Geoscientific
Model Development

Supplement of

Observational operator for fair model evaluation with ground NO$_2$ measurements

Li Fang et al.

Correspondence to: Jianbing Jin (jianbing.jin@nuist.edu.cn) and Hong Liao (hongliao@nuist.edu.cn)
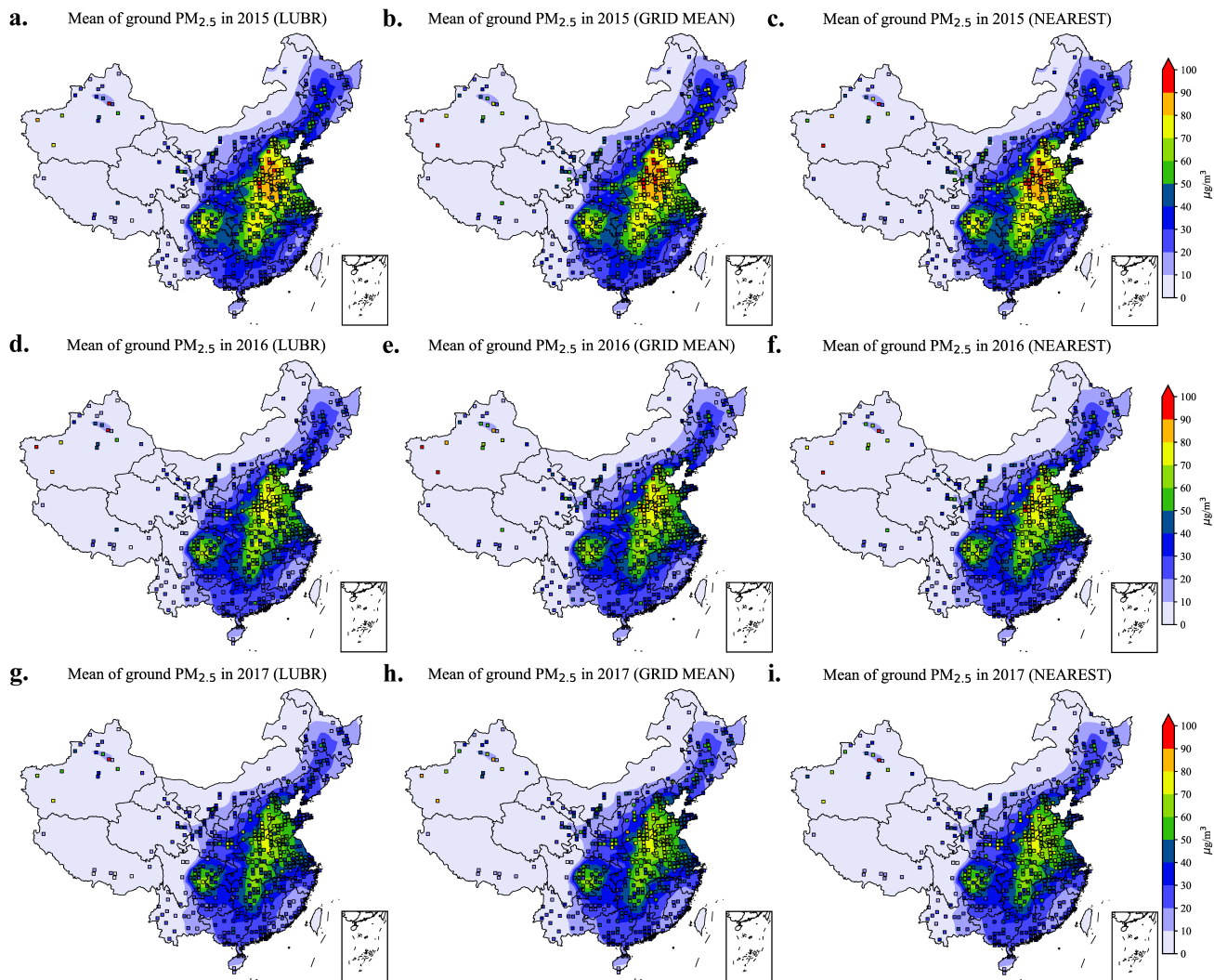
Figure S1. The annual averaged ground PM$_{2.5}$ from GEOS-Chem simulations (filled contours) and the represented observations of simulation grids (colored squares) from three methods. Panels a, d, and g present results using the LUBR method to represent grid PM$_{2.5}$ concentrations for 2015, 2016, and 2017, respectively. Panels b, e, and h present results using the grid mean method. Panels c, f, and i present results using the nearest search method.

S1 Formula of the statistic matrics

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}, \tag{S1}$$

1

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i|, \tag{S2}$$

$$\text{NMB} = \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)}{\sum_{i=1}^{n} \hat{y}_i} \times 100\%, \tag{S3}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2}, \tag{S4}$$

where n represents the size of the dataset, $\hat{y}_i$ and y_i denote the predicted values of GEOS-Chem and true observations, respectively, and $\bar{y}$ represents the mean of the observed values.
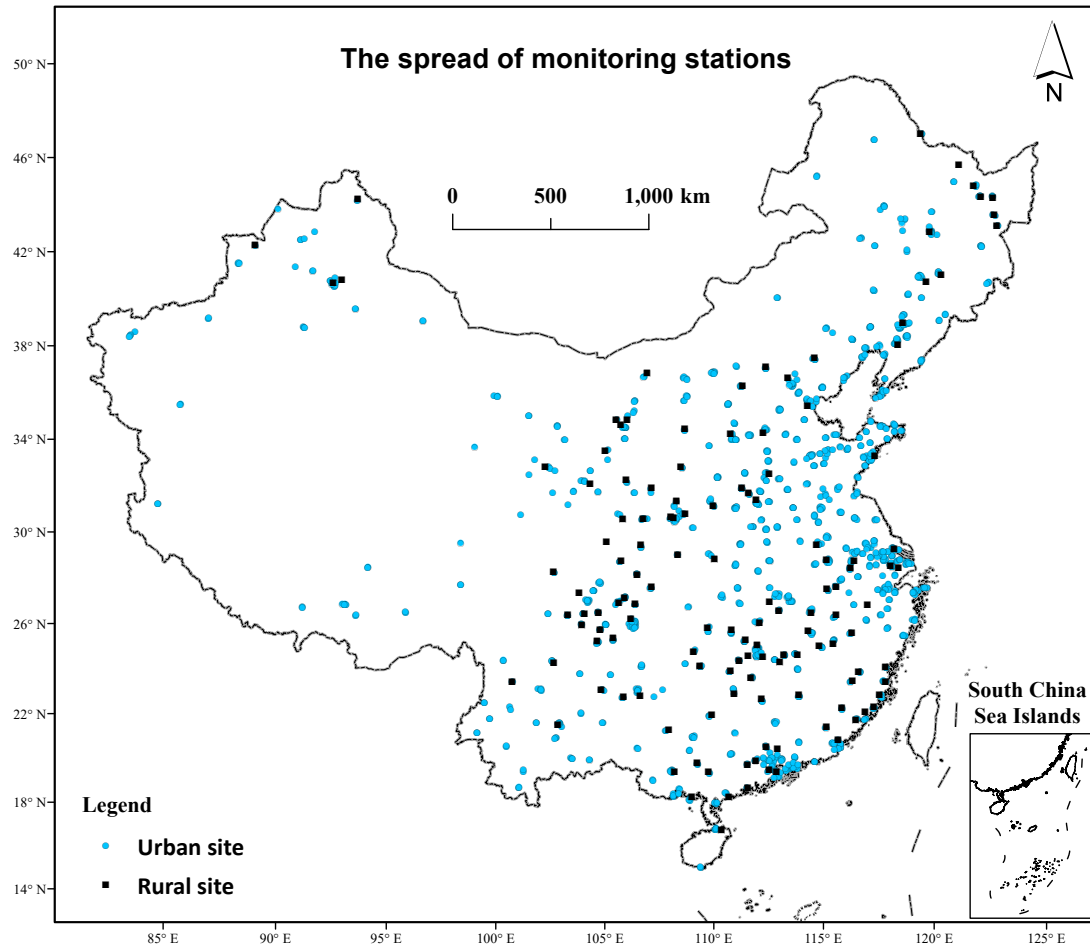
Figure S2. The spread of air quality monitoring stations in the study area. The turquoise dots and black squares represent urban sites and rural sites, respectively.
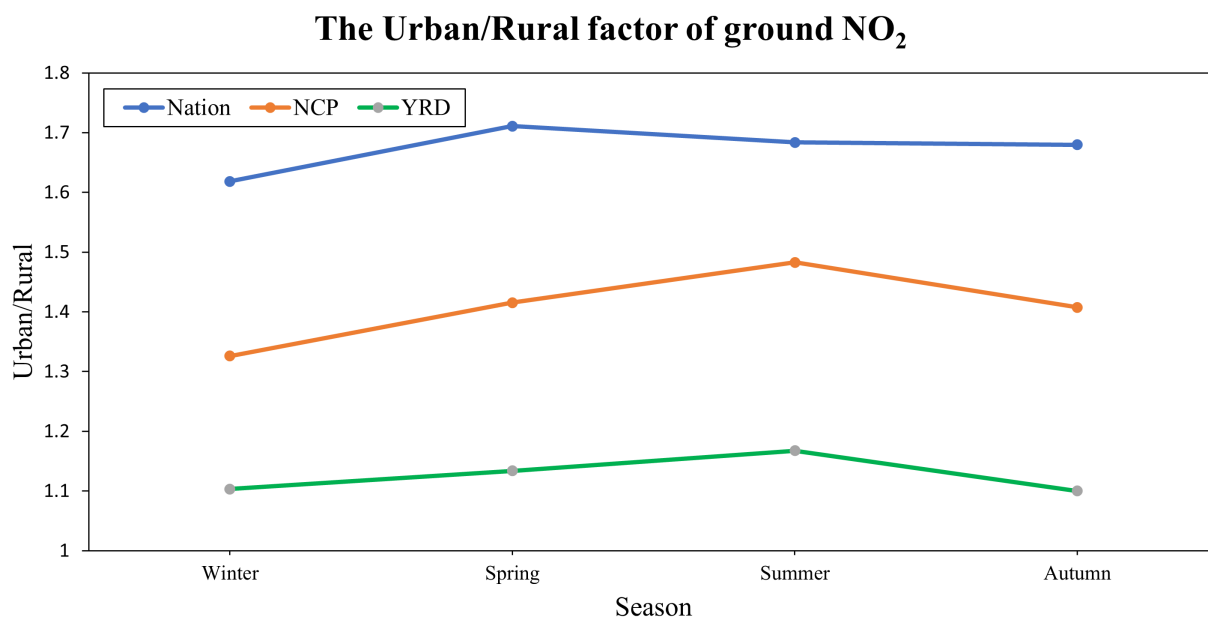
Figure S3. The seasonal Urban/rural factors of ground NO_2. Each season is calculated from the average of 2015-2017 data. The blue dots correspond to the national scale, while the orange and green dots represent the NCP and YRD regions, respectively.
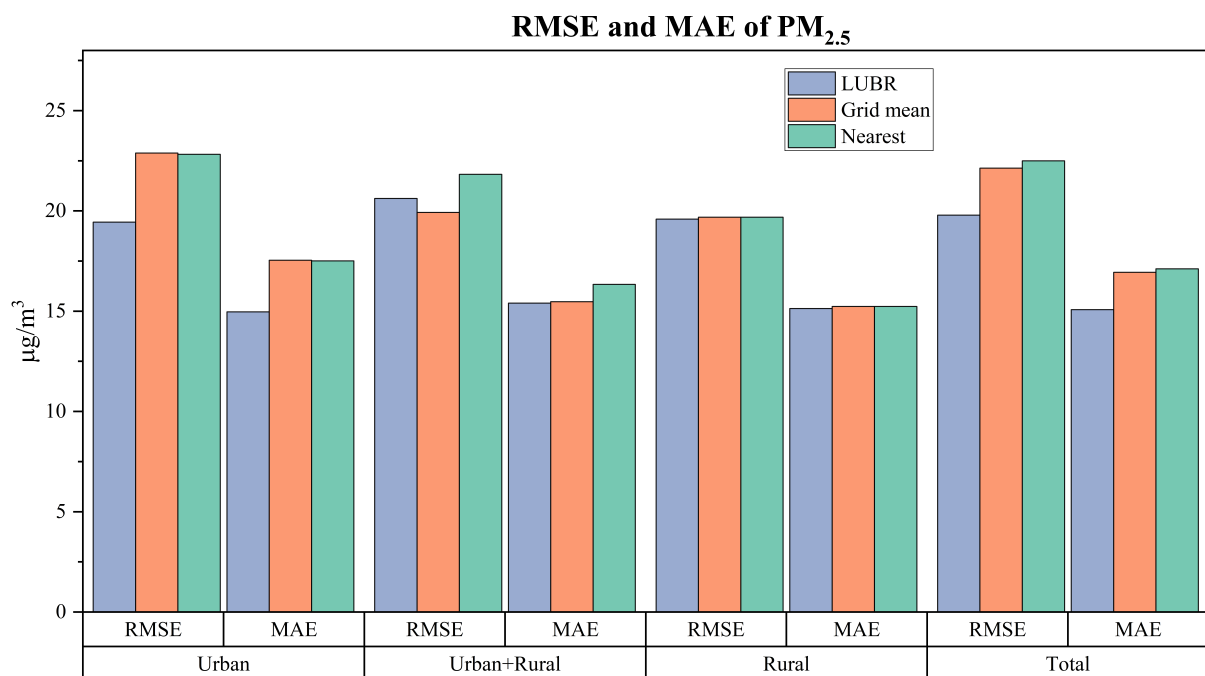
Figure S4. The comprehensive statistical results, encompassing RMSE and MAE, demonstrate the performance of the gridded PM$_{2.5}$ observations compared to the GEOS-Chem simulations. The colors ice blue, rosy red, and cyan represent the LUBR, nearest search, and grid mean methods, respectively. 'Urban,' 'Urban+Rural,' and 'Rural' categorize grids based on the presence of urban and rural sites. 'Urban' includes grids with exclusively urban sites, 'Urban+Rural' includes both urban and rural sites, and 'Rural' comprises grids with only rural sites. 'Total' aggregates results by calculating the average across all three categories.
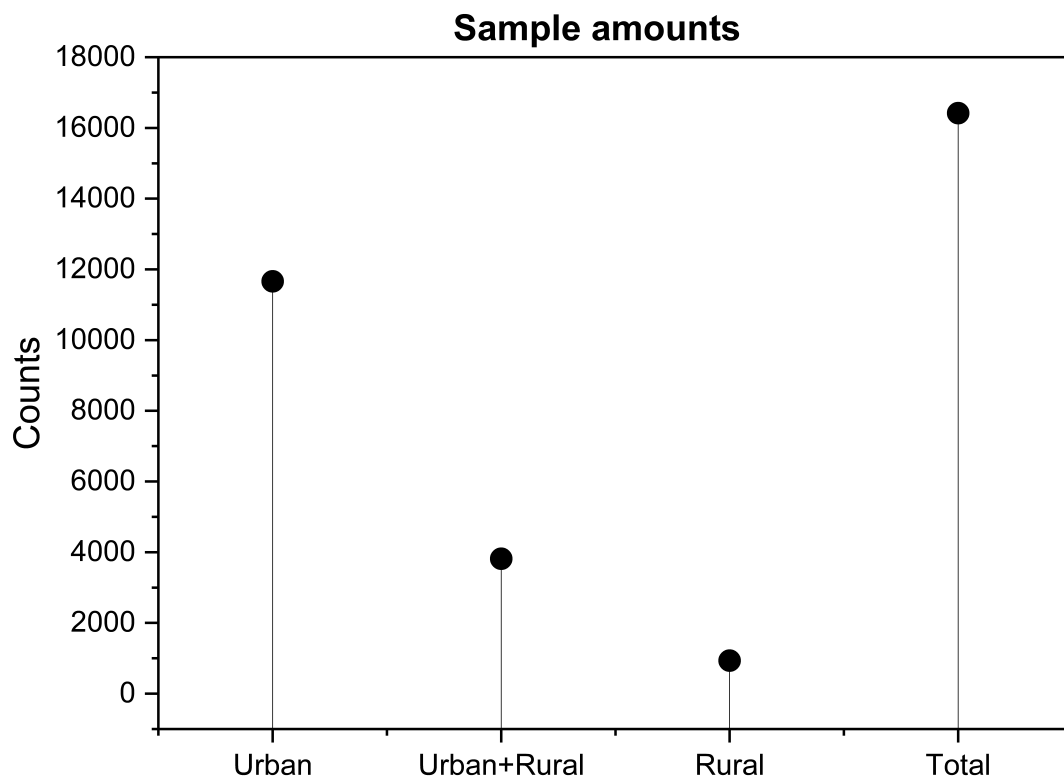
Figure S5. The sample amounts of the categorized grids of 'Urban,' 'Urban+Rural,' 'Rural,' and 'Total'.