**Real-time $PM_{2.5}$ forecast over Delhi: Performance of high resolution (400 m) WRF-Chem model integrated with data assimilation and dynamical downscaling**

Chinmay Jena[1], Sachin D. Ghude[1], Rajesh Kumar[2], Sreyashi Debnath[1,3], Vijay K. Soni[4], Santosh H. Kulkarni[5], G. Beig[1], Ravi S. Nanjundiah[1,6], M. Rajeevan[7]

[1]Indian Institute of Tropical Meteorology, Ministry of Earth Sciences, India
[2]National Center for Atmospheric Research, Boulder, CO, 80301, USA
[3]Dept. of Atmospheric and Space sciences, Savitribai Phule Pune University, Pune, India
[4]India Meteorological Department, New Delhi, Ministry of Earth Sciences, India
[5]Centre for Development of Advanced Computing, Pune 411 008, India
[6]Centre for Atmospheric and Oceanic Sciences, Indian Institute of Science, Bengaluru 560 012, India
[7]Ministry of Earth Sciences, Prithvi Bhavan, Lodhi Road, New Delhi 110003, India

## Abstract

We present a very high-resolution (400 m) operational air quality forecasting system developed to alert citizens of Delhi and the National Capital Region (NCR) about acute air pollution episodes. Such a high-resolution system has been developed for the first time and is evaluated during October 2019-February 2020. The system assimilates near real time aerosol observations from in situ and space-borne observations in the WRF-Chem model to produce a 72-h forecast every day in a dynamical downscaling framework. The assimilation of aerosol optical depth and surface $PM_{2.5}$ observations improves the initial condition for surface $PM_{2.5}$ by about 45 μg/m$^3$ (about 50%). The accuracy of the forecast degrades slightly with time as mean bias increases from +2.5 μg/m$^3$ on the first day to -17 μg/m$^3$ on the third day of forecast. Our forecasts are found to be very capable both for $PM_{2.5}$ concentration and unhealthy/ very unhealthy air quality indices categories.

## 1. Introduction

Being the second most populated megacity in the world, Delhi faces enormous challenges due to a range of issues including adverse air pollution episodes during the winter season (Ghude et al., 2008; Beig et al., 2019; Chandra et al., 2018; Hakkim et al., 2019; Takigawa et al., 2020), which pose a higher health risk as a result of substantial population density exposure (Beig et al., 2013; Chate et al., 2013; Ghude et al., 2016; Guttikunda and Goel, 2013; Spears et al., 2019). In recent years, particulate matter of aerodynamic diameter smaller than 2.5 µm($PM_{2.5}$) has dominated severe air pollution episodes frequently observed during the winter season, affecting daily life in Delhi (Chowdhury et al., 2019; Jethva et al., 2019). Thus, managing air quality with practical mitigation options has emerged as a complicated task without compromising the intended growth in the overall infrastructure development, industries, and service sectors.

The Government of India(GOI) is committed to enforce policy-driven measures to reduce the pollutant emissions. The National Clean Air Program(NCAP) initiated by the GOI targets significant reduction of surface $PM_{2.5}$ concentration by the year 2024. A Graded Response Action Plan(GRAP) has been designed for the National Capital Region(NCR) that allows pollution control authorities to reduce the magnitude of predicted air pollution for different air quality indices(AQI) categories. Activation of different GRAP measures requires information of forthcoming extreme air pollution episodes so that effective temporary control measures can be identified early and implemented in advance. Therefore, GOI mandate required the Ministry of Earth Science(MoES) to develop an operational high-resolution modelling for the NCR region. In response to the mandate, MoES institutions, Indian Institute of Tropical Meteorology(IITM) and India Meteorology Department(IMD), has developed a first very high-resolution (400 m) chemical weather forecasting capability in mutual collaboration with U.S. National Centre for Atmospheric Research(NCAR). The initial capability was developed for Delhi during 2018 based on weather research and forecasting model coupled with online chemistry module(WRF-Chem) at horizontal resolution of 2 km (Ghude et al., 2020). The first version of the forecasting system has been found to significantly improve air quality decision-making activity by reducing biases in 72-h $PM_{2.5}$ forecasts to a greater extent (Kumar et al., 2020). However, that system struggled to predict the absolute $PM_{2.5}$ levels during very acute air pollution episodes characterized by surface $PM_{2.5}$ mass concentrations greater than 350 μg/m$^3$ (Ghude et al., 2020).

To further enhance the air quality decision-making activity the modelling framework was further extended to produce forecast fields of $PM_{2.5}$ concentrations at much higher resolution of 400 meters. To the best of our knowledge, none of the operational centres are currently providing short-term operational air quality forecasts at a spatial scale of 400 x 400 m$^2$ covering approximately 50 km$^2$ areas. This is the first attempt to develop and evaluate the performance of $PM_{2.5}$ forecast in a highly polluted environment using integration of dynamical models with chemical data assimilation. This high-resolution forecasting system consists of a newly developed high-resolution (400 m) emission inventory for Delhi, assimilation of satellite aerosol optical depth(AOD) and surface $PM_{2.5}$ concentrations from a dense air quality monitoring network in Delhi to constrain the regional and local background of aerosols, and ingests near real-time fire emissions, and applies high-resolution dynamical downscaling into the WRF-Chem. $PM_{2.5}$ forecasts at 400 m resolution were made operational in October 2019.

Here, we provide a brief description of the operational high-resolution forecasting system, highlight the impact of data assimilation, and evaluate the quality of the $PM_{2.5}$ operational forecast for the latest winter season in Delhi. We show that the forecast falls within the expected uncertainties and, therefore is suitable for operational air quality early warning for the Delhi region.

## 2. Air quality forecasting system

The complete overview of the air quality early warning system is given in the schematic shown in Figure S1. The major components of the system are a) high-resolution numerical predication model WRF-Chem, b) gridpoint statistical interpolation (GSI) based three-dimensional variational(3D-var) data assimilation integrated with dynamical downscaling, c) observational data and emission pre-processer, d) post analysis, and e) public dissemination system(https://ews.tropmet.res.in/). The system operates overnight and disseminates the information in the morning for the next 72 hours.

### 2.1 Modelling setup

The core of the forecasting system consists of the regional WRF-Chem Version 3.9.1 configured in a three domain set-up with the outer domain covering northern part of the Indian subcontinent at a horizontal resolution of 10 km, the second domain covering the NCR and neighboring state at 2 km resolution, and the innermost domain covering Delhi at 400 m resolution (Figures S1 and S2). The meteorological initial and boundary conditions are based on the analysis and forecast product (Ensemble-Kalman filtering) produced by the IITM-Global Forecasting System (IITM-GFS, T1534) spectral model at 12.5 km grid resolution available at every three hours. The outer domain (D1) is provided with the six-hourly chemical boundary conditions from the MOZART-4 10-year climatology. However, chemistry output from the D1 domain every three-hour interval was dynamically used to establish a chemical boundary for the WRF-Chem inner domain (D2). Similarly, the output from the D2 domain was dynamically used to establish boundary conditions for the final 400 m domain (D3). Physical and chemical parameterizations used in the model are listed in Supplementary Table ST1. The forecast was run with MOZART-4 gas-phase chemistry linked to GOCART aerosol scheme(MOZCART). In the GOCART, aerosol nitrates and secondary organic aerosols are missing and are a source of uncertainty in the estimated $PM_{2.5}$. However, this scheme is computationally most efficient than the other advanced aerosol schemes (e.g., MOZAIC), and therefore it is suitable for the operational forecast.

### 2.2 Satellite AOD and Surface $PM_{2.5}$ data assimilation

We used three-dimensional variational method(3D-Var) component of the community Gridpoint Statistical Interpolation(GSI) system Version 3.5. The 3D-var scheme blends the information from the satellite AOD and surface $PM_{2.5}$ observations, and iteratively minimizes a cost function $J$ that depends on observation and background error covariance matrices as defined in equation (1).

$$\mathrm{J}(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x_b})^{\mathrm{T}}\mathbf{B}^{-1}(\mathbf{x} - \mathbf{x_b}) + \frac{1}{2}(\mathrm{H}(\mathbf{x}) - \mathbf{y})^{\mathrm{T}}\mathbf{R}^{-1}(\mathrm{H}(\mathbf{x}) - \mathbf{y}) \qquad (1)$$

where $\mathbf{x}$ represents the state vector required in AOD calculation, $\mathbf{x_b}$ represents the "a priori" information about $\mathbf{x}$ and is referred to as background, $\mathbf{B}$ is the background error covariance (BEC) matrix, H is the forward operator that transforms WRF-Chem aerosol chemical composition to AOD following Liu et al. (2011), $\mathbf{y}$ represents the MODIS AOD retrievals, and $\boldsymbol{R}$ is the observation error covariance matrix. BEC statistical parameters are calculated using two 24-h WRF-Chem forecast initialized at 00 z with different meteorological, anthropogenic emissions and biomass burning emissions to account for the uncertainties in both meteorology, anthropogenic, and biomass burning emissions (Kumar et al., 2019).

Observations from MODIS overpasses at 10:30 and 1:30 L.T. at 10 km resolutions and hourly mean surface $PM_{2.5}$ observations from the 37 monitoring stations (Figure S3) across Delhi are collected at the analysis time on 0900 UTC assimilation cycle. GOCART has sixteen aerosol species, and all are adjusted directly by AOD and surface $PM_{2.5}$ assimilation. More details on the AOD assimilation for the GOCART scheme can be found elsewhere (Ghude et al., 2020; Kumar et al., 2020). Every day, the chemical fields are initialized from the previous WRF-Chem forecast at 0900 UTC, aerosol initialization is updated through assimilation, and meteorology is refreshed using the IITM-GFS forecast. To understand how much the assimilation of surface $PM_{2.5}$ and satellite AOD observations changes the WRF-Chem $PM_{2.5}$ at the initialization time on every day towards the observed reality, averaged surface $PM_{2.5}$ simulated by the model before and after assimilation are compared with the observations at 0900 UTC assimilation cycle (Figure 1). Before assimilation, the simulated $PM_{2.5}$ at the initialization time significantly underestimates the observed $PM_{2.5,}$ and assimilation pushes WRF-Chem initial condition very close to the observations. The average observed $PM_{2.5}$ is about 133±88 µg/m$^3$, whereas simulated $PM_{2.5}$ before assimilation and after assimilation is estimated as 90±36 µg/m$^3$ (mean bias = -32%) and 135±83 µg/m$^3$ (mean bias = +1%), respectively. This indicates that assimilation on an average improves the initial condition for $PM_{2.5}$ by ~45 µg/m$^3$ (about 50%) at the assimilation cycle.

### 3.3. Emissions and observations

Two anthropogenic emission databases are included in the forecasting setup. We used 2010 EDGAR-HTAP emissions at 10 km resolution for the outer (D1) and intermediate (D2) domains and scaled to 2019 using scaling factors given in Venkataraman et al. (2018). This data was used only for the area outside the Delhi region. For Delhi itself (D3), we used 400 m resolution emission inventory for the year 2018 developed under the MoES(Figure S4)

System of Air Quality and Weather Forecasting and Research(SAFAR) project. For the outer and intermediate domain, the original 400 m emissions were processed using a mass-conserving approach to match the 10 km and 2 km grid spacing so that the total mass emitted is the same before and after re-grinding. We adopted diurnal variation in emissions from a recent study by Govardhan et al. (2019). We performed emission sensitivity simulations by keeping EDGAR emissions for the entire model domain and then replacing the Delhi region with high-resolution Delhi emission inventory. We find that $PM_{2.5}$ mass concentration simulated by the EDGAR largely under-predict (bias = -52%) the observed $PM_{2.5}$ mass concentrations in Delhi, whereas as high-resolution inventory over-predict by 36% during the winter season (Figure S5). Additional simulation where emissions are reduced by 40% shows better agreement with observations (bias =3%). Therefore, in the forecasting setup, we choose emissions reduced by 40% over the Delhi region.

Model of Emissions of Gases and Aerosols from Nature (MEGAN; Guenther, 2007) is used to include interactive biogenic emissions. Dust emissions are based on the online AER/AFWA scheme (Jones and Creighton, 2011). Our recent study shows that post-monsoon biomass burning emission significantly affects the air quality in Delhi (Kulkarni et al., 2020). Therefore, an accurate representation of fire emissions is essential for the quality of the forecast. Most of the near real-time biomass burning emission estimate is available with a time lag of one day and thus operational air quality forecasts are forced to assume persistent fire emissions over the forecast cycle. Here, we have developed a pre-processor based on the high-resolution Fire INventory from NCAR (FINN, Wiedinmyer et al., 2011) to derive fire emissions for the forecast day instead of using the estimates from a previous day. We first developed a historical daily gridded (at the model grid spacing) data set of biomass burning emissions from 2010 to 2018. Secondly, daily fire location information is obtained from the near real-time MODIS-C6 active fire data from FIRMS (https://firms.modaps.eosdis.nasa.gov/). Finally, the historical fire emissions are used to represent fire emissions in all the model grids showing the fire activity. To include the fire emissions for the next two days, we calculated the historical fire frequency for each day and for each model grid based on a 10-year gridded data set. We include the fire emissions only in those grid cells where frequency of fire count is more than 50%.

The $PM_{2.5}$ data set used in the study are obtained in near real-time from the 37 air quality monitoring stations operated by the Central Pollution Control Board(CPCB), Delhi Pollution Control Committee(DPCC), and IITM. The instruments are calibrated, and the measurements are quality controlled (https://cpcb.nic.in/quality-assurance-quality-control/). In addition to the CPCB quality control, we apply additional filters to remove single spikes, $PM_{2.5}$ measurements above 1500 $\mu g/m^3$, observations from the period corresponding to instrument malfunction (Kumar et al. 2020). The details of these monitoring locations are given in Table ST2 in the supplement, and the geographical locations are shown in Figure S3. Statistical evaluation metrics such as mean bias(MB), Pearson's correlation coefficient(r), normalized mean bias(NMB), normalized mean fractional error(NME) (Yu et al., 2006; Zhang et al., 2006) are used to evaluate the performance of $PM_{2.5}$ forecasts.

## 3. Result

### 3.1 Performance of the $PM_{2.5}$ forecast

The air quality forecast verification period (21 October 2019–02 February 2020) selected here for Delhi was dominated by the large-scale open biomass burning in October and November followed by wintertime stable meteorological conditions conducive for build-up of

$PM_{2.5}$ pollution in Delhi. An example of the spatial distribution of average $PM_{2.5}$ concentration from day 1 forecast is shown in Figure 2a. As expected, the 400 m domain resolves emission sources in Delhi much better than the 10 km and 2 km domains (Figure S6). At finer grid resolution, $PM_{2.5}$ hotspots associated with the industrial, dense residential, major traffic junctions, and high-density vehicular traffic roads can be clearly distinguished from Figure 2a. Daily forecast (Figure 2b) tracts the daily variation of mean observations quite well in the NCR region. However, model performance for the $PM_{2.5}$ forecast for individual monitoring stations (supplementary table ST2) generally shows that the NMB varies from -46% to 85% among the 37 stations located across the NCR region. Out of 37 stations, 24 individual stations (65% of stations) show NMB within ±30%, and 5 stations (13% of stations) show NMB more than 50%.

To more closely examine the model performance for the $PM_{2.5}$ forecast specific to the NCR, we compared the hourly time series of mean $PM_{2.5}$ forecast with observations. The NCR region contains 37 monitoring stations, and model forecasts from the locations of all these monitoring stations are extracted and average to get $PM_{2.5}$ time series for Delhi. The comparison for the first, second, and third day of forecast at 400 m resolution domain is depicted in Figure 3. The forecast captures the temporal variability in $PM_{2.5}$ observation quite well on most of the days. On two occasions, 2–4 November and 12–16 November, the model largely failed to capture extremely high $PM_{2.5}$ values observed in the NCR region. However, very-poor pollution events occurred on 20–22 November, and very-poor to sever air quality events on 5–12 December, 19–22 December, 19–21 December, 29 December 2019–5 January 2020 were captured very well by the forecast. The sudden drop in $PM_{2.5}$ levels (e.g., on 14 December, 7 January, 24 January, etc.) followed by the very-poor to sever events were also captured very well by the model. The temporal variation in observed $PM_{2.5}$ over Delhi and pollution episodes are driven mainly by the frequent large scale open biomass burning and wintertime synoptic-scale meteorological condition in combination with the large anthropogenic emissions in this region. Figure 3 shows that the model has a very good ability to capture this variability and indicates that the forecast system is very capable of issuing $PM_{2.5}$ forecasts associated with urban pollution and synoptic-scale meteorological events.

The performance statistics for hourly mean $PM_{2.5}$ forecast for the first, second and third day is evaluated by examining MB, RMSE, r, NMFB, and NMFE (Table 1). Following Morris et al. (2005), we have adopted three levels of performance criteria for fractional bias and error to evaluate the forecast performance (supplementary table ST3). It can be seen that the magnitude of MB on the first day of hourly forecast was quite low (2.5 $\mu g/m^3$), and slightly under predicted by about -9 $\mu g/m^3$ and -17 $\mu g/m^3$ on the second and third day of the forecast, respectively. Table 1 reveals that $PM_{2.5}$ forecast performed close to excellent criteria on day 1 since the NMFB and NMFE were within 1.3% and 36.3%, respectively. The model performance was quite good on day 2 and day 3 with fairly low NMFB (<±10%), and NMFE was within ±40%. Performance statistics on daily mean $PM_{2.5}$ time series show fairly excellent performance on all the three days of forecast since NMFB and NMFE were within ±10% and 29%, respectively. The correlation coefficient (r) is around 0.5 for the hourly forecast and 0.6 for the daily mean forecast. It appears from Figure S7 that there is a considerable scatter between observed and predicted $PM_{2.5}$. Forecasts generally show a tendency to under predict the higher values and slightly over predict the lower values, consistent with the other studies (Eder et al., 2006, 2010). Efforts are underway to investigate why this underestimation in higher values occurs and whether it is the result of errors in the meteorological, boundary condition, emission, or chemical factors.

### 3.2 Performance of the Air Quality Index (AQI) forecast

The main interest in the application of high-resolution forecast is to provide timely air quality alerts to the public in the NCR region, particularly for the Poor to Severe AQI category. As per the CPCB guidelines, AQI category is classified as poor for AQI range 201-300, very-poor for AQI range 301-400, and severe for AQI range 401 and above. Therefore, it is essential to evaluate the applicability of the system to simulate the correct AQI values, as these are typically disseminated to the general public rather than the actual $PM_{2.5}$ mass concentrations. Accordingly, hourly AQI values for $PM_{2.5}$ based on 24-h $PM_{2.5}$ standard was calculated based on National Ambient Air Quality Standard(NAAQS), and break-point concentration suggested in the CPCB notification (see supplementary Table ST4). It can be seen in Table 1 that the magnitude of MB for overall AQI values (0–500) was slightly higher (about 22 units) on the first day of forecast compared to MB observed on the second day (10 unit) and third day (<1 unit). However, following the Morris et al., (2005) criteria, the AQI forecast on all three days performed excellently, since NMFB and NMFE were within ±6% and 18%, respectively. The statistical performance (Supplementary Table ST5) indicates that the forecast over-predicts poor air quality category by about 22% on day one and by about 19% and 16% on days two and three, respectively. AQI forecast on all three days in poor category performed fairly good since NMFB and NMFE were within ±22% and 22%, respectively. For a very-poor category, the AQI forecast performed excellent on all three days (NMFB <5% and NMFE <9%) of the forecast. In comparison, statistics show that forecast under-predicts the severe category by about 14% on day one and by about 17% and 21% on day two and day three, respectively. On the first day, the AQI forecast for the severe category performed excellently (NMFB <14% and NMFE <17%), while it performed reasonably well on day two and day three (NMFB <21% and NMFE <21%). Overall, the forecast falls within the expected uncertainties and, therefore, is suitable for operational air quality forecast for the Delhi region.

### 3.3 Skill score for categorical AQI forecast

To assess the skill of real-time forecast, that is, whether the forecast will fall in unhealthy (AQI >201), or very-unhealthy (AQI >301) or critical category (AQI >401), false alarm rate(FAR), probability of detection(POD) or hit rate, critical success index(CIS) and accuracy, was calculated according to Kang et al. (2005) and Eder et al. (2010). Supplementary table ST6 describes the equations that are used to calculate the skill score of the categorical AQI forecast. Table 1 presents the skill score for the unhealthy, very-unhealthy, and critical category of AQI for winter season. The forecast accuracy for unhealthy category (i.e., forecast that correctly predicted the unhealthy or no-unhealthy AQI) showed the accuracy values to be >88% on all the three days. The skill score for the POD and CSI is relatively promising with a value grater that ~0.9, which indicates that the model has reasonable good predictive accuracy in predicting the unhealthy air quality conditions with respect to a total number of the observed air quality hours. Also note that FAR is quite low (~10%) for the unhealthy category, which indicates that the performance of the real-time high-resolution forecast was excellent for both unhealthy category and non-unhealthy category of air quality. For the very-unhealthy category, although the skill score of POD (>0.9) is excellent, compared to the unhealthy category, the skill score for the CSI is a little low (~0.7), and FAR is a little high (20-30%) on all three days of forecast. However, the skill score overall indicates excellent performance for predicting air quality in the very-unhealthy category. On the other hand, the skill score shows moderate performance for the severe category of events on all three days. Table 1 indicates that compared to the unhealthy and very-unhealthy category, the skill score for POD and CSI is low (<0.35) on day one and day two of forecast and further declines on day three. Note that the accuracy of the forecast is about 80%, and the skill score for FAR does not show a significant increase compared to the other two categories, which indicates the moderate ability of the forecasting system to predict the extremely high pollution events in NCR region correctly. Some of the causes of unsuccessful prediction of accurate extremely high pollution events may include difficulties in simulating processes like boundary layer height, synoptic advection, and synoptic-scale conditions and choice of aerosol module. Our future studies will investigate the role of these processes in predicting the extreme $PM_{2.5}$ pollution events in the NCR region.

## 4. Conclusion

This study aims to demonstrate the efficacy of a prototype very high-resolution operational air quality forecasting system developed to issue timely warning to the citizens of Delhi about forthcoming air pollution episodes. The system is based on the WRF-Chem model integrated with satellite AOD and surface $PM_{2.5}$ data assimilation and dynamical downscaling. Performance of the system was evaluated both for $PM_{2.5}$ mass concentration and AQI categories to assist both local forecaster and air quality model developers. Both AOD and surface $PM_{2.5}$ data assimilation, on an average, improves the initial condition for $PM_{2.5}$ by about by ~45 $\mu g/m^3$ (about 50%) at the assimilation cycle. Model evaluation shows that finer-resolution $PM_{2.5}$ forecasting system is capable of issuing reliable forecasts for Delhi during winter season both for $PM_{2.5}$ concentration and AQI in particular for unhealthy and very-unhealthy AQI categories, within the expected uncertainties. On the other hand, verification statistics with reference to severe categories show moderate skill and may require further improvement in the forecast. Although finer-resolution $PM_{2.5}$ forecast captures the spatial variability in emissions, it is desirable to have an accurate forecast at every individual station with a high-resolution forecast. The model showed moderate performance to capture the accurate spatial concentration across the NCR region and requires further improvement. We recognize that the accuracy of the high-resolution emission inventory, choice of aerosol model, chemical mechanism, and boundary layer parameterization will continue to be a challenge to forecast the $PM_{2.5}$ at individual point monitoring locations in a highly polluted environment. Efforts are underway to explore the sensitivity of these parameters to the accuracy of location-specific $PM_{2.5}$ forecast. For the first time, the system was also used by the environmental pollution control authorities to make the decision on imposing/lifting the restriction on construction activities and regulating the heavy vehicle inflow in Delhi region during the pollution/no-pollution events. This has significantly contributed to the build-up the trust to the end-users and policy-makers for taking science-based well-informed decisions and actions for important public services in India.

**Acknowledgment:**

We acknowledge the use of surface $PM_{2.5}$ data from air quality monitoring network of the CPCB, India. Observational data on $PM_{2.5}$ measurements can be obtained from CPCB website on https://app.cpcbccr.com/ccr. This work was supported by the National Supercomputing Mission(NSM) program grant to the authors at C-DAC, and we are grateful to the Executive Director and the Director-General of C-DAC. This material is based upon work supported by the NCAR, which is a major facility sponsored by the NSF under Cooperative Agreement No. 1852977. MODIS AOD retrievals used for assimilation can be downloaded from this site (https://earthdata.nasa.gov/). The model data is available at Aditya (IITM) super-computer and can be provided upon request to corresponding author.

**Reference:**


Beig, G., Chate, D. M., Ghude, S. D., Mahajan, A. S., Srinivas, R., Ali, K., et al. (2013). Quantifying the effect of air quality control measures during the 2010 Commonwealth Games at Delhi, India. *Atmospheric Environment*, *80*, 455–463. https://doi.org/10.1016/j.atmosenv.2013.08.012

Beig, G., Srinivas, R., Parkhi, N. S., Carmichael, G. R., Singh, S., Sahu, S. K., et al. (2019). Anatomy of the winter 2017 air quality emergency in Delhi. *Science of the Total Environment*, *681*, 305–311. https://doi.org/10.1016/j.scitotenv.2019.04.347

Chandra, B. P., Sinha, V., Hakkim, H., Kumar, A., Pawar, H., Mishra, A. K., et al. (2018). Odd-even traffic rule implementation during winter 2016 in Delhi did not reduce traffic emissions of VOCs, carbon dioxide, methane and carbon monoxide. *Current Science*, *114*(6), 1318–1325. https://doi.org/10.18520/cs/v114/i06/1318-1325

Chate, D., Beig, G., Satpute, T., Sahu, S. K., Ali, K., Parkhi, N., & Ghude, S. (2013). Assessments of population exposure to environmental pollutants using air quality measurements during Commonwealth Games-2010. *Inhalation Toxicology*, *25*(6), 333–340. https://doi.org/10.3109/08958378.2013.788103

Chowdhury, S., Dey, S., Di Girolamo, L., Smith, K. R., Pillarisetti, A., & Lyapustin, A. (2019). Tracking ambient PM2.5 build-up in Delhi national capital region during the dry season over 15 years using a high-resolution (1 km) satellite aerosol dataset. *Atmospheric Environment*, *204*(July 2018), 142–150. https://doi.org/10.1016/j.atmosenv.2019.02.029

Eder, B., Kang, D., Mathur, R., Yu, S., & Schere, K. (2006). An operational evaluation of the Eta-CMAQ air quality forecast model. *Atmospheric Environment*, *40*(26), 4894–4905. https://doi.org/10.1016/j.atmosenv.2005.12.062

Eder, B., Kang, D., Rao, T., Mathur, R., Yu, S., Otte, T., et al. (2010). Using national air quality forecast guidance to develop local air quality index forecasts. *Bulletin of the American Meteorological Society*, *91*(3), 313–326. https://doi.org/10.1175/2009BAMS2734.1

Ghude S.D., Jain S.L., Arya B.C., Beig G., Ahammed Y.N., Arun Kumar, Tyagi B (2008). Ozone in ambient air at a tropical megacity, Delhi: characteristics, trends and cumulative ozone exposure indices. *Journal of Atmospheric Chemistry*, 60, 237-252. d0i 10.1007/s10874-009-9119-4.

Ghude, S. D., Chate, D. M., Jena, C., Beig, G., Kumar, R., Barth, M. C., et al. (2016). Premature mortality in India due to PM2.5 and ozone exposure. *Geophysical Research Letters*, *43*(9), 4650–4658. https://doi.org/10.1002/2016GL068949

Ghude, S. D., Kumar, R., Jena, C., Debnath, S., Kulkarni, R. G., Alessandrini, S., et al. (2020). Evaluation of PM2.5 forecast using chemical data assimilation in the WRF-Chem model: A novel initiative under the Ministry of Earth Sciences Air Quality Early Warning System for Delhi, India. *Current Science*, *118*(11), 1803–1815. https://doi.org/10.18520/cs/v118/i11/1803-1815

Govardhan, G., Satheesh, S. K., Moorthy, K. K., & Nanjundiah, R. (2019). Simulations of Black Carbon Over Indian Region: Improvements and Implications of Diurnality in Emissions. *Atmospheric Chemistry and Physics Discussions*, 1–25. https://doi.org/10.5194/acp-2019-152

Guenther, A. (2007). Erratum: Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature) (Atmospheric Chemistry and Physics (2006) 6 (3181-3210)). *Atmospheric Chemistry and Physics*, *7*(16), 4327. https://doi.org/10.5194/acp-7-4327-2007

Guttikunda, S. K., & Goel, R. (2013). Health impacts of particulate pollution in a megacity-

Delhi, India. *Environmental Development*, *6*(1), 8–20. https://doi.org/10.1016/j.envdev.2012.12.002

Hakkim, H., Sinha, V., Chandra, B. P., Kumar, A., Mishra, A. K., Sinha, B., et al. (2019). Volatile organic compound measurements point to fog-induced biomass burning feedback to air quality in the megacity of Delhi. *Science of The Total Environment*, *689*, 295–304. https://doi.org/10.1016/j.scitotenv.2019.06.438

Jethva, H., Torres, O., Field, R. D., Lyapustin, A., Gautam, R., & Kayetha, V. (2019). Connecting Crop Productivity, Residue Fires, and Air Quality over Northern India. *Scientific Reports*, *9*(1), 1–11. https://doi.org/10.1038/s41598-019-52799-x

Jones, S. and Creighton, G.: AFWA dust emission scheme for WRF/Chem-GOCART, 2011 WRF workshop, 20–24 June 2011, Boulder, CO, USA, 2011

Kang, D., Eder, B. K., Stein, A. F., Grell, G. A., Peckham, S. E., & McHenry, J. (2005). The New England Air Quality Forecasting Pilot Program: Development of an Evaluation Protocoland Performance Benchmark. *Journal of the Air & Waste Management Association, 55:12*, 1782-1796, DOI: 10.1080/10473289.2005.10464775

Kulkarni, S. H., Ghude, S. D., Jena, C., Karumuri, R. K., Sinha, B., Sinha, V., et al. (2020). How Much Does Large-Scale Crop Residue Burning Affect the Air Quality in Delhi? *Environmental Science & Technology*, *54*(8), 4790–4799. https://doi.org/10.1021/acs.est.0c00329

Kumar, R., Delle Monache, L., Bresch, J., Saide, P. E., Tang, Y., Liu, Z., et al. (2019). Toward Improving Short-Term Predictions of Fine Particulate Matter Over the United States Via Assimilation of Satellite Aerosol Optical Depth Retrievals. *Journal of Geophysical Research: Atmospheres*, *124*(5), 2753–2773. https://doi.org/10.1029/2018JD029009

Kumar, R., Ghude, S. D., Biswas, M., Jena, C., & Alessandrini, S. (2020). Enhancing Accuracy of Air Quality and Temperature Forecasts During Paddy Crop Residue Burning Season in Delhi Via Chemical Data Assimilation, *Journal of Geophysical Research Atmospheres*, 125. https://doi.org/10.1029/2020JD033019.

Liu, Z., Liu, Q., Lin, H. C., Schwartz, C. S., Lee, Y. H., & Wang, T. (2011). Three-dimensional variational assimilation of MODIS aerosol optical depth: Implementation and application to a dust storm over East Asia. *Journal of Geophysical Research Atmospheres*, *116*(23), 1–19. https://doi.org/10.1029/2011JD016159.

Morris, R. E., Mc Nally, D. E., Tesche, T. W., Tonnesen, G., Boylan, J. W., & Brewer, P. (2005). Preliminary Evaluation of the Community Multiscale Air Quality Model for 2002 over the Southeastern United States. *Journal of the Air and Waste Management Association*, *55*(11), 1694–1708. https://doi.org/10.1080/10473289.2005.10464765

Spears, D., Dey, S., Chowdhury, S., Scovronick, N., Vyas, S., & Apte, J. (2019). The association of early-life exposure to ambient PM2.5 and later-childhood height-for-age in India: An observational study. *Environmental Health: A Global Access Science Source*, *18*(1), 1–10. https://doi.org/10.1186/s12940-019-0501-7

Takigawa, M., Patra, P. K., Matsumi, Y., Dhaka, S. K., Nakayama, T., Yamaji, K., et al. (2020). Can Delhi's Pollution be Affected by Crop Fires in the Punjab Region? *Sola*, *16*(0), 86–91. https://doi.org/10.2151/sola.2020-015

Venkataraman, C., Brauer, M., Tibrewal, K., Sadavarte, P., Ma, Q., Cohen, A., et al. (2018). Source influence on emission pathways and ambient PM 2.5 pollution over India (2015-2050). *Atmospheric Chemistry and Physics*, *18*(11), 8017–8039. https://doi.org/10.5194/acp-18-8017-2018

Wiedinmyer, C., Akagi, S. K., Yokelson, R. J., Emmons, L. K., Al-Saadi, J. A., Orlando, J. J., & Soja, A. J. (2011). The Fire INventory from NCAR (FINN): A high resolution global model to estimate the emissions from open burning. *Geoscientific Model*

*Development*, *4*(3), 625–641. https://doi.org/10.5194/gmd-4-625-2011
Yu, S., Eder, B., Dennis, R., Chu, S.-H., & Schwartz, S. E. (2006). New unbiased symmetric metrics for evaluation of air quality models. *Atmospheric Science Letters*, *7*(1), 26–34. https://doi.org/10.1002/asl.125
Zhang, Y., Liu, P., Pun, B., & Seigneur, C. (2006). A comprehensive performance evaluation of MM5-CMAQ for the Summer 1999 Southern Oxidants Study episode-Part I: Evaluation protocols, databases, and meteorological predictions. *Atmospheric Environment*, *40*(26), 4825–4838. https://doi.org/10.1016/j.atmosenv.2005.12.043

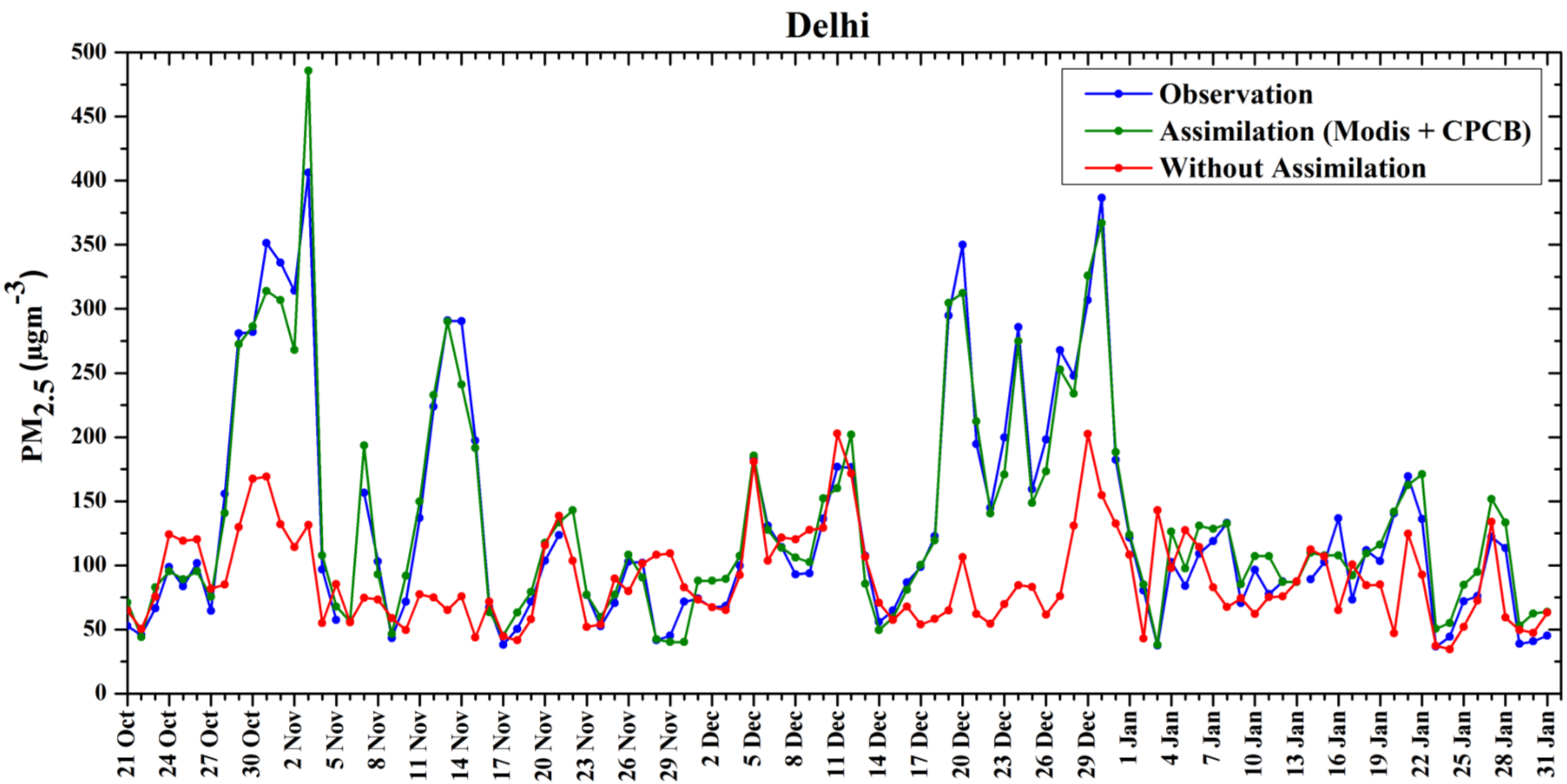


**Figure 1:** Averaged surface $PM_{2.5}$ simulated by the model before (red) and after (green) assimilation at 0900 UTC assimilation cycle (Figure 1) and its comparison with observed mean $PM_{2.5}$ over Delhi during 21 October 2019 to 01 February 2020.

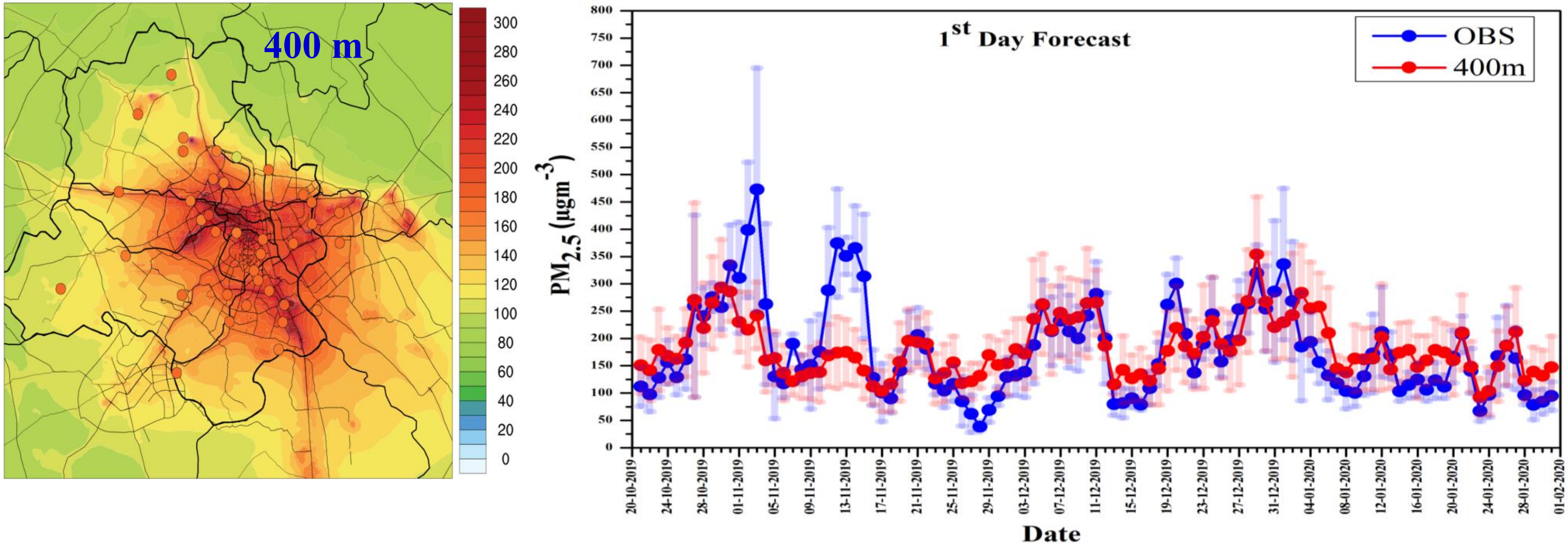


**Figure 2: (a)** Spatial distribution of averaged $PM_{2.5}$ at 400 m horizontal resolution (from day 1 forecast) overlay with mean $PM_{2.5}$ observed at different monitoring stations across Delhi during 21 October 2019 to 01 February 2020, **(b)** Comparisons between daily mean $PM_{2.5}$ forecast (red) and daily mean $PM_{2.5}$ observations (blue) over Delhi during 21 October 2019 to 01 February 2020 (vertical bar shows the standard deviation).

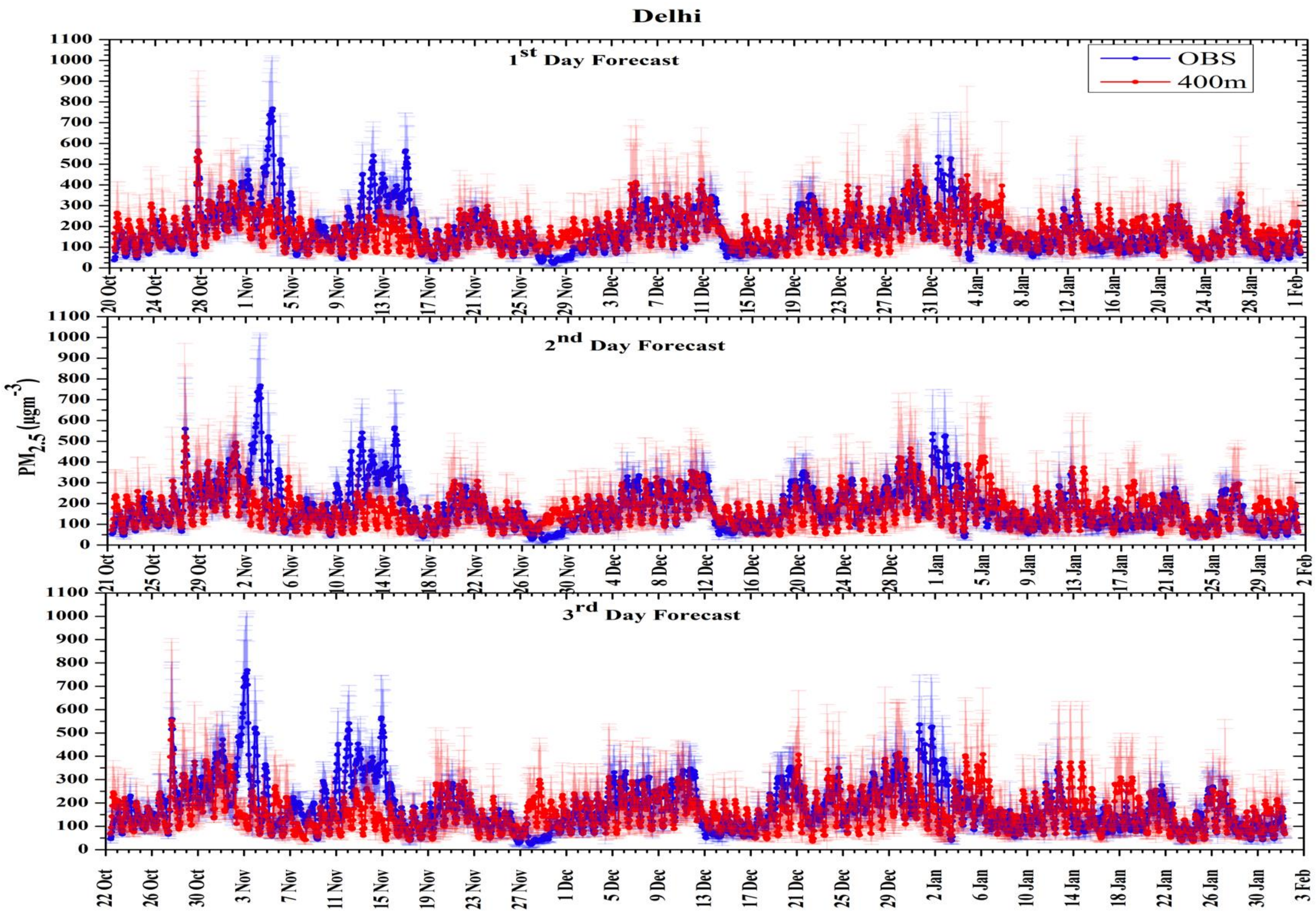


**Figure 3:** Comparisons between hourly mean $PM_{2.5}$ forecast (red) and hourly mean $PM_{2.5}$ observations (blue) on day one (top), day two (middle) and day three at 400 meter horizontal resolution over Delhi during 21 October 2019 to 01 February 2020 (vertical bar shows the standard deviation).

**Table 1:** Performance statistics for mean $PM_{2.5}$ forecast and skill score for different forecast AQI category over Delhi during 21 October 2019 to 01 February 2020.

.

| **Statistical performance** | | | | | |
|---|---|---|---|---|---|
| State Variables | Forecast Day | **400 meter** | | | |
| | | **MB** | **NMFB (%)** | **NMFE (%)** | **r** |
| $PM_{25}$_hourly | 1st day | 2.5 | 1.3 | 36.3 | 0.5 |
| | 2nd day | -8.4 | -4.8 | 38.1 | 0.5 |
| | 3rd day | -16.8 | -9.8 | 40.5 | 0.4 |
| $PM_{25}$_daily | 1st day | 1.8 | 1.0 | 25.6 | 0.6 |
| | 2nd day | -8.8 | -5.0 | 26.7 | 0.6 |
| | 3rd day | -17.3 | -10.1 | 29.5 | 0.5 |
| $PM_{25}$_AQI | 1st day | 21.7 | 6.5 | 16.5 | 0.7 |
| | 2nd day | 10.4 | 3.1 | 16.5 | 0.6 |
| | 3rd day | 0.3 | 0.1 | 17.8 | 0.5 |
| **Skill score for AQI** | | | | | |
| AQI range | Forecast Day | **FAR** | **POD** | **CSI** | **Accuracy** |
| Unhealthy (poor, very-poor, severe) | 1st day | 0.11 | 1.00 | 0.88 | 0.88 |
| | 2nd day | 0.09 | 0.99 | 0.90 | 0.90 |
| | 3rd day | 0.09 | 0.98 | 0.88 | 0.88 |
| Very Unhealthy (very poor, severe) | 1st day | 0.28 | 0.98 | 0.70 | 0.72 |
| | 2nd day | 0.25 | 0.94 | 0.71 | 0.75 |
| | 3rd day | 0.23 | 0.89 | 0.70 | 0.74 |
| Critical (severe) | 1st day | 0.35 | 0.34 | 0.29 | 0.82 |
| | 2nd day | 0.15 | 0.35 | 0.33 | 0.85 |
| | 3rd day | 0.25 | 0.21 | 0.19 | 0.82 |