atmosphere

MDPI

Supplementary Materials

Estimation of the Near-Surface Ozone Concentration With Full Spatiotemporal Coverage Across the Beijing-Tianjin-Hebei Re-Gion Based on Extreme Gradient Boosting Combined With Wrf-Chem Model

Xiaomin Hu [1], Jing Zhang [1,*], Wenhao Xue [2,*], Lihua Zhou [3], Yunfei Che [1,4] and Tian Han [1]

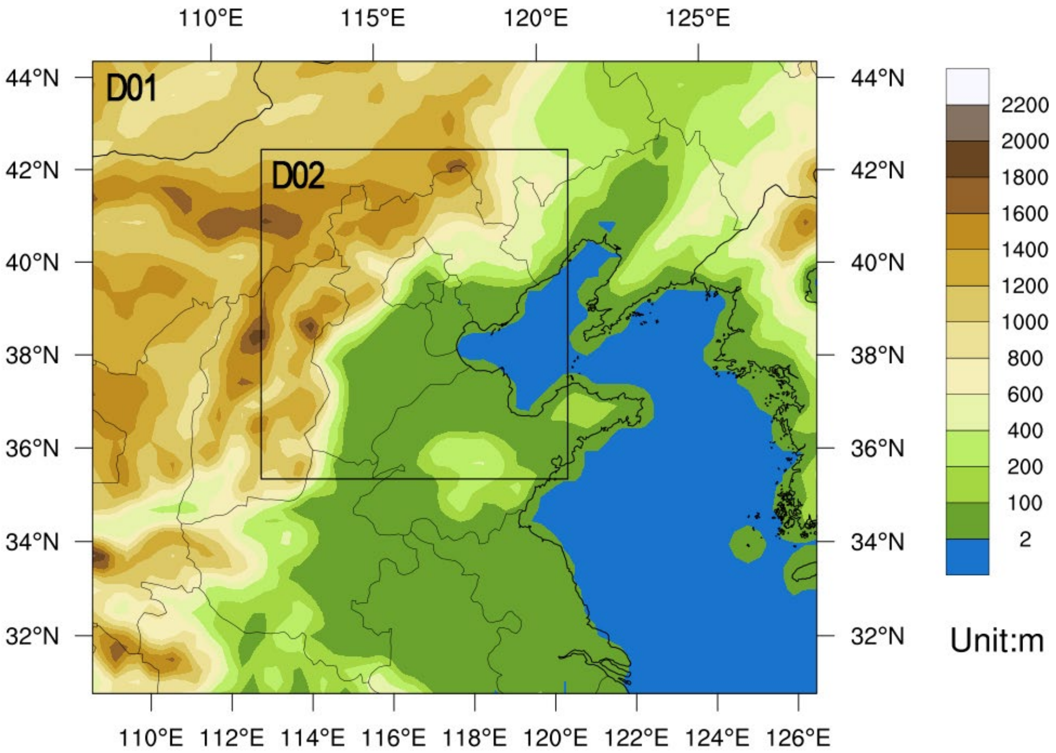

Figure S1. Double-layer nesting distribution in the WRF-Chem model, the color bar represents the altitude.

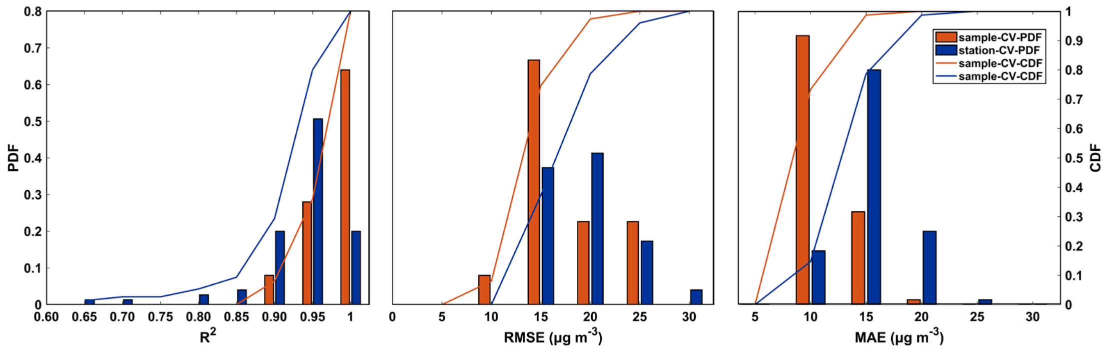

Figure S2. Probability density functions (PDFs) and cumulative density functions (CDFs) of the sample-based (red columns) and station-based (blue columns) 10-fold cross-validation.

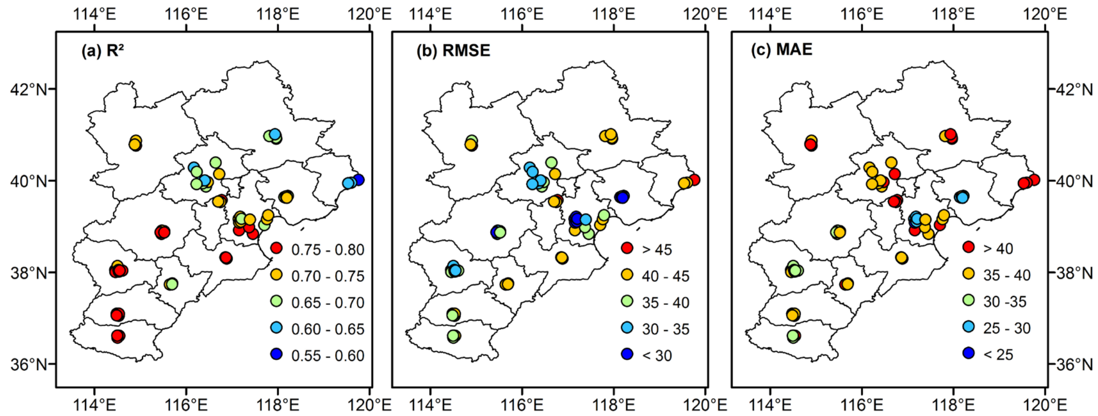

Figure S3. The validation results of the WRF-Chem simulations and site observations in the Beijing-Tian-Hebei region in 2018.

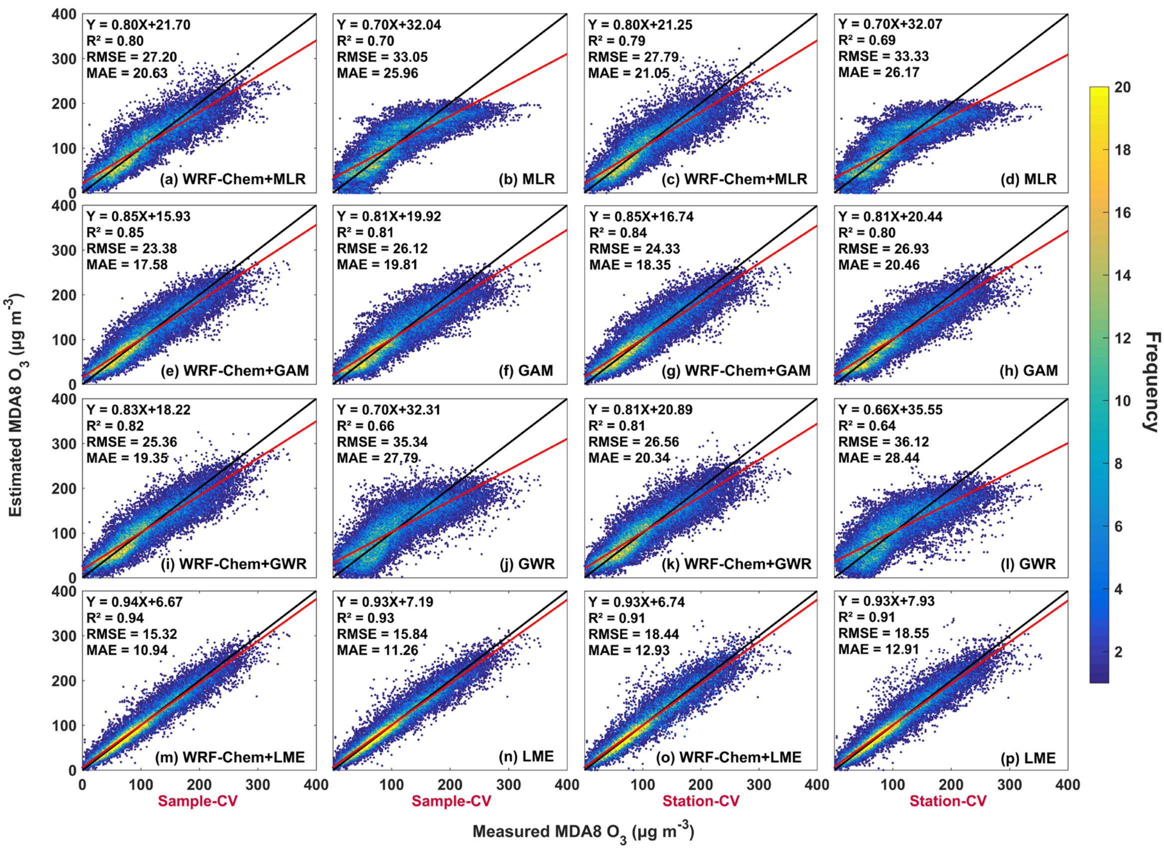

Figure S4. The scatter density plot of the final estimation accuracy of different traditional models with and without fusion of the WRF-Chem model.

Table S1. Comparison of MDA8 O_3 concentration estimation accuracy between integrating different machine learning algorithms with the WRF-Chem model and the WRFC-XGB model (KNN：K-Nearest Neighbor; SVR: Support Vector Regression; DT：Decision Tree; ET：Extra Tree; RF：Random Forest; GBM: Gradient Boosting Machine).

Model	sample-based 10-CV			station-based 10-CV		
	R^2	RMSE(μg/m^3)	MAE(μg/m^3)	R^2	RMSE(μg/m^3)	MAE(μg/m^3)
WRFC-KNN	0.90	18.86	12.86	0.85	23.57	16.45
WRFC-SVR	0.79	27.80	20.63	0.77	28.80	21.70
WRFC-DT	0.83	24.50	18.06	0.82	25.62	18.89
WRFC-ET	0.84	24.29	17.77	0.82	25.98	19.10
WRFC-RF	0.90	19.34	14.21	0.88	21.11	15.49
WRFC-GBM	0.90	18.77	13.92	0.88	20.65	15.29
WRFC-XGB	0.95	13.50	9.60	0.91	17.70	12.89