

A meteorology based particulate matter prediction model for megacity Dhaka

Sadia Afrin¹², Mohammad Maksimul Islam^{12*}, Tanvir Ahmed¹

¹ *Department of Civil Engineering, Bangladesh University of Engineering and Technology, Dhaka - 1000, Bangladesh*

² *Department of Civil, Construction, and Environmental Engineering, North Carolina State University, 2501 Stinson Dr, Raleigh, NC 27695, USA*

Contents of this file

Figures S1, S2, and S3

Table S1, S2, S3, S4, and S5

Introduction

This file contains additional tables and figures that support the article's main text. The locations of the city's air monitoring stations are shown in Fig. S1. The summary statistics of the observed variables measured at these three monitoring sites are listed in Table S1, while the time-series variation of these observed variables at CAMS-1 for a typical dry winter month (February 2015) is plotted in Fig. S2.

Table S2 includes the comparative model parameters for all the four different regression models schemes considered in this analysis. For the selected PM_{2.5} and PM_{2.5-10} models, we

* Corresponding author. Tel: 1-919-943-6639;

E-mail address: mislam7@ncsu.edu

conducted detailed performance evaluation and investigate the influence of season and months, which are summarized in Table S3 and Table S4 respectively. Additionally, we investigated the applicability of our developed models at the two other monitoring sites of Dhaka – CAMS-2 and CAMS-3. Fig. S3 presents the performance of both modeled $PM_{2.5}$ and $PM_{2.5-10}$ concentration in capturing the temporal variability and their association with observed concentrations of CAMS-3 and CAMS-2. Finally, the influence of previous days PM concentration in model performance is summarized in Table S5.

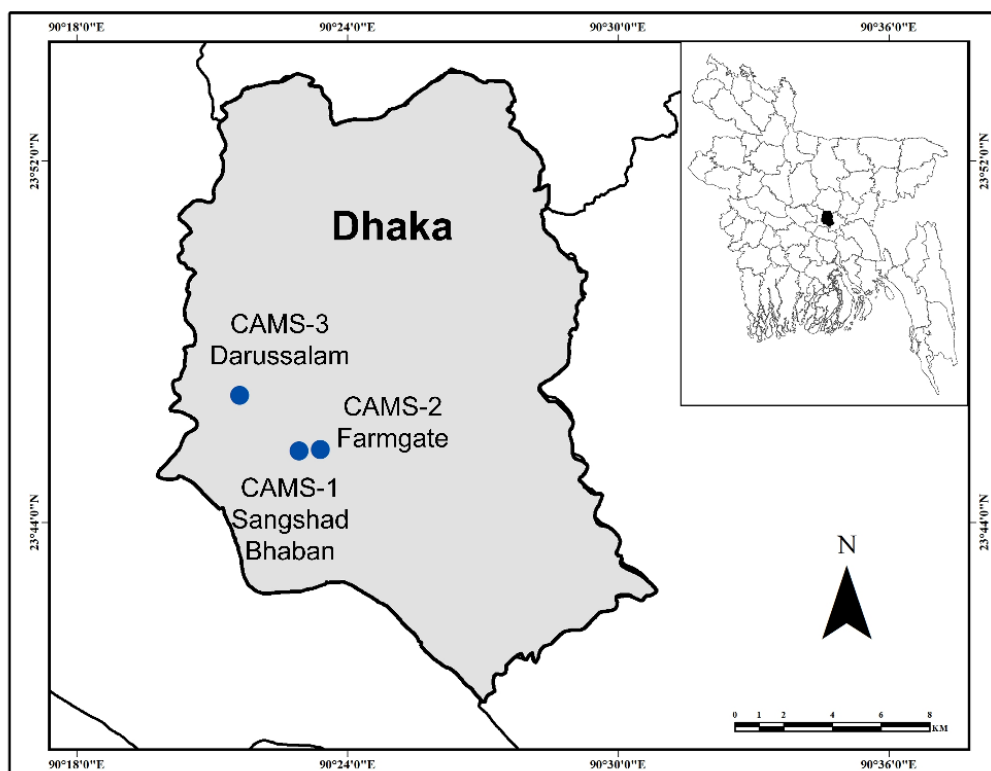


Figure S1. Geographic locations of the three Continuous Air Monitoring Stations (CAMSs): Sangshad Bhaban (CAMS-1), BARC (CAMS-2), and Darussalam (CAMS-3).

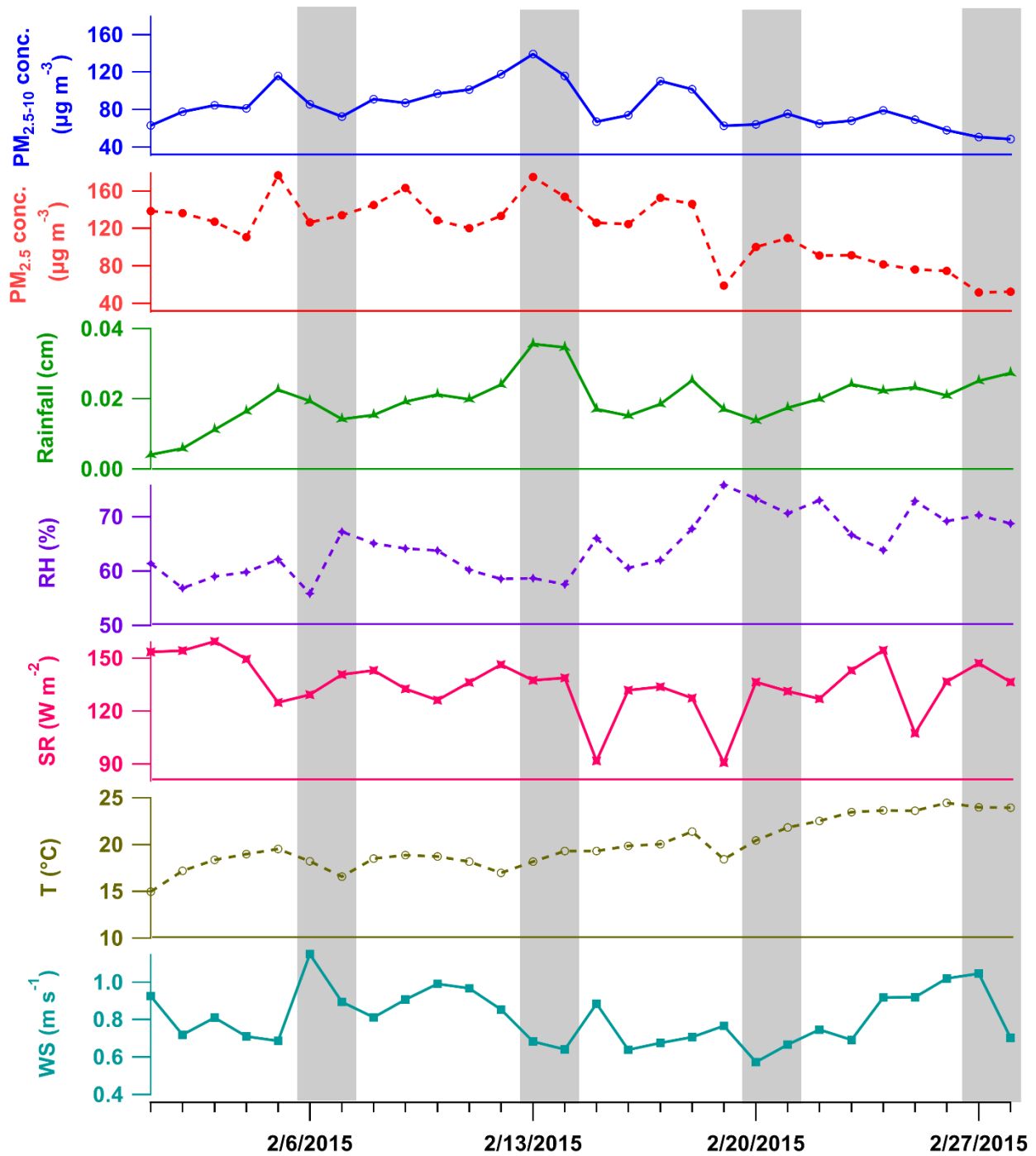


Figure S2. Daily variation of PM concentrations and meteorological variables considered in model development - $PM_{2.5-10}$, $PM_{2.5}$, rainfall, relative humidity, solar radiation, temperature, and wind speed- for a typical dry winter month in February 2015. The weekends during this month are shown as grey shades.

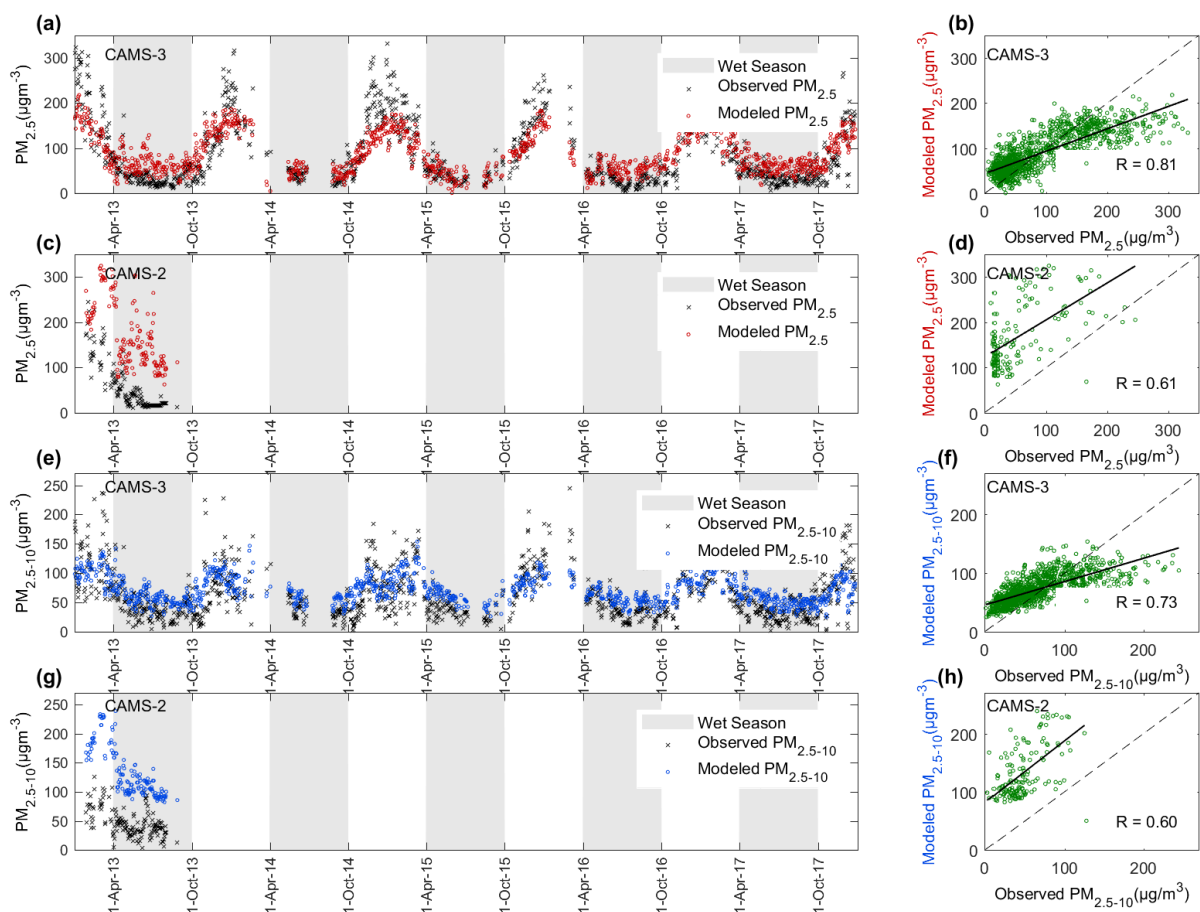


Figure S3. Comparison of observed and modeled 24-hr average PM concentrations at CAMS-1 and CAMS-3. The left panels show the time series plots for (a) $PM_{2.5}$ concentration at CAMS-3, (c) $PM_{2.5}$ concentration at CAMS-2, (e) $PM_{2.5-10}$ concentration at CAMS-3, and (g) $PM_{2.5-10}$ concentration at CAMS-2. The typical wet seasons (April to September) are shown in grey shading. The right panels show the correlation coefficient between observed and modelled – (b) $PM_{2.5}$ concentration at CAMS-3, (d) $PM_{2.5}$ concentration at CAMS-2, (f) $PM_{2.5-10}$ concentration at CAMS-3, and (h) $PM_{2.5-10}$ concentration at CAMS-2.

Table S1. Daily summary statistics of the model variables from 2013 to 2017.

| Variables | Unit | CAMS-1 | | | CAMS-3 | | | CAMS-2 | | |
|----------------------|-------------------|--------------------------|------|------|-------------|------|------|--------------------------|------|------|
| | | (2013-2017) [§] | | | (2013-2017) | | | (2013-2017) [¥] | | |
| | | Mean | Min | Max | Mean | Min | Max | Mean | Min | Max |
| PM _{2.5} | µg/m ³ | 75 | 0 | 305 | 90 | 5 | 332 | 56 | 10 | 246 |
| PM _{2.5-10} | µg/m ³ | 62 | 1 | 344 | 64 | 1 | 245 | 47 | 3 | 125 |
| T | °C | 24 | 10 | 33 | 26 | 10 | 39 | 20 | 3 | 30 |
| RH | % | 70 | 35 | 95 | 71 | 34 | 93 | 51 | 16 | 64 |
| SR | W/m ² | 138 | 8 | 1091 | 180 | 18 | 2356 | 411 | 227 | 517 |
| R | cm | 0.14 | 0.00 | 0.70 | 0.02 | 0.00 | 0.53 | 0.24 | 0.00 | 0.66 |
| WS | m/s | 1.32 | 0.41 | 3.51 | 1.97 | 0.49 | 5.17 | 1.36 | 0.17 | 2.57 |

§ Data was not available for the year 2016

¥ Representative data is available in 2013 only. For the rest of the period, data is either missing or having recording issue (e.g. PM_{2.5} is higher than PM₁₀)

Table S2. Summary of multilinear regression models from four modeling approaches[#] showing the adjusted R², AIC, significant predictors (highlighted) and their coefficients associated with each approach

| Predictors | PM _{2.5} | | | | PM _{2.5-10} | | | |
|---|-------------------|--------|--------|--------|----------------------|--------|--------|--------|
| | T | BE | FS | BD | T | BE | FS | BD |
| Intercept | 447.54 | 416.86 | 446.83 | 446.83 | 330.23 | 298.80 | 298.80 | 298.80 |
| R | 3.69 | | | | -0.87 | | | |
| T | -7.92 | -8.39 | -8.97 | -8.97 | -2.20 | -2.51 | -2.51 | -2.51 |
| RH | -1.56 | -1.20 | -1.41 | -1.41 | -2.60 | -2.27 | -2.27 | -2.27 |
| SR | -0.10 | | | | -0.10 | | | |
| WS | -11.52 | -11.41 | -11.13 | -11.13 | 1.12 | | | |
| SinWD | -11.41 | -10.08 | | | -3.68 | | | |
| CosWD | 11.46 | 11.98 | 11.19 | 11.19 | -2.69 | | | |
| Adjusted R² (%) | | | | | | | | |
| Model evaluation | 57.1 | 57.1 | 56.7 | 56.7 | 35.1 | 35.1 | 35.1 | 35.1 |
| Akaike informaiton criterion (AIC) | | | | | | | | |
| | 2468.8 | 2467.4 | 2469.3 | 2469.3 | 2242.6 | 2237.9 | 2237.9 | 2237.9 |

Four approaches include backward elimination (BE), forward selection (FS), bidirectional (BD) stepwise, and the traditional (T) multilinear regression approach (e.g. model with all covariates included)

Table S3. Detailed performance statistics of the PM_{2.5} model at CAMS-1

| Category | Time Period | # of Days with observation | Normalized Mean Bias (%) | Normalized Mean Error (%) | R-value | P-Value |
|------------------|--------------------|-----------------------------------|---------------------------------|----------------------------------|----------------|----------------|
| Overall | 2013-2017 | 703 | 48.1 | 58.7 | 0.81 | 0.00* |
| Yearly | 2013 | 184 | 24.1 | 48.8 | 0.83 | 0.00* |
| | 2014 | 219 | 42.9 | 51.1 | 0.81 | 0.00* |
| | 2015 | 236 | 65.2 | 68.1 | 0.88 | 0.00* |
| | 2016 | 9 | 11.4 | 19.6 | -0.16 | 0.69 |
| | 2017 | 55 | 152.3 | 152.3 | 0.13 | 0.34 |
| Seasonal | Dry ^a | 350 | 23.7 | 37.0 | 0.63 | 0.00* |
| | Wet ^a | 353 | 135.4 | 136.5 | 0.38 | 0.00* |
| Quarterly | JFM ^a | 222 | 19.2 | 34.6 | 0.56 | 0.00* |
| | AMJ ^a | 244 | 104.9 | 106.2 | 0.40 | 0.00* |
| | JAS ^a | 109 | 269.3 | 269.3 | 0.10 | 0.31 |
| | OND ^a | 128 | 32.4 | 41.8 | 0.75 | 0.00* |
| Monthly | Jan ^a | 75 | 0.6 | 25.3 | 0.04 | 0.74 |
| | Feb ^a | 79 | 34.8 | 42.5 | 0.44 | 0.00* |
| | Mar ^a | 68 | 36.6 | 43.1 | 0.43 | 0.00* |
| | Apr ^a | 100 | 74.4 | 76.9 | 0.17 | 0.09 |
| | May ^a | 81 | 113.8 | 113.8 | 0.27 | 0.02* |
| | Jun ^a | 63 | 199.1 | 199.1 | -0.04 | 0.73 |
| | Jul ^a | 79 | 278.0 | 278.0 | -0.03 | 0.79 |
| | Aug ^a | 28 | 246.8 | 246.8 | 0.44 | 0.02* |
| | Sep ^a | 2 | | | | |
| | Oct ^a | 35 | 125.4 | 125.4 | 0.32 | 0.06 |
| | Nov ^a | 41 | 30.6 | 42.0 | 0.52 | 0.00* |
| | Dec ^a | 52 | 13.4 | 23.8 | 0.36 | 0.01* |

^a All observed data points between 2013 to 2017 are considered in the analysis.

Table S4. Detailed performance statistics of the PM_{2.5-10} model at CAMS-1

| Category | Time Period | # of Days with observation | Normalized Mean Bias (%) | Normalized Mean Error (%) | R-value | P-Value |
|------------------|--------------------|-----------------------------------|---------------------------------|----------------------------------|----------------|----------------|
| Overall | 2013-2017 | 703 | 42.0 | 50.3 | 0.76 | 0.00* |
| Yearly | 2013 | 184 | 35.4 | 45.9 | 0.73 | 0.00* |
| | 2014 | 219 | 30.1 | 41.4 | 0.78 | 0.00* |
| | 2015 | 236 | 54.8 | 58.3 | 0.79 | 0.00* |
| | 2016 | 9 | 29.7 | 46.0 | -0.20 | 0.60 |
| | 2017 | 55 | 90.6 | 90.6 | 0.81 | 0.00* |
| Seasonal | Dry | 350 | 27.0 | 38.7 | 0.56 | 0.00* |
| | Wet | 353 | 70.6 | 72.3 | 0.79 | 0.00* |
| Quarterly | | | 22.5 | 33.7 | 0.57 | 0.00* |
| | JFM | 222 | | | | |
| | AMJ | 244 | 58.7 | 60.7 | 0.77 | 0.00* |
| | JAS | 109 | 120.0 | 120.5 | 0.37 | 0.00* |
| | OND | 128 | 37.4 | 50.0 | 0.31 | 0.00* |
| Monthly | Jan | 75 | 18.7 | 36.5 | 0.03 | 0.81 |
| | Feb | 79 | 19.4 | 29.3 | 0.72 | 0.00* |
| | Mar | 68 | 30.2 | 35.9 | 0.71 | 0.00* |
| | Apr | 100 | 45.6 | 49.2 | 0.76 | 0.00* |
| | May | 81 | 75.3 | 75.3 | 0.48 | 0.00* |
| | Jun | 63 | 70.6 | 70.8 | 0.52 | 0.00* |
| | Jul | 79 | 116.4 | 117.0 | 0.33 | 0.00* |
| | Aug | 28 | 127.6 | 127.6 | 0.54 | 0.00* |
| | Sep | 2 | | | | |
| | Oct | 35 | 143.5 | 143.5 | 0.61 | 0.00* |
| | Nov | 41 | 26.7 | 37.4 | 0.25 | 0.11 |
| | Dec | 52 | 16.4 | 33.9 | -0.23 | 0.10 |

^a All observed data points between 2013 to 2017 are considered in the analysis.

Table S5. Significant parameters of the PM_{2.5} backward elimination regression models varying the averaging window of lag PM_{2.5} concentrations

| Predictors* | PM _{2.5} | |
|-----------------------------|--------------------------------|--------------------------------|
| | average of previous 2 days' PM | average of previous 3 days' PM |
| Intercept | 135.228 | 138.57 |
| R | - | - |
| T | -4.261 | -4.393 |
| RH | - | - |
| SR | - | - |
| WS | - | - |
| SinWD | - | - |
| CosWD | - | - |
| PM lag | 0.706 | 0.704 |
| Adjusted R ² (%) | 0.73 | 0.71 |
| AIC | 845.424 | 856.179 |

*Predictors with no coefficients (represented by a horizontal dash: -) indicate that those did not appear significant ($p < 0.05$) in the respective models. In addition, no interaction terms were listed in the table as they were not significant.