Supplement of Atmos. Chem. Phys., 21, 2837–2860, 2021
https://doi.org/10.5194/acp-21-2837-2021-supplement

Supplement of

Improving regional air quality predictions in the Indo-Gangetic Plain – case study of an intensive pollution episode in November 2017

Behrooz Roozitalab et al.

Correspondence to: Behrooz Roozitalab (behrooz-roozitalab@uiowa.edu)
and Gregory R. Carmichael (gcarmich@engineering.uiowa.edu)

S.1. Effect of additional filters on CPCB data

We did not apply any filter to this data as we relied on quality control done by CPCB (https://cpcb.nic.in/quality-assurance-quality-control/). However, we studied how applying the following filters, done by Jena et al. (2020) and Kumar et al. (2020), change the dataset consisting of total 12768 hourly data points:

Filter 1: Remove less than 10 μgm^{-3} instances: removes 31 data-points

Filter 2: Remove the hourly difference between 100 (or 150 or 200) μgm^{-3} : removes 186 (or 71 or 31) hourly-data

Filter 3: Remove values more than 200 (400) μgm^{-3} right after NAN value: 33 (19). It basically removes data for Nov. 9th as it was applied after filter #2.

We found that the order of applying these filters is important. Figure S1 and Table S1 shows statistics and timeseries for different orders of filters. Order of filters (1,2,3) removes data for Nov. 9th and significantly improves the model performance over Delhi.

S.2 Sensitivity to changes in boundary conditions data

Figure S2 shows the hourly averaged $PM_{2.5}$ and PM_{10} concentration maps during the studied period using four different boundary conditions as described in methods section. The major difference between these maps is on the western parts of the domain. The conceptual model in Beig et al. (2019) suggested that long range transported dust coming from Pakistan and Middle East influenced air quality in northern India during this period. FINN_MERRA2 simulation had the highest values for both $PM_{2.5}$ and PM_{10}, which shows that some parts of the domain were affected by pollution from the boundaries. FINN_CAMS simulation shows lower concentrations, which can be attributed to CAMS assimilation technique. On the other hand, FINN_MOZART and FINN_CAMCHEM scenarios are very similar to each other. Overall, data assimilation as applied in MERRA-2 can improve the regional modeling features for the domains that get affected by long-range transported dust. However, pollutants coming from boundaries had small influences on Delhi region's air quality, during the studied period (Table S2)

S.3 Sensitivity to changes in dust emissions

Sensitivity tests using different boundary conditions showed that long range transported dust coming from Middle East did not majorly influence air quality in Delhi. However, our domain covers some desert regions in eastern Pakistan and their dust emission impacts were evaluated. Figure S3 shows the response of hourly averaged $PM_{2.5}$ and PM_{10} concentrations to changes in the dust emissions. Turning on the dust option affected $PM_{2.5}$ and PM_{10} concentrations in eastern parts of Pakistan and some parts of the borders between Pakistan and India, but did not affect Delhi (Fig. S4). In another experiment, we increased the total dust emissions by 5 times, which increased PM_{10} concentrations significantly over western parts of the domain, close to source. This small-range transport is due to the mass of large dust particles and accompanying higher dry deposition rates. It also increased $PM_{2.5}$ concentrations and influenced some western parts of India with smaller size aerosols. However, they did not reach Delhi region, as the statistics over Delhi show no improvements (Table S2). In another experiment, changing the allocation of total dust in different bins as explained in methods section changed the aerosol regime in the west parts of the domain. Specifically, larger areas were effected by small size aerosols. Changing allocation of dusts, directly affected $PM_{2.5}$ concentrations in Delhi during the extreme pollution episode. Specifically, it increased $PM_{2.5}$ concentrations by ~20 μgm^{-3} on Nov. 8th. However, it was less than 5% contribution (Table S2). Moreover, increasing dust emissions had both positive and negative effects on concentrations (e.g. positive effect on Nov. 20th and negative effect on Nov. 28th), which are due to indirect effects of aerosols (Fig. S4). We did not perform more experiments as these tests suggest that in-domain dust sources were not a major source of extreme pollution episode in Delhi during November 2017. It should be mentioned that dust experiments, had lowest correlation coefficients, since the fire emissions were significantly high for all the days in all of them.

S.4. Statistics metrics used in analysis

$$ME = \frac{1}{n}\sum_{i=1}^{n}|P_i - O_i| \tag{1}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(P_i - O_i)^2}{n}} \tag{2}$$

$$NMB = \frac{\sum_{i=1}^{n}(P_i - O_i)}{\sum_{i=1}^{n} O_i} \tag{3}$$

$$NME = \frac{\sum_{i=1}^{n}|P_i - O_i|}{\sum_{i=1}^{n} O_i} \tag{4}$$

$$R = \frac{\sum_{i=1}^{n}\left((P_i - \bar{P}) * (O_i - \bar{O})\right)}{\sqrt{\sum_{i=1}^{n}(P_i - \bar{P})^2 * \sum_{i=1}^{n}(O_i - \bar{O})^2}} \tag{5}$$

$$CRMSE = \sqrt{\frac{\sum_{i=1}^{n}\left((P_i - \bar{P}) - (O_i - \bar{O})\right)^2}{n}} \tag{6}$$

Where P_i is the i-th prediction value, O_i is the i-th observed value, $\bar{P}$ is the mean predicted value, $\bar{O}$ is the mean observed value, and n is total number of paired sample. Coordinates of ground measurement stations are provided in Table S2

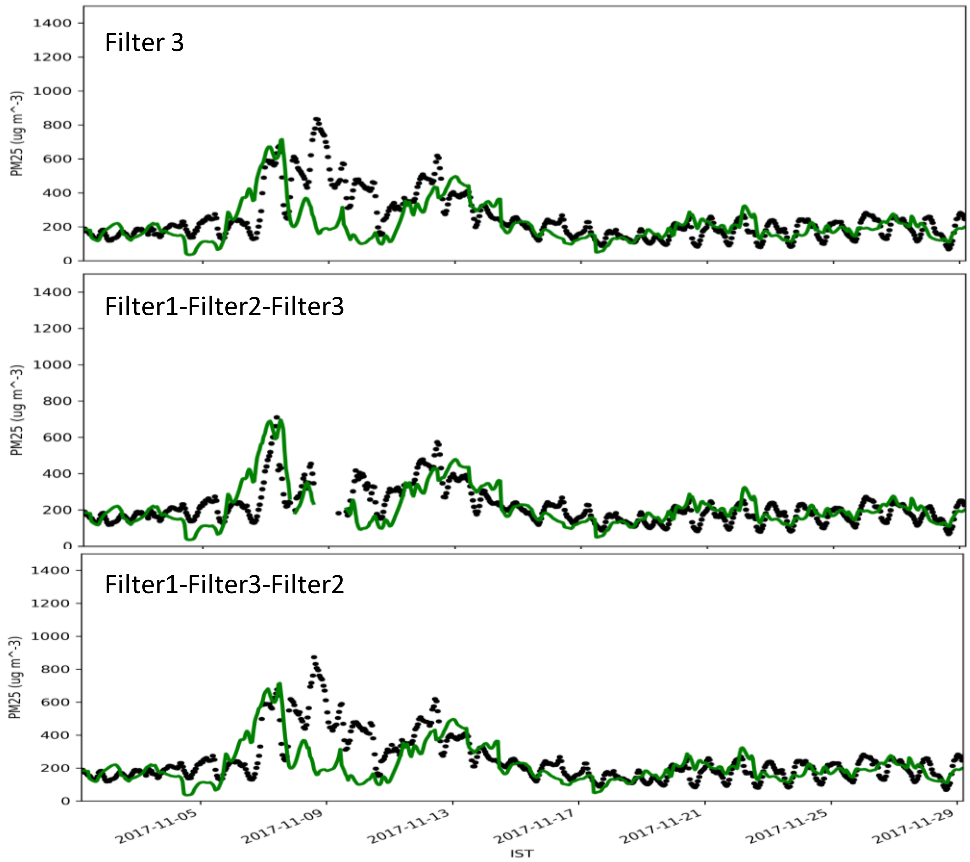

Figure S1 Effect of applying additional filters to CPCB data on averaged $PM_{2.5}$ timeseries in Delhi

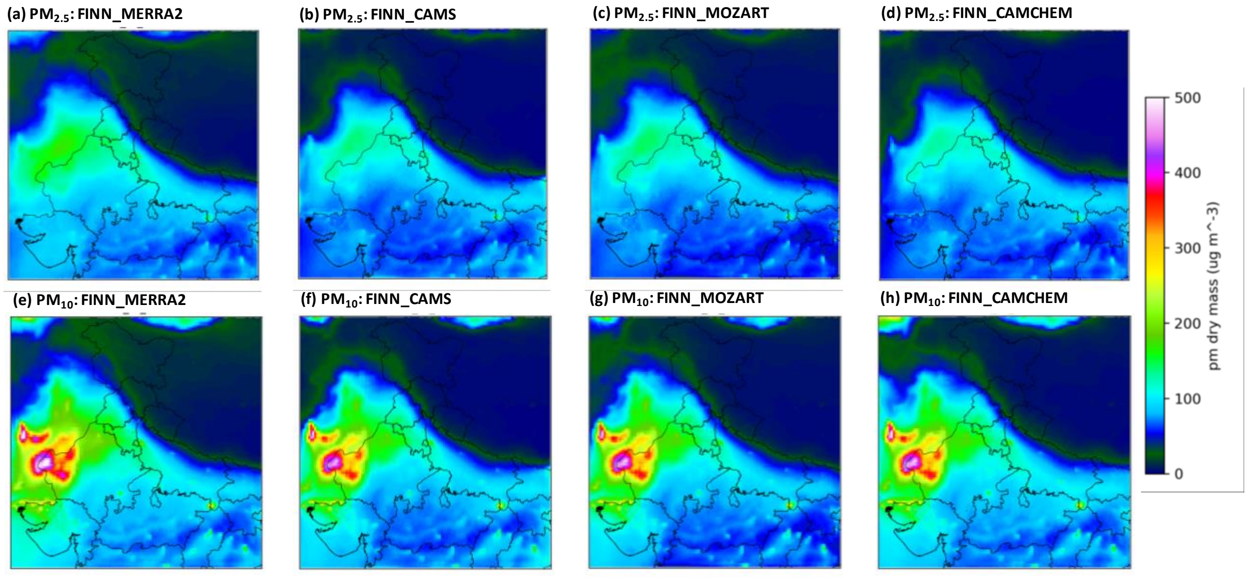

Figure S2 Responses of $PM_{2.5}$ and PM_{10} to changes in boundary conditions coming from: a, e) MERRA-2, b, f) CAMS, c, g) MOZART, and d, h) CAM-Chem.

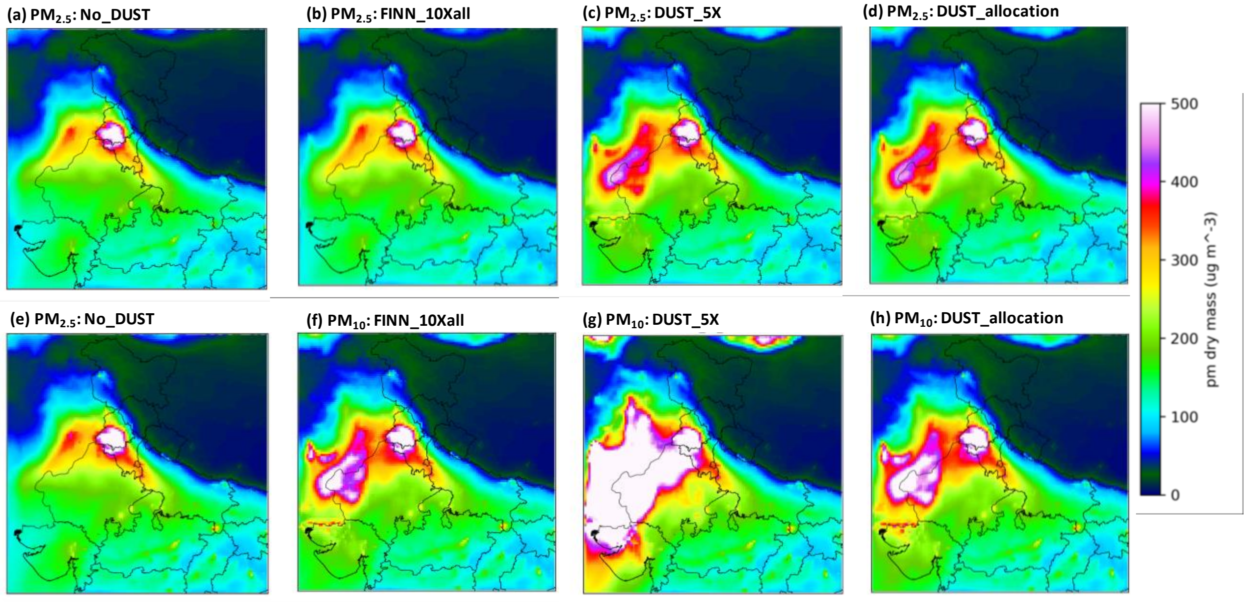

Figure S3 Responses of $PM_{2.5}$ and PM_{10} to changes in dust emissions: a, e) dust is turned off, b, f) dust is turned on, c, g) dust emissions are increased by 5 times, and d, h) dust emissions with different allocation in MOSAIC bins.

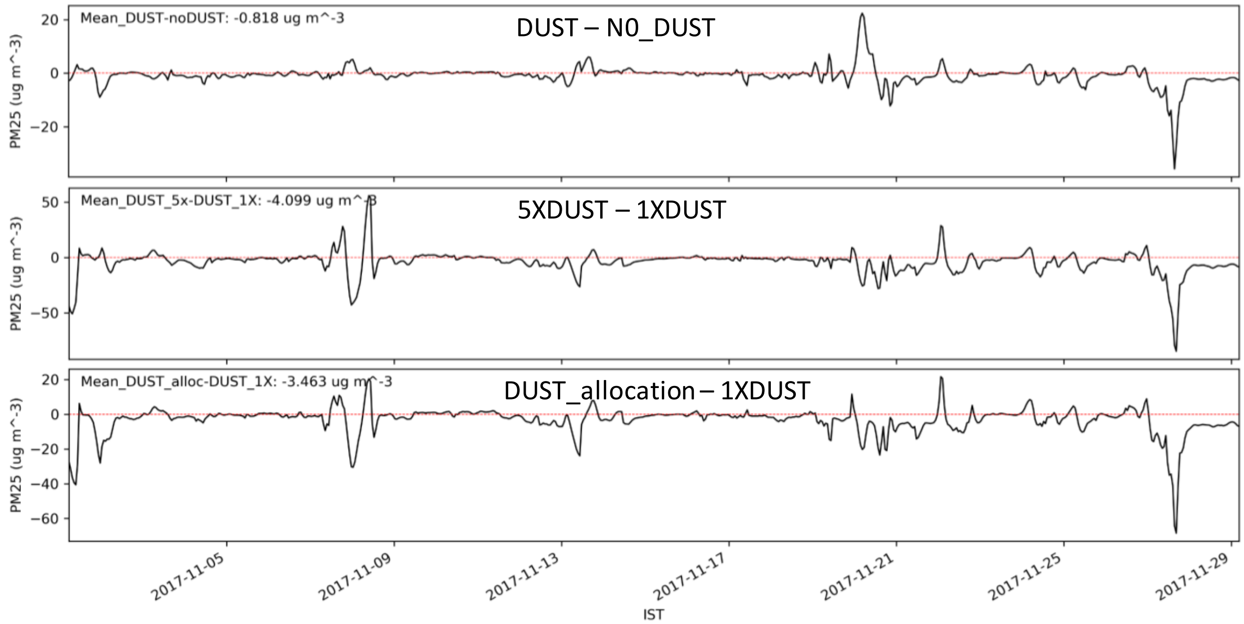

Figure S4 $PM_{2.5}$ concentrations' difference time series due to modifications in dust-scheme at the location of the US Embassy

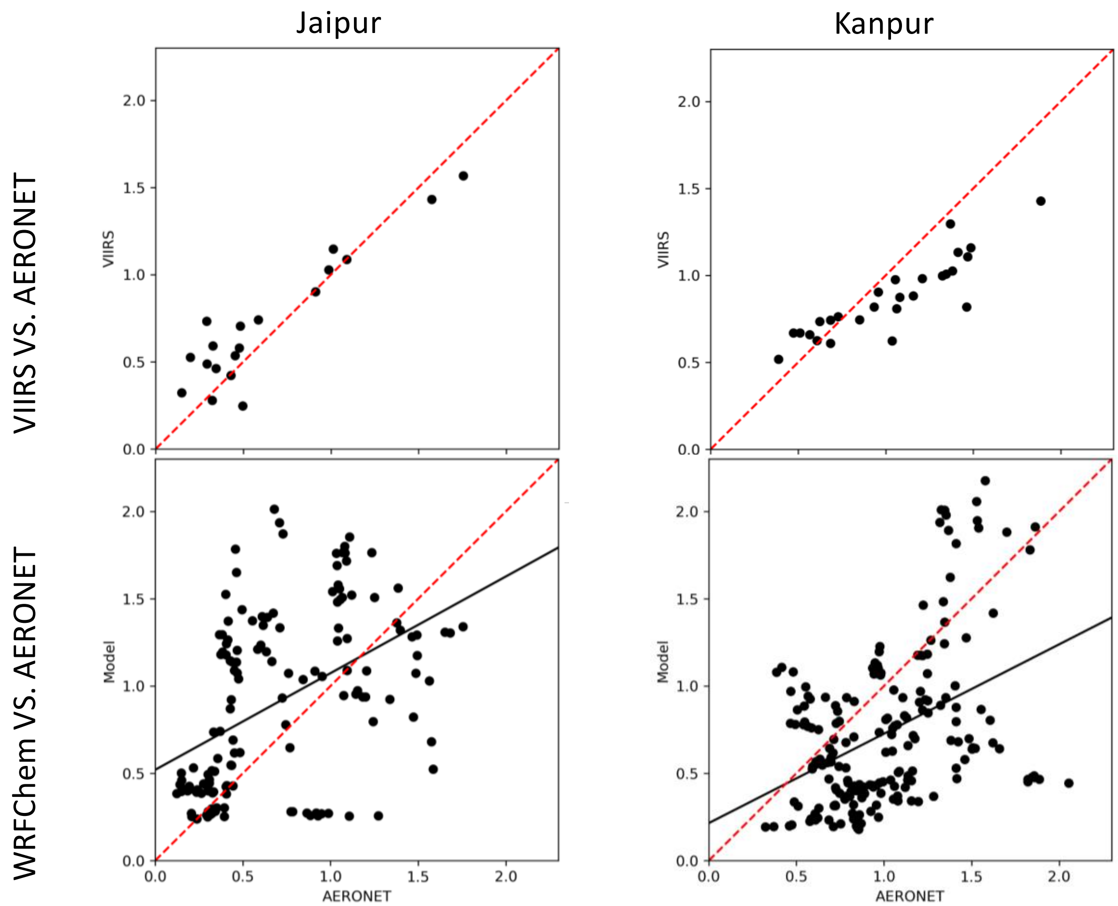

Figure S5 AOD550 scatter plot for VIIRS vs. AERONET (top row) and WRF-Chem vs AERONET (bottom row) for Jaipur (left column) and Kanpur (right column). Dashed lines shows 1-1 line

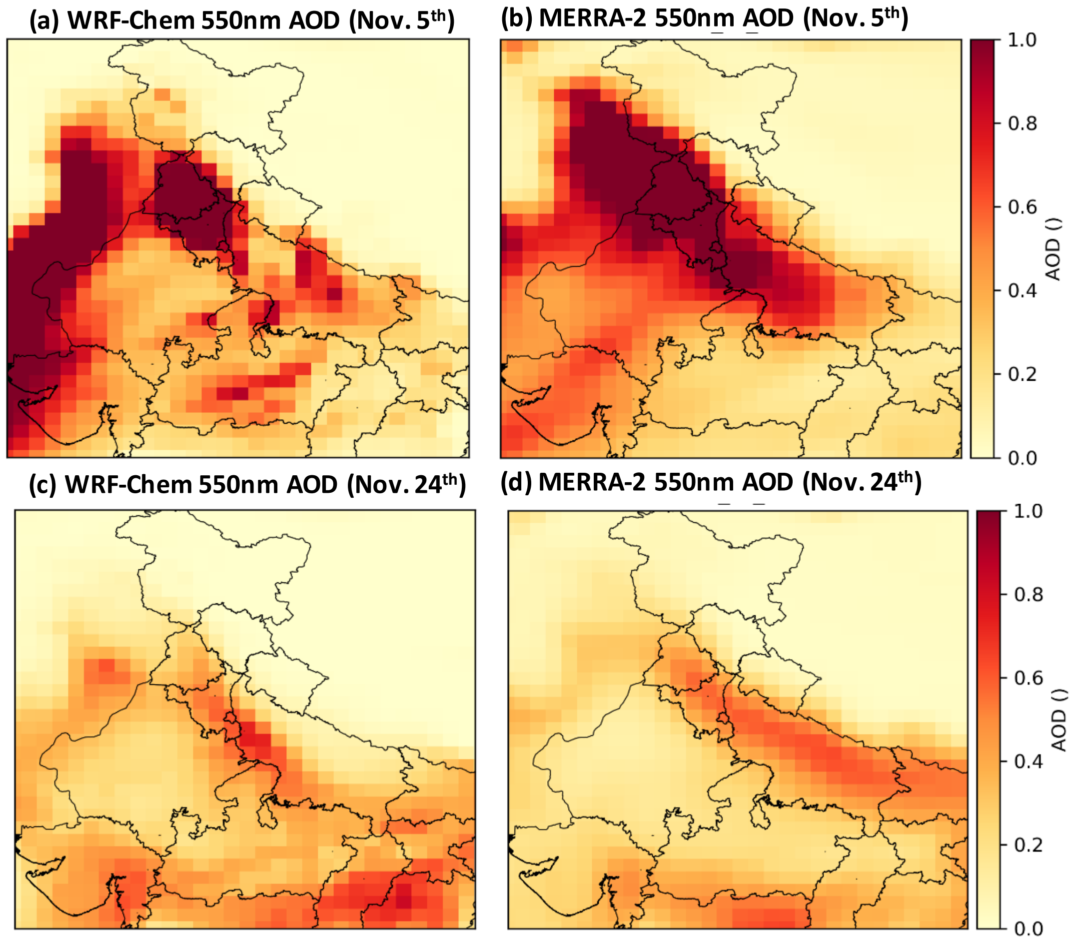

Figure S6 Modeled (left column) and MERRA-2 (right column) AOD maps at 550 nm for Nov 5th (panels: a, b) and Nov 20th (panels: c, d) daytime hours. Color bar is modified (compared to the plots in manuscript) to better represent low AOD values.

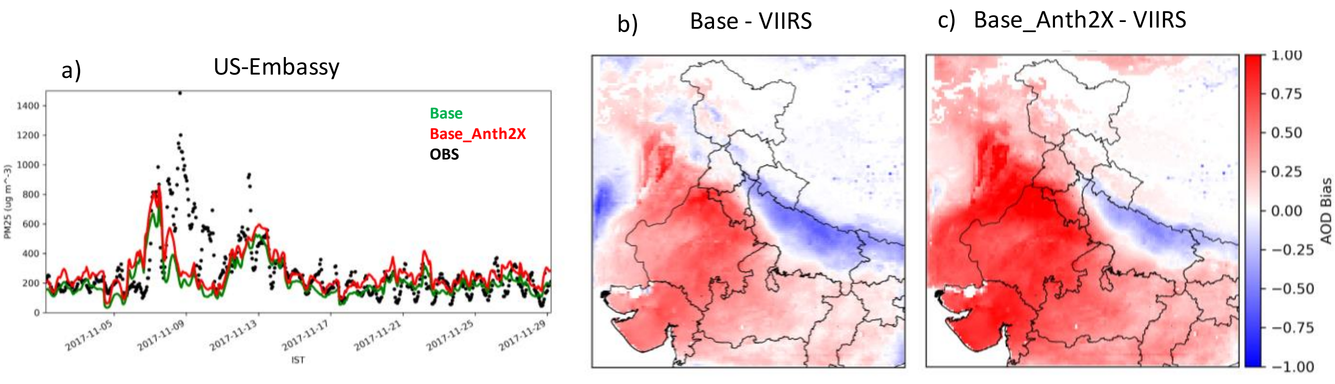

Figure S7 a) Timeseries for $PM_{2.5}$ concentration at the location of US embassy using Base scenario and BASE_ANTHRO2X scenario B) Bias of AOD at 550nm averaged over November 2017 base on b) base scenario c) base scenario with 2 times more anthropogenic particle emissions (ID: BASE_ANTHRO2X)

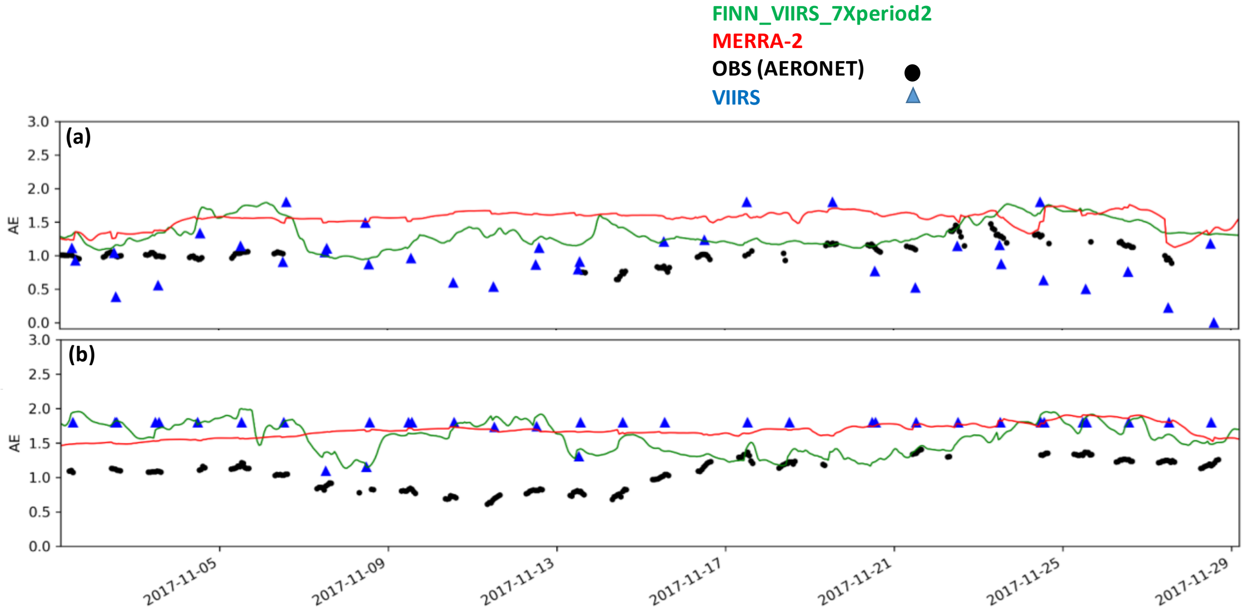

Figure S8 Time series of modeled (green line), VIIRS retrievals (blue triangle), MERRA-2 (red line), and AERONET (black dots) Angstrom Exponent during Nov. 2017 at a) Jaipur, b) Kanpur.

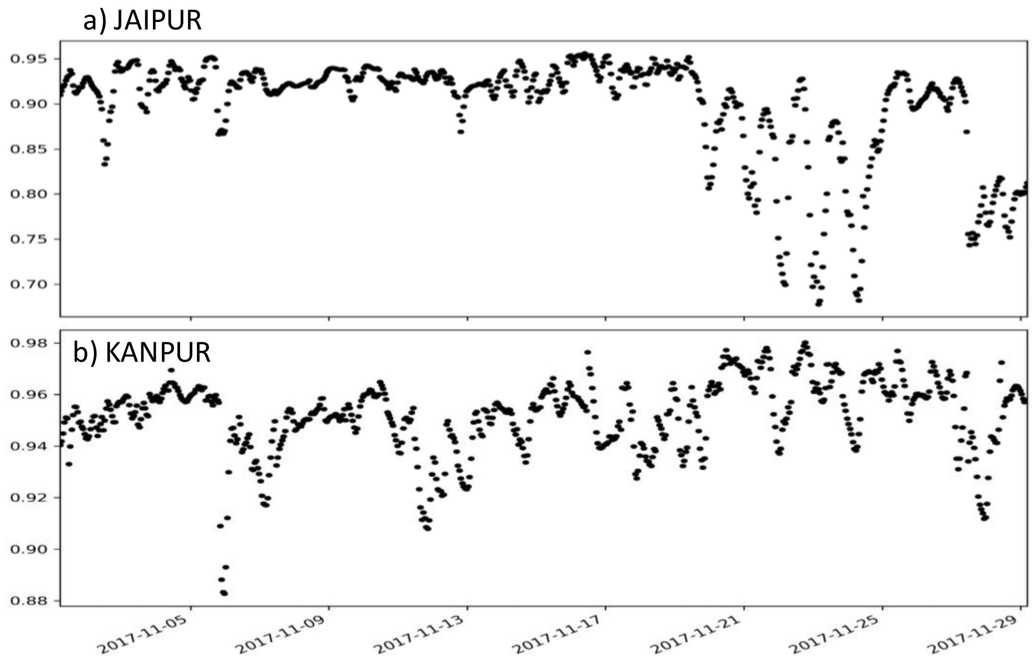

Figure S9 Modeled PM_{25}/PM_{10} ratio (Base scenario) at a) Jaipur and b) Kanpur

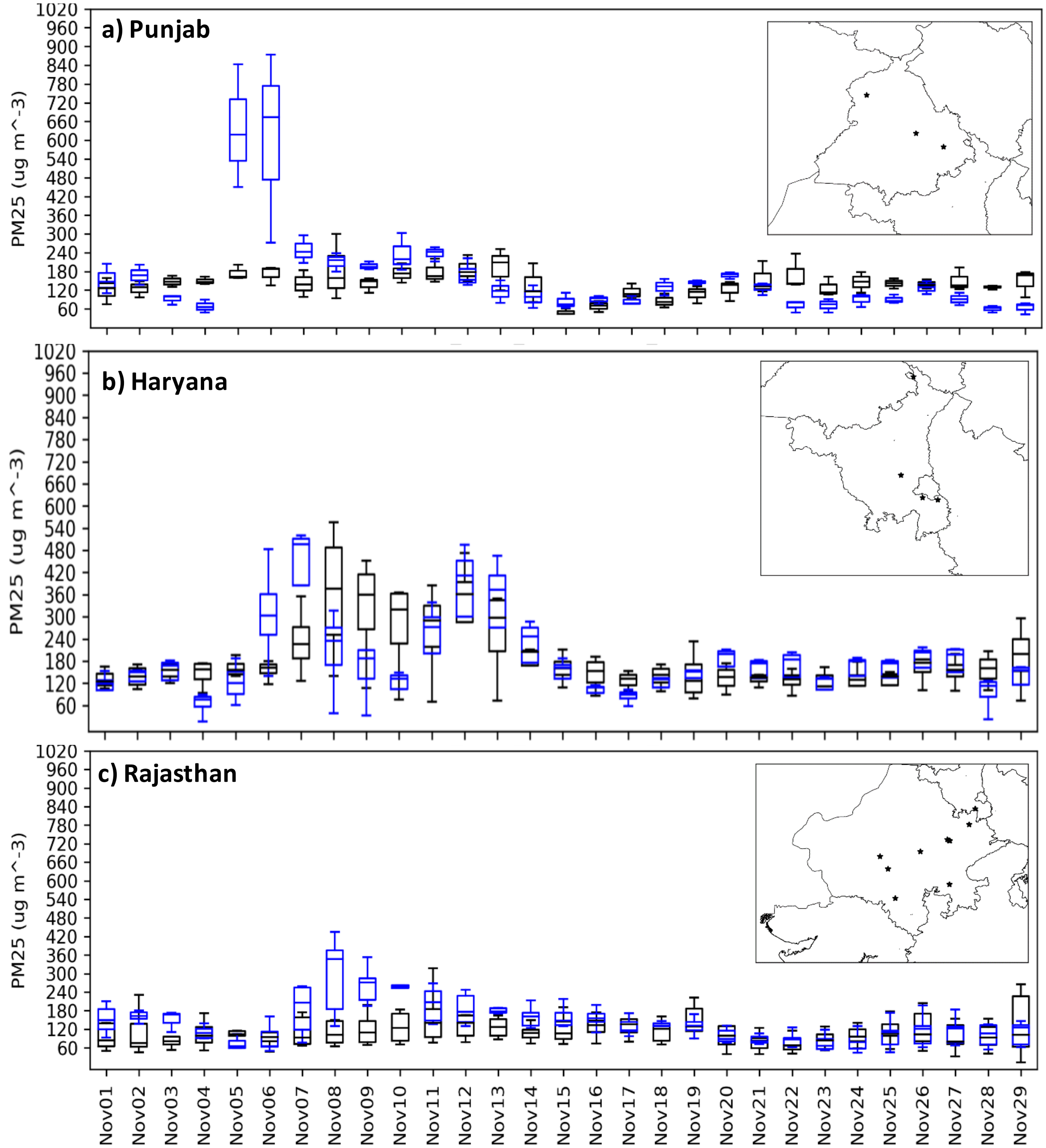

Figure S10 Box and Whisker plots of observed (black) and modeled (base scenario) daily $PM_{2.5}$ concentration averaged over all CPCB stations in: a) Punjab (3 stations), b) Haryana (4 stations), c) Rajasthan (10 stations). The inset maps show the location of stations in each province.

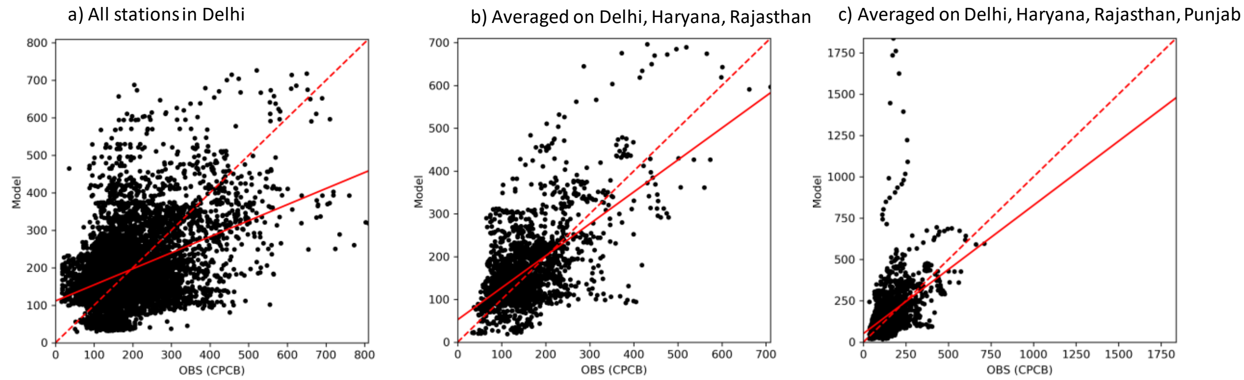

Figure S11 Scatter plots for a) all stations in Delhi combined b) averaged concentrations in Delhi, Haryana, and Rajasthan c) averaged concentrations in Delhi, Haryana, Rajasthan, and Punjab. Filters are applied to CPCB data.

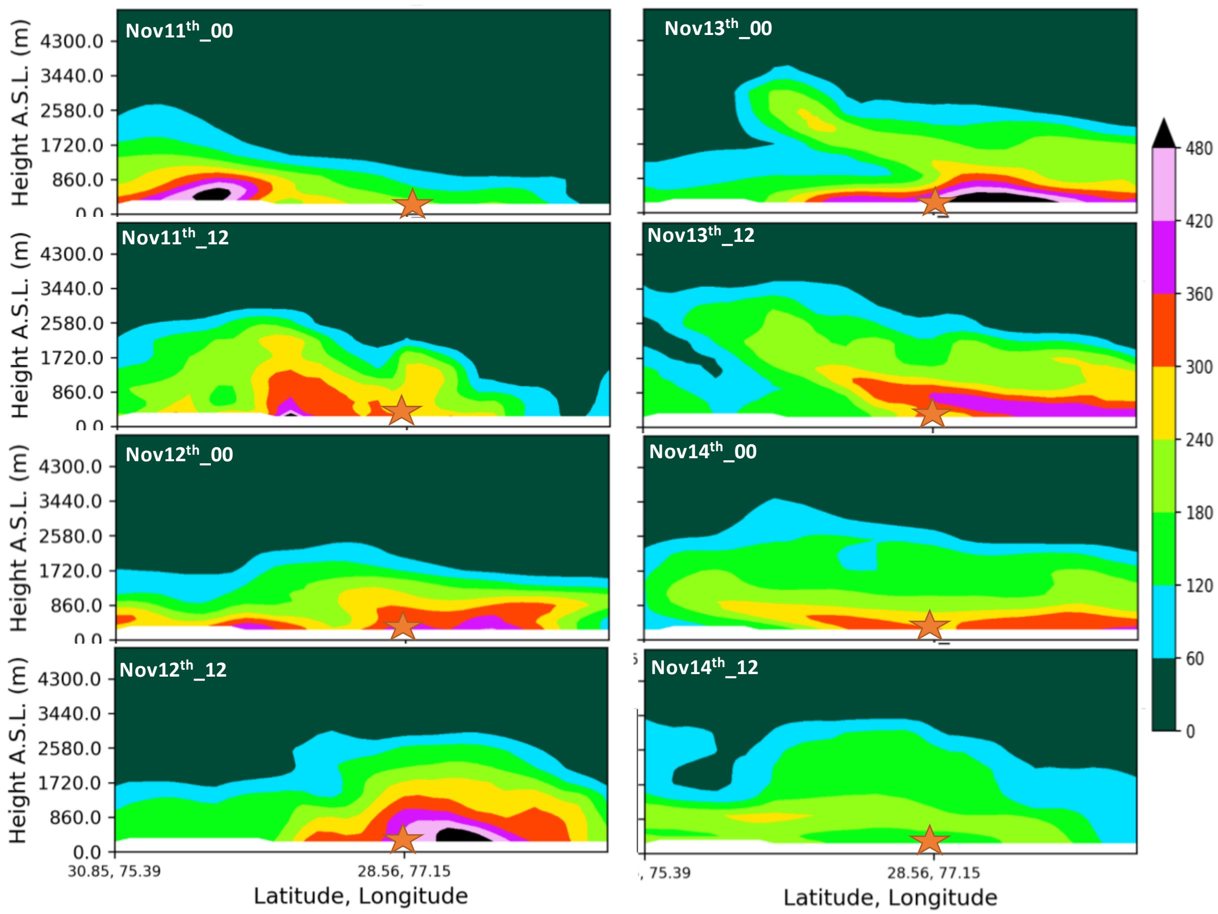

Figure S12 Vertical cross section of $PM_{2.5}$ concentration through the path shown in Fig. 1 for the days between Nov. 11th and Nov. 14th. For each day, two snapshots are shown at 00UTC (5:30AM local time) and 12UTC (5:30PM local time). The orange star shows the location of Delhi through the path.

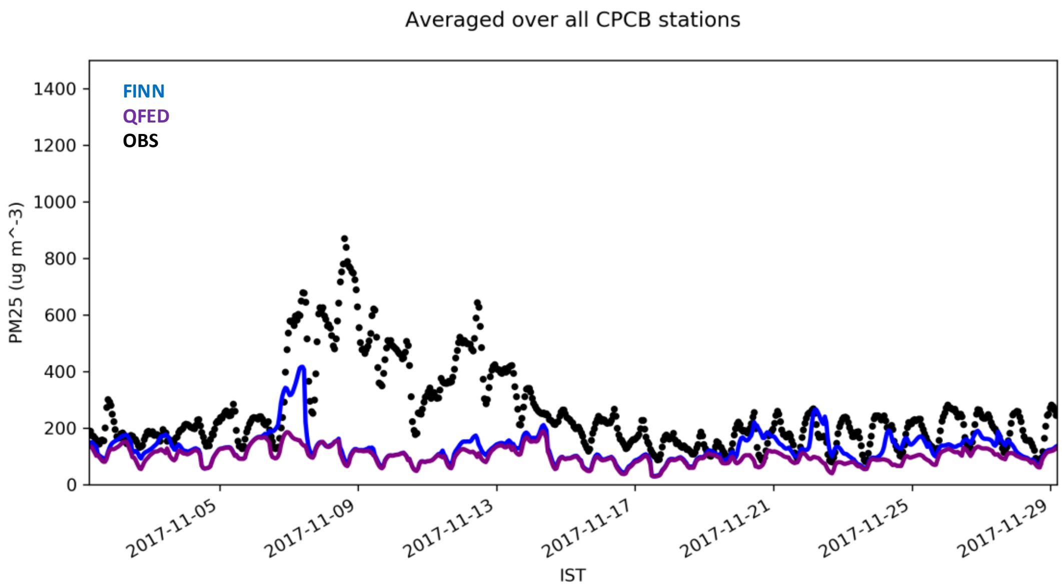

Figure S13 $PM_{2.5}$ time series using FINN (blue) and QFED (purple) biomass burning emission inventories averaged over all CPCB stations

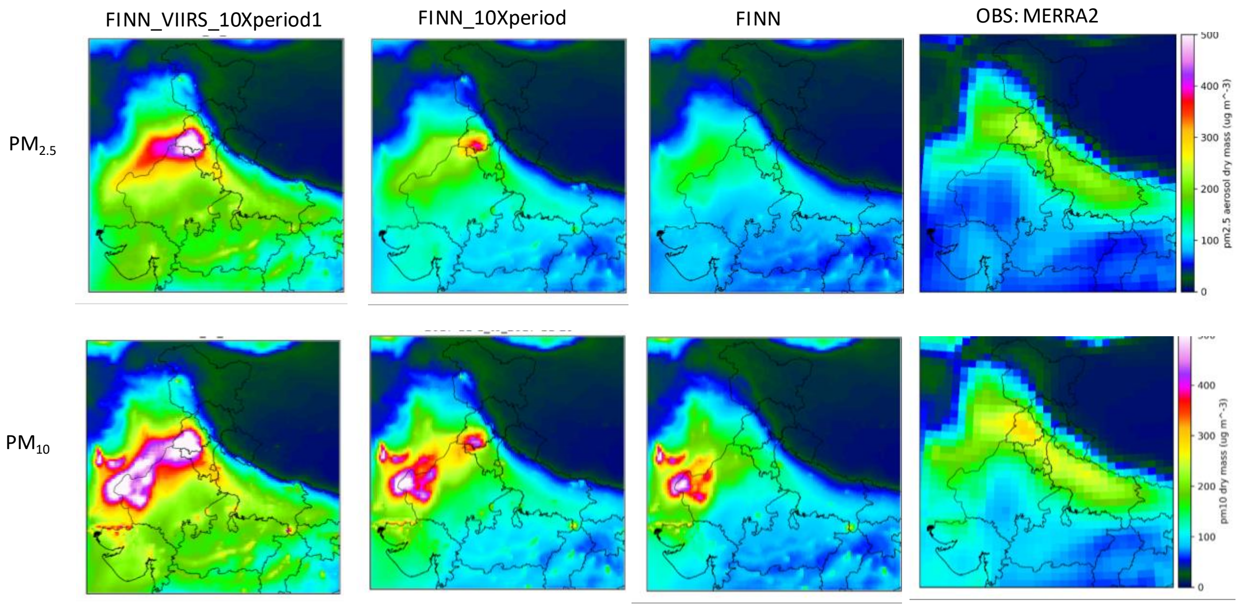

Figure S14 maps of $PM_{2.5}$ (top row) and PM_{10} (bottom row) concentration averaged in November 2017 (all hours) using different experiments on FINN biomass burning emission inventory

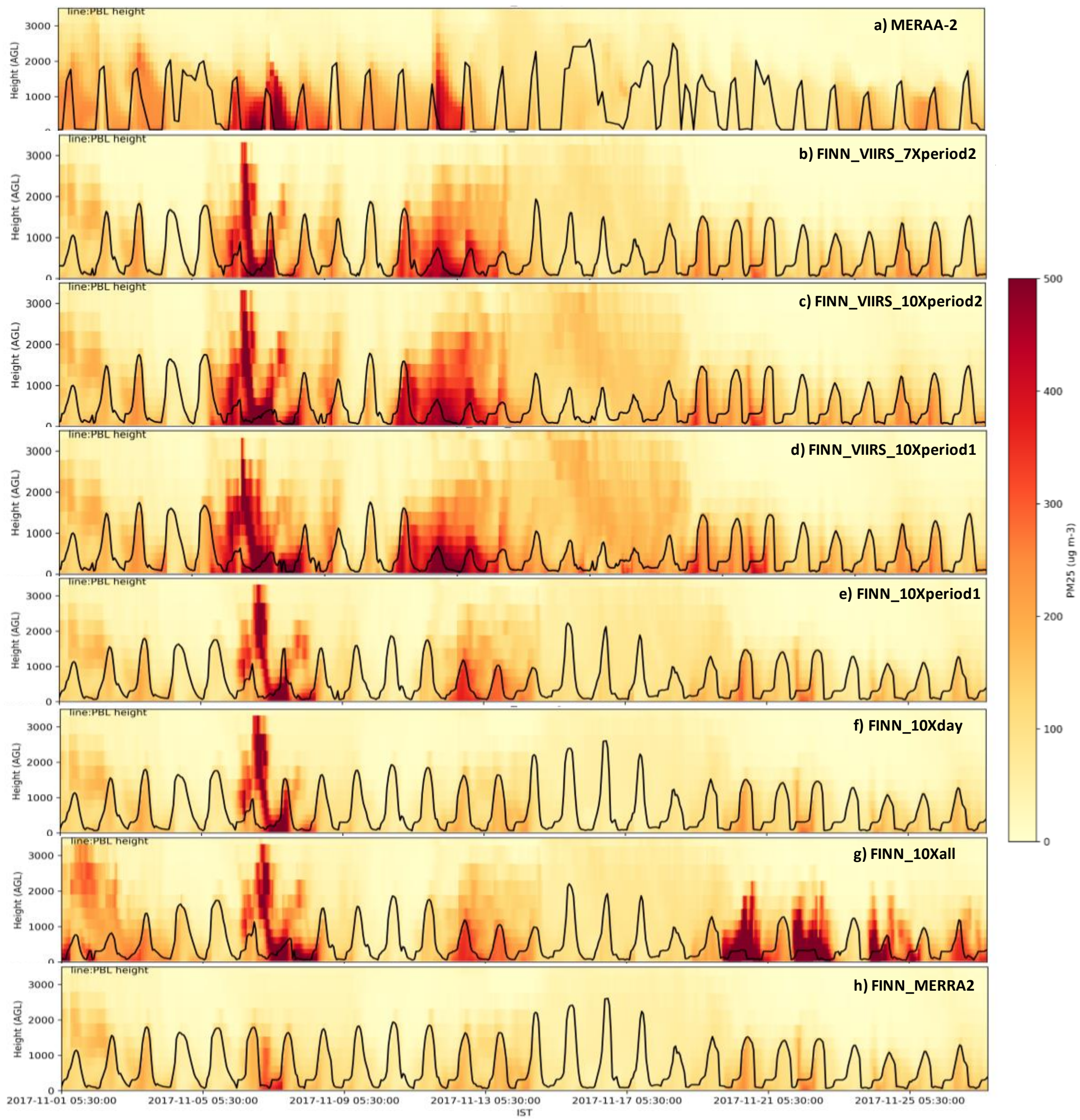

Figure S15 vertical cross section model $PM_{2.5}$ sensitivity to different experiments on FINN emission inventory at US Embassy coordinates

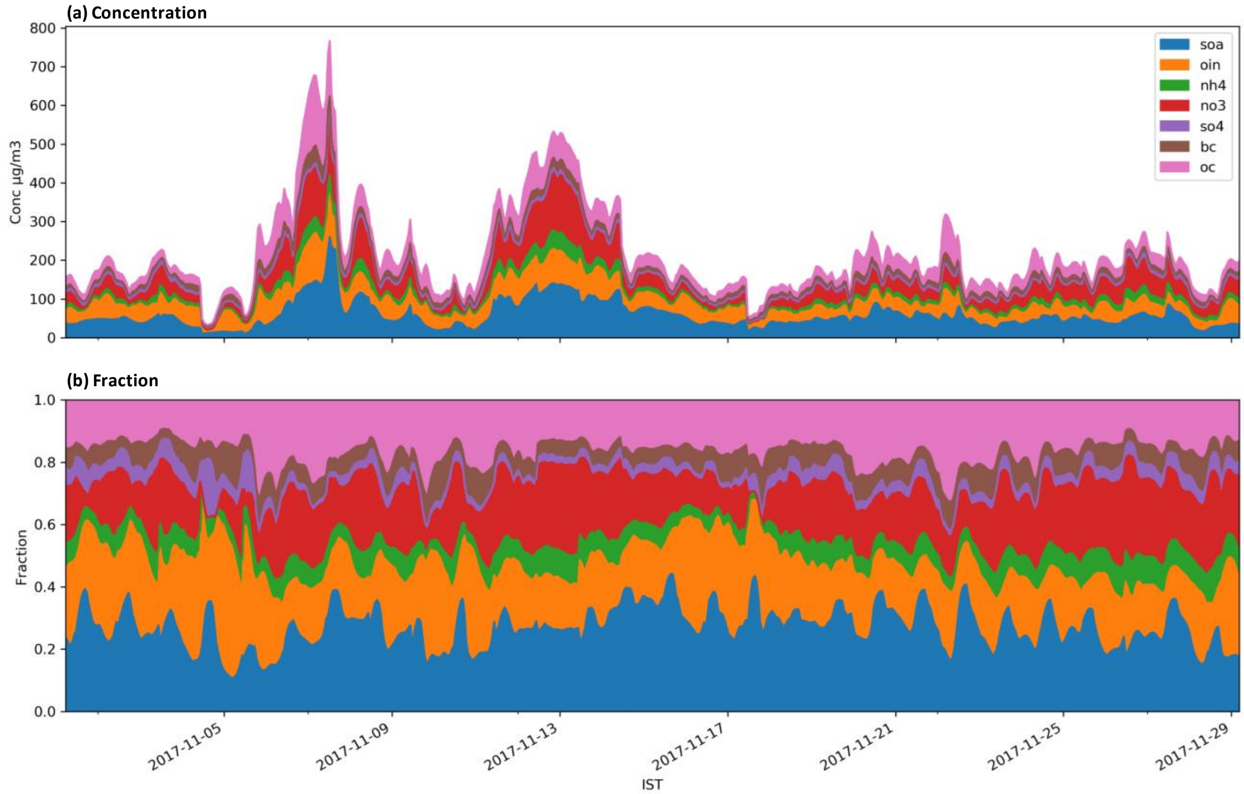

Figure S16 $PM_{2.5}$ composition at US Embassy coordinates in base scenario: a) Concentration values, b) Fractional values

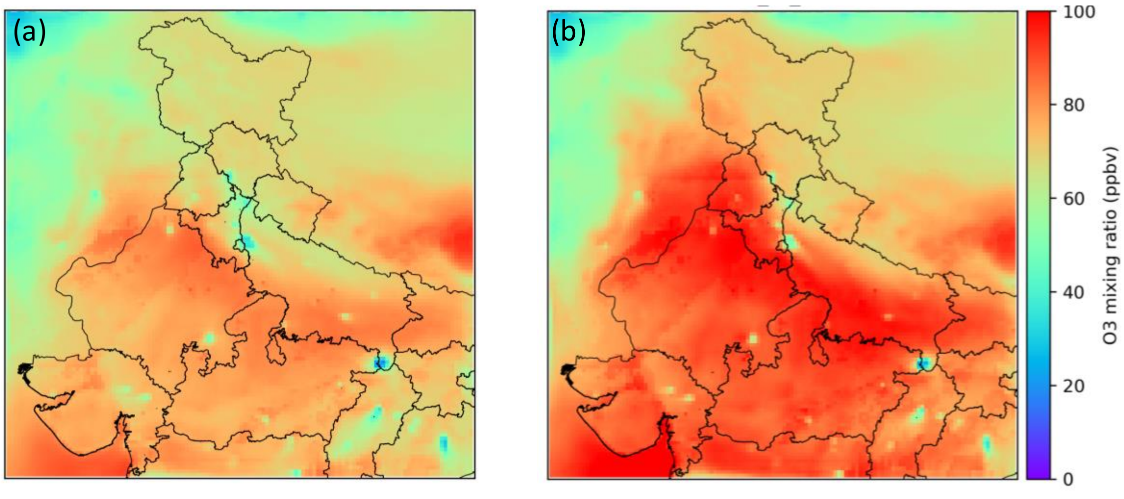

Figure S17 Daytime (8AM-6PM) ozone concentrations averaged during November 2017 for: a) a scenario without fire emission scaling (FINN_MERRA2) and b) base scenario with 7times higher fire emissions. Panel b is copied from the manuscript (Fig. 15) for easier comparison.

Table S1Effect of applying filters to CPCB data on $PM_{2.5}$ statistics in Delhi

Province	Hourly Obs. Mean (±std) (μgm^{-3})	Hourly Model Mean (±std) (μgm^{-3})	24-hours NMB (%)	24-hours NME (%)	24-hours R (%)
CPCB-Delhi	255.5 (±146.6)	213.9 (±113.9)	-16.6	27.6	0.48
Only filter 3	248.4 (±140.3)	214.5 (±114.5)	-13.9	26.4	0.49
Filter123	215.5 (±95.5)	214.8 (±115.2)	-1.9	23.6	0.64
Filter132	248.6 (±140.8)	214.6 (±114.5)	-13.9	26.4	0.49

Table S2 coordinates of CPCB ground measurement stations in Delhi used for statistical performance of experiments

Station	Latitude	Longitude
Alipur	28.8153	77.1530
Anand Vihar	28.6468	77.3160
Aya Nagar	28.4707	77.1099
Burari Crossing	28.7256	77.2011
CRRI Mathura Road	28.5512	77.2736
DTU	28.7500	77.1113
East Arjun Nagar	28.6556	77.2859
IGI Airport (T3)	28.5628	77.1180
IHBAS, Dilshad Garden	28.6812	77.3025
ITO	28.6317	77.2494
Lodhi Road	28.5918	77.2273
Mandir Marg	28.6364	77.2011
NSIT Dwarka	28.6091	77.0325
North Campus, DU	28.6574	77.1585
Punjabi Bagh	28.6740	77.1310
Pusa	28.6396	77.1463
R K Puram	28.5633	77.1869
Shadipur	28.6515	77.1473
Sirifort	28.5504	77.2159

Table S3 Statistics of all experiments for all days in November 2017 compared with data from CPCB stations: STD: Standard Deviation, R: Pearson Correlation Coefficient, RMSE: Root Mean Squared Error, NMB: Normalized Mean Bias, NME; Normalized Mean Error, MB: Mean Bias, ME: Mean Error

ITEM	Hourly Mean	Hourly STD	24-hours R	24-hours RMSE	24-hours NMB	24-hours NME	24-hours MB	24-hours ME
CPCB Obs. data	255.47	146.62	—	—	—	—	—	—
FINN_VIIRS_7Xperiod2	213.86	113.87	0.48	118.47	-16.6	27.63	-42.38	70.54
BASE_ANTHRO2X	280.24	125.60	0.50	114.20	9.73	32.79	24.85	83.70
FINN_VIIRS_10Xperiod2	254.15	149.44	0.51	112.52	-1.06	29.56	-2.72	75.46
FINN_VIIRS_10Xperiod1	276.2	136.42	0.54	104.11	7.27	29.51	18.56	75.34
FINN_10Xperiod1	174.33	95.92	0.44	136.21	-32.06	35.62	-81.86	90.94
FINN_10Xday	151.73	85.99	0.39	155.43	-40.61	42.13	-103.68	107.54
FINN_10Xall	295.48	279.81	0	279.68	15.57	66.47	37.21	169.69
NO_DUST	294.84	279.96	0	279.8	14.31	66.42	36.54	169.57
DUST_5X	298.85	280.76	-0.01	281.62	15.95	66.81	40.71	170.55
DUST_allocation	298.25	280.39	0	281.03	15.71	66.79	40.1	170.5
FINN_MERRA2	141.94	55.93	0.33	167.88	-44.32	45.82	-113.14	116.98
FINN_MOZART	130.39	53.23	0.3	176.91	-48.73	49.19	-124.4	125.58
FINN_CAMS	128.45	53.6	0.31	178.2	-49.59	49.96	-126.59	127.54
FINN_CAMCHEM	127.03	51.93	0.27	180.72	-50.24	50.49	-128.25	128.9
QFED_CAMCHEM	101.15	29.25	0.41	196.33	-60.11	60.11	-153.44	153.44

Table S4 Same as Table S3 except after excluding extreme days of Nov. 7th, 8th, 9th, 10th.

Scenario	Hourly Mean	Hourly Standard Deviation	24-hours R	24-hours RMSE	24-hours NMB	24-hours NME	24-hours MB	24-hours ME
CPCB Obs data	215.26	97.58						
FINN_VIIRS_7Xperiod2	209.91	104.94	0.7	55.11	-2.44	18.96	-5	38.94
BASE_ANTHRO2X	274.4	112.27	0.7	84.96	26.95	29.53	55.36	29.53
FINN_VIIRS_10Xperiod2	241.75	123.19	0.66	88.56	13.98	25.56	28.71	52.5
FINN_VIIRS_10Xperiod1	264.67	109.84	0.65	76.87	20.93	25.63	42.99	52.64
FINN_10Xperiod1	166.16	64.42	0.57	81.43	-24.3	28.73	-49.92	59.01
FINN_10Xday	143.82	58.12	0.42	101.02	-32.74	34.62	-67.25	71.12
FINN_10Xall	301.38	287.4	-0.03	261.96	31.22	68.94	64.12	141.62
NO_DUST	300.63	287.54	-0.02	262.09	30.89	68.89	63.45	141.51
DUST_5X	305.6	288.48	-0.03	263.76	33.05	69.26	67.88	142.25
DUST_allocation	304.7	288.02	-0.03	263.29	32.59	69.29	67.13	142.32
FINN_MERRA2	141.58	51.65	0.42	107.42	-35.26	37.13	-72.43	76.27
FINN_MOZART	130.44	49.2	0.4	116.37	-40.21	40.78	-82.58	83.76
FINN_CAMS	128.27	49.17	0.4	117.8	-41.18	41.64	-84.59	85.53
FINN_CAMCHEM	129.56	48.87	0.38	117.98	-40.85	41.21	-83.91	84.64
QFED_CAMCHEM	100.23	29.29	0.5	146.14	-55	55	-112.98	112.98

References

Beig, G., Srinivas, R., Parkhi, N. S., Carmichael, G., Singh, S., Sahu, S. K., Rathod, A., and Maji, S.: Anatomy of the winter 2017 air quality emergency in Delhi, Science of The Total Environment, 681, 305-311, 2019.

Jena, C., Ghude, S. D., Kulkarni, R., Debnath, S., Kumar, R., Soni, V. K., Acharja, P., Kulkarni, S. H., Khare, M., and Kaginalkar, A. J.: Evaluating the sensitivity of fine particulate matter (PM 2.5) simulations to chemical mechanism in Delhi, Atmospheric Chemistry and Physics Discussions, 1-28, 2020.

Kumar, R., Ghude, S. D., Biswas, M., Jena, C., Alessandrini, S., Debnath, S., Kulkarni, S., Sperati, S., Soni, V. K., and Nanjundiah, R. S.: 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, e2020JD033019, 2020.