Supplemental materials: Fusing Observational, Satellite Remote Sensing and Air Quality Model Simulated Data to Estimate Spatiotemporal Variations of $PM_{2.5}$ Exposure in China

Tao Xue[1], Ph.D; Yixuan Zheng[1], M.S.; Guannan Geng[1,2], M.S.;

Bo Zheng[2], M.S.; Xujia Jiang[2], Ph.D;

Qiang Zhang[1,3,*], Ph.D; Kebin He[2,3], Ph.D

1 Ministry of Education Key Laboratory for Earth System Modeling, Center for Earth System Science, Tsinghua University, Beijing 100084, China

2 State Key Joint Laboratory of Environmental Simulation and Pollution Control, School of Environment, Tsinghua University, Beijing 100084, China

3 The Collaborative Innovation Center for Regional Environmental Quality, Beijing 100084, China

* Corresponding author

Email: qiangzhang@tsinghua.edu.cn

Phone & Tax: +86-10-62795090

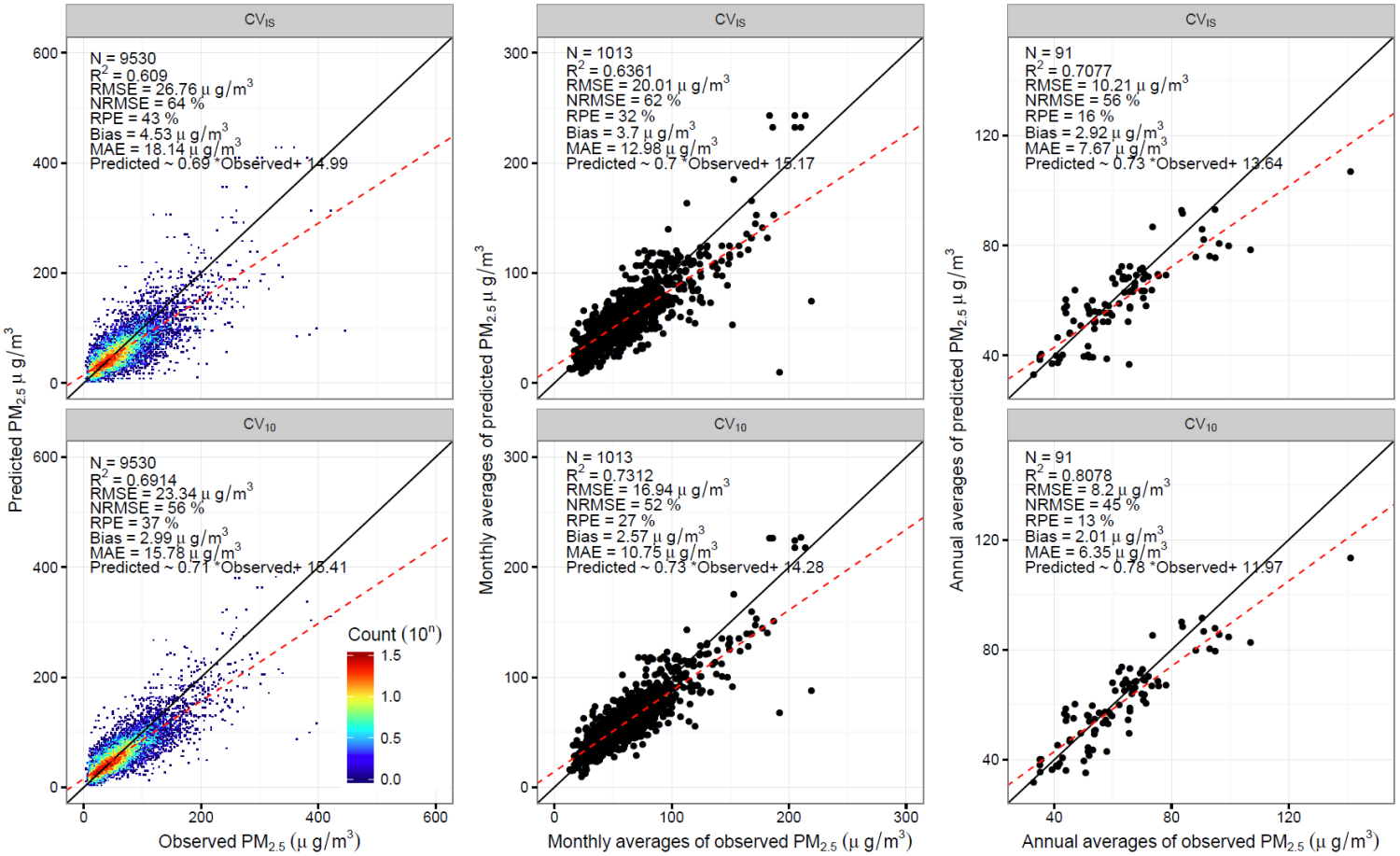

Figure S1 Scatterplots to compare CV_{10} and CV_{IS} using AOD-derived $PM_{2.5}$ from a LME model.

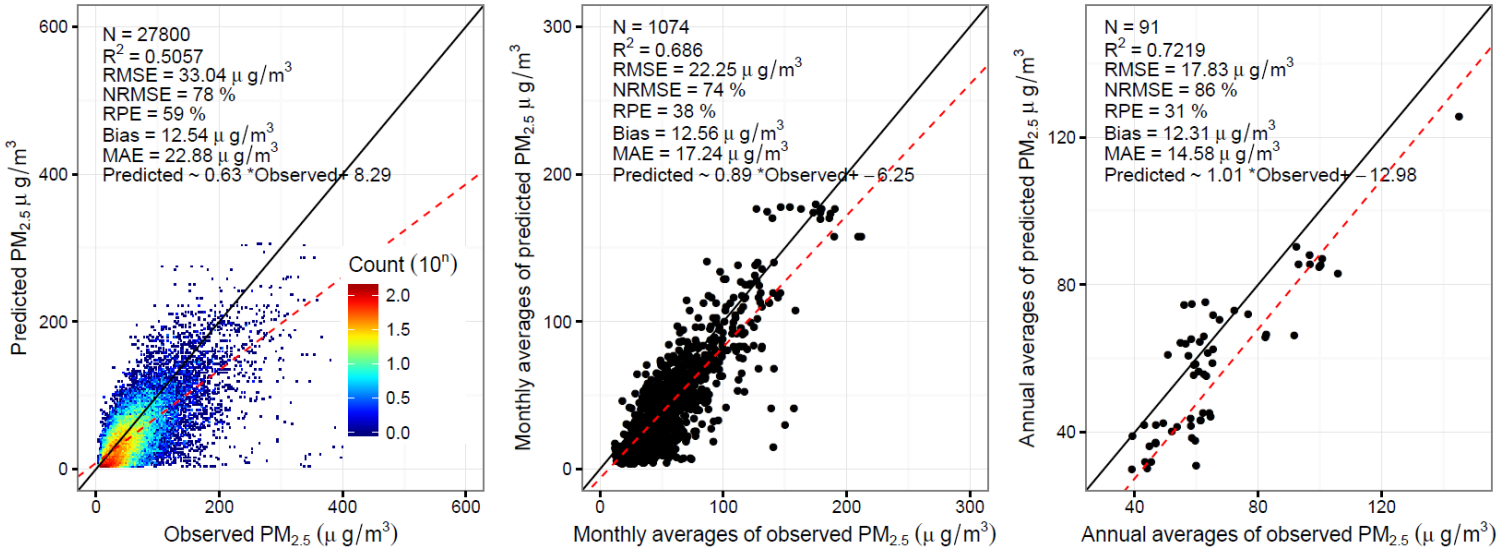

Figure S2 Scatterplots of cross-validated values and their monthly or annual averages for downscaled CMAQ $PM_{2.5}$ (0.1˚*0.1˚).

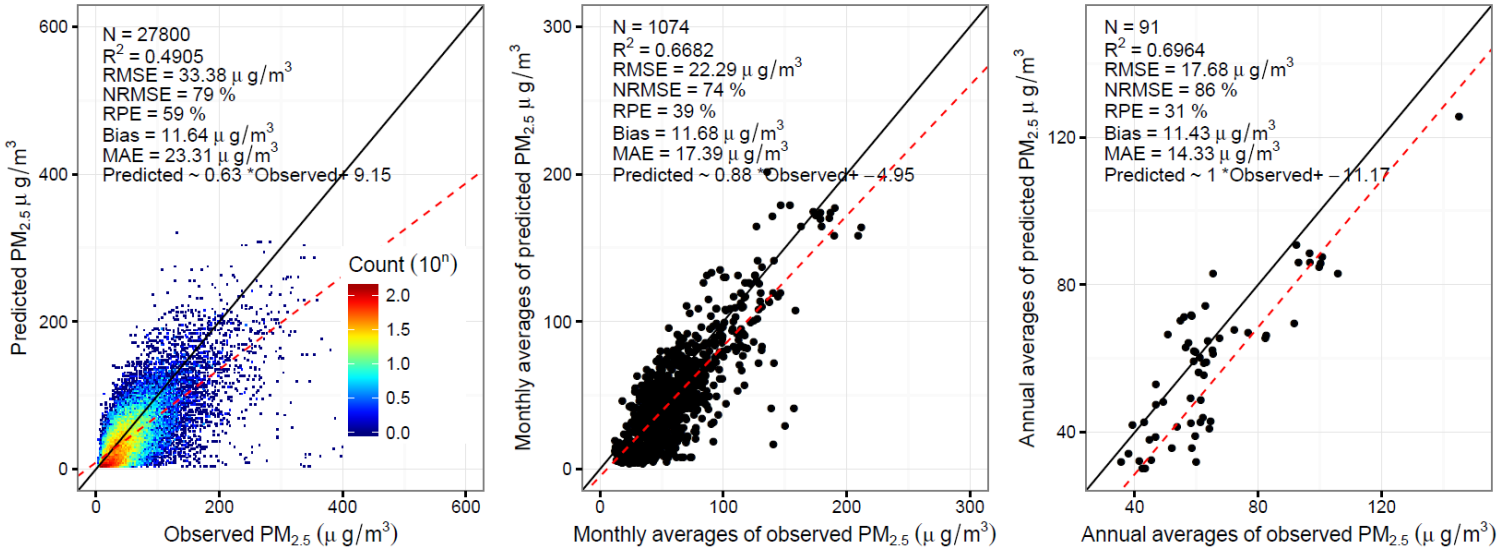

Figure S3 Scatterplots of cross-validated values and their monthly or annual averages for raw CMAQ $PM_{2.5}$ (36km*36km).

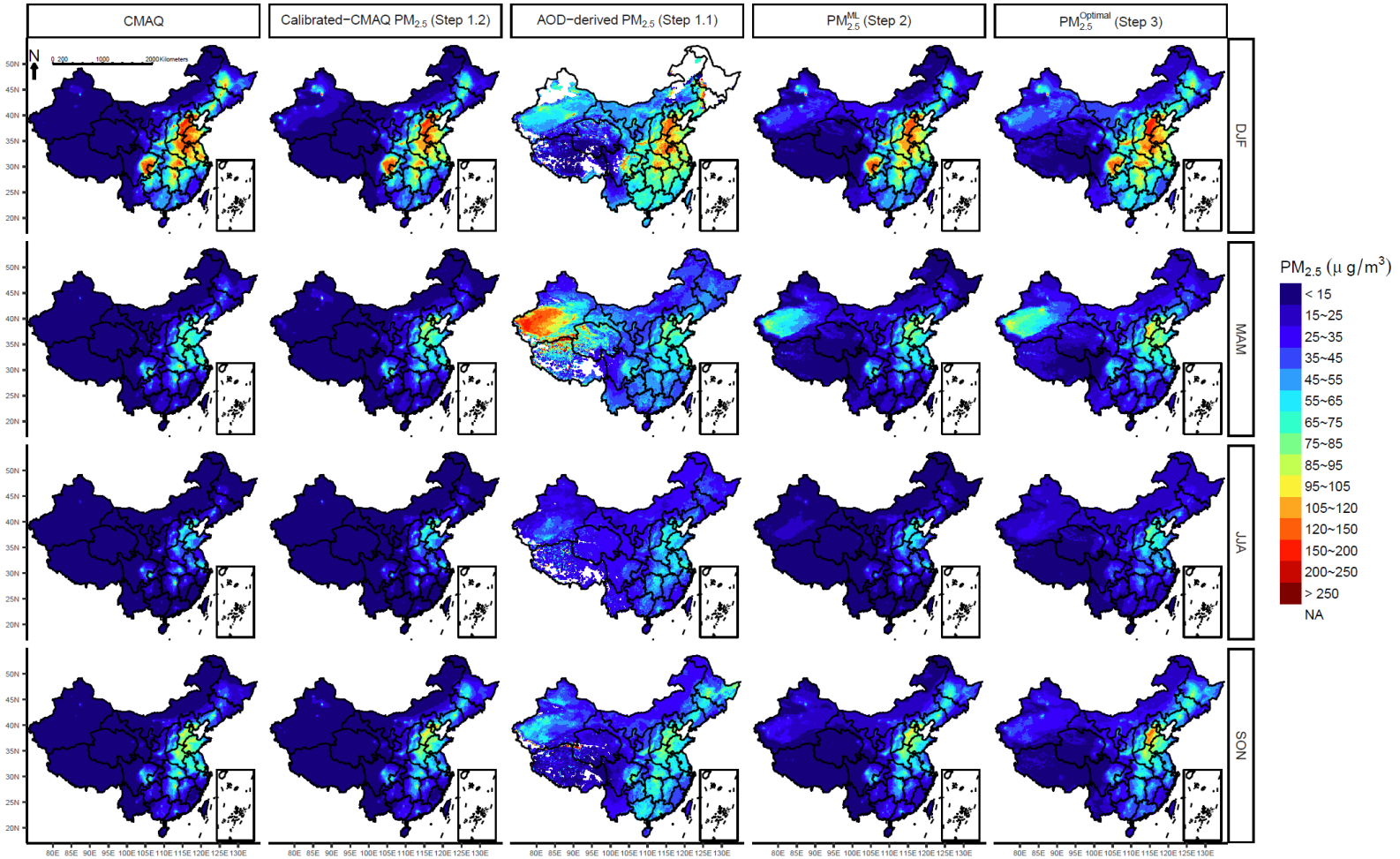

Figure S4 Seasonal maps of PM2.5 in 2014 over China, produced by CMAQ, intermediate and final estimators of the three-stage model.

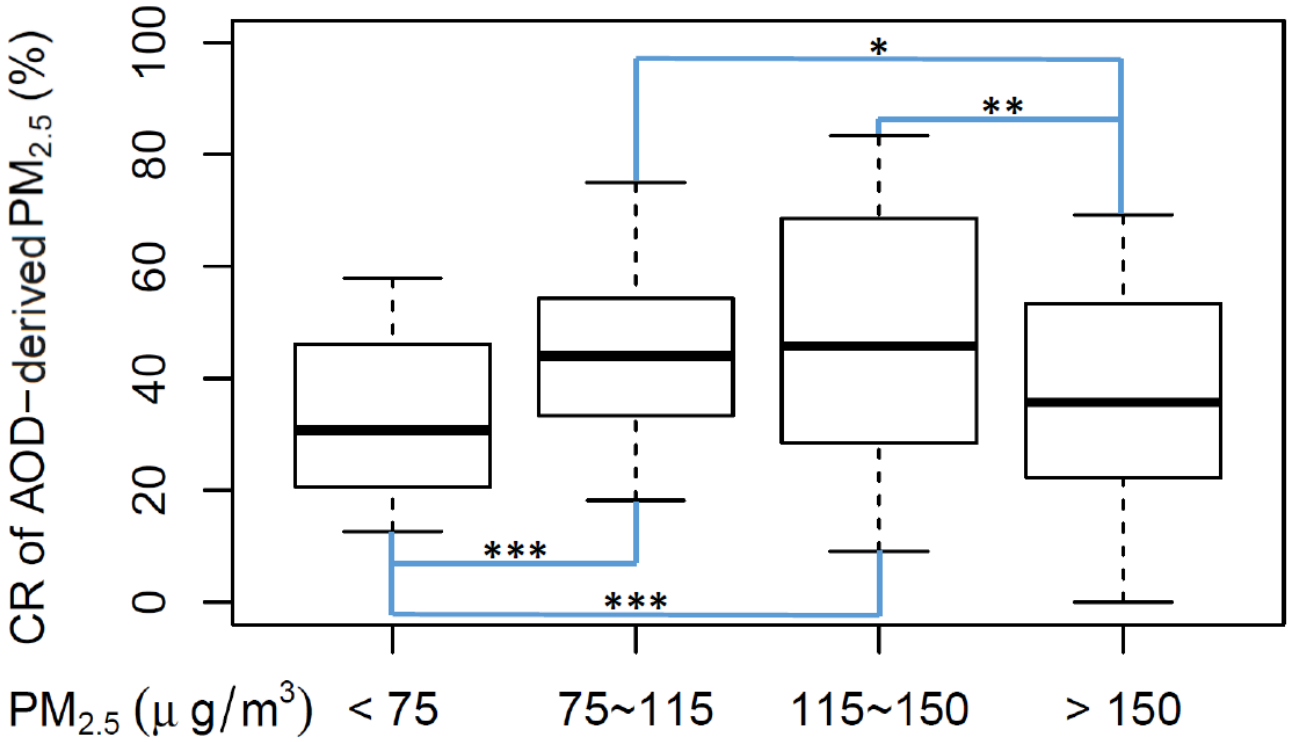

Figure S5 Comparisons of coverage rate (CR) of AOD-derived $PM_{2.5}$ by groups of observational $PM_{2.5}$ at the CV_{IS} testing sites. CR is defined by percentage of AOD-available days among all monitoring period for each site. Different CRs between groups are examined using ANOVA approach and P-values (***: $P < 0.01$; **: $0.01 \leq P < 0.05$; *: $0.05 \leq P<0.1$) for pairwise comparisons are calculated using Bonferroni correction.

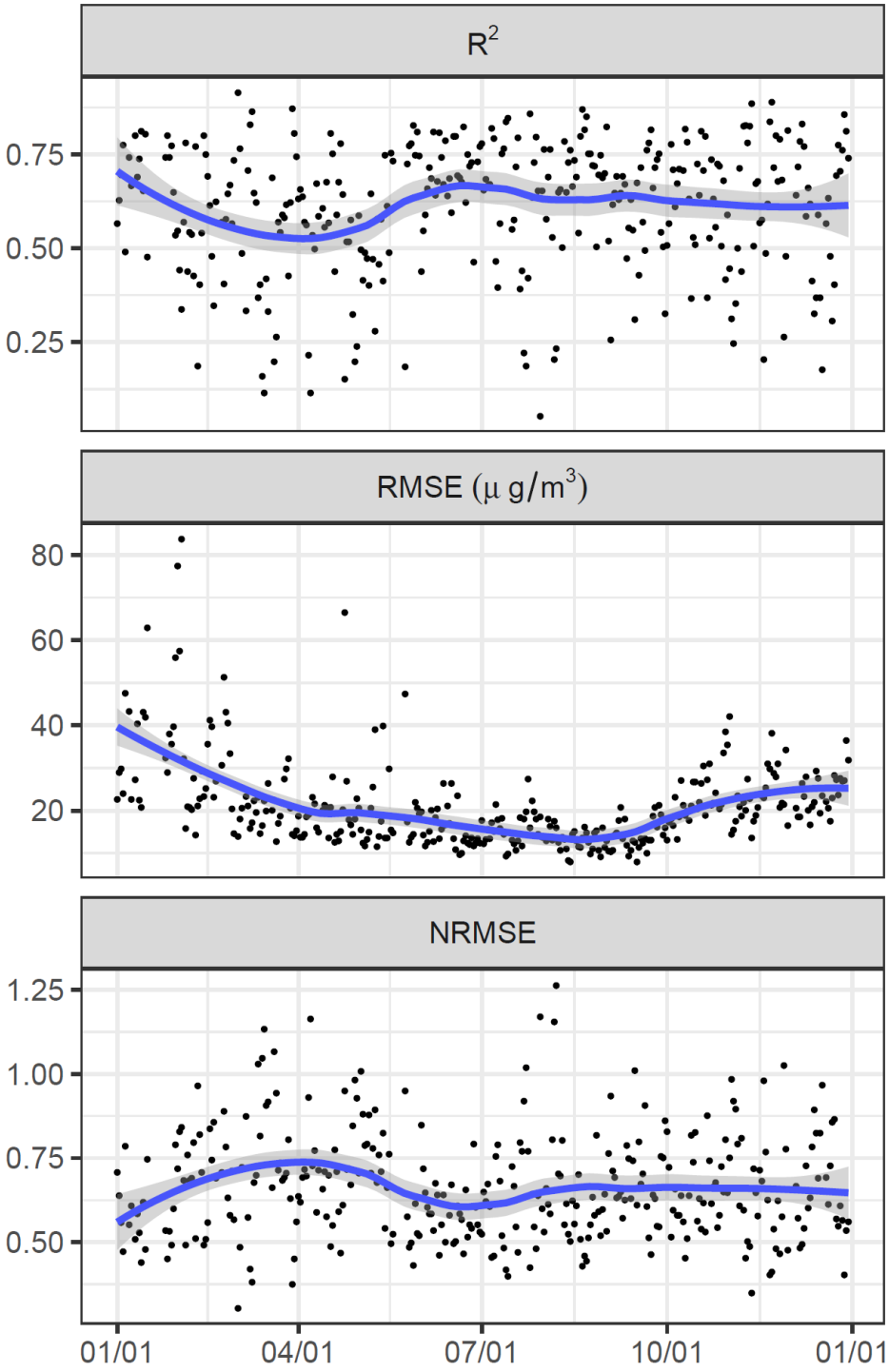

Figure S6 Temporal variations of CV results for the final estimator ($PM_{2.5}^{Optimal}$).

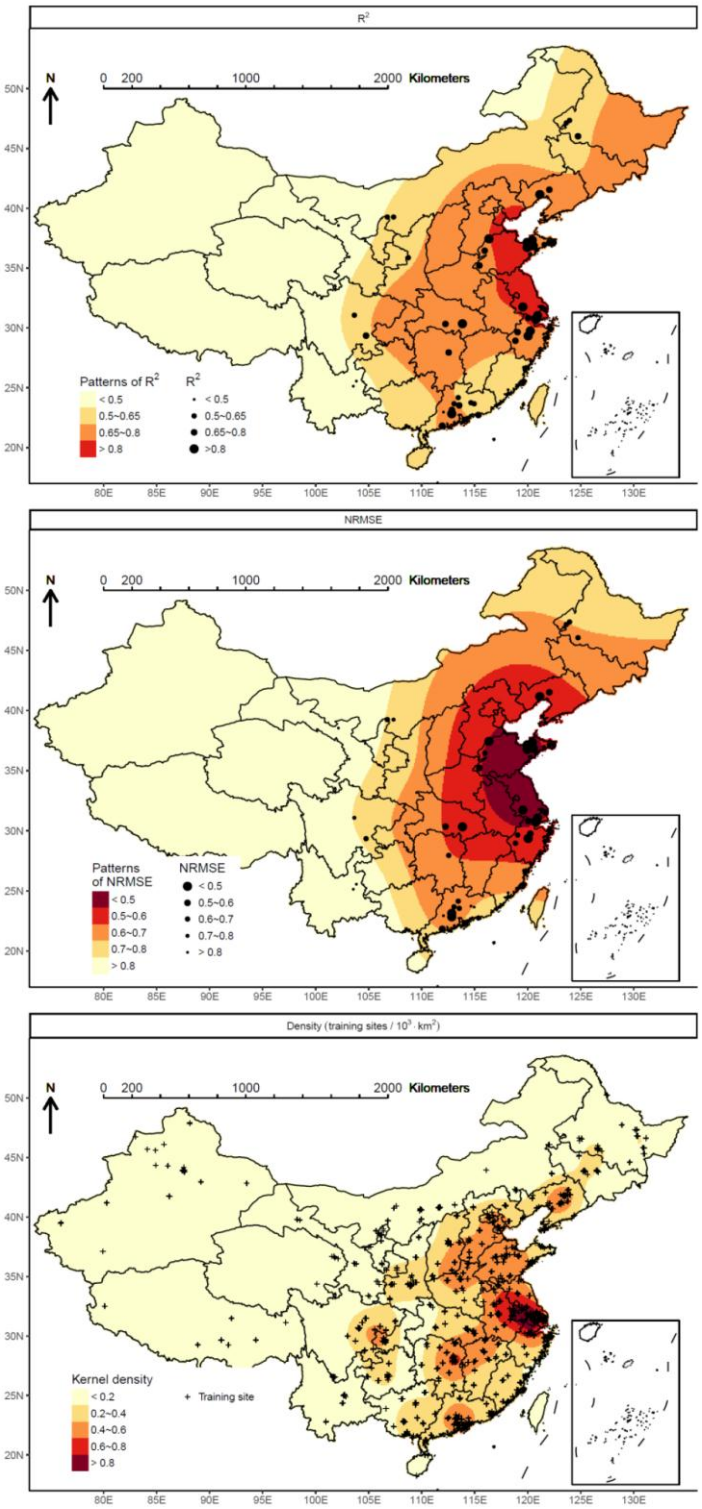

Figure S7 Spatial distributions of CV results for the final estimator ($PM_{2.5}^{Optimal}$). The spatial patterns of CV errors were fitted by a 2D-splines regression and are presented by colors. The density distribution for the training sites were estimated using a 2D-Kernel with a bandwidth of 50 km and are presented in the bottom panel.

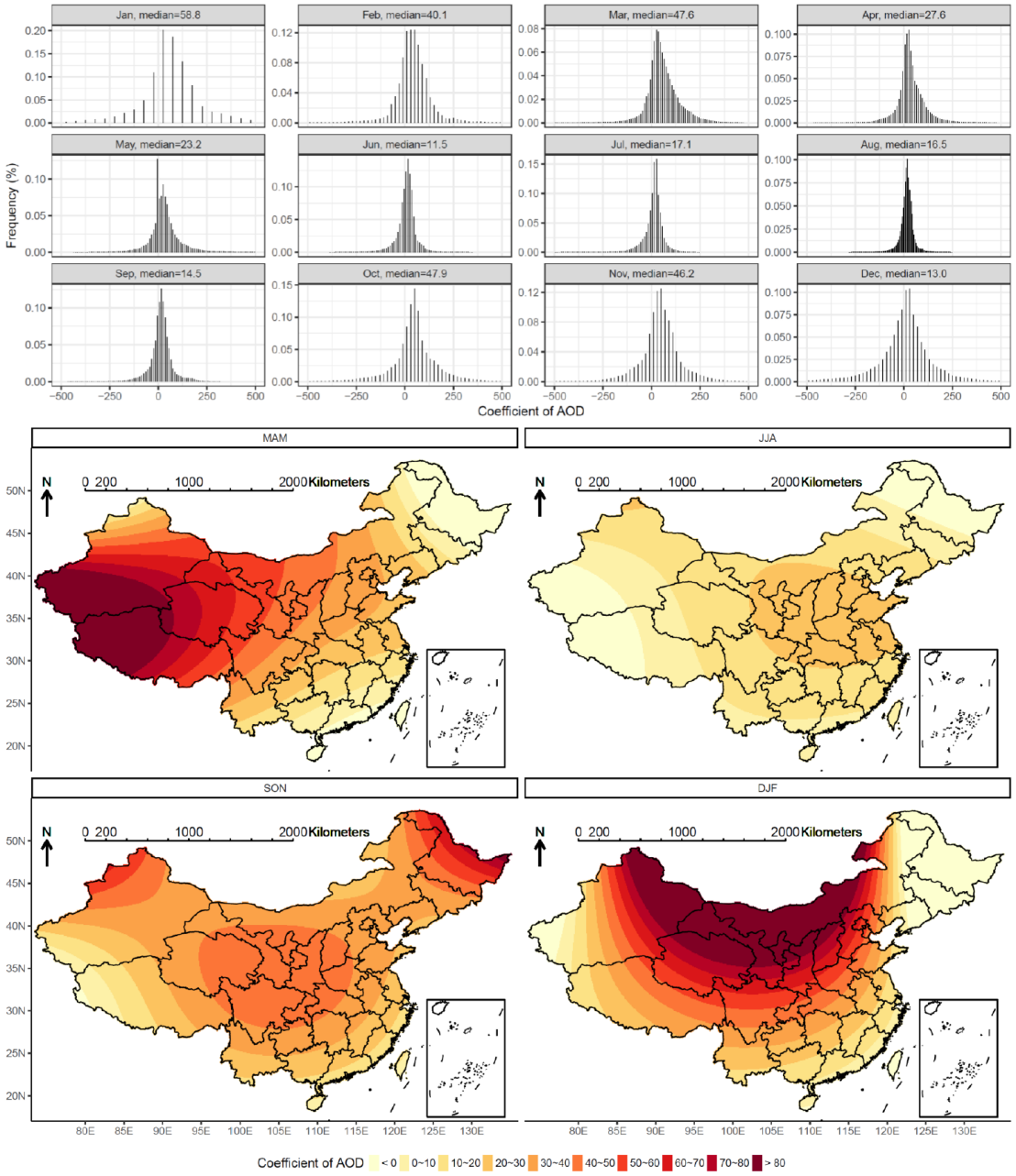

Figure S8 Distributions of coefficients for AOD by months (upper panel) and their spatial patterns by seasons (lower panel) in Equation (1).

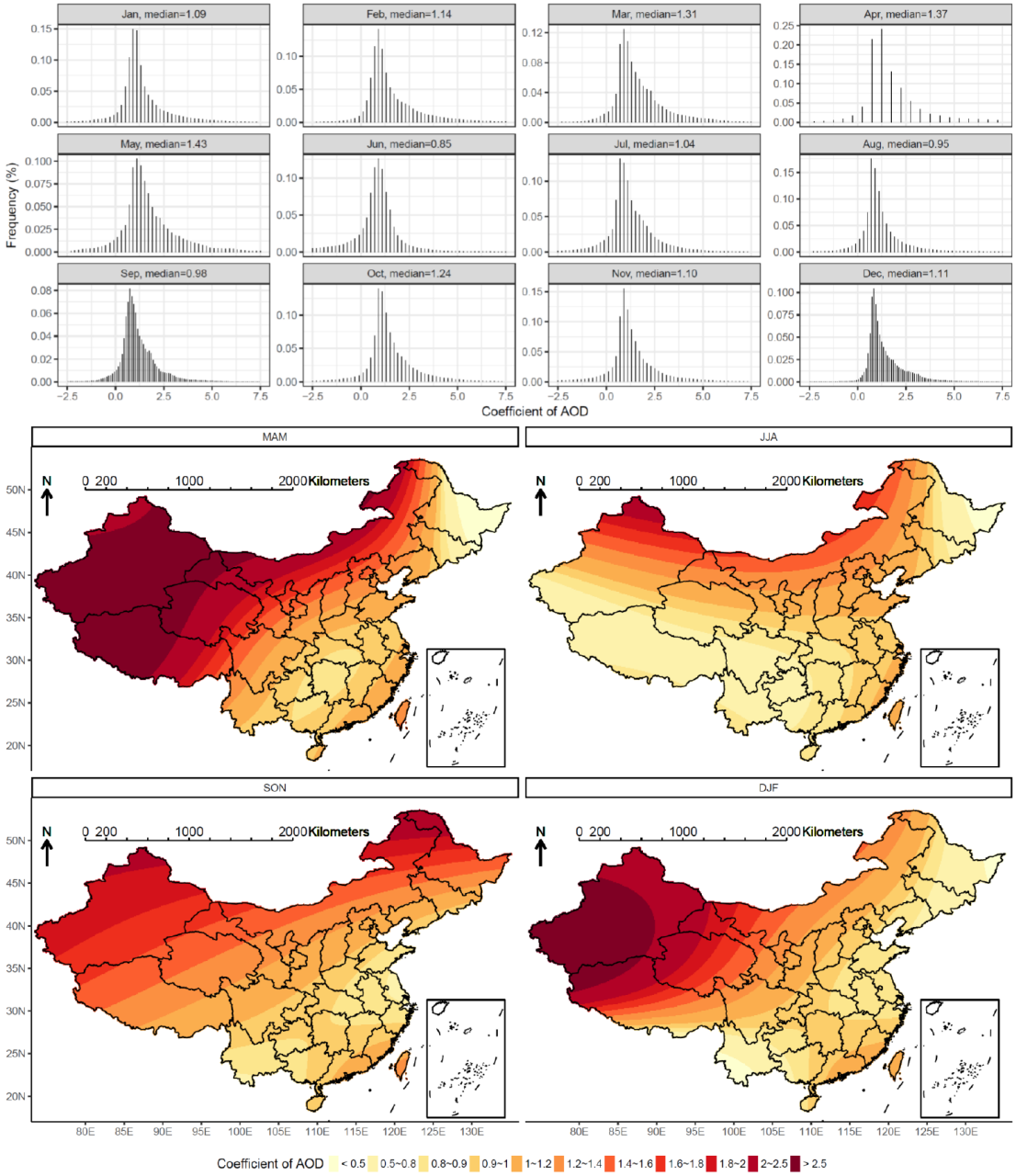

Figure S9 Distributions of coefficients for CMAQ-simulated $PM_{2.5}$ by months (upper panel) and their spatial patterns by seasons (lower panel) in Equation (2).