Supplemental Materials: Fusing Observational, Satellite Remote Sensing and Air Quality Model Simulated Data to Estimate Spatiotemporal Variations of PM2.5 Exposure in China

Tao Xue [1], Yixuan Zheng [1], Guannan Geng [1,2], Bo Zheng [2], Xujia Jiang [1], Qiang Zhang [1,3,*] and Kebin He [2,3]
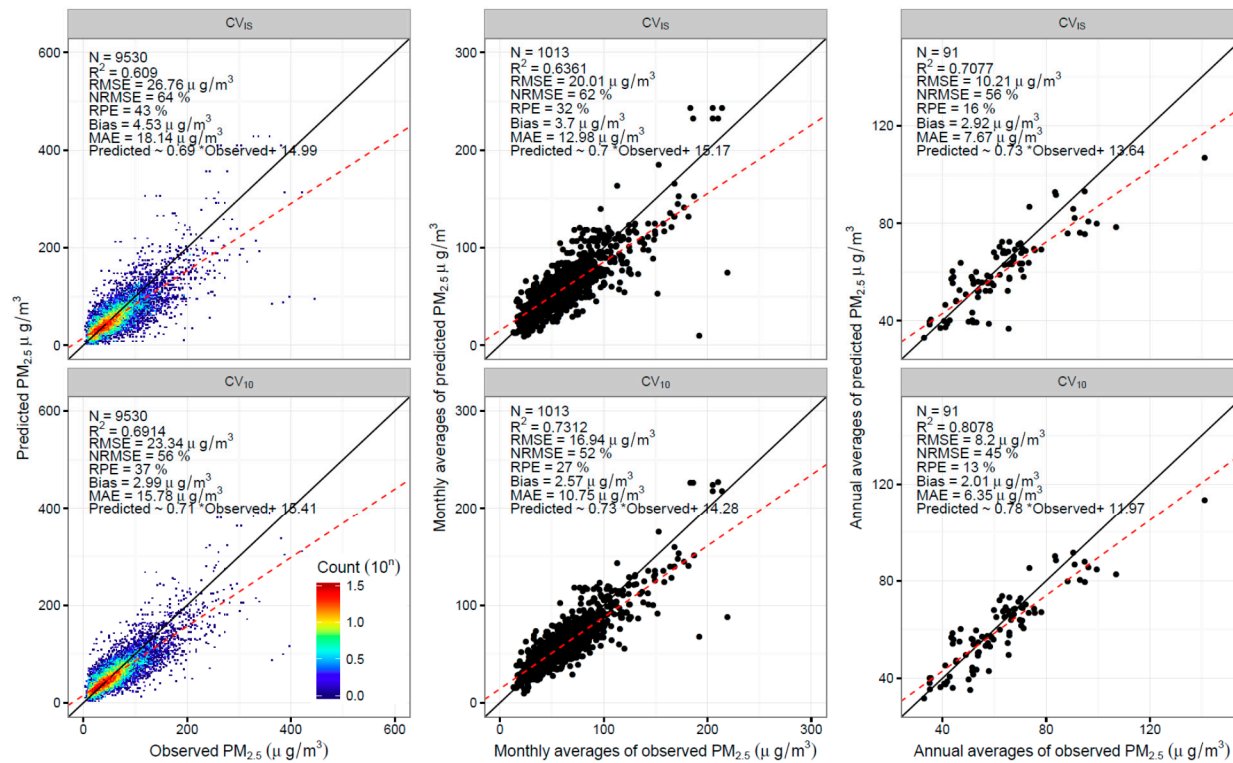
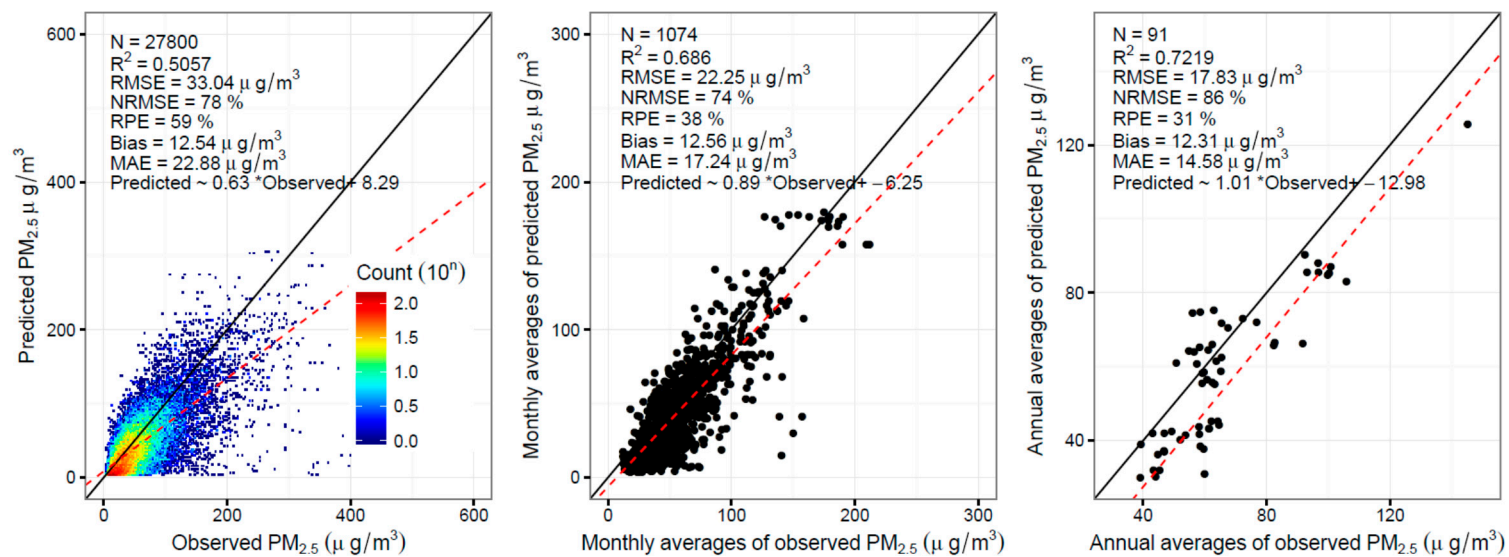Figure S1. Scatterplots to compare CV_{10} and CV_{IS} using AOD-derived PM2.5 from a LME model.

Figure S2 Scatterplots of cross-validated values and their monthly or annual averages for downscaled CMAQ PM₂.₅ (0.1°*0.1°).

Left panel:
N = 27800
R^2 = 0.4905
RMSE = 33.38 $\mu g/m^3$
NRMSE = 79 %
RPE = 59 %
Bias = 11.64 $\mu g/m^3$
MAE = 23.31 $\mu g/m^3$
Predicted ~ 0.63 *Observed+ 9.15

Count (10^n)
2.0
1.5
1.0
0.5
0.0

Predicted PM$_{2.5}$ $\mu g/m^3$
Observed PM$_{2.5}$ ($\mu g/m^3$)

Middle panel:
N = 1074
R^2 = 0.6682
RMSE = 22.29 $\mu g/m^3$
NRMSE = 74 %
RPE = 39 %
Bias = 11.68 $\mu g/m^3$
MAE = 17.39 $\mu g/m^3$
Predicted ~ 0.88 *Observed+ −4.95

Monthly averages of predicted PM$_{2.5}$ $\mu g/m^3$
Monthly averages of observed PM$_{2.5}$ ($\mu g/m^3$)

Right panel:
N = 91
R^2 = 0.6964
RMSE = 17.68 $\mu g/m^3$
NRMSE = 86 %
RPE = 31 %
Bias = 11.43 $\mu g/m^3$
MAE = 14.33 $\mu g/m^3$
Predicted ~ 1 *Observed+ −11.17

Annual averages of predicted PM$_{2.5}$ $\mu g/m^3$
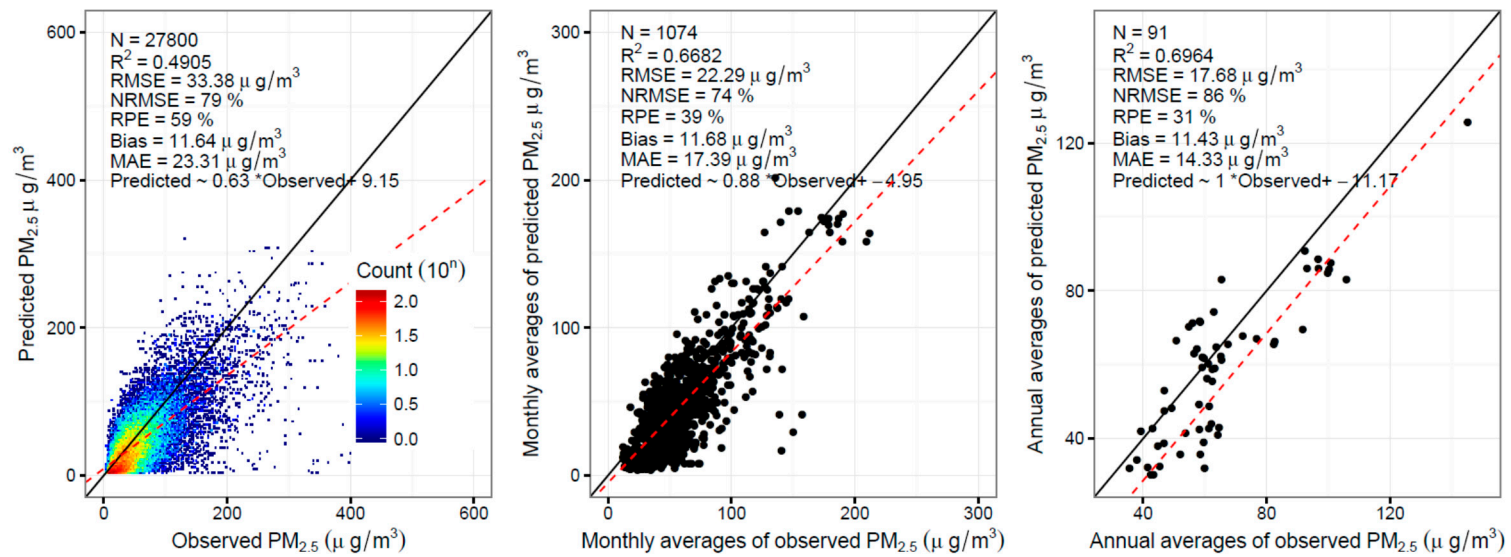Annual averages of observed PM$_{2.5}$ ($\mu g/m^3$)

Figure S3 Scatterplots of cross-validated values and their monthly or annual averages for raw CMAQ PM$_{2.5}$ (36km*36km).
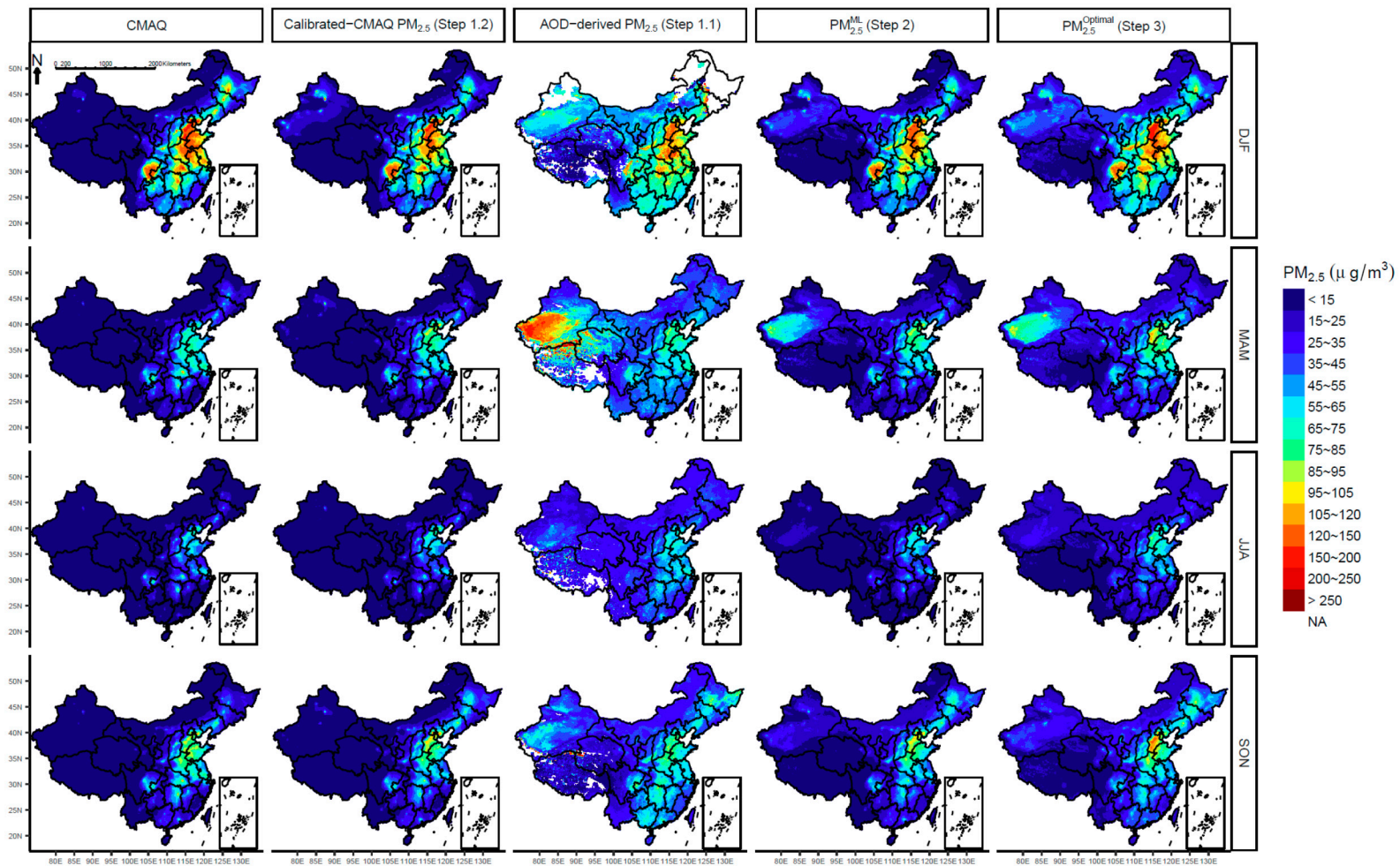
Figure S4 Seasonal maps of PM2.5 in 2014 over China, produced by CMAQ, intermediate and final estimators of the three-stage model.
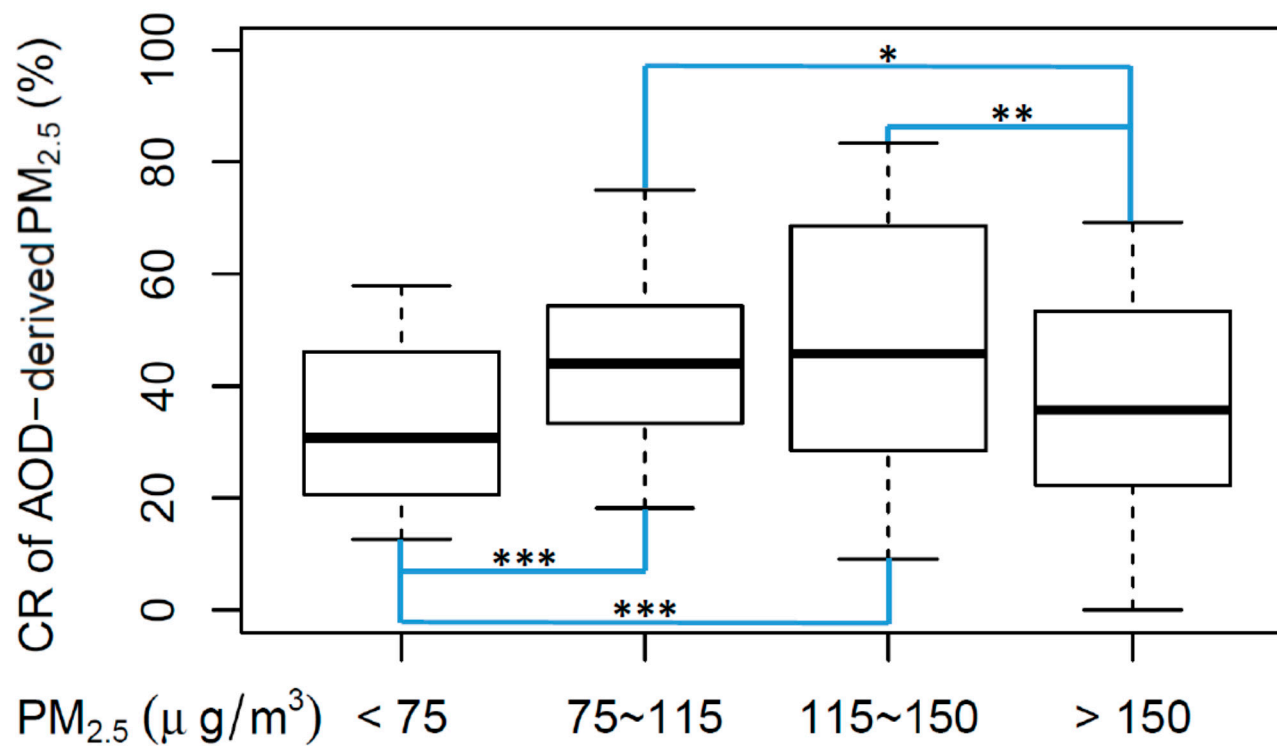
Figure S5 Comparisons of coverage rate (CR) of AOD-derived $PM_{2.5}$ by groups of observational $PM_{2.5}$ at the CV_{IS} testing sites. CR is defined by percentage of AOD-available days among all monitoring period for each site. Different CRs between groups are examined using ANOVA approach and P-values (***: P < 0.01; **: $0.01 \leq P < 0.05$; *: $0.05 \leq P < 0.1$) for pairwise comparisons are calculated using Bonferroni correction.
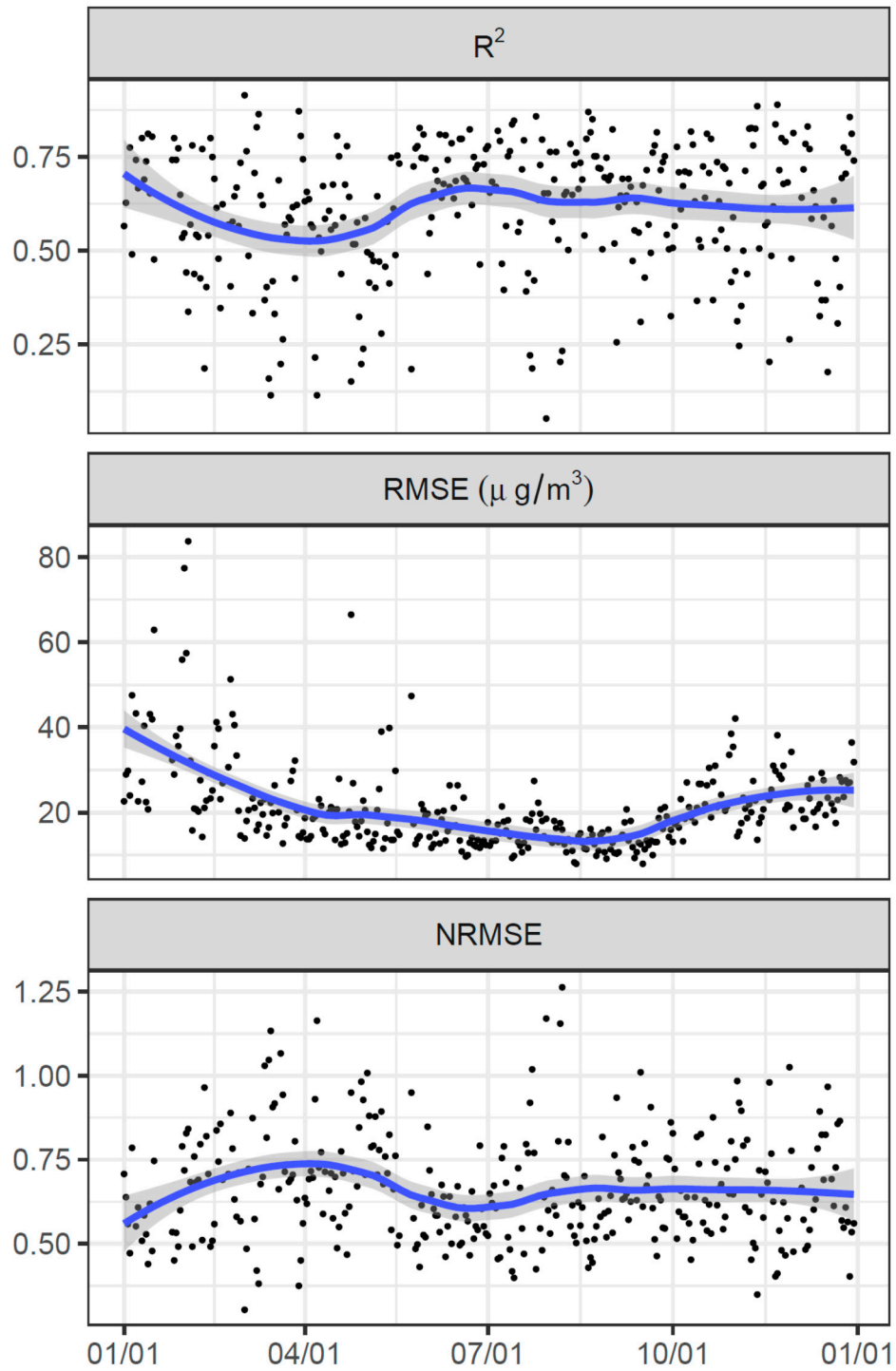
Figure S6 Temporal variations of CV results for the final estimator (PM$_{2.5}$^{Optimal}).
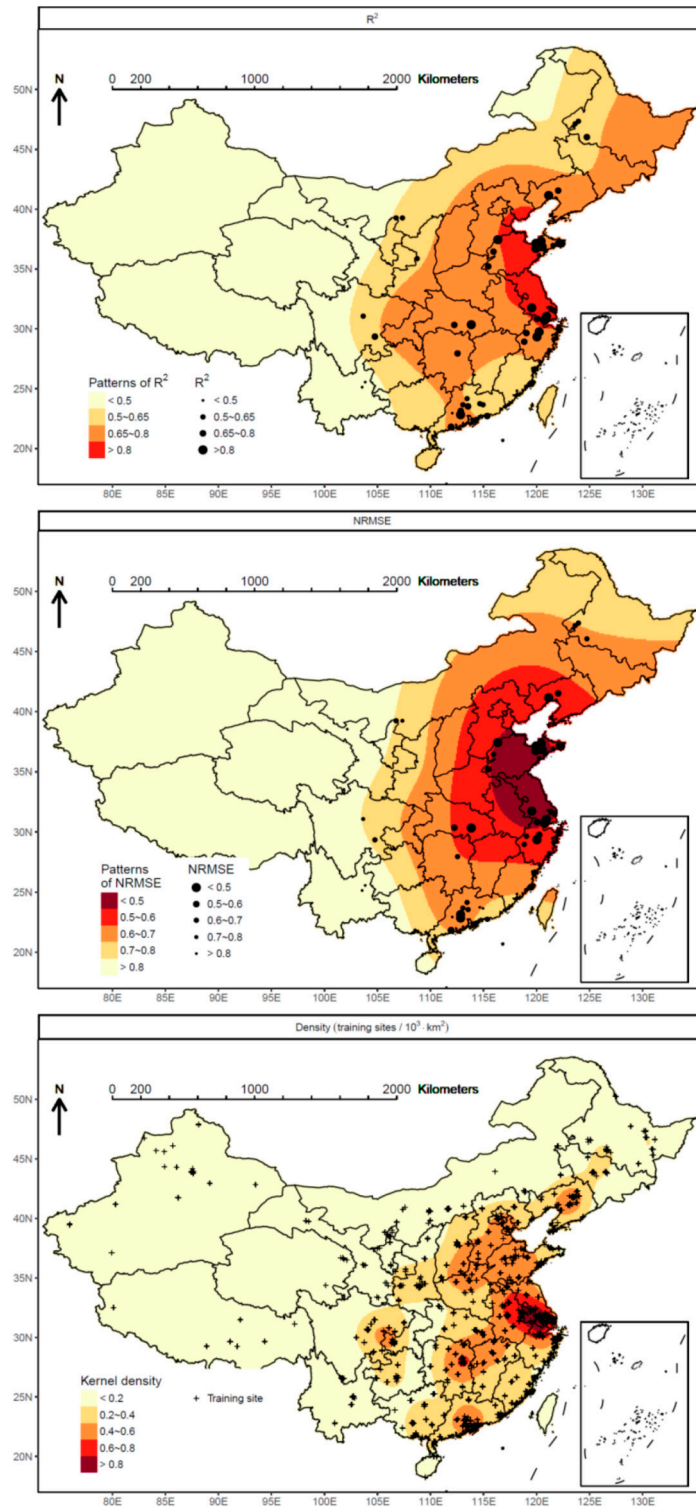
Figure S7 Spatial distributions of CV results for the final estimator ($PM_{2.5}^{Optimal}$). The spatial patterns of CV errors were fitted by a 2D-splines regression and are presented by colors. The density distribution for the training sites were estimated using a 2D-Kernel with a bandwidth of 50 km and are presented in the bottom panel.
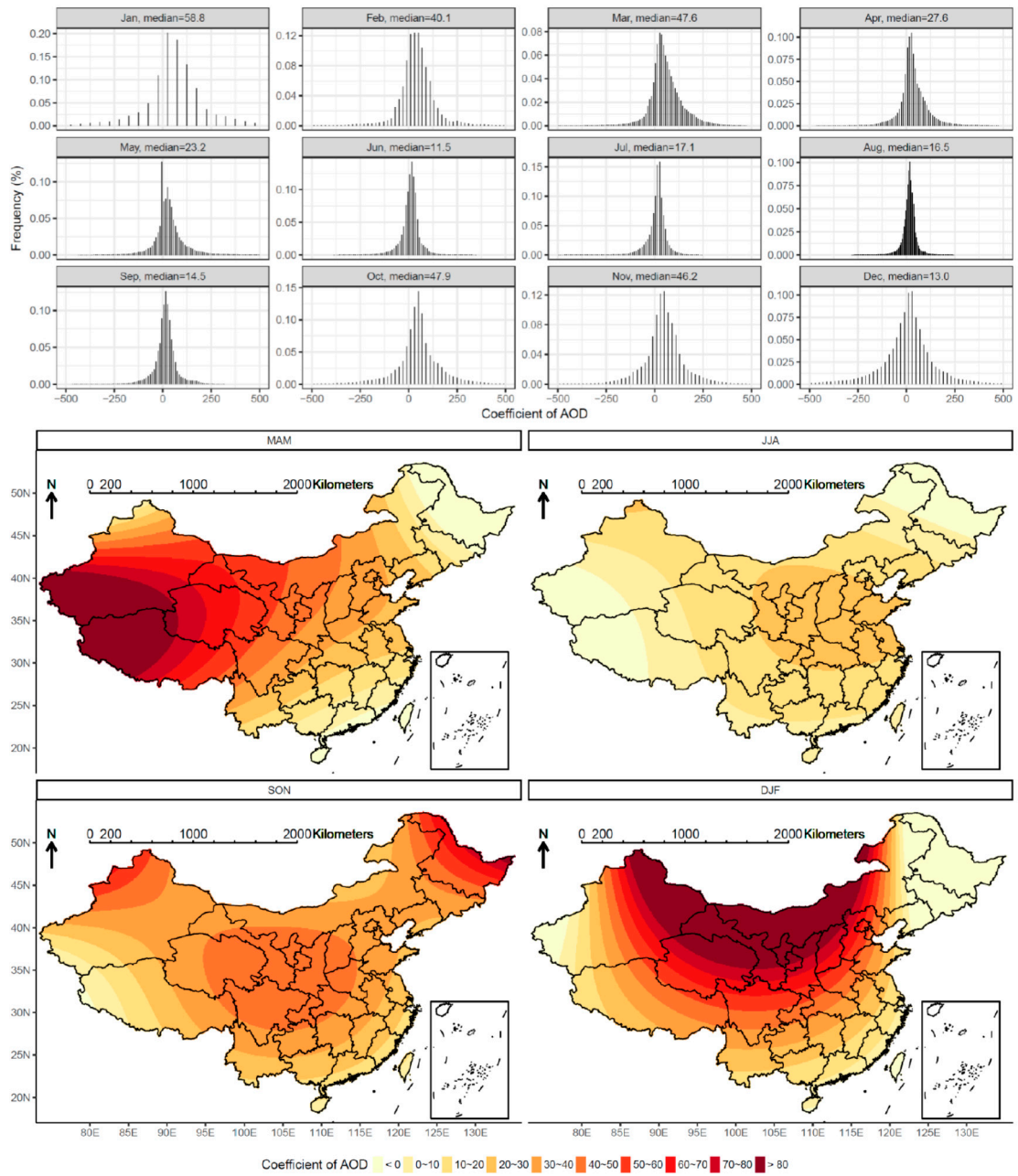
Figure S8 Distributions of coefficients for AOD by months (upper panel) and their spatial patterns by seasons (lower panel) in Equation (1).
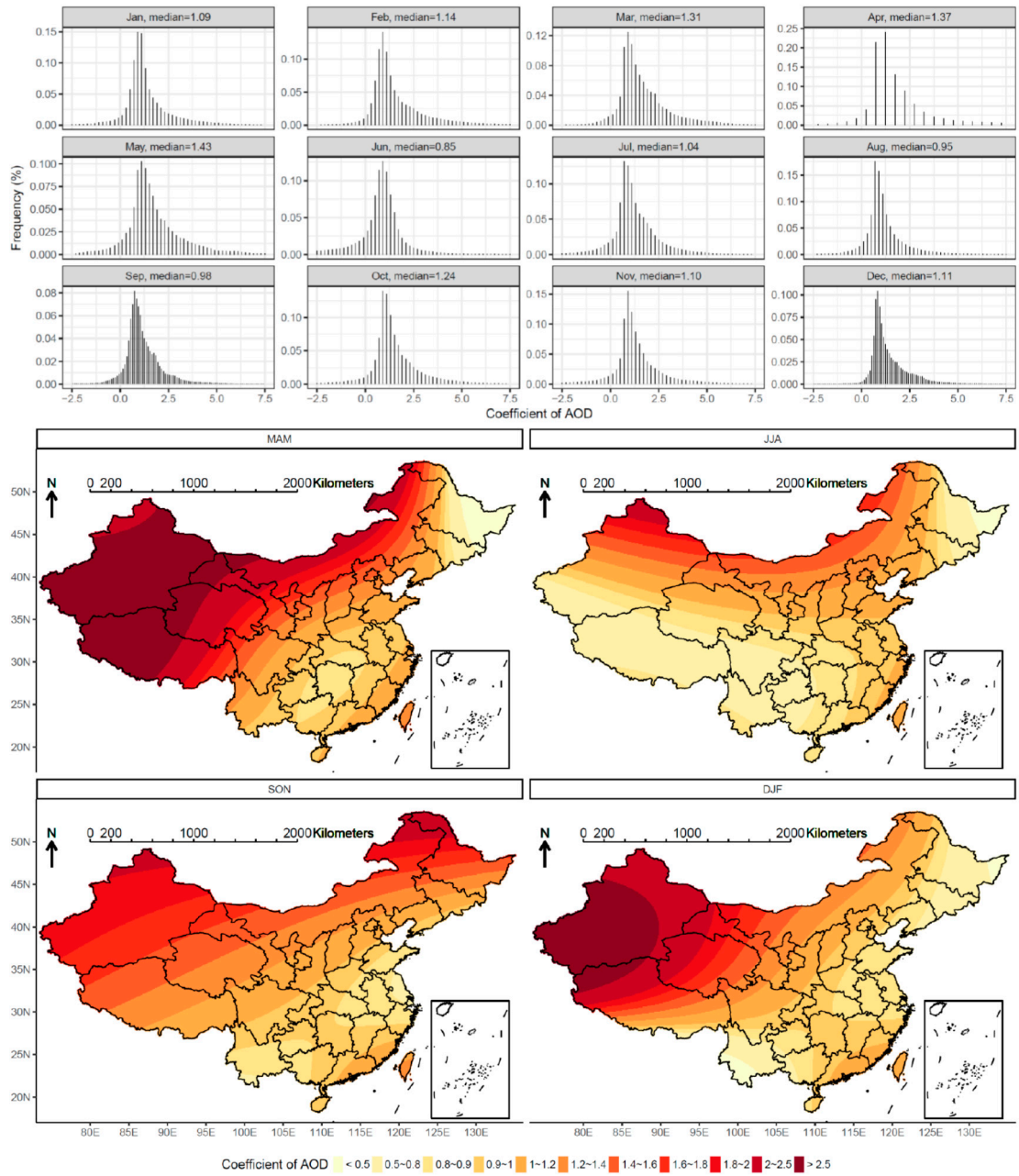
Figure S9 Distributions of coefficients for CMAQ-simulated PM$_{2.5}$ by months (upper panel) and their spatial patterns by seasons (lower panel) in Equation (2).