

Aerosol and Air Quality
Research

ORIGINAL RESEARCH

<https://doi.org/10.4209/aaqr.2020.05.0247>

Evolving Spiking Neural Network Model for PM_{2.5} Hourly Concentration Prediction Based on Seasonal Differences: A Case Study on Data from Beijing and Shanghai

Hengyuan Liu¹, Guibin Lu^{1*}, Yangjun Wang², Nikola Kasabov^{3,4}

¹School of Economics, Shanghai University, Shanghai 200444, China

²School of Environmental and Chemical Engineering, Shanghai University, Shanghai 200444, China

³School of Engineering, Computing and Mathematical Sciences, Auckland University of Technology, Auckland, New Zealand

⁴Intelligent Systems Research Centre, Ulster University, Londonderry, UK

ABSTRACT

In recent years, the dangers that air pollutants pose to human health and the environment have received widespread attention. Although accurately predicting the air quality is essential to managing pollution and developing control policies, traditional forecasting models have not been able to simulate the seasonal and diurnal variation in air pollutant concentrations. Furthermore, inadequate processing of the available spatio-temporal data has precluded the capture of predictive historical patterns. Therefore, we have developed a staging evolving spiking neural network (eSNN) model named *Staging-eSNN* that first employs a time series clustering algorithm to distinguish the seasonal from the diurnal variation in the PM_{2.5} concentration. We then predict the concentrations in Beijing and Shanghai 1, 3, 6, 12 and 24 hours in advance. Various evaluation indicators show that the Staging-eSNN model achieves higher performance than the support vector regression (SVR), random forest (RF) and other eSNN models.

Keywords: Air pollutant prediction, PM_{2.5} hourly concentration, Seasonality, Evolving spiking neural networks, Time series clustering

1 INTRODUCTION

Since the reform and opening up, China has made steady progress in the process of urbanization. The urbanization rate has increased from 17.92% before the reform and opening up to 58.52% in 2017. Beijing and Shanghai, as two key cities to support China's sustained economic growth, are also experiencing frequent effects of smog while developing rapidly. Smog is usually characterized by high PM_{2.5} concentration. According to Liu *et al.* (2019), there is a statistically significant correlation between short-term exposure to PM₁₀, PM_{2.5} and cardiovascular, respiratory mortality. In addition, PM_{2.5} has a negative impact on socioeconomic and climate change (Li *et al.*, 2016).

To reduce smog and improve air quality, researchers have proposed many methods to analyze and predict the concentration of air pollutants such as PM_{2.5}. Among the three commonly used methods, one is the deterministic method which is based on aerodynamic theory and physico-chemical processes. It establishes a numerical model of air pollution concentration diffusion, further predicts the dynamic changes of atmospheric pollutant concentration through high-speed calculation and simulation. The commonly used models are the the Community Multi-scale Air Quality (CMAQ) model (Chen *et al.*, 2014) and the WRF-Chem model (Saide *et al.*, 2011). However, in these models, various types of parameters need to be determined by experience,

OPEN ACCESS 

Received: May 19, 2020

Revised: August 11, 2020

Accepted: August 25, 2020

* Corresponding Author:

guibinlu@163.com

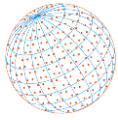
Publisher:

Taiwan Association for Aerosol
Research

ISSN: 1680-8584 print

ISSN: 2071-1409 online

 **Copyright:** The Author(s).
This is an open access article
distributed under the terms of the
[Creative Commons Attribution
License \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/), which permits
unrestricted use, distribution, and
reproduction in any medium,
provided the original author and
source are cited.



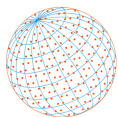
which results in limited accuracy (Xu *et al.*, 2017). And informative source data are required to be used in the models, which are difficult to obtain in practice (Stern *et al.*, 2008). The second method is the traditional time series analysis model represented by ARIMA (Auto Regressive Integrated Moving Average), which uses univariate time series modeling without the need for other feature variables (Jian *et al.*, 2012; Wang *et al.*, 2017; Zhang *et al.*, 2018), so it is common in the field of predicting air pollutants. However, the ARIMA model only uses univariate time information, lacks consideration of the influence of other variables, and has high requirements for the stability of the series.

Using machine learning to predict air pollutants reduces the requirements for good statistical properties (stationarity, normality) of the data, which is the third commonly used method in the field of air quality prediction, including support vector regression (SVR), random forests (RFs), etc. Sun and Sun (2017) used SVR to predict PM_{2.5} concentration based on principal component analysis (PCA) dimensionality reduction technology, which improved the accuracy of the single model of SVR. Siwek and Osowski (2012) discussed the relationship between meteorological parameters and pollutant concentration, combined with wavelet decomposition, integrated multiple neural predictors to predict the daily concentration of PM₁₀ in Warsaw. However, the spatio-temporal data composed of multivariate time series in air quality prediction often have autocorrelation, which means that the air quality of a certain period is affected by previous periods and is not independent in the time dimension. Therefore, these machine learning methods, which mainly focus on static data, lose time information during the prediction process. To deal with this kind of spatio-temporal data, several neural network models and their hybrid model have been applied to air quality prediction and achieved good accuracy, such as convolutional neural networks (CNNs), recurrent neural networks (RNNs) and long short-term memory (LSTM) networks (Maciąg *et al.*, 2019; Wen *et al.*, 2019; Zhai and Cheng, 2020). These models are based on multiple hidden layers with several neurons, using multi-feature time series as input for training, and learning spatio-temporal information from them. As another type of neural network, a spiking neural network (SNN) model was developed, called *NeuCube*, whose neurons can adjust whether they are active or inactive according to whether there is information transmission. The feature is inspired by some principles about how the human brain processes spatio-temporal information and therefore more biological (Kasabov, 2014). At present, SNN have been applied in fields such as market finance (Reid *et al.*, 2014) and ecological environment (Hartono *et al.*, 2014; Maciąg *et al.*, 2019). For example, Maciąg *et al.* (2019) used air pollutants and meteorological parameters as feature variables, and developed an integrated clustering-based ensemble model, CEeSNN, to predict the PM₁₀ and O₃ concentration in the United Kingdom, which improved the prediction accuracy of the classical *NeuCube* model.

Although these studies have achieved good performance in prediction accuracy, the issue of combining variation characteristics of PM_{2.5} has not been addressed properly. These characteristics include two aspects. On the one hand, there is a seasonal difference in PM_{2.5} concentration variations. It is generally believed that the concentration of PM_{2.5} is higher in the spring and winter seasons and the seasonal variations in different regions are different, which is reflected in inter-city difference and urban-rural difference (Feng *et al.*, 2013; Ma *et al.*, 2019; Wang *et al.*, 2019; Shen *et al.*, 2020). On the other hand, the diurnal variations of PM_{2.5} concentration is unimodal or bimodal (Zhao *et al.*, 2009; Carslaw and Beevers, 2013; Elangasinghe *et al.*, 2014; Kim and Kim, 2020). The combination of these two characteristics is called the *diurnal variation seasonality* of PM_{2.5} concentration in the paper. Specifically, for a certain city, there are different diurnal variation structures of PM_{2.5} concentration in different months (seasons). For different cities, due to differences in geographical location and climate, there is a diurnal variation difference in PM_{2.5} concentration in the same month (season).

In this context, in order to improve the accuracy of PM_{2.5} hourly concentration prediction, this paper proposes a staging evolving spiking neural network (eSNN) model, Staging-eSNN, and a whole day's pollution and meteorological data is used to predict PM_{2.5} concentration in the next 1, 3, 6, 12 and 24 hours in two cities (Beijing and Shanghai).

The structure of this paper is as follows. In the second section, we discuss the seasonal variation difference and diurnal variation characteristics of PM_{2.5} concentration based on the data of Beijing and Shanghai. The third section introduces the Staging-eSNN model. The fourth section evaluates the prediction performance of the SVR, RF, classical *NeuCube* eSNN (Plain-



eSNN), CEeSNN and Staging-eSNN model. The conclusion part summarizes the work and points to some directions for further research.

2 METHODS

2.1 Study Regions

Beijing is located in $115^{\circ}42'–117^{\circ}24'E$ and $39^{\circ}24'–41^{\circ}36'N$ which is in the northern part of the North China Plain and adjacent to Bohai Bay. It is China's economic decision center and political center. Beijing has a typical warm temperate semi-humid continental monsoon climate. Shanghai is located in $120^{\circ}52'–122^{\circ}12'E$, and $30^{\circ}40'–31^{\circ}53'N$, which is at the mouth of the Yangtze River. It is an important city in the Yangtze River Delta region of China. Due to the influence of the East Asian monsoon, Shanghai is a subtropical monsoon climate. Fig. 1(a) shows the geographic locations of the two cities.

2.2 Data Collection

This paper uses the hourly historical data of $PM_{2.5}$, NO_2 , and SO_2 air pollutants from January 1, 2017, to August 21, 2019 (for convenience of description, hereinafter referred to as “2017–2019”). These data were monitored by the national air quality monitoring stations in Beijing (36) and Shanghai (10), which can be downloaded through the China Urban Air Quality Real-Time Publishing Platform (<http://106.37.208.233:20035/>). The meteorological data is from NOAA's National Centers for Environmental Information (NCEI; <https://www.ncdc.noaa.gov/>), which provides half-hour data for Shanghai Pudong, Hongqiao airports and Beijing Capital International Airport (Figs. 1(b) and 1(c) show the locations of air quality monitoring stations and airports in Beijing and Shanghai respectively). This paper selects wind speed (WS; $m\ s^{-1}$), wind direction (WD; $^{\circ}$), and temperature (TEMP; $^{\circ}C$) from 2017 to 2019. For the three meteorological variables,

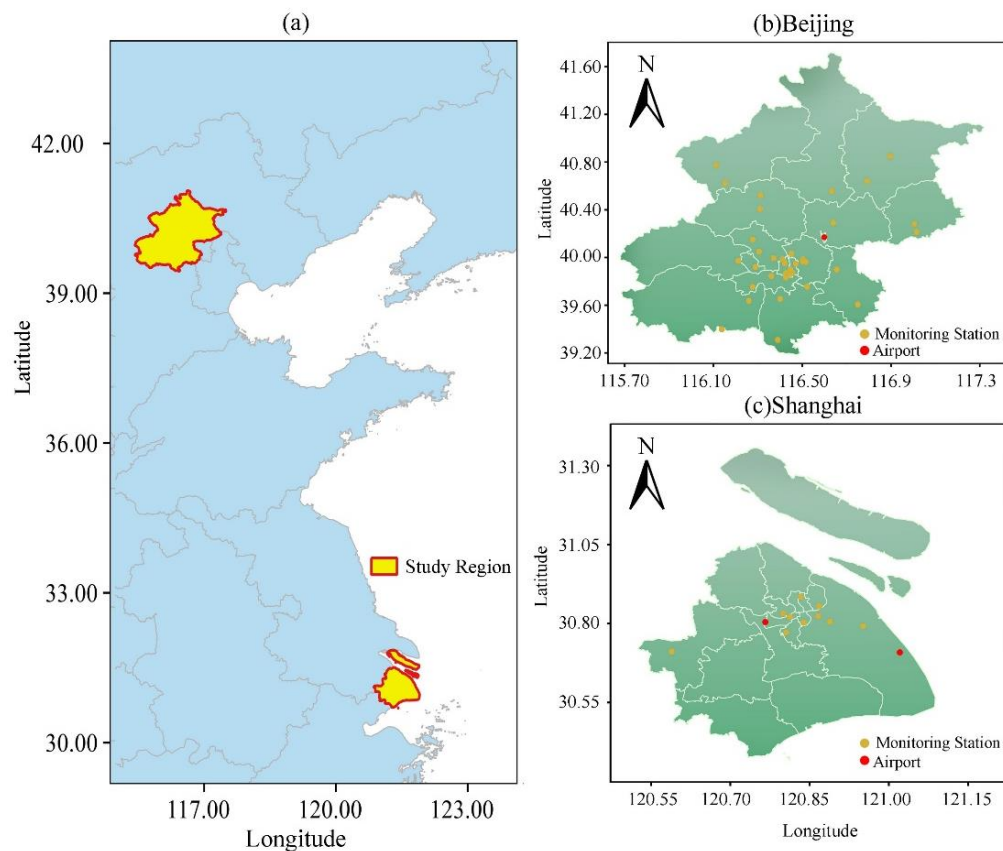
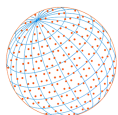


Fig. 1. (a) Geographical location of the study regions. National air quality monitoring stations and airports in (b) Beijing and (c) Shanghai.



if one half-hour data is missing, the other half-hour data is filled. The missing data for the entire hour is considered to be missing. The hourly data is finally obtained by averaging the half-hour data. For missing values in air pollutant and meteorological hourly data, the same treatment method is adopted, that is, if there are more than 12 missing values of any variable on a certain day, the day will be eliminated, and the missing values in the remaining days will be filled with the mean of the day. The final dataset of Beijing is obtained after averaging the pollution data of all air quality monitoring stations and the meteorological data of all airports, which contains hourly historical data of PM_{2.5}, NO₂, SO₂, TEMP, WS, and WD from 2017 to 2019 (Shanghai final dataset is obtained in the same way). Table 1 shows descriptive statistics of air pollutants and meteorological parameters in Beijing and Shanghai, and Fig. 2 shows their distribution.

2.3 Diurnal and Seasonal Variations of PM_{2.5} Concentration

Fig. 3 shows the trends of PM_{2.5} concentration in Beijing and Shanghai at different time scales from 2017 to 2019. The shading indicates a 95% confidence interval in the mean. From Fig. 3(a), it can be found that the diurnal variation of PM_{2.5} concentration in the two cities shows a bimodal change and the peak shape of Beijing is weaker. The specific performance is that the PM_{2.5} concentration on different days in a week will peak at around 09:00 and 20:00, which is the characteristic of the morning and evening peak due to the traffic emission in two cities. Fig. 3(b) shows the same data averaged over each hour, which further confirms the bimodal characteristics of diurnal variations of PM_{2.5} concentration in the two cities. Another characteristic of PM_{2.5} is reflected in the monthly trend and the concentration value in different months. It can be seen from Fig. 3(c) that the downward and upward trends of PM_{2.5} concentration are shown before and after August respectively. From the difference of the concentration values, the U-shaped structure of PM_{2.5} concentration reflects the seasonal difference, that is, the concentration during the cold seasons (spring and winter) will be greater than the warm seasons (summer and autumn).

Based on the above analysis, we believe that there are seasonal differences in the diurnal variations of PM_{2.5} concentration. The characteristics is called the *diurnal variation seasonality* of PM_{2.5} concentration, which can be used to divide the data into different periods (seasonality) for PM_{2.5} prediction. However, due to the differences in urban locations, pollution sources, natural climate and other factors, the seasonality of PM_{2.5} divided by the four seasons only reflects the

Table 1. Descriptive statistics of selected variables from 2017 to 2019 in Beijing and Shanghai.

	PM _{2.5} (μg m ⁻³)	NO ₂ (μg m ⁻³)	SO ₂ (μg m ⁻³)	TEMP (°C)	WS (s ⁻¹)	WD (°)
<i>Beijing</i>						
Mean	50.85	39.07	6.11	13.6	2.82	183.89
Std	52.78	24.1	7.36	12.43	2.1	92.44
Min	2	2	1	-15	0	10
Max	607	182	136	40.28	16.54	360
<i>Shanghai</i>						
Mean	37.58	41.24	9.99	17.71	4.5	174.95
Std	27.82	24.14	6.92	9.18	1.91	85.35
Min	5.4	1	1	-4.44	0	10
Max	235	184	89	39.44	15.42	360

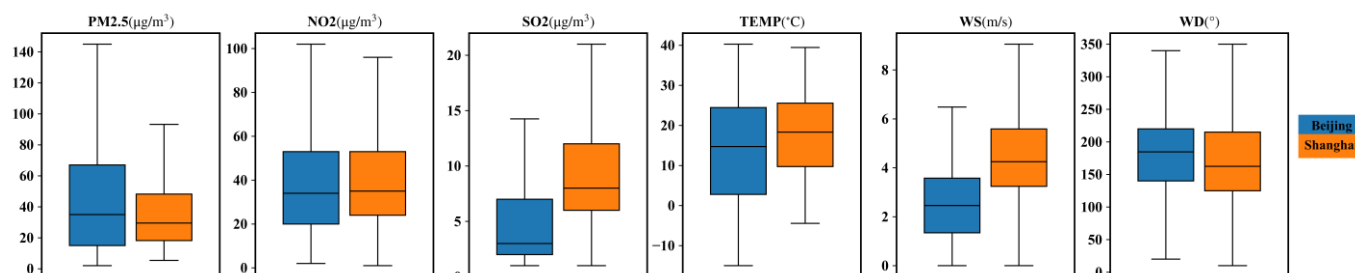


Fig. 2. The box plot for selected variables in Beijing and Shanghai from 2017 to 2019.

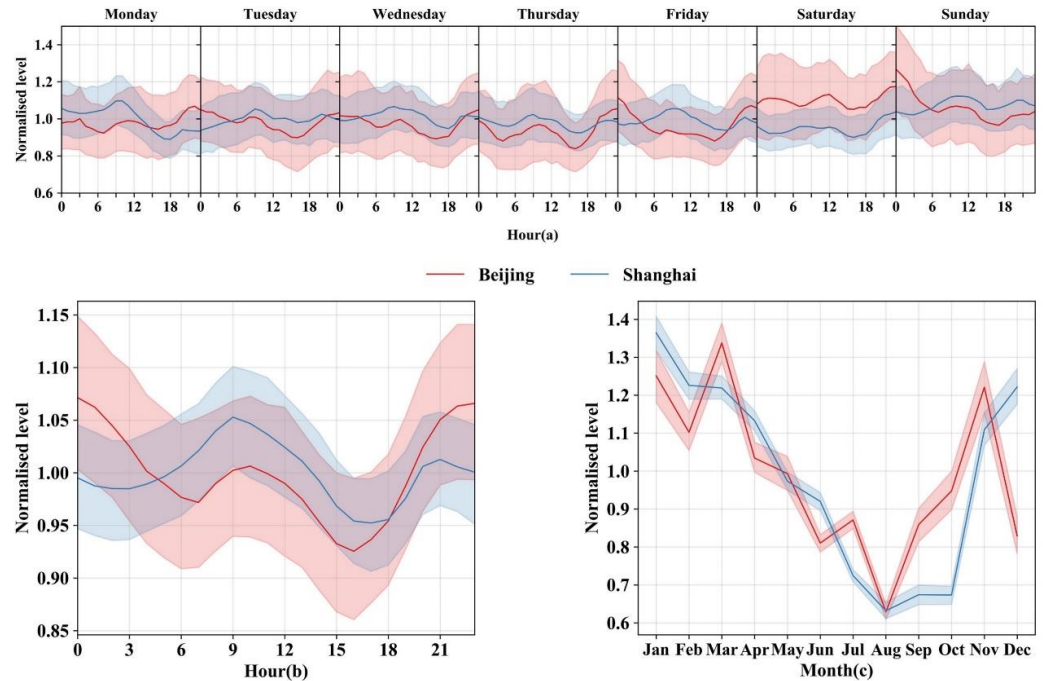
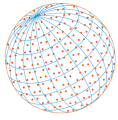


Fig. 3. PM_{2.5} concentration trends in Beijing and Shanghai at different time scales from 2017 to 2019 averaged for (a) each weekday and hour, (b) hour and (c) month.

difference in the climate. Therefore, a time series clustering technology is used to re-evaluate and re-define the seasonality in different cities, which to the extent possible control the impact of multiple factors (see the subsection “Clustering results”). Eventually, the diurnal variation of PM_{2.5} in the same period (called “hot”, “warm” or “cold” period) only reflects one variation pattern, which is driven by a potentially stable structure.

3 THEORY AND MODEL IMPLEMENTATION

A unique characteristic of spiking neural networks is that they learn temporal or spatio-temporal patterns in their connectionist structures that can be used to predict future temporal events.

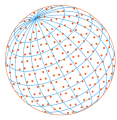
The first two parts of this section introduce SNN and the details of implementing eSNN in a known SNN architecture, NeuCube. The last part introduces our proposed Staging-eSNN model, which combines a time series clustering algorithm and eSNN models.

3.1 Spiking Neural Networks

In spiking neural networks, spatial and temporal information can be encoded as locations of synapses and neurons, and time of their spiking activity respectively (Kasabov, 2014; Kasabov, 2019). Temporal information is processed and transmitted by synaptic neurons to form memories. During the transmission process, synaptic connections are modified to more accurately reflect the correlation between different series (Tavanaei *et al.*, 2019).

One class of SNN, the evolving SNN, can learn incrementally new data and evolve new output neurons to capture different patterns from input data in an adaptive way (Kasabov, 2019). And that is why eSNN are selected as the basic model in this paper.

In recently published CEeSNN model, Maciąg *et al.* (2019) clustered all training samples for the eSNN model training, which lead to the fact that although each eSNN model receives pollutant series with similar variations, the variation’s causes of these series are different. For example, although the pollutant’s concentration variations on a day in January and a day in August are similar, they are affected obviously by different seasonality, which are reflected in the difference of pollution sources, natural climate and other factors. This means that the data processed by the CEeSNN model may be driven by multiple different patterns, and therefore it can be improved.



3.2 Implementing eSNN Model in NeuCube

NeuCube is a modular development system that provides a framework to build an eSNN model for data mining and prediction, especially spatio-temporal data (Kasabov, 2014). For the sake of brevity, the model built in the NeuCube hereafter is called “eSNN.” NeuCube contains four modules, as shown in Fig. 4.

Input module: This paper uses step-forward encoding to encode spatio-temporal data into binary temporal events (Kasabov, 2019).

- 1). 3D Cube module: In this paper, 1006 leaky integrate-and-fire (LIF) neurons are initialized in the cube and 6 input neurons (corresponding to 6 feature variables) among them are mapped into the corresponding spatial positions. Since there is no available neuron position information for feature variables in predicting PM_{2.5} problems, vectorization principle is used to measure the temporal similarity between variables to map variables with similar patterns to closer neurons in the cube (Tu *et al.*, 2017). The neurons in the cube are connected by inhibitory or excitatory synapses and the initial connection weight is set according to the small world principle (Kasabov *et al.*, 2016). After the initial cube is established, the encoded spatio-temporal data is input into the cube, and the neurons learn the spike sequences in an unsupervised mode through spike-timing-dependent plasticity (STDP) learning rule to capture spatio-temporal patterns in the neuronal connections. During the learning process, synaptic connection weights are modified to form the final SNN cube. In this module, the learned connectivity patterns in the SNN cube can be interpreted as deep knowledge representing deep spatio-temporal patterns in the data (Kasabov 2019).
- 2). Output module: The output module is a classifier/predictor trained in a supervised mode. The learned connectivity patterns can be interpreted for rule extraction (Kasabov, 2019).
- 3). Parameter optimization module: The prediction performance of the eSNN model can be easily improved by changing a large number of parameters. Different encoding parameters will significantly change the information density of spike sequences. Different “mod” and “drift” parameters will lead to different classification accuracy in the eSNN (Kasabov *et al.*, 2016). We use the grid search method to execute the parameter optimization module.

3.3 The Proposed Staging-eSNN Model

The input of the Staging-eSNN model is spatio-temporal data composed of 24-hour air pollutants and meteorological data in a day. The target value is PM_{2.5} concentration whose time point is 1, 3, 6, 12 or 24 hours after the input spatio-temporal sample. Fig. 5 shows the data structure. Compared with the sequence length of 12 in the CEeSNN model, the way of setting the sequence length to 24 provides a complete diurnal variation cycle of PM_{2.5} concentration for the Staging-eSNN model. Furthermore, shorter samples contain lower information density, making

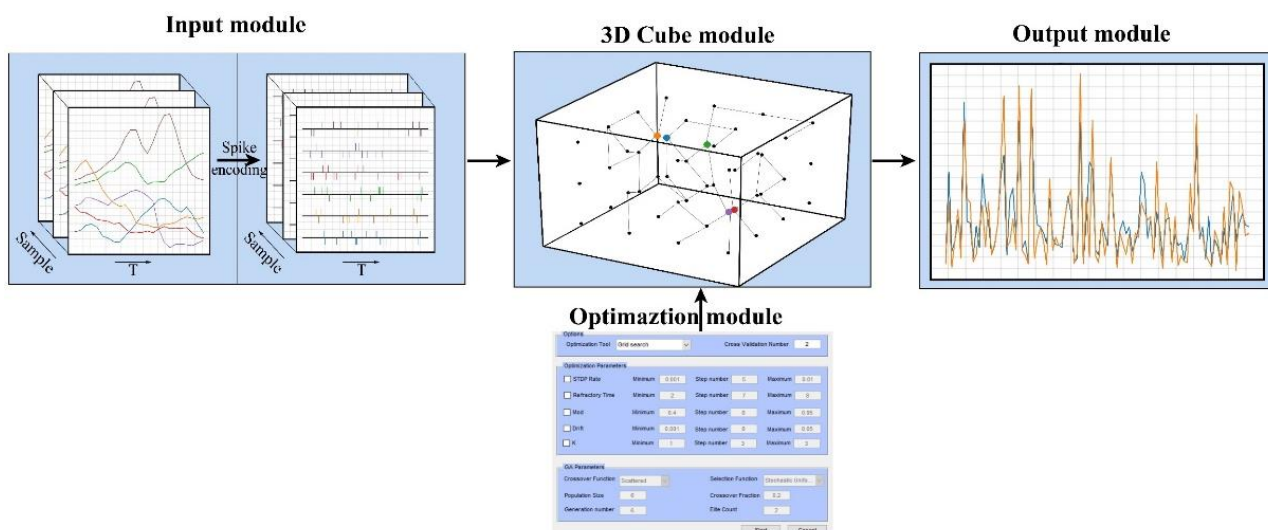
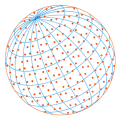


Fig. 4. Modules for implementing an eSNN model in NeuCube.



Date	PM2.5	NO2	SO2	TEMP	WS	WD
2019/1/25 0:00	57.8	75	6	8.19	2.68	173.42
...
2019/1/25 23:00	20.3	30	7	2.64	7.26	272.5
2019/1/26 0:00	15.8	-	-	-	-	-
2019/1/26 2:00	14.6	-	-	-	-	-
2019/1/26 5:00	17	-	-	-	-	-
2019/1/26 11:00	25.3	-	-	-	-	-
2019/1/26 23:00	24.4	-	-	-	-	-

The table is divided into two sections by brackets on the right. The first section, labeled '24 hours a day', includes the first three rows (2019/1/25 0:00 to 2019/1/25 23:00). The second section, labeled 'Target value', includes the last five rows (2019/1/26 0:00 to 2019/1/26 23:00).

Fig. 5. The data structure of an exemplar sample for the Staging-eSNN model.

it difficult for encoded binary events to activate neurons, thereby weakening the connectivity of neurons in the eSNN model.

Fig. 6 shows the framework of the Staging-eSNN model. The training process is divided into the following two steps which reveals the technical details of the proposed Staging-eSNN model.

- 1). The time series clustering algorithm is applied first to distinguish the seasonal differences in diurnal variations of PM_{2.5} concentration (see the subsection “Clustering results”). Specifically, for the time series of PM_{2.5} diurnal trends in 12 months, dynamic time warping (DTW) is applied to calculate the distance between them. A smaller DTW distance means that the two series are more aligned and similar, and have similar seasonal variations. On this basis, the clustering algorithm is used to cluster the monthly diurnal variation series of PM_{2.5} concentration to distinguish the changing structure in different periods. During the process, partitioning around medoids (PAM) is used as an update function of the centroid to reduce the possibility of failure to converge (Petitjean *et al.*, 2011). Eventually, in each cluster class, PM_{2.5} diurnal variations have the similar seasonality, which means that they share a common variation pattern.
- 2). A multi-model eSNN system is created that consists of several eSNN models, one for each seasonal period defined in Step 1. Based on the clustering results, samples from different periods are entered as input samples into the corresponding eSNN models implemented in NeuCube, and the prediction results of all eSNN models are finally integrated.

4 RESULTS AND DISCUSSION

4.1 Clustering Results

To observe the characteristic of the diurnal variation seasonality, we calculate the diurnal variation concentration of PM_{2.5} in each month from 2017 to 2019 to obtain 12 time series in Beijing and Shanghai respectively. Then a time series clustering algorithm (PAM-DTW) is further used to classify the seasonality. The number of candidate clusters is 2, 3, and 4. In this paper, silhouette index (Sil), Calinski–Harabasz (CH), and Davies–Bouldin index (DB) are selected as the basis for evaluating the clustering effect. A good clustering result corresponds to low DB value, high CH and Sil value (Arbelaitz *et al.*, 2013). Table 2 shows the evaluation results of clustering, indicating that Beijing’s seasonality can be divided into 2 categories, and Shanghai can be divided into 3 categories.

Based on the clustering results of PM_{2.5} concentration in Beijing and Shanghai shown in Figs. 7(a) and 7(b), we divide the seasonality of PM_{2.5} concentration diurnal variations in Beijing into the cold and warm period, of which the cold period (January–April, November) is mainly concentrated in spring and winter, while the warm period (May–October, December) is concentrated in summer and autumn, the shape is roughly inverted-V-shaped. There are obvious hierarchical structures between the clusters and similar trends within the clusters in Shanghai. We refer to the three periods of PM_{2.5} concentration variations in Shanghai as the “cold,” “warm,”

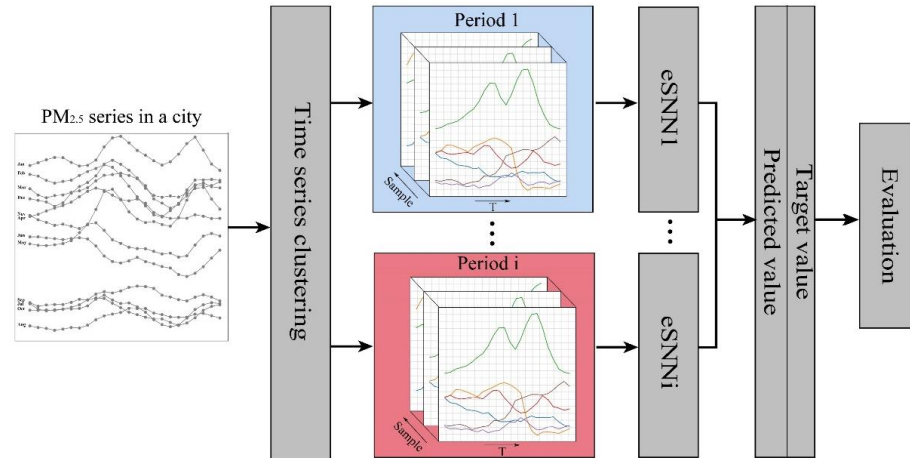
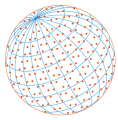


Fig. 6. The proposed Staging-eSNN model framework.

Table 2. Evaluation results of time series clustering algorithm (number of Clusters 2, 3, 4).

	k_2	k_3	k_4
<i>Beijing</i>			
Sil	0.61	0.47	0.3
CH	14.94	7.65	6.29
DB	0.49	0.83	1.31
<i>Shanghai</i>			
Sil	0.74	0.74	0.52
CH	25.75	32.05	20.13
DB	0.23	0.22	0.43

Note that good results are highlighted in bold.

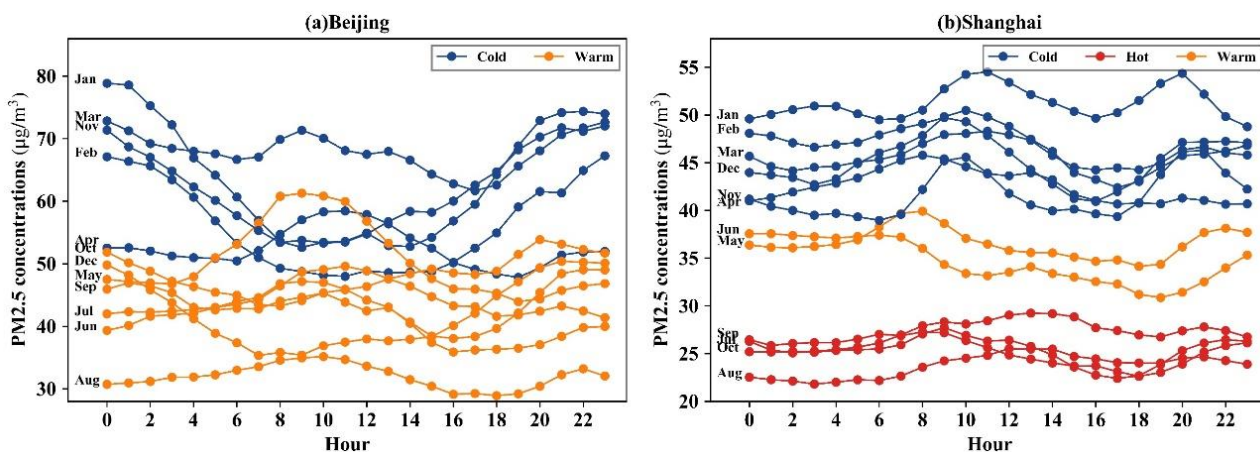
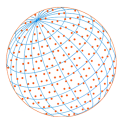


Fig. 7. Clustering results of PM_{2.5} concentration's diurnal variation and seasonality in (a) Beijing and (b) Shanghai from 2017 to 2019.

and “hot” period. The cold period (January–April, November–December) is mainly concentrated in spring and winter. The warm period (May–June) is the turn of spring and summer. The hot period (July–October) is concentrated in summer and autumn. In addition, the descriptive statistics of PM_{2.5} concentration in different periods are reported in Table 3.

4.2 Model Prediction Performance Evaluation

Based on the clustering results, we first randomly selected 150 and 100 samples from each period in Beijing and Shanghai, respectively. The samples from both cities contained 7200 hours



of air pollutants and meteorological data. Then, the eSNN model was implemented by NeuCube software to predict PM_{2.5} concentration in the next 1, 3, 6, 12 and 24 hours. Two-fold cross-validation was applied to model training and prediction. Finally, the predict results of each period are summarized in the two cities.

4.2.1 Definition of Predictive Performance Evaluation Indicators

The following indicators are used to evaluate the prediction results of models: the bias, root mean square error (RMSE), root mean square percentage error (RMSPE), index of agreement (IOA), Pearson correlation coefficient (*r*), fraction of predictions within a factor of two (FAC2). See Table 4 for definition and description.

4.2.2 The Staging-eSNN model predictive performance evaluation

Fig. 8 shows the evaluation results of the Staging-eSNN model for two cities in different periods. For the Shanghai region, the bias of the Staging-eSNN model is higher than 0 in the most predicted moments, which means that the predicted values are more likely to be overestimated than the target values. RMSE is lower than 15 μg m⁻³ in the hot period while it is above 18 μg m⁻³ in the cold period. However, due to the differences in samples from different periods, RMSE cannot effectively compare model performance in different periods. RMSPE measures the magnitude of the predicted deviation from the target mean, which does not show an obvious difference between the three periods. FAC2 reveals that more than 84% of the Staging-eSNN

Table 3. Descriptive statistics of PM_{2.5} in different periods.

Period	Mean	Std	Min	Max
<i>Beijing</i>				
Cold	60.50	64.83	2.00	607.00
Warm	43.05	38.81	2.00	570.00
<i>Shanghai</i>				
Cold	45.55	31.95	6.00	235.00
Warm	34.90	20.97	7.45	145.80
Hot	25.28	16.97	5.40	167.90

Table 4. Model prediction performance evaluation indicators (*O* and *M* represent target value and predicted value, respectively).

Definition	Description
$Bias = \frac{1}{N} \sum (M_i - O_i)$	Measure whether the predicted value is higher than the target value, where a negative value indicates underestimation.
$RMSE = \sqrt{\frac{1}{N} \sum (M_i - O_i)^2}$	Measure the difference between the predicted and target values.
$RMSPE = \frac{\sqrt{N \sum (M_i - O_i)^2}}{\sum O_i} \times 100$	Measure the size of RMSE relative to the target mean and can be used to compare different datasets (Hyndman and Koehler, 2006).
$IOA = 1 - \frac{\sum (M_i - O_i)^2}{\sum (M_i - \bar{O}_i + O_i - \bar{O}_i)^2}$	Measure the correlation between the predicted and target values. The ideal model is 1.
$r = \frac{\sum (M_i - \bar{M}_i)(O_i - \bar{O}_i)}{\sqrt{\sum (M_i - \bar{M}_i)^2 \sum (O_i - \bar{O}_i)^2}}$	Measure the correlation between the predicted and target values. The ideal model is 1.
FAC2 refers to the proportion of predicted values that satisfy $0.5 \leq \frac{M_i}{O_i} \leq 2.0$.	Measure the extent to which the predicted value deviates from the target value within a certain range. The ideal model is 1.

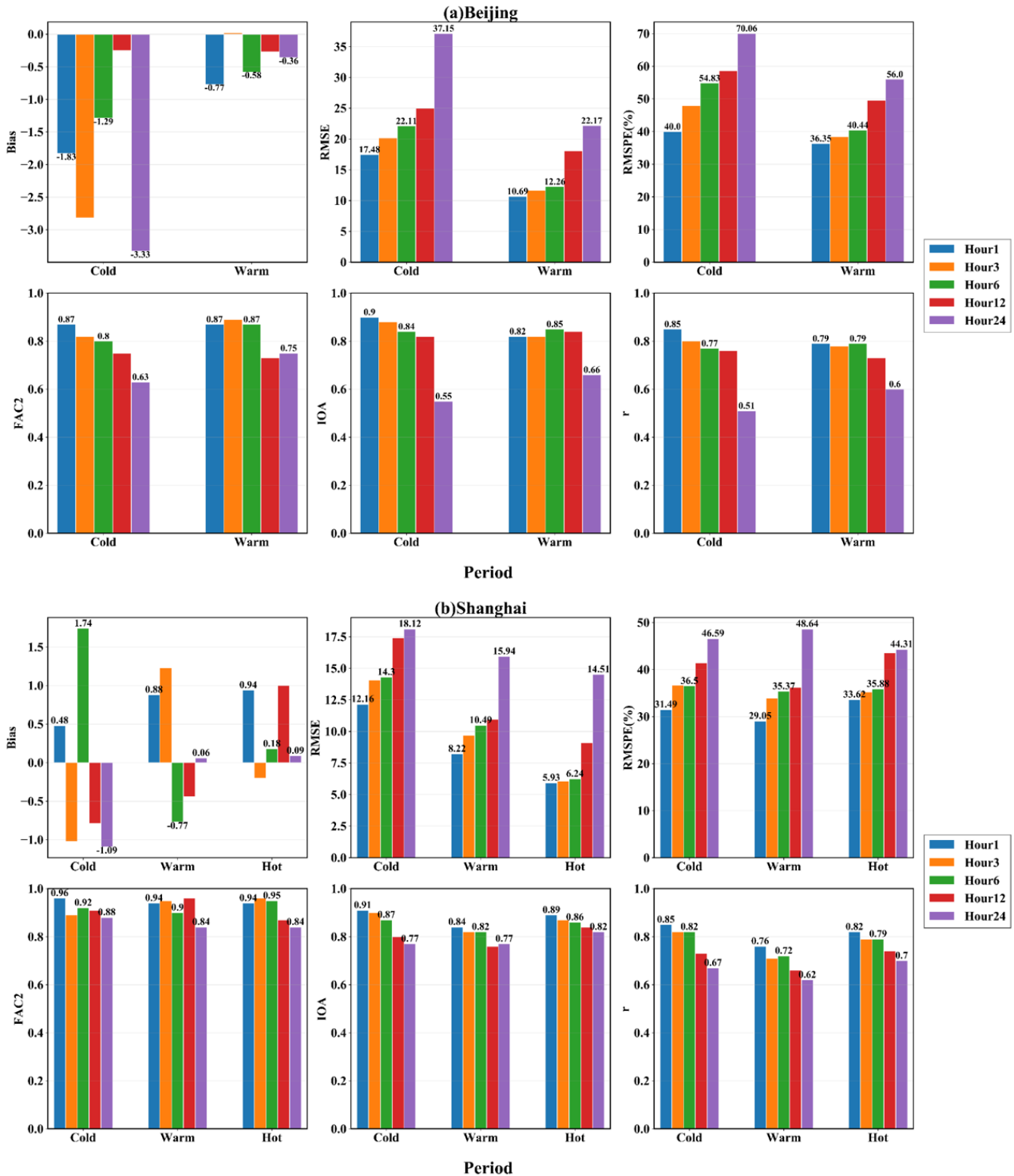
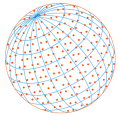
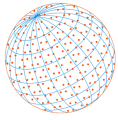


Fig. 8. Evaluation indicators of the Staging-eSNN model in any period in (a) Beijing and (b) Shanghai.

model predictions are within a factor of two of the target PM_{2.5} concentration regardless of any predicted moment in any period, and there is no significant difference in each period. The Pearson correlation coefficients of the cold and hot periods are almost above 0.8 in the first three predicted moments, while *r* in the warm period is lower than 0.76, indicating that the Staging-eSNN model predicts the warm period even worse. IOA shows results similar to *r*.



For Beijing, except for 3-hour-ahead prediction in the warm period, the bias is almost less than 0 in both periods, indicating that the Staging-eSNN model underestimates the target value, which is different from the prediction result in Shanghai. RMSPE and FAC2 indicate that the prediction result of the warm period is better than the cold period, while IOA and r are more optimistic about the performance of the Staging-eSNN model in the cold period.

The scatter plot of Fig. 9 shows the relationship between the target value and the predicted value of the Staging-eSNN model at five predicted moments in different periods. The diagonal solid line indicates that the predicted value and the target value are equal, and the dotted line indicates the critical condition of FAC2. For the Shanghai area, more than 85% of the scattered points fall within the range of the dotted line and are distributed near the diagonal. For high target values, the scatter is more likely to deviate from the diagonal, which means that the Staging-eSNN model has a poor prediction of high PM_{2.5} concentration. The prediction results in the Beijing area verify this conclusion. From Fig. 9(a), it can be found that the r and FAC2 indicators in the Beijing area perform worse than the Shanghai area, which is the result of the Beijing area having more high target values.

It should be noted that six indicators used reflect different aspects of prediction capabilities (Table 4); the performance of the Staging-eSNN model in different periods depends on what we focus on. But in general, the prediction results of both regions show that the Staging-eSNN model performance will deteriorate as the predicted moment moves backward. The conclusion can be explained by the proportion of outliers. Based on the definition of outliers (Wang *et al.*, 2017a), we compared the PM_{2.5} series in the sample to determine whether the target value is an outlier. The final results are listed in Table 5. It can be found that at any period in the two cities, as the predicted moment moves backward, the proportion of outliers in the target value will increase.

4.3 Predictive Performance Evaluation of Different Models

We compared the predictive performance of SVR, RF, Plain-eSNN, CEEsNN and Staging-eSNN models using the six indicators mentioned above. The Plain-eSNN model was implemented entirely based on the NeuCube. In order to prevent the impact of sample differences, the same samples were used for the three models. And for SVR and RFs, two-fold cross-validation was applied to model training and prediction using the sampled 1000-hour data. The evaluation results of the models for the prediction of the next 1, 3, 6, 12 and 24 hours are shown in Fig. 10.

When comparing the three SNN models, we found that no matter in Beijing or Shanghai,

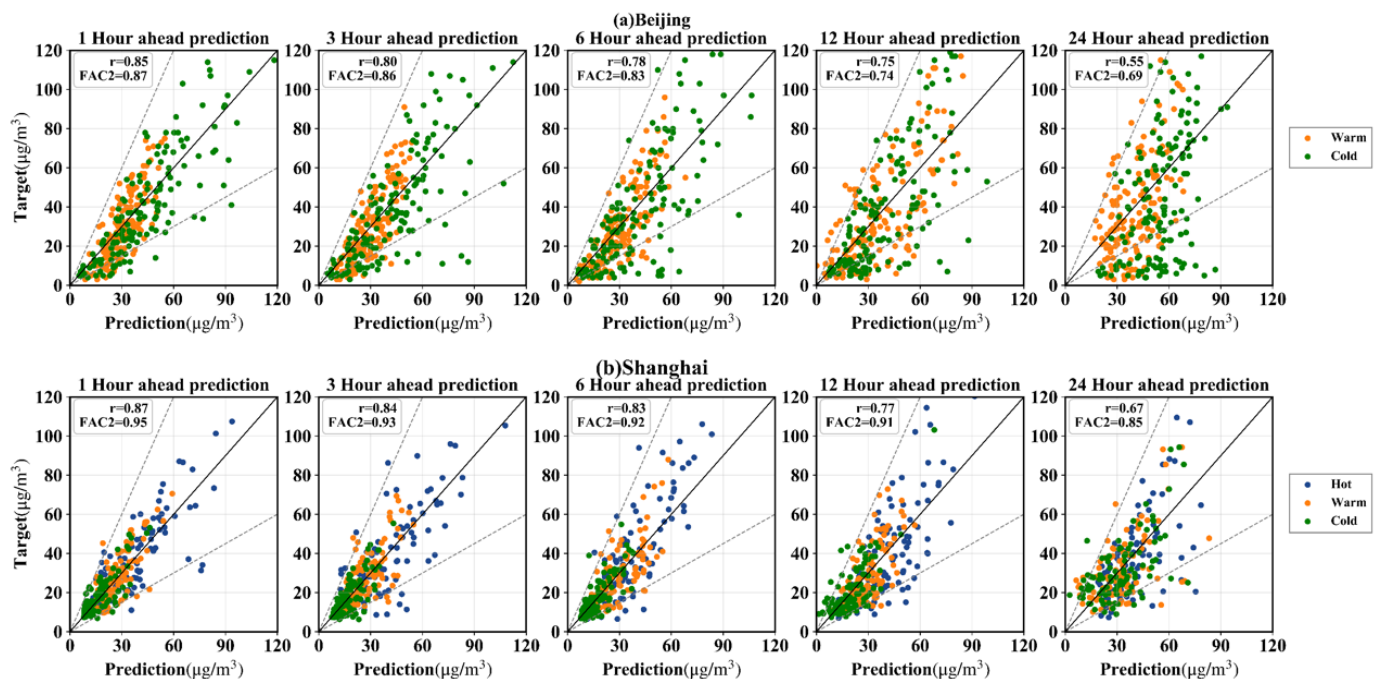


Fig. 9. The scatter plot of predicted and target values of the Staging-eSNN model at five predicted moments in different periods.

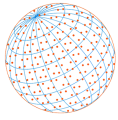


Table 5. Proportion of outliers in target values in different periods in Beijing and Shanghai.

Proportion	Hour 1	Hour 3	Hour 6	Hour 12	Hour 24
<i>Beijing</i>					
Cold	57%	60%	65%	71%	99%
Warm	51%	52%	52%	78%	91%
<i>Shanghai</i>					
Cold	9%	16%	19%	30%	33%
Warm	15%	19%	26%	28%	41%
Hot	9%	11%	17%	21%	33%

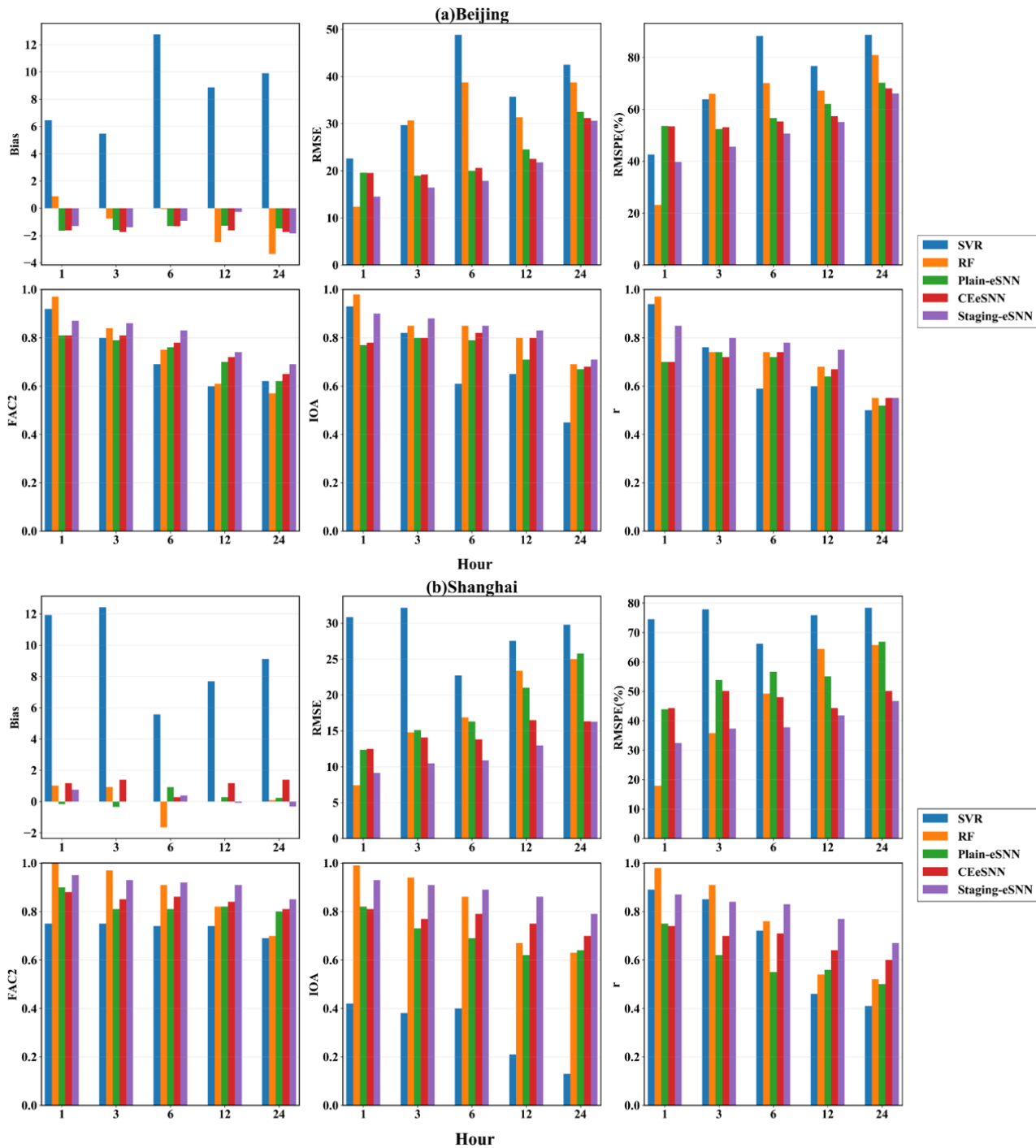
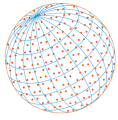


Fig. 10. Comparison of different models.



almost all the indicators show that the Staging-eSNN model has the best prediction performance. Specifically, taking Shanghai as an example, RMSE of the Staging-eSNN model is 9.14, 10.46, 10.85, 12.99 $\mu\text{g m}^{-3}$ at the first four predicted moments, which is lower than CEeSNN's 12.49, 14.04, 13.81, 16.49 $\mu\text{g m}^{-3}$ and Plain-eSNN's 12.38, 15.12, 16.29, 20.99 $\mu\text{g m}^{-3}$. The FAC2 and IOA are above 0.8 regardless of any predicted moment in the Staging-eSNN model, which are higher than the other two SNN models. In the performance of the r indicator, the correlation coefficient of the prediction result of the Plain-eSNN model is between 0.50–0.75, which is lower than r (0.60–0.74) in the CEeSNN model and is lower than r (0.67–0.87) in the Staging-eSNN model. For the prediction result of Beijing, except for the sign of the bias which is opposite to that of Shanghai in the most predicted moments, the other five evaluation indicators have similar trends to those of Shanghai. These facts show that the Staging-eSNN model improves the prediction accuracy of SNN models and has the best prediction performance in the same series of SNN models. On the other hand, the CEeSNN model also has a clear improvement in prediction accuracy compared to the Plain-SNN model.

When compared with the SVR and RF model, we found that Staging-eSNN model's prediction results are slightly worse than these two models in the first (Beijing) or the first two (Shanghai) predicted moments, while in the prediction at other moments, the Staging-SNN model almost all showed the best prediction performance compared to the other four models.

5 CONCLUSIONS

To better predict the $\text{PM}_{2.5}$ concentration, we developed an eSNN model, Staging-eSNN, that applies a time series clustering algorithm (PAM-DTW) to differentiate between seasonal and diurnal variation in the levels of this pollutant. Based on our forecasts for the concentrations in Beijing and Shanghai 1, 3, 6, 12 and 24 hours in advance, we drew the following conclusions.

- 1). When using the bias, RMSE, RMSPE, FAC2, IOA, and r -value as evaluation indicators, the Staging-eSNN model exhibited better performance than classical NeuCube and the recently proposed CEeSNNs, which indicates that Staging-eSNN offers good predictive accuracy and stability compared to other eSNN models.
- 2). When predicting the $\text{PM}_{2.5}$ concentration 1 or 3 hours in advance, the Staging-eSNN model performed slightly worse than the SVR and RF models.
- 3). Staging-eSNN incorporated seasonal factors to enhance the accuracy of its $\text{PM}_{2.5}$ forecasting. This capability merits further investigation and can be added to future models for predicting the concentrations of other pollutants.

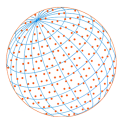
Although this study achieved several promising results, potential improvements include better mapping of the input variables to the 3D SNN architecture and further optimization of the algorithm's parameters. We also plan to use residual modeling to increase Staging-eSNN's predictive accuracy.

ACKNOWLEDGEMENTS

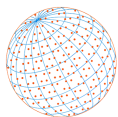
The work was financially supported by the National Natural Science Foundation of China (No. 41105102) and the National Key R&D Program of China (No. 2018YFC0213600).

REFERENCES

- Arbelaitz, O., Gurrutxaga, I., Muguerza, J., Pérez, J.M., Perona, I. (2013). An extensive comparative study of cluster validity indices. *Pattern Recognit.* 46, 243–256. <https://doi.org/10.1016/j.patcog.2012.07.021>
- Carslaw, D.C., Beevers, S.D. (2013). Characterising and understanding emission sources using bivariate polar plots and k-means clustering. *Environ. Modell. Software* 40, 325–329. <https://doi.org/10.1016/j.envsoft.2012.09.005>
- Chen, J., Lu, J., Avise, J.C., DaMassa, J.A., Kleeman, M.J., Kaduwela, A.P. (2014). Seasonal modeling of $\text{PM}_{2.5}$ in California's San Joaquin Valley. *Atmos. Environ.* 92: 182–190. <https://doi.org/10.1016/j.atmosenv.2014.04.030>



- Elangasinghe, M.A., Singhal, N., Dirks, K.N., Salmond, J.A., Samarasinghe, S. (2014). Complex time series analysis of PM₁₀ and PM_{2.5} for a coastal site using artificial neural network modelling and k-means clustering. *Atmos. Environ.* 94, 106–116. <https://doi.org/10.1016/j.atmosenv.2014.04.051>
- Feng, J., Li, M., Zhang, P., Gong, S., Zhong, M., Wu, M., Zheng, M., Chen, C., Wang, H., Lou, S. (2013). Investigation of the sources and seasonal variations of secondary organic aerosols in PM_{2.5} in Shanghai with organic tracers. *Atmos. Environ.* 79, 614–622. <https://doi.org/10.1016/j.atmosenv.2013.07.022>
- Hartono, R.N., Pears, R., Kasabov, N., Worner, S.P. (2014). Extracting temporal knowledge from time series: A case study in ecological data. 2014 International Joint Conference on Neural Networks (IJCNN), Beijing, 2014, pp. 4237–4243. <https://doi.org/10.1109/IJCNN.2014.6889918>
- Hyndman, R.J., Koehler, A.B. (2006). Another look at measures of forecast accuracy. *Int. J. Forecasting.* 22, 679–688. <https://doi.org/10.1016/j.ijforecast.2006.03.001>
- Jian, L., Zhao, Y., Zhu, Y.P., Zhang, M.B., Bertolatti, D. (2012). An application of ARIMA model to predict submicron particle concentration from meteorological factors at a busy roadside in Hangzhou, China. *Sci. Total Environ.* 426, 336–345. <https://doi.org/10.1016/j.scitotenv.2012.03.025>
- Kasabov, N.K. (2014). NeuCube: A spiking neural network architecture for mapping, learning and understanding of spatio-temporal brain data. *Neural Networks.* 52, 62–76. <https://doi.org/10.1016/j.neunet.2014.01.006>
- Kasabov, N., Scott, N.M., Tu, E., Marks, S., Sengupta, N., Capecci, E., Othman, M., Doborjeh, M.G., Murli, N., Hartono, R., Espinosa-Ramos, J.I., Zhou, L., Alvi, F.B., Wang, G., Taylor, D., Feigin, V., Gulyaev, S., Mahmoud, M., Hou, Z.G., Yang, J. (2016). Evolving spatio-temporal data machines based on the NeuCube neuromorphic framework: Design methodology and selected applications. *Neural Networks.* 78, 1–14. <https://doi.org/10.1016/j.neunet.2015.09.011>
- Kasabov, N.K. (2019). Time-space, spiking neural networks and brain-inspired artificial intelligence. Springer-Verlag Berlin Heidelberg. <https://doi.org/10.1007/978-3-662-57715-8>
- Kim, S.U., Kim, K.Y. (2020). Physical and chemical mechanisms of the daily-to-seasonal variation of PM₁₀ in Korea. *Sci Total Environ.* 712, 136429. <https://doi.org/10.1016/j.scitotenv.2019.136429>
- Li, G., Fang, C., Wang, S., Sun, S. (2016). The effect of economic growth, urbanization, and industrialization on fine particulate matter (PM_{2.5}) concentration in China. *Environ. Sci. Technol.* 50, 11452–11459. <https://doi.org/10.1021/acs.est.6b02562>
- Liu, C., Chen, R., Sera, F., Vicedo-Cabrera, A.M., Guo, Y., Tong, S., Coelho, M.S.Z.S., Saldiva, P.H.N., Lavigne, E., Matus, P., Valdes Ortega, N., Osorio Garcia, S., Pascal, M., Stafoggia, M., Scortichini, M., Hashizume, M., Honda, Y., Hurtado-Díaz, M., Cruz, J., ... Kan, H. (2019). Ambient particulate air pollution and daily mortality in 652 cities. *N. Engl. J. Med.* 381, 705–715. <https://doi.org/10.1056/NEJMoa1817364>
- Ma, X., Jia, H., Sha, T., An, J., Tian, R. (2019). Spatial and seasonal characteristics of particulate matter and gaseous pollution in China: implications for control policy. *Environ. Pollut.* 248, 421–428. <https://doi.org/10.1016/j.envpol.2019.02.038>
- Maciąg, P.S., Kasabov, N., Kryszkiewicz, M., Bembenik, R. (2019). Air pollution prediction with clustering-based ensemble of evolving spiking neural networks and a case study for London area. *Environ. Modell. Software* 118, 262–280. <https://doi.org/10.1016/j.envsoft.2019.04.012>
- Petitjean, F., Ketterlin, A., Gancarski, P. (2011). A global averaging method for dynamic time warping, with applications to clustering. *Pattern Recognit.* 44, 678–693. <https://doi.org/10.1016/j.patcog.2010.09.013>
- Reid, D., Hussain, A.J., Tawfik, H. (2014). Financial Time Series Prediction Using Spiking Neural Networks. *PLoS One* 9, e103656. <https://doi.org/10.1371/journal.pone.0103656>
- Saide, P.E., Carmichael, G.R., Spak, S.N., Gallardo, L., Osses, A.E., Mena-Carrasco, M.A., Pagowski, M. (2011). Forecasting urban PM₁₀ and PM_{2.5} pollution episodes in very stable nocturnal conditions and complex terrain using WRF-hem Co tracer model. *Atmos. Environ.* 45, 2769–2780. <https://doi.org/10.1016/j.atmosenv.2011.02.001>
- Shen, F., Zhang, L., Jiang, L., Tang, M., Gai, X., Chen, M., Ge, X. (2020). Temporal Variations of Six Ambient Criteria Air Pollutants from 2015 to 2018, Their Spatial Distributions, Health Risks and Relationships with Socioeconomic Factors During 2018 in China. *Environ. Int.* 137, 105556. <https://doi.org/10.1016/j.envint.2020.105556>



- Siwek, K., Osowski, S. (2012). Improving the accuracy of prediction of PM₁₀ pollution by the wavelet transformation and an ensemble of neural predictors. *Eng. Appl. Artif. Intell.* 25, 1246–1258. <https://doi.org/10.1016/j.engappai.2011.10.013>
- Stern, R., Builtjes, P., Schaap, M., Timmermans, R., Vautard, R., Hodzic, A., Memmesheimer, M., Feldmann, H., Renner, E., Wolke, R., Kerschbaumer, A. (2008). A model inter-comparison study focussing on episodes with elevated PM₁₀ concentrations. *Atmos. Environ.* 42, 4567–4588. <https://doi.org/10.1016/j.atmosenv.2008.01.068>
- Sun, W., Sun, J. (2017). Daily PM_{2.5} concentration prediction based on principal component analysis and LSSVM Optimized by cuckoo search algorithm. *J. Environ. Manage.* 188, 144–152. <https://doi.org/10.1016/j.jenvman.2016.12.011>
- Tavanaei, A., Ghodrati, M., Kheradpisheh, S.R., Masquelier, T., Maida, A. (2019). Deep learning in spiking neural networks. *Neural Networks* 111, 47–63. <https://doi.org/10.1016/j.neunet.2018.12.002>
- Tu, E., Kasabov, N., Yang, J. (2017). Mapping temporal variables into the neucube for improved pattern recognition, predictive modeling, and understanding of stream data. *IEEE Trans. Neural Networks Learn. Syst.* 28, 1305–1317. <https://doi.org/10.1109/TNNLS.2016.2536742>
- Wang, Y., Wang, C., Shi, C., Xiao, B. (2017). Short-term cloud coverage prediction using the arima time series model. *Remote Sens. Lett.* 9, 274–283. <https://doi.org/10.1080/2150704x.2017.1418992>
- Wang, Y., Zhang, H., Zhai, J., Wu, Y., Cong, L., Yan, G., Zhang, Z. (2019). Seasonal variations and chemical characteristics of PM_{2.5} aerosol in the urban green belt of Beijing, China. *Pol. J. Environ. Stud.* 29, 361–370. <https://doi.org/10.15244/pjoes/104358>
- Wen, C., Liu, S., Yao, X., Peng, L., Li, X., Hu, Y., Chi, T. (2019). A novel spatiotemporal convolutional long short-term neural network for air pollution prediction. *Sci. Total Environ.* 654, 1091–1099. <https://doi.org/10.1016/j.scitotenv.2018.11.086>
- Xu, Y., Du, P., Wang, J. (2017). Research and application of a hybrid model based on dynamic fuzzy synthetic evaluation for establishing air quality forecasting and early warning system: A case study in China. *Environ. Pollut.* 223, 435–448. <https://doi.org/10.1016/j.envpol.2017.01.043>
- Zhai, W., Cheng, C. (2020). A long short-term memory approach to predicting air quality based on social media data. *Atmos. Environ.* 237, 117411. <https://doi.org/10.1016/j.atmosenv.2020.117411>
- Zhang, L., Lin, J., Qiu, R., Hu, X., Zhang, H., Chen, Q., Tan, H., Lin, D., Wang, J. (2018). Trend analysis and forecast of PM_{2.5} in Fuzhou, China using the ARIMA model. *Ecol. Indic.* 95, 702–710. <https://doi.org/10.1016/j.ecolind.2018.08.032>
- Zhao, X., Zhang, X., Xu, X., Xu, J., Meng, W., Pu, W. (2009). Seasonal and diurnal variations of ambient PM_{2.5} concentration in urban and rural environments in Beijing. *Atmos. Environ.* 43, 2893–2900. <https://doi.org/10.1016/j.atmosenv.2009.03.009>