Estimating hourly ground-level aerosols using GEMS aerosol optical depth: A machine learning approach

Sungmin O[1], Ji Won Yoon[2,3], and Seon Ki Park[1,2,3,*]

[1]Department of Climate & Energy System Engineering, Ewha Womans University, Seoul, Republic of Korea
[2]Center for Climate/Environment Change Prediction Research, Ewha Womans University, Seoul, Republic of Korea
[3]Severe Storm Research Center, Ewha Womans University, Seoul, Republic of Korea
[*]corresponding author: Seon Ki Park (spark@ewha.ac.kr)

Contents

Supplementary figures
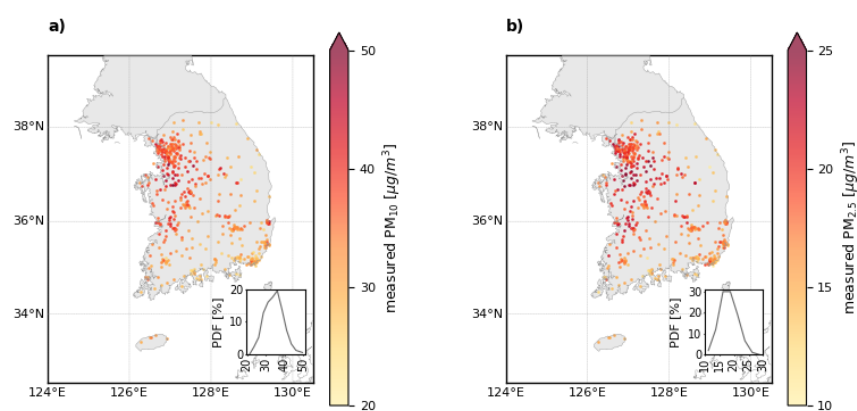
Figure S1. Average concentrations of (a) PM10 and (b) PM2.5 at urban air quality monitoring stations for the period from January 2022 to June 2023 in South Korea. All available PM measurements are used regardless of the GEMS AOD data availability.
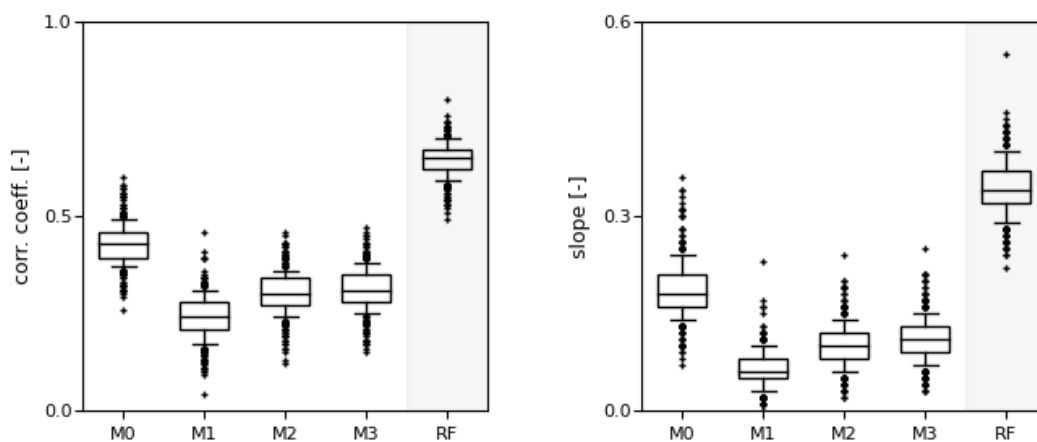
Figure S2. Comparison of model performance. For M0, we use the corresponding monthly average as estimated PM10 for a given month. M1, M2, and M3 show the performance of linear regression models with different sets of input variables; M1 with AOD data only, M2 with AOD, BLH, RH, and TEMP, and M3 with all meteorological variables. RF shows performance of random forest models that are used in the main analysis. All model performances are evaluated through the five-fold cross-validation.
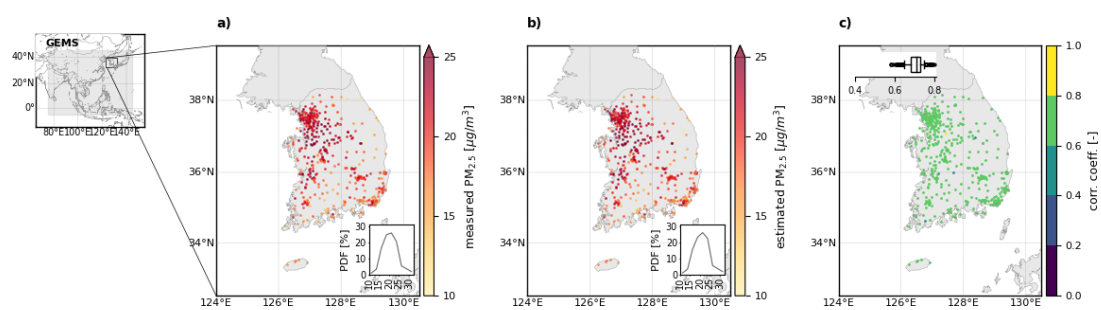
Figure S3. Same as in Fig. 1 but for PM2.5.
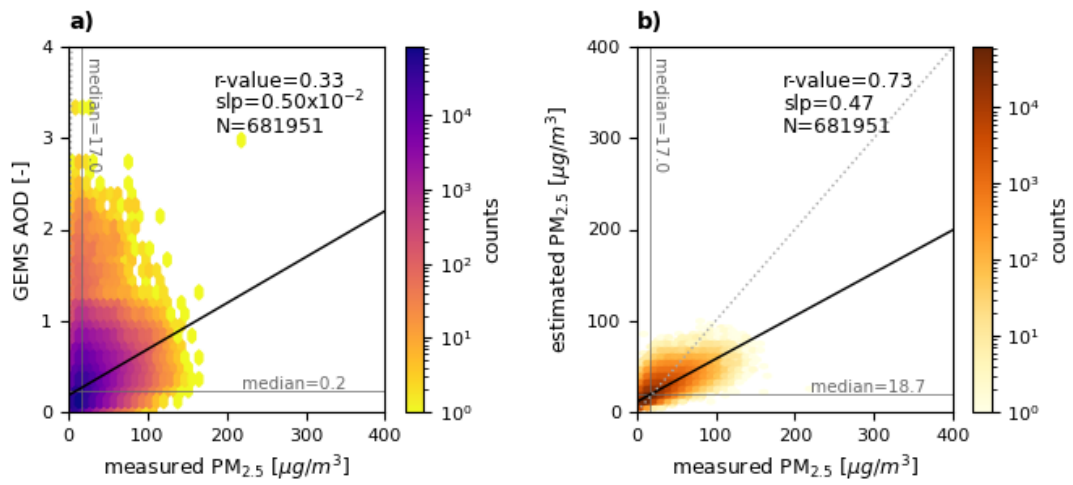
Figure S4. Same as in Fig. 2 but for PM2.5.
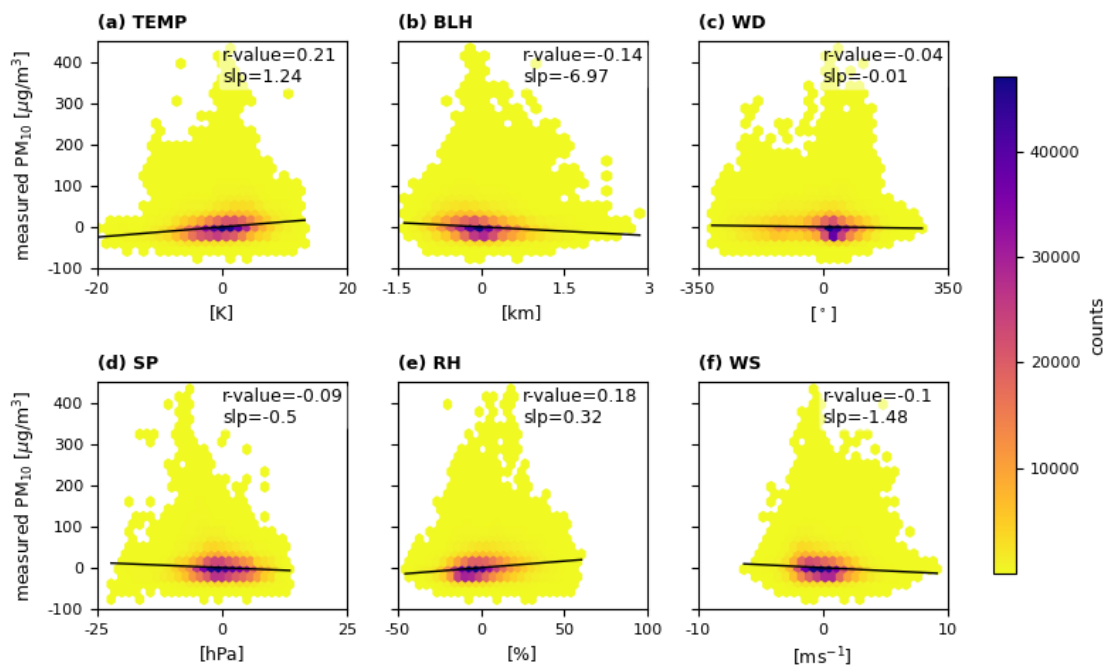
Figure S5. Comparison between anomalies of PM10 measurements and meteorological variables at all stations.
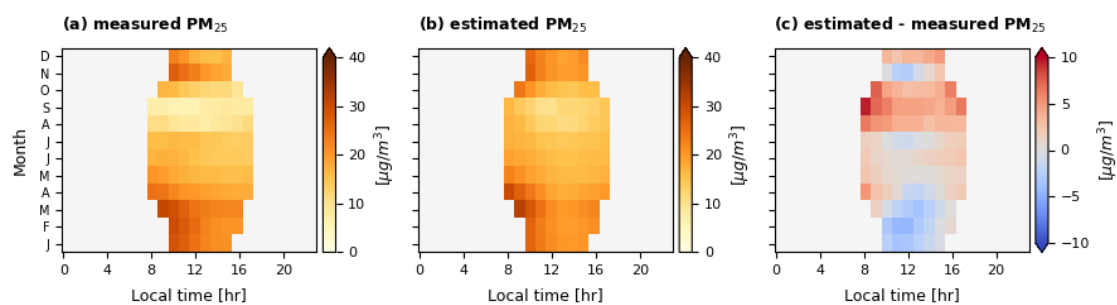
Figure S6. Same as in Fig. 4 but for PM2.5.
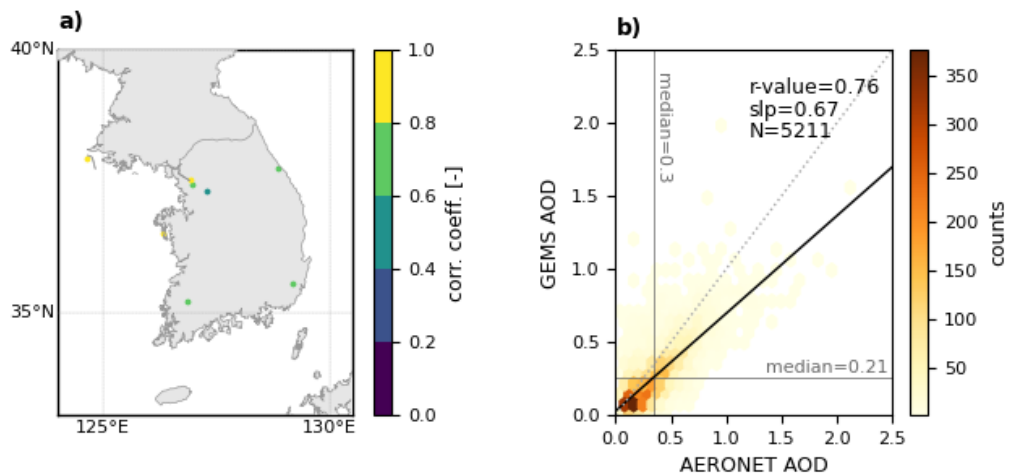
Figure S7. Comparison of the AERONET and GEMS AOD data used in Fig. 5 in the main analysis. (a) Pearson correlation coefficient values between the data at each AERONET site. (b) The density scatter plot between the AERONET and GEMS AOD data from all sites. Vertical and horizontal lines represent the corresponding mean values. The thick solid line is the regression line and the dotted diagonal line is the one-to-one line.