangsu province.**FIGURE 8** | O₃ ($\mu\text{g}/\text{m}^3$) forecasting in jiangsu province.

TABLE 6 | O₃ 2020 model performance statistics with different time period in Jiangsu Province.

Time	y	R	MSE	RMSE	MAE
15 Days	0.32x+18.33	0.47	40.22	6.34	5.14
1 Month	0.44x+15.11	0.84	86.27	9.29	7.81
3 Months	0.60x+17.30	0.49	95.96	9.80	7.81
6 Months	0.74x+14.36	0.62	136.59	11.69	8.98
1 Year	0.73x+13.74	0.66	128.28	11.33	8.84

O₃ Prediction

PFM provides superior results in both short-term and long-term O₃ prediction in Jiangsu Province. In 15 days prediction, the model predicts with $R = 0.47$, $RMSE = 6.34 \mu\text{g}/\text{m}^3$, and $MAE = 5.14 \mu\text{g}/\text{m}^3$ for O₃ forecasting. PFM has the best R value for O₃ (0.84) in 1 month prediction. The smallest and highest RMSE and MAE values are 6.34 and 11.69 $\mu\text{g}/\text{m}^3$ and 5.14 and 8.98 $\mu\text{g}/\text{m}^3$, respectively. Compared with other air pollutant factor prediction, PFM has superior performance for O₃ prediction. In 1-year prediction by PFM, the R, RMSE, and MAE values for O₃ are 0.66, 11.33 $\mu\text{g}/\text{m}^3$, and 8.84 $\mu\text{g}/\text{m}^3$, respectively. We can see that all statistical measures demonstrate that the model has adequate performance. The forecasting values for O₃ and the actual and predicted values are significantly fitted. The prediction results of PFM for O₃ in Jiangsu Province are presented in **Figure 8** and **Table 6**.

CONCLUSION

To the best of our knowledge, this is the first provincial study that predicts six air pollutant parameters based on the prophet forecasting model (PFM). To control the threat and mitigate the hazardous effects of air pollution, a crucial step is to predict accurate air pollution over both short-term and long-term intervals. This allows policymakers and lower-level authorities to make strategies and plan accordingly to control air pollution as early as possible. In this study, the PFM is used to predict air pollution factors, including PM₁₀, PM_{2.5}, SO₂, NO₂, CO, and O₃, using 5 years of data in Jiangsu Province. The results demonstrate that the model has the unique ability to accurately forecast both short-term and long-term air quality, supporting its effectiveness. Few studies have reported the use of PFM to predict air pollution, and in the field of environmental modeling, applications of the model are still unexplored (Ye, 2019; Shan et al., 2020). This work illustrates that PFM has a wide ability to predict air pollution, and due to the fast-training time (approximately 10 times faster than ARIMA) and lack of a complex system, it can be applied to other regions. Compared with other models, as discussed in Section 2, the PFM has a unique ability to forecast air pollution and the model provides superior results in predicting air pollutants in Jiangsu Province. This study provides useful information and credible outcomes for the Chinese administration, scientific community and policymakers to

mitigate air pollution problems and to plan accordingly in upcoming years.

Policy Implications, Limitations and Future Research Directions

It is best to stop air pollution at its source, but until that day comes, experts suggest the following. Avoid spending time on busy roads and places with pollution. When you walk or bike away from congested streets, you can reduce exposure by half by using backstreet routes. Even on busy streets, cyclists experience less pollution than drivers. Scientists recommend parents cover their buggies with covers to protect their infants. Make sure you get to work early before rush hour begins and pollution levels rise. Exercise indoors or reduce strenuous outdoor exercise when air pollution is high or if you have a lung condition such as asthma.

- (1) Comprehensively implement pollution reduction. Coordinate and promote structural emission reduction, engineering emission reduction and management emission reduction: It is strictly forbidden to build new capacity projects in industries with severe overcapacity, adjust and optimize the industrial structure, and promote industrial transformation and upgrading. Eliminate outdated thermal power units and cement production capacity, strictly control the newly added emissions of sulfur dioxide and nitrogen oxides in the power industry, and simultaneously build and put into operation flue gas desulfurization and denitrification facilities for newly built coal-fired units. Focus on the reduction of emissions in the thermal power industry.
- (2) Strengthen coal control and management: Speed up the elimination of small coal-fired boilers. Strengthen the replacement of clean energy, vigorously develop cogeneration and regional central heating, and adopt methods such as “coal to gas”, “coal to electricity”, and “coal to biomass” to promote the elimination of small coal-fired boilers.
- (3) Strengthen the control of industrial air pollution: Promote the prevention and control of industrial pollution, and promote the replacement of old dust collectors with high-efficiency wet electrostatic and wet desulfurization dust collectors.
- (4) Completely ban the burning of crop straws and improve the comprehensive utilization rate of straws. Comprehensively improve the comprehensive utilization level of straw, and promote the demonstration projects of comprehensive utilization of straw, technologies such as straw returning to the field and wood replacement, and energy utilization such as curing and molding.

In current study, we used the PFM model to predict six air pollutants (PM₁₀, PM_{2.5}, SO₂, NO₂, CO, and O₃) in Jiangsu Province, China. The model shows good performance and provides better results in predicting air pollutant concentrations. The current research emphasizes air quality, the model should also be used in other fields as a forecasting method, such as in environmental economics to evaluate economic variations and the impact of climate change on the market. After the identification of the COVID-19 pandemic, strict restrictions and

control measures have been taken to control its rapid spread by the affected countries. Many scholars reported that due to strict restrictions and control actions, the air quality of these areas and regions was improved at a significant level (Li et al., 2020; Hasnain et al., 2021; Islam et al., 2021). Future research can be conducted to predict air quality trends during the COVID-19 pandemic and the current trends. The results of both periods can be compared to find out the changes and obtain new findings for policy implications. The study can be further extended to explore the impact of the COVID-19 pandemic on economic and industrial activities. Moreover, the current study focuses to predict air pollutants; future work can be conducted to predict meteorological factors. In the future, the model should also be used to predict the impact of climate change on agriculture production. The next step of the research is to extend the PFM to other fields and regions to obtain new findings, and it will be useful in areas such as effective preparation, health alarms for liable categories, reduced monitoring expenditures, etc.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: CNEMC (2019). China national

environmental monitoring centre. <http://www.cnemc.cn/>. Accessed 08 August 2019.

AUTHOR CONTRIBUTIONS

AH: Conceptualization, Methodology, Data curation, Writing - original draft, Writing - review & editing, Visualization; YS: Supervision, Conceptualization, Resources, Investigation, Project administration, Funding acquisition; MZH: Supervision, Investigation, Writing - review & editing; UAB: Investigation, Data curation, Writing - review & editing; AH: Data curation, Writing - review & editing; MH: Data curation, Writing - review & editing; SM: Data curation, Writing - review & editing; SUB: Data curation, Writing - review & editing; MAH: Data curation, Writing - review & editing; MS: Data curation, Writing - review & editing; RAW: Data curation, Writing - review & editing; YZ: Supervision, Conceptualization, Resources, Investigation.

FUNDING

This research was supported by the Key Fund of National Natural Science Foundation of China (grant No. 41631175).

REFERENCES

- Aamir, M., Li, Z., Bazai, S., Wagan, R. A., Bhatti, U. A., Nizamani, M. M., et al. (2021). Spatiotemporal Change of Air-Quality Patterns in Hubei Province-A Pre- to Post-COVID-19 Analysis Using Path Analysis and Regression. *Atmosphere* 12 (10), 1338. doi:10.3390/atmos12101338
- AirVisual (2019). Airvisual-air Quality Monitor and Information You Can Trust. Available at: <https://www.airvisual.com/> (Accessed Aug 26, 2019).
- AirVisual (2018). World Air Quality Report-Region & City pm2.5 Ranking. Tech. Rep.
- Appel, K. W., Napelenok, S. L., Foley, K. M., Pye, H. O. T., Hogrefe, C., Luecken, D. J., et al. (2017). Description and Evaluation of the Community Multiscale Air Quality (CMAQ) Modeling System Version 5.1. *Geosci. Model Dev.* 10, 1703–1732. doi:10.5194/gmd-10-1703-2017
- Bashir Shaban, K., Kadri, A., and Rezk, E. (2016). Urban Air Pollution Monitoring System with Forecasting Models. *IEEE Sensors J.* 16 (8), 2598–2606. doi:10.1109/jsen.2016.2514378
- Bhatti, U. A., Huang, M., Wu, D., Zhang, Y., Mehmood, A., and Han, H. (2019). Recommendation System Using Feature Extraction and Pattern Recognition in Clinical Care Systems. *Enterp. Inf. Syst.* 13 (3), 329–351. doi:10.1080/17517575.2018.1557256
- Bhatti, U. A., Yan, Y., Zhou, M., Ali, S., Hussain, A., Qingsong, H., et al. (2021). Time Series Analysis and Forecasting of Air Pollution Particulate Matter (PM2.5): An SARIMA and Factor Analysis Approach. *IEEE Access* 9, 41019–41031. doi:10.1109/access.2021.3060744
- Bhatti, U. A., Zeeshan, Z., Nizamani, M. M., Bazai, S., Yu, Z., and Yuan, L. (2022). Assessing the Change of Ambient Air Quality Patterns in Jiangsu Province of China Pre-to Post-COVID-19. *Chemosphere* 288, 132569. doi:10.1016/j.chemosphere.2021.132569
- Bilal, M., Mhawish, A., Nichol, J. E., Qiu, Z., Nazeer, M., Ali, M. A., et al. (2021). Air Pollution Scenario over Pakistan: Characterization and Ranking of Extremely Polluted Cities Using Long-Term Concentrations of Aerosols and Trace Gases. *Remote Sen. Environ.* 264, 112617. doi:10.1016/j.rse.2021.112617
- Cabaneros, S. M., Calautit, J. K., and Hughes, B. R. (2019). A Review of Artificial Neural Network Models for Ambient Air Pollution Prediction. *Environ. Mod. Soft.* 119, 285–304. Available at: <http://www.sciencedirect.com/science/article/pii/S1364815218306352>. doi:10.1016/j.envsoft.2019.06.014
- Chen, D., and Li, R. Y. M. (2022). “Predicting Housing Price in Beijing via Google and Microsoft AutoML,” in *Current State of Art in Artificial Intelligence and Ubiquitous Cities* (Singapore: Springer), 105–115. doi:10.1007/978-981-19-0737-1_7
- Choi, H. D., Liu, H., Crawford, J. H., Considine, D. B., Allen, D. J., Duncan, B. N., et al. (2017). Global O₃-CO Correlations in a Chemistry and Transport Model during July-August: Evaluation with TES Satellite Observations and Sensitivity to Input Meteorological Data and Emissions. *Atmos. Chem. Phys.* 17 (13), 8429–8452. doi:10.5194/acp-17-8429-2017
- CNEMC (2019). China National Environmental Monitoring Centre. Available at: <http://www.cnemc.cn/> (Accessed Aug 08, 2019).
- Conibear, L., Butt, E. W., Knote, C., Arnold, S. R., and Spracklen, D. V. (2018). Residential Energy Use Emissions Dominate Health Impacts from Exposure to Ambient Particulate Matter in India. *Nat. Commun.* 9, 617. doi:10.1038/s41467-018-02986-7
- Dai, Q., Bi, X., Liu, B., Li, L., Ding, J., Song, W., et al. (2018). Chemical Nature of PM_{2.5} and PM₁₀ in Xi'an, China: Insights into Primary Emissions and Secondary Particle Formation. *Environ. Pollut.* 240, 155–166. doi:10.1016/j.envpol.2018.04.111
- Dehghani, M., Keshtgar, L., Javaheri, M. R., Zahra, D., Gea, O. C., Pietro, Z., et al. (2017). The Effects of Air Pollutants on the Mortality Rate of Lung Cancer and Leukemia. *Mol. Med. Rep.* 15, 3390–3397. doi:10.3892/mmr.2017.6387
- Deng, M., Xu, F., and Wang, H. (2018). Prediction of Hourly pm_{2.5} Using a Space-Time Support Vector Regression Model. *Atmos. Environ.* 181, 12–19. Available at: <http://www.sciencedirect.com/science/article/pii/S1352231018301535>. doi:10.1016/j.atmosenv.2018.03.015
- Deters, J. K., Zalakeviciute, R., Gonzalez, M., and Rybarczyk, Y. (2017). Modeling PM_{2.5} Urban Pollution Using Machine Learning and Selected Meteorological Parameters. *J. Electr. Comput. Eng.* 2017, 1–14. doi:10.1155/2017/5106045

- Duan, W., Wang, X., Cheng, S., Wang, R., and Zhu, J. (2021). Influencing Factors of PM_{2.5} and O₃ from 2016 to 2020 Based on DLNM and WRF-CMAQ. *Environ. Pollut.* 285, 117512. doi:10.1016/j.envpol.2021.117512
- Fajersztajn, L., Veras, M., Barrozo, L. V., and Saldiva, P. (2013). Air Pollution: A Potentially Modifiable Risk Factor for Lung Cancer. *Nat. Rev. Cancer* 13, 674–678. doi:10.1038/nrc3572
- Fan, H., Zhao, C., Ma, Z., and Yang, Y. (2020). Atmospheric Inverse Estimates of CO Emissions from Zhengzhou, China. *Environ. Pollut.* 267, 115164. doi:10.1016/j.envpol.2020.115164
- Fuller, G. W., Carslaw, D. C., and Lodge, H. W. (2002). An Empirical Approach for the Prediction of Daily Mean PM₁₀ Concentrations. *Atmos. Environ.* 36 (9), 1431–1441. doi:10.1016/s1352-2310(01)00580-5
- Garcia, J. M., Teodoro, F., Cerdeira, R., Coelho, L. M. R., Kumar, P., and Carvalho, M. G. (2016). Developing a Methodology to Predict PM₁₀ Concentrations in Urban Areas Using Generalized Linear Models. *Environ. Technol.* 37 (18), 2316–2325. doi:10.1080/09593330.2016.1149228
- Guo, L., Chen, B., Zhang, H., and Zhang, Y. (2020). A New Approach Combining a Simplified FLEXPART Model and a Bayesian-RAT Method for Forecasting PM₁₀ and PM_{2.5}. *Environ. Sci. Pollut. Res.* 27, 2165–2183. doi:10.1007/s11356-019-06605-w
- Hasnain, A., Hashmi, M. Z., Bhatti, U. A., Nadeem, B., Wei, G., Zha, Y., et al. (2021). Assessment of Air Pollution before, during and after the COVID-19 Pandemic Lockdown in Nanjing, China. *Atmosphere* 12, 743. doi:10.3390/atmos12060743
- He, B., Heal, M. R., and Reis, S. (2018). Land-Use Regression Modelling of Intra-Urban Air Pollution Variation in China: Current Status and Future Needs. *Atmosphere* 9 (4), 134.
- Huang, L., Zhou, L., Chen, J., Chen, K., Liu, Y., Chen, X., et al. (2016). Acute Effects of Air Pollution on Influenza-Like Illness in Nanjing, China: A Population-Based Study. *Chemosphere* 147, 180–187. doi:10.1016/j.chemosphere.2015.12.082
- Islam, S., Tusher, T. R., Roy, S., and Rahman, M. (2021). Impacts of Nationwide Lockdown Due to COVID-19 Outbreak on Air Quality in Bangladesh: A Spatiotemporal Analysis. *Air Qual. Atmos. Heal.* 14, 351–363. doi:10.1007/s11869-020-00940-5
- Ivanov, A., Voynikova, D., Stoimenova, M., Gocheva-Ilieva, S., and Iliev, I. (2018). Random Forests Models of Particulate Matter PM₁₀: a Case Study. *Am. Institue Phys. Conf. Proc.* 2025 (1), 162–166. doi:10.1063/1.5064879
- Kamińska, J. A. (2018). The Use of Random Forests in Modelling Short-Term Air Pollution Effects Based on Traffic and Meteorological Conditions: A Case Study in Wrocław. *J. Environ. Manage.* 217, 164–174. doi:10.1016/j.jenvman.2018.03.094
- Lee, M., Lin, L., Chen, C. Y., Tsao, Y., Yao, T. H., Fei, M. H., et al. (2020). Forecasting Air Quality in Taiwan by Using Machine Learning. *Sci. Rep.* 10, 4153. doi:10.1038/s41598-020-61151-7
- Li, H., and Shi, X. (2016). Data Driven Based PM_{2.5} Concentration Forecasting. *Adv. Biol. Sci. Res.* 3, 301–304. doi:10.2991/bep-16.2017.64
- Li, L., Li, Q., Huang, L., Wang, Q., Zhu, A., Xu, J., et al. (2020). Air Quality Changes during the COVID-19 Lockdown over the Yangtze River Delta Region: An Insight into the Impact of Human Activity Pattern Changes on Air Pollution Variation. *Sci. Total Environ.* 732, 139282. doi:10.1016/j.scitotenv.2020.139282
- Li, Y., Chen, Q., Zhao, H., Wang, L., and Tao, R. (2015). Variations in PM₁₀, PM_{2.5} and PM_{1.0} in an Urban Area of the Sichuan Basin and Their Relation to Meteorological Factors. *Atmosphere* 6 (1), 150–163. doi:10.3390/atmos6010150
- Liu, C., Yin, P., Chen, R., Meng, X., Wang, L., Niu, Y., et al. (2018). Ambient Carbon Monoxide and Cardiovascular Mortality: a Nationwide Time-Series Analysis in 272 Cities in China. *Lancet. Planet. Health.* 2 (1), e12. doi:10.1016/S2542-5196(17)30181-X
- Liu, D., and Sun, K. (2019). Short-Term PM_{2.5} Forecasting Based on CEEMD-RF in Five Cities of China. *Environ. Sci. Pollut. Res.* 26, 32790–32803. doi:10.1007/s11356-019-06339-9
- Maleki, H., Sorooshian, A., Goudarzi, G., Baboli, Z., Tahmasebi Birgani, Y., and Rahmati, M. (2019). Air Pollution Prediction by Using an Artificial Neural Network Model. *Clean. Technol. Environ. Policy* 21 (6), 1341–1352. doi:10.1007/s10098-019-01709-w
- Pasero, E., and Mesin, L. (2010). *Artificial Neural Networks for Pollution Forecast*. INTECH Open Access Publisher.
- Petetin, H., Sauvage, B., Smit, H. G. J., Gheusi, F., Lohou, F., Blot, R., et al. (2018). A Climatological View of the Vertical Stratification of RH, O₃ and CO within the PBL and at the Interface with Free Troposphere as Seen by IAGOS Aircraft and Ozonesondes at Northern Mid-Latitudes over 1994–2016. *Atmos. Chem. Phys.* 18 (13), 9561–9581. doi:10.5194/acp-18-9561-2018
- Qayyum, M., Ali, M., Nizamani, M. M., Li, S., Yu, Y., and Jahanger, A. (2021). Nexus between Financial Development, Renewable Energy Consumption, Technological Innovations and CO₂ Emissions: The Case of India. *Energies* 14 (15), 4505. doi:10.3390/en14154505
- Qayyum, M., Yu, Y., Nizamani, M. M., Raza, S., Ali, M., and Li, S. (2022). Financial Instability and CO₂ Emissions in India: Evidence from ARDL Bound Testing Approach. *Energy Environ.*, 0958305X2110650. doi:10.1177/0958305X211065019
- Salonen, H., Salthammer, T., and Morawska, L. (2019). Human Exposure to NO₂ in School and Office Indoor Environments. *Environ. Int.* 130, 104887. doi:10.1016/j.envint.2019.05.081
- Sayegh, A. S., Munir, S., and Habeebullah, T. M. (2014). Comparing the Performance of Statistical Models for Predicting PM₁₀ Concentrations. *Aerosol Air Qual. Res.* 14 (3), 653–665. doi:10.4209/aaqr.2013.07.0259
- Scher, S., and Messori, G. (2019). How Global Warming Changes the Difficulty of Synoptic Weather Forecasting. *Geophys. Res. Lett.* 46 (5), 2931–2939. doi:10.1029/2018gl081856
- Shairsingh, K. K., Brook, J. R., Mihele, C. M., and Evans, G. J. (2021). Characterizing Long-Term NO₂ Concentration Surfaces across a Large Metropolitan Area through Spatiotemporal Land Use Regression Modelling of Mobile Measurements. *Environ. Res.* 196, 111010. doi:10.1016/j.envres.2021.111010
- Shan, Y., Wang, X., Wang, Z., Liang, L., Li, J., and Sun, J. (2020). The Pattern And Mechanism of Air Pollution in Developed Coastal Areas of China: From the Perspective of Urban Agglomeration. *PLoS One* 15 (19), e0237863.
- Shen, J., Valagolam, D., and McCalla, S. (2020). Prophet Forecasting Model: a Machine Learning Approach to Predict the Concentration of Air Pollutants (PM_{2.5}, PM₁₀, O₃, NO₂, SO₂, CO) in Seoul, South Korea. *PeerJ* 8, e9961. doi:10.7717/peerj.9961
- Su, X., Tie, X., Li, G., Cao, J., Huang, R., Feng, T., et al. (2017). Effect of Hydrolysis of N₂O₅ on Nitrate and Ammonium Formation in Beijing China: WRF-Chem Model Simulation. *Sci. Total Environ.* 579, 221–229. doi:10.1016/j.scitotenv.2016.11.125
- Tang, G., Zhang, J., Zhu, X., Song, T., Munkel, C., Hu, B., et al. (2016). Mixing Layer Height and its Implications for Air Pollution over Beijing, China. *Atmos. Chem. Phys.* 16, 2459–2475. doi:10.5194/acp-16-2459-2016
- Taylor, S. J., and Letham, B. (2017). Forecasting at Scale. *Am. Statistician* 72 (1), 37–45. doi:10.1080/00031305.2017.1380080
- The State Council of China (2013). Air Pollution Prevention and Control Action Plan. Available at: http://www.gov.cn/jrzq/2013-09/12/content_2486918.htm. (Accessed January 2, 2022).
- Wang, H., Gao, Z., Ren, J., Liu, Y., Chang, L. T.-C., Cheung, K., et al. (2019). An Urban-Rural and Sex Differences in Cancer Incidence and Mortality and the Relationship with PM_{2.5} Exposure: An Ecological Study in the Southeastern Side of Hu Line. *Chemosphere* 216, 766–773. doi:10.1016/j.chemosphere.2018.10.183
- Wang, J., He, L., Lu, X., Zhou, L., Tang, H., Yan, Y., et al. (2022). A Full-Coverage Estimation of PM_{2.5} Concentrations Using a Hybrid XGBoost-WD Model and WRF-Simulated Meteorological Fields in the Yangtze River Delta Urban Agglomeration, China. *Environ. Res.* 203, 111799. doi:10.1016/j.envres.2021.111799
- Wang, X. (2019). Historical Data of Air Quality in China. Available at: <http://beijingair.sinaapp.com/> (Accessed Aug 04, 2019).
- WHO-World Health Organization (2018). Ambient (Outdoor) Air Quality and Health. Available at: <http://www.who.int/mediacentre/factsheets/fs313/en/> (Accessed Jan 6, 2018).
- Wu, X., Guo, J., Wei, G., and Zou, Y. (2020). Economic Losses and Willingness to Pay for Haze: The Data Analysis Based on 1123 Residential Families in Jiangsu Province, China. *Environ. Sci. Pollut. Res.* 27, 17864–17877. doi:10.1007/s11356-020-08301-6

- Xi, X., Wei, Z., Xiaoguang, R., Yijie, W., Xinxin, B., Wenjun, Y., et al. (2015). "A Comprehensive Evaluation of Air Pollution Prediction Improvement by a Machine Learning Method," in 2015 IEEE International Conference on Service Operations and Logistics, And Informatics (SOLI), Yasmine Hammamet, Tunisia, 15-17 November 2015 (Piscataway: IEEE). doi:10.1109/soli.2015.7367615
- Yang, B., Jahanger, A., and Ali, M. (2021). Remittance Inflows Affect the Ecological Footprint in BICS Countries: Do Technological Innovation and Financial Development Matter? *Environ. Sci. Pollut. Res. Int.* 28 (18), 23482–23500. doi:10.1007/s11356-021-12400-3
- Yang, B., Ali, M., Hashmi, S. H., and Jahanger, A. (2022). Do Income Inequality and Institutional Quality Affect CO2 Emissions in Developing Economies? *Environ. Sci. Pollut. Res.* 29, 42720–42741. doi:10.1007/s11356-021-18278-5
- Ye, Z. (2019). Air Pollutants Prediction in Shenzhen Based on Arima and Prophet Method. *E3S Web Conf.* 136, 05001. doi:10.1051/e3sconf/201913605001
- Zhang, H., Di, B., Liu, D., Li, J., and Zhan, Y. (2019). Spatiotemporal Distributions of Ambient SO₂ across China Based on Satellite Retrievals and Ground Observations: Substantial Decrease in Human Exposure during 2013-2016. *Environ. Res.* 179, 108795. doi:10.1016/j.envres.2019.108795
- Zhang, T., Liu, P., Sun, X., Zhang, C., Wang, M., Xu, J., et al. (2020). Application of an Advanced Spatiotemporal Model for PM_{2.5} Prediction in Jiangsu Province, China. *Chemosphere* 246, 125563. doi:10.1016/j.chemosphere.2019.125563
- Zhao, Z., Qin, J., He, Z., Li, H., Yang, Y., and Zhang, R. (2020). Combining Forward with Recurrent Neural Networks for Hourly Air Quality Prediction in Northwest of China. *Environ. Sci. Pollut. Res.* 27, 28931–28948. doi:10.1007/s11356-020-08948-1
- Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.
- Copyright © 2022 Hasnain, Sheng, Hashmi, Bhatti, Hussain, Hameed, Marjan, Bazai, Hossain, Sahabuddin, Wagan and Zha. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.