Supplement of Atmos. Chem. Phys., 26, 33–58, 2026
https://doi.org/10.5194/acp-26-33-2026-supplement

Atmospheric Chemistry and Physics
Open Access
EGU

Supplement of

NMVOC emission optimization in China through assimilating formaldehyde retrievals from multiple satellite products

Canjie Xu et al.

Correspondence to: Jianbing Jin (jianbing.jin@nuist.edu.cn)

Sect. S1 Emission Uncertainty

Table S1. Uncertainty assumptions for different emission sectors.

	Anthropogenic	Biogenic	Biomass burning
VOC	150%	200%	300%

Following previous studies (Souri et al., 2020), sector-specific prior uncertainties for anthropogenic, biogenic, and biomass burning emissions can be combined into a total uncertainty using a weighted approach.

$$\sigma^2_{\text{total}} = f^2_{\text{anthro}} \cdot \sigma^2_{\text{anthro}} + f^2_{\text{biogenic}} \cdot \sigma^2_{\text{biogenic}} + f^2_{\text{bioburn}} \cdot \sigma^2_{\text{bioburn}} \quad \text{(Eq. S1)}$$

Applying this method to the uncertainty values reported in earlier work, we obtained a total prior uncertainty of 120.22%. Accordingly, we set the standard deviation of the multiplicative factor to 0.4 in this study. To assess the validity of this simplification, we compared the sector-weighted prior emissions (Figure S1 (a)) with the prior emissions uniformly scaled by 120% (Figure S1 (b)). The two results are generally consistent, supporting the reasonableness of adopting a uniform total uncertainty in this study. Sector-specific inversion will be considered in future work.

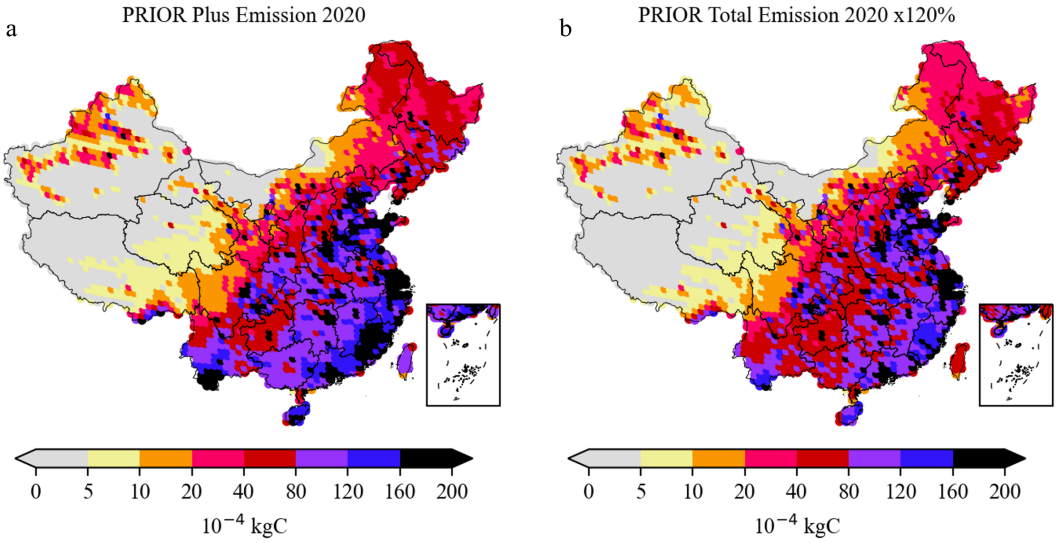

Figure S1. Comparison of prior NMVOC emissions estimated by combining sector-specific uncertainties (a) and by uniformly scaling the prior total emissions by 120% (b) in 2020. The two distributions are generally consistent, supporting the applicability of the total uncertainty assumption used in this study.

Sect. S2 Minimization of the Cost Function in 4DEnVar

The minimization of the cost function follows the 4DEnVar processes. An ensemble of emission inventory is generated randomly using the prior emission vector f and the assumed emission error covariance $\mathbf{B}$:

$$[f_1, \ldots, f_N] \quad \text{(Eq. S2)}$$

An ensemble of GEOS-Chem model simulations is then forward run with the ensemble emission inventories in parallel:

$$[\mathcal{M}(f_1),\ldots,\mathcal{M}(f_N)] \tag{Eq. S3}$$

Denote the emission ensemble perturbation matrix by:

$$F' = \frac{1}{\sqrt{N-1}}[f_1-\bar{f},\ldots,f_N-\bar{f}] \tag{Eq. S4}$$

and the mean of ensemble simulation by:

$$\mathcal{M}(\bar{f}) = \frac{1}{N}\sum_{i=1}^{N}\mathcal{M}(f_i) \tag{Eq. S5}$$

where $\bar{f}$ is the mean of the ensemble emission inventories. In the 4DEnVar assimilation algorithm, the optimal emission f is defined as a weighted sum of the columns of the perturbation matrix F' using weights from a control variable vector w:

$$f = \bar{f} + F'w \tag{Eq. S6}$$

The cost function could then be reformulated as:

$$\mathcal{J}(w) = \frac{1}{2}w^T w + \frac{1}{2}\left\{H\mathcal{M}'w + H\mathcal{M}(\bar{f}) - y\right\}^T \mathbf{O}^{-1}\left\{H\mathcal{M}'w + H\mathcal{M}(\bar{f}) - y\right\} \tag{Eq. S7}$$

where $\mathcal{M}$ is the linearization of the GEOS-Chem formaldehyde simulating model required for cost function minimization, and is approximated by:

$$\mathcal{M}'F' \approx \frac{1}{\sqrt{N}}\left[\mathcal{M}(f_1)-\mathcal{M}(\bar{f}),\ldots,\mathcal{M}(f_N)-\mathcal{M}(\bar{f})\right] \tag{Eq. S8}$$

With the uncertainty in emission transferred into the observation space, the minimum of the cost function in Equation Eq. S7 could then be directly calculated, and the posterior emission f subsequently updated.

$$m_z = \frac{M_z^m - B_z}{M^m - B} \tag{Eq. S9}$$

Here M_z^m represents the modeled concentration of formaldehyde at altitude z, B_z is the background concentration of formaldehyde at the same altitude, M^m represents the total modeled concentration of formaldehyde in the atmosphere, and B is the total background concentration.

$$A_z^a = \frac{1}{N}\frac{\hat{X}^a - B}{\hat{X}^l - B} \tag{Eq. S10}$$

Here X_z^a represents the a priori (or assumed) concentration of formaldehyde at altitude z, B_z is again the background concentration at the same altitude, $\hat{X}^a$ is the total a priori concentration, and N is a normalization factor ensuring the matrix A_z^a sums correctly to account for all altitudes.

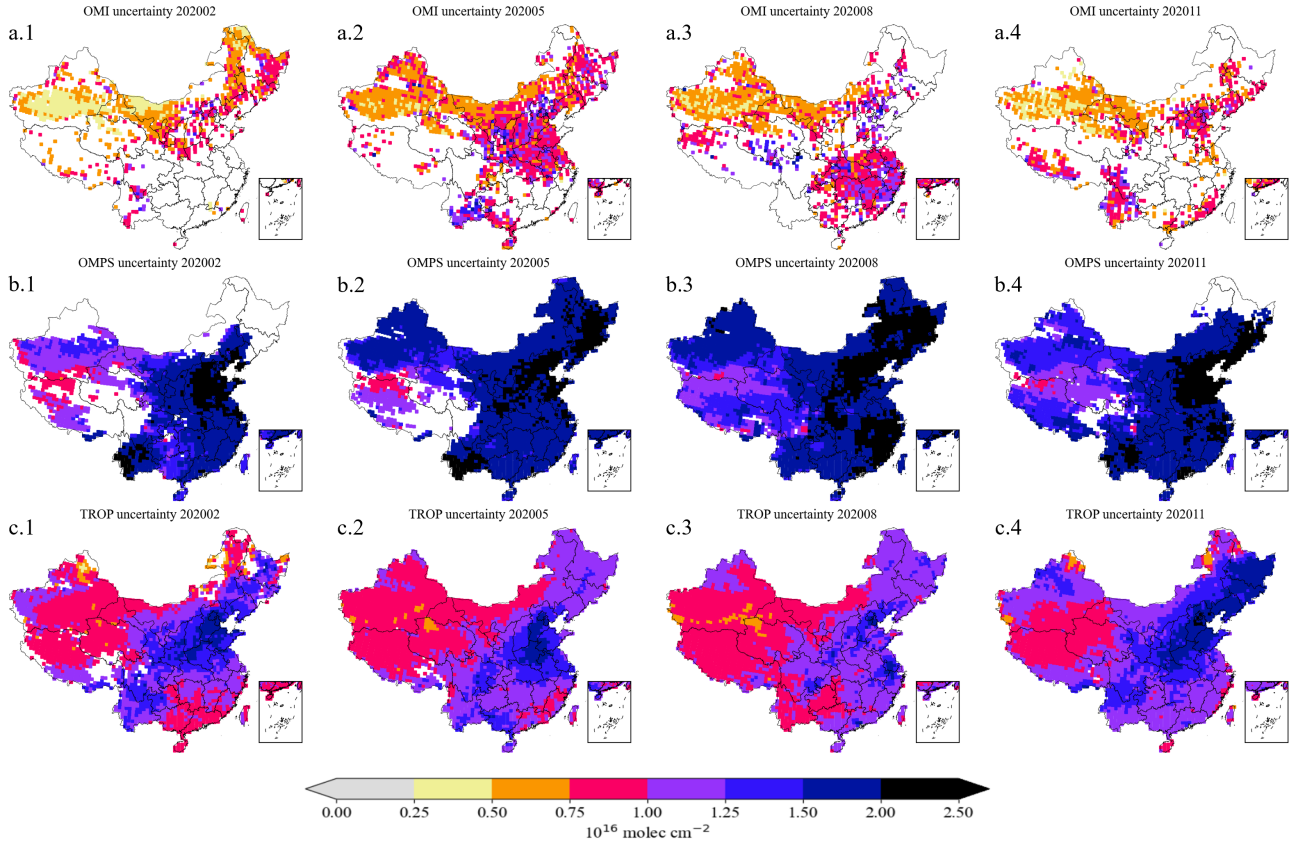

Figure S2. Uncertainties of formaldehyde columns from OMI (a.1-a.4), OMPS (b.1-b.4), and TROPOMI (c.1-c.4) in February, May, August, and November 2020.

References

Souri, A. H., Nowlan, C. R., González Abad, G., Zhu, L., Blake, D. R., Fried, A., Weinheimer, A. J., Wisthaler, A., Woo, J.-H., Zhang, Q., Chan Miller, C. E., Liu, X., and Chance, K.: An inversion of NO x and non-methane volatile organic compound (NMVOC) emissions using satellite observations during the KORUS-AQ campaign and implications for surface ozone over East Asia, Atmospheric Chemistry and Physics, 20, 9837–9854, 2020.

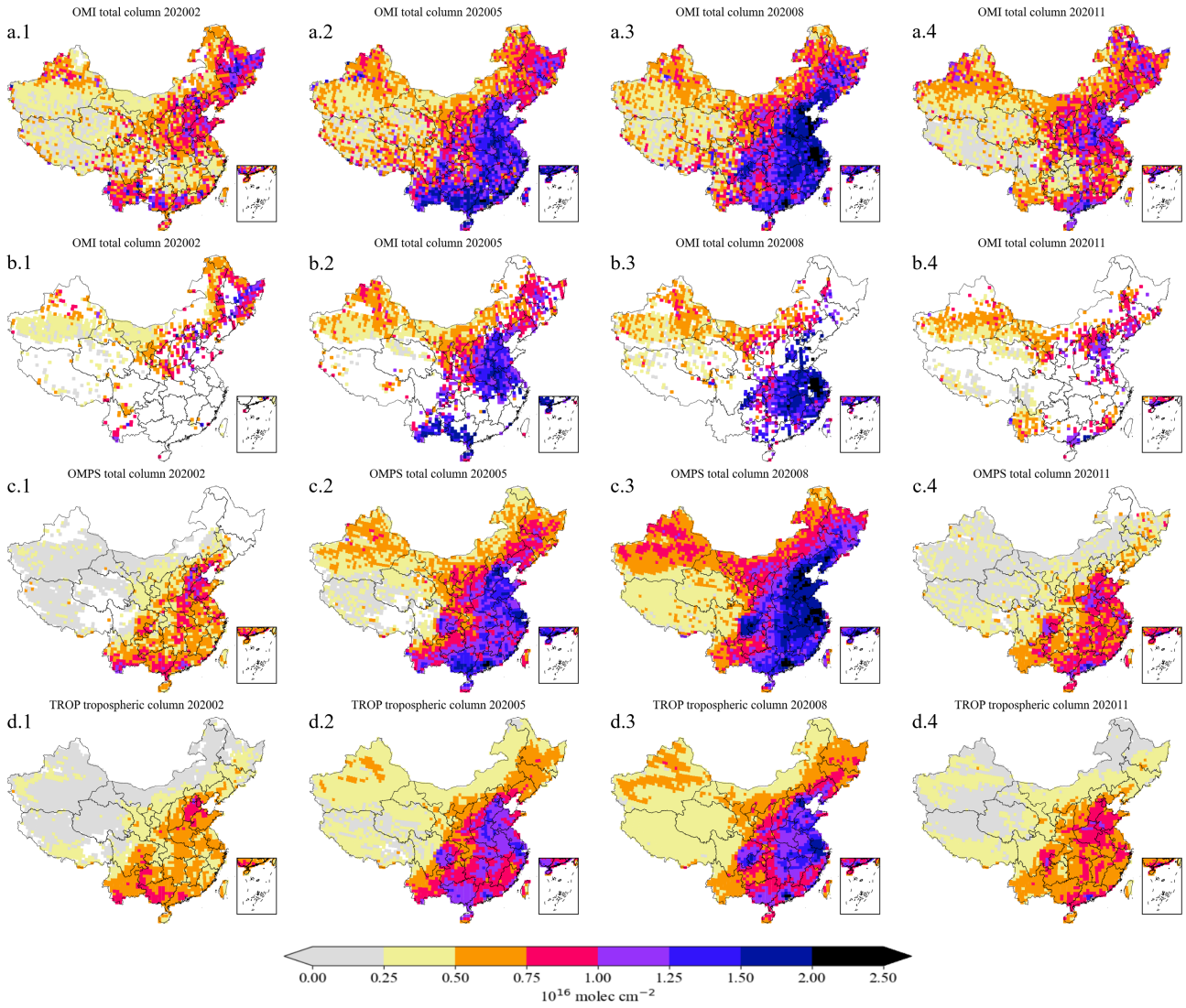

Figure S3. Comparison of monthly mean formaldehyde columns in February, May, August, and November 2020 after applying different data filtering thresholds. Panels (a.1-a.4), (c.1-c.4), and (d.1-d.4) show OMI, OMPS, and TROPOMI results, respectively, after removing grid cells with fewer than 10 observations. Panels (b.1-b.4) show OMI results after removing grid cells with fewer than 50 observations.

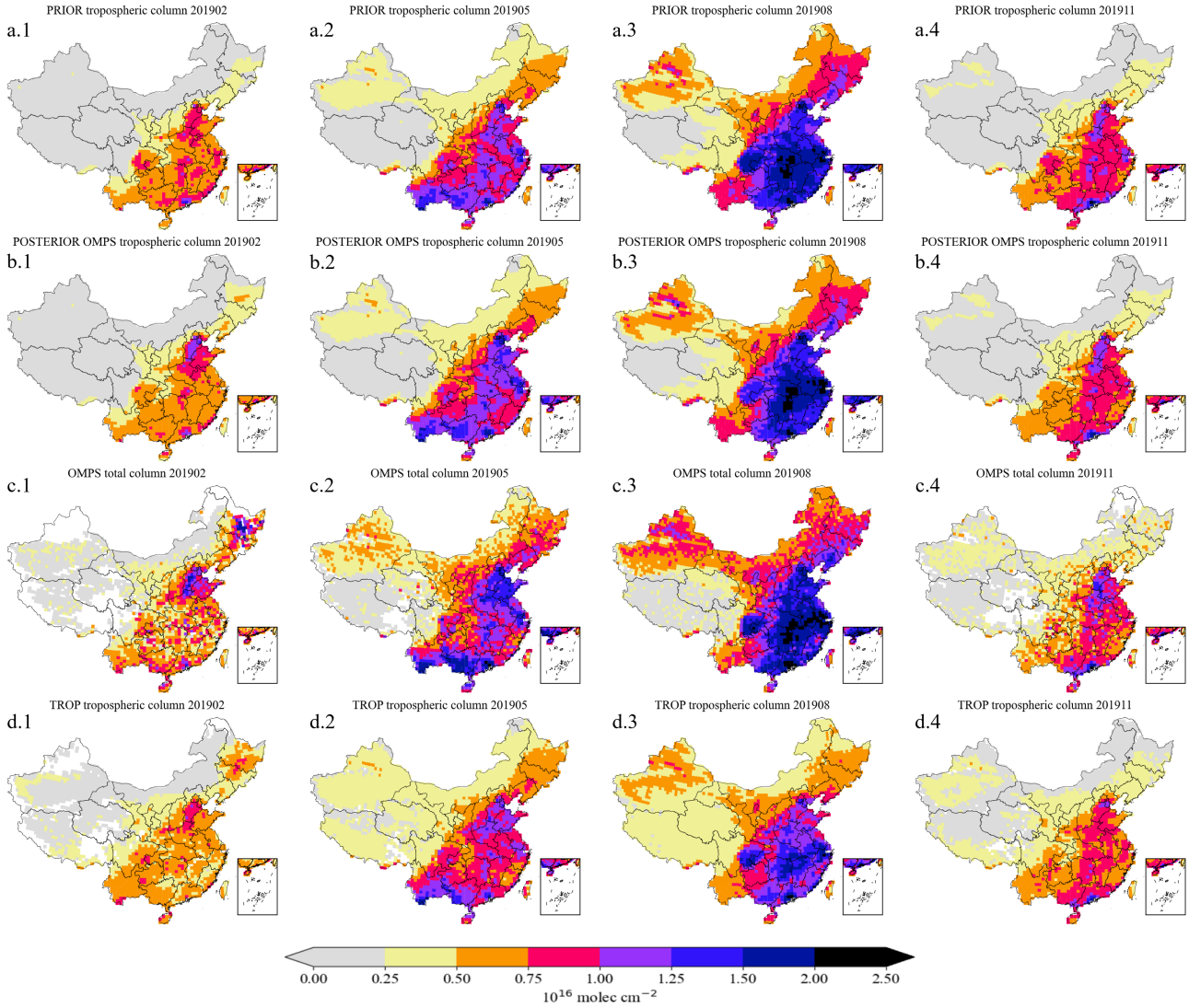

Figure S4. Spatial distributions of formaldehyde columns from GEOS-Chem model-simulated prior tropospheric columns (a) and posterior tropospheric columns constrained by OMPS assimilation (b), satellite observations of OMPS total columns (c), and satellite observations of TROPOMI tropospheric columns (d), both reprocessed to be consistent with the GEOS-Chem vertical profile. Panels (a.1–d.1), (a.2–d.2), (a.3–d.3), and (a.4–d.4) show February, May, August, and November of 2019, respectively.

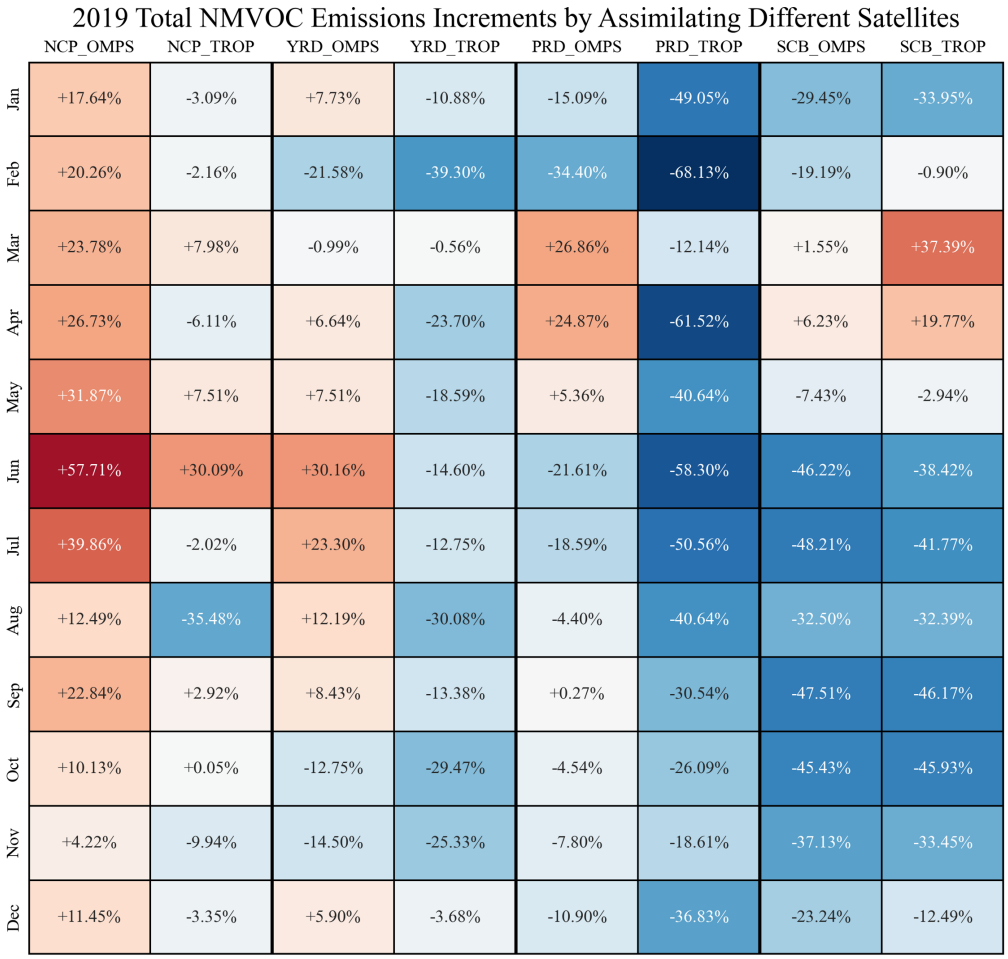

Figure S5. Monthly increments in total NMVOC emissions between the posterior and prior simulations derived from the assimilation of OMPS and TROPOMI formaldehyde observations over four key regions of China: the North China Plain, Yangtze River Delta, Pearl River Delta, and Sichuan Basin in 2019. Positive values indicate an increase in posterior emissions relative to the prior, while negative values indicate a decrease.

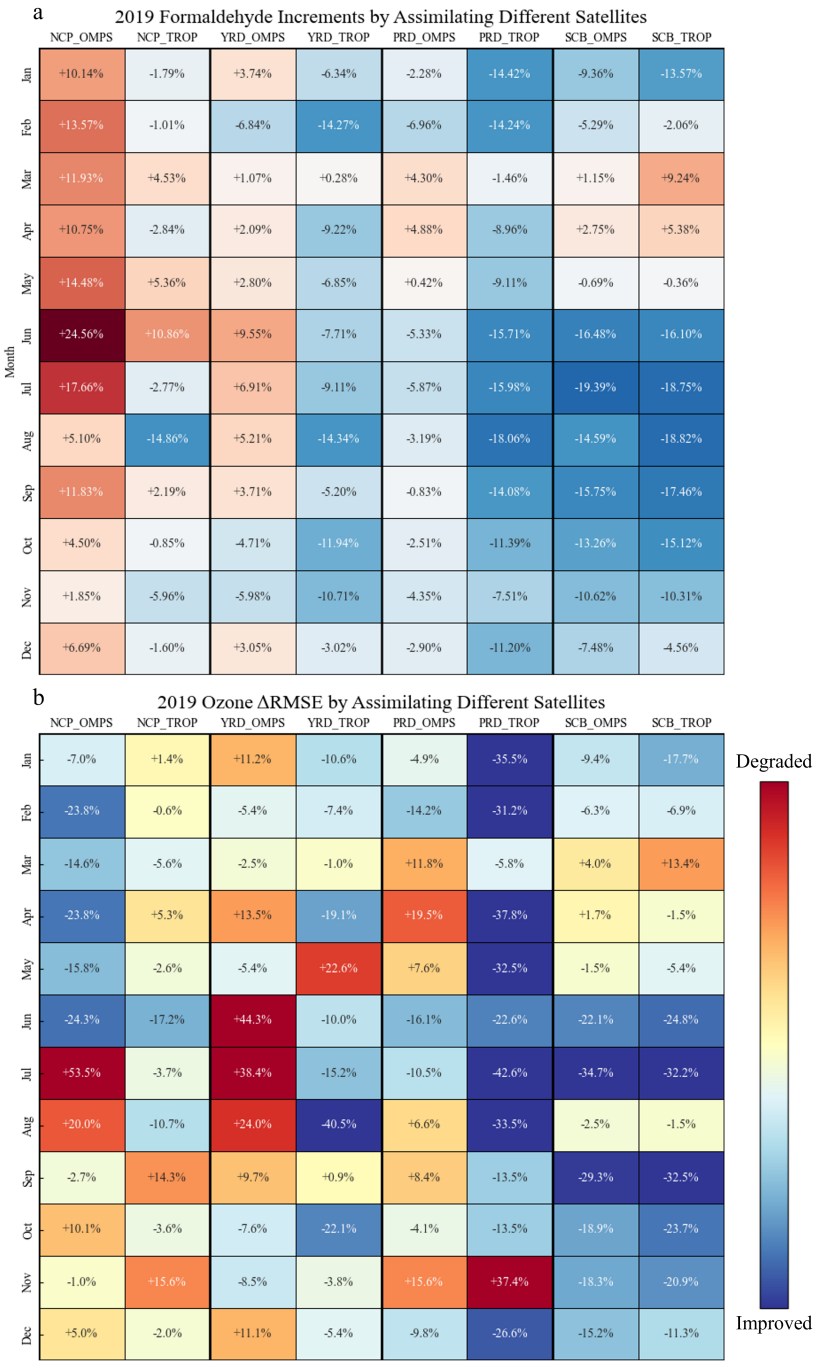

Figure S6. Monthly increments in (a) formaldehyde columns between posterior and prior simulations and (b) the relative changes in MDA8 ozone RMSE (ΔRMSE) after assimilating OMPS and TROPOMI observations in 2019. Results are shown for the North China Plain, Yangtze River Delta, Pearl River Delta, and Sichuan Basin. Positive values indicate an increase relative to the prior, while negative values indicate a decrease.

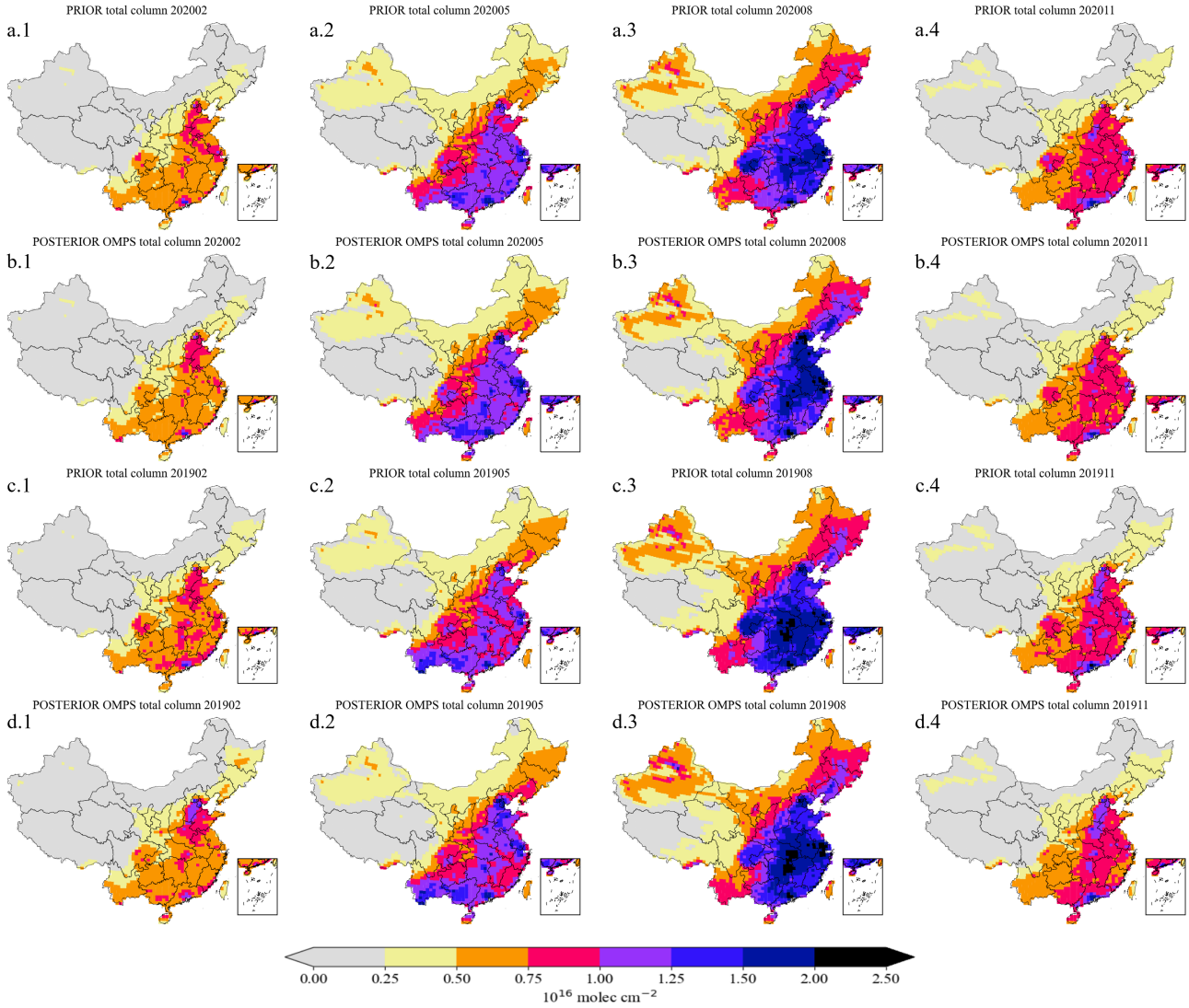

Figure S7. Spatial distributions of the formaldehyde total columns from GEOS-Chem model-simulated prior (a,c) and posterior (b,d) results in February (a.1-d.1), May (a.2-d.2), August (a.3-d.3), November (a.4-d.4) for 2020/2019.