



# A Development of PM2.5 Forecasting System in South Korea Using Chemical Transport Modeling and Machine Learning

Youn-Seo Koo<sup>1</sup> · Hee-Yong Kwon<sup>2</sup> · Hyosik Bae<sup>3</sup> · Hui-Young Yun<sup>1</sup> · Dae-Ryun Choi<sup>1</sup> · SukHyun Yu<sup>4</sup> · Kyung-Hui Wang<sup>1</sup> · Ji-Seok Koo<sup>2</sup> · Jae-Bum Lee<sup>5</sup> · Min-Hyeok Choi<sup>5</sup> · Jeong-Beom Lee<sup>5</sup>

Received: 29 August 2022 / Revised: 18 December 2022 / Accepted: 3 January 2023 / Published online: 13 January 2023  
© The Author(s) under exclusive licence to Korean Meteorological Society and Springer Nature B.V. 2023

## Abstract

Ambient exposure to PM2.5 can adversely affect public health, and forecasting PM2.5 is essential for implementing protection measures in advance. Current PM2.5 forecasting systems are primarily based on the chemical transport model of Community Multiscale Air Quality (CMAQ) modeling systems and the Weather Research and Forecasting (WRF) model. However, the forecasting accuracies of these models are substantially constrained by uncertainties in the input data of anthropogenic emissions and meteorological fields, as well as inherent limitations in the models. The PM2.5 forecasting system developed in this study aimed at overcoming the limitations of CMAQ predictions by utilizing advanced machine learning algorithms. The proposed system was developed using forecast data from CMAQ and WRF, as well as observed PM2.5 concentrations and meteorological variables at monitoring stations in China and South Korea. It was then applied to national PM2.5 forecasting in South Korea. This study focused on developing secondary input data and machine learning models that can reflect the long-range transport in Northeast Asia. The proposed system can forecast 6-h average PM2.5 concentrations up to two days in advance at 19 forecast regions in South Korea. To evaluate the performance of the proposed models, a real-time machine learning-based forecasting system was applied to 19 forecasting regions from January 2020 to April 2021. Herein, the four machine learning algorithms applied, including deep neural network, recurrent neural network, convolutional neural network, and Ensemble, could reduce the over-prediction of the CMAQ forecast by decreasing the normal mean bias and improving the index of agreement. The reduced false alarm rates and high prediction accuracy confirm the feasibility of these models for practical applications.

**Keywords** PM2.5 · forecast · machine learning · DNN · RNN · CNN

## 1 Introduction

Particulate matter (PM) in the atmosphere can adversely affect natural ecosystems and public health and lead to climate change. In particular, exposure to particulate matter with an aerodynamic diameter  $<2.5 \mu\text{m}$  (PM2.5) increases the risk of cardiovascular morbidity and mortality (Bourdrel et al. 2017; Du et al. 2016). According to the Global Burden of Disease Project, PM2.5 exposure may have contributed to 16,135 deaths in South Korea in 2017 (IHME 2018).

South Korea has undergone rapid urbanization and industrialization, leading to severe air pollution. Consequently, the problem of exposure to high concentrations of PM has received increased attention nationwide, making it a social and environmental problem. Forecasting the air quality in terms of PM concentration has become increasingly important as environmental authorities and citizens require air quality

---

Responsible Editor: Chang-Keun Song.

Youn-Seo Koo  
koo@anyang.ac.kr

<sup>1</sup> Department of Environmental and Energy Engineering, Anyang University, Anyang, South Korea

<sup>2</sup> Department of Computer Engineering, Anyang University, Anyang, South Korea

<sup>3</sup> Nextsoft, Anyang, South Korea

<sup>4</sup> Department of Information, Electrical & Electronic Engineering, Anyang University, Anyang, South Korea

<sup>5</sup> National Institute of Environmental Research, Incheon, Republic of Korea



information in advance to implement appropriate public health protection measures during pollution episodes.

The air quality forecasting system typically used to detect high-PM episodes in South Korea primarily includes physicochemical models based on the chemical transport model (CTM) of CMAQ, weather forecasting model of WRF, and emission processing model of Sparse Matrix Operator Kernel Emissions (SMOKE). For example, a Korean air quality forecasting system (KAQFS) developed by Koo et al. (2012, 2015) has been operating since 2007 with open access to citizens (<http://www.kaq.or.kr>). However, this numerical model has structural constraints that affect the forecasting accuracy because of uncertainties in meteorological and emission inventory data as well inherent limitations of the model.

Various machine learning algorithms have been developed to overcome the drawbacks of the physicochemical models used in PM2.5 forecasting. The transport of PM2.5 from neighboring regions, including long-range transport (LRT), and the secondary formation of PM2.5 through complex homogeneous and heterogeneous chemical reactions in the atmosphere must be reflected in the machine learning model to ensure effective PM2.5 forecasting. LRT is defined as the transport characteristic of air pollutants, and the secondary formation of PM2.5 as the formation characteristic. Because air pollutants are transported from upwind to downwind regions along the wind direction, geographical conditions and velocity fields are important factors in determining the PM2.5 transport. In contrast, the emission of air pollutants in the forecast region and meteorological conditions such as wind speed, temperature, humidity, and insolation are also important factors in the PM2.5 formation processes.

Various spectra of the machine learning algorithms, such as feed forward and recurrent networks, have been developed for forecasting regional air quality. Cabaneros et al. (2019) reviewed 139 papers published in international peer-reviewed journals for artificial neural network models in ambient air pollution forecasting from 2001 to February 2019. Li et al. (2022) also examined the application of machine learning techniques in developing air quality forecasting systems during the last two decades (2000–2019) by scientometric and content analysis.

Previous studies have shown that the accuracy of machine learning methods using the monitoring station data decreases significantly as the prediction time interval increases, making long-term predictions up to or greater than 24 h challenging. A method to ensure longer forecast time is to use the PM concentrations and meteorological data forecasted by the CTM such as CMAQ and WRF as the input data, which can reflect the transport and formation mechanisms in the machine learning algorithm. Various studies (Ho et al. 2021; Kim et al. 2019; Lee et al. 2022; Lightstone et al. 2017, 2021) have used the CMAQ and WRF model outputs to incorporate

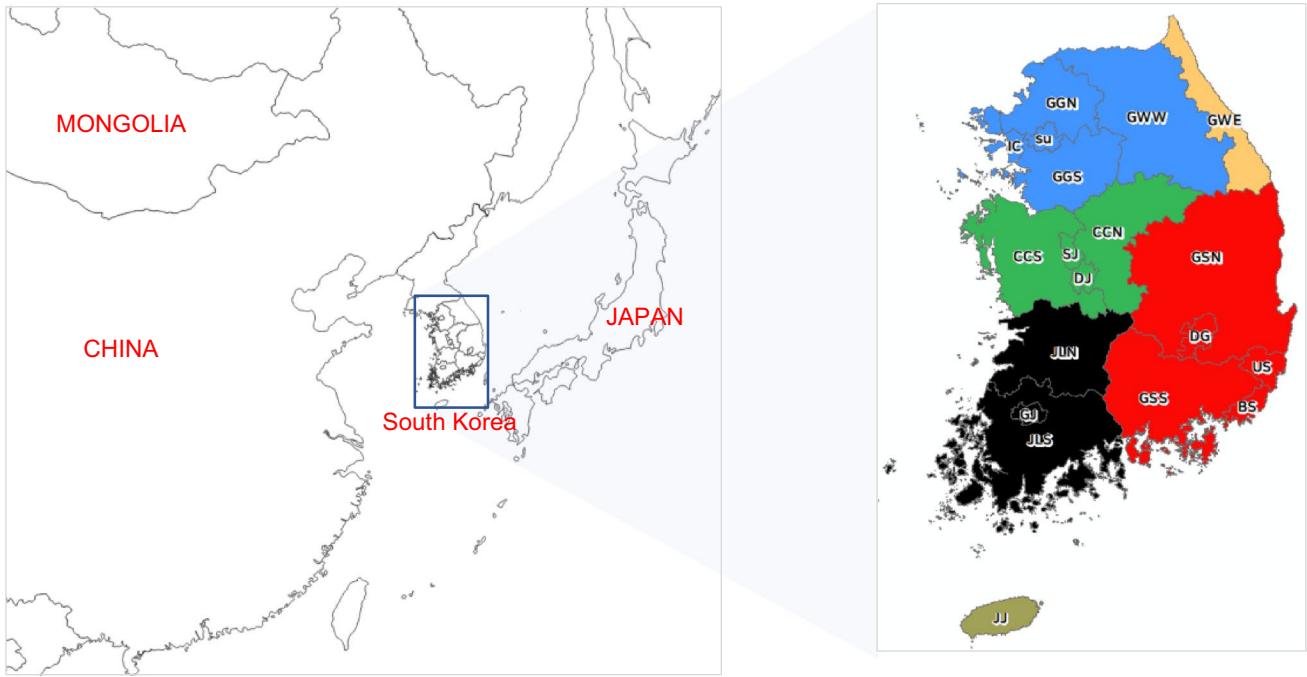
the spatiotemporal variations in the forecasting. These studies showed that machine learning could improve the forecasting accuracy by correcting the biases from the CTM.

To overcome the drawback of the KAQFS and utilize the recent advancements in machine learning algorithms, a PM2.5 forecasting system using machine learning algorithms with CTM forecasts was developed in this study and initial models were validated with real-time operational testing for regional PM2.5 forecasting in South Korea. The PM2.5 concentrations highly depend on the LRT from China and local emissions; therefore, input feature data and machine learning algorithms were designed to reflect the spatial relationship between the source and receptor regions as well as the local temporal variation in PM2.5 in forecast regions.

In this paper, we first demonstrate a method for the organization of input data that can accurately represent the spatial relationship between the source and forecast regions. We then discuss the development of deep neural network (DNN), recurrent neural network (RNN), and convolutional neural network (CNN) models, and analyze the forecasting performance of the proposed forecasting system in real-time operation.

## 2 Forecast Regions and Evaluation Metrics

Figure 1 presents a map of Northeast Asia and the location of 19 forecast regions in South Korea. Air pollutants in China are transported along the northwesterly wind to South Korea after one or two days. Previous studies have shown that high-PM episodes, such as haze events, in South Korea are initiated by the LRT of these pollutants from China and further deteriorated by the accumulation of emission from local sources. Numerous studies (Kim et al. 2016; Koo et al. 2018; Lee et al. 2011a, b, 2013; Oh et al. 2015, 2020) have reported that the LRT from China and stagnant high-pressure system over Korea are the main inducers of the haze events occurring in the Seoul Metropolitan Area (SMA). Koo et al. (2018) demonstrated that a haze event in the SMA during January 12–18, 2013 was an extension of an extraordinary haze event that occurred in northern China during that period, triggered by PM transport by the northwesterly wind. Kim et al. (2017) showed that the LRT of PM10 and PM2.5 from China contributed to 64% and 70% of these pollutants, respectively, in the SMA during a PM episode in late February 2014. The LRT from China and PM emissions from local sources are the most important factors in determining the PM concentrations in South Korea. To reflect the LRT from China to South Korea in machine learning, air quality and weather monitoring data were designed to be used as input data from China and South Korea as described in Section 3. The identification of a spatial relationship of PM transport between China and South Korea is crucial for developing an air-pollution forecasting system with machine learning.



**Fig. 1** Geographic map of Northeast Asia and 19 forecast regions with 6 clusters in South Korea. The clusters are color-coded for grouping the source cluster with the corresponding forecast regions as follows: Cluster 1: SU (Seoul), GGS (South Gyeonggi), GGN (North Gyeonggi), GWW (West Gangwon), IC (Incheon); Cluster

2: DJ (Daejeon), SJ (Sejong), CCS (Chungnam), CCN (Chungbuk); Cluster 3: GSS (Gyeongnam), GSN (Gyeongbuk), DG (Daegu), BS (Busan), US (Ulsan); Cluster 4: GJ (Gwangju), JLS (Jeonnam), JLN (Jeonbuk); Cluster 5: GWE (East Gangwon); and Cluster 6: JJ (Jeju Island)

The 19 forecast regions in South Korea are also shown in Fig. 1. To prevent overfitting caused by a lack of machine learning input data, the 19 forecast regions were classified into 6 metropolitan areas, which show similar patterns in air quality and weather variations with time. Cluster analysis was performed by applying a hierarchical algorithm and Ward's method (Anderberg 2014; Batagelj 1988). The forecast regions were classified into clusters using the similarity of observed PM2.5 concentrations, as shown in Fig. 1. The forecast regions in the same metropolitan areas shared the input data for the machine learning models.

Additionally, the mean bias (MBIAS), normalized mean bias (NMB), index of agreement (IOA), and correlation coefficient ( $r$ ) were used as statistical evaluation metrics to determine the PM2.5 forecasting performance.

$$MBIAS = \frac{1}{N} \sum_1^N (\underline{Model} - \underline{Obs}) \quad (1)$$

$$NMB = \frac{\sum_1^N (\underline{Model} - \underline{Obs})}{\sum_1^N \underline{Obs}} \times 100 \quad (2)$$

$$IOA = 1 - \frac{\sum_1^N (\underline{Model} - \underline{Obs})^2}{\sum_1^N (|\underline{Model} - \underline{Obs}| - |\underline{Obs} - \underline{Obs}|)^2} \quad (3)$$

$$r = \frac{\sum_1^N (\underline{Model} - \underline{Model}) \times (\underline{Obs} - \underline{Obs})}{\sqrt{\sum_1^N (\underline{Model} - \underline{Model})^2} \times \sqrt{\sum_1^N (\underline{Obs} - \underline{Obs})^2}} \quad (4)$$

where  $N$  indicates the sample size,  $\underline{Model}$  represents the PM2.5 concentration determined by the applied model, and  $\underline{Obs}$  indicates the observed PM2.5 concentration. The underlined  $Model$  and  $Obs$  represent average values.

The air quality index for PM2.5 in South Korea was classified into four categories based on the national ambient PM2.5 standard and results from public health studies according to the following criteria: good, 0 to  $15 \mu\text{g m}^{-3}$ ; moderate, 16 to  $35 \mu\text{g m}^{-3}$ ; bad, 36 to  $75 \mu\text{g m}^{-3}$ ; and very bad,  $76 \mu\text{g m}^{-3}$  or higher. The categorical performance of the PM2.5 index forecast was analyzed by accuracy (ACC), probability of detection (POD) for bad and very bad indices, and false alarm rate (FAR) for bad and very bad indices (Lee et al. 2022). The statistical and categorical analyses were used to evaluate the forecasting performance of the CMAQ and machine learning models.

### 3 Input Data

The input feature data were classified into primary and secondary data, as shown in Fig. 2. The primary input features were based on the observations and model forecast

predicted by physiochemical models of WRF and CMAQ, whereas the secondary data were derived by processing and combining the primary data to provide the spatiotemporal linkage between the remote source emission regions in the upwind and forecast regions.

### 3.1 Primary Input Data

#### 3.1.1 Observations

Observed air quality and meteorological values over the northeastern region were used as the primary data; their monitoring locations are shown in Fig. 3. The hourly air quality data comprised the PM10, PM2.5, NO<sub>2</sub>, CO, and SO<sub>2</sub> concentrations and the hourly meteorological variables of wind speed, wind direction, humidity, surface pressure, temperature, and precipitation rate.

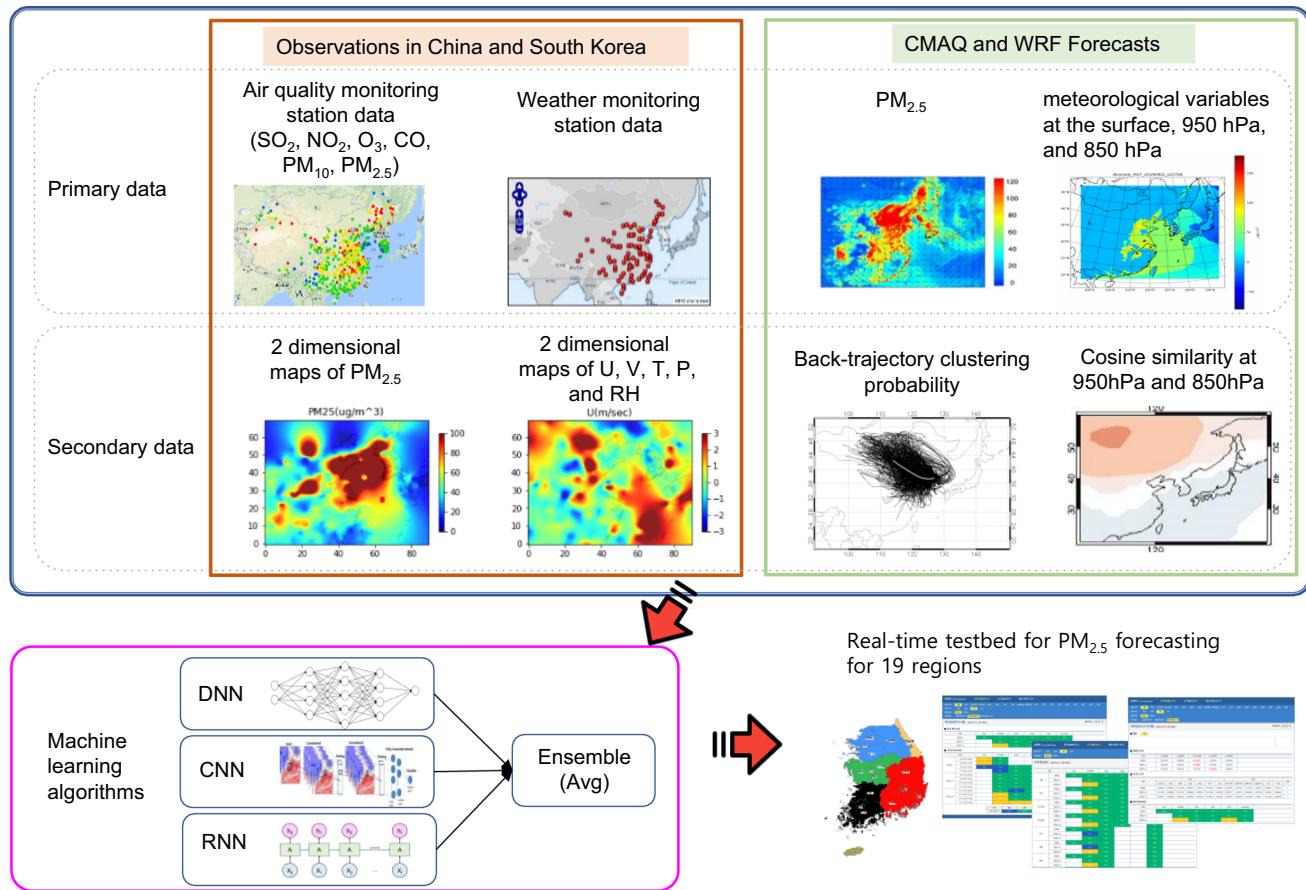
The air quality data for South Korea were obtained from the AirKorea website ([www.airkorea.or.kr](http://www.airkorea.or.kr)) and those for China were from the PM25.in website (pm25.in, non-operational since June 2022). The meteorological data for China and South Korea

were acquired from the Korea Meteorological Administration URL-API external service website (<http://203.247.66.28/>). To account for missing monitoring data in a specific period, the average values available for the corresponding period were used as the data for the entire period. The number of air quality and meteorological monitoring stations is increasing yearly. As of 2020, 1,196 air quality monitoring stations are active in China, while 416 are available in South Korea. Further, there are 590 meteorological stations in China and 93 in South Korea.

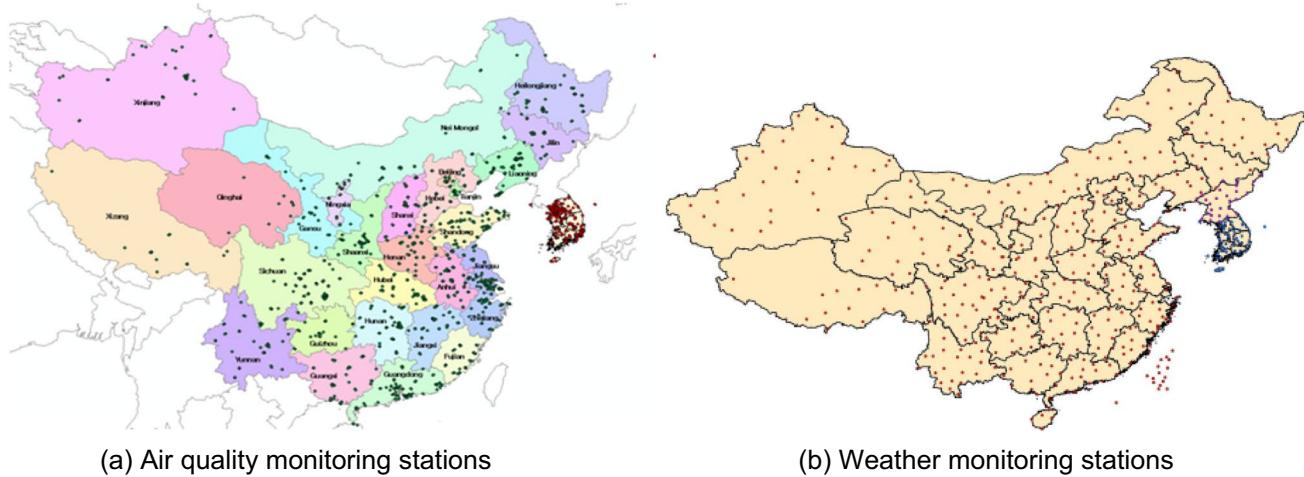
The observed air quality and meteorological data were used as the primary historical features in the DNN and RNN models. The station data were converted to two-dimensional distribution maps over the modeling regions (including China and South Korea) using the kriging method. The maps were then used as CNN input data, which define the spatial relationship for PM2.5 transport between China and South Korea. The two-dimensional maps are discussed in detail in Sect. 3.2.3.

#### 3.1.2 CMAQ and WRF Data

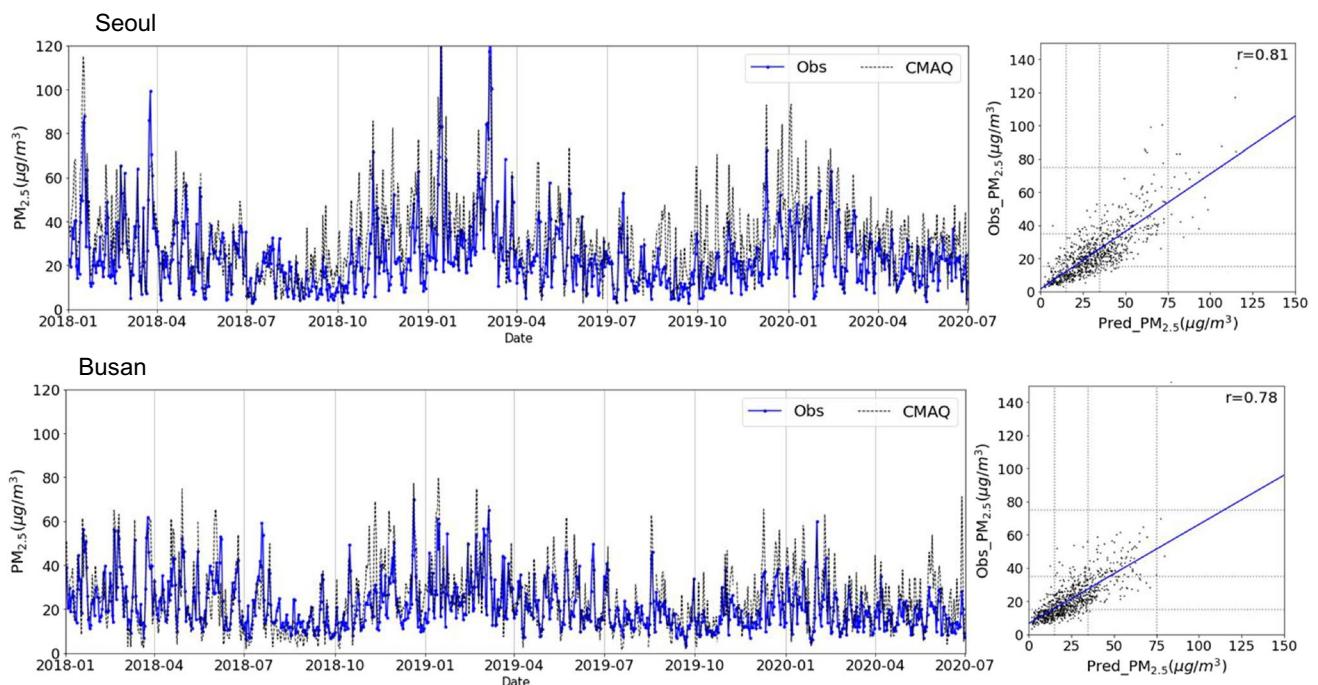
The KAQFS was used to establish the forecast data of PM2.5 concentrations and meteorological variables. It consists of the



**Fig. 2** Outline of the PM<sub>2.5</sub> forecasting system using machine learning algorithms of DNN, RNN, CNN, and ESB method



**Fig. 3** Locations of air quality and meteorological stations in China and South Korea



**Fig. 4** Time-series comparison and scatter plots of the observed daily average PM<sub>2.5</sub> concentrations and the CMAQ forecast values for the first day forecast at Seoul and Busan from 2018–01 to 2020–06

WRF model for meteorological modeling, the SMOKE system for emission data processing, and the CMAQ model for chemical transport simulation over Northeast Asia. The modeling configurations of the KAQFS with CMAQ and WRF models with emission inventories are provided in Supplementary Material 1, and further details are described in Koo et al. (2012, 2015, 2018).

Time-series comparison and scatter plots of the observed daily average PM<sub>2.5</sub> concentrations with the CMAQ forecasts for the first day (D+1) and the second day (D+2) forecasts at Seoul (capital of South Korea) and Busan (second

largest city in the southeast of the Korean Peninsula) from 2018–01 to 2020–06 are shown in Figs. 4 and 5, respectively. The statistical performance of the PM<sub>2.5</sub> forecasting system is summarized in Table 1.

The temporal variations in the forecasted PM<sub>2.5</sub> concentrations could depict the variations of the observed PM<sub>2.5</sub>, and the forecasted values were in good agreement with the observed values with IOAs of 0.84 and 0.85 for D+1 in Seoul and Busan, respectively. However, the forecasted PM<sub>2.5</sub> concentration in Seoul was highly over-predicted, with an NMB of 34.27% for

D+1. This over-prediction results in high POD of 89.03% with an FAR of 59.05% and low ACC of 57.00%, as shown in Table 2. The PM<sub>2.5</sub> forecast for Busan also exhibited a similar trend as that for Seoul, but the degree of over-prediction was relatively low (NMB of 14.92%, ACC of 68.13%, POD of 78.18%, and FAR of 53.01%). The high concentration predicted in South Korea was attributed to the use of overestimated emission data in the CTM.

This implies that the CMAQ forecast could detect the high-PM<sub>2.5</sub> episode in advance, but the accuracy and FAR need to be improved to increase the public reliability of advance forecasting. We expect machine learning could improve the over-prediction of the PM<sub>2.5</sub> forecast of the CMAQ.

### 3.2 Secondary Input Data

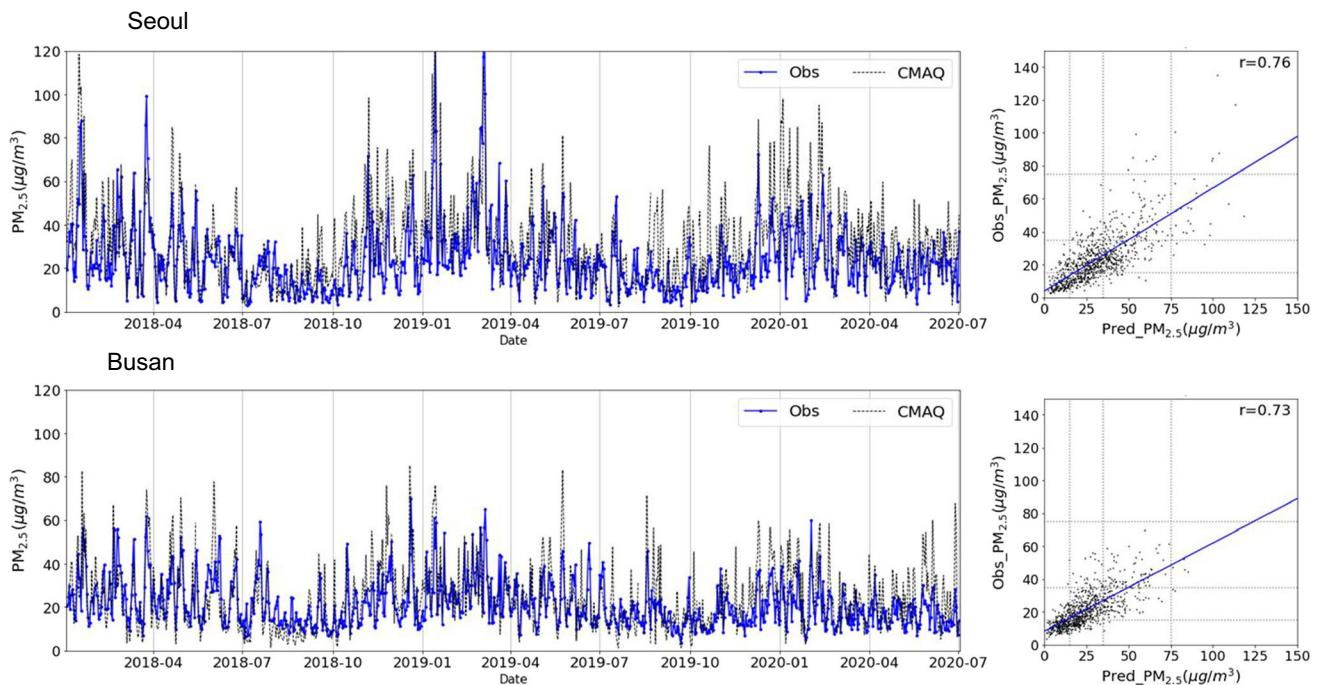
Secondary data were organized in this study to reflect the spatial relationships between the emission regions in China

(sources from which LRT occurs) and forecast regions in South Korea (sources of regional emissions) using the machine learning algorithm. The FLEXible PARTicle dispersion model (FLEXPART) backward trajectory, cosine similarities of the meteorological variables, and two-dimensional distributions of observed PM<sub>2.5</sub> concentrations and meteorological variables were selected as the secondary data.

#### 3.2.1 FLEXPART Backward Trajectory

The backward trajectories based on the FLEXPART (Stohl et al. 2005; Fast and Easter 2006) were used to identify the LRT from China and influence of local emissions in South Korea. The backward trajectories at forecast regions starting at a height of 500 m were estimated using the WRF predicted meteorological data from the KAQFS, which operates twice daily.

Hoi et al. (2021) performed a cluster analysis on the FLEXPART back-trajectory data to identify the origin of



**Fig. 5** Time-series comparison and scatter plots of the observed daily average PM<sub>2.5</sub> concentrations, and the CMAQ forecast values for the second day forecast at Seoul and Busan from 2018–01 to 2020–06

**Table 1** Summary of PM<sub>2.5</sub> forecast performance of the CMAQ system for the first day (D+1) and the second day (D+2) forecasts at Seoul and Busan from 2018–01 to 2020–06

Region	classification	IOA	MBIAS	NMB	RMSE	r	ACC	POD	FAR
Seoul	D+1	0.84	8.16	34.27	13.50	0.81	57.00 (509/893)	89.03 (138/155)	59.05 (199/337)
	D+2	0.82	7.81	32.65	14.77	0.76	55.94 (499/892)	81.53 (128/157)	60.12 (193/321)
Busan	D+1	0.85	3.19	14.92	9.56	0.78	68.13 (607/891)	78.18 (86/110)	53.01 (97/183)
	D+2	0.82	2.94	13.81	10.50	0.73	65.73 (585/890)	71.30 (77/108)	57.46 (104/181)

**Table 2** Input features from the observations and CMAQ and WRF predictions applied to the DNN model (Lee et al. 2022)

Input type	Features
Observed parameters at a monitoring station	SO <sub>2</sub> , NO <sub>2</sub> , O <sub>3</sub> , CO, PM10, PM2.5, V, U, RN_ACC, RH, Td, Pa, Rad, and Ta
WRF output data	V, U, RN_ACC, RH, Pa, Ta, MH, 925 hPa_gpm, 925 hPa_V, 925 hPa_U, 850 hPa_gpm, 850 hPa_V, 850 hPa_U, 850 hPa_RH, 850 hPa_Ta, and Temp_850–925 hPa
CMAQ forecast	PM2.5
SO <sub>2</sub> (Sulfur dioxide), NO <sub>2</sub> (Nitrogen dioxide), O <sub>3</sub> (Ozone), CO (Carbon monoxide), PM10 and PM2.5 (Particulate matter with aerodynamic diameters less than 10 μm and 2.5 μm, respectively), V and U (horizontal velocities), RN_ACC (accumulative precipitation), RH (relative humidity), Td (dew point temperature), Pa (Pressure), Rad (Solar radiation), Ta (temperature), MH (Mixing height), gpm (geopotential height), and Temp_850–925 hPa (Potential temperature difference between 850 and 925 hPa)	
850 hPa and 925 hPa preceding the variables indicate the altitudes at which the variable was extracted from the three-dimensional WRF prediction	

the air masses transported to the forecast regions in South Korea. They classified the transport patterns into five clusters: south, local, long-northwest (NW), short-NW, and north. They also demonstrated that the back-trajectory patterns were closely related to the PM2.5 concentration in South Korea. To incorporate these back-trajectory clusters into the machine learning data, their occurrence probabilities were used based on the Euclidean distance (Ho et al. 2021).

### 3.2.2 Cosine Similarities of the Meteorological Variables

This study considered the meteorological patterns in Northeast Asia, which are key variables determining the LRT from China to South Korea, and the chemical formations of secondary organic and inorganic aerosols, such as sulfate and nitrate, accounting for > 60% of the air pollutants in South Korea, especially during PM episodes (Koo et al. 2018).

The high-pressure system over the Korean Peninsula has favorable meteorological conditions for the accumulation of air pollutants that can induce high-PM episodes (Lee et al. 2011a, b). The stagnant high-pressure patterns over the West Sea between East China and the Korean Peninsula also result in high-PM episodes in Korea (Koo et al. 2018). To reflect the synoptic patterns in the machine learning algorithm, the cosine similarities ( $\cos^t$ ) were calculated using the following equation (Ho et al. 2021).

$$\cos^t = \frac{A_i^t \cdot A_i^{clm-high}}{\|A_i^t\| \cdot \|A_i^{clm-high}\|} \quad (5)$$

where  $i$  and  $t$  indicate the  $i$ th grid point and  $t$ th time, respectively.

### 3.2.3 Spatial Distributions of Observed PM2.5 Concentrations and Meteorological Variables

In the CNN model, to reflect the influence of LRT, the PM2.5 concentrations and meteorological variables

obtained from the air quality and weather monitoring stations in Chain and South Korea (as shown in Fig. 2) were converted to two-dimensional distribution data using the kriging method.

The kriging method was used to grid variables from the observed point data to generate two-dimensional maps. The grid size was  $142 \times 122$  with grid dimension of 27 km, which was the same as the CMAQ grid. The kriging interpolation method for mapping the point values to grid values is an unbiased optimal sample estimation method based on the spatial distances between the measured monitoring stations and surrounding grid cells (Mahmoudabadi and Briggs 2016; Meng 2021).

The grid value  $Z(x_0)$  at  $x_0$  was estimated by a linear combination of the observed  $Z(x_i)$  for the neighboring  $n$  observation points.

$$Z(x_0) = \sum_{i=1}^n \lambda_i Z(x_i) \quad (6)$$

where  $\lambda_i$  is the weight of the observation point  $x_i$ . The weight depends on the distance between the grid and observation point, as well as spatial distribution of observed values over the grid domain, that is, the spatial covariance and variance.

The kriging equation and calculation of variance are as follows:

$$\sum_{i=0}^n \lambda_i \gamma(x_i, x_j) + \mu = \gamma(x_i, x), \sum_i^n \lambda_i = 1, (i = 1, 2, \dots, n) \quad (7)$$

$$\sigma_k^2 = \sum_{i=1}^n \lambda_i \gamma(x_i, x) - \gamma(x, x) + \mu \quad (8)$$

where  $\gamma(x_i, x_j)$  is the semi-variogram between observation points  $x_i$  and  $x_j$ ,  $\gamma(x_i, x)$  is the semi variogram between grid  $x_i$  and observation point  $x$ , and  $\gamma(x, x)$  is the semi variogram between grid  $x$  and observation point  $x$ .  $\mu$  is the Lagrange multiplier related to variance minimization. The interpolated grid values can be calculated considering a linear semi-variogram by estimating the semi-variogram and variance data



using the observed values and distances between the observation and grid points.

Figure 6 presents the two-dimensional distributions of observed PM2.5, u-velocity, v-velocity, pressure, relative humidity, and temperature over Northeast Asia using the kriging method with a linear semi-variogram. The 6-h average of observed PM2.5 and meteorological variables at the monitoring stations in South Korea and China were applied to generate the two-dimensional maps. Figure 6 shows that a high concentration of PM2.5 exists over the West Sea. The generated maps were then used as the key input features in the CNN models to reflect the LRT from China to the forecast regions in South Korea.

## 4 Machine Learning Algorithms

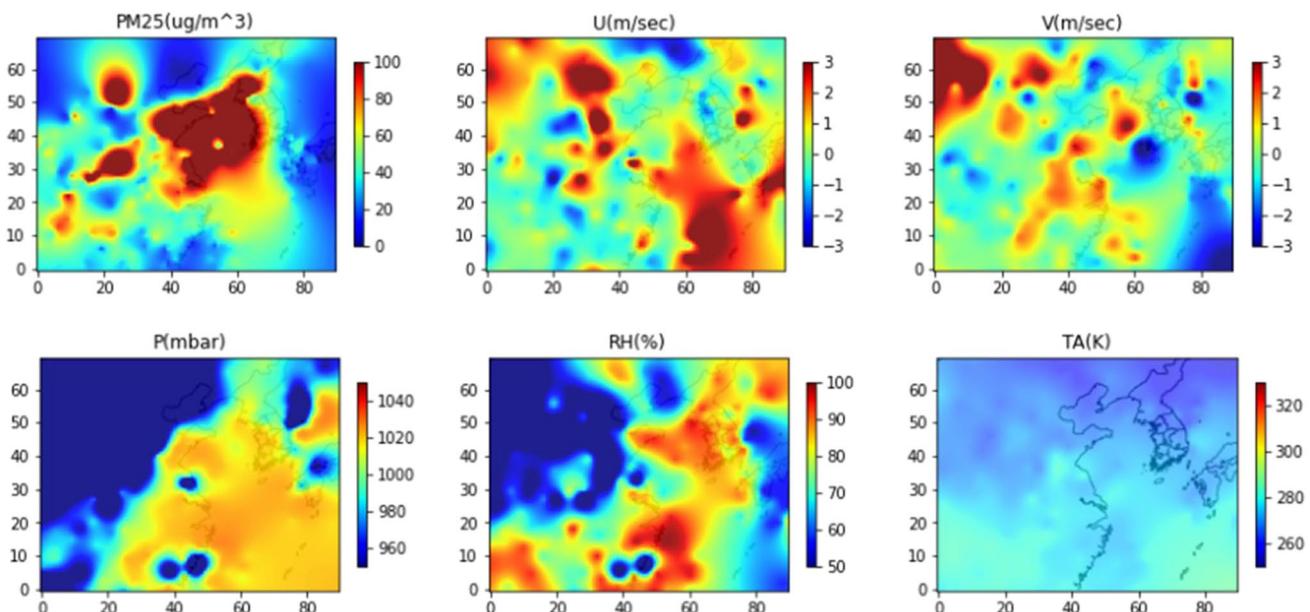
Advanced machine learning models have been applied to PM2.5 forecasting by developing methods that reflect the transport and formation characteristics in the corresponding algorithms. The feed forward neural network (Bai et al. 2016; Lee et al. 2022; Lightstone et al. 2017, 2021; Perez and Reyes 2006) and RNN (Ho et al. 2021; Kim et al. 2019; Zhao et al. 2019; Zhang et al. 2022) models are the basic algorithms used to simulate the temporal variations in PM2.5 concentrations by depicting the formation characteristics in the forecast region. CNN and graph neural network (GNN) models have also been developed using the observed data from monitoring stations in the forecast region and neighboring areas to directly reflect transport characteristics (Park et al. 2020; Mao and Lee 2019; Qi et al. 2019;

Wen et al. 2019; Wan et al. 2019; Zhang et al. 2022; Gilik et al. 2022). These models can effectively represent the spatial correlation between the forecasting region and emission sources located upwind. Hybrid networks combining CNN and GNN with the temporal characteristics of LSTM, such as CNN-LSTM and GNN-LSTM, can reflect temporal variations in the forecast region as well as the transport from the upwind region. Theoretically, the network structure of convolutional LSTM (ConvLSTM) makes it an ideal algorithm to simultaneously incorporate transport and formation characteristics; however, it cannot accurately predict these characteristics for more than 12 h (Wang et al. 2021).

We developed a forecasting model by applying the most traditional algorithms—DNN, RNN, and CNN—to real-time forecasting operations using a testbed as shown in Fig. 2. In this section, we describe the methodology adopted in developing the proposed machine learning algorithms. As the findings of DNN (Lee et al. 2022) and RNN (Ho et al. 2021) models have already been reported in existing journals, we describe them briefly with a focus on the main features.

### 4.1 DNN

Multilayer perceptron, that is, DNN, is a typical feedforward model that extracts the complex and non-linear relationships between the input features and target PM2.5 forecast. Lee et al. (2022) and Lightstone et al. (2017, 2021) applied DNN models with CMAQ and WRF data as the key input features. Their results showed that the DNN modeling using forecast



**Fig. 6** Two dimensional distributions of observed PM2.5 concentrations and meteorological variables at 10–02–2020 00:00 LST

variables could improve the forecasting accuracy compared to CMAQ predictions. The DNN analyzed for real-time operation in this study was based on the DNN by Lee et al. (2022). The key parameters are presented herein, and further details can be found in the study by Lee et al. (2022).

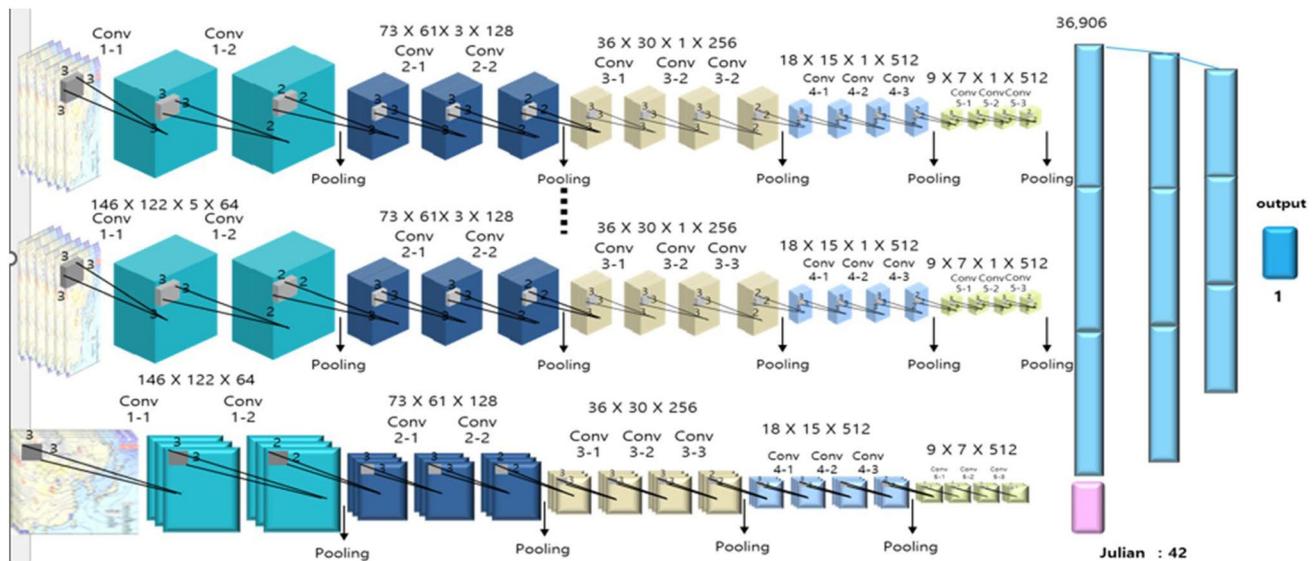
The input features used in the DNN were observed air quality and meteorological data at the monitoring stations as well as the predicted CMAQ and WRF variables, as listed in Table 2. The upper meteorological variables at 850 and 925 hPa vertical levels predicted by the WRF were included as key features, and five stacked layers of the DNN were considered in this analysis.

## 4.2 RNN

In contrast to the feedforward DNN, an RNN facilitates feedback, and it is a state-of-the-art algorithm for generating sequential data. The model framework of RNN comprises two connected RNNs with encoder and decoder networks. Ho et al. (2021) had previously analyzed the RNN network structure for PM2.5 forecasting in South Korea and suggested the use of double-stacked layers with GRU cells combined with input features in the forecasting system (Table 3). Further details on the network structure and optimal hyper parameters of the RNN can be found in Ho et al. (2021).

**Table 3** Input features from the observations as well as CMAQ and WRF predictions used in the NN model (Ho et al. 2021)

Input type	Features
Observation at a monitoring station	SO <sub>2</sub> , NO <sub>2</sub> , O <sub>3</sub> , CO, PM10, PM2.5 V, U, RH, Td, and Ta
WRF output data	V, U, RH, Ta, Cosine similarity of Z, T, RH, U, V, W, and Back-trajectory clustering probability
CMAQ forecast	PM2.5
Please refer to Table 1 for the description of these variables	



**Fig. 7** Structure and parameters of the CNN network for PM2.5 forecasting

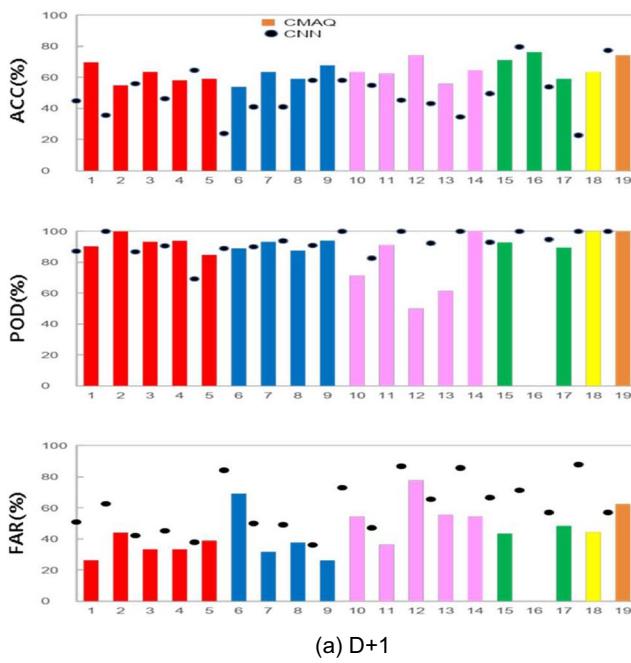
**Table 4** Input features from the observations and CMAQ and WRF predictions used in the CNN

Input type	Features
Spatial grid data	Maps of observed PM2.5, V, U, RH, and Ta Map of forecasted PM2.5 using CMAQ
Julian date data	Membership function of Julian date to reflect the monthly change characteristics of PM2.5 concentration (Lee et al. 2022)

### 4.3 CNN

CNN, a class of artificial neural networks that have been dominantly used in various computer vision applications, can automatically and adaptively learn spatial hierarchies of features through backpropagation by using multiple building blocks, such as convolution, pooling, and fully connected layers. The CNN extracts the required feature from the spatial input data, and this algorithm can efficiently represent the transport characteristics from the two-dimensional maps of the PM2.5 concentrations and meteorological variables as shown in Fig. 6.

Soh et al. (2018) used the CNN and LSTM to extract spatiotemporal relationships between the monitoring stations in forecasting PM2.5 concentrations in Taiwan. Han et al. (2022) developed a hybrid deep learning framework that combined the CNN with LSTM to describe the spatiotemporal characteristic of PM2.5 transport from the neighboring stations in Hong Kong and Beijing. Wang et al. (2021) also demonstrated that the ConvLSTM network could effectively represent the complex spatiotemporal correlations in forecasting PM2.5 concentrations in Beijing. They used the observed air quality and meteorological data at the monitoring stations with stationary data, such as normalized difference vegetation index and traffic networks.

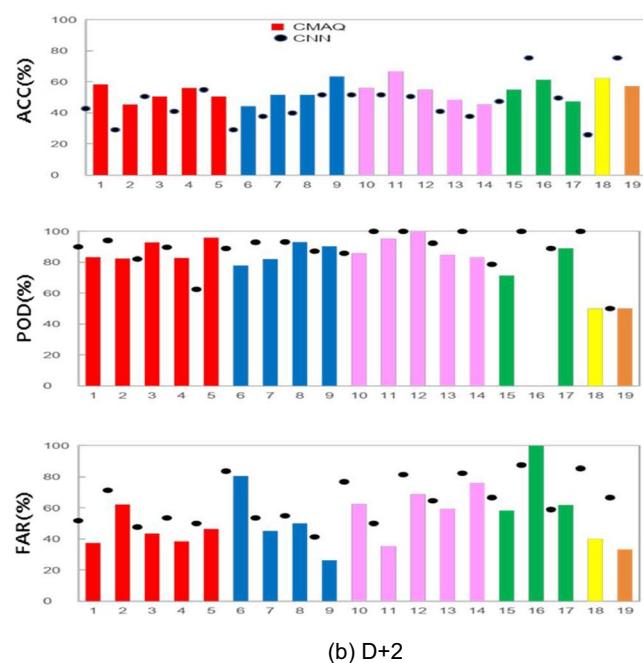


The CNN in this study was developed by applying two-dimensional spatial distribution maps of observed meteorological variables and PM2.5 concentrations forecasted by the CMAQ over Northeast Asia as machine learning input data to simulate the effects of LRT from China on the PM2.5 concentrations in South Korea.

The DNN and RNN described in the previous sections cannot directly reflect the influence of pollutant transport in the surrounding areas, especially from China to South Korea, because they only used the measurement and forecast data within the forecasting region. A model based on CNN combined with DNN in the fully connected layers was developed to reflect the LRT from China as well as characterize the local PM2.5 formation in the forecast region considered in this study.

The PM2.5 forecast algorithm was developed using the VGG-16 proposed by Simonyan and Zisserman (2014). The original VGG-16 has three channels of 224 length  $\times$  224 width. The CNN model for PM2.5 forecast was modified into five channels of 146 width  $\times$  122 length. The five channels in input data, which are consecutive time series of spatial data, can play an important role in representing the sequence module in PM2.5 forecasting.

The network and parameters of the CNN forecast model based on the VGG network structure are illustrated in Fig. 7. In the fully connected layer, the Julian data were combined



**Fig. 8** Comparison of the daily average PM2.5 concentration forecasted by the CNN and CMAQ models for 19 forecast regions (Bars indicate the CNN forecast, whereas the black dots represent the CMAQ forecast. Numbers represent forecast regions in

Fig. 1: 1 = SU, 2 = IC, 3 = GGN, 4 = GGS, 5 = GWW, 6 = DJ, 7 = SJ, 8 = CCN, 9 = CCS, 10 = BS, 11 = DG, 12 = US, 13 = GSN, 14 = GSS, 15 = GJ, 16 = JLN, 17 = JLS, 18 = GWE, 19 = JJ)

to reflect the temporal patterns in the forecast regions. Table 4 lists the input variables used in the CNN.

The PM2.5 forecast was performed for 19 regions using the proposed CNN model at 15:00 LST. The training data were used from January 01, 2015 to December 31, 2019, and validation data were from January 1, 2020 to February 28, 2021.

The performance of the CNN model for 19 forecast regions was compared with that of the CMAQ model, as shown in Fig. 8, and the categorical evaluation results are

provided in Table 5. The CNN model exhibits better performance with an increase in the ACC and decrease in the FAR in almost all forecast regions. The CMAQ forecast indicates a higher POD and FAR than the CNN due to the over-prediction of the CMAQ forecast, as described in Sect. 3.1.2. The CNN forecast can correct for the over-prediction of the CMAQ forecast, which shows that the overall forecasting performance improved, although the POD decreased slightly. This finding is similar to those

**Table 5** Categorical performance of the CNN model and CMAQ for the 19 forecast regions in South Korea

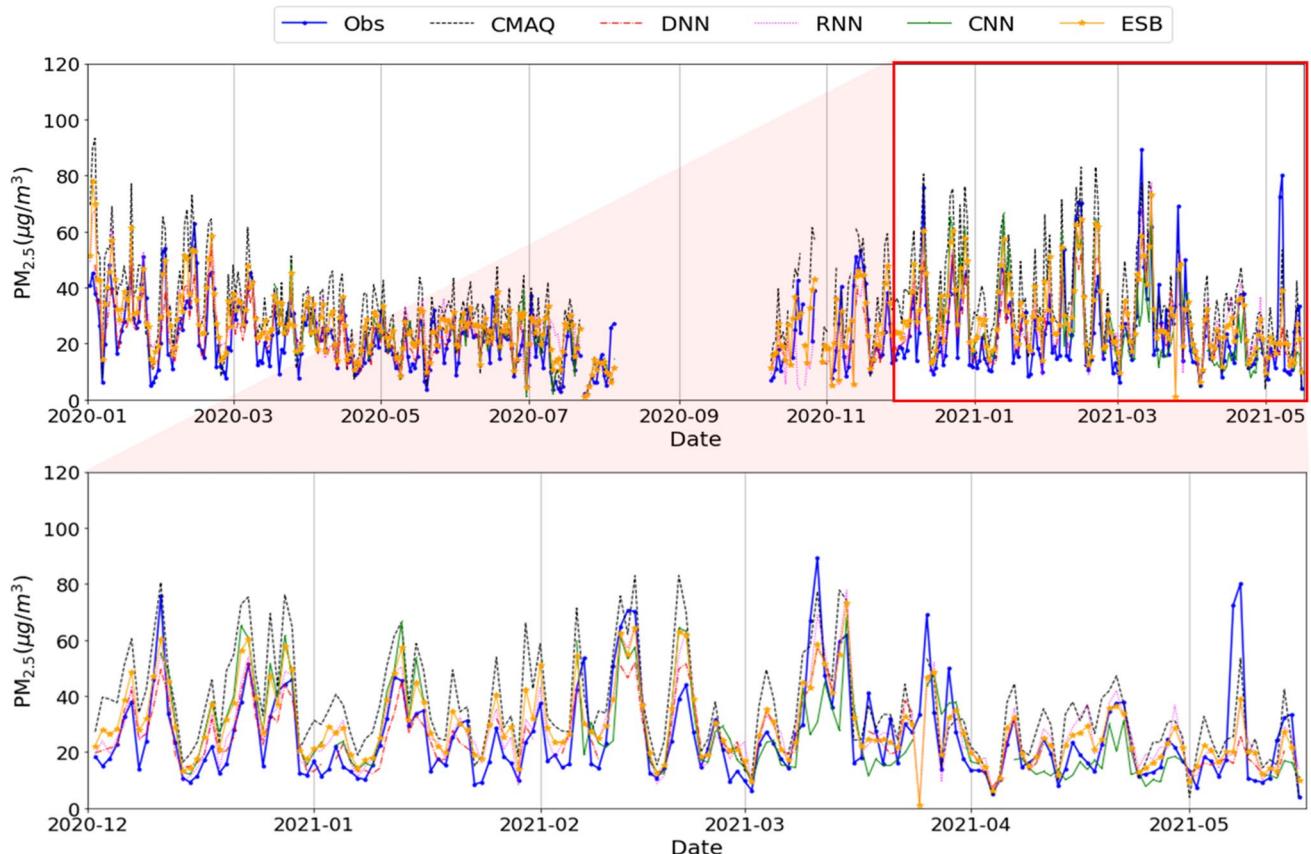
Region	Day	CNN				CMAQ			
		ACC	POD	FAR		ACC	POD	FAR	
SU	D+1	69.8 (67/96)	90.3 (28.31)	26.3 (10/38)		44.8 (43/96)	87.1 (27/31)	50.9 (28/55)	
	D+2	58.3 (56/96)	83.3 (25/30)	37.5 (15/40)		42.7 (41/96)	90 (27/30)	51.8 (29/56)	
IC	D+1	54.8 (51/93)	100 (19/19)	44.1 (15/34)		35.5 (33/93)	100 (19/19)	62.7 (32/51)	
	D+2	45.2 (42.93)	82.4 (14/17)	62.2 (23/37)		29 (27/93)	94.1 (16/17)	71.4 (40/56)	
GGN	D+1	63.4 (59/93)	93.3 (28/30)	33.3 (14/42)		55.9 (52/93)	86.7 (26/30)	42.2 (19/45)	
	D+2	50.5 (47/93)	92.9 (26/28)	43.5 (20/46)		50.5 (47/93)	82.1 (23/28)	47.7 (21/44)	
GGS	D+1	58.1 (54/93)	93.8 (30/32)	33.3 (15/45)		46.2 (43/93)	90.6 (29/32)	45.3 (24/53)	
	D+2	55.9 (52.93)	82.8 (24/29)	38.5 (15/39)		40.9 (38/93)	89.7 (26/29)	53.6 (30/56)	
GWW	D+1	59.1 (55/93)	84.6 (22/26)	38.9 (14/36)		64.5 (60/93)	69.2 (18/26)	37.9 (11/29)	
	D+2	50.5 (47/93)	95.8 (23/24)	46.5 (20/43)		54.8 (51/93)	62.5 (15/24)	50 (15/30)	
DJ	D+1	53.8 (50/93)	88.9 (8/9)	69.2 (18/26)		23.7 (22/93)	88.9 (8/9)	84.3 (43/51)	
	D+2	44.1 (41/93)	77.8 (7/9)	80.6 (29/36)		29 (27/93)	88.9 (8/9)	83.7 (41/49)	
SJ	D+1	63.4 (59/93)	93.3 (28/30)	31.7 (13/41)		40.9 (38/93)	90 (27/30)	50 (27/54)	
	D+2	51.6 (48/93)	82.1 (23/28)	45.2 (19/42)		37.6 (35/93)	92.9 (26/28)	53.6 (30/56)	
CCN	D+1	59.1 (55/93)	87.5 (28/32)	37.8 (17/45)		40.9 (38/93)	93.8 (30/32)	49.2 (29/59)	
	D+2	51.6 (48/93)	93.1 (27/29)	50 (27/54)		39.8 (37/93)	93.1 (27/29)	55 (33/60)	
CCS	D+1	67.7 (63/93)	93.9 (31/33)	26.2 (11/42)		58.1 (54/93)	90.9 (30/33)	36.2 (17/47)	
	D+2	63.4 (59/93)	90.3 (28/31)	26.3 (10/38)		51.6 (48/93)	87.1 (27/31)	41.3 (19/46)	
BS	D+1	63.4 (59/93)	71.4 (5/7)	54.5 (6/11)		58.1 (54/93)	100 (7/7)	73.1 (19/26)	
	D+2	55.9 (52/93)	85.7 (6/7)	62.5 (10/16)		51.6 (48/93)	85.7 (6/7)	76.9 (20/26)	
DG	D+1	62.4 (58/93)	91.3 (21/23)	36.4 (12/33)		54.8 (51/93)	82.6 (19/23)	47.2 (17/36)	
	D+2	66.7 (62/93)	95.2 (20/21)	35.5 (11/31)		51.6 (48/93)	100 (21/21)	50 (21/42)	
US	D+1	74.2 (69/93)	50 (2/4)	77.8 (7/9)		45.2 (42/93)	100 (4/4)	86.7 (26/30)	
	D+2	54.8 (51/93)	100 (5/5)	68.8 (11/16)		50.5 (47/93)	100 (5/5)	81.5 (22/27)	
GSN	D+1	55.9 (52/93)	61.5 (8/13)	55.6 (10/18)		43 (40/93)	92.3 (12/13)	65.7 (23/35)	
	D+2	48.4 (45/93)	84.6 (11/13)	59.3 (16/27)		40.9 (38/93)	92.3 (12/13)	64.7 (22/34)	
GSS	D+1	64.5 (60/93)	100 (5/5)	54.5 (6/11)		34.4 (32/93)	100 (5/5)	85.7 (30/35)	
	D+2	45.2 (42/93)	83.3 (5/6)	76.2 (16/21)		37.6 (35/93)	100 (6/6)	82.4 (28/34)	
GJ	D+1	71 (66/93)	92.9 (13/14)	43.5 (10/23)		49.5 (46.93/)	92.9 (13/14)	66.7 (26/39)	
	D+2	54.8 (51/93)	71.4 (10/14)	58.3 (14/24)		47.3 (44/93)	78.6 (11/14)	66.7 (22/33)	
JLN	D+1	76.3 (71/93)	0 (0/2)	-1 (0/2)		79.6 (74/93)	100 (2/2)	71.4 (5/7)	
	D+2	61.3 (57/93)	0 (0/1)	100 (3/3)		75.3 (70/93)	100 (1/1)	87.5 (7/8)	
JLS	D+1	59.1 (55/93)	89.5 (17/19)	48.5 (16/33)		53.8 (50/93)	94.7 (18/19)	57.1 (24/42)	
	D+2	47.3 (44/93)	88.9 (16/18)	61.9 (26/42)		49.5 (46/93)	88.9 (16/18)	59 (23/39)	
GWE	D+1	63.4 (59/93)	100 (5/5)	44.4 (4/9)		22.6 (21/93)	100 (5/5)	87.8 (36/41)	
	D+2	62.4 (58/93)	50 (3/6)	40 (2/5)		25.8 (24/93)	100 (6/6)	85.4 (35/41)	
JJ	D+1	74.2 (69/93)	100 (3/3)	62.5 (5/8)		77.4 (72/93)	100 (3/3)	57.1 (4/7)	
	D+2	57 (53/93)	50 (2/4)	33.3 (1/3)		75.3 (70/93)	50 (2/4)	66.7 (4/6)	

obtained using the DNN and RNN models (as described in Sects. 4.1 and 4.2).

## 5 Test Bed in Real-Time Forecasting

The real-time machine learning system shown in Fig. 2 was operated to examine the data flow using machine learning models and obtain insights into the forecast in real-time. The input data from air quality and weather monitoring stations and the forecasted results of the CMAQ and WRF were collected in real-time. The two-dimensional maps of PM<sub>2.5</sub> concentrations and meteorological variables (u and V wind velocities, relative humidity, pressure, and temperature) were generated using the kriging method. The additional secondary data of cosine similarities and anomalies at 950 and 850 mbar were calculated using the interface programs. The DNN, RNN, and CNN models were operated based on the primary and secondary data twice a day at 03:00 and 15:00 LST, respectively. The results of the DNN, RNN, CNN, and CMAQ forecast were averaged and used as the ESB forecast.

Figure 9 shows the time-series comparison of the forecasts by the CMAQ, DNN, RNN, CNN, and ESB models for D + 1 from January 1, 2020 to April 31, 2021 by operating the real-time machine learning forecasting system in Seoul. The red boxes in Figs. 9 and 11 represent a period during which five models were successfully operated simultaneously. The ESB forecast is the average of the CMAQ, DNN, RNN, and CNN predictions. The blank period in Fig. 9 is the period during which the real-time forecast system could not operate due to system failure. The scatter plots of the PM<sub>2.5</sub> concentrations for D + 1 forecasted by the CMAQ model and various machine learning algorithms are presented in Fig. 10. The comparisons of the real-time observations with the corresponding model forecasts for D + 2 based on the time-series and scatter plots of PM<sub>2.5</sub> concentrations are presented in Figs. 11 and 12, respectively. The CMAQ and machine learning models could depict the time variations of the observation for D + 1 and D + 2, respectively. The results of machine learning models are often better than CMAQ, but the opposite is sometimes seen. To explain this, we have distinguished two opposite cases. The first is the case where the



**Fig. 9** Time series comparison of PM<sub>2.5</sub> forecasts by the CMAQ, DNN, RNN, CNN, and ESB models with the real-time observations for D + 1 in Seoul from January 1, 2020 to April 31, 2021

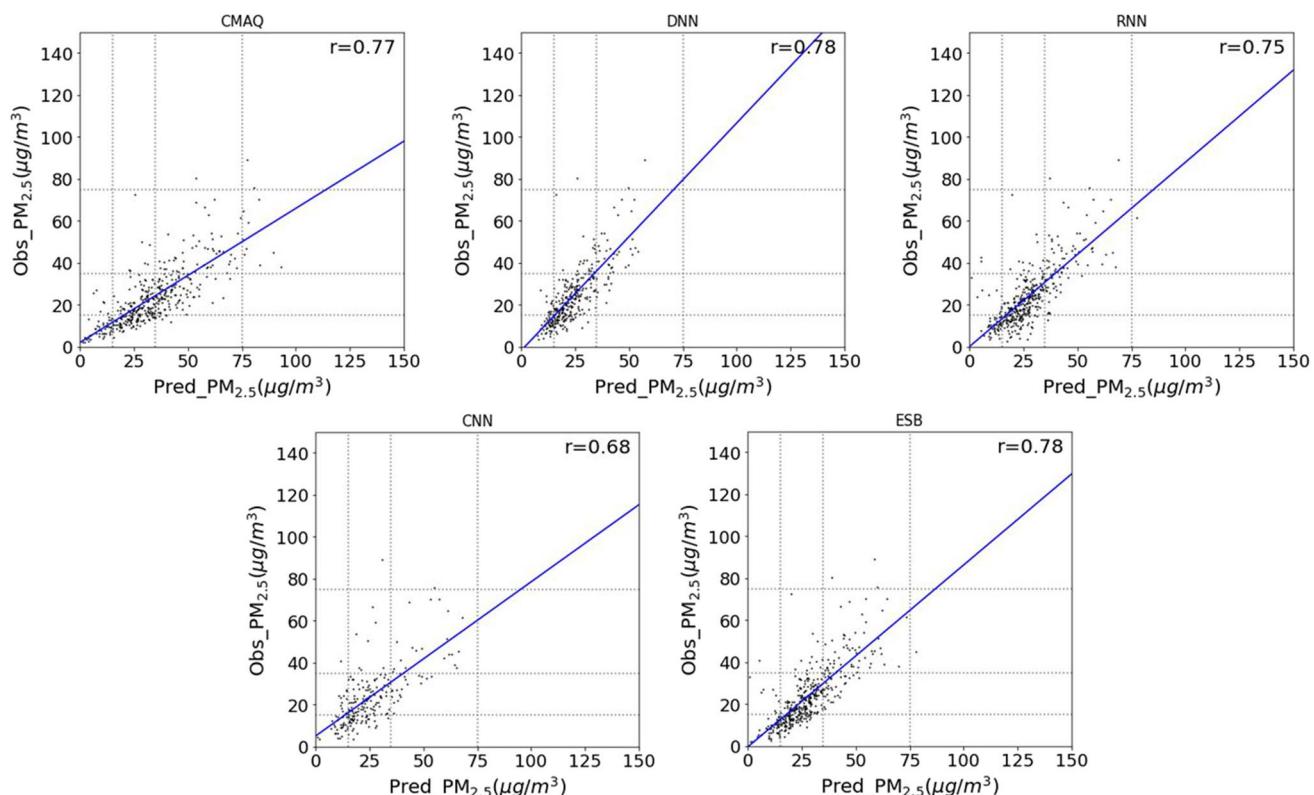
CMAQ concentration is higher than the measured value, as can be seen in box "A" in Fig. 11; the second is the case where the measured value is high, but the CMAQ concentration is low, as shown in box "B". In machine learning, measurements are used as historical input data previous to forecasting while CMAQ is a numerical prediction model, in which concentrations calculated for each future forecast time are used as input data. Therefore, for the machine learning model that predicts future concentrations, the CMAQ forecast value has a higher weight in determining the model's forecast value as it moves away from the forecast. In both cases, the machine learning model sits between the CMAQ predictions and measurements. This shows that the machine learning model can correct the over- or under-simulated results of the CMAQ.

The results of the statistical and categorical analyses for D+1 and D+2 forecasts are summarized in Table 6.

All the machine learning models developed in this study and the CMAQ forecast could successfully simulate the time-series variation of observed PM2.5 concentrations with IOAs of 0.78 (CMAQ), 0.85 (DNN and RNN), 0.81 (CNN), and 0.86 (ESB) for D+1. The ACC also improved in the proposed models from 51.1% in CMAQ to 70.8% (DNN), 69.2% (RNN), 64.0% (CNN), and 68.1% (ESB). The NMB improved from 42.61% in CMAQ to -1.09% in

DNN, 3.16% in RNN, 6.58% in CNN, and 17.17% in ESB, as shown in Table 6. In Table 6, the r in CMAQ is higher than that of RNN while ACC is lower than that of RNN. This is because the evaluation was divided into quantitative statistical metrics, such as r and RMSE, and categorical evaluation using ACC and POD, which evaluates interval accuracy by classifying PM2.5 concentration into 4 categorical classes. The number of data in each class is largely different. For example, there are many data in the good and moderate PM2.5 classes, while the frequency of data for bad or very bad classes consists of less than 10% in the total number of case days. If r is good but ACC is low, it means that the evaluation for good and moderate classes is good, but the degree of matching is poor in the bad and very bad classes. In overall sense, the four machine learning algorithms (DNN, RNN, CNN, and ESB) could correct the over-prediction of the CMAQ forecast by decreasing the NMB and increasing the IOA.

The CMAQ forecast for D+1 exhibited the best performance in detecting high-PM2.5 episodes with 90.1% POD, but the practical application of this model is difficult owing to the high FAR of 62.6%. This implies that nine out of ten high-PM2.5 episodes can be predicted, but more than six of these forecasts are false alarms. Actual field forecasting requires a lower FAR while maintaining



**Fig. 10** Comparison of the scatter plots of PM2.5 forecasts by the CMAQ, DNN, RNN, CNN, and ESB models with the real-time observations for D+1 in Seoul from January 1, 2020 to April 31, 2021

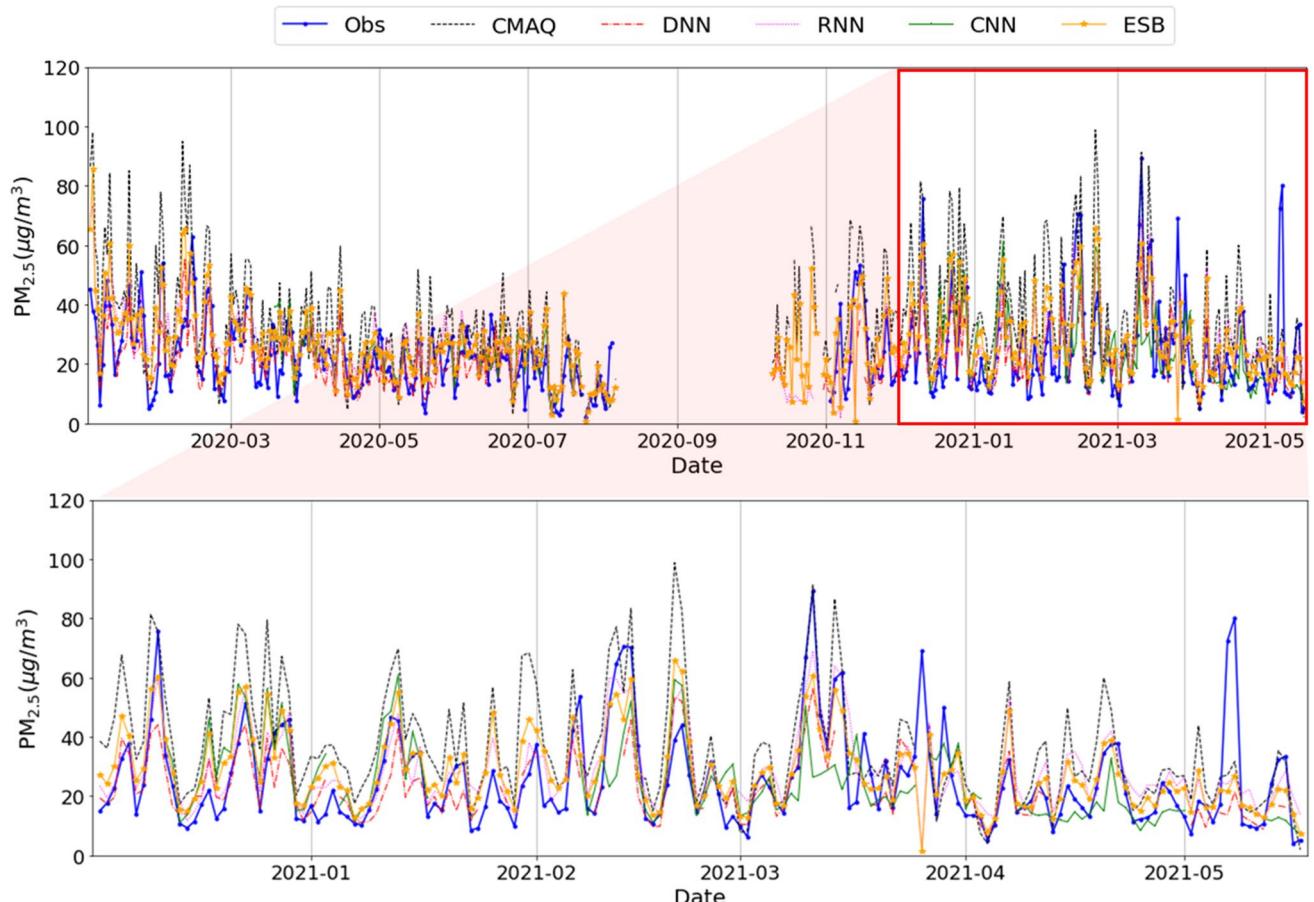
the POD at a high level. The machine learning methods of DNN, RNN, CNN, and ESB successfully decreased the FAR to 19.5, 32.9, 47.5, and 37.2% with PODs of 60.0, 79.2, 70.0, and 81.9%, respectively. The performance of the second day's forecast ( $D+2$ ) presented in Figs. 11 and 12 also exhibited a similar trend to the first day's forecast ( $D+1$ ). The performance evaluation metrics for all 19 forecast regions, including Seoul, are presented in Figs. 13, 14, 15, 16. Although differences were observed in the forecasting regions, the machine learning models could effectively increase the prediction accuracy and decrease the FAR, compared with the CMAQ models except the JJ region. This is due to underfitting problem in the machine learning model in the Jeju Island (JJ region) because the number of high concentration cases is relatively low compared with other regions.

The PODs in the forecasts using machine learning were slightly lower than that of the CMAQ forecasts; however, the proposed models were more practical for field application in PM2.5 forecasting as they improved the overall accuracy of forecasting by increasing the temporal

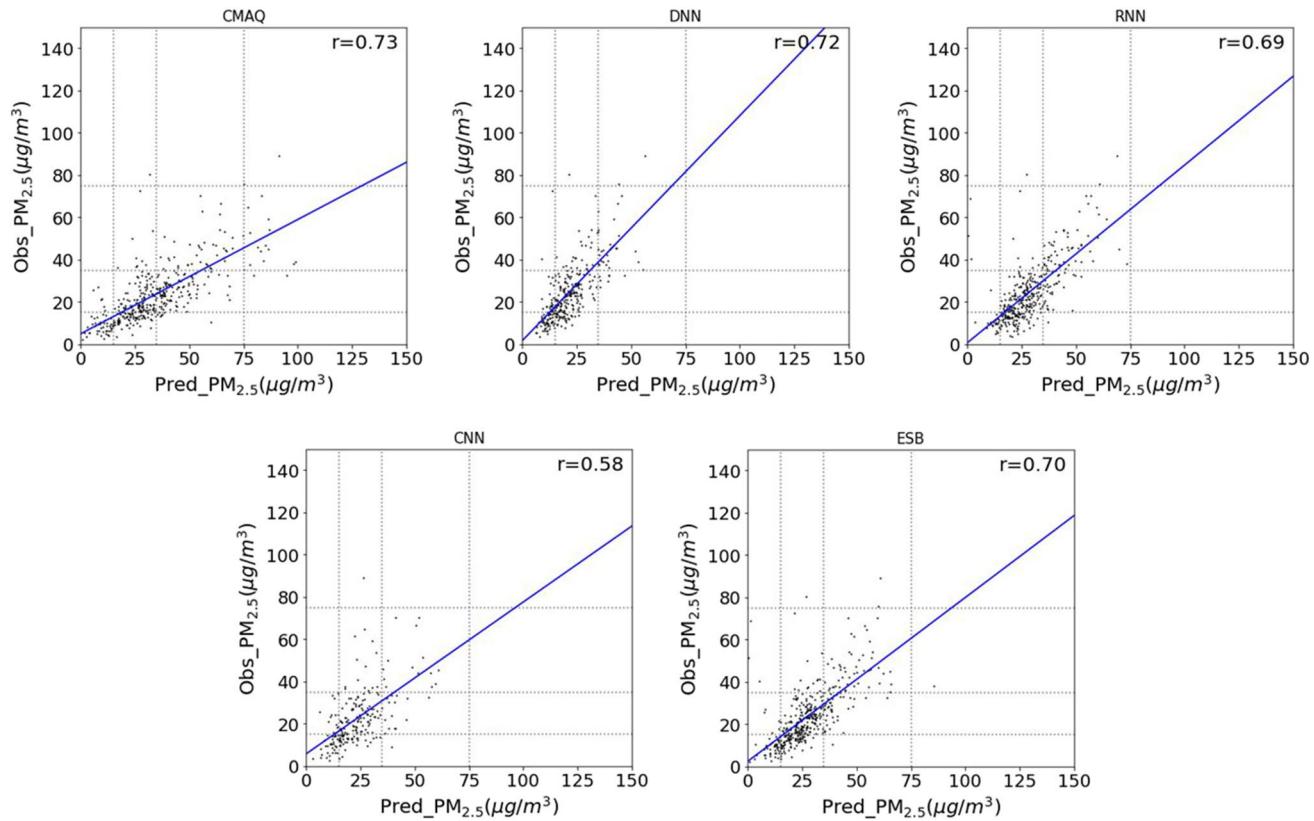
correlations with observations and decreasing the FAR. Future studies on machine learning models using CMAQ forecast data should focus on developing a method to decrease the FAR, while maintaining the POD level of the CMAQ forecast.

## 6 Conclusions

Various machine learning methods were developed to improve the uncertainty of numerical forecasting based on the CTM currently used in PM2.5 forecasting. DNN, RNN, CNN, and ESB algorithms were proposed and their performance validated in a real-time forecasting operation. The primary data for the machine learning models were the observed air quality and meteorological data at the monitoring stations in China and South Korea. The forecasted PM2.5 concentrations and meteorological data using CMAQ and WRF, which were incorporated to identify the local PM2.5 formation in the forecast region, as well as to simulate the pollutant transport from the outside regions,



**Fig. 11** Time series comparison of PM2.5 forecasts by the CMAQ, DNN, RNN, CNN, and ESB models with the corresponding observations for  $D+2$  in Seoul from January 1, 2020, to April 31, 2021



**Fig. 12** Comparison of the scatter plots of PM2.5 forecasts by the CMAQ, DNN, RNN, CNN, and ESB models with the observations for D + 1 in Seoul from January 1, 2020 to April 31, 2021

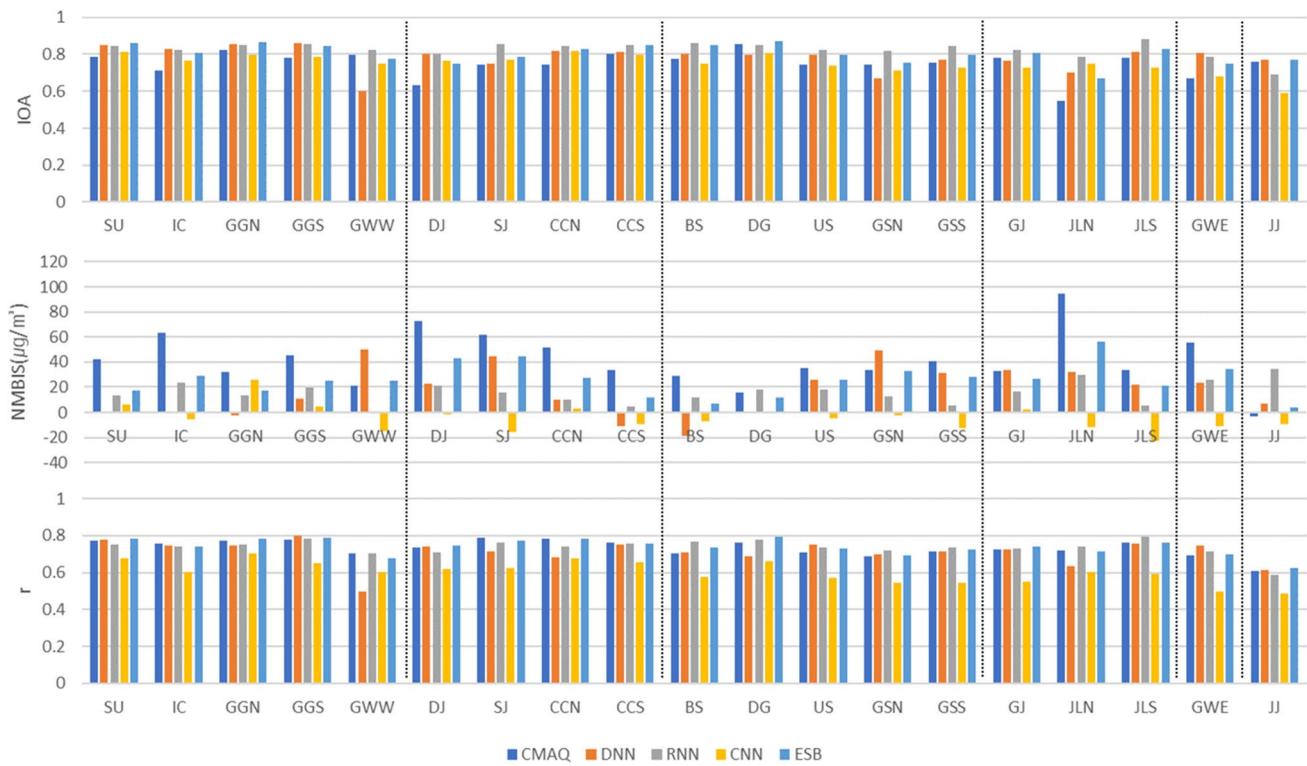
**Table 6** Results of the statistical and categorical performance analyses for PM2.5 forecasts by the CMAQ, DNN, RNN, CNN, and ESB models with real-time observations for D + 1 and D + 2 in Seoul from January 1, 2020 to April 31, 2021

	classification	IOA	MBIAS	NMB	RMSE	r	ACC	POD	FAR
D + 1	CMAQ	0.78	10.03	42.61	14.65	0.77	51.1 (214/419)	90.1 (64/71)	62.6 (171/107)
	DNN	0.85	-0.26	-1.09	8.63	0.78	70.8 (233/329)	60.0 (33/55)	19.5 (41/8)
	RNN	0.85	3.16	13.29	9.75	0.75	69.2 (287/415)	79.2 (57/72)	32.9 (85/28)
	CNN	0.81	1.55	6.58	11.02	0.68	64.0 (142/222)	70.0 (21/30)	47.5 (40/19)
	ESB	0.86	4.05	17.17	9.68	0.78	68.1 (288/423)	81.9 (59/72)	37.2 (94/35)
D + 2	CASE4	0.74	10.76	45.88	16.58	0.73	48.6 (198/407)	84.8 (56/66)	66.3 (166/110)
	DNN1	0.80	-3.08	-12.55	9.99	0.72	68.1 (220/323)	43.9 (25/57)	21.9 (32/7)
	RNN1	0.80	3.81	15.64	10.94	0.69	66.1 (252/381)	82.1 (55/67)	38.9 (90/35)
	CNN1	0.74	0.83	3.67	11.34	0.58	62.9 (139/221)	50.0 (14/28)	51.7 (29/15)
	ESB1	0.81	3.70	15.67	10.98	0.70	66.4 (273/411)	75.4 (52/69)	42.2 (90/38)

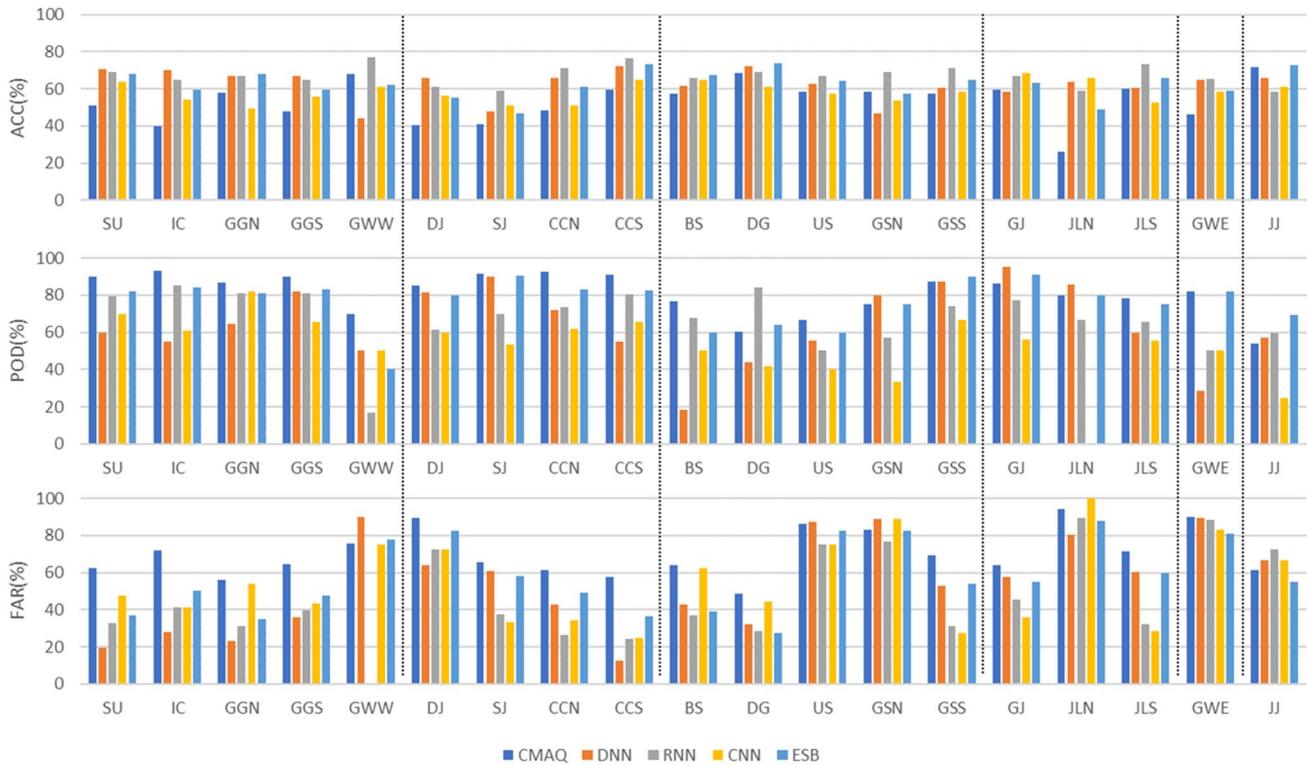
were used as the primary input features in the machine learning algorithms. The secondary data in these models were used to characterize the meteorological patterns over Northeast Asia, which had an important influence on the local PM2.5 formation, and to reflect the LRT of PM2.5 from China to South Korea by the northwest wind. Two-dimensional maps of the observed PM2.5 concentrations and meteorological variables were generated using the

kriging method and applied in the CNN models to reflect the LRT.

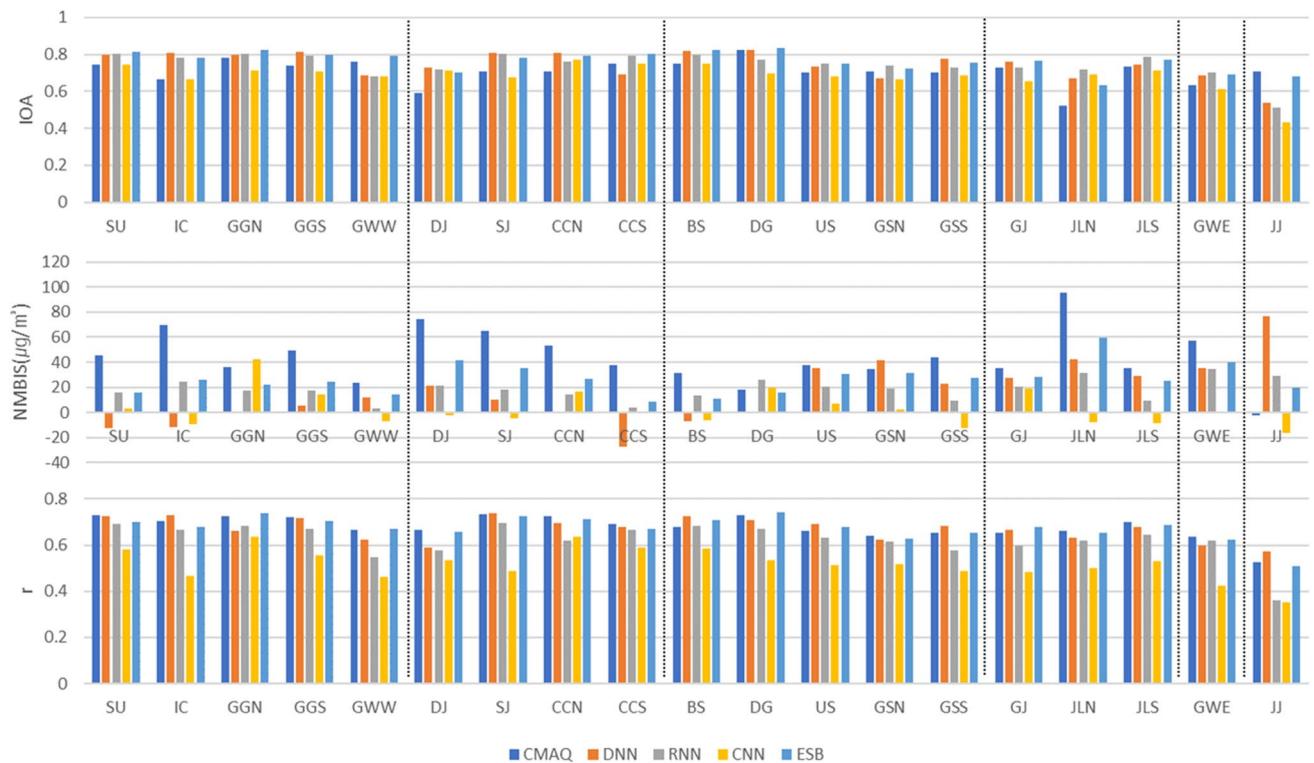
The real-time machine learning forecasting system was finally developed, and the performance of the machine learning models was evaluated for 19 forecasting regions from January 2020 to April 2021. According to the real-time operation analysis, the machine learning algorithms could effectively reduce the bias in the CMAQ forecast by



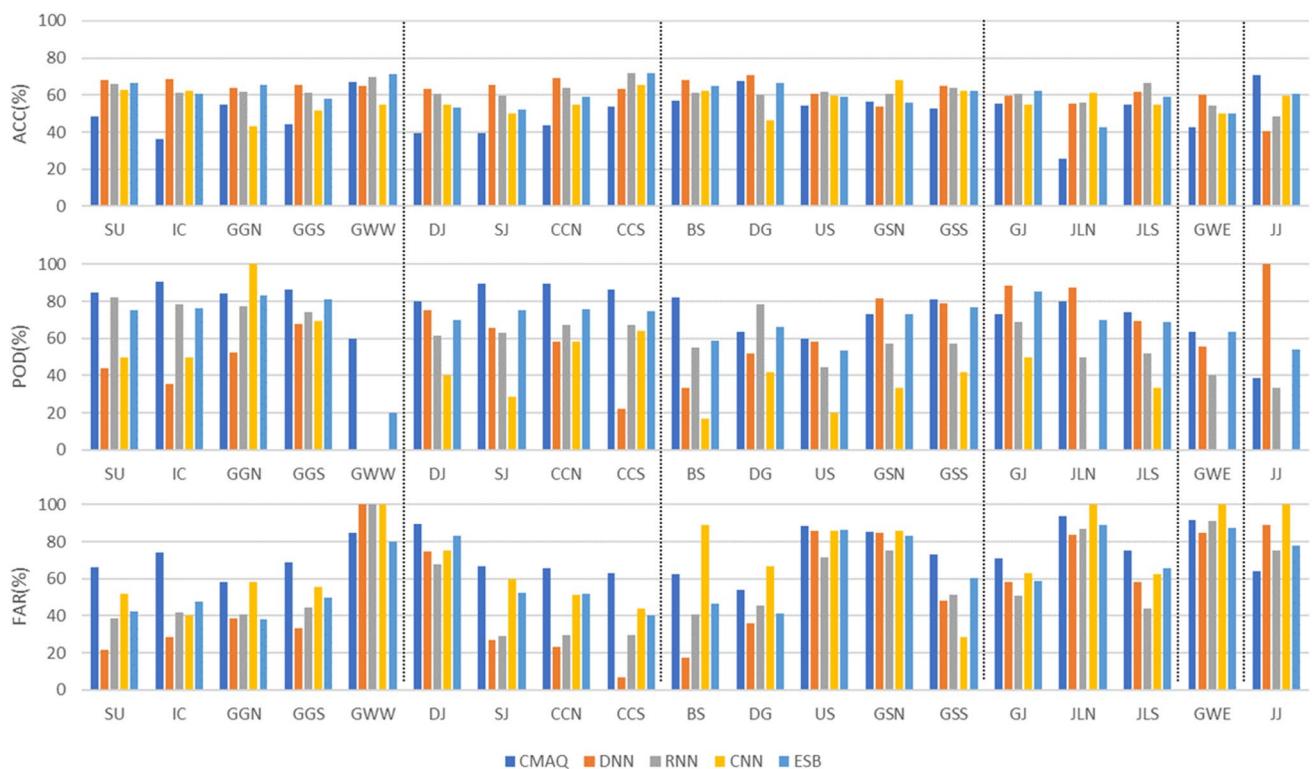
**Fig. 13** Statistical performance of PM2.5 forecasts with the CMAQ, DNN, RNN, CNN, and ESB models based on real-time observations for D+1 in the 19 forecast regions from January 1, 2020 to April 31, 2021



**Fig. 14** Categorical performance of PM2.5 forecasts with the CMAQ, DNN, RNN, CNN, and ESB models based on real-time observations for D+1 in the 19 forecast regions from January 1, 2020 to April 31, 2021



**Fig. 15** Statistical performance of PM2.5 forecasts with the CMAQ, DNN, RNN, CNN, and ESB models based on real-time observations for D+2 in the 19 forecast regions from January 1, 2020 to April 31, 2021



**Fig. 16** Categorical performance of PM2.5 forecasts with the CMAQ, DNN, RNN, CNN, and ESB models based on real-time observations for D+1 in the 19 forecast regions from January 1, 2020 to April 31, 2021

improving the accuracy and reducing the FAR, even though the POD was slightly lower than that of the CMAQ model. The ESB model, which is an average of the DNN, RNN, CNN, and CMAQ forecasting model, demonstrates application potential of the proposed system in field forecasting.

Future studies should focus on improving the proposed network-based algorithms and input data to ensure a high POD and ACC while maintaining a low FAR.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s13143-023-00314-8>.

**Acknowledgements** This study was supported in part by the "Experts Training Graduate Program for Particulate Matter Management" from the Ministry of Environment and by the "Development of Short-Term Prediction Tools for Particulate Matter Using Artificial Intelligence (V)" (project number: NIER-2022-04-02-068) from the National Institute of Environmental Research in South Korea.

**Author's Contributions** Youn-Seo Koo: supervision, writing—original draft, review & editing; Hee-Yong Kwon and SukHyun Yu: CNN design and test; Hyosik Bae: test bed development and real-time operation; Hui-Young Yun: test-bed operation, writing -original draft, review & editing, data analysis; Dae-Ryun Choi: CMAQ forecasting; Kyung-Hui Wang: writing—original draft, review & editing; Ji-Seok Koo: Kriging data.; Jae-Bum Lee, Min-Hyeok Choi, and Jeong-Beom Lee: test bed management and evaluation.

**Data Availability** Not applicable.

**Code Availability** Not applicable.

## References

- Anderberg, M.R.: Cluster Analysis for Applications. Academic Press, New York (2014)
- Bai, Y., Li, Y., Wang, X., Xie, J., Li, C.: Air pollutants concentrations forecasting using back propagation neural network based on wavelet decomposition with meteorological conditions. *Atmos. Pollut. Res.* **7**(3), 557–566 (2016). <https://doi.org/10.1016/j.apr.2016.01.004>
- Batagelj, V.: Generalized ward and related clustering problems. In: Classification and Related Methods of Data Analysis, Eds.: Bock, H.H. 67–74 (1988)
- Borge, R., Alexandrov, V., del Vas, J.J., Lumbreiras, J., Rodriguez, E.: A comprehensive sensitivity analysis of the WRF model for air quality applications over the Iberian Peninsula. *Atmos. Environ.* **42**(37), 8560–8574 (2008). <https://doi.org/10.1016/j.atmosenv.2008.08.032>
- Bourdrel, T., Bind, M.A., Béjot, Y., Morel, O., Argacha, J.F.: Cardiovascular effects of air pollution. *Arch. Cardiovasc. Dis.* **110**(11), 634–642 (2017). <https://doi.org/10.1016/j.acvd.2017.05.003>
- Byun, D.W., Ching, J.K.S.: Science algorithms of the EPA Models-3 Community Multiscale Air Quality (CMAQ) modeling system. U.S. Environmental Protection Agency, Washington D.C., EPA/600/R-99/030 (NTIS PB2000–100561) (1998)
- Byun, D., Schere, K.L.: Review of the governing equations, computational algorithms, and other components of the Models-3 Community Multiscale Air Quality (CMAQ) modeling system. *Appl. Mech. Rev.* **59**(2), 51–77 (2006). <https://doi.org/10.1115/1.2128636>
- Cabaneros, S.M., Calautit, J.K., Hughes, B.R.: A review of artificial neural network models for ambient air pollution prediction. *Environ. Model. Softw.* **119**, 285–304 (2019). <https://doi.org/10.1016/j.envsoft.2019.06.014>
- Du, Y., Xu, X., Chu, M., Guo, Y., Wang, J.: Air particulate matter and cardiovascular disease: the epidemiological, biomedical and clinical evidence. *J. Thoracic Dis.*, 8(1), E8–E19 (2016) <https://doi.org/10.3978%2Fj.issn.2072-1439.2015.11.37>
- Fast, J.D., Easter, R.C.: A Lagrangian particle dispersion model compatible with WRF. In: 7th Annual WRF User's Workshop, Boulder, CO, 19–22 June (2006)
- Gilik, A., Ogrenç, A.S., Ozmen, A.: Air quality prediction using CNN+LSTM-based hybrid deep learning architecture. *Environ. Sci. Pollut. Res.* **29**(8), 11920–11938 (2022). <https://doi.org/10.1007/s11356-021-16227-w>
- Han et al.: developed a hybrid deep learning framework that combines the CNN with LSTM to describe the spatial-temporal nature of PM<sub>2.5</sub> transport from the neighbor stations in Hong Kong and Beijing (2022)
- Ho, C.-H., Park, I., Oh, H.-R., Gim, H.-J., Hur, S.-K., Kim, J., Choi, D.-R.: Development of a PM<sub>2.5</sub> prediction model using a recurrent neural network algorithm for the Seoul metropolitan area, Republic of Korea. *Atmos. Environ.*, 245, 118021 (2021) <https://doi.org/10.1016/j.atmosenv.2020.118021>
- IHME: Findings from the Global Burden of Disease Study 2017. Institute for Health Metrics and Evaluation, Seattle, WA (2018)
- Kim, H.C., Kim, S., Son, S.-W., Lee, P., Jin, C.-S., Kim, E., Kim, B.-U., Ngan, F., Bae, C., Song, C.-K., Stein, A.: Synoptic perspectives on pollutant transport patterns observed by satellites over East Asia: Case studies with a conceptual model. *Atmospheric Chemistry and Physics Discussions*, 1–30 [preprint] (2016). <https://doi.org/10.5194/acp-2016-673>
- Kim, B.-U., Bae, C., Kim, H.C., Kim, E., Kim, S.: Spatially and chemically resolved source apportionment analysis: Case study of high particulate matter event. *Atmos. Environ.* **162**, 55–70 (2017). <https://doi.org/10.1016/j.atmosenv.2017.05.006>
- Kim, H.S., Park, I., Song, C.H., Lee, K., Yun, J.W., Kim, H.K., Jeon, M., Lee, J., Han, K.M.: Development of a daily PM<sub>10</sub> and PM<sub>2.5</sub> prediction system using a deep long short-term memory neural network model. *Atmos. Chem. Phys.*, 19(20), 12935–12951 (2019) <https://doi.org/10.5194/acp-19-12935-2019>
- Koo, Y.-S., Kim, S.-T., Cho, J.-S., Jang, Y.-K.: Performance evaluation of the updated air quality forecasting system for Seoul predicting PM10. *Atmos. Environ.* **58**, 56–69 (2012). <https://doi.org/10.1016/j.atmosenv.2012.02.004>
- Koo, Y.-S., Choi, D.-R., Kwon, H.-Y., Jang, Y.-K., Han, J.-S.: Improvement of PM10 prediction in East Asia using inverse modeling. *Atmos. Environ.* **106**, 318–328 (2015). <https://doi.org/10.1016/j.atmosenv.2015.02.004>
- Koo, Y.-S., Yun, H.-Y., Choi, D.-R., Han, J.-S., Lee, J.-B., Lim, Y.-J.: An analysis of chemical and meteorological characteristics of haze events in the Seoul metropolitan area during January 12–18, 2013. *Atmos. Environ.* **178**, 87–100 (2018). <https://doi.org/10.1016/j.atmosenv.2018.01.037>
- Lee, D.-G., Lee, Y.-M., Jang, K.-W., Yoo, C., Kang, K.-H., Lee, J.-H., Jung, S.-W., Park, J.-M., Lee, S.-B., Han, J.-S., Hong, J.-H., Lee, S.-H.: Korean national emissions inventory system and 2007 air pollutant emissions. *Asian J. Atmos. Environ.* **5**(4), 278–291 (2011a). <https://doi.org/10.5572/ajae.2011.5.4.278>
- Lee, S., Ho, C.-H., Choi, Y.-S.: High-PM<sub>10</sub> concentration episodes in Seoul, Korea: Background sources and related meteorological conditions. *Atmos. Environ.* **45**(39), 7240–7247 (2011b). <https://doi.org/10.1016/j.atmosenv.2011.08.071>
- Lee, S., Ho, C.-H., Lee, Y.G., Choi, H.-J., Song, C.-K.: Influence of trans-boundary air pollutants from China on the high-PM<sub>10</sub> episode in Seoul, Korea for the period October 16–20, 2008. *Atmos. Environ.* **77**, 430–439 (2013). <https://doi.org/10.1016/j.atmosenv.2013.05.006>
- Lee, J.-B., Lee, J.-B., Koo, Y.-S., Kwon, H.-Y., Choi, M.-H., Park, H.-J., Lee, D.-G.: Development of a deep neural network for

- predicting 6 h average PM<sub>2.5</sub> concentrations up to 2 subsequent days using various training data. *Geosci. Model Dev.*, **15**(9), 3797–3813 (2022) <https://doi.org/10.5194/gmd-15-3797-2022>
- Li, Y., Guo, J., Sun, S., Li, J., Wang, S., Zhang, C.: Air quality forecasting with artificial intelligence techniques: A scientometric and content analysis. *Environ. Model. Softw.* **149**, 105329 (2022). <https://doi.org/10.1016/j.envsoft.2022.105329>
- Lightstone, S.D., Moshary, F., Gross, B.: Comparing CMAQ forecasts with a neural network forecast model for PM<sub>2.5</sub> in New York. *Atmosphere*, **8**(9), 161 (2017) <https://doi.org/10.3390/atmos8090161>
- Lightstone, S., Gross, B., Moshary, F., Castillo, P.: Development and assessment of spatially continuous predictive algorithms for fine particulate matter in New York State. *Atmosphere* **12**(3), 315 (2021). <https://doi.org/10.3390/atmos12030315>
- Mahmoudabadi, H., Briggs, G.: Directional kriging implementation for gridded data interpolation and comparative study with common methods. American Geophysical Union, Fall Meeting, abstracts G21B-1006 (2016).
- Mao, Y., Lee, S.: Deep convolutional neural network for air quality prediction. *J. Phys: Conf. Ser.* **1302**(3), 032046 (2019). <https://doi.org/10.1088/1742-6596/1302/3/032046>
- Meng, J.: Raster data projection transformation based-on Kriging interpolation approximate grid algorithm. *Alex. Eng. J.* **60**(2), 2013–2019 (2021). <https://doi.org/10.1016/j.aej.2020.12.006>
- Oh, H.-R., Ho, C.-H., Kim, J., Chen, D., Lee, S., Choi, Y.-S., Chang, L.-S., Song, C.-K.: Long-range transport of air pollutants originating in China: A possible major cause of multi-day high-PM<sub>10</sub> episodes during cold season in Seoul, Korea. *Atmos. Environ.* **109**, 23–30 (2015). <https://doi.org/10.1016/j.atmosenv.2015.03.005>
- Oh, H.-R., Ho, C.-H., Koo, Y.-S., Baek, K.-G., Yun, H.-Y., Hur, S.-K., Choi, D.-R., Jhun, J.-G., Shim, J.-S.: Impact of Chinese air pollutants on a record-breaking PMs episode in the Republic of Korea for 11–15 January 2019. *Atmos. Environ.* **223**, 117262 (2020) <https://doi.org/10.1016/j.atmosenv.2020.117262>
- Ohara, T., Akimoto, H., Kurokawa, J., Horii, N., Yamaji, K., Yan, X., Hayasaka, T.: An Asian emission inventory of anthropogenic emission sources for the period 1980–2020. *Atmos. Chem. Phys.* **7**(16), 4419–4444 (2007). <https://doi.org/10.5194/acp-7-4419-2007>
- Park, Y., Kwon, B., Heo, J., Hu, X., Liu, Y., Moon, T.: Estimating PM<sub>2.5</sub> concentration of the conterminous United States via interpretable convolutional neural networks. *Environ. Pollut.*, **256**, 113395 (2020) <https://doi.org/10.1016/j.envpol.2019.113395>
- Perez, P., Reyes, J.: An integrated neural network model for PM<sub>10</sub> forecasting. *Atmos. Environ.* **40**(16), 2845–2851 (2006). <https://doi.org/10.1016/j.atmosenv.2006.01.010>
- Qi, Y., Li, Q., Karimian, H., Liu, D.: A hybrid model for spatiotemporal forecasting of PM<sub>2.5</sub> based on graph convolutional neural network and long short-term memory. *Sci. Total Environ.*, **664**, 1–10 (2019) <https://doi.org/10.1016/j.scitotenv.2019.01.333>
- Soh, P.-W., Chang, J.-W., Huang, J.-W.: Adaptive deep learning-based air quality prediction model using the most relevant spatial-temporal relations. *IEEE Access* **6**, 38186–38199 (2018). <https://doi.org/10.1109/ACCESS.2018.2849820>
- Stohl, A., Forster, C., Frank, P., Seibert, P., Wotawa, G.: Technical note: The Lagrangian particle dispersion model FLEXPART version 6.2. *Atmos. Chem. Phys.*, **5**(9), 2461–2474 (2005) [10.5194/acp-5-2461-2005](https://doi.org/10.5194/acp-5-2461-2005)
- Simonyan K, Zisserman A.: Very deep convolutional networks for large-scale image recognition. arXiv preprint, [arXiv:1409.1556](https://arxiv.org/abs/1409.1556) (2014) <https://doi.org/10.48550/arXiv.1409.1556>
- Wan, R., Mei, S., Wang, J., Liu, M., Yang, F.: Multivariate temporal convolutional network: A deep neural networks approach for multivariate time series forecasting. *Electronics* **8**(8), 876 (2019). <https://doi.org/10.3390/electronics8080876>
- Wang, W., Mao, W., Tong, X., Xu, G.: A novel recursive model based on a convolutional long short-term memory neural network for air pollution prediction. *Remote Sensing* **13**(7), 1284 (2021). <https://doi.org/10.3390/rs13071284>
- Wen, C., Liu, S., Yao, X., Peng, L., Li, X., Hu, Y., Chi, T.: A novel spatiotemporal convolutional long short-term neural network for air pollution prediction. *Sci. Total Environ.* **654**, 1091–1099 (2019). <https://doi.org/10.1016/j.scitotenv.2018.11.086> [CrossRef]
- Zhang, Q., Han, Y., Li, V.O.K., Lam, J.C.K.: Deep-AIR: A hybrid CNN-LSTM framework for fine-grained air pollution estimation and forecast in metropolitan cities. *IEEE Access*, **10**, 55818–55841 (2022) <https://doi.org/10.1109/ACCESS.2022.3174853>
- Zhao, J., Deng, F., Cai, Y., Chen, J.: Long short-term memory - Fully connected (LSTM-FC) neural network for PM<sub>2.5</sub> concentration prediction. *Chemosphere*, **220**, 486–492 (2019) <https://doi.org/10.1016/j.chemosphere.2018.12.128>

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

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