IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received March 31, 2022, accepted May 4, 2022, date of publication May 9, 2022, date of current version May 16, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3173734

Revealing Influence of Meteorological Conditions on Air Quality Prediction Using Explainable Deep Learning

YUTING YANG[1], GANG MEI[1], AND STEFANO IZZO[2]
[1]School of Engineering and Technology, China University of Geosciences, Beijing 100083, China
[2]Department of Mathematics and Applications, University of Naples Federico II, 80126 Naples, Italy

Corresponding author: Gang Mei (gang.mei@cugb.edu.cn)

This work was supported in part by the 2021 Graduate Innovation Fund Project of China University of Geosciences, Beijing, under Grant ZD2021YC009; and in part by the Major Program of Science and Technology of Xinjiang Production and Construction Corps under Grant 2020AA002.

ABSTRACT Meteorological conditions have a strong influence on air quality and can play an important role in air quality prediction. However, due to the ''black-box'' nature of deep learning, it is difficult to obtain trustworthy deep learning models when considering meteorological conditions in air quality prediction. To address the above problem, in this paper, we reveal the influence of meteorological conditions on air quality prediction by utilizing explainable deep learning. In this paper, (1) the source data from air pollutant datasets, including $PM_{2.5}$, PM_{10}, SO_2 hourly concentration, and the meteorological condition datasets measuring the temperature, humidity, and atmospheric pressure are obtained; (2) the Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) models are established for air quality prediction in 4 conditions; (3) the SHapley Additive exPlanation (SHAP) method is employed to analyze the explainability of the air quality prediction models. We find that the prediction accuracy is not improved by considering only meteorological conditions. However, when combining meteorological conditions with other air pollutants, the prediction accuracy is higher than considering other air pollutants. In addition, the largest contribution to air quality prediction is atmospheric pressure, followed by humidity and temperature. The reason for the different accuracies of the prediction may because of the interaction between meteorological conditions and other air pollutants. The investigated results in this paper can help improve the prediction accuracy of air quality and achieve trusted air quality predictions.

INDEX TERMS Explainable deep learning, air quality prediction, meteorological condition, long short-term memory (LSTM), gate recurrent unit (GRU).

I. INTRODUCTION

The continuous acceleration of global urbanization and industrialization has brought environmental problems. One of the serious environmental problems is air quality induced by the development of urbanization and industrialization [1], [2]. Due to the needs of transportation, production, and life, energy production and consumption processes, such as power plants, factories, and automobile exhaust emissions have ultimately led to the continuous deterioration of global air quality [3]. Air pollution can cause various respiratory diseases and may even lead to the occurrence of cancer, which seriously threatens people's lives and health [4].

The main air pollutants include $PM_{2.5}$, PM_{10}, and SO_2, etc. Among them, $PM_{2.5}$ is a fine particle with a diameter smaller than 2.5 microns. Compared with larger particulate pollutants, $PM_{2.5}$ particles are more active, meaning that they can easily carry substances that affect human health and the environment, as well as remain in the air for a long time and spread quickly. $PM_{2.5}$ is one of the most important sources of air pollution [2]. Due to its small particle size, it can enter the nasal cavity and throat of the human body, and then easily cause asthma, bronchial or cardiovascular diseases [5]. Air pollution poses a great threat to people's health [4], [6]. Being in an environment with severe air pollution

The associate editor coordinating the review of this manuscript and approving it for publication was Liandong Zhu.

for a long time may cause various respiratory diseases and even decreased cardiopulmonary function problems. The incidence of various diseases will dramatically increase, which will overdraft people's health, affect people's living and happiness indices, and increase mortality [7]. Air pollution also damages the ecosystem, affects its diversity and stability, and harms the environment [1].

Frequent air pollution incidents not only cause serious harm to human health but also cause huge economic losses and many social problems [8]. Therefore, based on air pollution parameters, timely scientific analysis, accurate prediction of air quality and effective protection and treatment can help relevant departments and related groups take preventive measures in advance, as well as more reasonably arrange travel. People's health could be ensured, and the occurrence of diseases could be prevented [9].

In addition, the prediction of air quality can also provide reliable information for the prevention and control of air pollution. Through further understanding of the influencing factors and changing trends of air pollutants, effective evaluation and prediction of air quality changes are helpful for the control and prevention of air pollution, which would then enable the environment and human health to be better protected [10]. Air quality prediction is also conducive to relevant departments to understand the air quality status, and thus, a valuable theoretical basis can be provided for it. In addition, air pollution prevention and control policies can be formulated according to specific conditions. It also provides constructive opinions and suggestions for decision-makers to take more economical and efficient measures to improve air quality in the future [11].

Some of the related research work is as follows. For example, Kumar and Goyal [12] employed three statistical models: auto regressive integrated moving average (ARIMA), principal component regression (PCR), and combination of both to predict the daily Air Quality Index (AQI) for each season. Cobourn [13] proposed an enhancement $PM_{2.5}$ prediction model on the basis of back-trajectory concentrations and nonlinear regression (NLR), which has lower mean absolute errors. The process of predicting air quality using physical and chemical methods is more complex. Rajput *et al.* [14] proposed an approach to assess and represent quality status through an AQI, which can be useful for better forecasting of air quality parameters. But the model is a time-frequency prediction model, which is better for shorter time periods only.

Moreover, Liu *et al.* [15] predicted PM_{10}, NO_2, and SO_2 in seven locations in Guangzhou based on sample selection and backpropagation (BP) neural network, and achieved satisfactory prediction results. It can provide reliable and accurate air quality prediction and warnings in practical applications. Complex correlation calculations were performed when considering the influence of meteorological conditions. Xiao *et al.* [16] collected meteorological data from 1980 to 2012 in Baoji, China, to investigate the trend of visibility changes. $PM_{2.5}$ was measured, and the influencing factors and reasons for visibility reduction were analyzed based on the improved equation. But only past trends in visibility changes were analyzed, and no projections of future trends were made. Qi *et al.* [8] analyzed the relationship between the meteorological conditions and the concentration of air pollutants in Beijing. It was demonstrated that meteorological conditions have a corresponding effect on air quality. However, the air quality was not predicted, and the effect of meteorological conditions on the air quality prediction was not studied.

Air quality prediction can be performed using the traditional statistical methods and traditional machine learning methods as well as the latest deep learning methods. Traditional statistical methods and traditional machine learning methods generally use air quality data from a period of time in the past and predict air quality for a period of time in the future based on the characteristic connections and association rules between the data, which are generally more complex, computationally intensive and inefficient [17], [18]. Studies on the influence of meteorological conditions on air quality have generally considered only the relationship between air quality and the measured meteorological conditions, and rarely consider the effects on air quality in future trends.

In recent years, deep learning has been widely employed in various fields [19], [20]. Deep learning has advantages in capturing complex data relationships. Moreover, because of the large amount of historical monitoring data for air quality, applying deep learning methods to air quality prediction does not require extremely detailed pollution emission parameters and meteorological data, which can reduce the computational effort and improve the computational efficiency [21]–[23].

For example, Kuo *et al.* [24] applied deep learning methods (Recurrent Neural Network) to predict air quality in Taipei, Taiwan. The results indicated that the RNN using the Gaussian process is better than the backpropagation neural network and the basic RNN. Although meteorological conditions are taken into account in air quality predictions, the effect of meteorological conditions on air quality predictions cannot be explained. Chang *et al.* [25] proposed an aggregate long short-term memory model (ALSTM). The results showed that the aggregation model can improve the accurate of prediction effectively. There are still some shortcomings in the data pre-processing which leads to some problems in the accuracy of $PM_{2.5}$ prediction. Zhang *et al.* [26] suggested a semi-supervised model including empirical mode decomposition (EMD) and bidirectional long short-term memory (BiLSTM) neural networks to predict $PM_{2.5}$ concentration. It reduces the accumulation of errors in multi-step $PM_{2.5}$ prediction and can achieve a higher accuracy rate. However, the predictions are made through $PM_{2.5}$ levels only, ignoring the influence of other factors on the predicted results.

Deep learning methods are suitable for air quality prediction, and the research of interpretable deep learning methods has far-reaching implications for future relevance research in air quality prediction. However, deep learning methods lack explainability due to their "black box" nature. To build trustworthy deep learning models, much research work has been conducted on the explainability of deep learning.

For example, Lundberg and Lee [27] proposed the Shapley additive explanation (SHAP) method that can explain the contribution of each feature in the machine learning model to each predicted value. SHAP method can explain complex machine learning models. However, only the SHAP method is proposed, and the application of the SHAP method to air quality prediction is not implemented. Navares and Aznarte [28] employed the LSTM to predict air quality in the Madrid region. And the comprehensive deep network configuration model was used to predict, indicating that a single comprehensive model might be a better option than multiple individual models. Although meteorological conditions are taken into account when predicting air quality, there is no corresponding analysis of how meteorological conditions affect air quality predictions.

In addition, Arrieta *et al.* [29] summarized the existing literature and related research in the XAI field, provided a new definition of interpretable machine learning, and a more comprehensive interpretable machine learning classification method. There is no mention of deep learning interpretable analysis operations in terms of the influence of meteorological conditions on air quality prediction. Zhang *et al.* [30] used partial correlation diagram (PDP), SHAP method, linear regression (LR), and decision tree (DT) methods to study the interpretability of machine learning models on the thermal comfort of smart buildings. It can be extended to explainable deep learning analysis for prediction of air quality.

However, currently, while several deep learning models utilize meteorological conditions for air quality prediction, meteorological conditions are only used as input data, and there is little research work on the influence of meteorological conditions on air quality prediction. In this case, the influence of meteorological conditions on air quality prediction in deep learning models is not yet well understood, such as how it affects air quality prediction. This is because the deep learning model has the common "black box" nature, i.e., the weak explainability. Although it is possible to combine meteorological condition data with air quality data, and then use the deep learning model's powerful fitting advantage for complex data relationships to predict air quality. There are still many difficulties in analyzing the influence of meteorological condition data on air quality prediction and their correlations.

To address the above problems, in this paper, we reveal the impact of meteorological conditions on air quality prediction using explainable deep learning and explain how meteorological conditions affect air quality prediction accordingly. By revealing the influence of meteorological conditions on the prediction of air quality, the accuracy is further improved. Deep learning models for air quality prediction with higher accuracy and credibility can be obtained. Thus, it can be better applied in practice. This can help people plan their travel arrangements reasonably and take corresponding preventive measures on time to protect their health. Through the advanced understanding of the air quality status, corresponding prevention and control measures are adopted to realize timely and effective environmental management.

The contributions of this paper could be summarized as follows.

1) LSTM and GRU models were employed for air quality prediction under four different conditions and achieved positive results.
2) The influence of meteorological conditions on air quality prediction is revealed using explainable deep learning methods.
3) How meteorological conditions affect air quality prediction is revealed using the SHAP method, which is beneficial to further improve the accuracy of air quality prediction.

The rest of this paper is organized as follows. Section II describes the dataset and methods used in this article. Section III analyzes the results of the air quality prediction and uses the SHAP method to conduct an explainability analysis of the deep learning model. Section IV discusses the use of explainable deep learning methods as well as possible future work. Section V is a conclusion of the paper.

II. MATERIALS AND METHODS

A. OVERVIEW

In this paper, we employ the explainable deep learning method SHapley Additive exPlanations to reveal the impact of meteorological conditions on air quality prediction. Our objective is to discuss the influence of meteorological conditions on air quality prediction to improve the air quality prediction accuracy of the deep learning model; see the workflow of the research work in Figure 1.

First, we collect a large amount of hourly concentration data of air pollutants, as well as meteorological data from the same location and time from publicly available websites. The raw data is cleaned up. Second, deep learning models suitable for air quality prediction are established. Since both LSTM and GRU models are classical time-series prediction models and compared with Recurrent Neural Network (RNN) models, LSTM models can learn long-term dependence information, and GRU models are variants of LSTM models, which are simpler in structure compared with LSTM models and can also achieve better results. Both models are RNN-based versions. Instead of using the original RNN model, the research work uses the more effective LSTM and GRU to establish deep learning models (stand-alone LSTM model and stand-alone GRU model) for air quality prediction. The prediction first separately considers meteorological conditions and other air pollutants, before combining the meteorological conditions and the other air pollutants for comparison and analysis. Third, the SHapley Additive exPlanations (SHAP) method conducts a single analysis and an analysis for all features of the deep learning model. Additionally, the influence of meteorological conditions on

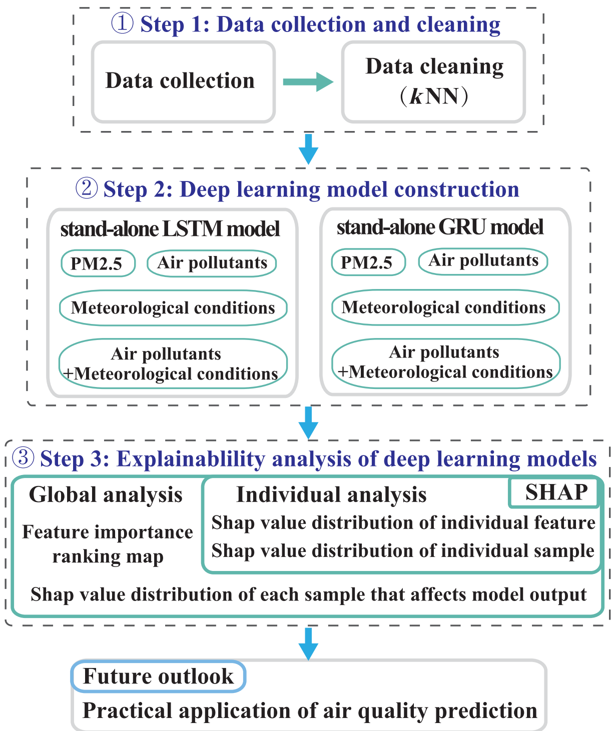


FIGURE 1. Flowchart of revealing the influence of meteorological conditions on air quality prediction by explainable deep learning.

air quality prediction is revealed. The foundation for the practical application of air quality prediction is expressed.

B. STEP 1: DATA COLLECTION AND CLEANING

The dataset for this research work is from the Harvard Dataverse [31], including hourly average concentrations of regulated air pollutants data collected from 35 air quality monitoring stations and hourly average meteorological parameter data collected from 18 meteorological stations in Beijing from January 30th, 2017 to January 31st, 2018. Air quality data is provided by the Ministry of Environmental Protection (MEP) of China. Hourly averaged meteorological data in the same period was accessed from The National Oceanic and Atmospheric Administration (NOAA).

The obtained dataset name and available URL are as follows:

Dataset: Air pollution and meteorological data in Beijing 2017–2018.

URL: https://doi.org/10.7910/DVN/USXCAK

In this paper, the objective is to reveal the influence of meteorological conditions on air quality prediction to improve the accuracy of the deep learning model for the practical application of air quality prediction.

The hourly average air pollutant concentration data and meteorological data at the same location should be selected for the investigation. By comparing the latitude and longitude position information of 35 air quality monitoring stations and 18 meteorological stations, the two closest stations were found, and the influence of meteorological conditions on air quality prediction was explored. In addition, feature selection is also required. High-quality feature selection can help improve the accuracy of air quality prediction. Therefore, corresponding analysis and selection of air quality prediction features are required. If using high-quality features, better air quality prediction effects can be obtained, and model performance can be improved.

In the collected data, there are 4,405 missing values, 54,696 irrelevant values, and 11,440 redundant values, and the final dataset has a total of 57,083 available data.

The processing methods for dealing with missing values include the direct deletion method, global constant replacement method, statistical number filling method (including mean value filling, median filling, mode filling, etc.), interpolation and kNN filling method, etc. [32], [33]. Since this paper is time-series data, the kNN filling method (K-Nearest Neighbors) is employed to achieve a better data filling effect.

The kNN filling method is derived from the kNN algorithm and inherits the idea of the kNN algorithm. It employed distance measurement to identify neighboring points and recognizes k samples that are similar in space in the dataset [34]. These "k" samples, that is, neighboring observations, are employed to estimate the value of the missing data point [35]. Missing values for each sample are interpolated employing the weighted average of the adjacent observations contained in the "k" neighborhood found in the dataset.

The selection of k in the kNN algorithm is also very important [36]. If the selected k value is too large, it will lead to over-simplification of the model, and the fitting results will be influenced by the farther point due to too many samples selected for fitting, which is also easy to produce prediction errors. In practical applications, a smaller k value is generally chosen. Since the missing values in this paper are relatively small, the existing data are relatively complete, and the dataset used in this paper is time-series data, the effect of the farther apart time on the vacancy value is small. Therefore, to improve the accuracy of the final fitting results, a relatively small k value, $k = 3$, is chosen to achieve the required missing value filling results.

For irrelevant and redundant values, the direct deletion method is employed for processing. The final dataset of this paper contains a total of 61,488 values. The missing values are some of the blank values of the parameters in the corresponding dates, and a total of 4,405 missing values were filled. Irrelevant values are the values of other influencing factors that are not relevant to the study of this paper, and a total of 54,696 irrelevant values were removed. Redundant values are the values measured at times outside the study time, and a total of 11,400 redundant values were removed.

C. STEP 2: CONSTRUCTION OF DEEP LEARNING MODELS

Deep learning models for time-series data analysis can be employed for air quality prediction. In this paper, two deep learning models, i.e., the LSTM model and the GRU model, are employed for air quality prediction. The LSTM model,

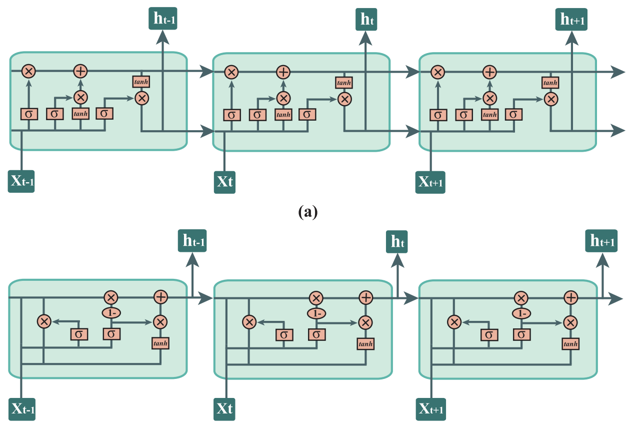


FIGURE 2. LSTM and GRU structure, (a) is the structure of LSTM model, (b) is the structure of GRU model.

known as Long short-term memory, is first proposed by Hochreiter and Schmidhuber [37]. It is a variant of RNN with a long-term memory function. The GRU model, known as Gated Recurrent Unit, is a variant of the LSTM model. It is proposed by Cho, and Merrienboer [38]. Both the LSTM model and the GRU model are typical time-series prediction models that can be employed to predict air quality. The structure of the LSTM and GRU models is illustrated in Figure 2.

In this paper, $PM_{2.5}$ predictions are first performed by only $PM_{2.5}$ data. Then, meteorological conditions and other air pollutant data are considered separately for air quality prediction. Finally, meteorological condition data and other air pollutants are combined to make air quality predictions, as illustrated in Table 1. The MAE, MSE, RMSE, and R^2 values of the LSTM and GRU models for air quality prediction are compared and analyzed for each of the above categories.

D. STEP 3: EXPLAINABILITY ANALYSIS OF DEEP LEARNING MODELS

Traditional feature importance analysis methods only identify which feature is important, but are not clear about how that feature affects the prediction results. SHAP is a game-theoretic approach to explain the output of any machine learning model. It connects optimal credit allocation with local explanations by the classical Shapley values from game theory and the related extensions [27]. The SHAP value reflects the influence of the features in each sample and shows the positivity and negativity of the influence, which belongs to the post hoc explanation of the model and can be interpreted for complex machine learning models. The principle of the SHAP method is illustrated in Figure 3. The base value (e.g., -0.08) in Figure 3 indicates the average value of the model prediction and the output indicates the model prediction output value of a single sample. In Figure 3, the SHAP method explains the final Prediction Output, which is

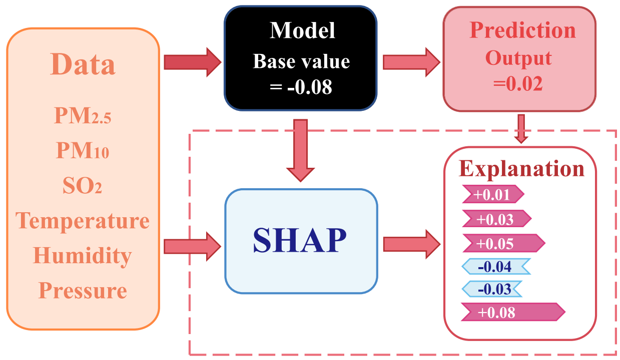


FIGURE 3. Principle of SHAP method.

due to the influence of $PM_{2.5}$, PM_{10}, and other factors on the final prediction result of +0.01, +0.03, etc., resulting in the final model output of 0.02.

Therefore, the SHAP method is chosen in this paper to conduct deep learning model explainability research to reveal the influence of meteorological conditions on air quality prediction. First, single-sample analysis is performed to analyze the distribution of SHAP values for individual features and individual samples. Analysis for all features is then performed to analyze the importance distribution of all features in the air quality prediction to have a more intuitive understanding of the influence of meteorological conditions on air quality prediction.

III. RESULTS

A. EXPERIMENTAL ENVIRONMENT

The software and hardware environment configurations used in this paper are listed in Table 2 and Table 3.

B. EXPERIMENTAL DATA

After comparing latitude and longitude, air pollutant data from the closest located Wanliu air quality monitoring station (39.987°N, 116.287°E) and meteorological condition data from the Hadian meteorological monitoring station (39.986°N, 116.291°E) are utilized for air quality prediction. Since the monitoring times of the two sites do not exactly overlap, the data between January 30th, 2017, 16:00, and January 31st, 2018, 15:00 are taken for the training test. The kNN interpolation method is used to fill in the missing values, and the irrelevant values are removed by direct deletion. $PM_{2.5}$ is selected for air quality prediction because it has a strong influence on human health and is of the highest concern.

$PM_{2.5}$ predictions are first conducted by only $PM_{2.5}$ data, where the first 70% was taken as the training data, and the last 30% as the test data. Then, based on the reading of the relevant literature and the analysis of the results of the pre-experiments, PM_{10} ($\mu g/m^3$), and SO_2 ($\mu g/m^3$) were taken from the air pollutant concentration data for the prediction of the air pollutant $PM_{2.5}$ ($\mu g/m^3$). And meteorological conditions, including temperature(°),

TABLE 1. Input and output of the prediction.

Case	Input	Output
1	$PM_{2.5}$ concentration	$PM_{2.5}$ concentration
2	$PM_{2.5}$ concentration + meteorological conditions(Temperature, Humidity, Atmospheric pressure)	$PM_{2.5}$ concentration
3	$PM_{2.5}$ concentration + other air pollutants concentration (PM_{10}, SO_2)	$PM_{2.5}$ concentration
4	$PM_{2.5}$ concentration + meteorological conditions(Temperature, Humidity, Atmospheric pressure) + other air pollutants concentration (PM_{10}, SO_2)	$PM_{2.5}$ concentration

TABLE 2. Software environment configurations used in this paper.

Software	Details
OS	Windows 10 Professional
Programming language	Python
Deep learning framework	PyTorch
Dependent library	Torch, SHAP, CUDA, PIL etc.

TABLE 3. Hardware environment configurations used in this paper.

Hardware	Details
CPU	Intel Xeon Gold 5118 CPU
CPU Frequency (GHz)	2.30
CPU core	48
CPU RAM (GB)	128
GPU	Quadro P6000
GPU memory (GB)	24
CUDA cores	3840
CUDA version	V10.2

humidity(%) and pressure (hPa) were employed for the prediction of the air pollutant $PM_{2.5}$ ($\mu g/m^3$). The first 70% was taken as the training data and the last 30% as the test data. Meteorological conditions, including temperature(°), humidity(%) and pressure (hPa), are then combined with $PM_{2.5}$($\mu g/m^3$), PM_{10}($\mu g/m^3$), and SO_2($\mu g/m^3$), for prediction of the air pollutant $PM_{2.5}$($\mu g/m^3$). Again, the first 70% was taken as the training data and the last 30% as the test data.

C. EXPERIMENTAL RESULTS AND ANALYSIS

The results of air quality prediction employing the LSTM model and GRU model are illustrated in Figure 4. Figure 4 (a) shows the line graph of the real data. In Figure 4 (b) $\sim$ (i), the red line represents the real raw data, the blue line represents the model training prediction value, and the black line represents the test value.

The mean absolute error (MAE), mean squared error (MSE), root mean squared error (RMSE), coefficient of determination (R^2) predicted by the LSTM and GRU models for 500 epochs are calculated, and the comparative analysis is illustrated in Figure 5 and Table 4. It can be seen that the results of LSTM and GRU models are close, but generally, LSTM can achieve better prediction results compared to GRU model, because although the MAE of LSTM model is larger compared to GRU, both MSE and RMSE have a larger decrease compared to GRU, and R2 has also improved compared to GRU.

In the LSTM model, considering only meteorological conditions, the MAE and MSE values are the largest, and the RMSE and R^2 values are also larger and compared to the prediction using only $PM_{2.5}$ concentration, the prediction results do not improve but decrease. Considering only air pollutants, MAE, MSE, and RMSE decrease compared to meteorological conditions only and $PM_{2.5}$ concentration only, and R^2 increases, the predictions improve. Considering only air pollutants, MAE, MSE, and RMSE decrease compared to considering only meteorological conditions and only $PM_{2.5}$ concentration, and R^2 increase, the predictions improve. The combination of meteorological conditions with air pollutants for air quality prediction is even better than considering only air pollutants, MAE, MSE, and RMSE values all decrease further and R^2 values increase further.

In the GRU model, the MAE decreases when only meteorological conditions are considered compared to only the $PM_{2.5}$ concentration are considered, while other values do not change significantly. The MAE increases when only meteorological conditions are considered, but the MSE and RMSE decrease, and the R^2 value increases when only meteorological conditions are considered. Air quality prediction is also better when meteorological conditions are considered in combination with other air pollutants, but not as much as the LSTM enhancement. This indicates that meteorological conditions cannot achieve better prediction results when employed directly for $PM_{2.5}$ prediction, meteorological conditions may become an interference factor in air quality prediction and interfere with the accuracy of air quality forecasting while combining with other air pollutants can lead to better prediction results than employing air pollutant data only, which can have a good influence on air quality prediction.

D. INFLUENCE OF METEOROLOGICAL CONDITIONS ON AIR QUALITY PREDICTION REVEALED BY SHAP

1) ANALYSIS FOR SINGLE SAMPLE

Two samples from the data (labeled sample 1 and sample 2) are randomly selected for the analysis of SHAP values, and the distribution of an individual sample SHAP values is illustrated in Figures 6 $\sim$ 7 (the eigenvalue data in the figure is normalized). The red part indicates features that make positive contributions to the predicted value, while the blue part indicates features that make negative contributions to the predicted value. Each segment indicates the contribution of a specific feature, and the length of the segment indicates

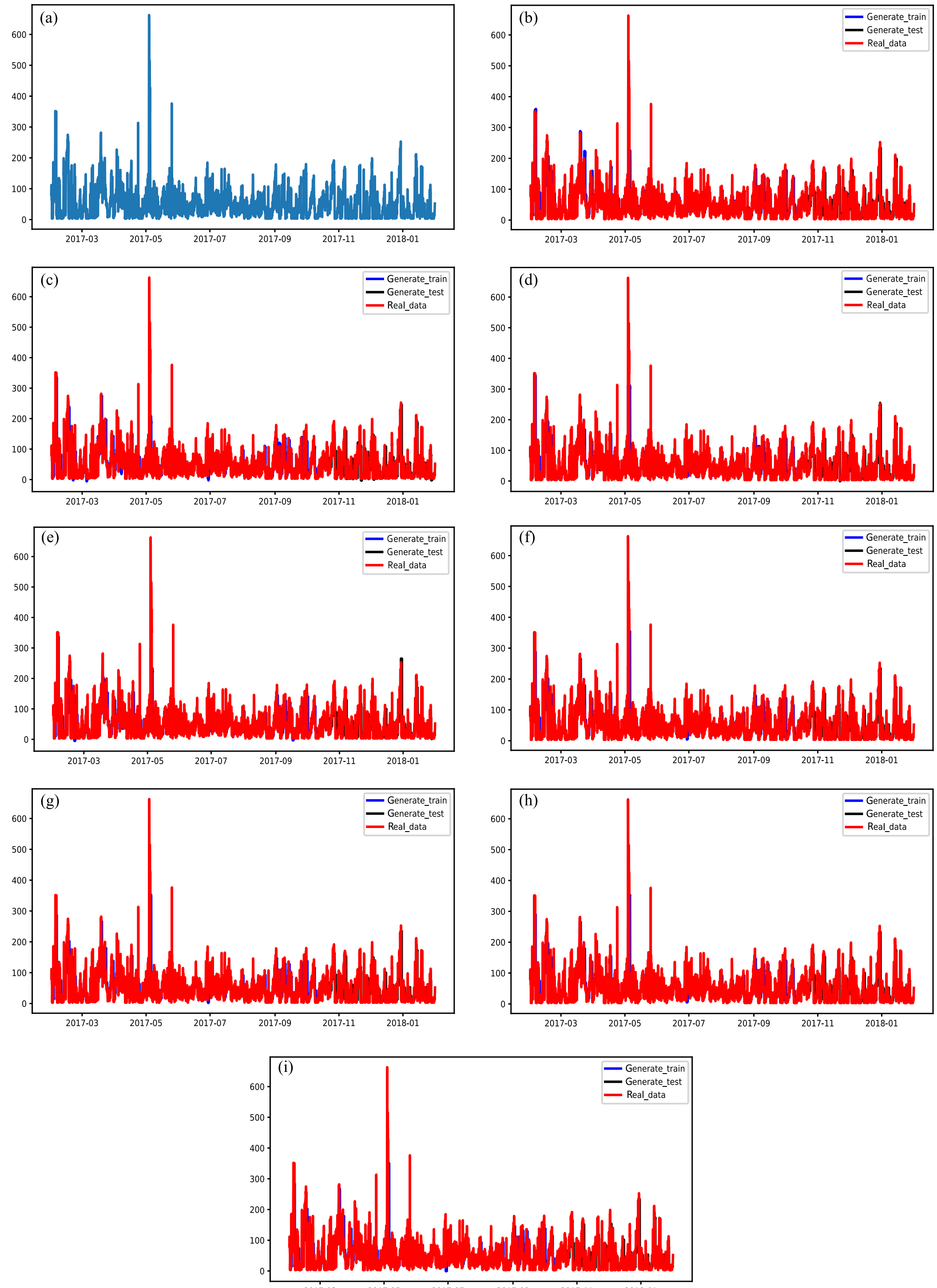


FIGURE 4. The results of air quality prediction employing the LSTM model and GRU model. (a) is the real data line, (b) ∼ (e) is the result of air quality prediction with $PM_{2.5}$ data, meteorological condition data, other air pollutants data, and combination of other air pollutants and meteorological condition data by using LSTM model, (f) ∼ (i) is the result of air quality prediction with $PM_{2.5}$ data, meteorological condition data, other air pollutants data, and combination of other air pollutants and meteorological condition data by using GRU model.

the contribution of the current feature to the current sample SHAP value (for features whose contribution is too small, it is not indicated due to the image space. In all of the subsequent figures, pressure means atmospheric pressure, and all of the later sections are replaced by pressure for atmospheric pressure.)

TABLE 4. Experimental results of LSTM and GRU models.

Model	Case	MAE	MSE	RMSE	R^2
LSTM	$PM_{2.5}$ concentration	0.4396	0.5752	0.7219	0.8504
	$PM_{2.5}$ + meteorological conditions	0.4401	0.5759	0.7214	0.8504
	$PM_{2.5}$ + air pollutants	0.4316	0.5624	0.7217	0.8534
	$PM_{2.5}$ + air pollutants	0.4306	0.5567	0.7085	0.8551
GRU	$PM_{2.5}$ concentration	0.4270	0.5922	0.7379	0.8459
	$PM_{2.5}$ + meteorological conditions	0.4262	0.5922	0.7382	0.8457
	$PM_{2.5}$ + air pollutants	0.4269	0.5896	0.7373	0.8465
	$PM_{2.5}$ + air pollutants	0.4263	0.5888	0.7359	0.8469

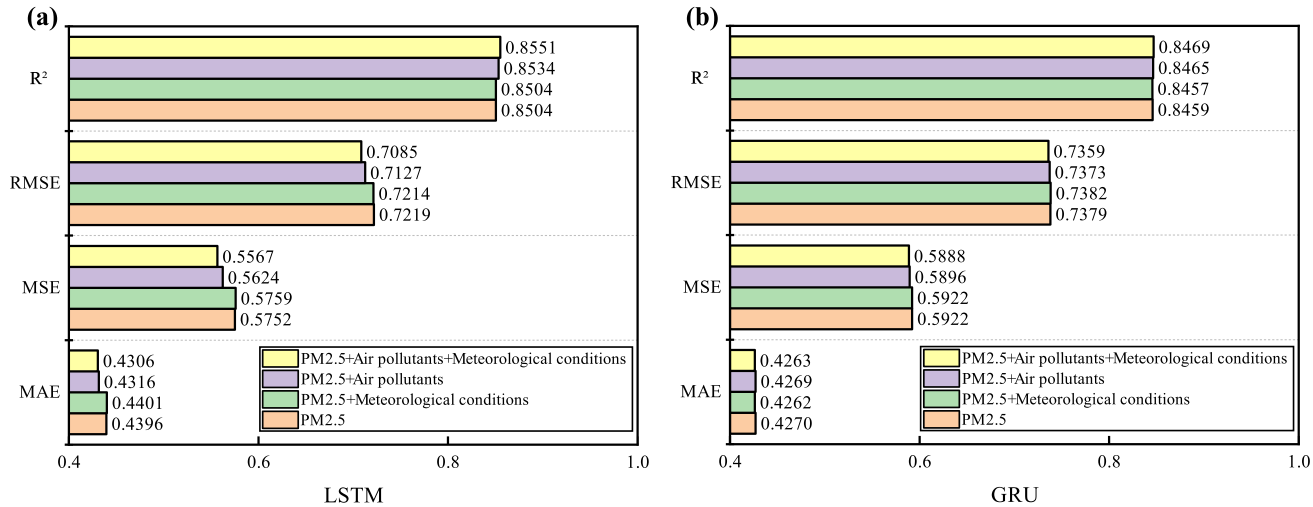


FIGURE 5. Comparison of LSTM and GRU model experimenta results. (a) is the experimenta results of LSTM model, (b) is the experimenta results of GRU model.

For the LSTM model, it can be seen in Figure 6 that when only meteorological conditions are considered, the base value is -0.1018, sample 1 output is -0.01, and sample 2 is -0.17. In sample 1, pressure makes the largest positive contribution, followed by humidity and temperature, which make a negative contribution. The contribution of $PM_{2.5}$ is small, and it makes positive contribution. In sample 2, pressure and humidity both make positive contributions, while $PM_{2.5}$ and temperature make negative contributions, with $PM_{2.5}$ making the largest contribution. In both samples, all three types of meteorological conditions are able to produce large contribution values, weakening the contribution of $PM_{2.5}$, which may be one of the reasons for the lack of improvement in air quality prediction when considering only meteorological conditions.

When only considering other air pollutant data, the base value is -0.0706, the sample 1 output is 0.93, and the sample 2 output is 1.00. The features $PM_{2.5}$, PM_{10}, and SO_2 all make positive contributions, with SO_2 making the largest contribution in both samples 1 and 2, and for the predictions of samples 1 and 2, SO_2 was more important.

When combining the other air pollutant data with the meteorological condition data, the base value is -0.06008, and the predicted value for both samples 1 and 2 is 0.17. In sample 1, pressure and SO_2 make positive contributions, humidity and temperature make negative contributions, and other features are ignored in the visualization due to their small contribution. In sample 1, the pressure made the largest contribution. In sample 2, pressure, humidity, $PM_{2.5}$, and SO_2 all made positive contributions, and temperature made a negative contribution. In both samples, the contribution of meteorological conditions for air quality prediction is larger and has a greater influence on air quality prediction.

For the GRU model, it can be seen in Figure 6 that when only meteorological conditions are considered, the base value is -0.2435, the output value of sample 1 is 0.04 and the output value of sample 2 is -0.47. In sample 1, pressure and $PM_{2.5}$ make positive contributions, temperature and humidity make negative contributions, and pressure makes the largest contribution. In sample 2, pressure and humidity make positive contributions, temperature and $PM_{2.5}$ make negative contributions, and $PM_{2.5}$ makes the largest contribution.

When only other air pollutants are considered, the base value is -0.1997, the output value of sample 1 is 0.99, and that of sample 2 is 1.0. In sample 1, PM_{10} and SO_2 makes a positive contribution, $PM_{2.5}$ makes a negative contribution, and SO_2 provides the maximum contribution value.

When combining the other air pollutants with the meteorological conditions, the base value is -0.3097, and the sample 1 output value is -0.26. Pressure and PM_{10}

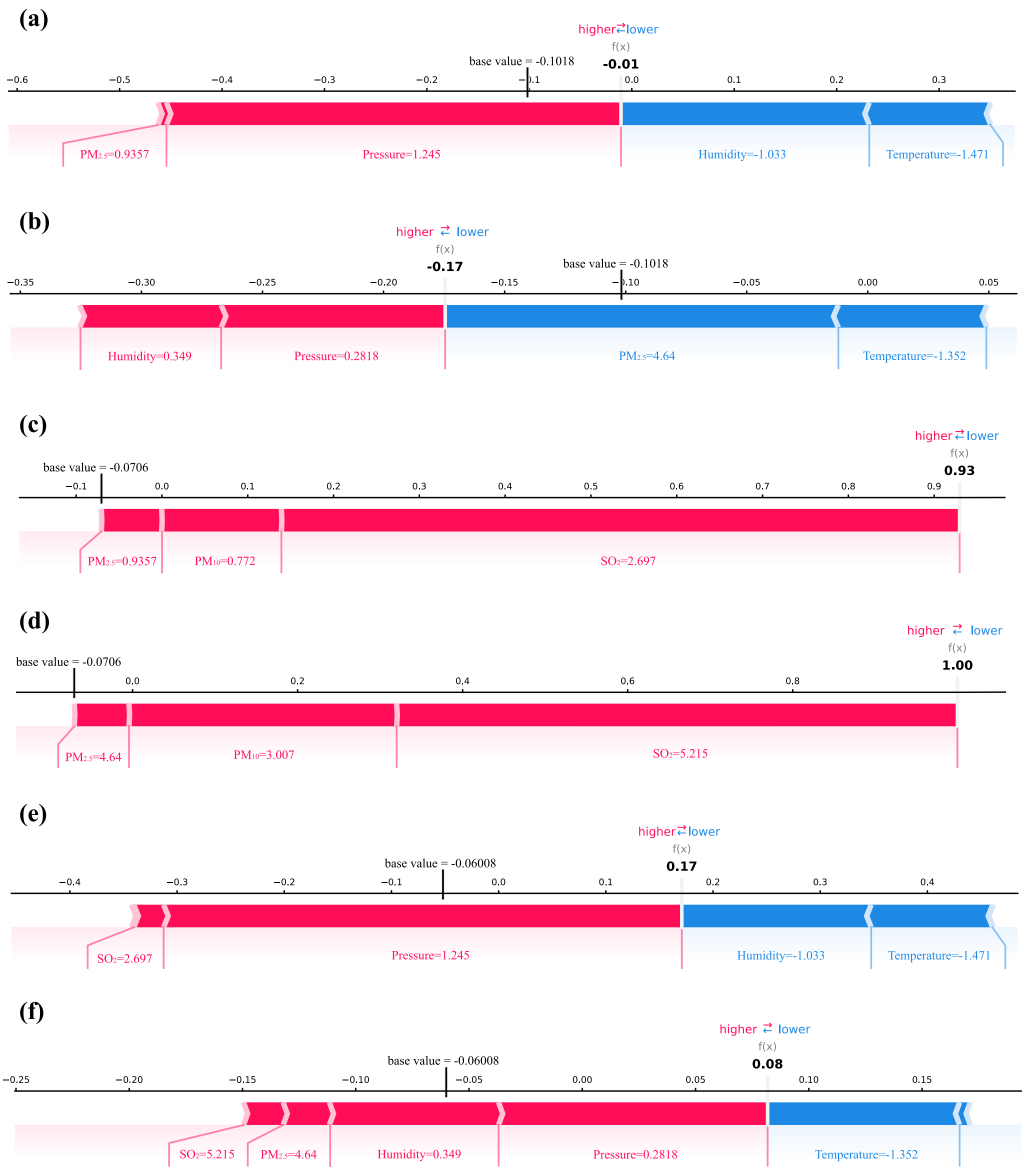


FIGURE 6. The distribution of individual sample SHAP values in LSTM model. (a) is the distribution of sample 1 SHAP value consider meteorological conditions, (b) is the distribution of sample 2 SHAP value consider meteorological conditions, (c) is the distribution of sample 1 SHAP value consider other air pollutants, (d) is the distribution of sample 2 SHAP value consider other air pollutants, (e) is the distribution of sample 1 SHAP value consider combination of other air pollutants and meteorological conditions, (f) is the distribution of sample 2 SHAP value consider combination of other air pollutants and meteorological conditions.

make positive contributions, temperature, SO_2, and humidity make negative contributions, and pressure makes the largest contribution. The output value of sample 2 is -0.6, with PM_{10}, pressure and humidity making positive contributions and SO_2, temperature, and $PM_{2.5}$ making negative contributions. The SO_2 contribution is the largest.

From Figures 6 $\sim$ 7, it can be investigated that meteorological conditions make a certain contribution to the predicted values and can reach larger SHAP values in some samples.

2) ANALYSIS OF ALL SAMPLES

The summary graph of SHAP before and after the addition of meteorological conditions is illustrated and ranked according to the importance of the features. The horizontal axis is the value of SHAP, as illustrated in Figure 8. The graph is wider at the point aggregation and thinner at the point dispersion, the redder color indicates a higher feature value, and the bluer color indicates a lower feature value.

When only meteorological conditions are considered, for the LSTM model, the larger the feature value of the pressure,

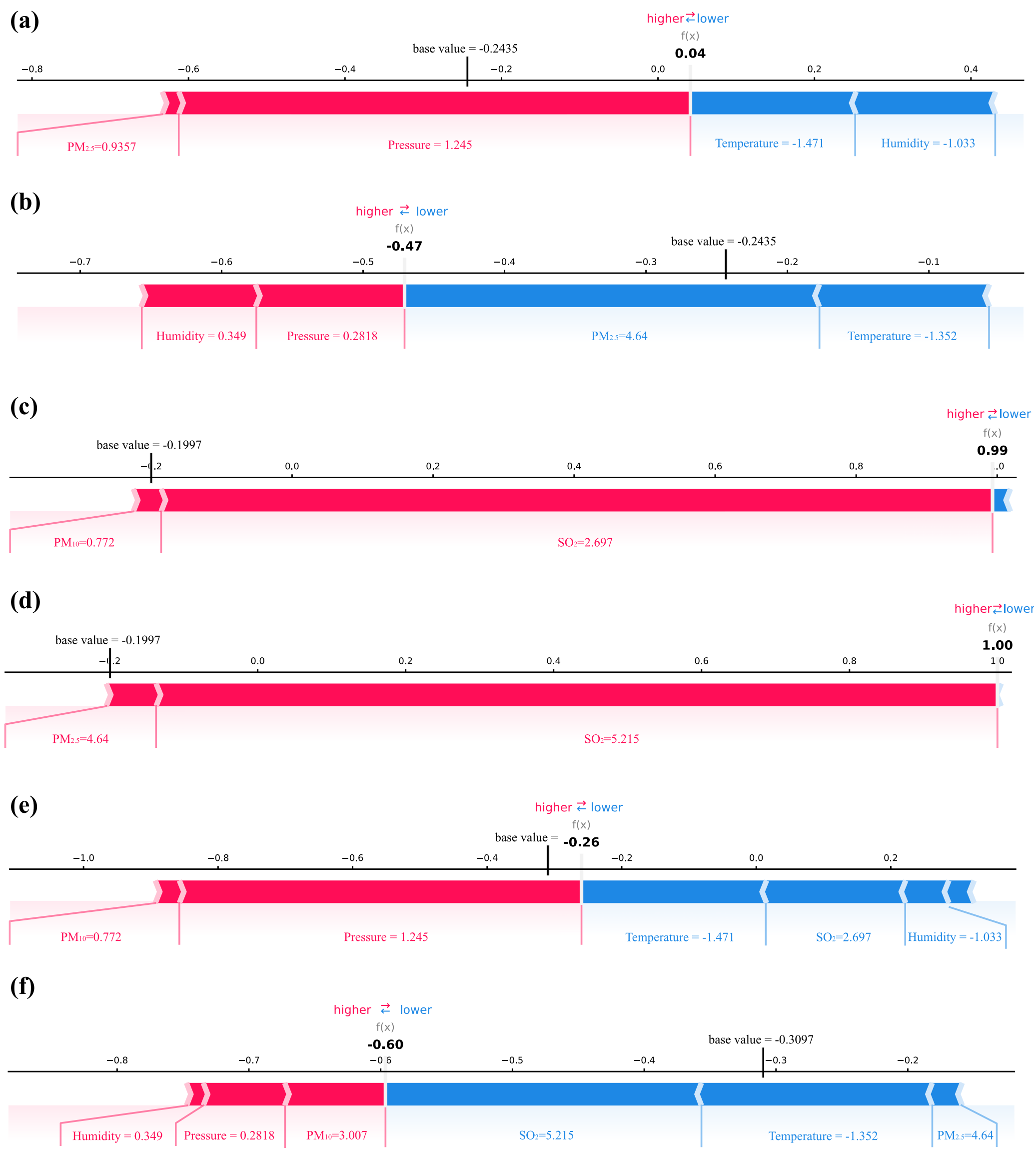


FIGURE 7. The distribution of individual sample SHAP values in GRU model. (a) is the distribution of sample 1 SHAP value consider meteorological conditions, (b) is the distribution of sample 2 SHAP value consider meteorological conditions, (c) is the distribution of sample 1 SHAP value consider other air pollutants, (d) is the distribution of sample 2 SHAP value consider other air pollutants, (e) is the distribution of sample 1 SHAP value consider combination of other air pollutants and meteorological conditions, (f) is the distribution of sample 2 SHAP value consider combination of other air pollutants and meteorological conditions.

humidity, and temperature, the larger the SHAP value, with $PM_{2.5}$ clustered at approximately 0 and unable to provide a large contribution. The GRU model indicates a similar pattern to the LSTM model, but the smaller the feature of $PM_{2.5}$, the larger the SHAP value, although it also mostly clustered at approximately 0, providing a limited contribution to the model prediction, though also detrimental to the model prediction.

Considering other air pollutants, for the LSTM model, all three features exhibit a positive effect on the model, and as the feature value becomes larger, the SHAP value also increases. Most of the points with small feature values have a negative impact on the model, and a small number of points with small feature values also have a positive impact on the model. For the GRU model, large SO_2 values are accompanied by large SHAP values, indicating that large SO_2 values have a positive effect on the predicted values and that a higher SO_2 is more likely to increase the predicted $PM_{2.5}$ values. The distribution characteristics of the PM_{10} and SO_2 SHAP values are similar, but there are individual cases of high feature values and low SHAP values, which may be caused by the influence of other factors. Most of the $PM_{2.5}$ SHAP values are clustered at

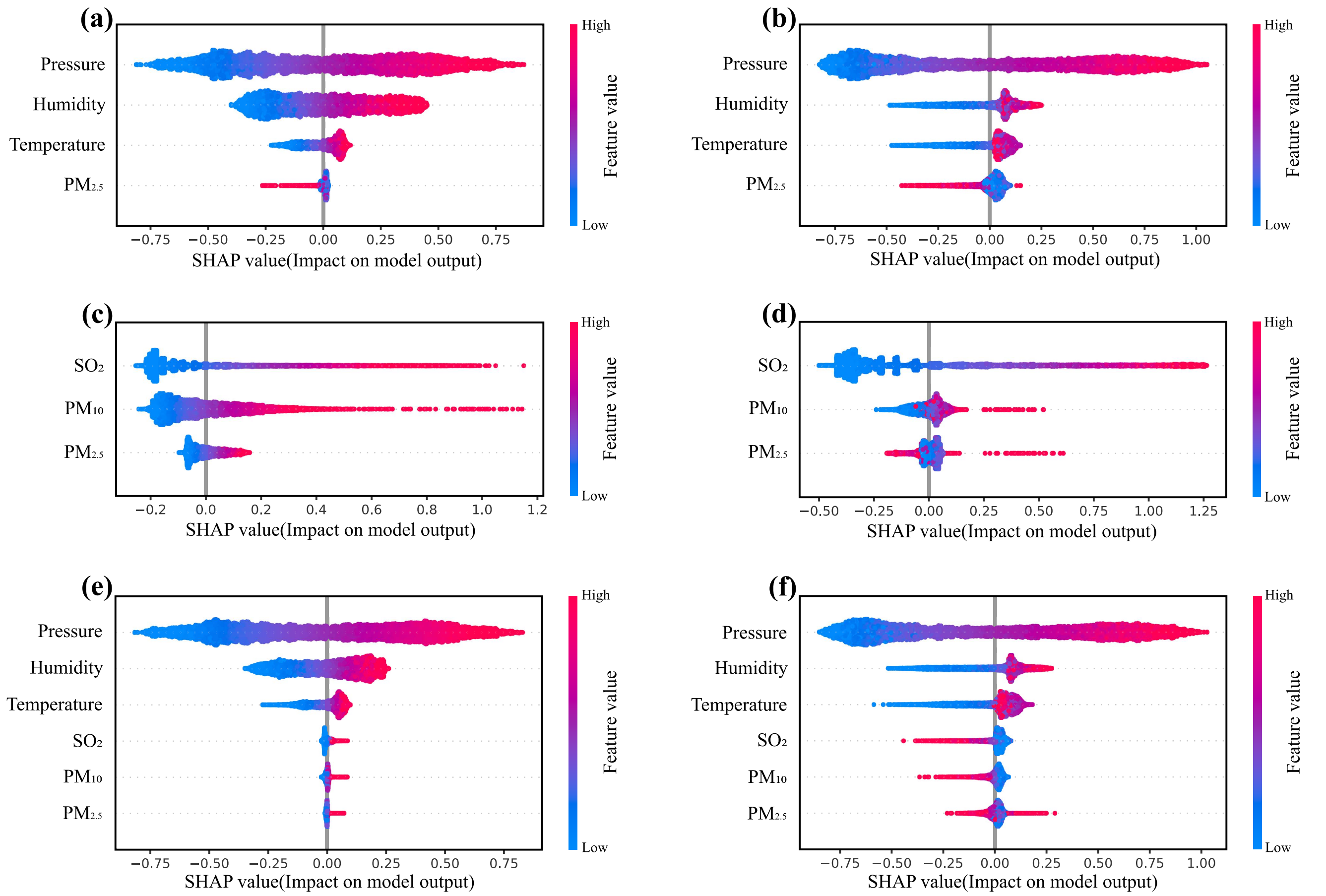


FIGURE 8. SHAP summary map, (a) ∼ (b) is LSTM and GRU model consider meteorological condition, (c) ∼ (d) is LSTM and GRU model consider other air pollutants, (e) ∼ (f) is LSTM and GRU model consider combination of other air pollutants and meteorological conditions.

approximately 0. Some of the larger features include both samples with positive and negative effects on the predicted values. Thus, their effects on the prediction results do not have very obvious characteristics.

After combining meteorological conditions with other air pollutant conditions, the three characteristics of meteorological conditions have a greater impact on the predicted values in both the LSTM and GRU models, ranking in the top three. Pressure contributes substantially more to the prediction of the LSTM and GRU models than other features. When the pressure is larger, it has a greater positive effect on the predicted value, and when the pressure is smaller, it has a greater negative effect on the model. In the LSTM model, humidity has a smaller range of SHAP value distribution compared to pressure, but it also has a larger effect relative to other features, and the larger the feature is, the greater the positive effect. SO_2, PM_{10}, and $PM_{2.5}$ have the same influence pattern as meteorological conditions, but their SHAP values are smaller, and most of them are clustered around SHAP value = 0.

In the GRU model, when the feature values of humidity and temperature are large, the SHAP value is positive and has a positive impact on the predicted results, but the impact value is not large. There are also some values with smaller feature values that also have a positive impact on the model, while most values with smaller feature values have a negative impact on the predicted values. SO_2, $PM_{2.5}$, and PM_{10} with large feature values have a negative effect on the predicted values, which may be due to the limitations of the model itself. SO_2, $PM_{2.5}$, and PM_{10} are mostly clustered around SHAP value = 0. The SHAP values are all small and have limited influence on the prediction results. Additionally, the air quality prediction is mainly influenced by meteorological conditions.

3) ANALYSIS OF SINGLE FEATURE

The SHAP value distribution of each feature is plotted as illustrated in Figures 9 ∼ 14. The horizontal coordinate is the normalized feature value, and the vertical coordinate is the SHAP value that corresponds to the feature.

In the LSTM model, when only meteorological conditions are considered, the SHAP value increases as the temperature, humidity and pressure increase. As $PM_{2.5}$ increases, the SHAP value decreases. This is not in accordance with

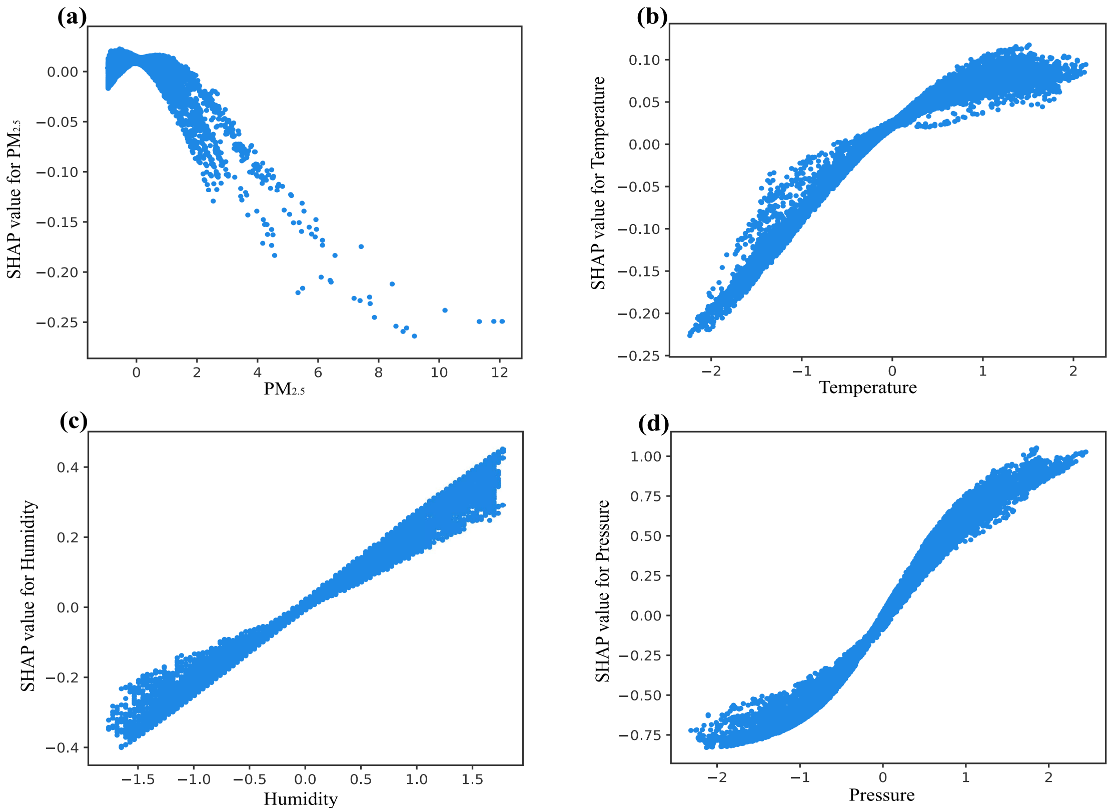


FIGURE 9. The SHAP value distribution of each feature consider meteorological condition by LSTM model. (a) is the SHAP value for $PM_{2.5}$, (b) is the SHAP value for Temperature, (c) is the SHAP value for Humidity, (d) is the SHAP value for Pressure.

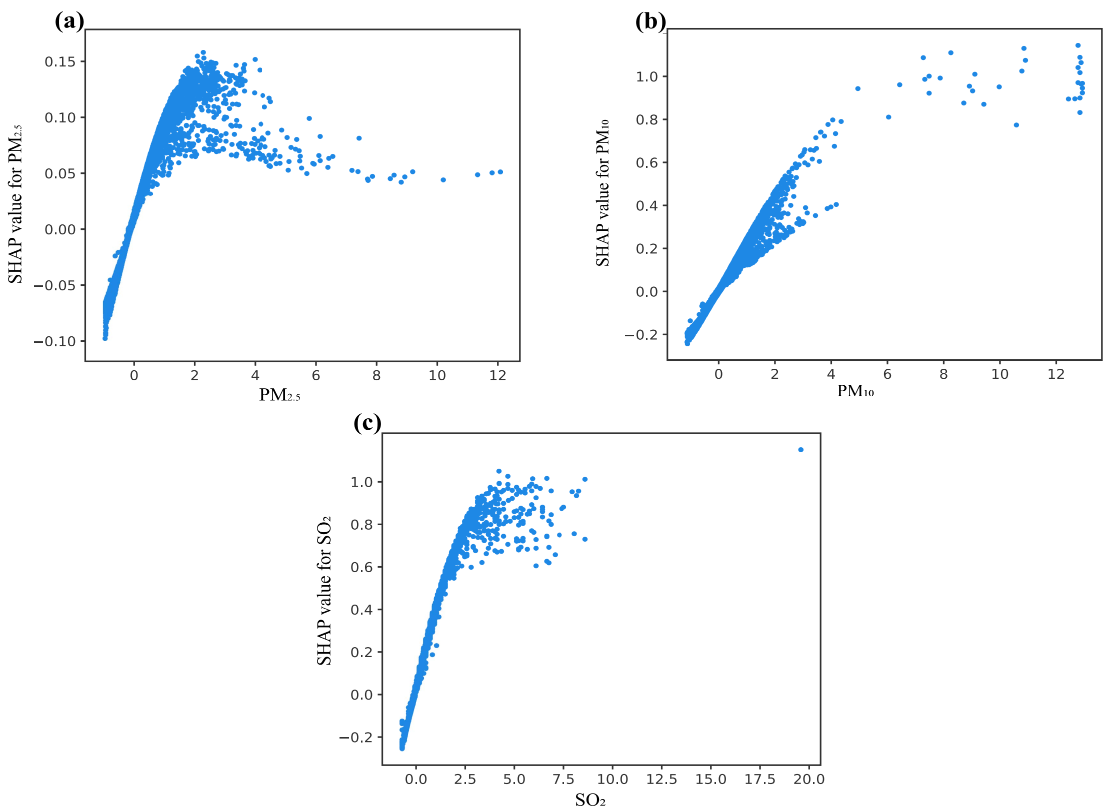


FIGURE 10. The SHAP value distribution of each feature consider other air pollutants by LSTM model. (a) is the SHAP value for $PM_{2.5}$, (b) is the SHAP value for PM_{10}, (c) is the SHAP value for SO_2.

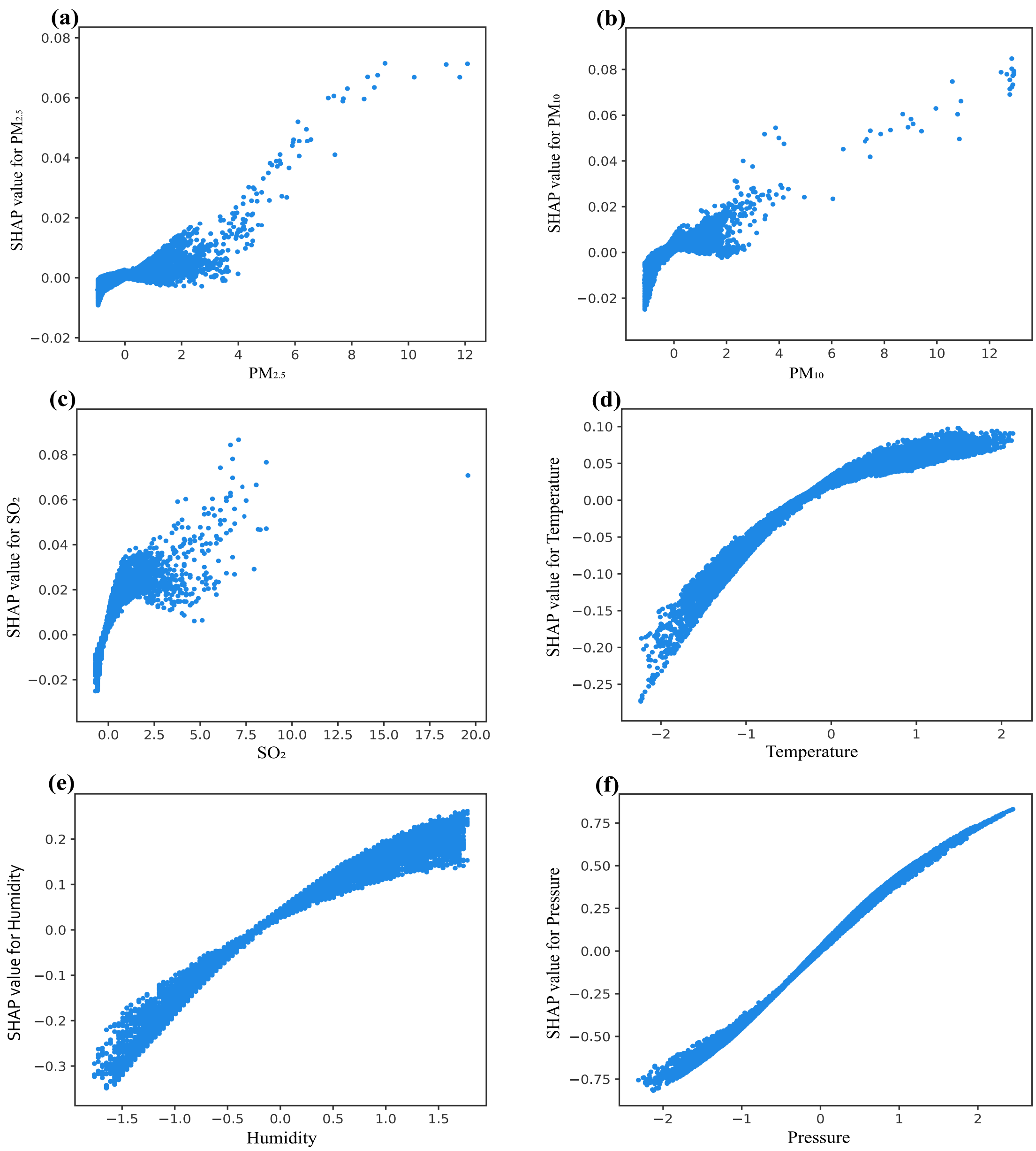


FIGURE 11. The SHAP value distribution of each feature consider combination of air pollutant and meteorological condition by LSTM model. (a) is the SHAP value for $PM_{2.5}$, (b) is the SHAP value for PM_{10}, (c) is the SHAP value for SO_2, (d) is the SHAP value for Temperature, (e) is the SHAP value for Humidity, (f) is the SHAP value for Pressure.

the prediction law and may be due to the addition of meteorological conditions, which have an effect on the SHAP value of $PM_{2.5}$, and thus, interfere with the prediction results.

In the LSTM model, when considering other air pollutants, the SHAP value distribution diagram of each feature is illustrated in Figure 10. As seen in Figure 10, the SHAP value of $PM_{2.5}$ increases with the increase of the $PM_{2.5}$ feature value, and when it reaches a certain value, it indicates a trend of a decreasing value with the increase of a feature value. However, it is not obvious and gradually tends to be stable. The SHAP value of PM_{10} increases with an increasing feature value, and the SHAP value is more scattered and has fewer data points when the feature value is larger. The SHAP value of SO_2 increases linearly with the increase in the SO_2 feature, and the linear relationship is more obvious.

The SHAP value of each feature considering the combination of other air pollutants and meteorological conditions in the LSTM model is illustrated in Figure 11. As seen from Figure 11, the SHAP values of $PM_{2.5}$, PM_{10}, and SO_2 decrease substantially, and the linear relationship with feature value is not obvious, though it still retains the

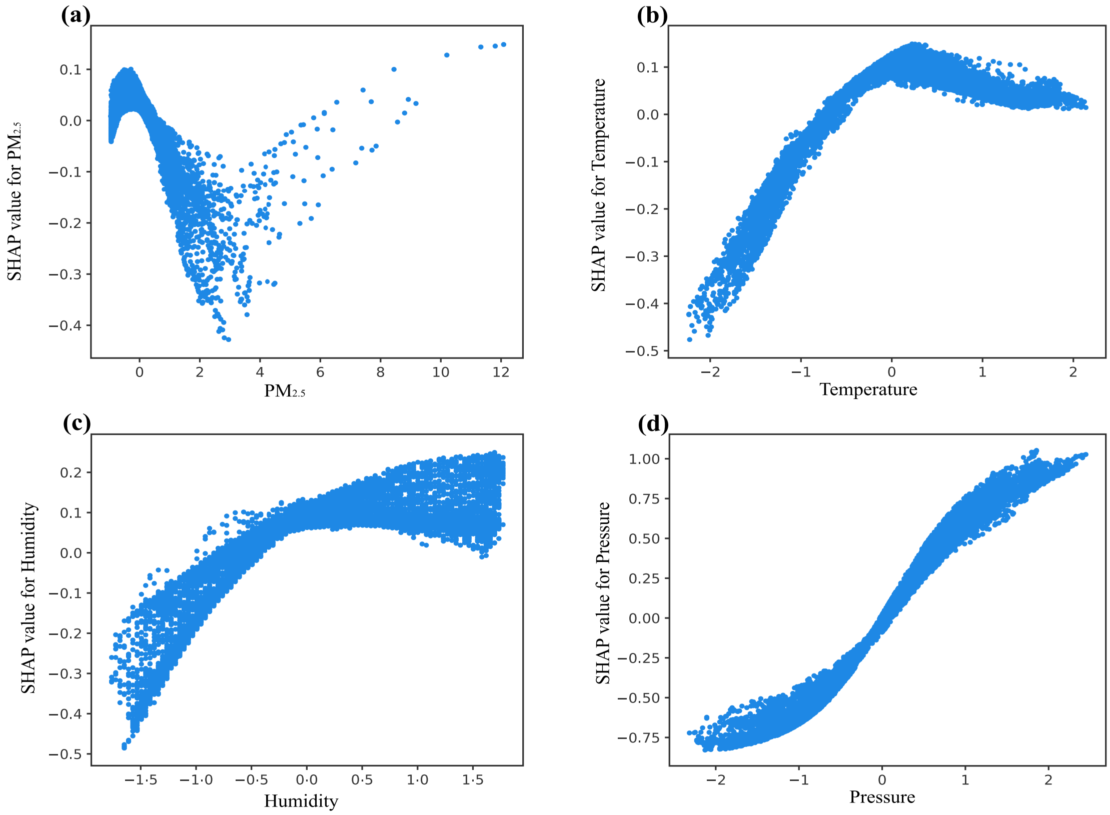


FIGURE 12. **The SHAP value distribution of each feature consider meteorological conditions by GRU model. (a) is the SHAP value for $PM_{2.5}$, (b) is the SHAP value for Temperature, (c) is the SHAP value for Humidity, (d) is the SHAP value for Pressure.**

trend of increasing with the feature. The SHAP values of the temperature, humidity, and pressure all have a more obvious linear relationship with the feature value. The larger the feature is, the larger the SHAP values of temperature, humidity, and pressure. The SHAP values of temperature, humidity, and pressure all tend to smooth out after reaching a certain value.

In the GRU model, similar to the LSTM model, the SHAP values of temperature, humidity, and pressure also tend to increase with increasing feature values when only meteorological conditions are considered, but tend to decrease slightly after reaching a certain value. $PM_{2.5}$, however, shows a decreasing SHAP value as the feature value increases.

For the GRU model, the distribution of the SHAP value for each feature when considering other air pollutants is illustrated in Figure 13. As illustrated in Figure 13, there is a fluctuating effect of $PM_{2.5}$ concentration on the predicted effect of $PM_{2.5}$. There is a certain linear relationship when the $PM_{2.5}$ value is small, and the linear relationship weakens as the $PM_{2.5}$ value keeps increasing, but it indicates a trend where while the SHAP value increases, the $PM_{2.5}$ value also increases. The relationship between PM_{10} and the predicted results is not obvious, but it generally also indicates that the SHAP value increases with an increasing PM_{10} value. The linear relationship of SO_2 is more obvious. With an increasing SO_2 value, the SHAP value increases, and its contribution to the predicted value increases.

The distribution of SHAP values for each feature considering the combination of other air pollutants and meteorological conditions to the GRU model is illustrated in Figure 14. As illustrated in Figure 14, the SHAP value of the air pollutant concentration data changes after the addition of meteorological condition data. With the change in $PM_{2.5}$, PM_{10}, and SO_2 feature values, the changing pattern of the SHAP value is not obvious. For $PM_{2.5}$ and SO_2, the general trend of the SHAP value decreases with an increasing feature value, while the general distribution of PM_{10} is loose with no obvious trend.

From the above descriptions, it is clear that meteorological conditions have a greater influence on air quality prediction results, and when only meteorological conditions are considered for air quality prediction, the accuracy of air quality prediction cannot be improved. This may be caused by the fact that meteorological conditions affect the prediction contribution of $PM_{2.5}$. The SHAP value of the contribution of meteorological conditions to air quality prediction is generally high. Additionally, the contribution is high and increases with the increasing feature. When meteorological conditions are considered in combination with other air pollutants, meteorological conditions interfere less with the SHAP value of $PM_{2.5}$ due to the interaction between other

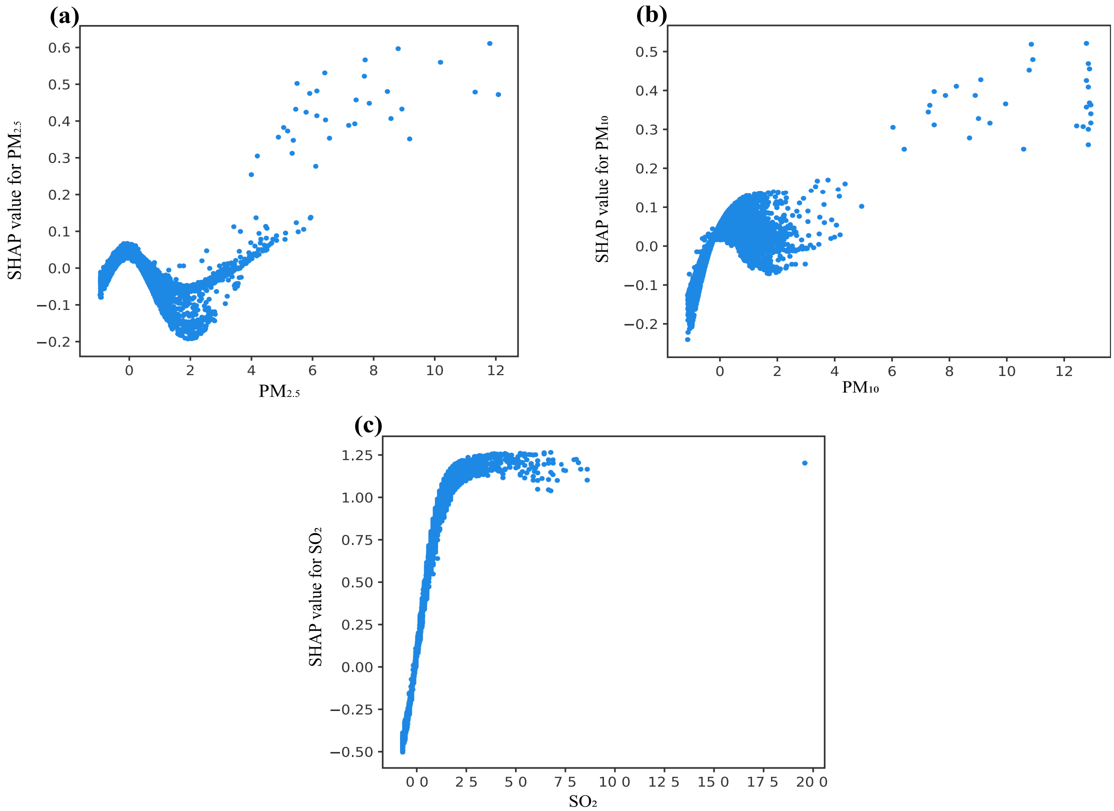


FIGURE 13. The SHAP value distribution of each feature consider other air pollutants by GRU model. (a) is the SHAP value for $PM_{2.5}$, (b) is the SHAP value for PM_{10}, (c) is the SHAP value for SO_2.

air pollutants and meteorological conditions, and can play a positive role themselves, thus facilitating the prediction of air quality.

4) ANALYSIS OF ALL FEATURES

The SHAP value of all of the samples of each feature is first found as an absolute value and then averaged. The mean | SHAP value | of each feature is obtained as the importance of each feature and ranked in descending order from largest to smallest. The feature importance before and after adding meteorological conditions to the LSTM and GRU models is illustrated in Figure 15.

As seen fin Figure 15, the LSTM and GRU models have essentially the same ranking of feature importance. When only meteorological conditions are considered, the meteorological conditions pressure, temperature, and humidity all have higher importance features than $PM_{2.5}$, weakening the importance of $PM_{2.5}$ for air quality prediction, which in turn leads to unsatisfactory prediction results.

When only other air pollutant conditions are considered, SO_2 is the most important feature influencing the predicted values, followed by PM_{10} and $PM_{2.5}$. In the GRU model, the mean | SHAP value | of SO_2 is much larger than the other two features, while the mean | SHAP value | of PM_{10} and $PM_{2.5}$ in the LSTM is smaller than that of SO_2, but it still occupies a certain proportion. To some extent, it explains why the LSTM model effect is slightly better than the GRU model effect in this experiment.

After combining meteorological conditions with other air pollutant conditions, meteorological conditions remain the top three most important factors influencing predictions in the LSTM and GRU models, but this condition plays a positive role. The mean | SHAP value | of pressure is the largest and much larger compared to other factors, indicating that the contribution of pressure to air quality prediction is the largest. The mean | SHAP value | of humidity and temperature compared to the air pollutant concentration is also much larger. For the hourly air pollutant concentration feature, the LSTM and GRU models are ranked slightly differently. In the LSTM, the mean | SHAP value | is SO_2, PM_{10}, and $PM_{2.5}$ from largest to smallest, while in the GRU model, the mean | SHAP value | is SO_2, $PM_{2.5}$, and PM_{10} from largest to smallest. However, these three features are less important in both models.

In summary, for each air quality prediction data feature, the larger the average absolute value of SHAP is, the higher the importance of the feature and the greater the influence on the prediction results. After the meteorological condition data is added, the SHAP method verifies that the average absolute values of SHAP for all three types of added meteorological condition data are high and could even reach the top three in the importance ranking of all samples, which is a factor that

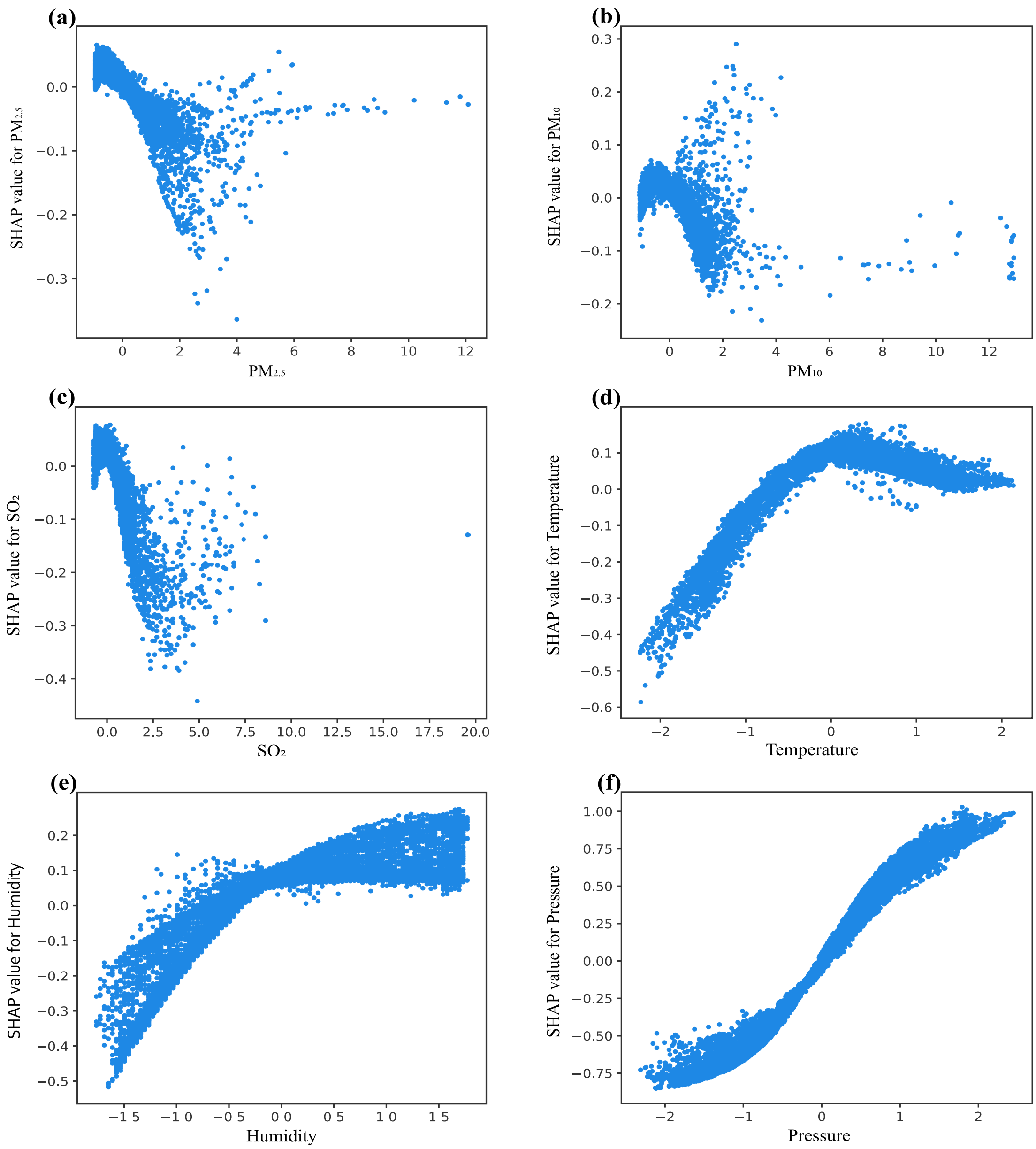


FIGURE 14. The SHAP value distribution of each feature considering the combination of other air pollutants and meteorological conditions by GRU model. (a) is the SHAP value for $PM_{2.5}$, (b) is the SHAP value for PM_{10}, (c) is the SHAP value for SO_2, (d) is the SHAP value for Temperature, (e) is the SHAP value for Humidity, (f) is the SHAP value for Pressure.

highly contributes to air quality prediction. It is demonstrated that meteorological condition data has a large impact on air quality prediction results. Adding meteorological condition data to air quality prediction can make the prediction results more accurate.

Compared with existing work, our research focuses on how meteorological conditions affect air quality prediction by employing the explainable deep learning method SHapley Additive exPlanations (SHAP) to reveal the impact of meteorological conditions on air quality prediction. This is beneficial to improve the air quality prediction accuracy of deep learning models in practical applications, and to establish deep learning models with higher accuracy and efficiency for air quality prediction to protect people's lives and health.

IV. DISCUSSION

In this paper, we reveal the influence of meteorological conditions on air quality prediction by utilizing the explainable deep learning method (SHAP). In this way, more accurate air quality predictions can be achieved, and more effective environmental protection strategies can be formulated.

By employing the SHAP method to explain the deep learning model for air quality prediction, we can not

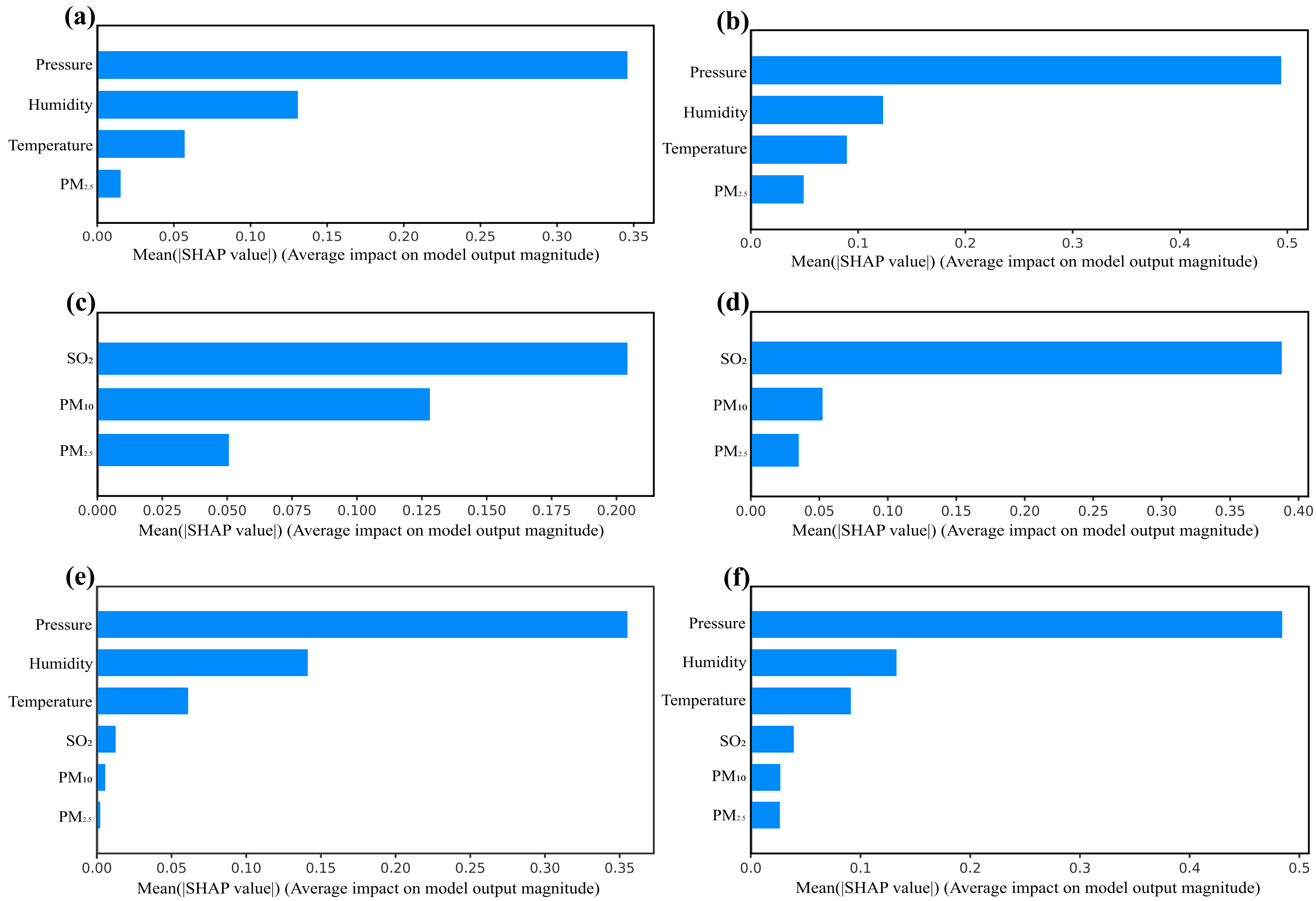


FIGURE 15. Feature importance before and after adding meteorological conditions. (a) ∼ (b) is the feature importance ranking considering meteorological condition by LSTM model and GRU model, (c) ∼ (d) is the feature importance ranking considering other air pollutants by LSTM and GRU model, (e) ∼ (f) is the feature importance ranking considering the combination of other air pollutants and meteorological conditions.

only identify which features are of high importance for air quality prediction but also explain how the features affect the air quality prediction results and obtain the magnitude of the contribution of each type of feature to the prediction results. This helps us to have a clearer understanding of the influence of meteorological conditions on air quality prediction in deep learning models and to have more knowledge about air quality prediction deep learning models to employ them more appropriately for air quality prediction.

However, there are several limitations of the research work described in this paper. The SHAP method has high computational effort, large computational memory required, and low computational efficiency. In our work, we had to ignore several features and reduce the computational dimension for research purposes due to the huge computational volume and excessive computation time, as well as the huge storage space required of SHAP. This reduces the air quality prediction performance of the model to a certain degree. In the air pollutant dataset, we selected three out of six pollutants for prediction. In the meteorological condition dataset, we also selected three meteorological conditions with a higher correlation with air quality from the six types of meteorological conditions data. To achieve the purpose of the study, we had to lose some prediction accuracy, which is an aspect we need to study further in-depth in the future.

In the future, the SHAP algorithm needs to be further optimized to reduce computational effort, improve computational efficiency, and better explain how meteorological conditions affect air quality prediction. It is hoped that the working principle of deep learning models for air quality prediction can also be further investigated by revealing the influence of meteorological conditions on air quality prediction, and optimization and improvement of air quality prediction deep learning model can be carried out, and the air quality prediction accuracy and efficiency of the deep learning model can be further improved. Ultimately, a more accurate and trustworthy deep learning model for air quality prediction will be built to be applied in realistic air quality prediction to protect people's lives and health.

V. CONCLUSION

In this paper, we employ the explainable deep learning method, SHapley Additive exPlanations, to reveal the influence of meteorological conditions on air quality prediction. The essential idea is to use the SHAP interpretation method to interpret the established LSTM and GRU air quality prediction models and analyze the influence of meteorological conditions on air quality prediction.

The results show that (1) in both the LSTM and GRU models, the prediction accuracy is not improved by considering only meteorological conditions. However, when considering other air pollutants, the prediction accuracy is improved, and when combining meteorological conditions with other air pollutants, the prediction accuracy is even higher. (2) Whether only considering meteorological conditions or combining meteorological conditions and other air pollutants for $PM_{2.5}$ prediction, in both the LSTM and GRU models, the meteorological conditions have a high contribution and importance to air quality prediction, meaning that they are all in the top three in terms of contribution. The largest contribution to air quality prediction is made by atmospheric pressure, the second by humidity, and the third by temperature. When considering only air pollutants, SO_2 contributes the most to the air quality prediction. (3) However, when only meteorological conditions are considered for air quality prediction, the high contribution of meteorological conditions to the prediction interferes with the results and makes the results more inaccurate. When meteorological conditions are considered in combination with other air pollutants, the high contribution of meteorological conditions to the prediction facilitates the prediction of air quality and leads to better results. (4) The reason for the different accuracies of the final prediction may be that the SHAP value is different in different conditions, meaning that the contribution to the prediction result is different. This is caused by the interaction of meteorological conditions with other air pollutants.

Compared with other research, this paper employs an explainable deep learning method, the SHAP method, to analyze how meteorological conditions affect air quality prediction. This facilitates the in-depth analysis and understanding of the deep learning models for air quality prediction and improves the trustworthiness of the deep learning models. In the future, we plan to build deep learning models with higher accuracy and trustworthiness for air quality prediction, which can be applied to realistic air quality prediction.

ACKNOWLEDGMENT

The authors would like to thank the editors and the reviewers for their valuable comments.

REFERENCES

[1] H. Kan, R. Chen, and S. Tong, "Ambient air pollution, climate change, and population health in China," *Environ. Int.*, vol. 42, pp. 10–19, Jul. 2012.

[2] H. Zhang, S. Wang, J. Hao, X. Wang, S. Wang, F. Chai, and M. Li, "Air pollution and control action in Beijing," *J. Cleaner Prod.*, vol. 112, pp. 1519–1527, Jan. 2016.

[3] P. Kumar, L. Morawska, C. Martani, G. Biskos, M. Neophytou, S. Sabatino, M. Bell, L. Norford, and R. Britter, "The rise of low-cost sensing for managing air pollution in cities," *Environ. Int.*, vol. 75, pp. 199–205, Feb. 2015.

[4] E. D. Schraufnagel, R. J. Balmes, T. C. Cowl, S. D. Matteis, S.-H. Jung, K. Mortimer, R. Perez-Padilla, B. M. Rice, H. Riojas-Rodriguez, A. Sood, D. G. Thurston, T. To, A. Vanker, and J. D. Wuebbles, "Air pollution and noncommunicable diseases a review by the forum of international respiratory societies' environmental committee, part 2: Air pollution and organ systems," *Chest*, vol. 155, no. 2, pp. 417–426, 2019.

[5] Y.-F. Xing, Y.-H. Xu, M.-H. Shi, and Y.-X. Lian, "The impact of PM2.5 on the human respiratory system," *J. Thoracic Disease*, vol. 8, no. 1, pp. E69–E74, 2016.

[6] J. J. West, A. Cohen, F. Dentener, B. Brunekreef, T. Zhu, B. Armstrong, M. L. Bell, M. Brauer, G. Carmichael, D. L. Costa, and D. W. Dockery, "What we breathe impacts our health: Improving understanding of the link between air pollution and health," *Environ. Sci. Technol.*, vol. 50, no. 10, pp. 4895–4904, 2016.

[7] X. Zhang, X. Zhang, and X. Chen, "Happiness in the air: How does a dirty sky affect mental health and subjective well-being?" *J. Environ. Econ. Manage.*, vol. 85, pp. 81–94, Sep. 2017.

[8] X. Qi, G. Mei, S. Cuomo, C. Liu, and N. Xu, "Data analysis and mining of the correlations between meteorological conditions and air quality: A case study in Beijing," *Internet Things*, vol. 14, Jun. 2021, Art. no. 100127.

[9] S. Al-Janabi, M. Mohammad, and A. Al-Sultan, "A new method for prediction of air pollution based on intelligent computation," *Soft Comput.*, vol. 24, no. 1, pp. 661–680, Jan. 2020.

[10] A. Alimissis, K. Philippopoulos, C. G. Tzanis, and D. Deligiorgi, "Spatial estimation of urban air pollution with the use of artificial neural network models," *Atmos. Environ.*, vol. 191, pp. 205–213, Oct. 2018.

[11] H. Li, J. Wang, R. Li, and H. Lu, "Novel analysis-forecast system based on multi-objective optimization for air quality index," *J. Cleaner Prod.*, vol. 208, pp. 1365–1383, Jan. 2019.

[12] A. Kumar and P. Goyal, "Forecasting of daily air quality index in Delhi," *Sci. Total Environ.*, vol. 409, no. 24, pp. 5517–5523, Nov. 2011.

[13] W. G. Cobourn, "An enhanced PM2.5 air quality forecast model based on nonlinear regression and back-trajectory concentrations," *Atmos. Environ.*, vol. 44, no. 25, pp. 3015–3023, Aug. 2010.

[14] T. S. Rajput and N. Sharma, "Multivariate regression analysis of air quality index for Hyderabad city: Forecasting model with hourly frequency," *Int. J. Appl. Res.*, vol. 3, no. 8, pp. 443–447, 2017.

[15] Y. Liu, Q. Zhu, D. Yao, and W. Xu, "Forecasting urban air quality via a back-propagation neural network and a selection sample rule," *Atmosphere*, vol. 6, no. 7, pp. 891–907, Jul. 2015.

[16] S. Xiao, Q. Y. Wang, J. J. Cao, R.-J. Huang, W. D. Chen, Y. M. Han, H. M. Xu, S. X. Liu, Y. Q. Zhou, P. Wang, J. Q. Zhang, and C. L. Zhan, "Long-term trends in visibility and impacts of aerosol composition on visibility impairment in Baoji, China," *Atmos. Res.*, vol. 149, pp. 88–95, Nov. 2014.

[17] Y. Qi, Q. Li, H. Karimian, and D. Liu, "A hybrid model for spatiotemporal forecasting of PM2.5 based on graph convolutional neural network and long short-term memory," *Sci. Total Environ.*, vol. 664, pp. 1–10, May 2019.

[18] C. Wen, S. Liu, X. Yao, L. Peng, X. Li, Y. Hu, and T. Chi, "A novel spatiotemporal convolutional long short-term neural network for air pollution prediction," *Sci. Total Environ.*, vol. 654, pp. 1091–1099, Mar. 2019.

[19] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Netw.*, vol. 61, pp. 85–117, Jan. 2015.

[20] Z. Ma and G. Mei, "Deep learning for geological hazards analysis: Data, models, applications, and opportunities," *Earth-Sci. Rev.*, vol. 223, Dec. 2021, Art. no. 103858.

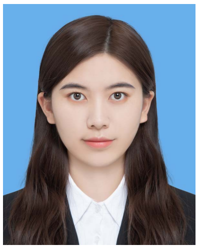
[21] S. Du, T. Li, Y. Yang, and S.-J. Horng, "Deep air quality forecasting using hybrid deep learning framework," *IEEE Trans. Knowl. Data Eng.*, vol. 33, no. 6, pp. 2412–2424, Jun. 2021.

[22] Y. Zhou, F.-J. Chang, L.-C. Chang, I.-F. Kao, and Y.-S. Wang, "Explore a deep learning multi-output neural network for regional multi-step-ahead air quality forecasts," *J. Clean Prod.*, vol. 209, pp. 134–145, Feb. 2019.

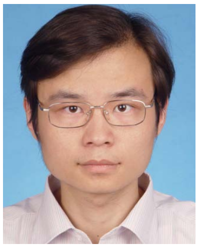
[23] Y. Jiao, Z. Wang, and Y. Zhang, "Prediction of air quality index based on LSTM," in *Proc. IEEE 8th Joint Int. Inf. Technol. Artif. Intell. Conf. (ITAIC)*, May 2019, pp. 17–20.

[24] R. J. Kuo, B. Prasetyo, and B. S. Wibowo, "Deep learning-based approach for air quality forecasting by using recurrent neural network with Gaussian process in Taiwan," in *Proc. IEEE 6th Int. Conf. Ind. Eng. Appl. (ICIEA)*, Apr. 2019, pp. 471–474.

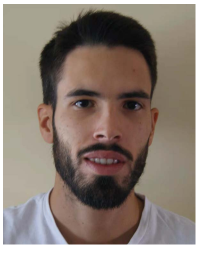
[25] Y.-S. Chang, H.-T. Chiao, S. Abimannan, Y.-P. Huang, Y.-T. Tsai, and K.-M. Lin, "An LSTM-based aggregated model for air pollution forecasting," *Atmos. Pollut. Res.*, vol. 11, no. 8, pp. 1451–1463, Aug. 2020.
[26] L. Zhang, P. Liu, L. Zhao, G. Wang, W. Zhang, and J. Liu, "Air quality predictions with a semi-supervised bidirectional LSTM neural network," *Atmos. Pollut. Res.*, vol. 12, no. 1, pp. 328–339, Jan. 2021.
[27] M. S. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in *Proc. 31st Annu. Conf. Neural Inf. Process. Syst. (NIPS)*, vol. 30, 2017, pp. 1–10.
[28] R. Navares and J. L. Aznarte, "Predicting air quality with deep learning LSTM: Towards comprehensive models," *Ecol. Informat.*, vol. 55, Jan. 2020, Art. no. 101019.
[29] A. B. Arrieta, N. Díaz-Rodríguez, J. Del Ser, A. Bennetot, S. Tabik, A. Barbado, S. Garcia, S. Gil-Lopez, D. Molina, R. Benjamins, R. Chatila, and F. Herrera, "Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI," *Inf. Fusion*, vol. 58, pp. 82–115, Jun. 2020.
[30] W. Zhang, Y. Wen, K. J. Tseng, and G. Jin, "Demystifying thermal comfort in smart buildings: An interpretable machine learning approach," *IEEE Internet Things J.*, vol. 8, no. 10, pp. 8021–8031, May 2021.
[31] H. Wang. (2019). *Air Pollution and Meteorological Data in Beijing 2017–2018*. [Online]. Available: https://doi.org/10.7910/DVN/USXCAK
[32] K.-F. Jea, C.-W. Hsu, and L.-Y. Tang, "A missing data imputation method with distance function," in *Proc. Int. Conf. Mach. Learn. Cybern. (ICMLC)*, Jul. 2018, pp. 450–455.
[33] C. Zhang, H. Feng, K. Jing, and T. Yang, "Nearest neighbour filling algorithm for missing data based on cluster analysis," *Comput. Appl. Softw.*, vol. 31, no. 5, pp. 282–284, 2014.
[34] L. Beretta and A. Santaniello, "Nearest neighbor imputation algorithms: A critical evaluation," *BMC Med. Informat. Decis. Making*, vol. 16, no. S3, pp. 197–208, Jul. 2016.
[35] D. Zeng, D. Xie, R. Liu, and X. Li, "Missing value imputation methods for TCM medical data and its effect in the classifier accuracy," in *Proc. IEEE 19th Int. Conf. e-Health Netw., Appl. Services (Healthcom)*, Oct. 2017, pp. 1–4.
[36] K. T. Chui, B. B. Gupta, and P. Vasant, "A genetic algorithm optimized RNN-LSTM model for remaining useful life prediction of turbofan engine," *Electronics*, vol. 10, no. 3, p. 285, Jan. 2021.
[37] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
[38] K. Cho, B. van Merrienboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using RNN Encoder–Decoder for statistical machine translation," in *Proc. Conf. Empirical Methods Natural Lang. Process. (EMNLP)*, 2014, pp. 1–15.



YUTING YANG is currently pursuing the Ph.D. degree with the China University of Geosciences, Beijing. Her main research interests include deep learning, data mining, and geological hazards analysis.



GANG MEI received the bachelor's and master's degrees from the China University of Geosciences, Beijing, and the Ph.D. degree from the University of Freiburg, Germany, in 2014. He is currently an Associate Professor in numerical modeling and simulation in civil engineering with the China University of Geosciences. He has published more than 70 research papers in journals and academic conferences. His main research interests include numerical simulation and computational modeling, including FEM analysis, GPU computing, data mining, deep learning, and network science and applications. He has been serving as an Academic Editor for the journals *PeerJ Computer Science*, IEEE ACCESS, and *SN Applied Sciences* and as a Guest Editor for the journal IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS.



STEFANO IZZO received the master's degree in mathematics from the University of Naples Federico II, where he is currently pursuing the Ph.D. degree in applied mathematics and computer science with the Department of Mathematics and Applications "R. Caccioppoli." His research interests include deep learning and physics-informed learning methodologies. He is a member of the Mathematical mOdelling and Data AnaLysis Research Group.