PM$_{2.5}$ concentration prediction in Lanzhou, China, using hyperchaotic cuckoo search—extreme learning machine

Wenjing Yang$^1$ · Juan Du$^1$ · Bowen Jia$^2$ · Ling Yang$^1$ · Renze Dong$^3$ · Xinyi Li$^4$

Accepted: 25 July 2022 / Published online: 17 August 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract
High concentrations of PM$_{2.5}$ cause environmental problems and many serious health effects, including heart and lung disease. To assess the PM$_{2.5}$ level with high accuracy, an advanced extreme learning machine (ELM) model is proposed, which combines the advantages of the hyperchaotic system and the cuckoo search algorithm. It improves the accuracy of the original ELM and avoids manual adjustment of parameters. The model examines the PM$_{2.5}$ concentration in Lanzhou, China, with daily predictions at four stations in 2018. It is also compared with different methods, such as the original ELM, multiple linear regression, and long short-term memory, and the results show that the proposed model obtains better forecasting performance than the others in terms of root mean squared error and the coefficient of determination ($R^2$), respectively. In addition, the model is applied to make short-term forecasts for four stations, predicting hourly PM$_{2.5}$ concentrations over the next week based on fourteen days of monitoring data. They are in high agreement with the monitored PM$_{2.5}$ concentrations. Our research indicates that the proposed model can facilitate effective measures to avoid exposure to high concentrations of PM$_{2.5}$. Meanwhile, it also provides a novel way to predict air pollution.

Keywords PM$_{2.5}$ prediction · Extreme learning machine · Hyperchaotic system · Cuckoo search

1 Introduction
Air pollution has become a growing problem in the 20th century. PM$_{2.5}$ is an atmospheric particulate matter with an aerodynamic equivalent diameter of up to 2.5 microns that can have an important impact on air quality. Many studies have confirmed that asthma, chronic obstructive pulmonary disease, and lung cancers are closely linked to PM$_{2.5}$, which can also have an impact on the cardiovascular system (Coleman et al. 2020; Ye et al. 2018). In China, rapid urban and industrial development in recent years has led to environmental problems of poor air quality in some regions. Lanzhou, a major city in Northwest China, is also plagued by air pollution. Therefore, this research focuses on air quality in Lanzhou and predicts long-term and short-term PM$_{2.5}$ concentrations. The long-term prediction refers to the assessment of daily levels of PM$_{2.5}$ during a year based on different variables, including temperature, humidity, sulfur dioxide, nitrogen dioxide, etc. The short-term forecast predicts the hourly PM$_{2.5}$ concentrations during a week. Achieving early warnings of air pollution by accurately predicting PM$_{2.5}$ concentrations is an
important step in the prevention and control of air pollution.

In addition to deterministic modeling based on physics and chemistry (Gilani et al. 2016), some scholars have also adopted statistical methods for air quality prediction. Compared to the former, the statistical modeling approach has received increasing attention, as it does not need to take into account complex chemical reactions and movement processes such as the dispersion of pollutants. The statistical models not only consist of traditional statistical models but also include artificial neural network (ANN) models and composite models. Traditional statistical models are mostly used to make predictions by analyzing patterns in the data, such as autoregressive (Zhou and Goh 2017) and multiple linear regression (MLR) models (Lesar and Filipcic 2021). Variations in PM$_{2.5}$ levels are influenced by a variety of factors, so simple linear models have limitations in assessing PM$_{2.5}$ concentrations.

Artificial neural networks, with their nonlinear mapping and adaptive learning capabilities, have been widely used in air environment prediction, such as feed-forward neural networks (Fu et al. 2015), recurrent neural networks (Biancofiore et al. 2017), and deep neural networks (Lightstone et al. 2021). With the development of computer science, the performance of composite models has been improved considerably, which has attracted attention in the development of various prediction models. For example, convolutional back-propagation neural networks (BPNNs) (Kow et al. 2020; Park et al. 2020), and weighted bagging based neural network combining image contrast-sensitive features, can be used to predict air pollution (Qiao et al. 2020). Hybrid convolutional neural networks combining high-level features extracted from convolutional layers with ground truth PM$_{2.5}$ can learn haze-related classification mappings and train support vector regression, indicating cheapness, quickness, and convenience (Li et al. 2020). Moreover, neural networks can derive a kind of multivariate Bayesian uncertainty processor to predict the probabilities of PM$_{2.5}$ (Zhou et al. 2020) or develop air quality early-warning systems through phase-space reconstruction and multi-objective optimization (Wang et al. 2020). In addition, the aerosol optical depth-PM$_{2.5}$ model (Xu et al. 2021) used a combination of MLR, BPNN, classification and regression trees, and random forest estimation methods to forecast PM$_{2.5}$ concentrations in eastern China, and the accuracy was improved by adding meteorological elements that vary over time and height.

The deterioration of air quality has affected public health in China (Wang et al. 2020; Jin et al. 2020; Schraufnagel et al. 2019), so a number of studies on the prediction of PM$_{2.5}$ concentrations in Chinese cities have been presented. Chinese researchers (Xiong et al. 2020) proposed a hybrid model that can rapidly estimate particulate pollution based on a data-driven ANN. At the regional level, Zhao et al. (2020) used recurrent neural networks to analyze hourly air quality in Northwest China. In addition, Wang et al. (2021) exploited an ANN to predict and analyze the PM$_{2.5}$ concentration in Chongqing, China. In terms of small-scale predictions, Tong et al. (2020) and Xu and Liu (2020) investigated machine learning to analyze PM$_{2.5}$ concentrations in households in Hong Kong and near the Beijing railway station, respectively.

Most existing algorithms are aimed at predicting short-term or long-term pollutant concentrations. There are rarely any algorithms that can perform well in both long-term and short-term PM$_{2.5}$ concentration forecasting. Therefore, it is meaningful to put forward a model that can accurately predict both the long-term and short-term PM$_{2.5}$ concentrations. The neural network is complicated to calculate, occupies many system resources, and has poor prediction accuracy for PM$_{2.5}$ concentrations. There is a correlation between the PM$_{2.5}$ concentration and time series, so an extreme learning machine (ELM) approach is employed to predict the PM$_{2.5}$ concentration. Avoiding manual adjustment, the cuckoo search (CS) algorithm can be applied to calculate the optimal number of hidden layer neurons of the ELM. For the adequate convergence and randomness of hyperchaotic systems (Li et al. 2019), its joint can reduce the computational cost of the CS algorithm and improve the accuracy of ELM (Jia et al. 2020). Inspired by the above, this research proposes the hyperchaotic CS-ELM (HCCS-ELM) model to predict both long-term and short-term PM$_{2.5}$ concentrations.

The main contributions of this research are as follows: (1) a new machine learning algorithm is proposed for measuring the daily and hourly concentrations of PM$_{2.5}$, which can predict pollutants levels for the coming year and make short-term predictions for the coming week; (2) the CS algorithm is optimized by a hyperchaotic system, and the improved CS algorithm and hyperchaotic series are applied to the original ELM. This research innovatively applies the hyperchaotic system to PM$_{2.5}$ concentration prediction; (3) the model improves the accuracy of pollutant concentration predictions, while effectively simulating the trend of pollutants over time.

The remainder of this article is organized as follows. Section 2 describes the relevant methods and models, which also contain the research area and data sources. Section 3 presents the results of the experiments comparing HCCS-ELM with other models. The experimental results are discussed in Sect. 4. Section 5 concludes the findings of this research.
2 Data and method

The ELM is used for PM$_{2.5}$ daily and hourly concentration prediction considering the close connection between pollutant levels and time series. In addition, the hyperchaotic system and the CS algorithm are used to improve the performance of the ELM. The study area and data sources are introduced in the following subsections. Then, the hyperchaotic system and the application of the hyperchaotic system to the CS algorithm and the ELM are described in detail. Finally, the improved CS algorithm and the hyperchaotic ELM are combined to form the HCCS-ELM shown in the last subsection.

2.1 Study Area

Lanzhou is the capital city of Gansu, located at 36°03′N and 103°40′E. The terrain is high in the southwest and low in the northeast. Surrounded by mountains in the north and south of the urban area, the Yellow River passes through the city, forming a narrow long valley between gorges and basins. The frequencies of calm winds and temperature inversions in the Lanzhou River Basin are high, which is not conducive to the diffusion of endogenous pollutants. The transfer of external dust into the surrounding area can lead to high concentrations of PM$_{2.5}$.

Based on the air quality records of Lanzhou from 1 January 2016 to 31 December 2021, the annual average PM$_{2.5}$ concentrations for Lanzhou from 2016 to 2021 were 54 μg/m$^3$, 45 μg/m$^3$, 38 μg/m$^3$, 37 μg/m$^3$, 35 μg/m$^3$, and 32 μg/m$^3$, respectively; the annual number of days with good or excellent air quality (air quality index less than 100) were 240, 237, 249, 298, 310 and 296, respectively. In Fig. 1, the red dashed line represents the Chinese PM$_{2.5}$ concentration limits, and the blue line represents Lanzhou’s PM$_{2.5}$ concentration, which has severely exceeded China’s standards in recent years.

2.2 Air quality monitoring and meteorological data

The meteorological data are the CMA dataset (available from http://data.cma.cn/), and the pollution data are the MEE dataset (available from http://www.mee.gov.cn/). In this research, the data were collected from four stations whose geographic locations are shown in Fig. 2. Station 1 is located in Xigu District, Lanzhou City, next to Lanzhou Petrochemical, and has relatively low traffic. Station 2 is located around the city of Lanzhou, with low traffic and no nearby factories. Stations 3 and 4 are located in the residential area of Lanzhou city, with heavy traffic and no nearby factories. The four stations do not affect the surrounding environment, which means they are hidden. For example, if people know where the stations are, they may go away, reducing traffic flow and resulting in inaccurate data. Daily and hourly meteorological data such as wind direction, wind speed, temperature, relative humidity, and pollutant (CO, SO$_2$, NO$_2$, O$_3$, and PM$_{2.5}$) data from 2016 to 2018 are selected as the study samples (Feng 2020; Barmpadimos et al. 2012; Lou et al. 2017).

2.3 Hyperchaotic system

In this section, a hyperchaotic real system and its corresponding complex system are used to generate the input weights and hidden layer biases of the HCCS-ELM to obtain better prediction results than comparable models. The hyperchaotic real system is given as follows (Li et al. 2019; Jia et al. 2020):

$$\begin{align*}
\dot{\xi}_1 &= -a\xi_1 + \xi_2\xi_3 + \xi_4 \\
\dot{\xi}_2 &= -\xi_1 + c\xi_2 \\
\dot{\xi}_3 &= \xi_1\xi_2 - b\xi_3 \\
\dot{\xi}_4 &= d\xi_1 + e\xi_2
\end{align*}$$

where $a$, $b$, $c$ and $d$ are real constants, and $\xi_1$, $\xi_2$, $\xi_3$ and $\xi_4$ are real variables. The dots denotes derivatives with respect to time $t$. The corresponding hyperchaotic complex system is given by

$$\begin{align*}
\dot{\eta}_1 &= -a\eta_1 + \eta_2\eta_3 + \eta_4 \\
\dot{\eta}_2 &= -\eta_1 + c\eta_2 \\
\dot{\eta}_3 &= \frac{1}{2}(\bar{\eta}_1\eta_2 + \eta_1\bar{\eta}_2) - b\eta_3 \\
\dot{\eta}_4 &= d\eta_1 + e\eta_2
\end{align*}$$

where $\eta_1 = u_1 + ju_2$, $\eta_2 = u_3 + ju_4$, $\eta_3 = u_5$, and $\eta_4 = u_6 + ju_7$ are complex variables, and $j = \sqrt{-1}$. The over bar represents the complex conjugate of the variables. The real version of System (2) is given as follows:

$$\begin{align*}
\dot{u}_1 &= -au_1 + u_2u_3 + u_6 \\
\dot{u}_2 &= -au_2 + u_4u_5 + u_7 \\
\dot{u}_3 &= -u_1 + cu_3 \\
\dot{u}_4 &= -u_2 + cu_4 \\
\dot{u}_5 &= u_1u_3 + u_2u_4 - bu_5 \\
\dot{u}_6 &= du_1 + eu_3 \\
\dot{u}_7 &= du_2 + eu_4
\end{align*}$$

The hyperchaotic sequences in System (1) are shown in Figure 3 for the $\xi_3$-$\xi_2$-$\xi_1$ and $\xi_2$-$\xi_1$-$\xi_4$ planes.
2.4 Hyperchaotic cuckoo search

The CS algorithm is a behavior simulation proposed by Yang and Deb (2009) based on interesting breeding behavior such as brood parasitism of certain species of cuckoos. Meanwhile, they introduced the breeding behavior of cuckoos and the characteristics of Lévy flights of some birds and fruit flies. In nature, cuckoo birds fly in random or similar random ways to find the location of the nest suitable for laying eggs. The CS algorithm has three ideal conditions as the premise (Yang and Deb 2009; Gandomi et al. 2011):

1. Each cuckoo lays only one egg at a time and randomly selects a nest location for hatching.
2. In a randomly selected set of nests, the high-quality nests will be retained for the next generation.
3. The number of available nests is fixed, and the probability of the bird’s nest owner finding the foreign egg is \( P_a \in [0, 1] \). When the owner finds a foreign cuckoo egg, he will throw it away or build a new nest.
Based on the above, the operation steps of the hyperchaotic cuckoo algorithm are as follows:

1. Setting parameters such as population size, searching space dimension, and maximum number of iterations. The position of the nest is initialized randomly as $X_i$. 

![Diagram](image-url)  

**Fig. 3** Phase diagrams for System (1) in the (a) $\xi_3$-$\xi_2$-$\xi_1$ plane, (b) $\xi_2$-$\xi_1$-$\xi_4$ plane
After position updating, random numbers \( r \) and \( P_a \) are used for comparison. If \( r > P_a \) is used, the nest position is randomly updated once. Otherwise, it remains unchanged.

If the maximum number of iterations or search accuracy requirements are met, the next step will proceed, or it will go back to Step (3).

Outputting the location of the globally optimal nest.

### 2.5 Hyperchaotic extreme learning machine

The hyperchaotic ELM is divided into three parts: input, hidden, and output layers. In the network, the neurons of the input, hidden and output layers are completely connected. Assume that the model has \( n \) input layer neurons and its corresponding input has \( n \) data variables, \( j \) neurons in the hidden layer, and \( m \) in the output layer so that its corresponding output has \( m \) data variables (Huang et al. 2006). \( W \) is the connection weight matrix of the input and hidden layers. It is represented using the following matrix:

\[
W = \begin{bmatrix}
    w_{11} & w_{12} & \cdots & w_{1n} \\
    w_{21} & w_{22} & \cdots & w_{2n} \\
    \vdots & \vdots & \ddots & \vdots \\
    w_{ni} & w_{n2} & \cdots & w_{nn}
\end{bmatrix}
\]  

(4)

where \( w_{ij} \) is the connection weight between the \( i^{th} \) input and the \( j^{th} \) hidden neurons.

\( \beta \) is the connection weight between the hidden and output layers as shown in Eq. 5:

\[
\beta = \begin{bmatrix}
    \beta_{11} & \beta_{12} & \cdots & \beta_{1m} \\
    \beta_{21} & \beta_{22} & \cdots & \beta_{2m} \\
    \vdots & \vdots & \ddots & \vdots \\
    \beta_{ni} & \beta_{n2} & \cdots & \beta_{nm}
\end{bmatrix}
\]  

(5)

where \( \beta_{ij} \) is the weight of the hidden neuron \( j \) to the output neuron \( k \). Both \( W \) and the status of neuron threshold in the hidden layer are generated using the above mentioned hyperchaotic system.

Hyperchaotic sequences generated using the above hyperchaotic functions are used to replace the hidden layer threshold matrix randomly generated using ELM. The running steps of the hyperchaotic ELM are as follows:

1. Importing data and starting training.
2. Generating the connection weight matrix between the input and hidden layers randomly.
3. Generating the threshold matrix of the hidden layer using the hyperchaotic function.
4. Calculating the connection weight matrix between the hidden and output layers using the least square method.
5. Importing the test data into the hyperchaotic ELM and outputting the prediction results after the training.

### 2.6 HCCS-ELM

Figure 4 shows the structure diagram of the HCCS-ELM. First, the training data are imported. Then, the chaos equation is used to generate the hyperchaotic sequence. Sequence 1 is selected to replace the Lévy flight in the cuckoo algorithm. It realizes the jump by randomly selecting the value in sequence without repetition. Sequence 2 is selected to replace the randomly generated connection weight matrix of the input and hidden layers. Sequence 3 is used to generate the hidden-layer threshold matrix in the ELM. The improved cuckoo algorithm is used to traverse the sample data to generate the number of hidden layer neurons in the ELM. After the training process of ELM, the test data are imported, and the prediction results are obtained.

To reduce the computational cost of the system, the hyperchaotic system is used to generate a sequence instead of the Lévy flight. The hyperchaotic system is only required to generate a sequence once, and then each iteration randomly selects a value from the sequence, which greatly reduces the resource occupation. In addition, the CS algorithm is exploited to traverse the samples to generate the number of hidden layer neurons, which effectively improves the adaptability of the algorithm and avoids repeated parameter adjustment. Considering the convergence phenomenon of chaos, the hyperchaotic sequence is used to replace the hidden layer threshold matrix in ELM, and regression can be understood as a convergence phenomenon. Therefore, the use of convergent sequences instead of random sequences improves the accuracy of ELM, yields highly accurate results in \( PM_{2.5} \) concentration prediction, and saves the use of system resources.
3 Experiments and results

In this section, the HSSC-ELM is used to predict the four stations in terms of the daily concentrations of PM$_{2.5}$ and compares with other models for forecasting performance in one year and each quarter. Then, the HSSC-ELM is used for hourly PM$_{2.5}$ concentration prediction to obtain the prediction results of the model. Finally, the HSSC-ELM is analyzed for daily pollutant level forecasting in terms of time and memory consumption compared with the ELM.

3.1 Long-term prediction

In this research, HCCS-ELM, BPNN, MLR, ELM, long short-term memory (LSTM), and Prophet are used to predict the PM$_{2.5}$ concentration at four observation stations. In addition, the results are compared and analyzed similarly to Amanollahi and Ausati (2019). Seventy percent of the historical observation data are used to train the model, and the remaining 30% of the data are used to test the model. In this research, PM$_{2.5}$ levels are predicted for different seasons, since the PM$_{2.5}$ concentration varies across seasons. For instrument failure and other irresistible making...
historical observations incompletable, invalid data are deleted if one or more kinds of variables are missing. However, special days, such as rainy days, holidays, or nights, are not criteria for filtering data. Thus, only two months in each season are selected for prediction. Table 1 presents the amount of filtered data. There are 500 samples in the training set of PM$_{2.5}$ prediction at Station 1 and 215 samples in the test set. There are 511 samples in the training set of PM$_{2.5}$ predictions of Stations 2, 3, and 4 and 219 samples in the test set. The input parameters include eight variables: temperature, humidity, sulfur dioxide, nitrogen dioxide, ozone, carbon monoxide, instantaneous wind speed, and wind direction. The predicted PM$_{2.5}$ concentration is output at the output layer. Both the HCCS-ELM and ELM have the same parameters, which are the sigmoid function, “regression” model, and 10 hidden layer neurons. However, the BPNN has 10 hidden layer neurons, 100 epochs, a 0.1 learning rate, and 0.4 goals. All parameters of the six models are kept constant during the experiment. Root mean squared error (RMSE) and the coefficient of determination ($R^2$) are selected as the evaluation criteria. The formulas of RMSE and $R^2$ are given as follows:

\[
RMSE = \sqrt{\frac{\sum_{i=1}^{n} (X_{\text{obs},i} - X_{\text{model},i})^2}{n}}
\]

\[
R^2 = 1 - \frac{\text{RMSE}}{\text{Var} (\cdot)}
\]

where $n$ is the number of samples, $X_{\text{obs},i}$ is the observed value of PM$_{2.5}$ concentration, $X_{\text{model},i}$ is the predicted value of PM$_{2.5}$ concentration, and $\text{Var}(\cdot)$ is the variance. The closer $R^2$ is to 1 and the smaller RMSE is, the better the prediction effect is.

MLR, LSTM, and Prophet are run with Python. Figure 5 shows the results predicted by HCCS-ELM. Tables 2 and 3 present the RMSE and $R^2$ of HCCS-ELM, ELM, BPNN, MLR, LSTM and Prophet, respectively. The analysis of the above chart shows that the RMSE of HCCS-ELM is smaller than the other five models, and the $R^2$ is larger than the other models, which could predict the concentration and the trend of PM$_{2.5}$ more accurately. Table 2 shows the results obtained by the daily concentration from 2016 to 2017 as the training set and the 2018 daily concentration as the test set. Table 3 uses the first two months of each quarter in 2018 as the training set and the last month as the test set.

### 3.2 Short-term prediction

The short-term extremely high PM$_{2.5}$ concentration is usually the object of the attention of the meteorological department. High PM$_{2.5}$ concentrations are harmful to human health. In this research, data with PM$_{2.5}$ concentrations exceeding China’s national standard are selected, and the data division is consistent with the above. Figure 6 shows that the HCSS-ELM is still efficient for predicting short-term extremely high PM$_{2.5}$ concentrations. Due to missing data, the hourly concentration data are incomplete. In Fig. 6, the test data are the hourly concentration of the date in the figure, and the training data are the effective value of the hourly concentration fourteen days before the test date.

### 3.3 Performance evaluation

The running time and memory usage are used to evaluate the performance of the HCCS-ELM and ELM. Tables 4 and 5 present the time and memory results. Table 4 shows the performance of the two algorithms in processing the daily concentration data of each station from 2016 to 2018, and Table 5 shows the performance of the two algorithms in processing the daily concentration data of four stations in each quarter. This shows that the HCCS-ELM needs more running time than the ELM, while they have similar occupied memory. That is because the HCCS-ELM needs to generate hyperchaotic sequences during operation. The differences in running time and memory usage of the same model can be attributed to the software.

### 4 Discussion

From the perspective of environmental protection, people should improve the environmental protection system, adopt an environmentally friendly lifestyle, and insist on the harmonious coexistence between humans and nature. For the improvement of the global ecological environment, it is necessary to set up and practice the concept that clear waters and green mountains are as valuable as mountains of gold and silver.
Machine learning has fast and accurate prediction performance and is commonly used to predict the trend of pollutant concentration transformation. Accurate prediction of pollutant concentrations can help local governments take more effective preventive measures. Therefore, the HCCS-ELM is obtained using two steps based on the ELM: (1) the chaotic sequences that are generated by the hyperchaotic system, are randomly chosen instead of Lévy flights for iteration to reduce the computational effort of the CS algorithm; (2) the improved CS algorithm is used for the number of hidden layer neurons of ELM, which can avoid manual adjustment and improve the accuracy of the ELM.

![Daily PM2.5 concentration prediction results of HSSC-ELM in a year](image)

**Fig. 5** Daily PM$_{2.5}$ concentration prediction results of HSSC-ELM in a year

**Table 2** Comparison of RMSE and $R^2$ of different methods (Year)

|                  | Station 1 | Station 2 | Station 3 | Station 4 |
|------------------|-----------|-----------|-----------|-----------|
|                  | RMSE      |           |           |           |
| HCCS-ELM         | 4.6255    | 5.7935    | 4.6431    | 4.7172    |
| ELM              | 9.3645    | 11.8354   | 8.9009    | 8.4336    |
| BPNN             | 15.7328   | 17.8923   | 24.6076   | 20.3576   |
| MLR              | 17.7866   | 21.3147   | 15.7897   | 19.1319   |
| LSTM             | 16.5247   | 18.5391   | 13.6778   | 16.2429   |
| Prophet          | 16.5238   | 21.1809   | 15.5725   | 17.6423   |
|                  |           |           |           |           |
|                  | $R^2$     |           |           |           |
| HCCS-ELM         | 0.9627    | 0.9686    | 0.9470    | 0.9625    |
| ELM              | 0.8842    | 0.8497    | 0.8244    | 0.8931    |
| BPNN             | 0.6079    | 0.6627    | 0.6607    | 0.6853    |
| MLR              | 0.4131    | 0.4296    | 0.3524    | 0.3979    |
| LSTM             | 0.4858    | 0.5626    | 0.5006    | 0.5423    |
| Prophet          | 0.4929    | 0.4366    | 0.3601    | 0.4588    |
**Table 3** Comparison of RMSE and $R^2$ of different methods (Season)

| Season | Station 1 | Station 2 | Station 3 | Station 4 |
|--------|-----------|-----------|-----------|-----------|
| RMSE   |           |           |           |           |
| HCCS-ELM | 5.0233 | 9.1717 | 19.1538 | 20.9775 | 16.7506 | 18.1760 | 0.9371 | 0.8384 | 0.6513 | 0.1443 | 0.2827 | 0.1386 |
| ELM    | 7.1416 | 9.7955 | 13.7006 | 26.1891 | 17.3320 | 19.0321 | 0.8987 | 0.8513 | 0.7429 | 0.0236 | 0.2708 | 0.1075 |
| BPNN   | 4.8586 | 15.2012 | 10.1559 | 18.5039 | 14.5123 | 15.4039 | 0.9236 | 0.6557 | 0.8090 | 0.0760 | 0.2200 | 0.1156 |
| MLR    | 5.2844 | 14.7320 | 8.8194 | 22.7076 | 16.6815 | 17.3568 | 0.9278 | 0.7330 | 0.8115 | 0.0135 | 0.2602 | 0.1880 |
| LSTM   | 3.4992 | 5.9359 | 5.1225 | 10.7228 | 10.6803 | 9.9579 | 0.9111 | 0.7087 | 0.8219 | 0.0551 | 0.0420 | 0.1275 |
| Prophet | 3.6244 | 7.8195 | 7.2272 | 14.7589 | 11.8178 | 9.9890 | 0.9165 | 0.7533 | 0.7443 | 0.0161 | 0.0637 | 0.1125 |
| Summer |           |           |           |           |           |           |           |           |           |           |           |           |
| RMSE   |           |           |           |           |           |           |           |           |           |           |           |           |
| HCCS-ELM | 6.5007 | 11.3300 | 13.9920 | 13.7498 | 17.1170 | 17.4268 | 0.9106 | 0.7194 | 0.6930 | 0.5252 | 0.1574 | 0.1389 |
| ELM    | 8.1300 | 13.2583 | 19.2205 | 29.0382 | 23.7319 | 24.5825 | 0.9543 | 0.8987 | 0.8456 | 0.4161 | 0.5688 | 0.5220 |
| BPNN   | 7.9112 | 13.8340 | 14.1030 | 21.6965 | 14.2278 | 15.3461 | 0.9164 | 0.7875 | 0.8581 | 0.3124 | 0.6511 | 0.5855 |
| MLR    | 6.1445 | 14.5434 | 15.9177 | 21.5627 | 17.2101 | 18.1260 | 0.9555 | 0.7517 | 0.7966 | 0.4725 | 0.6487 | 0.6050 |
| LSTM   | 2.2604 | 4.5818 | 4.4368 | 11.1462 | 9.5097 | 8.7015 | 0.9481 | 0.7903 | 0.8476 | 0.032 | 0.0127 | 0.1496 |
| Prophet | 2.2680 | 5.7048 | 4.3199 | 10.1234 | 8.9888 | 8.1486 | 0.9360 | 0.7222 | 0.7889 | 0.0081 | 0.0007 | 0.1425 |
| Autumn |           |           |           |           |           |           |           |           |           |           |           |           |
| RMSE   |           |           |           |           |           |           |           |           |           |           |           |           |
| HCCS-ELM | 7.2779 | 12.7341 | 10.5278 | 17.0671 | 20.9554 | 20.6743 | 0.9457 | 0.7307 | 0.8011 | 0.5285 | 0.1664 | 0.1717 |
| ELM    | 8.4257 | 20.8565 | 23.9630 | 19.5162 | 24.5324 | 22.5306 | 0.9271 | 0.7488 | 0.8243 | 0.5172 | 0.0012 | 0.1391 |
| BPNN   | 3.8454 | 12.7414 | 16.3453 | 17.1244 | 18.5874 | 17.4310 | 0.9590 | 0.7347 | 0.7582 | 0.2818 | 0.0216 | 0.1277 |
| MLR    |           |           |           |           |           |           |           |           |           |           |           |           |
| LSTM   | 5.9452 | 13.5708 | 13.3862 | 22.1096 | 22.2022 | 20.4694 | 0.9344 | 0.7172 | 0.8402 | 0.3632 | 0.0277 | 0.1453 |
| Prophet |           |           |           |           |           |           |           |           |           |           |           |           |

Fig. 6 Hourly PM$_{2.5}$ concentration prediction results of HSSC-ELM in a week
The proposed model has the following advantages in analyzing PM\textsubscript{2.5} concentrations in Lanzhou. First, HCCS-ELM simplifies the analysis process, while traditional numerical weather forecast methods are very complex and require many calculations. It also shows a reduction in computational complexity when predicting PM\textsubscript{2.5} compared to deterministic numerical prediction methods, for example, by modeling the relationship between satellite aerosol optical depth and ground-based PM\textsubscript{2.5} observations. In the HCCS-ELM, the CS algorithm and hyperchaotic system need some time to operate. As a result, compared to ELM, the running time increases, but the accuracy is improved and the prediction error is reduced. The HCCS-ELM is similar to other neural networks in that it is independent of the complex relationship between the parameters and outputs. However, it depends on constant changes in weights so that the parameters and outputs are closely related and tedious mathematical modeling is avoided. Second, the HCCS-ELM has excellent generalization. Since there is no mathematical modeling process, combined with similar and available data, HCCS-ELM can theoretically be applied to many cities in China that set up air quality monitoring stations without being affected by differences in the geographical environment and the level of economic development. Finally, considering the data processing characteristics of HCCS-ELM, the concentration of other pollutants can be predicted by simply transforming the columns in the data matrix, avoiding the complex adjustment process in the meteorological equations. Furthermore, the addition of a hyperchaotic system provides new ideas for improving the performance of neural networks. There are some limitations to this research. The model considers the temporal variation in PM\textsubscript{2.5}, but the spatial distribution is ignored. Furthermore, the model needs to be improved in terms of operational cost and efficiency.

### 5 Conclusion

This research proposes a novel model, HCCS-ELM, which combines the CS algorithm and hyperchaotic system with ELM. The ELM is conducted at each of the four stations to predict daily and hourly PM\textsubscript{2.5} concentrations based on monitored levels of other pollutants. The CS algorithm can calculate the optimal number of hidden layer neurons of ELM without manual adjustment, while the hyperchaotic system reduces the computational cost of the CS algorithm and improves the accuracy of ELM due to its adequate convergence and randomness. The prediction results show that HCCS-ELM achieves better performance than the original ELM, MLR, BP, LSTM, and Prophet in predicting daily concentrations of pollutants. Furthermore, the PM\textsubscript{2.5} model predictions of hourly PM\textsubscript{2.5} concentrations during a week are highly consistent with the monitored data. With its generalizability and high prediction accuracy, the HCCS-ELM model is of great significance for air pollution forecasting. Future work will focus on adding the spatial characteristics of pollutants to the HCCS-ELM to obtain more detailed information on the distribution of PM\textsubscript{2.5}.

#### Funding

This work was supported in part by the Fundamental Research Funds for the Central Universities (Grant No.lzujkby-2019) and The Natural Science Foundation of Gansu Province (20JR10RA606).

#### Data source

Meteorological data: http://data.cma.cn/. Pollutant data: http://www.mee.gov.cn/
Declarations

Conflict of 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.

Experiment platform CPU Intel i7-8700K, GPU Nvidia RTX2060, RAM 16GB.

References

Amanollahi J, Ausati S (2019) PM$_{2.5}$ concentration forecasting using ANFIS, EEMD-GRNN, MLP, and MLR models: a case study of Tehran, Iran. Air Qual Atmos Health 13(2):161–171. https://doi.org/10.1007/s11869-019-00779-5

Barmepadimos I, Keller J, Oderbolz D, Hueglin C, Prévôt ASH (2012) One decade of parallel fine (PM$_{2.5}$) and coarse (PM$_{10}$–PM$_{2.5}$) particulate matter measurements in Europe: trends and variability. Atmos Chem Phys 12(7):3189–3203. https://doi.org/10.5194/acp-12-3189-2012

Bianconi F, Busilacchio M, Verdecchia M, Tomassetti B, Arruffo E, Bianco S, Di Tommaso S, Colangelo C, Rossetti G, Di Carlo P (2017) Recursive neural network model for analysis and forecast of PM$_{10}$ and PM$_{2.5}$. Atmos Pollut Res 8(4):652–659. https://doi.org/10.1016/j.apr.2016.12.014

Coleman NC, Burnett RT, Ezzati M, Marshall JD, Robinson AL, Pope CA (2020) Fine particulate matter exposure and cancer incidence: Analysis of seer cancer registry data from 1992–2016. Environ Health Perspect 128(10):107004. https://doi.org/10.1289/ehp7246

Feng R (2020) Investigating wintertime air pollution in Hangzhou, China. Air Qual Atmos Health 13(3):321–328. https://doi.org/10.1007/s11869-020-00794-x

Fu M, Wang W, Le Z, Khorrman MS, Coleman NC, Burnett RT, Ezzati M, Marshall JD, Robinson AL, Pope CA (2020) Fine particulate matter concentrations by developed feed-forward neural network with rolling mechanism and gray model. Neural Comput Appl 26(1789):8. https://doi.org/10.1007/s00521-015-1853-8

Gandomi AH, Yang X-S, Alavi AH (2011) Cuckoo search algorithm: a metaheuristic approach to solve structural optimization problems. Eng Comput 29(1):17–35. https://doi.org/10.1007/s00366-014-0241-y

Gilani O, McKay L, Gregoire T, Guan Y, Leaderer B, Holford T (2016) Spatiotemporal calibration and resolution refinement of output from deterministic models. Stat Med 35(24):4224–4242. https://doi.org/10.1002/sim.6867

Huang G-B, Zhu Q-Y, Siew C-K (2006) Extreme learning machine: theory and applications. Neurocomput 70(1–3):489–501. https://doi.org/10.1016/j.neucom.2005.12.024

Jin X-B, Yang N-X, Wang X-Y, Bai Y-T, Su T-L, Kong J-L (2020) Deep hybrid model based on EMD with classification by frequency characteristics for long-term air quality prediction. Mathematics 8(2):214. https://doi.org/10.3390/math8020214

Kow P-Y, Wang Y-S, Zhou Y, Kao I-F, Issermann M, Chang L-C, Chang F-J (2020) Seamless integration of convolutional and back-propagation neural networks for regional multi-step-ahead PM$_{2.5}$ forecasting. J Clean Prod 261:121285. https://doi.org/10.1016/j.jclepro.2020.121285

Lesar TT, Filipic A (2021) The hourly simulation of PM$_{2.5}$ particle concentrations using the multiple linear regression (MLR) model for sea breeze in split, croatia. Water, Air, Soil Pollut 232(7):1–9. https://doi.org/10.1007/s11270-021-05209-w

Li P, Du J, Li S, Zheng Y (2019) Modulus synchronization of a novel hyperchaotic real system and its corresponding complex system. IEEE Access 7:109577–109584. https://doi.org/10.1109/access.2019.2943120

Li P, Du J, Li S, Zheng Y, Jia B (2019) The synchronization of n cascade-coupled chaotic systems. Complexity 2019:1–10. https://doi.org/10.1155/2019/2709820

Li K, Ma J, Li H, Han Y, Yue X, Chen Z, Yang J (2020) Discern depth under foul weather: estimate PM$_{2.5}$ for depth inference. IEEE Trans Ind Inform 16(6):3918–3927. https://doi.org/10.1109/tii.2019.2943631

Lightstone S, Gross B, Moshary F, Castillo P (2021) Development and assessment of spatially continuous predictive algorithms for fine particulate matter in New York State. Atmosphere 12(3). https://doi.org/10.3390/atmos12030315

Lou C, Liu H, Li Y, Peng Y, Wang J, Dai L (2017) Relationships of relative humidity with PM$_{2.5}$ and PM$_{10}$ in the Yangtze River Delta. China. Environmental Monitoring and Assessment 189(11). https://doi.org/10.1007/s10661-017-6281-z

Park Y, Kwon B, Heo J, Xu L, Liu Y, Moon T (2020) Estimating PM$_{2.5}$ concentration of the conterminous united states via interpretable convolutional neural networks. Environ Pollut 256:113395. https://doi.org/10.1016/j.envpol.2019.113395

Qiao J, He Z, Du S (2020) Prediction of PM$_{2.5}$ concentration based on weighted bagging and image contrast-sensitive features. Stoch Environ Res Assess 34:561–573. https://doi.org/10.1007/s00477-020-01787-7

Schaufnagel DE, Balms Jr, Cowl CT, Mattes SD, Jung S-H, Mortimer K, Perez-Padilla R, Rice MB, Riojas-Rodrigue H, Sood A, Thurston GD, To T, Vanker A, Wuebbles DJ (2019) Air pollution and noncommunicable diseases. Chest 155(2):409–416. https://doi.org/10.1016/j.chest.2018.10.042

Tong X, Ho JM, Li Z, Lai K-H, Kwok TCY, Tsoi KKF, Ho SF (2020) Prediction model for air particulate matter levels in the households of elderly individuals in Hong Kong. Sci Total Environ 717:135323. https://doi.org/10.1016/j.scitotenv.2019.135323

Wang Z, Chen L, Ding Z, Chen H (2020) An enhanced interval PM$_{2.5}$ concentration forecasting model based on BEMD and MLP with influencing factors. Atmos Environ 223:117200. https://doi.org/10.1016/j.atmosenv.2020.117200

Wang Y, Wang J, Li Z (2020) A novel hybrid air quality early-warning system based on phase-space reconstruction and multi-objective optimization: A case study in China. J Clean Prod 261:121027. https://doi.org/10.1016/j.jclepro.2020.121027

Wang X, Yuan J, Wang B (2021) Prediction and analysis of PM$_{2.5}$ in fuling district of chongqing by artificial neural network. Neural Comput Appl 33(2):517–524. https://doi.org/10.1007/s00521-020-04962-z

Xiong J, Yao R, Wang W, Yu W, Li B (2020) A spatial-and-temporal-based method for rapid particle concentration estimations in an urban environment. J Clean Prod 256:120331. https://doi.org/10.1016/j.jclepro.2020.120331

Xu Y, Huang Y, Guo Z (2021) Influence of AOD remotely sensed products, meteorological parameters, and AOD–PM$_{2.5}$ models on the PM$_{2.5}$ estimation. Stoch Environ Res Assess 35:893–908. https://doi.org/10.1007/s00477-020-01941-7

Xu Y, Liu H (2020) Spatial ensemble prediction of hourly PM$_{2.5}$ concentrations around Beijing railway station in China. Air Qual Atmos Health 13(5):563–573. https://doi.org/10.1007/s11869-020-00817-7

Yang X-S, Deb S (2009) Cuckoo search via lévy flights. In: 2009 World Congress on Nature & Biologically Inspired Computing
Ye D, Klein M, Mulholland JA, Russell AG, Weber R, Edgerton ES, Chang HH, Sarnat JA, Tolbert PE, Ebelt Sarnat S (2018) Estimating acute cardiovascular effects of ambient PM$_{2.5}$ metals. Environ Health Perspect 126(2):027007. https://doi.org/10.1289/EHP2182

Zhao Z, Qin J, He Z, Li H, Yang Y, Zhang R (2020) Combining forward with recurrent neural networks for hourly air quality prediction in Northwest of China. Environ Sci Pollut Res 27(23):28931–28948. https://doi.org/10.1007/s11356-020-08948-1

Zhou Y, Chang L-C, Chang F-J (2020) Explore a multivariate bayesian uncertainty processor driven by artificial neural networks for probabilistic PM$_{2.5}$ forecasting. Sci Total Environ 711:134792. https://doi.org/10.1016/j.scitotenv.2019.134792

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

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.