

Research Article

Application of Offshore Visibility Forecast Based on Temporal Convolutional Network and Transfer Learning

Zhenyu Lu^{1,2}, Cheng Zheng^{1,3} and Tingya Yang⁴

¹School of Artificial Intelligence, Nanjing University of Information Science and Technology, Nanjing, China

²Jiangsu Collaborative Innovation Center on Atmospheric Environment and Equipment, Nanjing, China

³School of Electronic and Information Engineering, Nanjing University of Information Science and Technology, Nanjing, China

⁴Jiangsu Meteorological Observatory, Nanjing, China

Correspondence should be addressed to Zhenyu Lu; luzhenyu76@163.com

Received 22 June 2020; Revised 3 September 2020; Accepted 5 October 2020; Published 20 October 2020

Academic Editor: Nian Zhang

Copyright © 2020 Zhenyu Lu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Visibility forecasting in offshore areas faces the problems of low observational data and complex weather. This paper proposes an intelligent prediction method of offshore visibility based on temporal convolutional network (TCN) and transfer learning to solve the problem. First, preprocess the visibility data sets of the source and target domains to improve the quality of the data. Then, build a model based on temporal convolutional network and transfer learning (TCN_TL) to learn the visibility data of the source domain. Finally, after transferring the knowledge learned from a large amount of data in the source domain, the model learns the small data set in the target domain. After completing the training, the model data of the European Mid-Range Weather Forecast Center (ECMWF) meteorological field were selected to test the model performance. The method proposed in this paper has achieved relatively good results in the visibility forecast of Qiongzhou Strait. Taking Haikou Station in the spring and winter of 2018 as an example, the forecast error is significantly lower than that before the transfer learning, and the forecast score is increased by 0.11 within the 0–1 km level and the 24 h forecast period. Compared with the CUACE forecast results, the forecast error of TCN_TL is smaller than that of the former, and the TS score is improved by 0.16. The results show that under the condition of small data sets, transfer learning improves the prediction performance of the model, and TCN_TL performs better than other deep learning methods and CUACE.

1. Introduction

Atmospheric visibility is an indicator used to judge the transparency of the atmosphere. It refers to the maximum horizontal distance that the person with normal vision can distinguish the outline of the target from the background when observing the black target with the sky as the background under the weather conditions at that time [1]. There are many climatic factors that affect the visibility of offshore waters, such as fog, haze, smoke, dust, precipitation, etc. The most important factor is fog [2, 3]. Low visibility often affects the travel safety of offshore vessels and can easily cause accidents at sea. Therefore, how to predict the occurrence of low-visibility weather in offshore areas as much as possible is a problem that researchers are concerned about. Here, the offshore visibility prediction method based on transfer

learning proposed in this paper aims to improve the current situation of offshore visibility prediction.

The visibility forecast used to rely on traditional numerical forecasting methods. The weather conditions in the future were calculated through numerical forecasting methods based on the theoretical basis of fluid mechanics, atmospheric dynamics, and thermodynamics. In recent years, with the continuous development and improvement of machine learning and deep learning, the application of machine learning and deep learning to visibility prediction has also become a hot spot for researchers [4–6].

As a commonly used meteorological forecasting method, traditional numerical forecasting is widely used. The National Center for Atmospheric Research (NCAR) has jointly developed a multiscale, multiprocess model system that is online and fully coupled. It is a widely used

meteorological–chemical coupled model named weather research and forecast model coupled with chemistry (WRF-Chem) [7, 8]. The US Environmental Protection Agency (EPA) has also developed a universal multiscale air quality model named congestion mitigation and air quality (CMAQ) [9, 10]. These two forecasting models have been localized in various regions in China, and the WRF-Chem model has been introduced in East China to establish a numerical forecast model named Beijing regional environmental meteorology prediction system (BREMPs) in East China. It is applied to visibility prediction in East China [11, 12]. The introduction of the CMAQ model in the southern region established the Southern China regional environmental meteorological numerical forecasting model named Guangdong Regional Assimilation Chemistry Environmental System (GRACES), which was applied to the operational forecasting of the Pearl River Delta [9]. At the same time, the domestic environmental meteorological numerical model is also continuously researched. The Chinese Academy of Meteorological Sciences has independently developed a national-level numerical haze forecasting system named CMA Unified Atmosphere Chemistry Environment (CUACE). CUACE online coupled with the chemical weather model, which has been used in environmental meteorological operations nationwide [13, 14].

In recent years, researchers have achieved a lot of results in machine learning and deep learning [15–19] and have applied them to various fields [20–24]. Among them, machine learning and deep learning also show advantages in weather forecasting such as visibility forecasting and air quality forecasting [25–28]. Machine learning algorithms have excellent performance in visibility forecasting. The Support Vector Machine (SVM) method was used to select multiple kernel functions for the forecasting modeling experiment of low-visibility weather in Shuangliu Airport and study the impact of various meteorological elements on visibility [29]. Long short-term memory–fully connected (LSTM-FC) neural network was used to predict the pollutants in Beijing area and achieved good results [30]. Li et al. used a hybrid CNN-LSTM model developed by combining the convolutional neural network (CNN) with the long short-term memory (LSTM) neural network for forecasting the next 24 h PM2.5 concentration in Beijing [31]. BP neural network was used to construct the visibility forecast model of the Bohai Rim city to reduce the visibility forecast error [32]. In terms of visibility prediction for offshore waters, people used classification and tree regression to establish a forecast model of sea fog along the coast of Qingdao [33]. However, these methods are not ideal for forecasting offshore visibility. They can only predict the general trend of visibility. The accuracy of low-visibility forecasting and visibility-level forecasting needs to be improved.

Both machine learning and deep learning network models are inseparable from large amounts of data. Using deep learning to predict visibility requires a large amount of meteorological data to train the network. However, offshore observation sites are scarce, so using deep learning to face offshore visibility forecasting is faced with the problem of insufficient data. Therefore, this article aims to use transfer learning to solve this problem.

Transfer learning is a machine learning method and a problem-solving idea [34]. In recent years, it has been widely used in various research fields to solve the difficulties caused by insufficient data and achieved excellent performance in building energy prediction and text prediction [35–37]. It is a way to achieve the knowledge transfer between the source and target domains by modeling the distribution of data in the source and target domains, thereby improving the performance of the algorithm [38]. The transfer learning method can better solve the problem of poor forecast performance caused by less data.

In order to improve the forecast accuracy of offshore visibility under the condition of a small data set, this paper uses time convolutional network (TCN) and transfer learning to establish a forecast model to achieve an objective forecast of visibility in offshore areas such as Qiongzhou Strait. At the same time, it improves the ability of forecasting and early warning services for the haze and sea fog in the strait, provides technical support for the low-visibility weather forecast in this area, and ensures the safety of ships traveling.

2. Materials and Methods

2.1. Data Source and Data Preprocessing. The climatic conditions in the coastal area of South China are subtropical monsoon climate, which belongs to the East Asian monsoon region. The meteorological background is greatly affected by relative humidity, and sea fog occurs frequently in winter and spring, so low-visibility weather occurs more frequently. The training data used in this article is divided into two parts. The first part is the meteorological data and environmental data of the stations in Leizhou Peninsula and northern Hainan, which is the source domain data. The second part is a small amount of meteorological data and environmental protection data of the weather stations on both sides of the Qiongzhou Strait, which is the target domain data. The specific data of the first part include routine ground observation data and high-altitude data of Leizhou Peninsula area and northern Hainan area from 2016 to 2018, such as data on wind speed, relative humidity, temperature, pollutants, etc., and visibility observation data. The specific data of the second part include visibility observation data and meteorological data and environmental protection data of the offshore sites of the Qiongzhou Straits on both sides of the north and south sides of the Qiongzhou Strait from January to April 2016–2018. The verification data used in this paper is the European Meteorological Forecast Center (ECMWF) meteorological field model data in 2018. In this paper, four inland sites are selected as source domain site data, namely, Haikang (59750), Danxian (59845), Chengmai (59843), and Anding (59851). The data of four stations on the seashore were selected to conduct a forecast experiment on the visibility of Qiongzhou Strait, namely, Xu Wen (59754), Haikou (59758), Lingao (59842), and Qiongshan (59757). Figure 1 shows the distribution of the eight sites.

Because the weather forecasting factors are composed of different parts, it is necessary to compose the data for time series analysis and time-space matching and then compose the input data that meets the requirements. At the same time, due to various factors in the real-time observation of

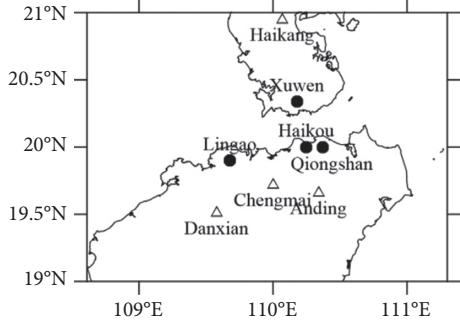


FIGURE 1: Distribution of representative sites of Qiongzhou Strait.

the meteorological conditions by the meteorological station, the meteorological observation data will often have some observations missing and abnormal, so the data should be cleaned. This paper uses Lagrange interpolation to fill in missing values. Lagrange interpolation formula is

$$y = a_0 + a_1 x + a_2 x^2 + \cdots + a_{n-1} x^{n-1}, \quad (1)$$

$$\begin{aligned} L(x) = & y_1 \frac{(x - x_2)(x - x_3) \cdots (x - x_n)}{(x_1 - x_2)(x_1 - x_3) \cdots (x_1 - x_n)} \\ & + y_2 \frac{(x - x_1)(x - x_3) \cdots (x - x_n)}{(x_2 - x_1)(x_2 - x_3) \cdots (x_2 - x_n)} \\ & + \cdots y_n \frac{(x - x_1)(x - x_3) \cdots (x - x_n)}{(x_n - x_1)(x_n - x_3) \cdots (x_n - x_{n-1})}. \end{aligned} \quad (2)$$

Equation (2) solves the Lagrange interpolation polynomial $L(x)$ and then substitutes the point corresponding to the missing value x to obtain the approximate value of the missing value.

The visibility data sample is composed of two parts, meteorological data and environmental protection data, which have a large number of features. In order to improve efficiency, the characteristics of the data samples need to be screened. In this paper, the Pearson correlation coefficient method is used to measure the degree of correlation between two variables. The Pearson correlation coefficient method is used to calculate the correlation between the features that affect visibility and the observation of visibility, and the features are screened by comparing the correlations of the various impact features. The formula of the correlation coefficient is shown in formula (3). In this paper, X represented the meteorological feature in the data, and Y represented the value of visibility:

$$\rho_{XY} = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y} = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y}. \quad (3)$$

The forecast meteorological elements of visibility in the data include 40 different forecast meteorological elements such as temperature, relative humidity, wind shear, and wind speed. The meteorological elements of visibility forecast after screening by the Pearson coefficient method are shown in Table 1. The visibility forecast meteorological elements screened by the Pearson coefficient method are shown in Table 1. In this paper, 12 meteorological elements are selected

and input into the network as forecast factors. Table 1 shows the correlation coefficients between each forecasting factor and the observation of visibility. It can be clearly observed that the correlation coefficients of temperature and humidity are relatively high. Visibility in Haikou area is highly correlated with temperature, relative humidity, etc., which is a good proof of the reliability of using the correlation coefficient method to filter visibility prediction factors here [39]. The forecast feature data includes historical weather data and historical environmental data, which are matched in time and space to form the original visibility data. The raw data are cleaned to obtain the final visibility forecast data. The processing flow is shown in Figure 2.

2.2. Method

2.2.1. Temporal Convolutional Network. Because the Temporal Convolutional Network (TCN) is a network structure that can better handle time series data, this paper uses TCN to build a prediction model of the visibility of the source domain. TCN is the second architecture that can analyze temporal data in addition to the Recurrent Neural Network (RNN) architecture. The structure of TCN is shown in Figure 2. TCN has two main characteristics: First, there is a causal relationship between the layers of the convolutional network, which means that there will be no “bobble” historical information or future data. Even if a long short-term memory network (LSTM) with the same time series processing function has a memory gate, it cannot completely remember all the historical information, let alone if some information is useless in the LSTM will gradually be forgotten [40]. Second, architecture of TCN can be flexibly adjusted to any length, and it can be mapped to correspond to several interfaces according to the output terminal. This is the same as the RNN framework, which is very convenient. The TCN network adopts the form of convolution, which is mainly composed of causal convolution, hole convolution, residual module, and full convolution network. Among them, causal convolution is used to make the network suitable for sequence models. The output of the convolution layer at time t is only convolved with the elements of the current layer and the previous layer. The causal convolution calculation formula at x_t is

$$(F * X)(x_t) = \sum_{k=1}^K f_k x_{t-K+k}. \quad (4)$$

Among them, filter is $F = (f_1, f_2, \dots, f_k)$, and input sequence is $X = (x_1, x_2, \dots, x_T)$. The use of the hole convolution and residual modules allows the TCN network to remember history. The hole convolution kernel is

$$F(s) = (x * _d f)(s) = \sum_{i=0}^{k-1} f(i) \cdot X_{s-d-i}, \quad (5)$$

where d is the hole coefficient, k is the size of the filter, and $s - d - i$ is the past direction [40]. The structure of the hollow convolution is shown in Figure 3(a). When the filter is

TABLE 1: Forecast factors and correlation coefficients.

Predictor	Pearson's correlation coefficient absolute value
Temperature	0.563
925 hPa temperature	0.511
900 hPa temperature	0.498
950 hPa horizontal wind speed	0.489
Atmospheric pressure	0.473
925 hPa horizontal wind speed	0.397
Vertical wind speed	0.390
925 hPa relative humidity	0.379
900 hPa horizontal wind speed	0.372
900 hPa relative humidity	0.338
Depression of the dew point	0.333
Relative humidity	0.328
950 hPa vertical wind speed	0.323

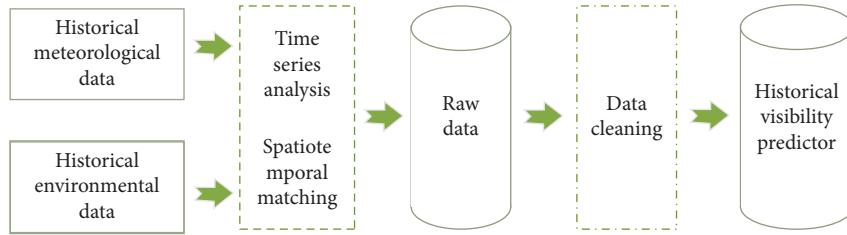


FIGURE 2: Data preprocessing flowchart.

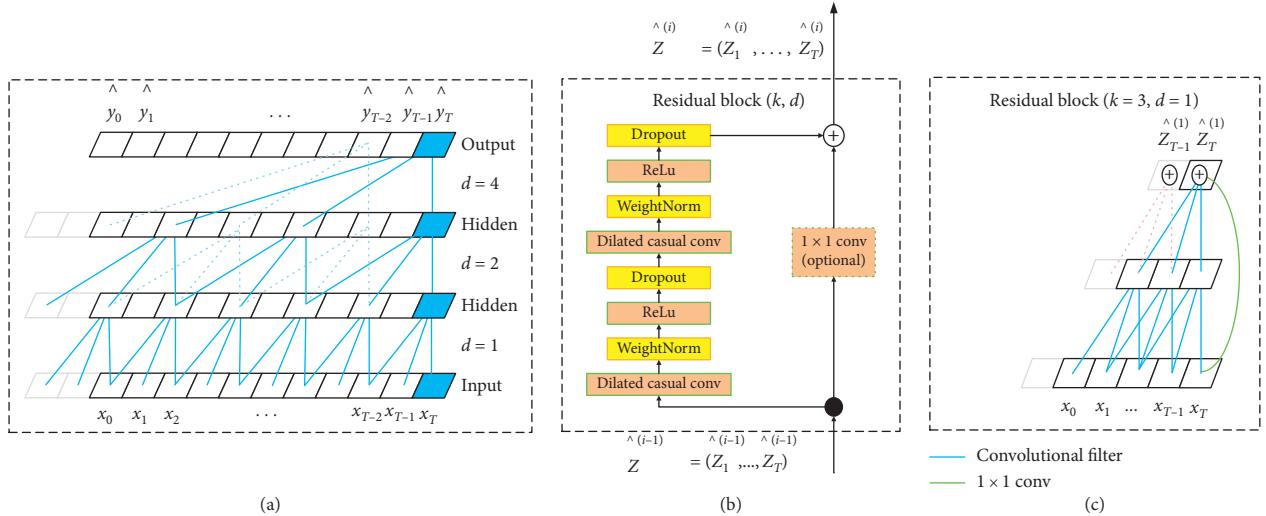


FIGURE 3: TCN architecture diagram. (a) Skip connection in TCN residual module. (b) Residual module of TCN. (c) Skip connection in TCN residual module.

$F = (f_1, f_2, \dots, f_k)$, input sequence is $X = (x_1, x_2, \dots, x_T)$, and the convolution of the holes at x_t is

$$(F *_d X)(x_t) = \sum_{k=1}^K f_k x_{t-(K-k)d}. \quad (6)$$

In formula (6), d can increase the receptive field, and the size is $(K-1)d + 1$. The formula for the residual module is

$$o = \text{activation}(x + \mathcal{F}(x)). \quad (7)$$

In formula (7), \mathcal{F} is a part of a series of transformations, and x is an input. The structure diagrams are shown in Figures 3(b) and 3(c).

Full convolutional networks make the output and input dimensions consistent, simplifying the network. It is these parts that make up TCN, which makes the network have the advantages of good parallelism, flexible receptive fields, small training memory, and adjustable input sequence length compared with LSTM.

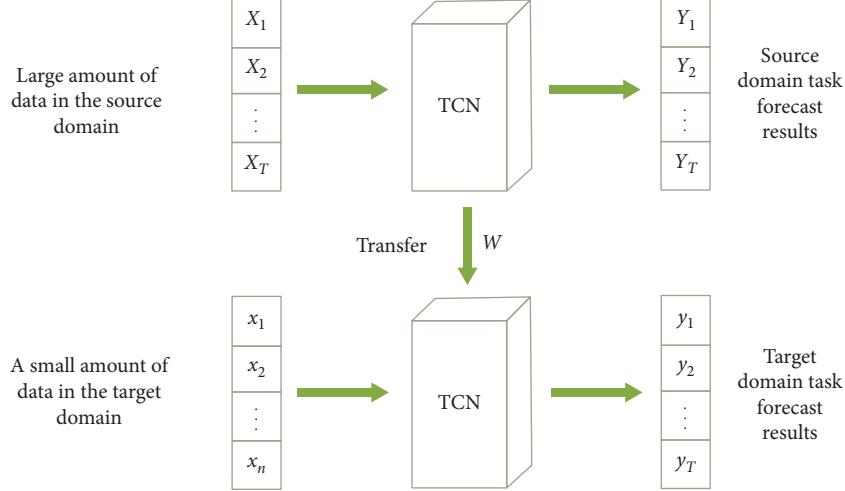


FIGURE 4: The transfer process of the source domain model to the target domain model.

2.2.2. Transfer Learning. Time convolutional networks belong to deep neural networks. During the training process of deep neural networks, the network will generate a large number of network parameters, so a large amount of data is needed to train the parameters. However, there is a problem of low data volume in offshore visibility, and the learning effect is not good. Therefore, the method of transfer learning is used to solve this problem.

However, there is a problem of low data volume in offshore visibility, and the learning effect is not good. Therefore, the method of transfer learning is used to solve this problem. Researcher defined the transfer learning in detail: given source domain $D_S = \{(X_1^{(S)}, Y_1^{(S)}), \dots, (X_{n_s}^{(S)}, Y_{n_s}^{(S)})\}$ and source domain learning tasks T_s , as well as target domain $D_t = \{X_1^{(t)}, \dots, X_{n_t}^{(t)}\}$ and target domain learning tasks T_t . The ultimate goal of transfer learning is to gradually improve the performance of the prediction function in the target domain D_s and the target domain task T_s by learning the knowledge in the source domain D_t and the source domain task T_s [41]. In this paper, a prediction model is established in the source domain task, and a large amount of source domain data is used as the training data of the model, and the pretrained model is saved after training. Transfer learning is used to make the target domain task network inherit the weights of the pretrained model. When performing a new task, a small amount of new visibility data is used as input, and the weight of the pretrained model loaded into the source domain task is used as the initial weight of the target domain network for training. The training process is shown in Figure 4.

2.3. Evaluation Method. From January to April 2018, there was frequent sea fog in the Qiongzhou Strait area with low visibility. The classification of visibility observation data during this period is relatively clear, and the selection of data during this period can make the experiment comprehensive effect better.

In this paper, there are two ways to evaluate the forecast performance, which are numerical test and classification test. The numerical test uses two indicators: root mean square

error (RMSE) and mean absolute error (MAE). RMSE is used to measure the deviation between the observed value and the predicted value. It usually reflects the precision of the measurement well. The smaller the value of RMSE, the higher the precision. MAE represents the average value of the absolute error between the predicted value and the observed value. The smaller the value of MAE, the higher the accuracy of prediction. MAE represents the average value of the absolute error between the predicted value and the observed value. The smaller the value of MAE, the higher the accuracy of prediction. The calculation formula of the two is as follows:

$$\text{RMSE}(X, P, O) = \sqrt{\frac{1}{m} \sum_{i=1}^m (p(x_i) - o_i)^2}, \quad (8)$$

$$\text{MAE}(X, P, O) = \frac{1}{m} \sum_{i=1}^m |p(x_i) - o_i|. \quad (9)$$

In formulas (8) and (9), m represents the number of samples, X represents the forecast factor of historical visibility, P represents the predicted value of the model for visibility, and O represents the observed value of historical visibility.

In this paper, the visibility level is mainly divided into four levels, namely, 0~1 km, 1~5 km, 5~10 km, and above 10 km. Among them, improving the forecast accuracy of the visibility level of 0 to 1 km has important practical guiding significance. The graded forecast test uses TS scoring rules to test forecast performance of the model for each visibility level. The formula for the test method is as follows:

$$\text{TS} = \frac{\text{NA}}{\text{NA} + \text{NB} + \text{NC}}. \quad (10)$$

In the formula, NA represents the number of correct forecasts, NB represents the number of empty forecasts, and NC represents the number of missed forecasts. Correct forecast means that the forecast level is the same as the live

level; empty forecast means that the forecast level is less than the live level; and missed forecast means that the forecast level is higher than the live level.

2.4. Forecasting Process. The forecast technology flow used in this paper is shown in Figure 5. This forecast model uses the TCN model and transfer learning method to forecast the Qiongzhou Strait. In the first step, the source domain prediction factor obtained after data preprocessing is used as the input feature of the TCN network to train the network. Based on the loss function result, adjust the parameters to optimize the model performance iteratively and save the optimal training model. In the second step, the transfer learning method based on parameter transfer is used to transfer the weight of the pretrained model in the source domain as the initial weight of the network and start learning the data in the target domain. Finally, input the EWMCF meteorological field data of the forecast period to check the model performance and output the forecast results.

3. Results and Discussion

3.1. Lab Environment. The experiment in this paper is implemented on TensorFlow 1.14.0 framework under Ubuntu 16.04 system and uses GPU to accelerate. The hardware configuration of the experimental platform is CPU: Intel Core i5-8600k, GPU: NVIDIA GTX 1080Ti, and memory is 16G.

3.2. Experimental Parameter Setting. Hyperparameters are the parameters set by the network model before learning, not the parameters obtained through training. Hyperparameter settings are crucial to the efficiency of model learning. Under normal circumstances, the model training should be selected before the start of learning. After observing the value of the loss function and the training status during the model learning process and adjusting the hyperparameters, the model learning efficiency is the highest. Among many hyperparameters, hyperparameters such as learning rate and batch size have the greatest impact on the efficiency and accuracy of learning. After repeated experiments, the model performs best when the learning rate is 0.002, the batch size is 60, and the hidden unit is 150.

3.3. Result Analysis

3.3.1. Comparison of Results before and after Transfer Learning. After establishing the prediction model named TCN_TL based on TCN and transfer learning, the network was tested and evaluated by using the offshore visibility data from January to April 2018. Figure 6 shows the MAE and RMSE of the forecasted visibility values of Haikou Station before and after the transfer learning in February 2018, including 24 h, 48 h, 72 h, and 96 h four-time error comparison. Figure 6 takes the Haikou Observation Station in February 2018 as an example and gives the visibility observation values and forecast values that change daily. It is not difficult to find that in each forecast period, the visibility before and after transfer learning can better reflect the change trend of visibility, but there is a

deviation between the two in the forecast value and magnitude. Table 2 shows the graded forecast score of Haikou Station in February 2018, named TS score. By comparing the RMSE and MAE of different forecast aging and the grading forecast scores of different aging, we can find that, regardless of migration, the error of short-term forecast is always smaller than the error of long-term forecast. And the TS score of short-term forecasts is always higher than the longer time-sensitive TS score. Therefore, it can be concluded that the performance of short-term forecast is better than that of long-term forecast.

Taking the 24-hour time-effect forecast as an example, the visibility forecast of the Haikou Station before and after transfer learning in February 2018 is analyzed. Figure 7 shows the forecast situation before and after transfer learning. The forecast value of TCN_TL for visibility is closer to the actual value of visibility than the forecast value of TCN. After transfer learning, both RMSE and MAE decreased significantly, RMSE decreased from 9 km to 6.2 km, and MAE decreased from 5.2 km to 2 km. As shown in Table 2, the accuracy of the grading forecast has also improved. At the 0~1 km level, the TS score increased from 0.23 to 0.35. At the 1~5 km level, the TS score increased from 0.41 to 0.5, and at the 5~10 km level, the TS score was 0.52, increased to 0.67. And the accuracy rate increased from 0.64 to 0.76 at the 10~35 km level. It is worth noting that although the TS score has increased in all levels, there is still room for improvement. Through transfer learning, the match between the predicted value of visibility and the observed value has been significantly improved.

3.3.2. Comparison and Analysis of Different Model Results. In order to better test the prediction performance of TCN_TL for visibility in offshore areas, this paper compares its experimental results with the experimental results of the other three models without transfer learning. It should be noted that the following experiments used historical visibility forecast data from 2016 to 2018 to forecast the visibility from January to April 2018 in the Qiongzhou Strait region. Figure 8 shows the errors between the predicted and observed values for the next 24 h, 48 h, 72 h, and 96 h under different models. The forecast errors of each model are given in the figure, which are the TCN_TL and CUACE models, the forecast model based on BP neural network, the forecast model based on LSTM network, and the forecast model based on TCN. Among them, the TCN_TL model uses two parts of the historical visibility observation data of the source domain and the target domain during training, while other models use the historical visibility observation data of the target domain for training.

From the data shown in Figure 8, it can be found that no matter what kind of forecast model, the longer the forecast time, the greater the error between the model's forecast and the observed value. Among them, the performance of the forecast model based on TCN_TL is significantly better than other forecast models. Figure 8(a) shows the root-mean-square error of the predicted and observed visibility of the model. Taking the 24 hr forecast period of validity as an example, the RMSE of the CUACE model is higher than other models. The LSTM network prediction model has a lower RMSE than the BP neural network, while the TCN prediction

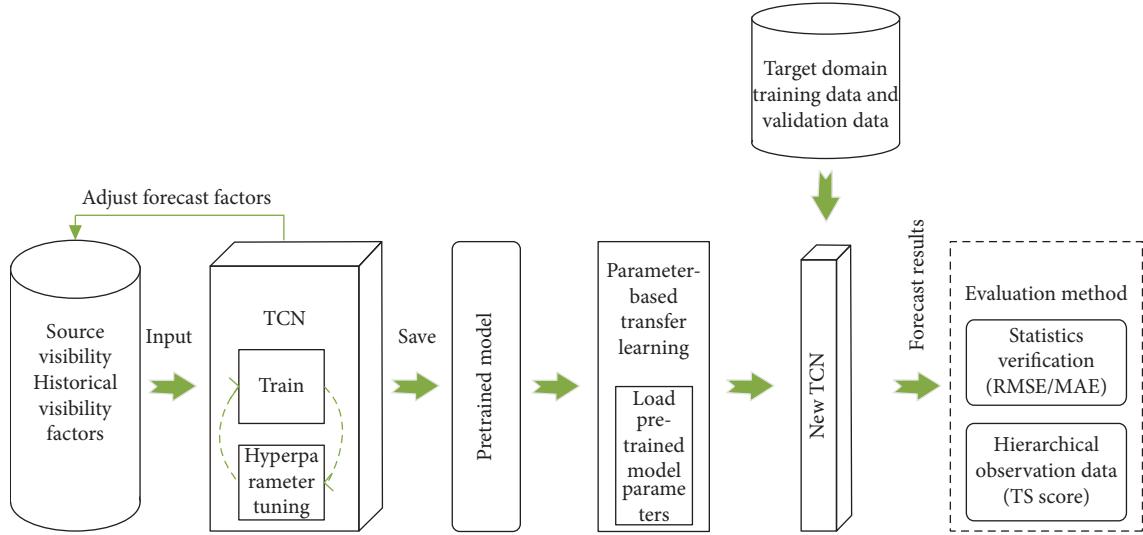


FIGURE 5: Technical process of visibility forecast based on TCN and transfer learning.

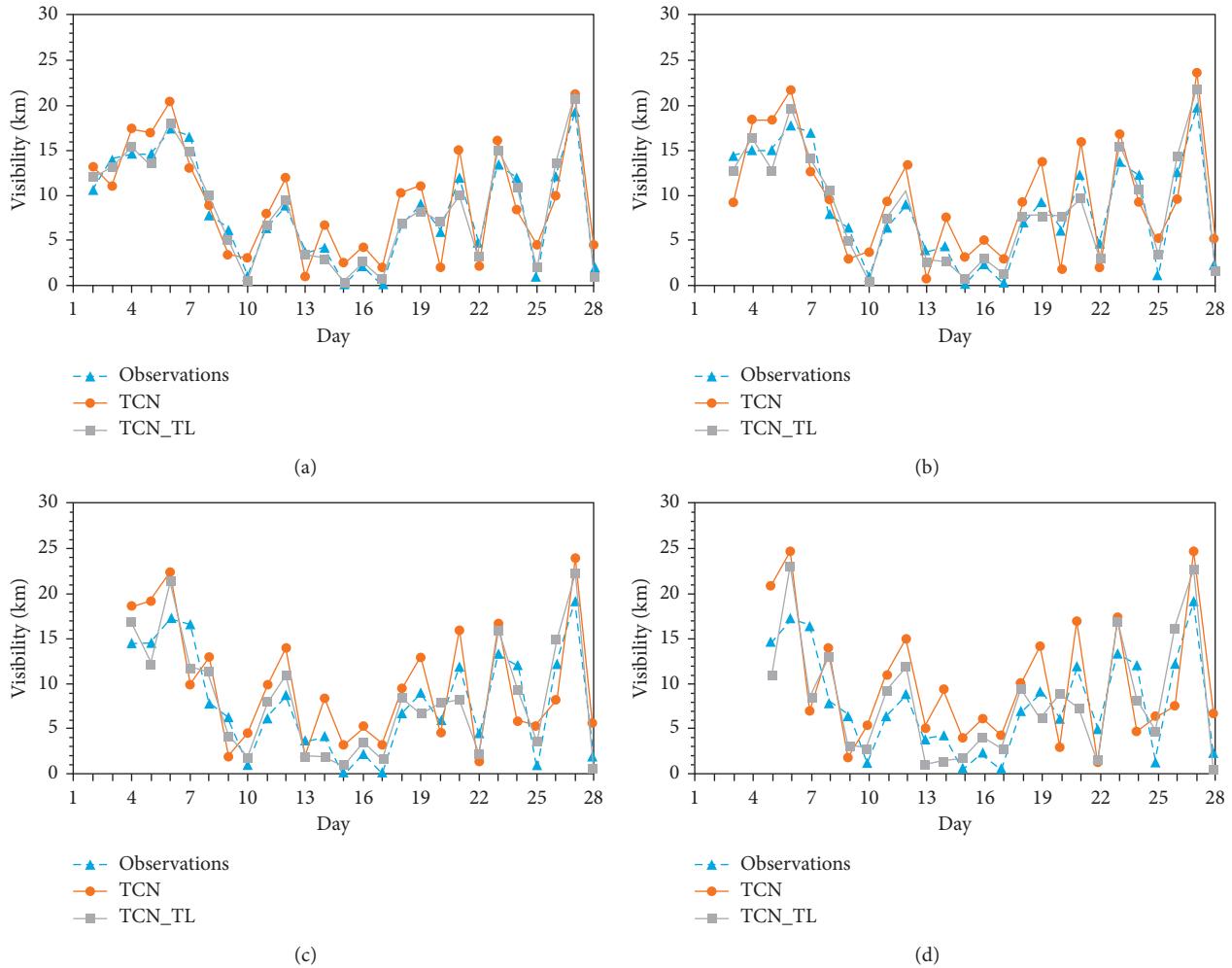


FIGURE 6: Daily changes of the visibility observation and forecast results of (a) 24 h, (b) 48 h, (c) 72 h, and (d) 96 h at the Haikou Station in February 2018.

TABLE 2: TS scores of visibility forecast in February 2018 at Haikou Station.

Classification (km)	0~24 h		24~48 h		48~72 h		72~96 h	
	TCN	TCN_TL	TCN	TCN_TL	TCN	TCN_TL	TCN	TCN_TL
[0, 1]	0.23	0.35	0.15	0.26	0.1	0.2	0.1	0.15
[1, 5]	0.41	0.5	0.36	0.54	0.3	0.45	0.23	0.31
[5, 10]	0.52	0.67	0.5	0.61	0.42	0.56	0.51	0.58
[10, 35]	0.64	0.76	0.6	0.7	0.56	0.7	0.63	0.63

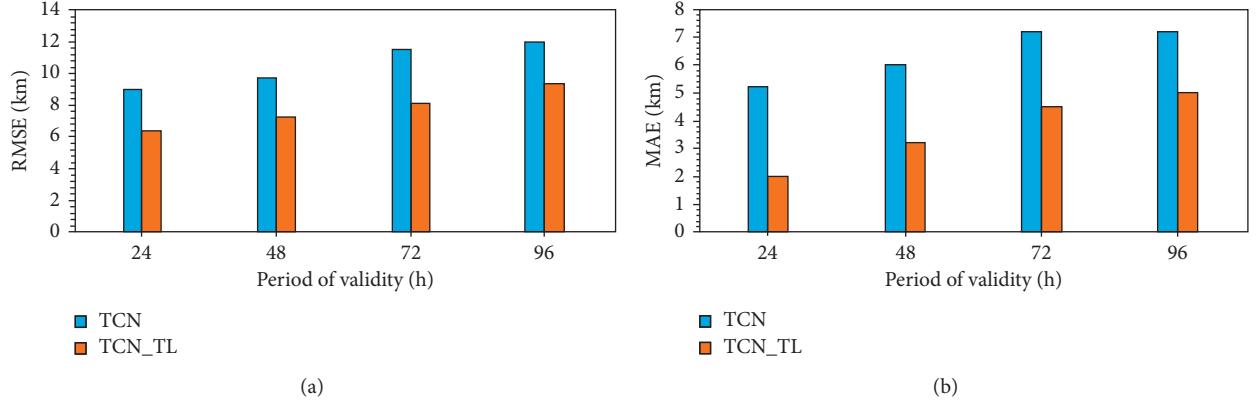


FIGURE 7: Statistical characteristics between the observed and forecasted visibility of February 2018 at Haikou Station. (a) RMSE. (b) MAE.

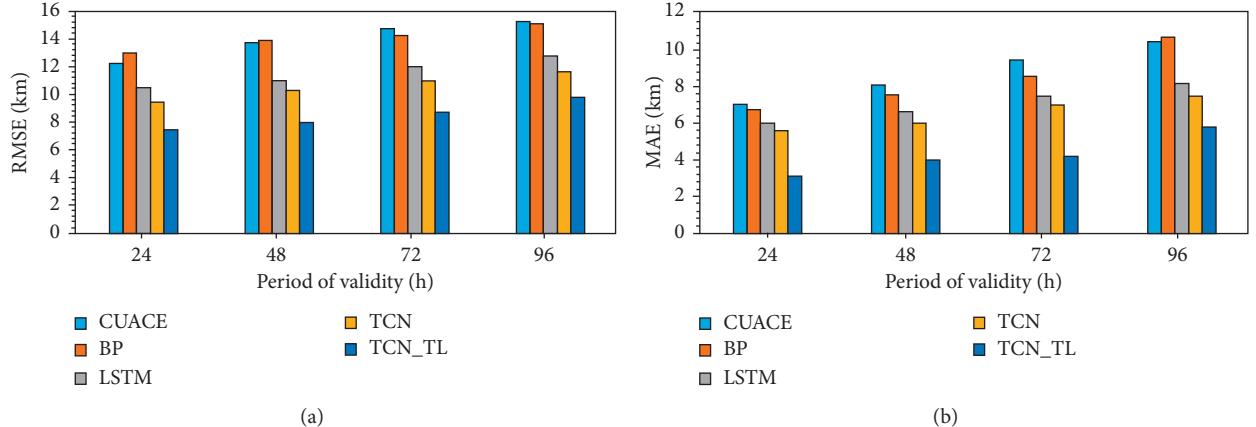


FIGURE 8: Statistical characteristics between the observed and forecasted visibility from February to April 2018 at Haikou Station. (a) RMSE. (b) MAE.

model has a slightly lower RMSE than LSTM. The RMSE of the TCN_TL prediction model using the transfer learning method is significantly lower than the RMSE of the TCN prediction model and other prediction models. The RMSE of the five forecasting models is 14.2 km, 14 km, 10.5 km, 9.4 km, and 7.4 km, respectively. Figure 8(b) shows the average absolute error between the predicted value and the observed value of the model. The MAE of the five forecasting models is 7 km, 6.7 km, 6 km, 5.6 km, and 3.1 km.

From the analysis in Figure 8, we can see that TCN_TL using the transfer learning method performs better than the other four models in both RMSE and MAE. Comparing TCN with the CUACE model, the BP neural network prediction model, and the LSTM prediction model, the performance of TCN is relatively good. From the

comparison of the errors between the TCN and TCN_TL models, the use of transfer learning to compensate for the problem of small data volume significantly reduces the error between the predicted visibility and the observed visibility and improves the prediction performance of the model.

Generally speaking, the weather with low-visibility level, especially the weather with a level of 0~1 km, appears less frequently than ordinary weather, and the opportunities for model learning are also much less. It is relatively low, but improving the accuracy of this level of forecasting has practical guiding significance. As shown in Table 3, taking the 24 hr forecast as an example, when the visibility is less than 1 km, the BP neural network prediction model is only 0.18. The TS score of TCN is slightly higher than LSTM, reaching 0.25. TCN_TL has the highest score, reaching 0.36. In

TABLE 3: 24 hr time-efficient grading forecast TS score.

Classification (km)	CUACE	BP	LSTM	TCN	TCN_TL
[0, 1]	0.14	0.18	0.21	0.25	0.36
[1, 5]	0.33	0.35	0.4	0.45	0.52
[5, 10]	0.38	0.37	0.51	0.58	0.67
[10, 35]	0.47	0.48	0.62	0.68	0.78

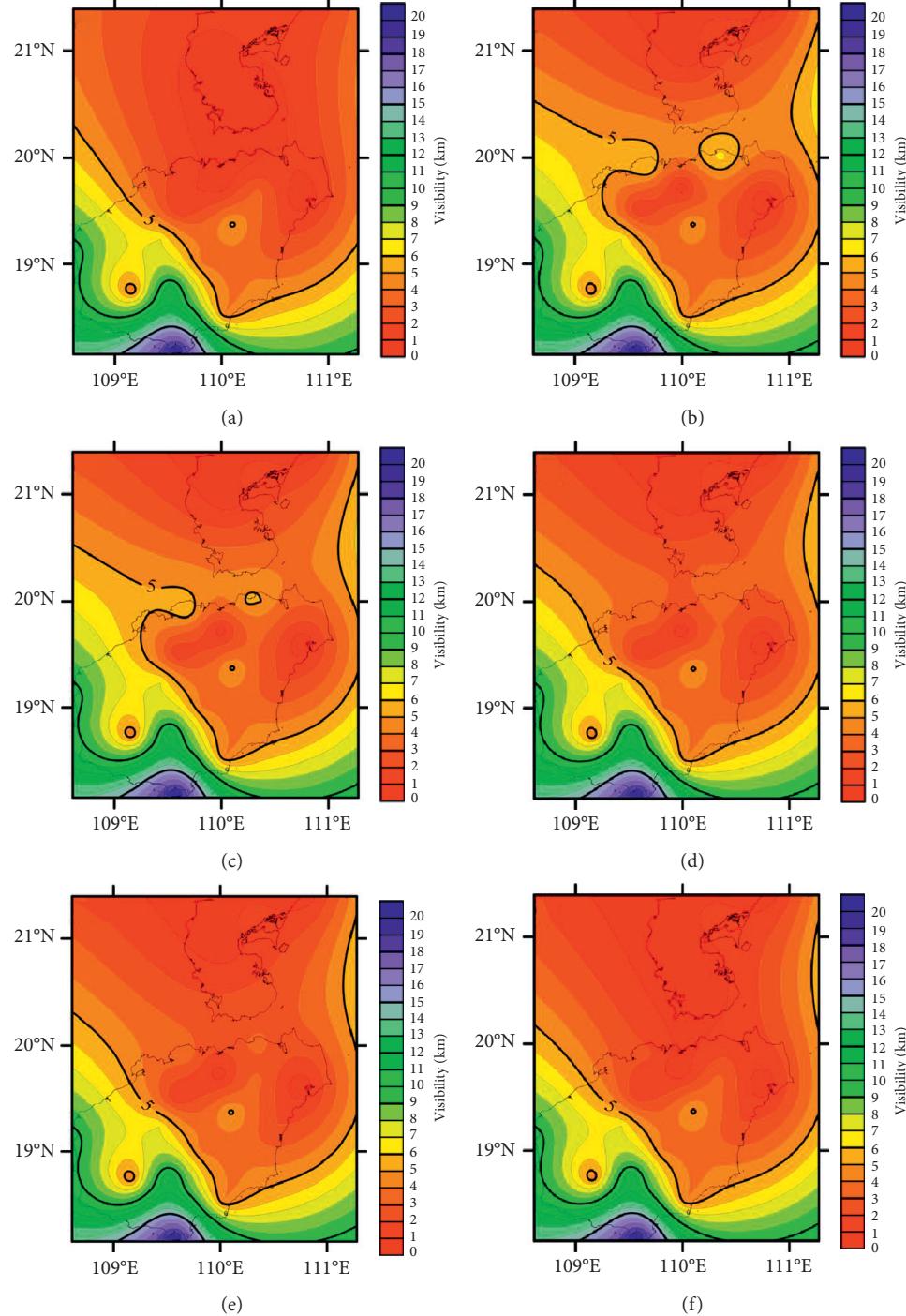


FIGURE 9: Visibility spatial distribution. (a) February 16th 8:00 measured map. (b) February 16th 8:00 CUACE model forecast map. (c) February 16th 8:00 BP model forecast map. (d) February 16th 8:00 LSTM model forecast map. (e) February 16th 8:00 TCN model forecast map. (f) February 16th 8:00 TCN_TL model forecast map.

addition, at the other three levels, the TS score of TCN_TL is also higher than the other four forecast models. Therefore, the results show that, compared with other traditional methods and machine learning methods, TCN_TL has more advantages in predicting the performance of offshore visibility.

3.4. Visibility Spatial Distribution Analysis. During the Spring Festival of 2018, the Qiongzhou Strait experienced persistent low-visibility weather. This section takes this event as an example and uses a 24 hr forecast model to perform forecast analysis. Figure 9 shows the visibility forecast of each model at 8 AM on February 16. Because the Kriging interpolation method considers the variation distribution of spatial attributes, it can effectively eliminate the errors caused by uneven sampling and make the results more in line with the actual situation [42]. Here, the Kriging interpolation method is used to interpolate the visibility spatial results of the Qiongzhou Strait, which better shows the spatial distribution of visibility. According to the spatial distribution of actual observation results of medium visibility in Figure 9(a), the Qiongzhou Strait is under low-visibility weather. Comparing the spatial distribution of visibility prediction results in Figures 9(b)–9(f), the overall spatial distribution state is gradually tending towards the actual spatial distribution state of visibility and the prediction result space of TCN_TL. The distribution is closest to the spatial distribution of actual observations. Comparing Figures 9(e) and 9(f), it can be found that the prediction result of TCN_TL is closer to the actual observation result than the prediction result of TCN. Therefore, the use of transfer learning improves the prediction accuracy of the model. Figure 9(b) is the spatial distribution of the CUACE forecast results. The forecast results are higher than the actual observation values, and the TCN_TL forecast results are more in line with the actual situation. Therefore, under the condition of small data set, the prediction performance of TCN_TL is better than that of CUACE.

4. Conclusions

The offshore visibility prediction method based on transfer learning proposed in this paper combines TCN and transfer learning. This paper uses TCN to establish a source domain forecast model and learns the knowledge of the source domain under the premise that the source domain has a large amount of data. And TCN_TL was used to forecast the visibility of the target domain offshore. The following conclusions can be obtained through experimental analysis:

- (1) This paper compares the results of TCN and TCN_TL for forecasting offshore visibility. The experimental results show that the model of transfer learning can be used to learn the visibility knowledge of the source domain under the condition of a small amount of offshore meteorological observation data

to improve the accuracy of the visibility forecast of the target domain.

- (2) In this paper, the TCN network is used to learn the source domain data. Compared with LSTM and BP neural networks, the RMSE and MAE between TCN prediction and observation are smaller, and the TS score is relatively higher than others' in each visibility forecast level. Therefore, TCN is more advantageous for learning and predicting time series like visibility data than others.
- (3) The 24-hour forecast of offshore visibility of Haikou Station from January to April 2018 was taken as an example. Under the conditions of small data sets, comparing the forecast results of TCN_TL and CUACE, the forecast error of TCN_TL is lower than that of CUACE, and the RMSE and the MAE are 6.8 km and 3.9 km. RMSE drops to 7.4 km, and MAE drops to 3.1 km. TS score of TCN_TL has also improved in each forecast level. At the level of 0~1 km, the TS score is 0.36, increased by 0.16. At the level of 1~5 km, the TS score is 0.52, increased by 0.19. At the level of 1~5 km, the TS score is 0.52, increased by 0.19. At the level of 5~10 km, the TS score is 0.67, increased by 0.29. At the level of greater than 10 km, the TS score is 0.78, increased by 0.39. Therefore, under the condition of a small data set, TCN_TL has an advantage in predicting the visibility of offshore waters than CUACE.
- (4) It is worth noting that, compared with the different aging prediction of each model, no matter which model, the prediction error of the short aging is lower than that of the longer aging. Therefore, the model method proposed in this paper is more suitable for the short aging prediction of the offshore visibility.

Data Availability

The data of this research work can be found through open data set. The ECMWF model data were obtained from the ECMWF website (<https://www.ecmwf.int/en/forecasts/datasets/browse-reanalysis-datasets>). The daily visibility data and the CUACE model data were provided by the National Meteorological Center.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

References

- [1] G. Fu, X. Li, and N. Wei, "Atmospheric visibility study," *Periodical of Ocean University of China*, vol. 39, no. 5, pp. 855–862, 2009.
- [2] A. D. Kappos, P. Bruckmann, T. Eikmann et al., "Health effects of particles in ambient air," *International Journal of Hygiene and Environmental Health*, vol. 207, no. 4, pp. 399–407, 2004.

- [3] Q. Weihong, C. L. Jeremy, C. Youli et al., "Applying anomaly-based weather analysis to the prediction of low visibility associated with the coastal fog at Ningbo-Zhoushan port in East China," *Advances in Atmospheric Sciences*, vol. 36, no. 10, pp. 1060–1077, 2019.
- [4] P. Li, S. Wang, K. Shang, B. Li, H. Zhu, and S. Zeng, "Visibility forecast in Beijing through artificial neural network based on hierarchical classification method," *Journal of Lanzhou University (Natural Sciences)*, vol. 48, no. 3, pp. 52–57, 2012.
- [5] D. Xia, Z. Wu, H. Tan, Z. Yuan, L. Chen, and P. Huang, "Analysis and correction of visibility measured by automatic observing system in Guangdong," *Meteorological Science and Technology*, vol. 42, no. 1, pp. 68–72, 2014.
- [6] F. Xia and C. Li, "Fog weather forecast experiments of Shandong province based three visibility schemes," *Journal of Meteorology and Environment*, vol. 34, no. 3, pp. 48–57, 2018.
- [7] G. A. Grell, S. E. Peckham, R. Schmitz et al., "Fully coupled “online” chemistry within the WRF model," *Atmospheric Environment*, vol. 39, no. 37, pp. 6957–6975, 2005.
- [8] K. Zhu, H. Zhang, and L. Zhang, "Multi-scale numerical modeling system of atmospheric environment and its application," *Journal of Agricultural Catastrophology*, vol. 4, no. 8, pp. 38–41, 2014.
- [9] X. Deng, T. Deng, D. Wu, D. Jiang, H. Tan, and F. Li, "The numerical forecast model system of air quality and visibility in the Pearl River Delta," *Guangdong Meteorology*, vol. 32, no. 4, pp. 18–22, 2010.
- [10] H. Liu, X. Rao, H. Zhang, M. Li, and Z. Zhang, "Comparative verification and analysis of environmental meteorology operational numerical prediction models in China," *Journal of Meteorology and Environment*, vol. 33, no. 5, pp. 17–24, 2017.
- [11] X. Chen, Y. Wang, and X. Ren, "Simulation and evaluation of haze days in Jiangsu province based on WRF/CMAQ in winter 2014," *Jounarl of Nanjing University (Natural Science)*, vol. 52, no. 6, pp. 961–976, 2016.
- [12] Y. Wang, W. Zhao, N. Xing, Z. Fu, and H. Li, "Visibility forecast based on RMAPS-CHEM products in Beijing," *Meteorology Monthly*, vol. 46, no. 3, pp. 403–411, 2020.
- [13] Z. Kang, H. Gui, C. Hua et al., "China's national environmental meteorological services and their developmental trend," *Advances in Meteorological Science and Technology*, vol. 6, no. 2, pp. 64–69, 2016.
- [14] M. Li, Z. Zhang, S. Li, X. Yu, and C. Jv, "Verification of CUACE air quality forecast in Urumqi," *Desert and Oasis Meteorology*, vol. 8, no. 5, pp. 63–68, 2014.
- [15] F. A. Gers, J. Schmidhuber, and F. Cummins, "Learning to forget: continual prediction with LSTM," in *Proceedings of the 9th International Conference on Artificial Neural Networks: ICANN*, pp. 850–855, Edinburgh, UK, September 1999.
- [16] A. T. Müller, J. A. Hiss, and G. Schneider, "Recurrent neural network model for constructive peptide design," *Journal of Chemical Information and Modeling*, vol. 58, no. 2, pp. 472–479, 2018.
- [17] L. Jiang, R. Hu, X. Wang, W. Tu, and M. Zhang, "Nonlinear prediction with deep recurrent neural networks for non-blind audio bandwidth extension," *China Communications*, vol. 15, no. 1, pp. 72–85, 2018.
- [18] J. Lei, C. Liu, and D. Jiang, "Fault diagnosis of wind turbine based on Long Short-term memory networks," *Renewable Energy*, vol. 133, pp. 422–432, 2019.
- [19] Y. A. Farha and J. Gall, "MS-TCN: multi-Stage temporal convolutional network for action segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3575–3584, Long Beach, CA, USA, June 2019.
- [20] C. Zhou, Z. Wu, and C. Liu, "A study on quality prediction for smart manufacturing based on the optimized BP-ada Boost model," in *Proceedings of the 2019 IEEE International Conference on Smart Manufacturing, Industrial and Logistics Engineering (SMILE)*, pp. 1–3, Hangzhou, China, January 2019.
- [21] R. Wang, Y. Dai, C. Han, K. Xu, and L. Dong, "Application of DPO–BP in Strength Prediction of Concrete," in *Proceedings of the IEEE 3rd Information Technology and Mechatronics Engineering Conference (ITOEC)*, pp. 1003–1006, Beijing, China, March 2017.
- [22] J. Yao, S. M. Raffuse, M. Brauer, G. J. Williamson, D. M. J. S. Bowman, and F. H. Johnston, "Predicting the minimum height of forest fire smoke within the atmosphere using machine learning and data from the CALIPSO satellite," *Remote Sensing of Environment*, vol. 206, pp. 98–106, 2018.
- [23] B. Cheng, L. Liu, Z. Qi, and H. Yang, "Prediction of continuous B-cell epitopes using long short term memory networks," in *Proceedings of the 2018 6th International Conference on Bioinformatics and Computational Biology (ICBCB)*, pp. 55–59, Chengdu, China, March 2018.
- [24] Z. Liu, K. Cheng, H. Li, G. Cao, D. Wu, and Y. Shi, "Exploring the potential relationship between indoor air quality and the concentration of airborne culturable fungi: a combined experimental and neural network modeling study," *Environmental Science and Pollution Research*, vol. 25, no. 4, pp. 3510–3517, 2018.
- [25] Y. Dai, Z. Lu, H. Zhang, T. Zhan, J. Lu, and P. Wang, "A correction method of environmental meteorological model based on long-short-term memory neural network," *Earth and Space Science*, vol. 6, no. 11, pp. 2214–2226, 2019.
- [26] X. Song, J. Huang, and D. Song, "Air quality prediction based on LSTM-Kalman model," in *Proceedings of the IEEE 8th Joint International Information Technology and Artificial Intelligence Conference (ITAIC)*, pp. 695–699, Chongqing, China, March 2019.
- [27] W. Yao, P. Huang, and Z. Jia, "Multidimensional LSTM networks to predict wind speed," in *Proceedings of the 37th Chinese Control Conference (CCC)*, pp. 7493–7497, Wuhan, China, September 2018.
- [28] X. Feng, Q. Li, Y. Zhu et al., "Artificial neural networks forecasting of PM2.5 pollution using air mass trajectory based geographic model and wavelet transformation," *Atmospheric Environment*, vol. 107, no. 4, pp. 118–128, 2015.
- [29] H. Feng, Y. Chen, Y. Cheng, and K. Luo, "Research on low visibility weather forecast method for Shuangliu Airport," *Journal of Applied Meteorology Science*, vol. 17, no. 1, pp. 94–99, 2006.
- [30] Z. Jiachen, F. Deng, Y. Cai, and J. Chen, "Long short-term memory-fully connected (LSTM-FC) neural network for PM2.5 concentration prediction," *Chemosphere*, vol. 220, pp. 486–492, 2019.
- [31] T. Li, M. Hua, and X. Wu, "A hybrid CNN-LSTM model for forecasting particulate matter (PM2.5)," *IEEE Access*, vol. 8, pp. 26933–26940, 2020.
- [32] H. Hu, H. Zhang, and Z. B. X. Chao, "Application analysis of neural network method in visibility forecast of coastal cities around Bohai Sea," *Journal of the Meteorological Sciences*, vol. 38, no. 6, pp. 798–805, 2018.
- [33] R. Gao, X. Li, Z. Ren, and J. Wang, "Research on forecast model of Qingdao coastal sea fog decision tree," *Marine Forecasts*, vol. 33, no. 4, pp. 80–87, 2016.

- [34] M. Talo, U. B. Baloglu, Ö. Yıldırım, and U. Rajendra Acharya, “Application of deep transfer learning for automated brain abnormality classification using MR images,” *Cognitive Systems Research*, vol. 54, pp. 176–188, 2019.
- [35] I. Ntinou, E. Sanchez, A. Bulat, M. Valstar, and G. Tzimiropoulos, “A transfer learning approach to heatmap regression for action unit intensity estimation,” 2004, <http://arxiv.org/abs/06657>.
- [36] C. Fan, Y. Sun, F. Xiao et al., “Statistical investigations of transfer learning-based methodology for short-term building energy predictions,” *Applied Energy*, vol. 262, Article ID 114499, 2020.
- [37] C. Raffel, N. Shazeer, A. Roberts et al., “Exploring the limits of transfer learning with a unified text-to-text transformer,” 2019, <http://arxiv.org/abs/10683>.
- [38] W. Guo, C. Xu, Q. Xiao, and M. Li, “Bn parameter learning algorithm based on dynamic weighted transfer learning,” *Application Research of Computers*, vol. 38, no. 1, pp. 1–6, 2019.
- [39] X. U. Feng, X. Tian-Zhu, W. Hui, and L. Ke-Xiu, “Comparative analysis on different data of air-sea temperature difference and analysis on variation characteristics for south China sea in the past 35 years,” *Journal of Tropical Meteorology*, vol. 23, no. 3, pp. 292–301, 2017.
- [40] S. Bai, J. Z. Kolter, and V. Koltun, “An empirical evaluation of generic convolutional and recurrent networks for sequence modeling,” 2018, <http://arxiv.org/abs/1803>.
- [41] S. J. Pan and Q. Yang, “A survey on transfer learning,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10, pp. 1345–1359, 2010.
- [42] W. Jiang, Guo, and C. Wang, “Temporal and spatial characteristics of visibility in Beijing from 2007 to 2015,” *Journal of Meteorology and Environment*, vol. 35, no. 1, pp. 45–52, 2019.