Statistical Seasonal Forecasting of Winter and Spring PM$_{2.5}$ Concentrations Over the Korean Peninsula

Dajeong Jeong$^1$ · Changhyun Yoo$^{1,*}$ · Sang-Wook Yeh$^2$ · Jin-Ho Yoon$^3$ · Daegyun Lee$^4$ · Jae-Bum Lee$^4$ · Jin-Young Choi$^4$

1 Department of Climate and Energy Systems Engineering, Ewha Womans University, 52 Ewhayeodae-gil, Seodaemun-gu, Seoul, South Korea
2 Department of Marine Sciences and Convergent Technology, Hanyang University ERICA, Ansan, South Korea
3 School of Earth Sciences and Environmental Engineering, Gwangju Institute of Science and Technology, Gwangju, South Korea
4 Air Quality Forecasting Center, National Institute of Environmental Research, Incheon, South Korea

Received: 15 December 2021 / Revised: 1 March 2022 / Accepted: 4 March 2022 / Published online: 28 March 2022
© The Author(s) under exclusive licence to Korean Meteorological Society and Springer Nature B.V. 2022

Abstract
Concentrations of fine particulate matter smaller than 2.5 μm in diameter (PM$_{2.5}$) over the Korean Peninsula experience year-to-year variations due to interannual variation in climate conditions. This study develops a multiple linear regression model based on slowly varying boundary conditions to predict winter and spring PM$_{2.5}$ concentrations at 1–3-month lead times. Nation-wide observations of Korea, which began in 2015, is extended back to 2005 using the local Seoul government’s observations, constructing a long-term dataset covering the 2005–2019 period. Using the forward selection stepwise regression approach, we identify sea surface temperature (SST), soil moisture, and 2-m air temperature as predictors for the model, while rejecting sea ice concentration and snow depth due to weak correlations with seasonal PM$_{2.5}$ concentrations. For the wintertime (December–January–February, DJF), the model based on SSTs over the equatorial Atlantic and soil moisture over the eastern Europe along with the linear PM$_{2.5}$ concentration trend generates a 3-month forecasts that shows a 0.69 correlation with observations. For the springtime (March–April–May, MAM), the accuracy of the model using SSTs over North Pacific and 2-m air temperature over East Asia increases to 0.75. Additionally, we find a linear relationship between the seasonal mean PM$_{2.5}$ concentration and an extreme metric, i.e., seasonal number of high PM$_{2.5}$ concentration days.

Keywords Seasonal prediction · PM$_{2.5}$ concentrations · Multiple linear regression model

1 Introduction
High concentrations of fine particulate matter smaller than 2.5 μm in diameter (PM$_{2.5}$) are a major environmental contributor to cardiovascular and respiratory diseases in East Asian countries (West et al. 2016; Song et al. 2017). Annual mean concentrations of PM$_{2.5}$ over the Korean Peninsula exceed 25 μg m$^{-3}$ (Lee 2014; Bae et al. 2020), more than double of the World Health Organization’s (WHO’s) air quality guideline value of 10 μg m$^{-3}$ annual mean of (Chae et al. 2021). On certain extreme days (e.g., March 3–5, 2019), the daily mean PM$_{2.5}$ over the Seoul metropolitan area reach approximately 110 μg m$^{-3}$ (Lee et al. 2019). This record also substantially exceeds the WHO’s daily mean guideline of 25 μg m$^{-3}$. To minimize citizens’ exposure to high PM$_{2.5}$ concentrations, the Korea Ministry of Environment (KME) issues advisories (warnings) when the 2-hour mean concentration exceeds or is expected to exceed 75 (150) μg m$^{-3}$.

The National Institute of Environmental Research (NIER) of KME conducts operational PM$_{2.5}$ concentration forecasts for lead times up to a week. Along with observation and meteorological field forecasts, their primary forecasting tool is the Community Multiscale Air Quality System (CMAQ) developed by the U.S. Environmental Protection Agency (Byun and Schere 2006). A recent report indicated their daily forecasts have an accuracy as high as 85% when
measured based on four categories: good (0–15 µg m⁻³), moderate (16–35 µg m⁻³), unhealthy (36–75 µg m⁻³), and very unhealthy (76 µg m⁻³ and above) (NIER 2019). However, demand to extend their forecast out to seasonal time scale, i.e., 1–3 months ahead, has been growing. Information regarding the air quality of upcoming seasons can play an important role in planning emission controls and bolstering vulnerable populations’ resilience against harmfully high concentration events.

Many operational centers provide seasonal climate predictions, such as temperature and precipitation forecasts. Such predictions are understood as boundary condition problems because a few months in the future, errors in initial atmospheric conditions become too substantial to provide useful information. Instead, these predictions rely on the sensitivity of the atmosphere to its slowly varying boundary conditions, such as sea surface temperature and soil wetness. In general, operational centers utilize two types of forecast guidance for their seasonal climate predictions: dynamical model forecast guidance and statistical model forecast guidance. Dynamical models provide numerical solutions for physical climate processes, which has become increasingly accurate as a result of advances in ensemble forecasting techniques and initializations of land and ocean. Statistical models, on the other hand, are based on empirical relationships between predictors and predictands. El Niño Southern Oscillation (ENSO) is, for example, one of the most important predictors of the globe’s seasonal climate; as a result, many institutions carefully monitor and predict it (e.g., https://iri.columbia.edu/our-expertise/climate/forecasts/enso/current/). One clear advantage of seasonal statistical models is that they have comparable accuracy to dynamical models and are much less expensive to develop and maintain.

Previous studies have found that various forms of inter-annual climate variability, such as ENSO and East Asian winter monsoons, strongly influence PM_{2.5} concentrations over East Asia (Shimadera et al. 2013; Jia et al. 2015; Yang et al. 2016; Jeong and Park 2017; Lang et al. 2017; Yu et al. 2019; Kumar et al. 2021). This suggests that the use of a statistical model to predict seasonal PM_{2.5} over the Korean Peninsula could be successful. For example, circulation changes associated with winter monsoons can result in anomalous PM_{2.5} transport that modulates surface-level concentrations (Jeong and Park, 2017). Applying this information regarding wintertime monsoons and Siberian Highs in a simple linear regression model generates a skillful PM_{2.5} concentration prediction for the same season with a correlation coefficient of approximately 0.72 (Jeong et al. 2021). However, due to the insufficiency of long-term surface PM_{2.5} observations, previous studies have relied heavily on either proxy data of PM_{2.5}, such as aerosol optical depth from satellite observations (e.g., Jia et al. 2015), or model-generated data, such as atmospheric chemical transport model simulations guided by atmospheric reanalysis data (Jeong et al. 2021). Also, while some studies have sought to generate short-term PM_{2.5} predictions (e.g., Lee et al. 2013; Xiao et al. 2020; Chae et al. 2021), models have not been developed yet for long lead times of 1–3 months and for different seasons, meaning our understanding of PM_{2.5} seasonal prediction remain limited. Lastly, extreme metrics, such as number of high concentration days, have only recently been investigated for fine particulate matter smaller than 10 µm in diameter, PM_{10} (e.g., Cho et al. 2021; Ku et al. 2021), and connections between seasonal mean PM_{2.5} concentrations and high concentration days are not examined.

Recognizing these gaps in existing research, this study seeks to make statistical seasonal predictions of PM_{2.5} concentrations over the Korean Peninsula based on atmospheric boundary conditions, which is the first attempt to the best of our knowledge. We use long-term in situ observation of PM_{2.5} to construct our model and verify its predictions. To understand the characteristics of the data, we explore seasonal distributions and connections between seasonal mean PM_{2.5} concentrations and high concentration days. This is followed by examination on boundary conditions by land and ocean as potential sources of prediction skill. Our analyses verify predictions for lead times from 1 to 3 months during winter and spring, respectively. The remainder of the study is organized as follows: Section 2 describes the data and model formulation; Section 3 presents the results of correlation analyses and model verifications; and Section 4 explains and discusses the conclusions of the study.

2 Data and Methods

2.1 PM_{2.5} Concentration Data

We use hourly PM_{2.5} concentration data observed at 25 stations in Seoul, South Korea, from 2005 to 2018, provided by Seoul Research Institute of Public Health and Environment (SRIPHE; https://data.seoul.go.kr/dataList/OA-15526/S/1/datasetView.do). This is the longest period of PM_{2.5} concentration observation in South Korea. We find that across the stations approximately from 40.1 to 62.8% of the hourly observations are missing. Unrealistic values are masked when they are either less than 0 µg m⁻³ or greater than 300 µg m⁻³, removing overall less than 0.1% of the entire data. We obtain a daily time series by taking averages for each day and over all the stations (black line in Fig. 1).

We also employ another hourly PM_{2.5} concentration dataset by KME (NIER 2019). Collection of this data began in 2015, and the dataset therefore covers a much shorter time period than the one collected by SRIPHE. However, the KME dataset has extensive spatial coverage
over South Korea, is quality controlled, and is publicly available after a few months delay (https://www.airkorea.or.kr). Measurements from 9 stations (Seoul, Busan, Daegu, Daejeon, Gwangju, Gangneung, Taean, Chungju, and Andong) are selected to cover temporally extensive length of the 2015–2019 period and spatially the Korean Peninsula. At each station, between 46.2 and 23.4% of data are missing, except Taean station that misses 76.9% of its data. As for the PM$_{2.5}$ of SRIPHE, we construct a daily time series without missing values by averaging the hourly concentration from KME over days and stations (red line in Fig. 1).

To benefit from both datasets, we construct a long-term observation by appending the KME dataset (2015–2019) to the SRIPHE dataset (2005–2014), dropping SRIPHE data of the overlapping 2015–2018 period. Note that between 2015 and 2018, the two datasets exhibit an extremely high correlation: correlation coefficient (R) = 0.88, which exceeds the 99% statistical significance level. Seasonal averages for DJF and MAM will further ensure that the differences between the two data are smoothed out. We find that the overall conclusions of this study remain the same when we use only the SRIPHE data for 2005–2018 (not shown). Therefore, in this study, we present results for 14 seasons based on the merged data during DJF of 2005/2006–2018/2019 and MAM of 2006–2019.

We define the number of high PM$_{2.5}$ concentration days by counting the total number of days for a season that daily mean concentrations exceeds 36 µg m$^{-3}$. In terms of air quality category of Korea, this includes both unhealthy (36–75 µg m$^{-3}$) or very unhealthy (76 µg m$^{-3}$ and above) air quality categories (NIER 2019). Also, the definition of high PM$_{2.5}$ concentration days resembles the definitions on high PM$_{10}$ concentration days in previous studies (Yun and Yoo 2019; Ku et al. 2021).

### 2.2 Climate Data

As potential predictors of seasonal PM$_{2.5}$ concentration, we examine five variables that provide boundary conditions of atmospheric circulation: sea surface temperature (SST), soil moisture, and 2-m air temperature ($T_{2m}$), along with sea ice concentration and snow depth. For SST, we employ monthly data obtained from Met Office Hadley Centre (Rayner et al. 2003). The data has a horizontal resolution of 1° and is fill valued when the value of the SST data is lower than $-1.8$ °C, indicating formation of sea ice. For soil moisture, $T_{2m}$, and snow depth, we use the monthly European Centre for Medium-Range Weather Forecasts (ECMWF) fifth generation of atmospheric reanalysis (ERA5; Hersbach et al. 2020) dataset. We integrate four volumetric soil water layers in the reanalysis, which ranges approximately 0–3 m, to represent soil moisture. The horizontal resolution of the datasets was originally at 0.25° but is interpolated to 1.5° using bilinear interpolation. For sea ice concentration, we take the data derived using measurements from the Scanning Multichannel Microwave Radiometer (SMMR) and from the Special Sensor Microwave/Imager (SSM/I) sensors through bootstrap algorithm (Comiso 2017). The sea ice data is at 25 km resolution on the polar stereographic grid.

For all the climatological variables, we first apply the seasonal averages. We then compute anomalies by removing the seasonal cycles, which are defined as the calendar month means. Lastly, we subtract the linear trends defined for each calendar month from the anomalies to contain variability of interannual time scale or longer without trends.
2.3 Probabilistic Forecast by Multiple Linear Regression Model

Multiple linear regression model is built for predictions of seasonal mean PM$_{2.5}$ concentrations and numbers of high PM$_{2.5}$ concentration days during DJF and MAM at lead times of 1–3 months. To identify the predictors, we apply the forward selection stepwise regression approach, which is described in Section 3 (see also Myoung et al. 2020). Note that the regression model is built after the linear trends are removed respectively from predictors and predictands for each season. Also note that here we limit the maximum number of predictors to be equal to or less than three, including the linear trend of the PM$_{2.5}$ concentration. This helps us avoid overfitting the model, given the short time period of observation. As a result, all models turn out to have three predictors, which make the equation of the model:

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + b_3x_3$$

where $\hat{y}$ is the predictands, such as DJF mean PM$_{2.5}$ concentration or number of high PM$_{2.5}$ concentration days during DJF; and $x_1$ and $x_2$ are two predictors based on boundary conditions (Section 3.2) with their lagged regression coefficients, $b_1$ and $b_2$. The model includes the linear trend of the PM$_{2.5}$ concentration $b_3$, which is defined for each season, and time $x_t$, as well as the intercept term, $b_0$.

To validate the prediction skill of the regression model, we apply a leave-one-year-out cross-validation approach. This means we construct the regression model for a certain year based on the data from all other years and verify the model using the data for the chosen year. We repeat this process each year for each model. Three categories can be defined for categorical forecast using lower and upper climatological terciles of the predictand distribution. Because only 14 seasonal values are available, we obtain the terciles from 30-day running averaged data for each season, instead of seasonal values, an approach we discuss further in the following section.

3 Results

3.1 Predictands

We first explore the probabilistic distributions of PM$_{2.5}$ concentrations by seasons to determine the characteristics of our chosen predictand variable. A histogram of daily averaged PM$_{2.5}$ concentrations reveals overall positive skewness of the data (Fig. 2). Many days of low and moderate PM$_{2.5}$ concentrations are clustered around the left tail of the distribution while some extremely high PM$_{2.5}$ concentrations can be seen in the right tail. Compared to the annual case (Fig. 2), the skewness is alleviated for the DJF and MAM distributions (Fig. 2), which is due to reductions in relative frequency of occurrence of low PM$_{2.5}$ concentrations.

In contrast, we find seasonally averaged PM$_{2.5}$ concentrations are expected to be much more symmetrical than the daily means. Because we only have data for 14 years, daily PM$_{2.5}$ concentrations are 90-day running averaged, instead of being seasonally averaged to examine possible Gaussianity in the seasonal mean field, assuming that all daily concentrations are independent from one another. Given that the decorrelation time scale of the PM$_{2.5}$ concentration over the Korean Peninsula is only a few days, the assumption seems reasonable. The distributions of the 90-day running average concentrations are bell shaped for all seasons, despite some greater-than-normal kurtosis for the DJF distributions (Fig. 2).

Similarly, 90-day running days can be used to estimate underlying distributions of high concentration days instead of seasonal number of days (Fig. 2). Positive skewness is apparent in the all-season case (Fig. 2). However, for DJF and MAM, the skewness is alleviated, and the distributions are roughly normal distributions with no pronounced kurtosis (Fig. 2). MAM show a larger number of high concentration days than DJF.

Next, we examine seasonally averaged PM$_{2.5}$ concentrations and seasonal high PM$_{2.5}$ days. Note that we use these metrics as predictands in our statistical forecasting models. Interestingly, a linear relationship can be seen between the quantities, and more strongly for MAM (Fig. 3). The Pearson correlation between the metrics for MAM is 0.95, suggesting that the same statistical model may be used to predict both predictands, particularly for MAM (Fig. 3), because the number of daily extreme events primarily determines the seasonal mean value. For DJF, the correlation between the metrics is slightly less, at 0.71. Despite some linearity, we observe a range of high PM$_{2.5}$ days near the seasonal mean PM$_{2.5}$ value of approximately 30 µg m$^{-3}$ (Fig. 3), suggesting that each predictand may have a different predictor. In terms of location of the predictands, the means of the seasonal values for DJF are 29.91 µg m$^{-3}$ and 8.17 days, respectively. These seasonal values are slightly higher than those from MAM (28.02 µg m$^{-3}$ and 7.36 days). For the spread, however, MAM exhibits a greater range of values in both mean concentration and number of days (see also Fig. 2).

3.2 Predictors

We begin selecting the models’ first predictors using SST data. As described earlier, we build respective models for DJF and MAM, first targeting seasonal mean PM$_{2.5}$ concentration predictions out to 3 months. Next, we select
predictors by exploring soil moisture, T_{2m}, sea ice concentration, and snow depth through stepwise regression approach.

In search of potential predictors in SST, we examine lagged correlation maps between seasonal mean concentrations and SSTs north of 30°S (Fig. 4). Note that in the maps, the three-month averaged SSTs are lagged for the PM$_{2.5}$ concentrations of DJF and MAM by 1–3 months. For DJF PM$_{2.5}$ concentrations, significant values can be observed over the North Atlantic, most pronouncedly over the equator (Fig. 4), with a hint of tripole-like SST mode in the North Atlantic including the area south of Greenland. The North Atlantic is known to influence East Asia through changes
in atmospheric circulation that exhibit poleward and then equatorward wave trains (e.g., Ok et al. 2017; Myoung et al. 2020). In addition, due to the slowly varying nature of the ocean, the correlation patterns remain largely unchanged from September to November (i.e., 3 months ahead) through October–December (2 months ahead) and November to January (1 month ahead), suggesting that the equatorial Atlantic SST exerts a significant influence at seasonal time scale.

Area averaged at the three regions with the largest correlation coefficients, we select the DJF SSTs as individual predictors in our regression models (Table 1). The model built with the equatorial Atlantic SST (10°S–10°N, 30°W–10°E; green boxed region in Fig. 4) shows a 0.52 correlation with detrended DJF PM$_{2.5}$ concentration. The same calculation using the western North Pacific (30°N–65°N, 110°E–150°E) and North Atlantic (50°N–70°N, 20°W–30°E) SSTs show smaller correlations (i.e., 0.36 and 0.24, respectively). Moreover, the strongly negative correlation we find between the equatorial Atlantic and western North Pacific SSTs suggests the predictors may be

![Figure 4](image_url)

Fig. 4 Lagged correlation coefficients between seasonal mean PM$_{2.5}$ concentrations and SST anomalies. Seasonal mean SSTs lead (a) DJF and (b) MAM concentrations by 1–3 months. Here, we define the anomaly as a deviation from the monthly climatology. The linear trend is also removed. Regions with correlation values that exceed the 95% confidence level of the two-tailed Student t-test are hatched.

### Table 1

| SST Region | R with detrended PM$_{2.5}$ | R between predictors | SST Region | R with detrended PM$_{2.5}$ | R between predictors |
|------------|----------------------------|----------------------|------------|----------------------------|----------------------|
| Eq. Atlantic (10°S–10°N, 30°W–10°E) | 0.52 | | N. Pacific (35°N–45°N, 160°E–180°) | 0.77* | |
| W. N. Pacific (30°N–65°N, 110°E–150°E) | 0.36 | -0.60* | Eq. Atlantic (20°N–30°N, 80°W–40°W) | 0.64* | -0.57* |
| N. Atlantic (50°N–70°N, 20°W–30°E) | 0.24 | -0.15 | Eq. E. Pacific (20°S–10°N, 140°W–50°W) | 0.47 | -0.71* |

We select three SST predictors for (left) DJF and (right) MAM, respectively, based on Fig. 4. Note that models are built no lag between the predictor and predictand but leave-one-year-out cross-validated. Models with the highest correlation are shown from the top, and the correlations that exceed the 95% confidence level of the two-tailed Student t-test are asterisked. Also shown are the correlations between a predictor and the most skillful predictor.
interdependent. As a result, we only select the equatorial Atlantic SST only as a predictor in the DJF concentration model.

For MAM, we choose North Pacific SST (35°N–45°N, 160°E–180°E; green boxed region in Fig. 4) as a first predictor in our models. As with the DJF models, we make this decision based on correlation maps (Fig. 4) and simple linear regression model results (Table 1). Note that through ENSO, the North Pacific SST is likely to be connected with the equatorial eastern Pacific SST (e.g., mode 1 of global SST in Fig. 2 of Messié and Chavez 2011), which also exhibits significant correlations. Our analysis shows significantly negative correlations in the North Pacific throughout the lags for MAM (Fig. 4). When built with this North Pacific predictor, the model demonstrates an astonishing accuracy, with R = 0.77 (Table 1). SSTs over the equatorial Atlantic (20°N–30°N, 80°W–40°W) and equatorial eastern Pacific (20°S–10°N, 140°W–50°W) also prove high skills (R = 0.64 and 0.47, respectively) in MAM concentration predictions, but we do not select them due to high cross correlations with the North Pacific SST (-0.57 and – 0.71, respectively).

After establishing the first predictors based on SSTs, we examine lagged relationships between seasonal mean PM$_{2.5}$ concentration and soil moisture (Fig. 5). Based on the statistical significance of lagged correlation values for each season, two potential regions are selected. During DJF, soil moistures over the eastern Europe (45°N–55°N, 40°E–50°E; green boxed region in Fig. 5) and the Gobi Desert (40°N–50°N, 90°E–100°E) show large correlation coefficients (Fig. 5). We examine these two candidate

![Fig. 5](image-url) As in Fig. 4, except that the lagged correlations with seasonal mean soil moisture are shown

| DJF Soil | R with detrended and SST predictor removed PM$_{2.5}$ | R between predictors | MAM Soil | R with detrended and SST predictor removed PM$_{2.5}$ | R between predictors |
|----------|-----------------------------------------------|----------------------|----------|-----------------------------------------------|----------------------|
| E. Europe (45°N–55°N, 40°E–50°E) | 0.82* | -0.07 | E. Europe (45°–60°N, 55°E–70°E) | 0.61* | -0.53 |
| Gobi Desert (40°N–50°N, 90°E–100°E) | 0.37 | 0.42 | Southern China (20°N–30°N, 90°E–110°E) | 0.57* | -0.38 |

Note that we linearly remove the SST predictor from the detrended PM$_{2.5}$ concentration before constructing and verifying the models for the chosen soil moisture predictor. Also shown are the correlations between the selected SST predictor (see main text and Table 1) and the soil moisture predictor.
regions using the stepwise procedure; after linearly removing the equatorial Atlantic SST predictor from the PM$_{2.5}$ data, regression models are built and verified using the soil moistures (Table 2). The resulting analysis highlights the benefit of using the eastern Europe soil moisture as a second predictor. The correlation is $R = 0.82$ for the DJF predictions of the detrended and equatorial Atlantic SST removed PM$_{2.5}$ concentration. The eastern Europe soil moisture exhibits a low correlation with the equatorial Atlantic SST, suggesting weak dependency between the predictors.

Repeating the process for MAM, we find that soil moisture over the eastern Europe ($45^\circ$–$60^\circ$N, $55^\circ$E–$70^\circ$E) and southern China ($20^\circ$N–$30^\circ$N, $90^\circ$E–$110^\circ$E) are good potential sources of predictability (Fig. 5). The correlations obtained from the models using these regions are reasonably high, $R = 0.61$ and $0.57$, respectively (Table 2). However, in contrast to the DJF case, both soil moisture predictors are highly correlated with the North Pacific SST ($R = -0.53$ and $-0.38$, respectively), making them less ideal as predictors.

Including the linear trend of each season, we have so far identified three and two predictors for DJF and MAM PM$_{2.5}$ concentration prediction models, respectively. To complement the model, we now examine T$_{2m}$ over the land as potential predictors. For MAM, high correlations can be seen over almost all of the Eurasian continent (Fig. 6), which we divide into three subdomains: East Asia ($25^\circ$N–$45^\circ$N, $110^\circ$E–$120^\circ$E; green boxed region in Fig. 6), Siberia ($45^\circ$N–$65^\circ$N, $60^\circ$E–$110^\circ$E), and western Europe ($35^\circ$–$50^\circ$N, $10^\circ$W–$20^\circ$E). Based on its extremely high accuracy and negligible correlation with the SST predictor, the T$_{2m}$ over East

![Fig. 6](image)

As in Fig. 4, except that the lagged correlations with seasonal mean T$_{2m}$ are shown

| DJF T$_{2m}$     | R with detrended and Predictor 1 removed PM$_{2.5}$ | R between predictors | MAM T$_{2m}$       | R with detrended and Predictor 1 removed PM$_{2.5}$ | R between predictors |
|------------------|---------------------------------------------------|----------------------|---------------------|---------------------------------------------------|----------------------|
| E. Europe        | 0.63*                                              | 0.23                 | East Asia           | 0.91*                                              | -0.10                |
| ($45^\circ$N–$55^\circ$N, $40^\circ$E–$50^\circ$E) |                                                   |                      | ($25^\circ$N–$45^\circ$N, $110^\circ$E–$120^\circ$E)|                                                   |                      |
| Siberia          | 0.56*                                              |                      |                     |                                                   |                      |
| ($45^\circ$N–$65^\circ$N, $60^\circ$E–$110^\circ$E)|                                                   |                      |                     |                                                   |                      |
| W. Europe        | 0.49                                               | -0.23                |                     |                                                   |                      |
| ($35^\circ$–$50^\circ$N, $10^\circ$W–$20^\circ$E) |                                                   |                      |                     |                                                   |                      |
Asia is selected as the final predictor for the MAM model (Table 3). Despite their reasonable accuracy and moderate correlations with the SST predictors, we leave aside the other two regions to limit the total number of predictors in our models. For DJF, a high correlation emerges over the eastern Europe (45°N–55°N, 40°E–50°E), but this region is already used for soil moisture.

The predictor selection process is reiterated for sea ice concentration and snow depth. However, the latter two do not make to our final selection because our analyses identify no large areas with significant correlations (Figs. S1–S2). This may be surprising given that some previous studies have shown these two factors to be good predictors of springtime rainfall over the Korean Peninsula (Myoung et al. 2020), which is an important wet removal mechanism of PM2.5 concentrations. Moreover, there are studies that find evidence of linkage between the first mode wintertime PM10 concentrations over the Korean Peninsula and Arctic sea ice (Kim et al. 2019; Lee et al. 2020). Although our results do not rule out the impact of sea ice and snow depth on PM2.5 concentrations, we speculate that physical and chemical processes other than precipitation may play an important role in PM2.5 concentrations. Also, the influence of sea ice and snow depth may be pronounced only during certain months of winter or spring, instead of the entire season (Kim et al. 2019).

To examine whether the circulation patterns modulated by the chosen predictors resemble the circulation anomalies associated with seasonal PM2.5 concentrations, we first inspect the interannual relationship identified using detrended PM2.5 concentrations with detrended seasonal mean geopotential height at 500-hPa (Fig. 7). This is compared with the 2-month lagged correlation maps between October–November–December (OND) SST anomalies and

![Fig. 7 Correlation coefficients using seasonal-mean detrended 500-hPa geopotential height anomalies and the PM2.5 concentrations of the Korean Peninsula for (a) DJF and (b) MAM. We obtain (c, e) lagged correlation coefficients between DJF detrended 500-hPa geopotential height anomalies and October–December (OND) SST anomalies averaged over equatorial Atlantic (10°S–10°N, 30°W–10°E) and soil moisture anomalies averaged over (45°N–55°N, 40°E–50°E). For (d, f), lagged correlation coefficients are computed using MAM detrended 500-hPa geopotential height anomalies and January–March (JFM) SST anomalies averaged over North Pacific (35°N–45°N, 160°E–180°) and T2m anomalies averaged over East Asia (25°N–45°N, 110°E–120°E). Regions with correlation values that exceed the 95% confidence level of the two-tailed Student t-test are hatched.](image-url)
D.J.F height field (Fig. 7) and also between January–February–March (JFM) predictors and MAM height field (Fig. 7). During DJF, we find zonally elongated areas of positive coefficients over Mongolia and in the east of the Korean Peninsula of R near 0.3 (Fig. 7). The coefficient values are not statistically significant by the 95% confidence level, which weakens their robustness. Nonetheless, this points to a decrease in the climatological wind speed as well as advection of PM$_{2.5}$ or relevant substances by anomalous southerly during high PM$_{2.5}$ concentration. The high anomaly over the Korean Peninsula also indicates a weakening of vertical ventilation due to lowered atmospheric boundary layer and increased static stability (e.g., Fig. 4 in Cho et al. 2021). The SST and soil moisture predictors of DJF likewise exhibit anomalously high coefficients over Mongolia and in the east of the Korean Peninsula, respectively (Fig. 7). Pattern correlation between the circulation anomalies associated with the PM$_{2.5}$ concentrations and equatorial Atlantic SST over 30°S and poleward is R = 0.63. However, the value is weak (R = 0.16), for the eastern Europe soil moisture—understanding the mechanisms of this factor will require further study.

During MAM, anomalously high correlations over Mongolia and China and a tripole structure over the North Pacific are pronounced in the correlation maps between the PM$_{2.5}$ concentration and the 500-hPa geopotential height (Fig. 7). The high anomaly over China forms a northwesterly to the Korean Peninsula, while the tripole indicates a weakening of the climatological jet stream. The MAM correlation map does not resemble that of DJF, suggesting that different mechanisms may work during the spring season, and further studies will be required to figure out the mechanisms for MAM PM$_{2.5}$ concentrations. Nonetheless, based on the importance of local T$_{2m}$ in the MAM predictions, we speculate that chemical processes may play an important role in MAM PM$_{2.5}$ concentrations. The lagged correlation maps for the SST and T$_{2m}$ predictors of MAM align closely with the pattern correlations of 0.76 and 0.73, respectively, over poleward of 30°S (Fig. 7).

### 3.3 Model Verification

To evaluate the performance of the chosen predictors, we first examine observed and forecast seasonal mean PM$_{2.5}$ concentration time series (Fig. 8). As discussed earlier, equatorial Atlantic SST, eastern Europe soil moisture, and the linear trend are selected for the DJF regression models (Fig. 8a), and North Pacific SST, East Asia T$_{2m}$, and the linear trend for the MAM models (Fig. 8). We find overall skills of the DJF forecasts were shown as R = 0.87, 0.84, and 0.69 in terms of correlations for lead times of 1–3 months, indicating that the measures of the accuracy gradually decrease with lead times. For the MAM models, the correlations are R = 0.93, 0.76, and 0.75, slightly higher for lag 1 and 3 months than the skills for the DJF models. Examining the time series, we find declining trends in the PM$_{2.5}$ concentrations of both seasons, indicating that setting the linear trend as a predictor has a positive effect. Interannual variation is more pronounced during MAM than DJF, and this is well captured by the predictors.

A scatterplot of the observed and forecast concentrations can reveal over- or under-forecasting biases and
refinements of forecast resolutions (Fig. 9). We use the scatterplot instead of a calibration function due to the very small sample size. Overall, the scattered values line up well with the diagonals, suggesting the forecasts have good accuracy, as already demonstrated by the correlations. The fact that the values are not one sided from the diagonals indicates that the forecasts have no pronounced bias. However, the DJF forecasts hint at some conditional bias, as the regression lines (colored lines) exhibit slopes less than 1 (Fig. 9). This indicates that the DJF models are underconfident and tend to predict smaller (greater) concentration values when higher (lower) seasonal concentrations are observed. The MAM models show better calibration than the DJF models despite featuring greater spread from the diagonal.

Because most operational climate forecasts provide categorical information, we reevaluate deterministic predictands as categorical forecasts. Three categories, i.e., below normal, normal, and above normal seasonal concentrations, are defined using the upper and lower terciles obtained from the 90-day running averaged daily PM$_{2.5}$ concentrations of each season (dashed black horizontal lines in Fig. 8). We adopt this approach because the terciles from 14 observed seasonal values alone are not likely to be reliable. We use the Heidke skill score (HSS; Eq. 8.22 in Wilks 2011), a common seasonal climate forecast measure, to quantify the accuracy of the three categorical forecasting. HSSs are 77.42%, 29.50%, and 29.50% for DJF and 78.63%, 100%, and 46.97% for MAM forecasts. The abrupt decline in the DJF skill is caused by the small interannual variation during this season (Figs. 2 and 3), which makes the normal category very narrow (between the horizontal dashed lines in Fig. 8a). Nevertheless, given that 100% HSS signifies a perfect forecast accuracy and 0% HSS represents the accuracy that can be expected from a random forecast, these are skillful predictions.

4 Conclusions and Discussion

We constructed prediction models for seasonal mean PM$_{2.5}$ concentrations and seasonal number of high PM$_{2.5}$ concentration days using multiple linear regression models based on slowly varying atmospheric boundary conditions and long-term trends in the predictands, targeting the winter and spring seasons over the Korean Peninsula at 1–3-months ahead. To build the statistical models, we extended the KME’s PM$_{2.5}$ observations from 2015 to 2019 by appending 2005–2014 observations from the SRIPHE. For this period, we found significant empirical relationships were found between the predictands and sea surface temperature, soil moisture, and 2-m air temperature. Skillful predictions were possible for seasonal mean PM$_{2.5}$ concentrations; the correlations between the observations and the forecasts at 3-months ahead reached 0.69 and 0.75 for the December–February (DJF) and March–May (MAM) seasons, respectively. We found slightly lower accuracy for the models of seasonal number of high PM$_{2.5}$ concentration days; the correlations were 0.64 and 0.58 for DJF and MAM, respectively.

As predictors, we chose equatorial Atlantic SST, eastern Europe soil moisture, and the linear term trend for DJF mean concentrations and North Pacific SST, East Asian 2-m temperature, and the linear trend for MAM mean concentrations. Our analysis show that anomalous circulation patterns accompanied with the predictors, which may influence the PM$_{2.5}$ concentrations through changes in atmospheric transport and ventilation. However, these processes were not demonstrated in this study. A recent clustering analysis-based study revealed that five different circulation patterns are responsible for the occurrence of extreme PM$_{10}$ concentration events over the Korean Peninsula (Ku et al. 2021). We speculate that...
PM$_{2.5}$ concentration is also governed by multiple patterns of meteorological conditions via its physical transport and chemical processes. Analyses of daily evolution are required because the processes may be obscured in seasonal mean fields.

Our models benefit from the downward trend in seasonal mean PM$_{2.5}$ concentrations (black lines in Fig. 8). Enhancement in forecast accuracy by the linear trend can be quantified if the model is built without the trend (not shown). For DJF, the accuracy of the forecast without the trend begins at 0.66 for 1-month lead and drops to 0.62 and 0.40 for 2–3 months leads. The values for MAM are 0.82, 0.62, and 0.60 for 1–3-month forecasts without the trend. By comparing the values, we find that an increase in skills of roughly 0.1–0.3 owes to the existence of linear trends, which are approximately $-0.24$ µg m$^{-3}$ year$^{-1}$ and $-0.33$ µg m$^{-3}$ year$^{-1}$ in DJF and MAM PM$_{2.5}$ concentrations, respectively.

We may apply the same procedure to configure prediction models for number of high PM$_{2.5}$ days. It is inherently more challenging to predict extreme metrics than seasonal mean values because the former targets the tails of distributions. However, the near-linear relationship between seasonal mean concentrations and numbers of high concentration days, particularly during MAM (Fig. 3b), indicates a potential of the model to skillful predictions of the number of high PM$_{2.5}$ days. For DJF, a reselection of predictors can be expected given the weak linear relationship between the two predictands (Fig. 3a). Nonetheless, the number of high PM$_{2.5}$ days in DJF is relatively well spread out unlike the DJF mean PM$_{2.5}$ concentration clustered near its climatological mean (Fig. 3a).

The immense impact of Coronavirus disease 2019 on regional air quality could undermine the accuracy of statistical model-based predictions. This is because the models do not directly take emission information into account, although one may argue that some of the emission information is concealed in the PM$_{2.5}$ concentration and high PM$_{2.5}$ day trends. Aerosol optical depth, a proxy for PM concentration, declined between February and March in 2020, even with stagnant atmospheric conditions, indicating a strong pause in anthropogenic emissions (Koo et al. 2020). Observations of PM$_{2.5}$ concentrations indeed showed unprecedented seasonal mean values of 23.64 µg m$^{-3}$ and 17.98 µg m$^{-3}$ over the Korean Peninsula during DJF 2019/2020 and MAM 2020, respectively. Similarly, the mean value was even lower at 21.80 µg m$^{-3}$ for DJF 2020/2021, while official records were not yet available for MAM 2021 during the time of this research. Statistical models may be trained directly using emission information. However, ensuring emission inventories up to date could take several years, making the use of emission information as a predictor not possible soon. This is also an obstacle for atmospheric chemistry models that require emission data as an input to initialize their simulations.

Previous climate prediction studies have outlined ways that regression model can be improved. One recommended improvement is to include uncertainty-related information when switching a deterministic forecast to a probabilistic forecast (Chang et al. 2021). Specifically, the variance of a forecast field can be estimated using the errors in the forecast, which can be used as a parameter to build a probabilistic distribution of the forecast field. This allows the model to provide probabilities for each category. Another way to improve statistical models is to bridge with numerical forecast. Kim et al. (2021) demonstrated how a numerical forecast of slowly varying atmospheric mode of variability, such as Arctic Oscillation, can be integrated into a statistical model to extend its surface air temperature forecast skill. Similarly, numerical forecasts of atmospheric blocking can be utilized to predict PM$_{10}$ concentrations several days ahead (Shin et al. 2021).

**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1007/s13143-022-00275-4.

**Acknowledgements** This research was supported by the National Institute of Environment Research of Korea through NIER-2020-01-02-023 and the National Research Foundation of Korea through grants NRF-2018R1A6A1A08025520 and NRF-2019R1C1C1003161.

**Declarations**

**Competing Interest** The authors declare no competing financial interests or personal relationships that influence the work reported in this paper.

**References**

Bae, C., Kim, B.-U., Kim, H.C., Yoo, C., Kim, S.: Long-range transport influence on key chemical components of PM2.5 in the Seoul Metropolitan Area, South Korea, during the Years 2012–2016. Atmosphere 11, 48 (2020). https://doi.org/10.3390/atmos11010048

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, 51–76 (2006). https://doi.org/10.1115/1.2128636

Chae, S., Shin, J., Kwon, S., Lee, S., Kang, S., Lee, D.: PM10 and PM2.5 real-time prediction models using an interpolated convolutional neural network. Sci. Rep. 11, 11952 (2021). https://doi.org/10.1038/s41598-021-91253-9

Chang, C.-H., Johnson, N.C., Yoo, C.: Evaluation of subseasonal impacts of the MJO/BSISO in the East Asian extended summer. Clim. Dynam. 56, 3553–3568 (2021). https://doi.org/10.1007/s00382-021-05656-5

Cho, M., Choi, Y.-S., Kim, H.-R., Yoo, C., Lee, S.-S.: Cold-season atmospheric conditions associated with sudden changes in PM10 concentration over Seoul, Korea. Atmos. Pollut. Res. 12 (2021). https://doi.org/10.1016/j.apr.2021.03.004
Comiso, J.C.: Bootstrap sea ice concentrations from nimbus-7 SMMR and DMSP SSM/I-SSMIS, Version 3. NASA National Snow and Ice Data Center Distributed Active Archive Center (2017)

Hersbach, H., et al.: The ERA5 global reanalysis. Quart. J. Roy Met. Soc. 146, 1999–2049 (2020). https://doi.org/10.1002/qj.3803

Jeong, J.I., Park, R.J., Yeh, S.-W., Roh, J.-W.: Statistical predictability of wintertime PM2.5 concentrations over East Asia using simple linear regression. Sci. Total Environ. 776, 146059 (2021). https://doi.org/10.1016/j.scitotenv.2021.146059

Jeong, J.I., Park, R.J.: Winter monsoon variability and its impact on aerosol concentrations in East Asia. Environ. Poll. 221, 285–292 (2017). https://doi.org/10.1016/j.envpol.2016.11.075

Jia, B., Wang, Y., Yao, Y., Xie, Y.: A new indicator on the impact of large-scale circulation on wintertime particulate matter pollution over China. Atmos. Chem. Phys. 15, 11919–11929 (2015). https://doi.org/10.5194/acp-15-11919-2015

Kim, J.-H., et al.: Possible Link Between Arctic Sea Ice and January PM10 Concentrations in South Korea. Atmosphere 10, 619 (2019). https://doi.org/10.3390/atmos10100619

Kim, M., Yoo, C., Sung, M.-K., Lee, S.: Classification of wintertime atmospheric teleconnection patterns in the Northern Hemisphere. J. Clim. 34, 1847–1861 (2021). https://doi.org/10.1175/jcli-d-20-0339.1

Koo, J.-H., et al.: The implication of the air quality pattern in South Korea after the COVID-19 outbreak. Sci. Rep. 10, 22462 (2020). https://doi.org/10.1038/s41598-020-80429-4

Ku, H.-Y. et al.: Classification of large-scale circulation patterns and their spatio-temporal variability during High-PM10 events over the Korean Peninsula. Atmos. Environ. 118632 (2021). https://doi.org/10.1016/j.atmosenv.2021.118632

Kumar, N., et al.: Contributions of international sources to PM2.5 in South Korea. Atmosphere. Environ. 261, 118542 (2021). https://doi.org/10.1016/j.atmosenv.2021.118542

Lang, J., et al.: Trends of PM2.5 and chemical composition in Beijing, 2000–2015. Aerosol Air. Qual. Res. 17, 412–425 (2017). https://doi.org/10.4209/aaqr.2016.07.0307

Lee, D., et al.: Analysis of a severe PM2.5 episode in the Seoul Metropolitan area in South Korea from 27 February to 7 March 2019: focused on estimation of domestic and foreign contribution. Atmosphere 10, 756 (2019)

Lee, E.-H., Ha, J.-C., Lee, S.-S., Chun, Y.: PM10 data assimilation over south Korea to Asian dust forecasting model with the optimal interpolation method. Asia-Pacific J. Atmos. Sci. 49, 73–85 (2013). https://doi.org/10.1002/aspj.2013143-013-0009-y

Lee, G., Ho, C.-H., Chang, L.-S., Kim, J., Kim, M.-K., Kim, S.-J.: Dominance of large-scale atmospheric circulations in long-term variations of winter PM10 concentrations over East Asia. Atmos. Res. 238, 104871 (2020). https://doi.org/10.1016/j.atmosres.2020.104871

Lee, M.: An analysis on the concentration characteristics of PM2.5 in Seoul, Korea from 2005 to 2012. Asia-Pacific J. Atmos. Sci. 50, 585–594 (2014). https://doi.org/10.1007/s13143-014-0048-z

Messié, M., Chavez, F.: Global modes of sea surface temperature variability in relation to regional climate indices. J Clim. 24, 4314–4331 (2011). https://doi.org/10.1175/2011JCLI3941.1

Myoung, B., Rhee, J., Yoo, C.: Long-lead predictions of warm season droughts in South Korea Using North Atlantic SST. J Clim. 33, 4659–4677 (2020). https://doi.org/10.1175/Jcli-D-19-0082.1

NIER: Annual Report of Air Quality in Korea, 384 pp (2019)

Ok, J., et al.: How does the SST variability over the western North Atlantic Ocean control Arctic warming over the Barents–Kara Seas? Environ. Res. Lett. 12, 034021 (2017). https://doi.org/10.1088/1748-9326/aa5f3b

Rayner, N.A., et al.: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. J. Geophys. Res. Atmos. 108, 4407 (2003). https://doi.org/10.1029/2002JD002670

Shimadera, H., Hayami, H., Morino, Y., Ohara, T., Chatani, S., Hasegawa, S., Kaneyasu, N.: Analysis of summertime atmospheric transport of fine particulate matter in Northeast Asia. Asia-Pacific J. Atmos. Sci. 49, 347–360 (2013). https://doi.org/10.1007/s13143-013-0033-y

Shin, U., Park, S.-H., Park, J.-S., Koo, J.-H., Yoo, C., Kim, S., Lee, J.: Predictability of PM2.5 in Seoul based on atmospheric blocking forecasts using the NCEP global forecast system. Atmos. Environ. 246, 118141 (2021). https://doi.org/10.1016/j.atmosenv.2020.118141

Song, C., et al.: Health burden attributable to ambient PM2.5 in China. Environ. Poll. 223, 575–586 (2017). https://doi.org/10.1016/j.envpol.2017.01.060

West, J.J., et al.: What we breathe impacts our health: improving understanding of the link between air pollution and health. Environ. Sci. Tech. 50, 4985–4904 (2016). https://doi.org/10.1021/acs.est.5b03827

Wilks, D.S.: Statistical methods in the atmospheric sciences, vol. 100, 704 pp. Academic (2011)

Xiao, F., Yang, M., Fan, H., Fan, G., Al-qaness, M.A.A.: An improved deep learning model for predicting daily PM2.5 concentration. Sci. Rep. 10, 20988 (2020). https://doi.org/10.1038/s41598-020-77757-w

Yang, Y., Liao, H., Lou, S.: Increase in winter haze over eastern China in recent decades: Roles of variations in meteorological parameters and anthropogenic emissions. J. Geophys. Res. Atmos. 121, 13,050 – 13,065 (2016). https://doi.org/10.1002/2016JD025136

Yu, X., Wang, Z., Zhang, H., Zhao, S.: Impacts of different types and intensities of El Niño events on winter aerosols over China. Sci. Total Environ. 655, 766–780 (2019). https://doi.org/10.1016/j.scitotenv.2018.11.090

Yun, S., Yoo, C.: The effects of spring and winter blocking on PM10 concentration in Korea. Atmosphere 10, 410 (2019). https://doi.org/10.3390/atmos10070410

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