

Haze Prediction Model Using Deep Recurrent Neural Network

Kailin Shang¹, Ziyi Chen¹, Zhixin Liu^{1,*}, Lihong Song², Wenfeng Zheng^{2,*}, Bo Yang², Shan Liu²
and Lirong Yin^{3,*}

¹ School of Life Science, Shaoxing University, Shaoxing 312000, China; KailinShang1@outlook.com (K.S.); ziyi.chen.cn@gmail.com (Z.C.)

² School of Automation, University of Electronic Science and Technology of China, Chengdu 610054, China; lihong.song.cn@outlook.com (L.S.); boyang@uestc.edu.cn (B.Y.); shanliu@uestc.edu.cn (S.L.)

³ Department of Geography and Anthropology, Louisiana State University, Baton Rouge, LA 70803, USA

* Correspondence: liuzhixin@usx.edu.cn (Z.L.); winfirms@uestc.edu.cn (W.Z.); lyin5@lsu.edu (L.Y.)

Abstract: In recent years, haze pollution is frequent, which seriously affects daily life and production process. The main factors to measure the degree of smoke pollution are the concentrations of PM_{2.5} and PM₁₀. Therefore, it is of great significance to study the prediction of PM_{2.5}/PM₁₀ concentration. Since PM_{2.5} and PM₁₀ concentration data are time series, their time characteristics should be considered in their prediction. However, the traditional neural network is limited by its own structure and has some weakness in processing time related data. Recurrent neural network is a kind of network specially used for sequence data modeling, that is, the current output of the sequence is correlated with the historical output. In this paper, a haze prediction model is established based on a deep recurrent neural network. We obtained air pollution data in Chengdu from the China Air Quality Online Monitoring and Analysis Platform, and conducted experiments based on these data. The results show that the new method can predict smog more effectively and accurately, and can be used for social and economic purposes.



Citation: Shang, K.; Chen, Z.; Liu, Z.; Song, L.; Zheng, W.; Yang, B.; Liu, S.; Yin, L. Haze Prediction Model Using Deep Recurrent Neural Network. *Atmosphere* **2021**, *12*, 1625. <https://doi.org/10.3390/atmos12121625>

Academic Editor: Ying Chen

Received: 9 November 2021

Accepted: 29 November 2021

Published: 6 December 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: haze prediction; PM_{2.5}/PM₁₀; deep recurrent neural network

1. Introduction

Haze is one of the common meteorological disasters in China, which has caused an extensive impact on all aspects of people's productivity and life. In terms of physical health, it affects people's respiratory systems, greatly increasing the probability of respiratory diseases, and hinders cardiovascular health. In haze weather, the inhalable particles in the air increase poor air mobility, resulting in bacteria, the slow spread of the virus, the concentration of the virus in the space range increases, making people more vulnerable to diseases. In terms of traffic, haze leads to low air visibility, easily causing traffic accidents.

Due to the many adverse effects of haze, many scientists have carried out relevant studies on haze. Lv, Z. et al. [1] analyzed the causes of the haze, believing that it is related to the exhaust emissions of vehicles and other vehicles. Wu, J. et al. [2] analyzed the causes of haze from multiple angles. Liu, Y. et al. [3] studied the problem of industrial development under the influence of haze, and pointed out the impact of haze on human production and life, especially the air control industry. Thus, the emergence of haze has produced a serious interference with our daily life [4–6].

Prediction of haze thus becomes extremely important to prevent the occurrence of severe haze. Baklanov, A. and Zhang, Y. [7] made a systematic introduction and summary of the development of modern atmospheric composition modeling and AQF (Air quality forecast) system in 2020. Gao et al. [8] evaluated the impacts of Global Positioning System Zenith Total Delay (GPS-ZTD) data assimilation on meteorology and aerosol simulations in this study using the WRF-CMAQ (the Weather Research and Forecasting model and Community Multiscale Air Quality) modeling system over the NCP during

1–31 December 2019. Hertwig et al. [9,10] developed the Lagrange dispersion model system in 2015 and applied it to haze prediction. The experiment proved that the system could provide a reliable prediction. Liu et al. [11,12] analyzed the fuchsia haze events in Nanjing in 2017. Singh, V. et al. [13] proposed a method based on co-kriging to predict air pollution. The method has been officially applied to the urban area of Milan.

In recent years, deep learning has been well applied in many fields and has gradually replaced traditional methods [14–17]. Compared with traditional technology, artificial neural networks [18,19] have been widely used in the field of haze prediction. Yin et al. [20] established a multi-convolution neural network haze prediction model to forecast haze in 2021. Avijoy Chakma et al. proposed a method that uses a deep Convolutional Neural Network (CNN) to classify natural images into different categories based on their PM2.5 concentrations in 2017 [21]. Compared with the two-dimensional data of image processing [22], haze is more suitable to be regarded as a kind of sequence data. Ziyang Zhang et al. proposed a haze prediction method based on 1D-CNN in 2021 [23]. As BP neural network has the characteristics of predicting nonlinear complex systems, BP neural network is also used in the field of haze prediction. Limei Ma et al. proposed a haze prediction method based on BP neural network in 2017 [24]. As the haze data is a time series, the above methods cannot well analyze the time relationship in the data. Therefore, RNN (Recurrent Neural Network) specially used for sequence analysis has been applied in the field of haze prediction [25,26].

When solving practical problems, the RNN model of a hidden layer has limited representation ability, and the results of haze prediction under multiple influencing factors are not excellent. In order to solve the above problems, a haze prediction system based on deep cyclic neural network is designed, the time-based back-propagation algorithm is selected to train the prediction model, and the PM2.5 and PM10 prediction experiments are completed. The results showed that the model has high feasibility and rationality and obtained accurate and reliable results, providing a valuable reference for improving air monitoring of relevant departments.

2. Materials and Method

In this study, we collected the air pollution data of 6 monitoring stations in Chengdu from the website of Chengdu air quality online monitoring and analysis platform <https://www.aqistudy.cn> (accessed on 17 March 2018).

The six monitoring stations are in Shahepu, Liangjia lane, Junpingjie, Caotangsi, Sanwayao and Shilidian in Chengdu. The data provided by the monitoring stations are the concentrations of PM2.5, PM10, O₃, CO, NO₂ and SO₂, of which the measurement unit of PM2.5, PM10, O₃ and SO₂ is (µg/m³), while NO₂ and CO are measured in (mg/m³). The monitoring data update frequency is once an hour. The geographical locations of the above six monitoring points in the urban area of Chengdu are shown in Figure 1.

By compiling crawler code using Python 2.70 by Python Software Foundation, this study collected the statistical data of the above monitoring points from 1 June 2014 to 30 June 2017. The average data of 6 monitoring points are used to represent the air quality in Chengdu, with a total of 26,120 data. The data format is shown in Table 1.

2.1. Mean Completion Data

As shown in Table 1, the original data is missing at a certain time. We use the mean interpolation method to complete some missing data. The specific method is shown in Equation (1):

$$X_t = \frac{1}{2}(X_{t-1} + X_{t+1}) \quad (1)$$

X_t represents the missing weather data at the current time, and X_{t-1} represents the weather data at the previous moment, X_{t+1} represents the weather data at a time after.

After interpolation, a total of 27,024 valid data were used in this study.

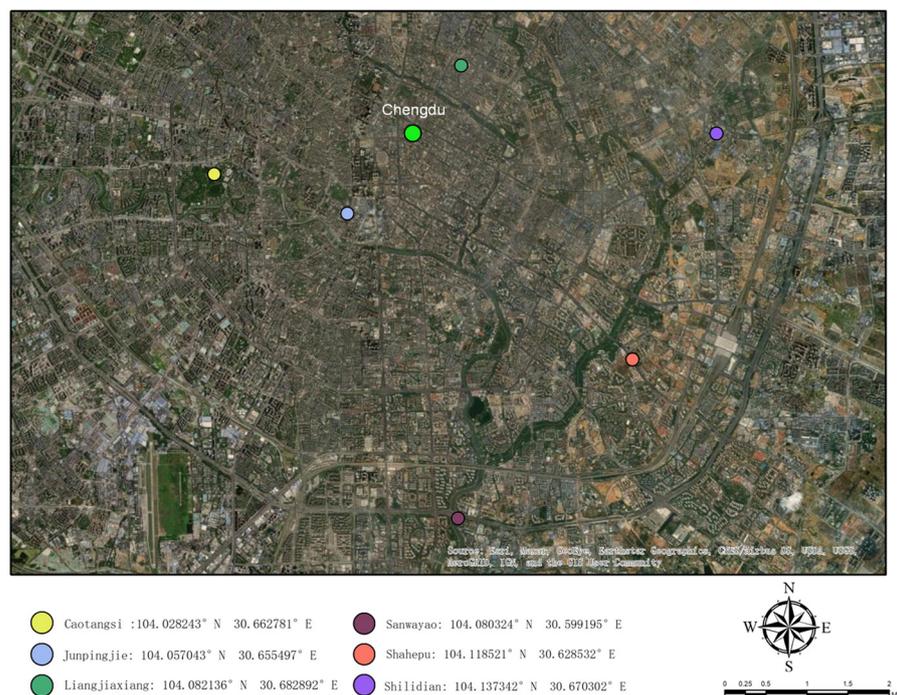


Figure 1. Geographical location of monitoring point.

Table 1. Data format.

Time	PM2.5 (µg/m ³)	PM10 (µg/m ³)	CO (mg/m ³)	NO ₂ (mg/m ³)	O ₃ (µg/m ³)	SO ₂ (µg/m ³)
1 June 2014 0:00	89	130	0.8	36	129	19
1 June 2014 1:00	86	120	0.9	54	91	16
1 June 2014 2:00	0	0	0	0	0	0
1 June 2014 3:00	71	111	0.8	35	97	17

2.2. Standardized Data Processing

It can be seen from Table 1 that there is a large difference in the numerical range of the six data collected. The data with larger values will increase the influence proportion of the model in the neural network and weaken the data characteristics with lower values. In order to avoid the error caused by different numerical ranges, the above data are normalized in this study. All data are normalized to between −1 and 1, as shown in Equation (2):

$$X' = \frac{X - \bar{X}}{X_{max} - X_{min}} \tag{2}$$

Among them, X' indicates the weather data set after the standardization process is completed. X represents the original weather data set, and \bar{X} indicates the mean of the weather data set. X_{max} indicates the maximum value of the weather dataset, X_{min} represents the minimum value of the weather dataset.

In the subsequent experiments, we will use the first 80% of the entire data set as the test set, the next 10% as the verification set, and the last 10% as the test set. Divided by time.

2.3. Methods

Recursive neural networks, simply RNN, are a class of neural networks used to process sequence data [27]. Sequence data refers to a kind of data that can reflect the degree of a phenomenon or something, and there is a close connection between before and after data. RNN includes input layer, hidden layer, output layer, hidden layer plays the most important role, hidden layer not only includes the output of the input layer at this time,

but also the output of its previous moment, RNN processing of sequence data is mainly reflected in the memory, data information characteristics, applied to the output, because RNN depth reflects the input and output process, it is also called depth network, the following is the model development diagram of RNN (Figure 2).

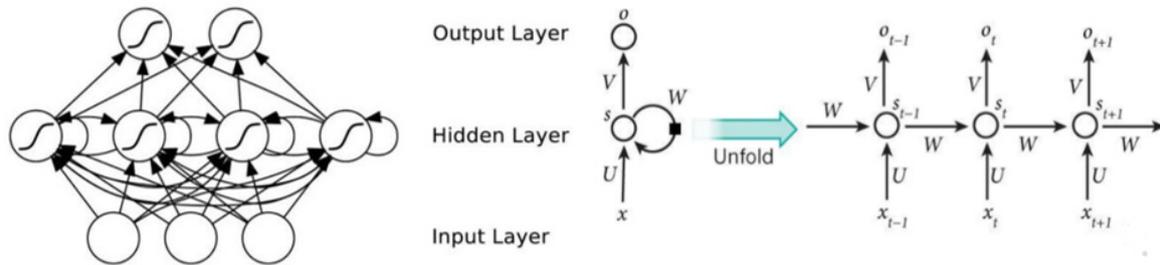


Figure 2. Model expansion diagram of recurrent neural network.

Unlike general neural networks that make connections only between layers, the RNN also connects between neurons between layers. We can use the hidden layer as the storage space for the entire network, and when the RNN is unfolded, we find that it can be used to supervise classification learning recurrent neural networks, introduce a directional neural network that remembers previous information and applies, which essentially distinguishes it from traditional feedforward neural networks. For traditional neural networks, it can train samples more efficiently, simulating complex feature relationships more realistically and credibly. RNN uses the reverse propagation method to calculate the forward propagation process based on the time and order of network propagation [28].

However, in the practical problem solving, RNN is still insufficient, expresses limited ability, and the hidden layer cannot express the whole model, so in this paper, in predicting the concentration scenario of PM2.5 and PM10, a deep recurrent neural network is designed to better complete the prediction of PM2.5 and PM10 concentration. To study three aspects of RNN deeper than other methods: input to hidden function; hidden to output function; and hidden to hidden transformation [29]. Since the depth of the deep recursive network model is mainly reflected in the hidden layer, we start from the hidden layer to determine the optimal model by changing the number and the units of the hidden layers [30] (Figure 3).

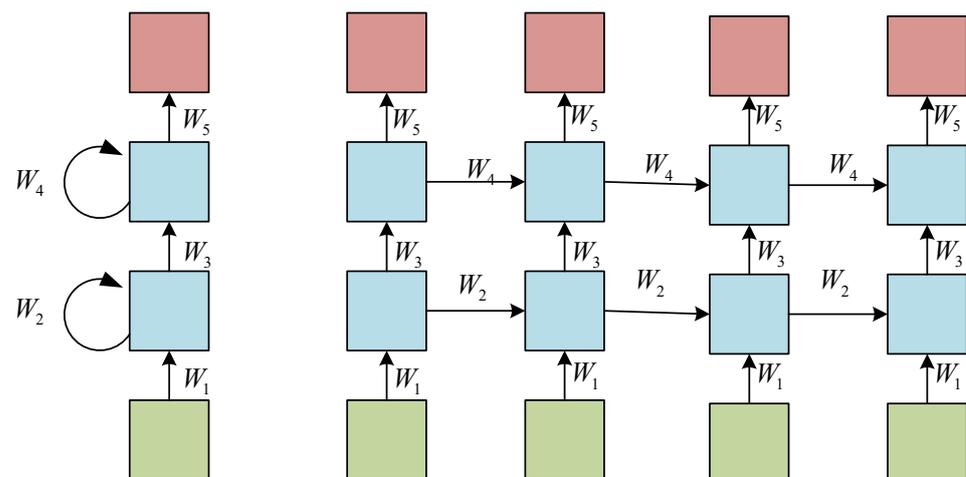


Figure 3. PM2.5 prediction model network structure of the deep recurrent neural network.

2.4. Deep Recurrent Neural Network Training Process

Recursive neural networks can encode long past information with the function of temporal storage. During the training of deep recurrent neural networks, the temporal backpropagation algorithm is trained to apply it to the recurrent layer. For the errors

generated by the backpropagation algorithm, we propagate in two directions for the former layer transmitted to the current layer and the previous time it transmits to:

- (1) Define annotations (Table 2)

Table 2. Annotation definition table.

Nerve Layer	Describe	Index Variable
$x(t)$	The input layer	i
$s(t - 1)$	Previous time hidden layer state	h
$s(t)$	Hide layer state at the current time	j
$y(t)$	Output layer	k
The weight matrix	Describes	Index Variables
V	Input layer to output layer	i, j
U	Previous time hidden layer to hidden layer	h, j
W	Hidden layer to output layer	j, k

According to Figure 4a, training the network requires the following processing.

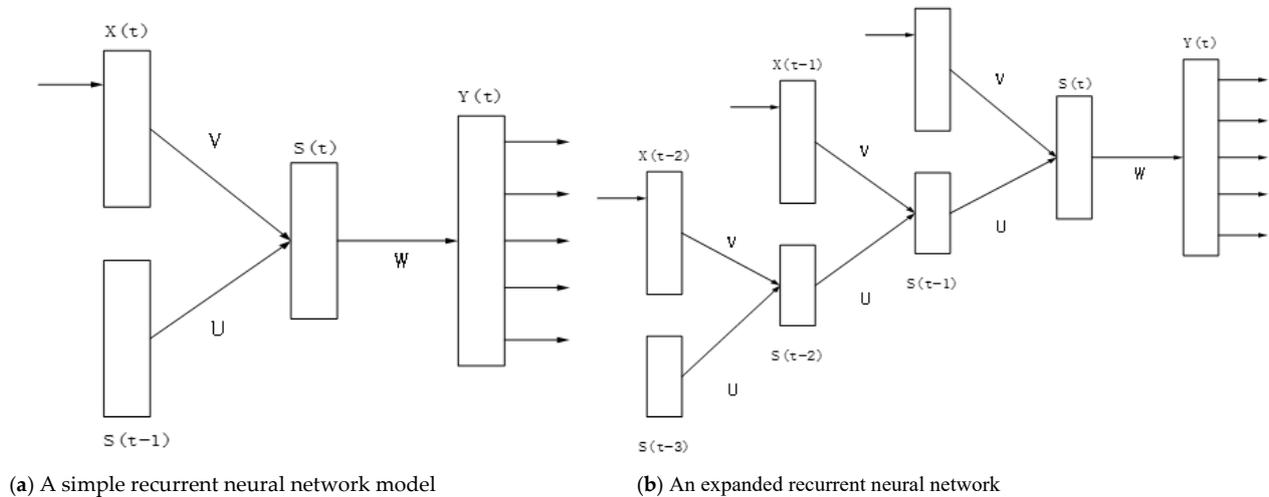


Figure 4. Classical recurrent neural network (RNN) models (a) A simple RNN; (b) An expanded RNN.

The input layer → the hidden layer. As shown in Equations (3) and (4):

$$s_j(t) = f(net_j(t)) \tag{3}$$

$$s_j(t) = f(net_j(t)) \tag{4}$$

The hidden layer → the output layer. As shown in Equations (5) and (6):

$$y_k(t) = g(net_k(t)) \tag{5}$$

$$net_k(t) = \sum_j^m s_j(t)w_{kj} + b_k \tag{6}$$

Among them, f and g are the hidden layer’s activation functions and the output layer, respectively. The activation function ensures the nonlinearity of the whole neural network, which significantly improves its expression ability. b_j and b_k are deviations.

- (2) Select the loss function and activation function

This paper chooses the variance sum function as the loss function of training the neural network, as shown in Equation (7):

$$C = \frac{1}{2} \sum_p^n \sum_k^o (d_{pk} - y_{pk})^2 \quad (7)$$

(3) Backpropagation algorithm along time

Design a neural network of more than 2 layers to capture longer historical information to transmit further the error (Figure 4b). Therefore, the backward propagation of error can be defined as Equations (8) and (9):

$$\delta = -\frac{\partial(C)}{\partial(net)} \quad (8)$$

$$\delta_{pj}(t-1) = \sum_h^m \delta_{ph}(t) u_{hj} f'(s_{pj}(t-1)) \quad (9)$$

where h is the hidden layer node index when the time step is t , and j is the hidden layer node index when the time step is $t-1$.

(4) Weight matrix update calculation

For brevity, the index p of the training sample is not added to our symbol here. Then, the error of the output layer is expressed as Equation (10):

$$e_o(t) = d(t) - y(t) \quad (10)$$

The output layer weights are expressed as Equation (11):

$$W(t+1) = W(t) + \eta s(t) e_o(t)^T \quad (11)$$

The error gradient propagates from the output layer to the hidden layer. As shown in Equation (12):

$$e_h(t) = d_h(e_o(t)^T V, t) \quad (12)$$

Obtain the error vector with the Equation (13):

$$d_{hj}(x, t) = x f'(net_j) \quad (13)$$

The weight matrix V is calculated as shown in the Equation (14):

$$V(t+1) = V(t) + \eta x(t) e_h(t)^T \quad (14)$$

The cyclic weight matrix W is calculated as shown in the Equation (15):

$$U(t+1) = U(t) + \eta s(t-1) e_h(t)^T \quad (15)$$

The research process predicts PM2.5 concentrations, the input for the past 24 h of PM2.5 density, and the current moment of PM10, O₃, CO, NO₂, and SO₂ output to predict the moment of PM2.5 density. As a result, the number of input layer neurons is 29. The number of output layer neurons is 1. The same as forecasting PM10 concentration, just change the input data of PM10 to PM2.5.

2.5. Evaluation

The main content of this study is to construct different deep neural network prediction models by changing the number of hidden layers and the number of neural units in each hidden layer, and to explore its impact on PM2.5/PM10 prediction accuracy.

In this study, we used root mean square error (RMSE) as the evaluation index [31–33]. The formula is as Equation (16):

$$RMSE = \sqrt{\frac{1}{m} \sum_i^m (T_i - P_i)^2} \quad (16)$$

where, i represents the number of test sample points and m represents the estimated total number of days. T_i is the actual concentration at the test sample point. P_i is the predicted concentration of the test sample point.

The experimental results are compared with the actual data at 360 times. Different hidden layers are selected to compare the effects of different hidden layers on the prediction accuracy of the model. For different hidden layers, the neuron number analysis prediction model is used.

In this research, the number of hidden layer nodes is determined according to the empirical formula $l = \sqrt{m + n} + \alpha$. The parameter m represents the number of input layer nodes, n represents the number of output layer nodes, α is a positive constant less than 10, and l represents the number of hidden layer neurons. Additionally, because the input of this model is 24 + 5 (24-h PM2.5 concentration value, and the other five pollutant concentration values at the current moment), the output node is 1.

Therefore, the number of hidden neurons is between 6–16. We chose 10 as the initial number of neurons in this model, and it decreases step by step according to the increase of the hidden layer.

The number of hidden layers and the distribution of neurons are shown in Table 3:

Table 3. Number of hidden layers and the distribution of neurons.

Hidden Layers	Neuron Distribution
1	10
2	10 9
3	10 9 8
4	10 9 8 7
5	10 9 8 7 6
6	10 9 8 7 6 5
7	10 9 8 7 6 5 4
8	10 9 8 7 6 5 4 3

Each group of experiments was conducted five times, and the mean value was selected as the experimental result for analysis.

The PM2.5/PM10 concentration values are divided into 6 grades in the process to explain the experimental results further. The classification of PM2.5/PM10 concentration value is shown in Table 4.

Table 4. Classification of PM2.5/PM10 concentration value.

Level	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6
Level range	(0–35)	(36–75)	(76–115)	(116–150)	(151–250)	(>250)

Suppose the predicted result and the actual measurements are in the same grade. In that case, the predicted result is judged to be superior. Suppose the predicted result and the actual measurements are in two adjacent grades. The predicted result is judged to be acceptable. Suppose the predicted results are more than two grades different from the actual measurements. The predicted results are judged to be unacceptable.

3. Results

The results and optimization of RNN model prediction are achieved by changing the number of hidden layers.

The prediction results are shown in Table 5. The results shows the RMSE and optimal evaluations with different number of hidden layer and neuron distribution.

Table 5. Deep RNN model prediction results.

Hidden Layers	Neuron Distribution	RMSE ($\mu\text{g}/\text{m}^3$)		Optimal		Acceptable		Unacceptable	
		PM2.5	PM10	PM2.5	PM10	PM2.5	PM10	PM2.5	PM10
1	10	12.59	19.29	78.89%	76.11%	20.83%	23.06%	0.28%	0.83%
2	10 9	10.26	18.91	81.39%	74.72%	18.33%	24.72%	0.28%	0.56%
3	10 9 8	9.93	17.11	80.83%	81.11%	18.89%	18.33%	0.28%	0.56%
4	10 9 8 7	9.31	17.00	83.33%	78.89%	16.67%	20.56%	0.00%	0.56%
5	10 9 8 7 6	8.89	16.87	85.83%	78.06%	14.17%	21.39%	0.00%	0.56%
6	10 9 8 7 6 5	8.81	16.62	85.83%	76.67%	14.17%	23.06%	0.00%	0.28%
7	10 9 8 7 6 5 4	8.69	16.62	86.11%	82.78%	13.89%	16.94%	0.00%	0.28%
8	10 9 8 7 6 5 4 3	8.63	15.99	85.00%	83.89%	15.00%	16.11%	0.00%	0.00%

Figures 5 and 6 are the comparison of the predictions of PM 2.5 by the RNN model at the number of 5 and 8 hidden layers, respectively.

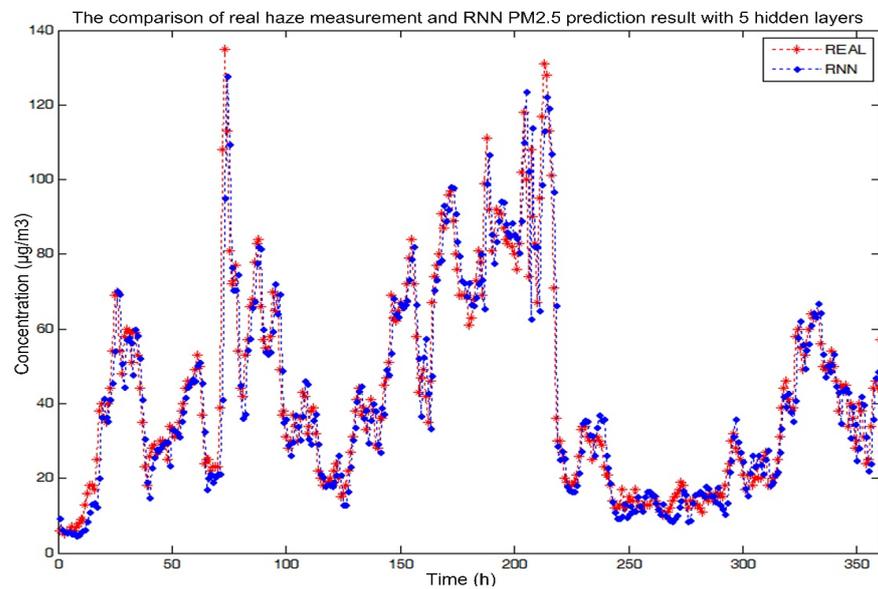


Figure 5. PM2.5 prediction results of 5 hidden layers of RNN network.

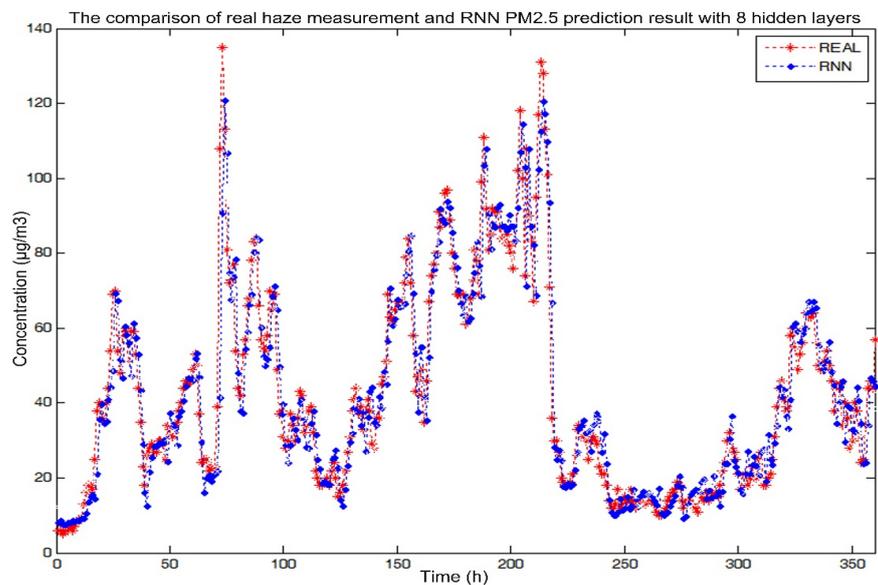


Figure 6. PM2.5 prediction results of 8 hidden layers of RNN network.

Figures 7 and 8 are the comparison of the PM 10 and actual results of the RNN model at the number of 5 and 7 hidden layers, respectively. After comparison, the model has small error and high prediction accuracy.

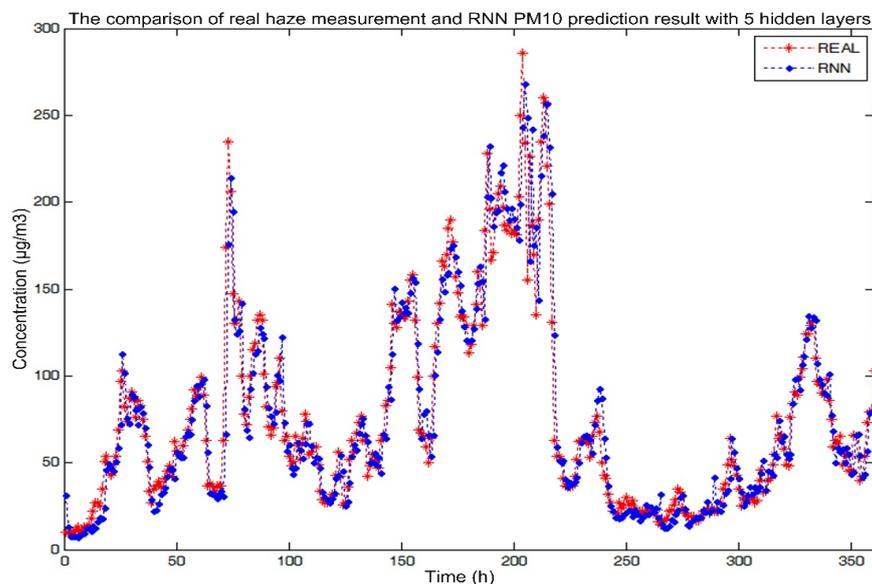


Figure 7. PM10 prediction results of 5 hidden layers of RNN network.

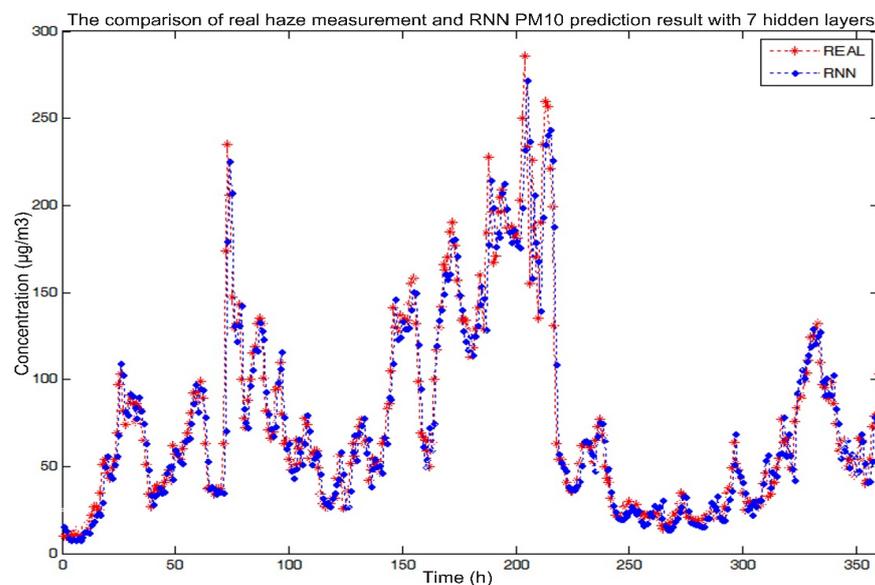


Figure 8. PM10 prediction results of 7 hidden layers of RNN network.

4. Discussion

Aiming at the PM_{2.5}/PM₁₀ concentration prediction problem, this paper builds a deep cyclic neural network composed of RNN. It predicts the PM_{2.5}/PM₁₀ concentration in Chengdu urban area. According to the prediction, the results are analyzed. The results show that under the same other conditions, the more hidden layers, the higher the prediction accuracy. When reaching a specific value, the accuracy is roughly the same; under the same network, the prediction accuracy of PM_{2.5} is significantly higher than that of PM₁₀.

Previous studies on PM_{2.5}/PM₁₀ prediction have mostly used linear regression methods, but in fact, the relationship between these factors is quite complex and often nonlinear. Therefore, in this paper, the deep learning theory is used to construct the haze prediction model. The deep cycling neural network prediction model and the deep,

long-term memory prediction model predict the PM_{2.5}/PM₁₀ concentration in the short term. Deep recurrent neural networks have many advantages, with stronger parameter prediction power than other neural networks and higher accuracy in the prediction results. RNN and long-term memory prediction models are examples of recurrent networks that take the hidden state of the previous layer as input and generate the hidden state of the current layer as output, improving its prediction accuracy and a larger processing range of data [34]. Compared with general models with ordinary structures, the dynamic properties of the data changing with time series are fully considered [35].

However, the research work is still in the preliminary stage. There are still many problems that can be further studied.

- (1) In the concentration prediction of PM_{2.5}/PM₁₀, this article gives up the meteorological data with the bottom correlation is abandoned. Instead, the highly correlated PM_{2.5}, PM₁₀, O₃, CO, NO₂, and SO₂ are selected as variables to train the prediction model. Therefore, the factors causing PM_{2.5} pollutions cannot be explored [36]. The scene of haze formation cannot be predicted with high accuracy, which has a certain impact on the experimental prediction results. How to improve the accuracy and study the experimental prediction using the factors causing PM_{2.5}/PM₁₀ pollution needs to be further improved.
- (2) In this paper, the prediction experiment is only given the Chengdu city PM_{2.5}/PM₁₀ concentration. Due to the geographical environment, the local climate and haze pollution situation is different, and the prediction results may not be the same, so in this experiment, the best model predicted results might not be applicable to other cities, to determine the best prediction model in other regions requires retraining in the mode [37]. Therefore, how to find a suitable multiregional haze prediction model still needs to be studied, thinking about further improving the scope of research to China, and making the prediction of PM_{2.5}/PM₁₀ concentration change.
- (3) The timescale chosen in this paper is small, the prediction of PM_{2.5}/PM₁₀ is only a short-term prediction, and the prediction model design is not sliding, so the prediction model effect is only suitable for the prediction of PM_{2.5}/PM₁₀ concentration content in short-term haze weather, and cannot be used to directly predict future long-term results. Instead, short-term predictions were used as variables to predict long-term concentrations of PM_{2.5}/PM₁₀ [38,39]. In the next step, we will choose the long-time scale and PM_{2.5} /PM₁₀ concentration content range to increase to the degree of heavy pollution, and conduct the time series analysis of the severe haze areas to study its development mechanism [40,41].

5. Conclusions

In this paper, interpolation completion and data standardization were carried out on the collected data. Correlation analysis was carried out on the processed data. The data were rounded up according to the correlation coefficient. The PM_{2.5}, PM₁₀, O₃, CO, NO₂, and SO₂ concentrations were finally selected as the input variables to predict PM_{2.5} and PM₁₀. The same training data set and test set are selected to build a deep recurrent neural network of RNN. The experimental conclusions are as follows:

- (1) The prediction results of PM_{2.5} and PM₁₀ are related to the number of hidden layers. The larger the number of hidden layers is, the higher the prediction accuracy is. Therefore, the prediction effect of PM_{2.5} is higher than that of PM₁₀, which is related to the data characteristics of PM_{2.5} and PM₁₀. As the concentration data of PM_{2.5} is lower than PM₁₀ and the concentration range is smaller, the RMS error is small, and the accuracy is high.
- (2) According to the prediction results of PM_{2.5}, the prediction accuracy is related to the number of hidden layers in the case of determining the number of nodes. The prediction results of the deep recurrent neural network are as follows: with the increase of the number of hidden layers, the higher the prediction accuracy is, and the deep recurrent neural network with eight hidden layers has the best prediction results.

- (3) According to the PM10 prediction results, as with the PM2.5 prediction results, prediction accuracy is related to the number of hidden layers. The larger the number of hidden layers is, the higher the prediction accuracy will be. The prediction results of a deep recurrent neural network with eight hidden layers are the best.
- (4) To ensure prediction accuracy, the most simplified neural network is selected to predict PM2.5/PM10. Good results can be achieved when PM2.5/PM10 is predicted using a deep recurrent neural network with seven hidden layers.

Author Contributions: Conceptualization, W.Z. and Z.L.; methodology, L.Y. and S.L.; software, Z.C. and B.Y.; formal analysis, B.Y. and L.S.; data curation, K.S., Z.C. and L.S.; writing—original draft preparation, K.S. and S.L.; writing—review and editing, L.Y. and W.Z.; funding acquisition, W.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This work was jointly supported by the Sichuan Science and Technology Program (2021YFQ0003).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data sets are available at Chengdu air quality online monitoring and analysis platform (<https://www.aqistudy.cn>, accessed on 17 March 2018) which are provided by The Environmental Protection Agency of Chengdu.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lv, Z.; Wang, X.; Deng, F.; Ying, Q.; Archibald, A.T.; Jones, R.L.; Ding, Y.; Cheng, Y.; Fu, M.; Liu, Y.; et al. Source–Receptor Relationship Revealed by the Halted Traffic and Aggravated Haze in Beijing during the COVID-19 Lockdown. *Environ. Sci. Technol.* **2020**, *54*, 15660–15670. [[CrossRef](#)] [[PubMed](#)]
2. Wu, J.; Zhang, P.; Yi, H.; Qin, Z. What Causes Haze Pollution? An Empirical Study of PM2.5 Concentrations in Chinese Cities. *Sustainability* **2016**, *8*, 132. [[CrossRef](#)]
3. Liu, Y.; Dong, F. How Industrial Transfer Processes Impact on Haze Pollution in China: An Analysis from the Perspective of Spatial Effects. *Int. J. Environ. Res. Public Health* **2019**, *16*, 423. [[CrossRef](#)] [[PubMed](#)]
4. Li, X.; Zheng, W.; Yin, L.; Yin, Z.; Song, L.; Tian, X. Influence of Social-economic Activities on Air Pollutants in Beijing, China. *Open Geosci.* **2017**, *9*, 314–321. [[CrossRef](#)]
5. Zheng, W.; Li, X.; Yin, L.; Wang, Y. Spatiotemporal heterogeneity of urban air pollution in China based on spatial analysis. *Rend. Lincei* **2016**, *27*, 351–356. [[CrossRef](#)]
6. Zheng, W.; Li, X.; Xie, J.; Yin, L.; Wang, Y. Impact of human activities on haze in Beijing based on grey relational analysis. *Rend. Lincei* **2015**, *26*, 187–192. [[CrossRef](#)]
7. Baklanov, A.; Zhang, Y. Advances in air quality modeling and forecasting. *Glob. Transit.* **2020**, *2*, 261–270. [[CrossRef](#)]
8. Gao, L.; Liu, Z.; Chen, D.; Yan, P.; Zhang, Y.; Hu, H.; Liang, H.; Liang, X. GPS-ZTD data assimilation and its impact on wintertime haze prediction over North China Plain using WRF 3DVAR and CMAQ modeling system. *Environ. Sci. Pollut. Res.* **2021**, 1–16. [[CrossRef](#)] [[PubMed](#)]
9. Yin, L.; Li, X.; Zheng, W.; Yin, Z.; Song, L.; Ge, L.; Zeng, Q. Fractal dimension analysis for seismicity spatial and temporal distribution in the circum-Pacific seismic belt. *J. Earth Syst. Sci.* **2019**, *128*, 22. [[CrossRef](#)]
10. Hertwig, D.; Burgin, L.; Gan, C.; Hort, M.; Jones, A.; Shaw, F.; Witham, C.; Zhang, K. Development and demonstration of a Lagrangian dispersion modeling system for real-time prediction of smoke haze pollution from biomass burning in Southeast Asia. *J. Geophys. Res. Atmos.* **2015**, *120*, 12605–12630. [[CrossRef](#)]
11. Zheng, W.; Li, X.; Yin, L.; Wang, Y. The retrieved urban LST in Beijing based on TM, HJ-1B and MODIS. *Arab. J. Sci. Eng.* **2016**, *41*, 2325–2332. [[CrossRef](#)]
12. Liu, D.; Liu, X.; Wang, H.; Li, Y.; Kang, Z.; Cao, L.; Yu, X.; Chen, H. A New Type of Haze? The December 2015 Purple (Magenta) Haze Event in Nanjing, China. *Atmosphere* **2017**, *8*, 76. [[CrossRef](#)]
13. Singh, V.; Carnevale, C.; Finzi, G.; Pisoni, E.; Volta, M. A cokriging based approach to reconstruct air pollution maps, processing measurement station concentrations and deterministic model simulations. *Environ. Model. Softw.* **2011**, *26*, 778–786. [[CrossRef](#)]
14. Zheng, W.; Liu, X.; Ni, X.; Yin, L.; Yang, B. Improving Visual Reasoning Through Semantic Representation. *IEEE Access* **2021**, *9*, 91476–91486. [[CrossRef](#)]
15. Ma, Z.; Zheng, W.; Chen, X.; Yin, L. Joint embedding VQA model based on dynamic word vector. *PeerJ Comput. Sci.* **2021**, *7*, e353. [[CrossRef](#)]

16. Zheng, W.; Liu, X.; Yin, L. Sentence Representation Method Based on Multi-Layer Semantic Network. *Appl. Sci.* **2021**, *11*, 1316. [[CrossRef](#)]
17. Tang, Y.; Liu, S.; Deng, Y.; Zhang, Y.; Yin, L.; Zheng, W. An improved method for soft tissue modeling. *Biomed. Signal Process. Control.* **2021**, *65*, 102367. [[CrossRef](#)]
18. Yao, L.; Lu, N.; Jiang, S. Artificial Neural Network (ANN) for multi-source PM2.5 estimation using surface, MODIS, and meteorological data. In Proceedings of the 2012 International Conference on Biomedical Engineering and Biotechnology, Macao, China, 28–30 May 2012; pp. 1228–1231.
19. Wu, X.; Liu, Z.; Yin, L.; Zheng, W.; Song, L.; Tian, J.; Yang, B.; Liu, S. A Haze Prediction Model in Chengdu Based on LSTM. *Atmosphere* **2021**, *12*, 1479. [[CrossRef](#)]
20. Yin, L.; Wang, L.; Huang, W.; Liu, S.; Yang, B.; Zheng, W. Spatiotemporal Analysis of Haze in Beijing Based on the Multi-Convolution Model. *Atmosphere* **2021**, *12*, 1408. [[CrossRef](#)]
21. Chakma, A.; Vizena, B.; Cao, T.; Lin, J.; Zhang, J. Image-based air quality analysis using deep convolutional neural network. In Proceedings of the 2017 IEEE International Conference on Image Processing (ICIP), Beijing, China, 17–20 September 2017; pp. 3949–3952.
22. Zhang, Z.; Wang, L.; Zheng, W.; Yin, L.; Hu, R.; Yang, B. Endoscope image mosaic based on pyramid ORB. *Biomed. Signal Process. Control.* **2022**, *71*, 103261. [[CrossRef](#)]
23. Zhang, Z.; Tian, J.; Huang, W.; Yin, L.; Zheng, W.; Liu, S. A Haze Prediction Method Based on One-Dimensional Convolutional Neural Network. *Atmosphere* **2021**, *12*, 1327. [[CrossRef](#)]
24. Ma, L.; Wang, F. The prediction of haze based on BP neural network and matlab. In Proceedings of the 2017 International Conference on Computer Network, Electronic and Automation (ICCNEA), Xi'an, China, 23–25 September 2017; pp. 210–220.
25. Chang-Hoi, H.; Park, I.; Oh, H.-R.; Gim, H.-J.; Hur, S.-K.; Kim, J.; Choi, D.-R. Development of a PM2.5 prediction model using a recurrent neural network algorithm for the Seoul metropolitan area, Republic of Korea. *Atmos. Environ.* **2021**, *245*, 118021. [[CrossRef](#)]
26. Tsai, Y.; Zeng, Y.; Chang, Y. Air pollution forecasting using RNN with LSTM. In Proceedings of the 2018 IEEE 16th Intl Conf on Dependable, Autonomic and Secure Computing, 16th Intl Conf on Pervasive Intelligence and Computing, 4th Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress (DASC/PiCom/DataCom/CyberSciTech), Athens, Greece, 12–15 August 2018; pp. 1074–1079.
27. Nanduri, A.; Sherry, L. Anomaly detection in aircraft data using Recurrent Neural Networks (RNN). In Proceedings of the 2016 Integrated Communications Navigation and Surveillance (ICNS), Herndon, VA, USA, 19–21 April 2016; pp. 5C2-1–5C2-8.
28. Scesa, V.; Henaff, P.; Ouezdou, F.B.; Namoun, F. Time window width influence on dynamic BPTT(h) learning algorithm performances: Experimental study. In *Artificial Neural Networks—ICANN 2006*; Kollias, S.D., Stafylopatis, A., Duch, W., Oja, E., Eds.; Springer: Berlin/Heidelberg, Germany, 2006; pp. 93–102.
29. Pascanu, R.; Gulcehre, C.; Cho, K.; Bengio, Y. How to construct deep recurrent neural networks. *arXiv* **2013**, arXiv:1312.6026.
30. Banerjee, I.; Ling, Y.; Chen, M.C.; Hasan, S.A.; Langlotz, C.P.; Moradzadeh, N.; Chapman, B.; Amrhein, T.; Mong, D.; Rubin, D.L.; et al. Comparative effectiveness of convolutional neural network (CNN) and recurrent neural network (RNN) architectures for radiology text report classification. *Artif. Intell. Med.* **2019**, *97*, 79–88. [[CrossRef](#)] [[PubMed](#)]
31. Li, Y.; Zheng, W.; Liu, X.; Mou, Y.; Yin, L.; Yang, B. Research and improvement of feature detection algorithm based on FAST. *Rendiconti Lincei. Scienze Fisiche e Naturali* **2021**, *32*, 775–789. [[CrossRef](#)]
32. Zheng, W.; Yin, L.; Chen, X.; Ma, Z.; Liu, S.; Yang, B. Knowledge base graph embedding module design for Visual question answering model. *Pattern Recognit.* **2021**, *120*, 108153. [[CrossRef](#)]
33. Zheng, W.; Liu, X.; Yin, L. Research on image classification method based on improved multi-scale relational network. *PeerJ Comput. Sci.* **2021**, *7*, e613. [[CrossRef](#)] [[PubMed](#)]
34. Zeebaree, D.Q.; Abdulazeez, A.M.; Abdullrhman, L.M.; Hasan, D.A.; Kareem, O.S. The Prediction Process Based on Deep Recurrent Neural Networks: A Review. *Asia J. Res. Comput. Sci.* **2021**, *11*, 29–45. [[CrossRef](#)]
35. Rao, K.S.; Devi, G.L.; Ramesh, N. Air quality prediction in visakhapatnam with lstm based recurrent neural networks. *Int. J. Intell. Syst. Appl.* **2019**, *11*, 18–24. [[CrossRef](#)]
36. Wu, Z.; Zhang, S. Study on the spatial-temporal change characteristics and influence factors of fog and haze pollution based on GAM. *Neural Comput. Appl.* **2019**, *31*, 1619–1631. [[CrossRef](#)]
37. Chang, L.; Wu, Z.; Xu, J. Potential impacts of the Southern Hemisphere polar vortices on central-eastern China haze pollution during boreal early winter. *Clim. Dyn.* **2020**, *55*, 771–787. [[CrossRef](#)]
38. Yin, Z.; Wang, H.; Duan, M. Outline of the realtime seasonal haze pollution prediction in China in recent years. *Trans. Atmos. Sci.* **2019**, *42*, 2–13.
39. Liu, Y.; Tian, J.; Zheng, W.; Yin, L. Spatial and temporal distribution characteristics of haze and pollution particles in China based on spatial statistics. *Urban Clim.* **2022**, *41*, 101031. [[CrossRef](#)]
40. Xu, J.; Liu, Z.; Yin, L.; Liu, Y.; Tian, J.; Gu, Y.; Zheng, W.; Yang, B.; Liu, S. Grey Correlation Analysis of Haze Impact Factor PM2.5. *Atmosphere* **2021**, *12*, 1513. [[CrossRef](#)]
41. Chen, X.; Yin, L.; Fan, Y.; Song, L.; Ji, T.; Liu, Y.; Tian, J.; Zheng, W. Temporal evolution characteristics of PM2.5 concentration based on continuous wavelet transform. *Sci. Total Environ.* **2020**, *699*, 134244. [[CrossRef](#)]