Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
The impact of COVID-19 on urban PM$_{2.5}$ —taking Hubei Province as an example$^\star$

Kun Yang$^{a,b,1}$, Changhao Wu$^{b,c,1}$, Yi Luo$^{a,b,7}$

$^a$ Faculty of Geography, Yunnan Normal University, Yunnan, 650500, China
$^b$ GIS Technology Research Center of Resource and Environment in Western China, Ministry of Education, Yunnan Normal University, Yunnan, 650500, China
$^c$ School of Information Science and Technology, Yunnan Normal University, Yunnan, 650500, China

ARTICLE INFO

Keywords:
PM$_{2.5}$
COVID-19
Anthropogenic activities
Air mass transport
Joint prevention

ABSTRACT

In January 2020, China implemented strict lockdown measures due to the invasion of the new coronavirus, which led to a sharp decline in the contribution of anthropogenic fine particulate matter (PM$_{2.5}$). The special period of COVID-19, especially in Hubei where the epidemic was the most severe, provides excellent research conditions for studying the contribution of anthropogenic activities to PM$_{2.5}$ pollution. We used an optimized deep learning model to predict PM$_{2.5}$ concentration during the epidemic period in the cities of Hubei Province. The contributions of local anthropogenic activities to PM$_{2.5}$ pollution were obtained by contrasting the predicted results with actual site observations.

However, a strange phenomenon was revealed that Yichang, a city with low local anthropogenic contribution to PM$_{2.5}$, was found to have severe haze in winter conflicting with our previous expectations. After further research, we found that an increased conversion of secondary aerosols caused by long-distance transport of pollutant gases from the northern region is the main cause of winter haze pollution in this city. This finding highlights the importance of joint regional prevention and control of air pollution.

1. Introduction

PM$_{2.5}$ refers to atmospheric fine particulate matter with an aerodynamic diameter less than or equal to 2.5 $\mu$m. Anthropogenic activities play an important role in the influence of atmospheric pollutants such as PM$_{2.5}$. With more than half of the world’s population living in cities (Lederbogen et al., 2011), urbanization, industrialization and increasing fossil-fuel consumption are the main causes of poor air quality (Raza et al., 2020), and urban areas have become significant contributors and victims of air pollution.

The contributions from urban anthropogenic activities to air pollution mainly include automobile exhaust, industrial processes (metal smelting, chemical plants, paper and pulp industries, and oil refineries) (Feng et al., 2015), anthropogenic construction dust, ground dust, catering industry emissions and other forms of emissions. Studies have shown that long-term exposure to PM$_{2.5}$ is significantly associated with respiratory-, lung cancer-, and cardiovascular-related mortality (Brook et al., 2010; Pun et al., 2017). More studies have shown that high concentrations of PM$_{2.5}$ are associated with high COVID-19 mortality (Coker et al., 2020). China is experiencing extremely rapid urbanization and has become a high-risk area for PM$_{2.5}$ pollution.

In mid-to-late January 2020, cases of new coronavirus infection were detected in Wuhan, China. Subsequently, epidemics broke out in countries around the world. Alongside China, with the successive implementation of blockade policies, reduced personnel flow, stagnant traffic operations, and shutdown of factories and enterprises, the most polluting countries such as India and the United States have reduced their PM$_{2.5}$ emissions due to COVID-19. Research shows that within 21 days of India’s lockdown, the PM$_{2.5}$ concentration in Kolkata is reduced by 34.52%, and 27.57% in Delhi, capital of India (Singh and Chauhan, 2020). The USA thus accounted for 33% and 27% of worldwide COVID-19 cases and deaths, respectively (Pata, 2020). It has been blocked for 26 days, but urban PM$_{2.5}$ has only decreased by 4.70% (Berman and Ebisu, 2020). In Hubei, China, PM$_{2.5}$ has been reduced by 29%, and Wuhan, the most severely affected city, has been locked down for 76 days, reducing it by 35% (Chu et al., 2021). Despite the difference...
in the degree of the outbreak, the time of policy formulation, the strictness of the lockdown, and the difference in regional production activities, COVID-19 has eased the local atmospheric pressure in most of these areas. However, lockdowns do not simply mean reduced pollution. Further research found that the mass percentages of sulfate (3.05%), organic carbon (2.48%) and secondary inorganic aerosols (8.70%) increased, which suggested an enhanced secondary formation of PM$_{2.5}$ (Zheng et al., 2020). Even in Shanghai during the COVID-19 outbreak, a puzzling haze event was observed (Chang et al., 2020). Some existing atmospheric chemical models, such as the Compositions of the Models-3 Community Multiscale Air Quality (CMAQ) and the Weather Research and Forecasting Model (WRF) (Binkowski and Roselle, 2003; Tuccella et al., 2012), can simulate the distribution and source of organic aerosols by taking primary organic aerosols into account. However, studies have shown that the simulation results of this type of models are greatly affected by simulation scale, emission inventory, driving-fields and chemical mechanism, making it difficult to apply to single satellite observation and extreme conditions (e.g. heavy pollution events) (Y Zhang and Li, 2015). Large bias errors are found not only in CMAQ but are also in many other air quality forecast models when predicting PM and its precursors (Solazzo et al., 2012; Djkalova et al., 2015). Therefore, more effective methods need to be implemented to combine multiple factors to reveal the causes of abnormal phenomena during COVID-19.

Studies have shown that nonparametric machine learning algorithms such as random forests are better than traditional regression models in estimating ground PM$_{2.5}$ concentrations (Chen et al., 2018). A generalized additive ensemble model was used to establish a prediction model, and the accuracy reached the monthly average data ($R^2 = 0.76$, RMSE = 29.16 μg/m$^3$) (Xiao et al., 2018). Some studies have proposed a geographically weighted gradient boosting machine (GW-GBM) method in predicting PM$_{2.5}$ ($R^2 = 0.76$, RMSE = 23.0 mg/m$^3$) (Zhan et al., 2017). In summary, machine learning models have reached high accuracy for estimating and predicting the concentration of PM$_{2.5}$ and have high use value.

The special period of COVID-19, especially in Hubei, where the epidemic was the most severe and implemented the most thorough blockade, provides excellent research conditions for studying the contributions from anthropogenic activities to PM$_{2.5}$ concentrations. Based on previous research (Luo, 2018; Shi, 2020; Yang, 2018, 2019, 2020a–c), we took the small and medium-scale cities as the research area to build an improved deep learning model to predict PM$_{2.5}$ pollution in Hubei under the assumption that there is no epidemic situation. Compared with the actual observations at the site, the approximation of the contributions from local anthropogenic activities to PM$_{2.5}$ pollution are obtained. Furthermore, we not only explore the differences in the contribution of anthropogenic activities to air pollution under different urban scales, and analyze the causes and mechanisms, we also use Yichang as an example to reveal the phenomenon of smog pollution caused by external transport. Our research provides the basis for the government’s air pollution prevention and control decision-making.
China has a large land area and diverse landforms. Pollution prevention and control must use complex systems engineering thinking. In the face of external and internal pollution, different measures must be taken to control from the source.

2. Study region and data

2.1. Study region

Hubei Province, whose capital is Wuhan city, is geographically located in central China and the middle reaches of the Yangtze River. It is in the transition zone from the second step to the third step of China’s topography. The area where the first confirmed cases of the novel coronavirus were detected in China was also the hardest hit area during COVID-19.

As shown in Fig. 1(a), (b), the terrain of Hubei Province is high on three sides, low in the middle, open to the south, and gaping in the north, showing an incomplete flat land slightly inclined from the northwest to the southeast. Hubei Province is located in the subtropical zone in a typical monsoon zone. The southeast monsoon from the Pacific Ocean prevails in summer, and the northwest wind from Mongolia-Siberia prevails in winter (Tang et al., 2021). Hubei Province has a relatively low proportion of primary industry (T Zhang and Zhang, 2018). See TextI for industrial structure.

2.2. Ground-level PM$_{2.5}$ measurements

This study used hourly data from 54 monitoring sites in 13 prefecture-level cities in Hubei Province obtained from the national urban air quality real-time release platform from January 2013 to July 2020. The site distribution is shown in Fig. 1.

2.3. Satellite observations

This study obtained monthly average MOD08 M3 and MYD08 M3 files (spatial resolution of 1° × 1°) from January 2013 to July 2020 from the atmospheric archive and distribution system to obtain aerosol optical depth (AOD) and selected the use of a dark blue land aerosol data product at 550 nm retrieved by the algorithm (parameter name: DeepBlue_Aerosol_Optical_Depth_550_Land_Mean_Mean). All data used are shown in Table S1.

2.4. Meteorological data

Meteorological data can effectively improve the accuracy of the AOD-PM$_{2.5}$ relationship model. Obtained from January 2013 to July 2020 from the European Centre for Medium-range Weather Forecast, the five kinds of monthly average meteorological data with a spatial resolution of 0.25° × 0.25° were used as covariates of the PM$_{2.5}$ concentration estimation model in this study, including 2 m temperature (T2M, K), east component of 10 m wind (10U, m s$^{-1}$), north component of 10 m wind (10V, m s$^{-1}$), boundary layer height (BLH, m), relative humidity (RH, %) and total precipitation (TP, m).

2.5. Other auxiliary data

2.5.1. GDP

The data selected are the Gross Domestic Product (GDP) data of Hubei Province in 2019. The data source was the “2019 Hubei Provincial Statistical Yearbook”. In the course of model training, to obtain the economic characteristics of cities, we used the GDP of various cities in Hubei province in 2019 before the outbreak of influenza as a proxy variable to annotate the data of different cities and achieve urban division indirectly. We used the GDP data of Hubei in 2018 to analyzes the industrial structure of cities in Hubei and provide the basis for the analysis of anthropogenic PM$_{2.5}$.

2.5.2. Normalised seasonal coefficient (NSC)

Marking the months with 1–12 ignores the periodicity and continuity of the time series. To make the model better capture the seasonal changes in PM$_{2.5}$ data, this study added a normalized seasonal index (NSC) to characterize the temporal characteristics of the data. The normalized seasonal index uses the position of the point where the sun illuminates the earth directly to represent different months. As shown Figure S1, by mapping the latitude 23.5° of the point of direct sunlight on summer solstice to 1, and the latitude −23.5° of the point of direct sunlight on winter solstice to −1, the NSC values at all time of the year can be obtained. Formula (1) shows the calculation principle of NSC, PDS (Point of Direct Sun) indicates the latitude of the point of direct sunlight which is positive in the northern hemisphere and negative in the southern hemisphere.

$$NSC = 2 \times \left( \frac{PDS + 23.5}{47} \right) - 1$$

(1)

2.6. Data preprocessing and matching

2.6.1. Data spatiotemporal matching

MOD08 is a three-level raster data product in MODIS products, and the data are stored in a 360 × 180 matrix. Each grid corresponds to 1° × 1° of latitude and longitude. According to the latitude and longitude of each station, the grid where each station is located can be located to extract the monthly average AOD of each station. The missing AOD values are supplemented by kriging interpolation and Text S2 shows this process. Similarly, meteorological data is also in raster data file format and the same method was used to match the grid with the monitoring station location. We used hourly PM$_{2.5}$ data to get monthly average data.

2.6.2. Correlation collinearity test

To remove the unit limit of the data, the data are converted into a dimensionless pure value through the normalization of the data, and the data are uniformly mapped to the [0, 1] interval; The 7 parameters, T2M, BLH, TP, RH, WS, AOD, and GDP, are normalized. Statistical Product and Service Solutions tool was used to analyze the correlation and multicollinearity between input parameters. Text S3 shows the analysis process, and finally decided to delete RH and use other parameters to join the model.

3. Methods

3.1. Wavelet period analysis

Due to the obvious seasonal heterogeneity of PM$_{2.5}$ concentrations, most areas of Hubei Province are affected by the subtropical monsoon climate and human production and management activities. To better construct the PM$_{2.5}$ concentration prediction model, the monthly average PM$_{2.5}$ concentration data from January 2013 to July 2020 at the 1325A air-quality monitoring station in Wuhan city were selected for DB1 wavelet cycle analysis to determine the periodic nature of PM$_{2.5}$ change in the region.

3.2. PM$_{2.5}$ concentration prediction

The back propagation neural network (BPNN) has the advantages of a high degree of self-learning and self-adapting ability, nonlinear mapping ability, generalization ability and fault tolerance. However, there are problems, such as slow convergence speed, sensitive initial weights leading to local minimization, and sample dependence. The Long Short Term Memory networks (LSTM) has a long-term memory function and has good predictive ability for time series. It mainly uses a gate mechanism to solve the problem of gradient explosion and gradient disappearance to a certain extent, but its performance is still lost due to its inability to parallelize. It is expressed by the following formula:
i = σ(Wixi + Whhi + bi)  
\[ \text{(2)} \]

f = σ(Wifxi + Whfhi + bf)  
\[ \text{(3)} \]

o = σ(Wiox + Whohi + bo)  
\[ \text{(4)} \]

c = fi ⊙ ci−1 + i ⊙ tanh(Wixi + Whhi + bi)  
\[ \text{(5)} \]

h = oi ⊙ tanh(ci)  
\[ \text{(6)} \]

- was the matrix element by dot product, b, b, and b are the deviation vectors in the output layer, the input layer and the forget gate respectively. σ(x) stood for sigmoid function. W was the weight matrix of the corresponding layer, W played the role of the weight matrix of the input layer to the forget gate, W played the role of weight matrix of the hidden layer to the input layer, W played the role of the weight matrix of the hidden layer to the output layer; ci played the role of update cell status (Yu et al., 2019). Formula (5) shows the process of information transmission control by forgetting gate, input gate and output gate.

Both models have advantages and disadvantages. Combining the two models can achieve the effect of optimizing the problem of slow convergence rate sample dependence and local minimization, stabilizing the prediction results of the model, and increasing the accuracy of the model prediction.

As shown in Fig. 2, the construction of the Gauss-LSTM & BP model includes 3 processes:

Step1: Multiple training

The LSTM and BP models are trained ten times, and ten LSTM prediction models and ten BP prediction models are obtained. L1, L2, ..., L10 and B1, B2, ..., B10, and record the training error of each model, which are denoted as EL1, EL2, ..., EL10 and EB1, EB2, ..., EB10. The prediction results of each model are denoted as RL1, RL2, ..., RL10 and RB1, RB2, ..., RB10.

Step2: Gaussian optimization

The normalization method is used (8) to map the model errors P(EL1, EL2 ..., EL10) and Q(EB1, EB2 ..., EB10) to the interval of [0,1]. Since we assume that a low value does not mean a value of 0, the minimum value normalization is set to 0.1. The normalized errors Nor-EL1, Nor-EL2, ..., Nor-EL10 and Nor-EB1, Nor-EB2, ..., Nor-EB10 are obtained.

The Gaussian kernel function (7) is used to calculate the weight of each original model, and the Gaussian coefficients GaussL and GaussB of each model are obtained through (9). When σ = 0.5, the weight of each model has a large difference, thus σ = 0.5 is chosen.

The Gaussian coefficient GaussL is multiplied by the predicted value of each model, and the weighted sum is used to obtain the Gauss LSTM model. In the same way, the Gauss_BP model is obtained. The results obtained by formula (10) are recorded as RGL and RGB. Model errors are recorded as EGL and EGB.

\[ f(x) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right) (\mu = 0, \sigma = 0.5) \]
\[ \text{(7)} \]

\[ \text{NorEL} = (1 - 0.1) \times \left(\text{EL} - \text{MIN}(\text{EL} \in P)\right) / \left(\text{MAX}(\text{EL} \in P) - \text{MIN}(\text{EL} \in P)\right) + 0.1 \]
\[ \text{(8)} \]

\[ \text{GaussL} = \frac{f(\text{NorEL})}{\sum_{i=1}^{10} f(\text{NorEL})} \quad \text{GaussB} = \frac{f(\text{NorEB})}{\sum_{i=1}^{10} f(\text{NorEB})} \]
\[ \text{(9)} \]

\[ \text{RGL} = \sum_{i=1}^{10} \text{GaussL} \times \text{RL}_i, \text{RGB} = \sum_{i=1}^{10} \text{GaussB} \times \text{RB}_i \]
\[ \text{(10)} \]

Step3: Gauss-LSTM&B

From EGL and EGB, the inverse proportional method and formula (11) are used to calculate the weights of the Gauss LSTM model and Gauss BP model. Formula (12) combines the Gauss LSTM model and Gauss BP model to calculate the Gauss LSTM&B prediction result RGL&B.

\[ \text{WGL} = \frac{1}{\text{WGL}} + \frac{1}{\text{WGB}} \quad \text{WGB} = \frac{1}{\text{WGL}} + \frac{1}{\text{WGB}} \]
\[ \text{(11)} \]

\[ \text{RGL&B} = \text{RGL} \times \text{WGL} + \text{RGB} \times \text{WGB} \]
\[ \text{(12)} \]

In order to avoid the over-fitting of the model, the accuracy of the model was evaluated by Cross Validation (CV). K-Fold method cannot be
simply used for CV of time series model prediction (G Zhang et al., 2020), but Rolling Origin Evaluation method is usually used. As shown in Figure S3, the model was verified 5 times, and each time the test set was moved forward by a time window, and the validation set and training set were moved accordingly. See Text S4 for verification method.

3.3. Kriging interpolation

This study uses the kriging algorithm for two purposes: 1) Meteorological data or AOD data are missing due to cloud cover or sensor problems. We use the valid values around the missing values to supplement the missing data with the Kriging interpolation method. 2) A heat map of PM contributions from anthropogenic activities in Hubei Province is made.

3.4. HYSPLIT-4 backward air-mass trajectory

The concentrations of PM$_{2.5}$ in Hubei Province are affected by internal source emissions and external transport. To explore the transport trajectory of PM$_{2.5}$ in Hubei Province and the impact of external air pollutants, the source of air pollution in Hubei Province is determined after adopting HYSPLIT-4 research on air-mass trajectory analysis.

3.5. Anthropogenic PM$_{2.5}$ index (APMI)

Through the prediction model, the PM$_{2.5}$ concentration that should be present in Hubei assuming no outbreak is obtained and then compared with the actual PM$_{2.5}$ concentration in Hubei Province after the outbreak. By calculating the difference between the actual result and the predicted result, the PM$_{2.5}$ emissions caused by anthropogenic activities are obtained. We define the monthly average of anthropogenic PM$_{2.5}$ emissions calculated from each city as their Anthropogenic PM$_{2.5}$ Index (APMI).
3.6. Green economy index (GEI)

This research proposes the Green Economy Index (GEI) based on APMI values and the GDP of Hubei Province in 2018. By calculating the GEI of each city and using the kriging interpolation method, the GEI heat map of Hubei Province is obtained. As formula (13) shows:

\[
GEI = \frac{APMI}{GDP} \quad \text{(Unit: } \mu g/m^3/100 \text{ million yuan)}
\]

(13)

4. Results and discussion

4.1. Periodic changes in PM$_{2.5}$ concentration

Figure S4 shows the result of wavelet analysis and the results were analyzed in Text S5. Through wavelet analysis, it can be seen that there is an obvious periodicity in the change in PM$_{2.5}$ concentrations. However, in December 2019, at the beginning of COVID-19, the periodic change in PM$_{2.5}$ led to mutation. The 12-months period scale is selected as the prediction time length for constructing the prediction model. Text S6 introduces the division method of the training set and test set of the machine learning model of this research.
4.2. Prediction model accuracy analysis

Text S7 illustrates the hyperparameters of the model. Fig. 3(c)-(i) shows the fitted scatter plots of the normal model and the improved model. The specific error results are shown in Table S6. From this we can see that the various indicators of the Gauss-LSTM & BP model perform best and there is no over-fitting phenomenon obviously. Gauss-LSTM & BP’s CV-R^2 = 0.76, GPR (CV-R^2 = 0.67) the worst performance, other models are about 0.7. The RMSE of Gauss-LSTM & BP is 8.06, the error is the smallest, followed by LSTM: 8.94 and LSBoost: 9.45, BP, SVM and GPR all exceed 10, showing poor performance. The mape of Gauss-LSTM & BP is 8.06, the error is the smallest, and the error of GPR: 0.35 is the largest. The MAE of Gauss-LSTM & BP is 6.24, which is better than all other models. Compared with other models, the average MAE of Gauss-LSTM & BP is reduced by 24%. In short, in the prediction of time series sequences, combining the memory characteristics of LSTM and the high degree of self-learning and self-adapting ability of BP, the accuracy of the model has been significantly improved. The actual prediction error reaches R^2 = 0.86 and the slope = 1.067, which has a high precision. It can be used for further predictive research.

4.3. Spatial characteristics of the impact of anthropogenic activities on \( PM_{2.5} \)

As shown in Fig. 3 (b) of GEI, in Huanggang, Huangshi, Ezhou, and Xianning, medium level economic growth is accompanied by extremely high anthropogenic PM2.5 emissions, which indicates that the industrial development model of the Eastern Yangtze River Economic Belt is still relatively uneven, and the situation of resource consumption and environmental pollution is not optimistic. In the eastern region, the high values decayed and diffused around the urban agglomeration area with the direction of passenger flow during the Spring Festival. Before the Spring Festival (mid-January), the population flowed out of Wuhan to other cities. During the Spring Festival, the population flowed into Wuhan from other cities, which caused this phenomenon. The analysis results of the central cities Jingmen, Jingzhou, and Xiangyang and the western cities Yichang, Shiyan, Enshi and Suizhou are in Text S8.

4.4. Time characteristics of the impact of anthropogenic activities on \( PM_{2.5} \)

The pre-epidemic model errors (from July 2019 to December 2019) are taken as the model’s own errors, and corrected predictive value is obtained after deducting model's own errors from the predicted results from January 2020 to July 2020. The predictive value caused by anthropogenic activities was indicated by the green line. As shown in Fig. 4, time is divided into B-COVID19, A-COVID19 and E-COVID19, representing the pre-outbreak period, peak period, and the end of COVID-19 phase, respectively. In the eastern city clusters such as Huangshi, Ezhou, Huanggang, Xian and Xiaogan, anthropogenic activity factors consistently appeared as two large and one small peak in February 2020 and June 2020, respectively. The anthropogenic activity here refers to the situation that will occur if there is no epidemic. Part of the peak in February is due to the Chinese Lunar New Year period, which is a relatively cold time period in winter and corresponded with the Spring Festival holiday and the resumption of production. There is a concentrated surge in traffic volume around the Spring Festival compared with normal times. Automobile exhaust emissions, the beginning of enterprise production, the pollutants emitted by fireworks and fireworks crackers into the atmosphere, and the increase in electricity and coal-burning consumption caused by the demand for heating in winter are all reasons for anthropogenic PM2.5 emissions during this period. Every June is the time for the concentrated holidays of various universities across the country. Various economic activities, such as tourism activities and service industries, are active and concentrated during this time. Therefore, there is a sharp increase in traffic and transportation during this period. This is the reason for the peak in June. Compared with A-COVID19, the model gradually recovers accuracy during E-COVID19. The anthropogenic PM2.5 value displayed at this time will also decrease. Wuhan is more unique. There are two large and one small peaks in the order of occurrence, and the time of occurrence is somewhat different from that in the five other high-value cities. The first two large peaks appear in January and March and appear Bimodal distribution, and the smaller peak appears in mid-May, which may be related to the particularity of Wuhan as the provincial capital and the direction of passenger flow during the Spring Festival. Before the Spring Festival (mid-January), the population flowed out of Wuhan to other cities. During the Spring Festival, the population of Wuhan was at a low point in the year. After the Spring Festival (early March), the population flowed into Wuhan from other cities, which caused this phenomenon. The analysis results of the central cities Jingmen, Jingzhou, and Xiangyang and the western cities Yichang, Shiyan, Enshi and Suizhou are in Text S8.

Fig. 4 shows that the impact of COVID-19 on the concentration of PM2.5 in Hubei Province is characterized by high values in the east and low values in the west. During the lockdown period (1, 2, and March), the impact in Huanggang has the largest decrease of 34.40%, the impact in Yichang has the smallest decrease of 4.21%, and the most severely affected area, Wuhan, is at 26.38%; Hubei Province has decreased by 22.61% on average. Although the impact of the epidemic on the PM2.5 concentration of Hubei and Wuhan in our study is slightly smaller than that of the Chu’s study (Chu et al., 2021), this might because we used the forecasting method and they used the method compared to previous years, but the impact on Wuhan is always greater than Hubei Province, production and management activities in the western region and the absorption of PM2.5 by vegetation have led to the formation of low-value APMI areas in the western region where anthropogenic activities cause minimal environmental pollution.

Fig. 1(d) shows that for Xiangyang and Yichang, GDP ranks second and third in the province. However, the GEIs of these two cities are at a low level in the province, especially Yichang city. Due to the transformation of the economic development mode and the absorption of natural conditions to pollution, the economic development of Yichang city and Xiangyang shows low anthropogenic pollution emissions.
which is consistent.

4.5. Analysis of external pollution

Yichang and Xiangyang, as the major economic cities in Hubei Province and even the central region, are the two major provincial sub-central cities. From the perspective of industrial output value, Xiangyang city completed an industrial output value of 207.7 billion yuan in 2019, of which the output value of heavy industry grew rapidly and accounted for a relatively large proportion. The industrial output value of Yichang city was 205.4 billion yuan, light-industry industries such as the chemical industry, new materials, food and biomedicine accounted for a relatively large proportion.

In these two cities, high levels of PM$_{2.5}$ pollution are detected every winter. As a result of winter biomass burning, primary emissions are the main source of PM$_{2.5}$ (Gelencsér et al., 2007). It is easy to think of this as the result of local industrial emissions and fossil-fuel combustion. However, the above research results show that due to the low energy consumption development mode adopted by Xiangyang and Yichang, anthropogenic PM$_{2.5}$ pollution is relatively low in the province, and the economic development model is a green economy. Therefore, it is speculated that the high concentrations of PM$_{2.5}$ pollution in Xiangyang and Yichang each winter may come from external transport.

By obtaining the monthly average PM$_{2.5}$ concentration of each city, a heat map of the PM$_{2.5}$ concentration in Hubei Province from November 2019 to July 2020 was produced using the kriging interpolation method. Figure S7 shows that starting in November 2019, the central and northern parts of Hubei Province, with Xiangyang as the center point, experienced medium-level air pollution. By December, high-concentration PM$_{2.5}$ pollution had already spread from Xiangyang city from north to south. The situation spread to Jingmen and Yichang. In January, high level of PM$_{2.5}$ pollution began to recede from south to north, and in February, the pollution basically subsided.

Fig. 6(a) shows that PM$_{2.5}$ and its 15-day moving average experienced a rapid increase, rapid decrease, rapid increase, and slow decrease in PM$_{2.5}$ concentration during the period from November 2019 to March 2020 in Yichang city. The trend of the 15-day moving average of (PM$_{2.5}$/SO$_2$) $\times$ 3, which symbolizes the mobile pollution source, is consistent with the trend of PM$_{2.5}$.

The simulation results of HYSPLIT backward air masses were combined to analyze the backward trajectory of Yichang city for approximately 80 h; on November 21, 2019, the cold and dry airflow from Siberia flowed to the northwestern region of Inner Mongolia, Shaanxi, etc., via Anqing city, Wuhan city, and transiting Yichang city. Fig. 6(b) shows the backward trajectory and some meteorological parameters in Yichang City. See Figure S7 for detailed meteorological parameters. Fig. 6(b) shows that more than 90% of the airflow comes from short-distance transport. Studies have shown that low temperature and high humidity in winter will promote the formation of sulfate, nitrate, and ammonium (SNA) aerosols (Chang et al., 2020). Since the main airflow contribution from source cities are nearby cities, the PM$_{2.5}$ concentration is low, and low humidity airflow from Siberia inhibits the formation of secondary aerosols from chemical reactions, such as sulfate and nitrate. Yichang city exhibits low PM$_{2.5}$ concentrations. PM/CO = 75 and PM/SO$_2$ = 5, which indicates a moderate proportion of secondary aerosols and a low proportion of mobile pollution.

On December 16, 2019, the airflow from eastern Siberia passed through the Beijing-Tianjin-Hebei region, carried large amounts of water vapor across the Bohai Sea and landed on the Shandong Peninsula via Dongying, Bozhou, and transiting Yichang. In Tangshan, RH = 99%, and TM = 0/1 °C (Format: TM = minimum temperature/maximum temperature °C); in Dongying and Bozhou, RH are 93% and 96%, respectively. The PM/CO values from Tangshan to Yichang are 32.8, 56.3, 40.9, and 102.5 in the order of air flow. This shows that when the air mass reaches Yichang, the proportion of secondary aerosols increases significantly. Aqueous-phase oxidation of SO$_2$ (via Text S10 reactions (1) to (2)) can be an important formation pathway of sulfate at high RH during haze events (Cheng et al., 2016; Xue et al., 2019), and the transformation of HNO$_3$ into the particle phase is generally enhanced at high RH and low temperature conditions (Liu et al., 2015).
provides good low OH conditions in the formation of NO$_3^-$ in urban areas of China, and the contribution of N$_2$O$_5$ hydrolysis to NO$_3^-$ formation is the major oxidation pathway (H W Xiao et al., 2020) (via Text S10 reactions (3), (4)). Low temperature in winter is also beneficial to the partitioning of ammonium nitrate in the particulate phase (Yao et al., 2002). Moreover, NaCl in the ocean will also react with NO$_2$ to produce sea salt aerosols (SSAs) (Fu et al., 2010; Karlsson and Ljungström, 1995) via Text S10 reaction (5). The Beijing-Tianjin-Hebei region, led by Tangshan city, contributes more than 90% of the air flow to Yichang city. High energy consumption industries account for a large proportion of the Beijing-Tianjin-Hebei region, and energy consumption increases rapidly. Automobile exhaust emissions and coal burning for heating (Duce et al., 2008; Galloway et al., 2004) produced NO$_x$ and SO$_x$ which coupled with the cold air flow in the northeast and water vapor on the ocean through long-distance transport and chemical reaction, explain the high concentration of PM$_{2.5}$ in Yichang city.

This result is consistent with the research conclusion of Chang (Chang et al., 2020), the abnormally high value of haze during COVID-19 is due to the generation of high secondary inorganic aerosol (mainly nitrate) in long-distance regional transport. They used chemical analysis, and we used machine learning to predict and find clues. According to the time of air mass transport, we made broken line charts of PM$_{2.5}$, PM$_{2.5}$/CO, PM$_{2.5}$/SO$_2$, and NO$_2$ concentrations along the four cities of air mass transport. Fig. 5 shows that in the process of airflow transport, Yichang city and the other three transit cities show opposite trends in PM$_{2.5}$, PM$_{2.5}$/CO and PM$_{2.5}$/SO$_2$. Text S9 analyzed the results in Fig. 5. We discover that the conversion rate of secondary aerosols did not decrease because of the longer distance from Yichang. In contrast, more serious haze events were observed in Yichang. These results are consistent with the pollution transfer order of PM$_{2.5}$ along the city under the effect of air transport and the theory of long-distance transport of secondary aerosol transformation enhancement.

Studies have shown that although the probability is small (Rincón et al., 2012), photochemical smog may appear in winter (Field et al., 2015). This kind of smog is a secondary pollutant formed by ultraviolet radiation of primary pollutants emitted by cars or factories (Dickerson

Fig. 6. (a) Change in air pollution composition. (b) backward air-mass trajectory in Yichang city (The area covered by the red pixel block means that this area contributes more than 90% of the airflow to the Receptor point area). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)
et al., 1997). Since volatile organic compounds (VOCs) and NOx react photochemically to produce ozone (B N Zhang and Oanh, 2002), elevated O3 concentrations can be used as a sign of photochemical smog pollution. Figure S6 shows the temporal variation in ozone in the four cities. The ozone concentration of the four cities in the process of air-mass transport showed a downward trend. Tangshan, Dongying, Yichang, the O3 trend line is convex, which may indicate photochemical pollution. In addition, the airflow from the north may form a cold front (Wang et al., 2013), which will hinder the vertical convection of air, form an inversion layer (He et al., 2018) and aggravate air pollution through zone transfer process in the downwind area (Li et al., 2019).

The analysis of January 6, 2020, January 15, 2020 and February 14, 2020 is in Text S11.

5. Conclusions

Compared with other studies on the impact of COVID-19 on the concentration of air pollutants, we used the epidemic blockade as the research condition, and used the most thorough blockade of Hubei Province as the research area. Using an improved deep learning algorithm, a nonlinear solution to the contribution of anthropogenic activities to PM2.5 concentration is realized. It provides ideas for quantifying and seeking approximate solutions for the contribution of anthropogenic activities to nature.

Research findings show that the improved hybrid prediction model, Gauss-LSTM&BP, has a very good performance in predicting PM2.5 concentration (CV-RMSE = 8.31, CV-R2 = 0.80). The current model algorithm does not make full use of the spatio-temporal information characteristics of PM2.5 and the process mechanism of its generation and transformation. In the future, using machine learning/deep learning algorithms combined with physical/chemical mechanism models, may be a better way to enhance inversion accuracy. At the same time, we also believe that with the improvement of sensor accuracy and reanalysis algorithms in the future, the auxiliary parameters of PM2.5 will have better data accuracy, which will play a key role in improving the accuracy of PM2.5 inversion.

Furthermore, we adopted a new macro-analysis method that uses the contribution of anthropogenic activities to air pollution rather than micro-scale chemical analysis. Taking Yichang as an example, we revealed the phenomenon of external pollution caused by inter-regional air mass transport in Hubei Province. China has a large land area and diverse landforms. We emphasize the use of complex systems engineering theory and source control thinking when formulating air pollution control policies.

The impact of COVID-19 on the concentration of PM2.5 in Hubei Province is characterized by high values in the east and low values in the west. During the lockdown period, the most severely affected area, Wuhan, had a decrease of 26.38% percent (17.3 µg/m^3) on concentration in Hubei Province decreased by 22.61% percent (14.4 µg/m^3) on average. Compared with Hubei Province, Wuhan has reduced more absolute values and percentages. This may be due to the fact that Wuhan City, as the political and economic center of Hubei Province, has a very large base of anthropogenic activity intensity. In addition, Wuhan has implemented more stringent and effective blockades and travel restrictions.

In summary, PM2.5 pollution has the characteristics of spatial heterogeneity and cross-regional transport, which will cause wrong decisions in regional quality prevention and control and industrial layout. Hubei Province has been used as an example to reveal the relationship between anthropogenic activities and PM2.5 and further analysis shows that the winter haze pollution in Yichang area originates from the air mass transport in the north. It provides a reference for the implementation of the government’s precise and quantitative air pollution control policies, which provides us with ideas and methods for the next step to carry out larger-scale (China or the world) research.

Author statement

Kun Yang: Conceptualization, Visualization, Investigation. Changhao Wu: Methodology, Software Data curation, Writing – original draft, Validation. Yi Luo: Visualization, Investigation, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The data of this study come from National Aeronautics and Space Administration (https://modis.gsfc.nasa.gov/data/), European Center for medium range weather forecasts (https://www.ecmwf.int/en/forecasters/datasets/era5), and China air quality real-time publishing platform (http://106.37.208.233:20035/). The experimental process data can be downloaded at https://zenodo.org/record/5709763#.YZYJ3MgznZs. This study was financially supported by the National Natural Science Foundation of China [41761084].

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envpol.2021.118633.

References

Berman, J.D., Ebisu, K., 2020. Changes in US air pollution during the COVID-19 pandemic. Sci. Total Environ. 739, 138646.

Binkowski, F.S., Roselle, S.J., 2003. Models-3 Community Multiscale Air Quality (CMAQ) model aerosol component I. Model description. J. Geophys. Res.: Atmosphere 108 (D6).

Brook, R.D., Rajagopalan, S., Pope III, C.A., Brook, J.R., Bhatnagar, A., Diez-Roux, A.V., Holguín, F., Hong, Y., Luepker, R.V., Mittleman, M.A., 2010. Particulate matter air pollution and cardiovascular disease: an update to the scientific statement from the American Heart Association. Circulation 121 (21), 2331–2378.

Cheng, Y., Huang, R.J., Ge, X., Huang, X., Hu, J., Duan, Y., Zou, Z., Liu, X., Lehmann, M., 2020. Puzzling haze events in China during the coronavirus (COVID-19) shutdown. Geophys. Res. Lett. 47 (12), e2020GL088523.

Chen, G., Li, S., Knibbs, L.D., Hamm, N.A., Cao, W., Li, T., Guo, J., Ren, H., Abramson, M.J., Guo, Y., 2018. A machine learning method to estimate PM2.5 concentrations across China with remote sensing, meteorological and land use information. Sci. Total Environ. 636, 52–60.

Cheng, Y., Zheng, G., Wei, C., Mu, Q., Zheng, B., Wang, Z., Gao, M., Zhang, Q., He, K., Carmichael, G., 2016. Reactive nitrogen chemistry in aerosol water as a source of sulfate during haze events in China. Sci. Adv. 2 (12), e1601530.

Chu, B., Zhang, S., Liu, J., Xu, Q., He, H., 2021. Significant concurrent decrease in PM2.5 and NO2 concentrations in China during COVID-19 epidemic. J. Environ. Sci. (China) 99, 346–353.

Coker, E.S., Cavalli, L., Fabrizi, E., Guanzella, G., Lippo, E., Parisi, M.L., Pontoriero, N., Rizzati, M., Varacca, A., Vergalli, S., 2020. The effects of air pollution on COVID-19 related mortality in northern Italy. Environ. Res. Resour. Econ. 76 (4), 611–634.

Dickerson, R.R., Kondragunta, S., Stenchikov, G., Cervera, K.L., Doddridge, B.G., Holben, B.N., 1997. The impact of aerosols on solar ultraviolet radiation and photochemical smog. Science 278 (5339), 827–830.

Djalalova, I., Delle Monache, L., Wilczak, J., 2015. PM2.5 analog forecast and Kalman filter post-processing for the Community Multiscale Air Quality (CMAQ) model. Atmos. Environ. 108, 76–87.

Duce, R.A., LaRoche, J., Aller, K., Arrigo, K.R., Baker, A.R., Capone, D., Cornell, S., Dentener, F., Galloway, J., Ganeshram, R.S., 2008. Impacts of atmospheric anthropogenic nitrogen on the open ocean. Science 320 (5878), 893–897.

Feng, X., Li, Q., Zhu, Y., Hou, J., Jin, L., Wang, J., 2015. Artificial neural networks forecasting of PM2.5 pollution using air mass trajectory based geographic model and wavelet transformation. Atmos. Environ. 107, 118–128.

Field, R.A., Sohls, J.J., Pérez-Ballesta, F., Grandesso, E., Montague, D.C., 2015. Distributions of air pollutants associated with oil and natural gas development measured in the Upper Green River Basin of Wyoming. Elem. Sci. Anth. 3, 000074.

Fu, Q., Zhuang, G., Li, J., Huang, K., Wang, Q., Zhang, R., Fu, J., Lu, T., Chen, M., Wang, Q., 2010. Source, long-range transport, and characteristics of a heavy dust pollution event in Shanghai. J. Geophys. Res.: Atmosphere 115 (D7).
