Authors' Response to Reviews of

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

Canjie Xu[1], Jianbing Jin*[1], Ke Li[1], Yinfei Qi[2], Ji Xia[1], Hai Xiang Lin[3,4], Hong Liao[1]
Atmospheric Chemistry and Physics, 10.5194/egusphere-2025-140

EC: ***Editor's Comments***, **RC:** ***Reviewers' Comment***, AR: Authors' Response, □ Manuscript Text

1. Overview

Response to Editor and Referee: We sincerely thank the Editor for the patience, and exceptionally insightful guidance throughout the entire revision process. We are particularly grateful for the meticulous attention to detail in identifying numerous subtle issues, and especially for the critical and decisive comments on the treatment of a priori profile as well as the mathematical equivalence between AMF recalculation and averaging-kernel application. These suggestions have profoundly improved the scientific rigour and consistency of our work, and we deeply admire the Editor's expertise and professionalism. We also extend our gratitude to referee for the thorough and constructive review, as well as for the final valuable comment on the literature review, which led to a much-needed and complete reorganisation of the Introduction. Thank you once again for the outstanding editorial handling and for guiding this work with such dedication.

2. Major comments

EC: ***1) Literature review in the introduction. As pointed out by the reviewer, the literature review in the introduction reads like a somewhat random collection of references related to VOC measurements and inversions. I suggest rewriting this section following a simple structure:***
• An overview of satellite HCHO measurements and inversions, sorted by time and methodological improvements
• A brief summary of other satellite measured VOC constraints (glyoxal, isoprene)
• A summary of the key findings of previous studies over the region of interest (China)
Please also refer to the suggestions and comments made by the other reviewer.

AR: We thank the reviewer for this helpful suggestion. We have completely rewritten the literature review in the Introduction with a clear, logical structure as recommended:
(1) a chronological overview of the development of satellite formaldehyde instruments and observations;
(2) the evolution of top-down approaches using formaldehyde observations to constrain NMVOC emissions, highlighting key methodological advances;
(3) a focused summary of previous top-down NMVOC emission studies over China based on formaldehyde assimilation, clearly identifying the gaps that motivate the innovations and improvements in the present work.
These revisions have greatly improved the coherence, logical flow, and relevance of the literature review.

Text in manuscript

1 Introduction

...

Remote sensing observations of these compounds typically rely on spectral channels in the ultraviolet-visible (UV-Vis) range, with their primary absorption features occurring between 330 and 460 nm (Platt, 1979; Lerot et al., 2010; De Smedt et al., 2012).

~~Compared to glyoxal, satellite products for formaldehyde are better established.. Satellite observations of glyoxal began later with initial identifications made using Aura Ozone Monitoring Instrument (OMI), and its retrieval is more challenging than that of formaldehyde (Kurosu et al., 2005; Kurosu et al., 2014). Recent years have seen further advancements in satellite observational instruments and algorithms for formaldehyde (Abad, 2022; Abad, 2017), leading to new satellite observation products with improvements in both accuracy and resolution. In contrast, while the glyoxal retrieval algorithm has been updated for OMI product (Alvarado et al., 2014), its satellite products continue to face substantial uncertainties. Notably, the latest version of Sentinel-5 Precursor Tropospheric Monitoring Instrument (TROPOMI) glyoxal products is no longer publicly available on the official website.~~

~~Formaldehyde measurements from instruments such as the Global Ozone Monitoring Experiment-2 (GOME-2) (De Smedt et al., 2012), OMI (González Abad et al., 2015), Ozone Mapping and Profiler Suite (OMPS) (Li et al., 2015) and TROPOMI (De Smedt et al., 2018) have been used widely for estimating NMVOC emissions through data assimilation. The core of the methodology is to calculate the most likely NMVOC emissions given the formaldehyde observations and the prior information. For instance, Fu et al. (2007) used six years of continuous satellite measurements of formaldehyde columns from GOME (1996-2001) to improve regional emission estimates of reactive NMVOCs, including isoprene, olefins, formaldehyde, and xylene, for East Asia and South Asia. Similarly, formaldehyde data from the GOME-2A satellite were used to constrain NMVOC emissions in India for 2009 (Chaliyakunnel et al., 2019). Souri et al. (2020) used observations from OMPS satellites during the KORUS-AQ campaign to estimate NO_x and NMVOC emissions in East Asia from May to June 2016. Kaiser et al. (2018) also utilized high-resolution formaldehyde retrieval data from OMI instrument to quantify isoprene emissions at the ecosystem scale in the southeastern United States during August-September 2013. Those promising results have demonstrated that formaldehyde measurements could be utilized to optimize the existing NMVOC emissions that were established in a bottom-up manner.~~

Satellite remote sensing of formaldehyde has made substantial progress since the atmospheric formaldehyde abundance was first retrieved in 1997 (Burrows et al., 1999). The earliest retrievals of formaldehyde vertical column densities were based on the Global Ozone Monitoring Experiment (GOME) (Thomas et al., 1998; Chance et al., 2000). Subsequently, the Scanning Imaging Absorption Spectrometer for Atmospheric Chartography (SCIAMACHY) served as an important transitional instrument between GOME and GOME-2, offering significantly improved spatial resolution compared to GOME (De Smedt et al., 2008). In 2004, the launch of NASA's Aura satellite carrying the Ozone Monitoring Instrument (OMI) provided high signal-to-noise-ratio ultraviolet–visible (UV–Vis) spectra that greatly advanced trace-gas retrieval studies (De Smedt et al., 2015; González Abad et al., 2015). Approximately four years after GOME effectively ceased operational observations, its successor, GOME-2, began routine operations in 2007 and started delivering formaldehyde data (De Smedt et al., 2012). In recent years, high-resolution formaldehyde observations have continued to emerge, including those from the Ozone Mapping and Profiler Suite (OMPS) onboard the Suomi National Polar-orbiting Partnership

(Suomi NPP) and NOAA-20 satellites (Li et al., 2015; González Abad et al., 2015, 2016; Nowlan et al., 2023), as well as from the Tropospheric Monitoring Instrument (TROPOMI) aboard the Sentinel-5 Precursor (Sentinel-5P) launched in 2017. TROPOMI's exceptional spatial resolution and near-daily global coverage have marked a new era in satellite formaldehyde monitoring (De Smedt et al., 2018, 2021). Furthermore, geostationary satellites now provide formaldehyde observations with high temporal resolution, including the Geostationary Environment Monitoring Spectrometer (GEMS) over East Asia (Kwon et al., 2019; Kim et al., 2020), the Tropospheric Emissions: Monitoring of Pollution (TEMPO) instrument over North America (Chance et al., 2019), and Sentinel-4, successfully launched on 1 July 2025, which is conducting geostationary formaldehyde observations over Europe (Gulde et al., 2017).

Glyoxal retrieval product from satellite platform began relatively late, with the first global differential optical absorption spectroscopy (DOAS) retrievals reported by Wittrock et al. (2006) using SCIAMACHY, followed by their application to constrain NMVOC emissions by Stavrakou et al. (2009a). Because glyoxal is retrieved in a longer wavelength range (∼435–460 nm) than formaldehyde (∼330–360 nm), it exhibits markedly lower sensitivity to molecular scattering, which in turn reduces the sensitivity of the measurement to the lower troposphere (Palmer et al., 2001; Chan Miller et al., 2014). Moreover, glyoxal optical depths are very weak (order of 10^{-4}–10^{-3}), rendering the retrieval highly susceptible to fitting residuals from stronger absorbers, uncertainties in absolute radiometric calibration, and spectral features in surface reflectivity (Sinreich et al., 2013; Alvarado et al., 2014). For instruments with comparatively modest spectral resolution and signal-to-noise ratios, such as OMI, these interference effects are further amplified, leading to larger retrieval uncertainties for glyoxal columns than for formaldehyde (Chan Miller et al., 2014; Cao et al., 2018). Consequently, glyoxal satellite observations remain considerably less suitable than formaldehyde for high-spatiotemporal-resolution assimilation studies. Beyond glyoxal and formaldehyde, retrievals of other VOCs are also progressing, as exemplified by Fu et al. (2019) and Wells et al. (2020, 2022), who derived isoprene columns from Cross-track Infrared Sounder (CrIS) observations, representing an important step toward next-generation satellite constraints on volatile organic compounds.

~~Compared with formaldehyde retrieval, satellite remote sensing of glyoxal began much later and remains considerably more challenging. Shim et al. (2005) assimilated formaldehyde observations from the GOME using a global Bayesian inversion to constrain isoprene emissions. Although China was included within their East Asia region, the analysis lacked region-specific focus and did not provide detailed characterization of emission patterns over China, and the coarse spatial resolution (4° × 5°) in that study further limited the ability to resolve subregional emission features. Stavrakou et al. (2016) conducted a regional inversion in Eastern China using multi-year satellite formaldehyde data from GOME and OMI to constrain VOC emissions during the post-harvest burning period, and they indicated that the crop burning fluxes of VOCs in June exceeded by a factor of two the combined emissions from other anthropogenic activities in the NCP region from 2005 to 2012. Cao et al. (2018) conducted a relatively systematic satellite-based emission inversion study over China, using a 4DVar method and assimilating OMI and GOME-2A formaldehyde products to estimate monthly NMVOC emissions in 2007, though the spatial resolution (4° × 5°) was still too coarse. Choi et al. (2022) applied a 4DVar system to assimilate TROPOMI formaldehyde over East Asia, demonstrating the capability of high-resolution satellite data to capture regional and seasonal variability in VOC emissions, but the analysis was conducted only for May-June. Beyond China, a number of important studies have advanced top-down VOC inversion methodologies: Palmer et al. (2003) pioneered the use of GOME formaldehyde observations in a Bayesian framework to constrain global isoprene emissions, laying the foundation for subsequent satellite-based VOC studies; Wells et al. (2020, 2022) further advanced~~

~~this field by retrieving isoprene emissions from CrIS measurements and providing high-resolution constraints on VOC oxidation chemistry at the global scale; and Oomen et al. (2024) derived weekly top-down VOC fluxes over Europe from TROPOMI formaldehyde data using the MAGRITTEv1.1 model, providing improved constraints on isoprene, biomass burning, and anthropogenic VOC emissions. Considering the increasingly stringent air pollution control in China (Wu et al., 2024), there is an urgent need for high-resolution top-down NMVOC emission optimization.~~

Top-down approaches, mainly assimilation techniques, with satellite formaldehyde columns have become the primary method for constraining NMVOC emissions. Palmer et al. (2003) pioneered to apply a Bayesian inversion framework with GOME formaldehyde observations for constraining isoprene emissions over North America. The approach was subsequently extended to global and European domains by Shim et al. (2005) and Dufour et al. (2009), respectively. With the availability of OMI and GOME-2 formaldehyde products, inversion algorithms were further refined. Stavrakou et al. (2009b) first introduced an adjoint-based inversion to optimize biogenic emissions and, in a companion study the same year, revealed substantial underestimation of continental glyoxal sources worldwide (Stavrakou et al., 2009a). Concurrently, Millet et al. (2008) used OMI formaldehyde and identified the underestimated isoprene emissions over the north-central United States, while Zhu et al. (2014) reported that anthropogenic emissions of highly reactive VOCs (HRVOCs) in the Houston area were underestimated by a factor of 4.8 ± 2.7 compared to the US Environmental Protection Agency inventory. Formaldehyde product with much higher spatial resolution were then available since the launch of TROPOMI and OMPS, and made the city-scale emission optimizations possible (González Abad et al., 2016; De Smedt et al., 2018, 2021). In recent years, studies leveraging these new-generation instruments have proliferated. Choi et al. (2022) assimilated OMPS and OMI observations into an updated 4DVar system for East Asia during May–June 2016. Their inversion revealed a 47% increase in VOC emissions across Northeast Asia relative to the prior inventory, indicating that isoprene emissions over South Korea and anthropogenic NMVOC emissions over eastern China were underestimated in the bottom-up inventory. Oomen et al. (2024) used weekly-averaged TROPOMI formaldehyde observations from 2018–2021 with the MAGRITTEv1.1 adjoint model to derive top-down biogenic, pyrogenic, and anthropogenic VOC fluxes over Europe, substantially correcting previous underestimates of isoprene emissions. Feng et al. (2024) applied an Ensemble Kalman Filter (EnKF) approach to optimize August 2022 NMVOC emissions over China, revealing overestimation of biogenic emissions during an extreme heatwave and demonstrating consequential impacts on summertime ozone simulations. The advent of geostationary satellites (e.g., GEMS, TEMPO) with high-frequency observations has enabled the incorporation of diurnal cycle information into algorithm frameworks, making daily-scale top-down emission optimization feasible (Kwon et al., 2019, 2021). Meanwhile, multi-species constraint is gaining traction; Opacka et al. (2025) developed a novel inversion technique that simultaneously optimizes monthly-mean VOCs and NOx emissions from 2019 TROPOMI observations, uncovering severe underestimation of both NOx and VOCs in prior inventories over Africa.

Although substantial progress has been made globally in satellite-based top-down constraints on NMVOC emissions, high-resolution top-down emission optimization studies specifically over China remain relatively scarce in recent years. Shim et al. (2005) first used GOME formaldehyde observations in a global Bayesian inversion framework to constrain isoprene emissions. While their domain encompassed East Asia including China, the study lacked a dedicated focus on China and was limited by coarse model resolution ($4^\circ \times 5^\circ$). Stavrakou et al. (2016) performed a regional inversion over eastern China using multi-year GOME and OMI formaldehyde columns, revealing that post-harvest agricultural burning in June contributed more than twice the VOC emissions of all other anthropogenic

sources combined over the North China Plain during 2005–2012. Cao et al. (2018) conducted one of the most systematic satellite-constrained inversions for China to date, applying a 4DVar assimilation of OMI and GOME-2A formaldehyde products to estimate monthly NMVOC emissions in 2007, yet the analysis was still constrained by the same coarse 4° × 5° resolution. Choi et al. (2022) assimilated OMPS and OMI formaldehyde columns into a regional 4DVar system over East Asia but only for May–June; similarly, the top-down optimization of Chinese NMVOC emissions by Feng et al. (2024) was limited to a single month (August 2022). Given the increasingly stringent air pollution control policies in China (Wu et al., 2024), there is an urgent need for high spatial- and temporal-resolution top-down NMVOC emission optimization to support more accurate air quality forecasting and effective regulatory strategies. In terms of the observation sources for assimilation, the OMI formaldehyde product remains one of the most widely used datasets in related studies to date. However, it has been affected globally by the persistent row anomaly since 2007 Kroon et al. (2011); Zhu et al. (2017), which degrades data quality in certain across-track positions and may reduce assimilation accuracy, particularly in high-resolution configurations. Although rigorous quality filtering and row anomaly flagging can mitigate this problem, the number of valid grid cells remaining after such screening is often severely limited, rendering OMI data insufficient for nationwide high-resolution emission inversion. In contrast, newer-generation instruments such as TROPOMI and OMPS provide formaldehyde products that are unaffected by the row anomaly, offer significantly higher spatial resolution and better data coverage (Section 2.3.3), and are therefore considerably more suitable for high-resolution top-down optimization of NMVOC emissions over China.

~~In terms of the assimilation observation sources, OMI product now remains as one of the most widely used formaldehyde products in related studies. However, the retrievals over China are reported to suffer from the so-called "row anomaly" issue (González Abad et al., 2015), which may reduce the assimilation accuracy especially in the high-resolution configuration. To ensure the quality of OMI data, rigorous filtering is usually required; however, the number of valid grid cells remaining after such screening is severely limited, making it insufficient for high-resolution assimilation. In contrast, the TROPOMI and OMPS formaldehyde products do not suffer from this issue and are therefore more suitable for high-resolution emission optimization at the national scale. Subsequently, the monthly NMVOC emission optimization in China is conducted. This is achieved by independently assimilating formaldehyde observations either from OMPS or from TROPOMI, based on the emission inversion system that couples the four-dimensional ensemble variational (4DEnVar) data assimilation algorithm and GEOS-Chem model. The effectiveness of this emission inversion system has been evaluated in our recent studies (Jin et al., 2023; Jin et al., 2025). In this study, we focus on the year 2020 for the main analysis, while results for 2019 are also presented in the Supplementary Information to provide additional context and support.~~

In this study, we conduct monthly optimization of anthropogenic NMVOC emissions over China at 0.5° latitude × 0.625° longitude horizontal resolution. It is achieved based on an emission inversion system that couples the four-dimensional ensemble variational (4DEnVar) data assimilation algorithm with the nested version of the GEOS-Chem model. The effectiveness of this emission inversion system has been evaluated in our recent studies of ammonia (Jin et al., 2023; Xia et al., 2025). Two independent assimilation experiments are performed: one assimilating OMPS total formaldehyde columns and the other assimilating TROPOMI tropospheric formaldehyde columns. In both cases, the satellite retrievals have been harmonized with the model by replacing the original shape profiles with GEOS-Chem profiles before assimilation. We focus on the year 2020 for the main analysis, while results for 2019 are also presented in the Supplementary Information to provide additional context and support. This paper

is organized as follows: Section 2 describes the dataset and methodology, focusing on GEOS-Chem model, input emission sources (anthropogenic, biogenic, and biomass burning), and the satellite and ground-based observations utilized. Section 3 provides an analysis of the assimilation results, including the estimation of posterior NMVOC emissions and the validation of both formaldehyde columns and ground-level ozone simulations. Section 4 summarizes the key findings and concludes the study.

EC: ***2) Tropospheric and total columns: In the manuscript, you distinguish at many places between the total HCHO columns of OMPS and the tropospheric HCHO columns from TROPOMI. However, there is no difference: Both retrievals work in a very similar way, and there is no removal of a possible stratospheric component in the TROPOMI HCHO retrieval. Any such stratospheric HCHO column would also be very small compared to the uncertainties of the retrievals. The discussion and differentiation provided in the manuscript are confusing at best.***

AR: We thank the reviewer for this important and accurate comment. We fully agree that the stratospheric formaldehyde contribution is negligible compared to tropospheric signals and retrieval uncertainties. To faithfully represent the original satellite data products, we retain the distinction in the Methods section (OMI/OMPS provide total columns, TROPOMI provides tropospheric columns). We have added an explicit statement, supported by a priori vertical profiles and relevant references, to clearly show that the total and tropospheric columns are essentially identical over the study region. Accordingly, in the Results and Discussion sections, we no longer differentiate between the two and uniformly use the term "formaldehyde vertical columns" to avoid confusion caused by alternating terminology. This clarification eliminates potential misunderstanding while preserving the precise description of the original datasets.

Text in manuscript

2.1 Model simulation

...

Since the satellite overpasses China mainly between 12:00 and 14:00 local time, the model outputs within this time window are sampled to calculate the formaldehyde columns for fair comparison with the satellite observations. ~~Our GEOS-Chem model outputs both total and tropospheric formaldehyde column concentrations, enabling comparison with OMPS total column data and TROPOMI formaldehyde column measurements as will be introduced later.~~ After this temporal collocation and post-processing, samples of the formaldehyde tropospheric column simulation are presented in Figure 1 (a).

...

2.3.2 Sentinel-5P TROPOMI

...

This criterion ensures the exclusion of error flags and requires that the cloud radiance fraction at 340 nm is below 0.5, the solar zenith angle (SZA) does not exceed 70°, the surface albedo is below 0.2, no snow or ice warning is present, and the air mass factor (AMF) is larger than 0.1. ~~The TROPOMI product provides vertical information on 34 layers, but the retrieval is primarily sensitive to the troposphere and thus reports the formaldehyde tropospheric column.~~The operational TROPOMI HCHO Level-2 product provides tropospheric vertical columns together with averaging kernels and a priori profiles defined on 34 vertical layers (from the surface to ∼0.1 hPa). Because stratospheric formaldehyde is negligible (De Smedt et al., 2018, 2021), we directly use the reported tropospheric columns in this

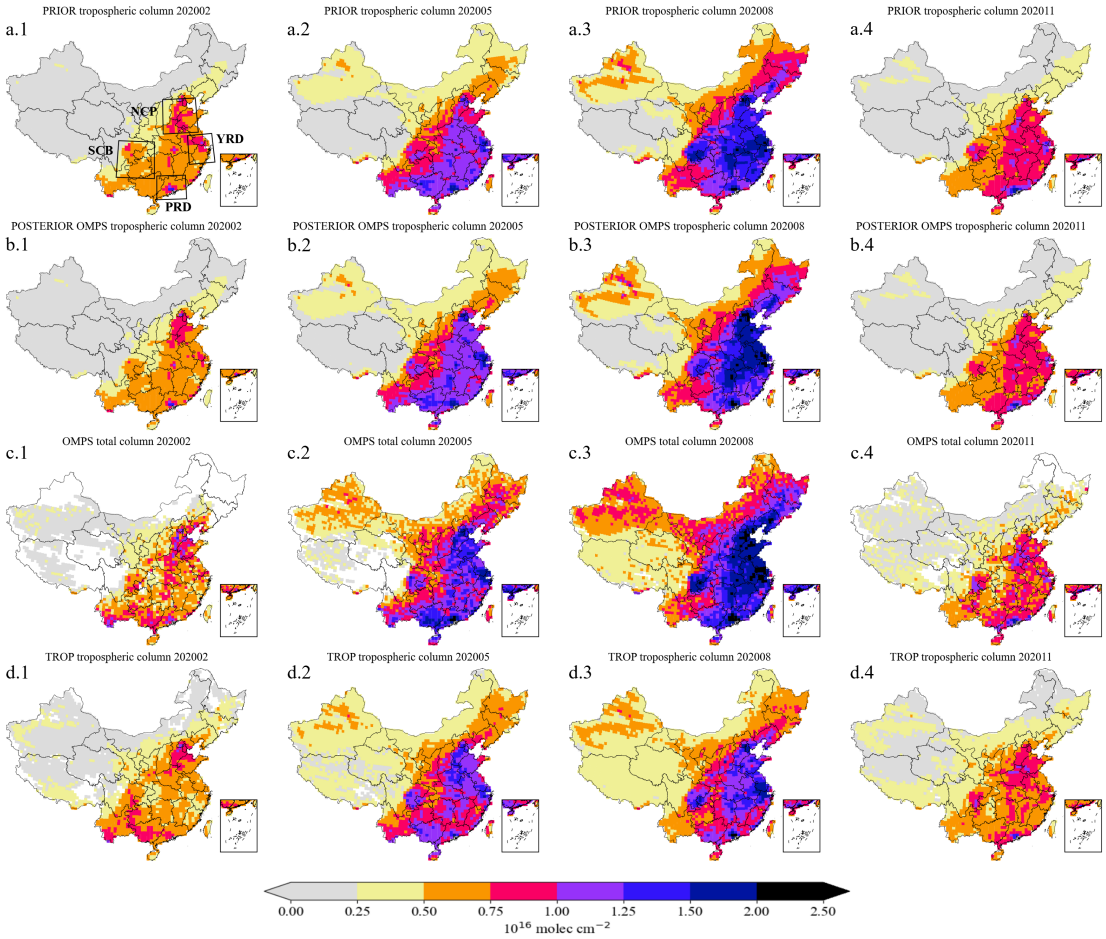

Figure 1. Spatial distributions of formaldehyde columns from GEOS-Chem model-simulated prior tropospheric columns (a) and posterior tropospheric columns constrained by OMPS assimilation (b), satellite observations of OMPS total columns (c), and satellite observations of TROPOMI tropospheric columns (d),both reprocessed to be consistent with the GEOS-Chem shape profile. ~~in February (a.1-d.1), May (a.2-d.2), August (a.3-d.3), and November (a.4-d.4) of 2020.~~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 2020, respectively.

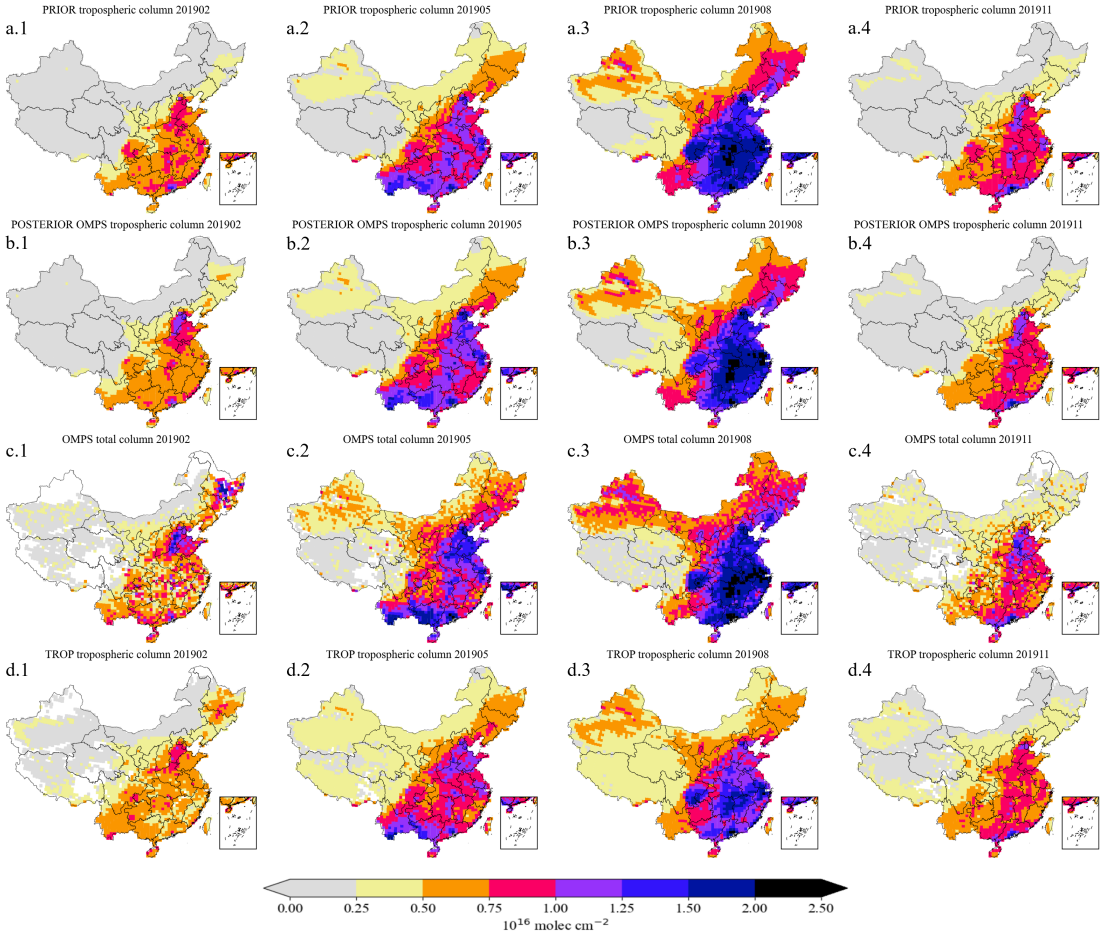

Figure 2. Spatial distributions of formaldehyde columns from GEOS-Chem model-simulated prior tropospheric columns (a) and posterior tropospheric columns constrained by OMPS assimilation (b), satellite observations of OMPS total columns (c), and satellite observations of TROPOMI tropospheric columns (d),both reprocessed to be consistent with the GEOS-Chem vertical profile. ~~in February (a.1-d.1), May (a.2-d.2), August (a.3-d.3), and November (a.4-d.4) of 2020.~~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.

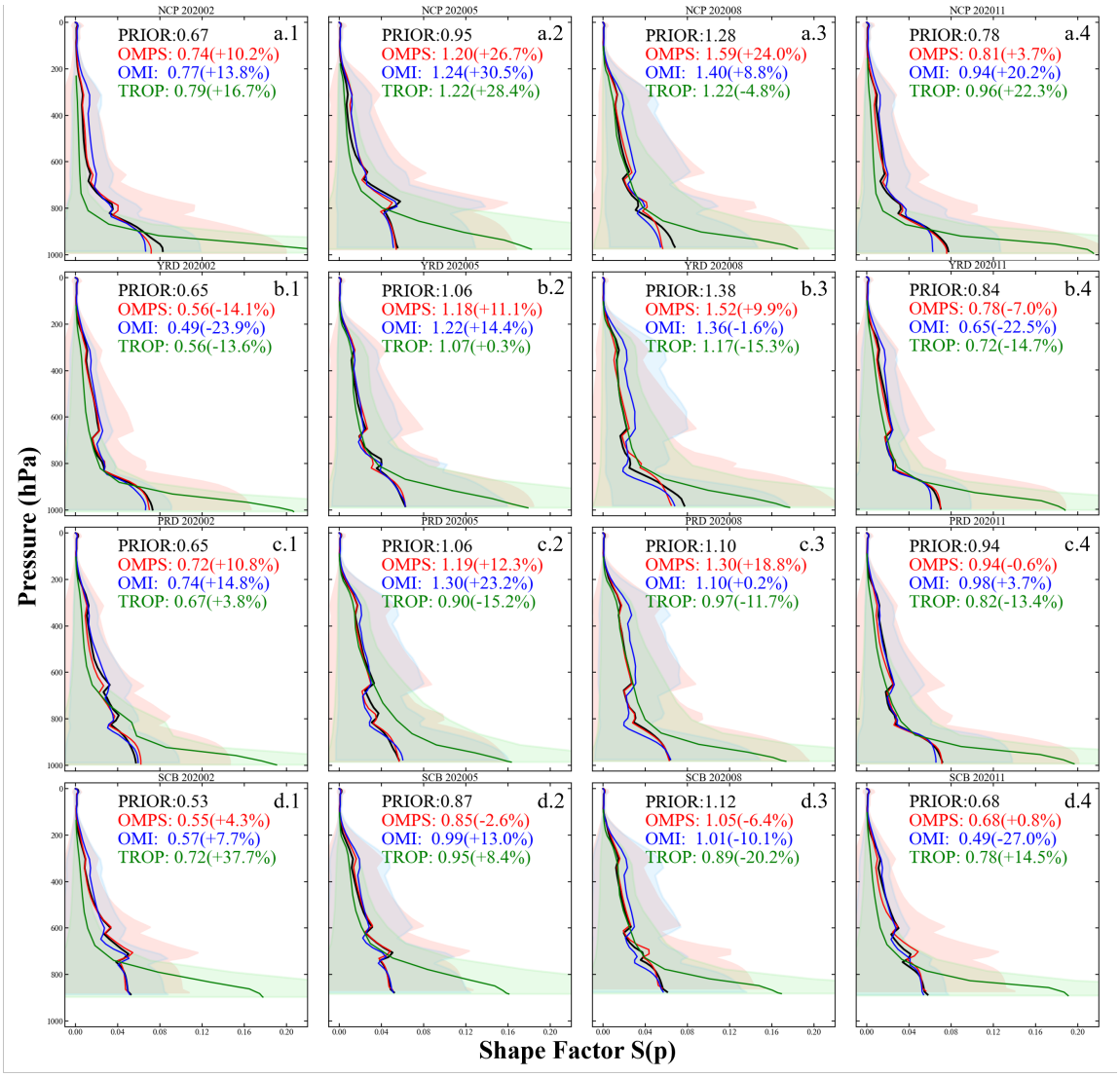

Figure 3. ~~Vertical distributions of the regional mean formaldehyde columns~~Shape factors of regional-mean formaldehyde columns, as derived from the a priori profiles, followed by normalization, from GEOS-Chem model-simulated prior (black) and satellite observations by OMPS (blue), TROPOMI (red), and OMI (green). Panels (a)-(d) correspond to the North China Plain, Yangtze River Delta, Pearl River Delta, and Northeast China, respectively. Sub-panels (a.1-d.1), (a.2-d.2), (a.3-d.3), and (a.4-d.4) represent February, May, August, and November 2020, respectively. Values in parentheses indicate the biases of satellite observations relative to the prior simulation. Shaded areas denote the observational uncertainties.

study without reconstructing total columns.

...

~~The processed TROPOMI retrievals were assimilated as tropospheric columns, which are presented in Figure 1 (d.1-d.4),~~The TROPOMI tropospheric columns were assimilated after application of GEOS-Chem averaging kernels and shape factors to ensure consistency with the model vertical profile. These reprocessed tropospheric columns are shown in Figure 1 (d.1–d.4), with their vertical shape factors shown in Figure 3 (green line) to illustrate the normalized contribution of each pressure layer to the tropospheric columns. We adopted tropospheric rather than total columns because the retrieval product itself provides tropospheric columns. ~~and recalculating total columns would introduce uncertainties.~~

...

3.1 Satellite data evaluation

...

Additionally, Figure 3 clearly shows that stratospheric formaldehyde (> ∼200 hPa) can be largely neglected in both the GEOS-Chem simulation and the OMPS and TROPOMI retrievals (González Abad et al., 2015; De Smedt et al., 2018). Because the stratospheric contribution is negligible and no explicit stratospheric correction is applied in either retrieval, we hereafter use the term "formaldehyde column" without distinguishing between total and tropospheric columns in subsequent discussion.

...

3.3 Formaldehyde columns evaluation

The spatial distributions of formaldehyde columns in February, May, August, and November 2020 are shown in Figure 1. Panels (a.1-a.4) display the prior simulations of ~~tropospheric~~formaldehyde columns, (b.1-b.4) present the posterior simulations of ~~tropospheric~~formaldehyde columns assimilated by OMPS, (c.1-c.4) show the OMPS satellite observations of ~~total~~formaldehyde columns, and (d.1-d.4) illustrate the TROPOMI satellite observations of ~~tropospheric~~formaldehyde columns. In addition, the prior and posterior simulations of ~~total~~formaldehyde columns for 2020 are also provided in the Supplementary Figure S7. ~~As indicated by the vertical profiles in Figure 3, formaldehyde is mainly distributed below the tropopause. Comparisons between the prior and posterior results show that the differences between total and tropospheric columns are relatively small.~~ Regarding the spatial patterns, high formaldehyde ~~values~~columns in February are concentrated in the NCP, YRD, and PRD regions, with the posterior results showing an expanded high-value area in the NCP but a reduced coverage in the YRD. In May, overall ~~concentrations~~formaldehyde columns increase nationwide, with particularly pronounced growth in the NCP and PRD. In August, ~~concentrations~~formaldehyde columns increase in the NCP, YRD, and PRD, while they decrease in the SCB. In November, the changes are modest, but all four regions exhibit reduced ~~concentrations~~formaldehyde columns.

The prior and OMPS-driven posterior simulations of formaldehyde ~~tropospheric~~ columns were compared with the TROPOMI formaldehyde ~~tropospheric~~ columns to evaluate the changes in formaldehyde.

EC: ***3) Column concentrations: As remarked in an earlier round of comments, columns are not concentrations (different units). Please use "columns" throughout the text, supplement, captions, and in the figures (Figs. 3, 7).***

AR: We thank the reviewer for pointing this out again. We have carefully checked the entire manuscript, including

the main text, Supplement, all figures, and replaced every inappropriate use of "formaldehyde column concentrations" with the correct term "formaldehyde columns".

Text in manuscript

2.1 Model simulation

...

Since the satellite overpasses China mainly between 12:00 and 14:00 local time, the model outputs within this time window are sampled to calculate the formaldehyde columns for fair comparison with the satellite observations. ~~Our GEOS-Chem model outputs both total and tropospheric formaldehyde column concentrations, enabling comparison with OMPS total column data and TROPOMI formaldehyde column measurements as will be introduced later.~~ After this temporal collocation and post-processing, samples of the formaldehyde tropospheric column simulation are presented in Figure 1 (a).

...

2.3.1 NOAA-20 OMPS

...

All screened data were then averaged to a spatial resolution of 0.5° latitude × 0.625° longitude on a monthly basis, consistent with the GEOS-Chem model configuration. To make a fair comparison between the observed and simulation formaldehyde column ~~concentration~~ in the assimilation, we further imposed constraints on the number of observations within each grid cell. Specifically, two filtering schemes were tested, in which grid cells with fewer than 10 or fewer than 50 original observations were excluded.

...

2.5 Assimilation algorithm

...

The prior estimate $\boldsymbol{f}_b$ is from the inventories described in Section 2.2, and the formaldehyde ~~concentration~~ observations $\boldsymbol{y}$ are described in Section 2.3. Mathematically, assimilation is performed via minimizing the cost function J as follows:

...

3 Results and discussion

Posterior of the NMVOC emission, formaldehyde ~~concentration~~column results and the impact on ozone simulation are discussed. A consistency analysis is introduced to assess the reliability of the two posterior emission.

...

3.3 Formaldehyde columns evaluation

...

Regarding the spatial patterns, high formaldehyde ~~values~~columns in February are concentrated in the

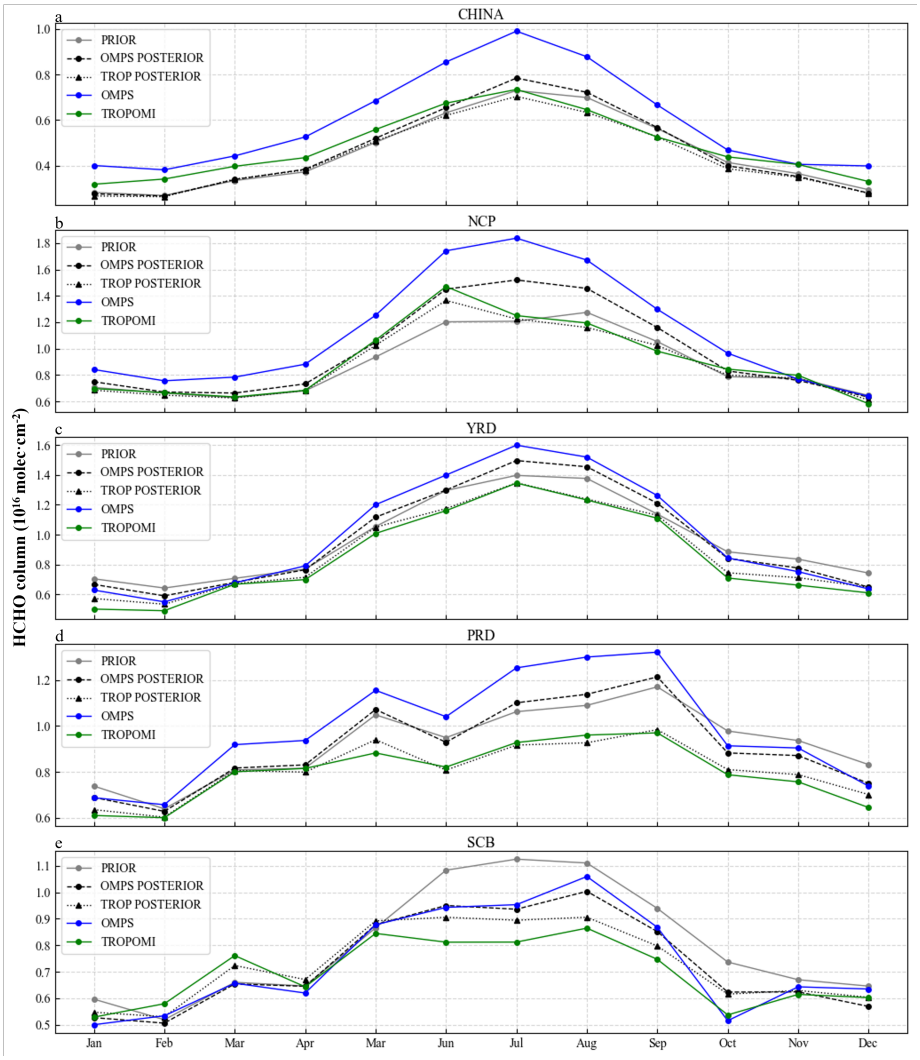

Figure 4. Monthly mean formaldehyde columns ~~concentrations~~ in 2020 from the prior simulation (gray), posterior simulations constrained by assimilating OMPS (black dashed) and TROPOMI (black dotted) observations, and satellite observations from OMPS (blue) and TROPOMI (green). Panels show results over China (a), the North China Plain (b), the Yangtze River Delta (c), the Pearl River Delta (d), and the Sichuan Basin (e).

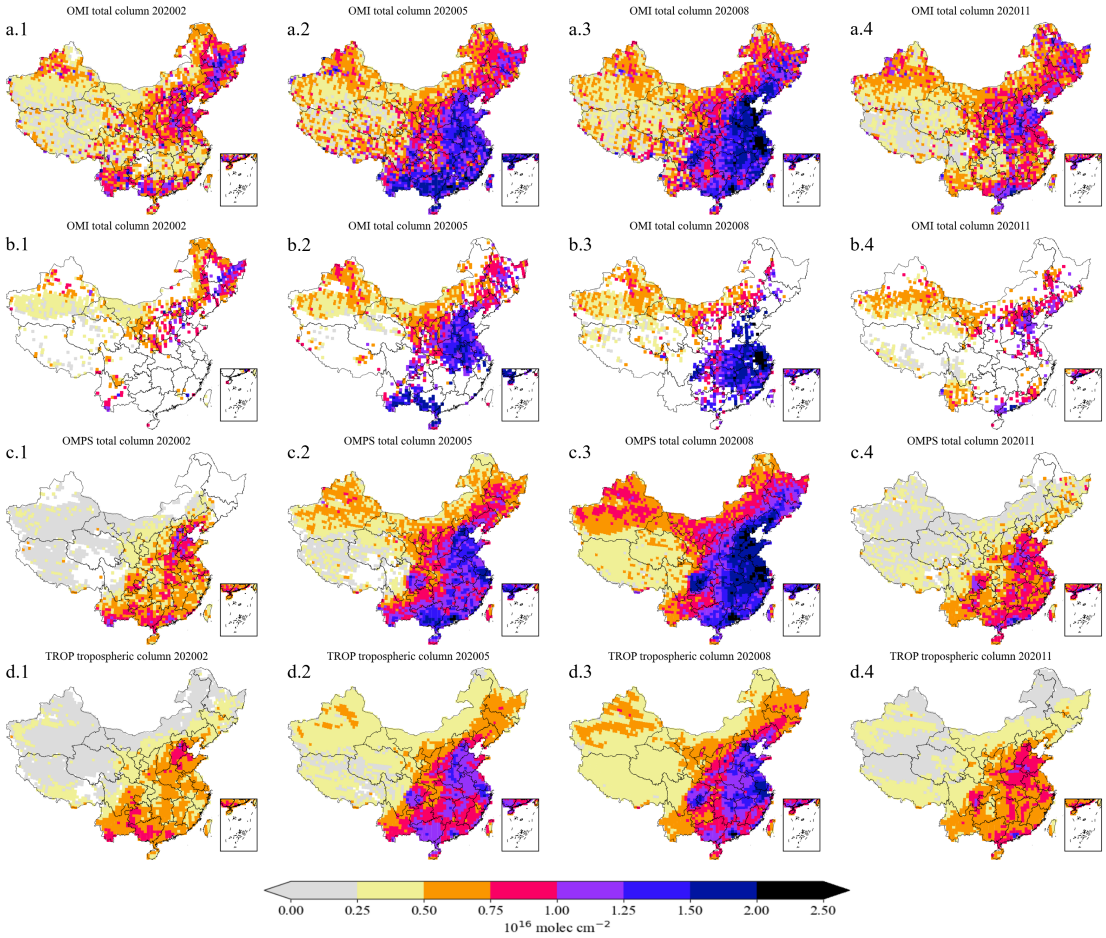

Figure 5. Comparison of monthly mean formaldehyde columns ~~concentrations~~ in February, May, August, and November 2020 after applying different data filtering thresholds. Panels (a.1-a.4), (c.1-c.4), and (d.1-d.4) show OMI, OMPS, and TROPOMI results, respectively, after removing grid cells with fewer than 10 observations. Panels (b.1-b.4) show OMI results after removing grid cells with fewer than 50 observations.

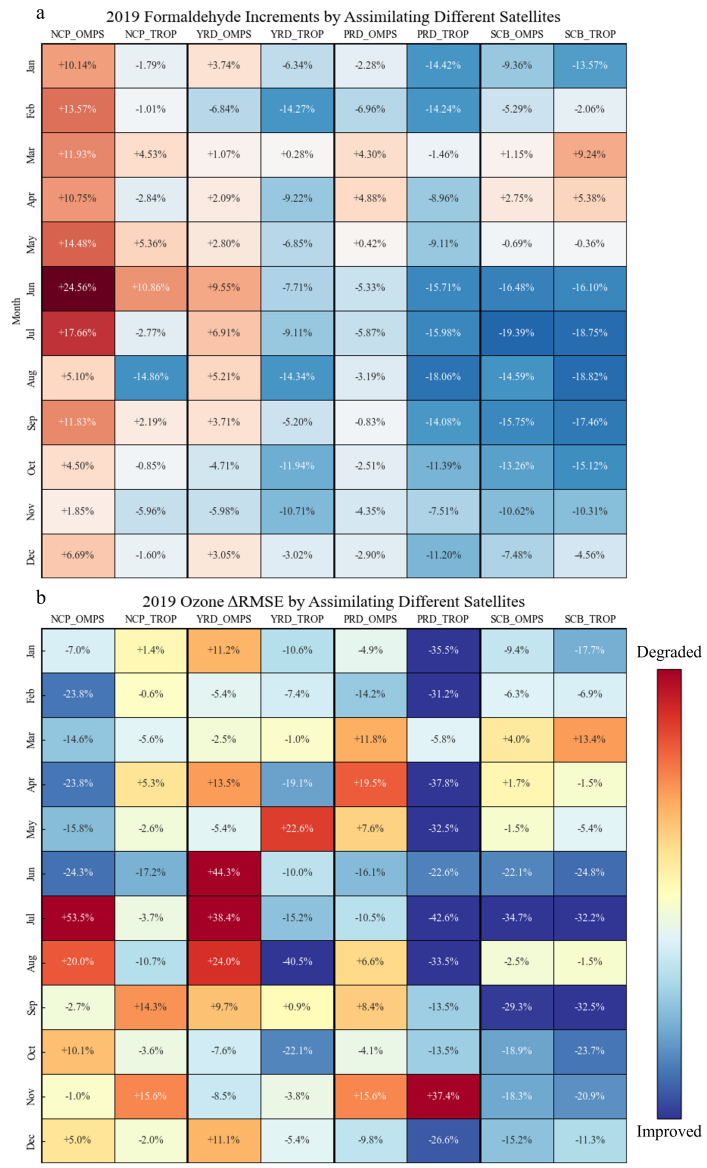

Figure 6. Monthly increments in (a) formaldehyde columns ~~concentrations~~ between posterior and prior simulations and (b) the relative changes in MDA8 ozone RMSE (ΔRMSE) after assimilating OMPS and TROPOMI observations in 2019. Results are shown for the North China Plain, Yangtze River Delta, Pearl River Delta, and Sichuan Basin. Positive values indicate an increase relative to the prior, while negative values indicate a decrease.

NCP, YRD, and PRD regions, with the posterior results showing an expanded high-value area in the NCP but a reduced coverage in the YRD. In May, overall ~~concentrations~~formaldehyde columns increase nationwide, with particularly pronounced growth in the NCP and PRD. In August, ~~concentrations~~formaldehyde columns increase in the NCP, YRD, and PRD, while they decrease in the SCB. In November, the changes are modest, but all four regions exhibit reduced ~~concentrations~~formaldehyde columns.

4 Summaryandconclusion

...

To further test the robustness of our approach, OMPS and TROPOMI satellite observations were independently assimilated to constrain NMVOC emissions for 2019 (Figure S4). The spatial distribution of formaldehyde hotspots is similar to 2020 but with overall higher ~~concentrations~~formaldehyde columns.

...

EC: ***4) Figure 4: a) I'm not sure what is shown in this figure. On the x-axis, you write "column concentration", which I assume should be column. But a column does not depend on altitude, so maybe partial columns are shown? Or are these concentrations or mixing ratios? Please check and correct. b) Independently of this, the profiles are not "satellite observations". The satellite products are columns, and the only profile I can think of is the a priori profile used for the AMF calculation. However, these profiles are model profiles, not measurements. The differences between the profiles you are reporting are then just the differences between the a priori used in the different products: GEOS-Chem for OMI and OMPS, TM5 for TROPOMI. It is therefore also not surprising that GEOS-Chem agrees well with the OMI and OMPS a priori, as these are just different runs of the same basic model. This misunderstanding needs to be fixed in the figure and in the discussion in the text (page 11, line 10, page 15, first paragraph, and other places).***

AR: We sincerely thank the reviewer for these insightful comments on the figure. We understand the reviewer is referring to Figure 3, as the description and the x-axis label "HCHO column concentration" match that figure rather than Figure 4 (Figure 3 in this response).
After careful reconsideration and to ensure maximum consistency with widely adopted terminology in the previous studies (Zhu et al., 2016, 2020), we have revised the presentation as follows:
a) The x-axis now reads "Shape Factor S(p)". The plotted quantity is the shape factor used in the air mass factor (AMF) calculation of each satellite product. This is exactly the same definition used in the references above.
b) We fully agree that these are not satellite-measured profiles. They are the a priori shape factors (derived from the respective a priori profiles) that enter the AMF computation of each retrieval algorithm. In our calculations, the shape factors are obtained by multiplying the a priori profiles by the layer pressure thickness (Δp), followed by normalization.
OMI and OMPS-NOAA20 both use a priori profile derived from GEOS-Chem simulations, which explains their close similarity to the GEOS-Chem profile simulated in this work. TROPOMI uses the a priori profile derived from the TM5-MP model, leading to the visible differences. This misunderstanding has been fully corrected in the revised figure, figure caption, and throughout the manuscript

Text in manuscript

Abstract.

...

~~This study evaluates the vertical profile structure, qualified-data volume and monthly mean biases of formaldehyde products from OMI, OMPS and TROPOMI satellites, emphasizing the enhanced capabilities of the latest OMPS and TROPOMI retrievals.~~This study evaluates the shape factor, filtered data volume, and monthly mean biases of OMI, OMPS, and TROPOMI formaldehyde products, with the latest OMPS and TROPOMI retrievals offering substantially higher effective spatiotemporal coverage.

...

2.3.2 Sentinel-5P TROPOMI

...

This criterion ensures the exclusion of error flags and requires that the cloud radiance fraction at 340 nm is below 0.5, the solar zenith angle (SZA) does not exceed 70°, the surface albedo is below 0.2, no snow or ice warning is present, and the air mass factor (AMF) is larger than 0.1. ~~The TROPOMI product provides vertical information on 34 layers, but the retrieval is primarily sensitive to the troposphere and thus reports the formaldehyde tropospheric column.~~The operational TROPOMI HCHO Level-2 product provides tropospheric vertical columns together with averaging kernels and a priori profiles defined on 34 vertical layers (from the surface to ∼0.1 hPa). Because stratospheric formaldehyde is negligible (De Smedt et al., 2018, 2021), we directly use the reported tropospheric columns in this study without reconstructing total columns.

...

3.1 Satellite data evaluation

~~The results shown in Figure 3 are based on formaldehyde columns before the application of model vertical profile corrections. The total columns of OMI, OMPS and the GEOS-Chem simulations are presented. Their vertical structures generally decrease with altitude but exhibit characteristic minima and maxima at several pressure levels: a trough at approximately 750-850 hPa, followed by a first peak about 30 hPa above; a second pronounced peak around 600-700 hPa; and a third peak at 350 hPa, which is more prominent in April and August but weaker in November and January. Above 350 hPa, formaldehyde partial column gradually diminish to near zero. These features are broadly consistent with previously reported formaldehyde profile distributions (Zhu et al., 2020). Overall, OMPS exhibits a closer agreement with the model profiles than OMI, whereas OMI shows displaced or poorly defined peaks in several regions, with particularly anomalous enhancements in the upper levels over NCP.~~

~~In contrast, the TROPOMI product provides tropospheric columns and displays markedly different vertical structures compared to the other three datasets. Its profiles decrease monotonically with altitude, with only localized perturbations over SCB (around 300 hPa) and PRD (around 800 hPa). The overall trend resembles a logarithmic decay, with especially high near-surface concentration (Vigouroux et al., 2020). These discrepancies highlight the importance of applying vertical profile corrections in satellite-model comparisons Eskes and Boersma (2003). Specifically, the similarity in vertical structures between OMI, OMPS, and GEOS-Chem supports the use of the AMF approach for correction, while the distinct TROPOMI profiles necessitate the application of AVK.~~

Figure 3 shows the vertical profiles of formaldehyde shape factors used to compute the reported satellite vertical columns before any model-based profile correction is applied. Both the OMI and OMPS retrievals use GEOS-Chem model outputs as their a priori profiles (González Abad et al., 2015, 2016; Nowlan et al., 2023). Consequently, their shape factors are highly similar. formaldehyde shape factors generally decrease with altitude but exhibit characteristic peaks and troughs: a minimum at 750–850 hPa, a first peak ∼30 hPa above it, a second prominent peak at 600–700 hPa, and a third peak near 350 hPa that is strongest in April and August and weaker in January and November. Above 350 hPa, concentration decay toward zero. These features agree well with previously reported formaldehyde profile shapes over China (Zhu et al., 2016, 2020). Overall, the OMPS profile most closely matches GEOS-Chem, whereas OMI shows slight peak shifts or spurious upper-level enhancements in some regions, particularly during May and August.

In contrast, the operational TROPOMI retrieval uses the a priori profiles from the TM5-MP model (De Smedt et al., 2018, 2021), which places substantially more mass near the surface. This results in a markedly different vertical structure: an approximately logarithmic monotonic decrease with altitude, with only minor perturbations over SCB (∼300 hPa) and PRD (∼800 hPa) and very high near-surface shape factor. These differences in a priori profile shape are the primary reason why profile correction is essential for meaningful satellite–model comparisons (Eskes and Boersma, 2003).

Additionally, Figure 3 clearly shows that stratospheric formaldehyde (> ∼200 hPa) can be largely neglected in both the GEOS-Chem simulation and the OMPS and TROPOMI retrievals (González Abad et al., 2015; De Smedt et al., 2018). Because the stratospheric contribution is negligible and no explicit stratospheric correction is applied in either retrieval, we hereafter use the term "formaldehyde column" without distinguishing between total and tropospheric columns in subsequent discussion.

EC: ***5) Use of AMF and Averaging Kernels: In the manuscript, you differentiate between the use of GEOS-Chem based AMFs for OMPS and AVK for TROPOMI. However, there is no difference between these two approaches. The AVK is just the height-dependent AMF (sometimes also called weighting function or scattering weights) divided by the total AMF used. If you compute the AMF based on GEOS-Chem profiles and apply it to the slant column, this is identical to applying the column AVK and the GEOS-Chem vertical profile to the vertical column from the product. The only complication is the proper treatment of the background correction, but this applies to both approaches. Please reconsider and clarify in the text. For the figures, it is not clear whether the satellite columns you show are the original or the AMF / AVK corrected ones. Please clarify.***

AR: We appreciate the reviewer for pointing out this important issue. We fully agree that the GEOS-Chem-based scattering weights/AMF approach used for OMI/OMPS and the averaging kernel (AVK) approach used for TROPOMI are mathematically equivalent when the same a priori profile is applied. Following the reviewer's suggestion, we have clarified this point explicitly in the revised Section 2.3.2 (Sentinel-5P TROPOMI). The text now states that both methods are fundamentally identical, but because the official TROPOMI product directly provides column averaging kernels while OMI/OMPS provide scattering weights, the implementation formulas differ slightly in practice. Additionally, to remove any remaining ambiguity, we now clearly state in text and in the captions of all relevant figures that all satellite formaldehyde columns presented in the manuscript are the fully profile-corrected columns.

Text in manuscript

2.3.1 NOAA-20 OMPS

...

~~The processed OMPS satellite observations were ultimately assimilated as total columns, which are presented in Figure 1 (c.1-c.4).~~ The OMPS observations were assimilated as total columns after re-calculation of the air mass factor using GEOS-Chem shape profiles for consistency with the model vertical profile. These reprocessed total columns are shown in Figure 1 (c.1–c.4). The original a priori shape factors used in the official satellite products (before reprocessing) are displayed as red lines in Figure 3.

...

2.3.2 Sentinel-5P TROPOMI

...

~~Beyond the recommended quality screening, a key consideration when comparing TROPOMI formaldehyde retrievals with model outputs is the dependence on the assumed a priori vertical profile. Traditionally, studies have relied on AMF-based corrections, in which AMF is recalculated using model-derived profiles to reduce such discrepancies (Palmer et al., 2001; Palmer et al., 2004; Palmer et al., 2016; Palmer et al., 2020). More recently, the availability of averaging kernel (AVK) information in the TROPOMI product has allowed a more consistent comparison by accounting for the impact of the assumed vertical profile shape in the retrieval, following the approach introduced in the IASI NH_3 version 4 product (Clarisse et al., 2023; Clarisse et al., 2025). In this study, we apply AVK-based correction for TROPOMI formaldehyde by projecting the model profiles onto the satellite pressure grid, thereby achieving a more harmonized comparison with GEOS-Chem simulations. The corrected column is calculated as:~~

Beyond the recommended quality filtering, a critical consideration when comparing TROPOMI formaldehyde retrievals with model simulations is the sensitivity of the retrieved columns to the a priori vertical profile assumed in the retrieval algorithm. In this study, OMPS and OMI formaldehyde products are harmonized with the model by recalculating the AMF using GEOS-Chem shape factors, following the conventional approach. For TROPOMI, the officially provided averaging kernels (AVK) are applied instead. Importantly, these two correction strategies are mathematically equivalent. The averaging kernel represents the ratio of the altitude-resolved sensitivity to the total AMF used in the operational retrieval. Consequently, convolving the model profile with the AVK and adding the same background column yields identical results to recalculating the total AMF with the model profile and applying it to the slant column, provided that the background correction and a priori profile replacement are handled consistently (Palmer et al., 2001; Eskes and Boersma, 2003; Boersma et al., 2004; De Smedt et al., 2021). Both approaches achieve the same objective: removing the influence of the satellite's a priori profile and replacing it with the GEOS-Chem profile, thereby ensuring observational–model consistency prior to assimilation. The AVK application for TROPOMI employed here follows the methodology established for the IASI NH_3 version 4 product (Clarisse et al., 2023; Xia et al., 2025). The corrected column is calculated as:

...

~~The processed TROPOMI retrievals were assimilated as tropospheric columns, which are presented in Figure 1 (d.1-d.4),~~The TROPOMI tropospheric columns were assimilated after application of GEOS-Chem shape profiles. These reprocessed tropospheric columns are shown in Figure 1 (d.1–d.4), with

their vertical shape factors shown in Figure 3 (green line) to illustrate the normalized contribution of each pressure layer to the tropospheric columns. We adopted tropospheric rather than total columns because the retrieval product itself provides tropospheric columns. ~~and recalculating total columns would introduce uncertainties.~~

...

2.3.3 Aura OMI

...

The vertical profile correction of OMI formaldehyde was conducted using the same approach as applied to OMPS, by recalculating AMF with model-simulated vertical profiles.The resulting OMI columns after profile correction and the two data-volume filters are shown in Figure S3.

EC: ***6) Abstract, line 6: See major comment on vertical structure – you are comparing the vertical structure of the a priori used, not of the satellite data.***

AR: We thank the reviewer for this precise comment. Following the reviewer's earlier major comment on vertical structure, we have revised the Abstract to clearly reflect that the comparison refers to the a priori profiles used in the retrievals rather than satellite-measured vertical distributions.

Text in manuscript

Abstract.

...

~~This study evaluates the vertical profile structure, qualified-data volume and monthly mean biases of formaldehyde products from OMI, OMPS and TROPOMI satellites, emphasizing the enhanced capabilities of the latest OMPS and TROPOMI retrievals.~~This study evaluates the shape factor, filtered data volume, and monthly mean biases of OMI, OMPS, and TROPOMI formaldehyde products, with the latest OMPS and TROPOMI retrievals offering substantially higher effective spatiotemporal coverage.

...

EC: ***7) Abstract, line 7: "emphasizing the enhanced capabilities" – not sure what this is referring to.***

AR: Thank you for pointing this out. We agree that the original phrase was vague. We have revised the sentence to explicitly state that the improvement refers to the substantially higher effective spatiotemporal coverage provided by the latest OMPS and TROPOMI retrievals.

Text in manuscript

Abstract.

...

~~This study evaluates the vertical profile structure, qualified-data volume and monthly mean biases of formaldehyde products from OMI, OMPS and TROPOMI satellites, emphasizing the enhanced capabilities of the latest OMPS and TROPOMI retrievals.~~This study evaluates the shape factor, filtered data volume, and monthly mean biases of OMI, OMPS, and TROPOMI formaldehyde products, with the

latest OMPS and TROPOMI retrievals offering substantially higher effective spatiotemporal coverage.

...

EC: ***8) Page 1, line 10: Add reference for GFED.***

AR: Thank you for highlighting this point. We have now added the reference for GFED.

Text in manuscript

1 Introduction

...

For biomass burning, widely used inventories include the Global Fire Emissions Database (GFED) (Van Der Werf et al., 2017) and the Fire INventory from NCAR (FINN) (Wiedinmyer et al., 2011).

...

EC: ***9) Page 1, line 19: Add reference for GFAS.***

AR: The reviewer's comment is appreciated. The manuscript now includes the GFAS citation.

Text in manuscript

1 Introduction

...

Biomass burning emissions also show large discrepancies across inventories (e.g., GFED, FINN, GFAS) (Kaiser et al., 2012), largely driven by uncertainties in burned area, fuel loading, and emission factors (He et al., 2011; Hua et al., 2024).

...

EC: ***10) Page 3, line 7: I do not know why you think the TROPOMI glyoxal product is discontinued. You will find the data at https://data-portal.s5p-pal.com/products/chocho.html.***

AR: We thank the reviewer for the correction. We acknowledge the mistake regarding the status of the TROPOMI glyoxal product and have removed the incorrect statement accordingly.

Text in manuscript

1 Introduction

...

~~Notably, the latest version of Sentinel-5 Precursor Tropospheric Monitoring Instrument (TROPOMI) glyoxal products is no longer publicly available on the official website.~~

Glyoxal retrieval product from satellite platform began relatively late, with the first global differential optical absorption spectroscopy (DOAS) retrievals reported by Wittrock et al. (2006) using SCIAMACHY, followed by their application to constrain NMVOC emissions by Stavrakou et al. (2009a). Because glyoxal is retrieved in a longer wavelength range (∼435–460 nm) than formaldehyde

(∼330–360 nm), it exhibits markedly lower sensitivity to molecular scattering, which in turn reduces the sensitivity of the measurement to the lower troposphere (Palmer et al., 2001; Chan Miller et al., 2014). Moreover, glyoxal optical depths are very weak (order of 10^{-4}–10^{-3}), rendering the retrieval highly susceptible to fitting residuals from stronger absorbers, uncertainties in absolute radiometric calibration, and spectral features in surface reflectivity (Sinreich et al., 2013; Alvarado et al., 2014). For instruments with comparatively modest spectral resolution and signal-to-noise ratios, such as OMI, these interference effects are further amplified, leading to larger retrieval uncertainties for glyoxal columns than for formaldehyde (Chan Miller et al., 2014; Cao et al., 2018). Consequently, glyoxal satellite observations remain considerably less suitable than formaldehyde for high-spatiotemporal-resolution assimilation studies. Beyond glyoxal and formaldehyde, retrievals of other VOCs are also progressing, as exemplified by Fu et al. (2019) and Wells et al. (2020, 2022), who derived isoprene columns from Cross-track Infrared Sounder (CrIS) observations, representing an important step toward next-generation satellite constraints on volatile organic compounds.

EC: ***11) Page 3, line 32: This paper does not include TROPOMI data.***

AR: We appreciate the reviewer's careful reading. You are correct that this paper does not include TROPOMI data. We have corrected the citation and ensured its accuracy in the revised manuscript.

Text in manuscript

1 Introduction

...

~~Choi et al. (2022) applied a 4DVar system to assimilate TROPOMI formaldehyde over East Asia, demonstrating the capability of high-resolution satellite data to capture regional and seasonal variability in VOC emissions, but the analysis was conducted only for May–June.~~

Choi et al. (2022) assimilated OMPS and OMI observations into an updated 4DVar system for East Asia during May–June 2016. Their inversion revealed a 47% increase in VOC emissions across Northeast Asia relative to the prior inventory, indicating that isoprene emissions over South Korea and anthropogenic NMVOC emissions over eastern China were underestimated in the bottom-up inventory.

EC: ***12) Page 4, line 8: The OMI row anomaly is not limited to China but affects data globally.***

AR: We thank the reviewer for the valuable reminder. The manuscript has been updated to reflect that the OMI row anomaly affects data globally.

Text in manuscript

1 Introduction

...

~~However, the retrievals over China are reported to suffer from the so-called "row anomaly" issue (González Abad et al., 2015), which may reduce the assimilation accuracy especially in the high-resolution configuration.~~

However, it has been affected globally by the persistent row anomaly since 2007 Kroon et al. (2011); Zhu et al. (2017), which degrades data quality in certain across-track positions and may reduce

assimilation accuracy, particularly in high-resolution configurations.

EC: ***13) Page 5, line 2: Which "boundary conditions" are you referring to?***

AR: We appreciate the reviewer's question. We have clarified in the revised manuscript that the global simulation provides lateral chemical boundary conditions to the nested Asia domain, which are updated every 3 hours.

Text in manuscript

2.1 Model simulation

...

~~The global simulation has a horizontal resolution of 2° latitude × 2.5° longitude, with boundary conditions updated every 3 hours.~~The global simulation is run at 2° × 2.5° resolution and provides lateral chemical boundary conditions to the nested Asia domain updated every 3 hours.

EC: ***14) Page 5, line 9: The Lin et al. paper referred to describes the implementation of a new and fast chemical solver into KPP3.0; it is not related to the chemical scheme or the oxidant-aerosol reactions used in your version of GEOS-Chem.***

AR: We thank the reviewer for carefully pointing out this error. The manuscript has been corrected to remove the inaccurate statement regarding the Lin et al. paper.

Text in manuscript

2.1 Model simulation

...

To better simulate oxidant-aerosol reactions in the ~~stratosphere and~~ troposphere,~~ GEOS-Chem v14.1.1 introduced a new KPP 3.0 chemical solver (Lin et al., 2023), including chemical mechanisms for isoprene, aromatics, ethylene, and acetylene (Bates and Jacob, 2019; Bates and Jacob, 2021; Bates and Jacob, 2021).~~ GEOS-Chem v14.1.1 includes state-of-the-science tropospheric chemistry with recent updates to the oxidation mechanisms of isoprene (Bates and Jacob, 2019), aromatics (Bates et al., 2021), ethylene, and acetylene (Kwon et al., 2021).

EC: ***15) Page 12, line 1: See major comment on tropospheric and total columns.***

AR: We thank the reviewer for this remark. We fully agree that any stratospheric formaldehyde column is negligible compared to retrieval uncertainties. The original sentence has therefore been removed from the revised manuscript. **Text in manuscript**

2.3.2 Sentinel-5P TROPOMI

...

We adopted tropospheric rather than total columns because the retrieval product itself provides tropospheric columns.~~ and recalculating total columns would introduce uncertainties.~~

EC: ***16) Page 12, line 9: Zhang et al. is not a good general reference for OMI.***

AR: We appreciate the reviewer's suggestion. The manuscript has been updated to replace the Zhang et al. reference with more appropriate and authoritative citations.

Text in manuscript

> **2.3.2 Aura OMI**
>
> The Ozone Monitoring Instrument (OMI) is an important satellite instrument onboard the Aura satellite, launched on July 15, 2004, with the objective of monitoring atmospheric gases, aerosols, and clouds to improve our understanding of atmospheric chemistry and climate change. OMI provides daily global coverage with a wide swath of 2600 km and a spatial resolution of approximately 13 $\times$ 24 km at nadir, with an equator crossing time of about 13:45 LT. The sensor contains three spectral channels (UV-1, UV-2, and VIS), covering the wavelength ranges of 264-311 nm, 307-383 nm, and 349-504 nm, respectively, which enable the retrieval of key trace gases including O_3, NO_2, SO_2, and formaldehyde ~~(Zhang et al., 2019)~~(Levelt et al., 2006, 2018).

EC: ***17) Page 16, line 32: Note that different AMF or vertical profiles do not apply, as you have corrected for them.***

AR: We thank the reviewer for this comment. The sentence in question has been removed entirely from the revised manuscript, as the differences in AMF/AVK and vertical profiles have already been corrected for and do not contribute to the discrepancies discussed.

Text in manuscript

> **3.2 NMVOCemissions**
>
> ...
>
> These differences may arise from variations in emission characterization during summer, marked by strong convection, high humidity, and elevated cloud and aerosol content, which differentially impact the retrieval of optical depth and columns by OMPS and TROPOMI. ~~Additionally, differences in AMF and vertical profiles may further contribute to these discrepancies (Lorente et al., 2017; Lorente et al., 2021).~~

EC: ***18) Figure 6: I understand that the colour code is somehow linked to probability, but a probability can not be larger than 1 by definition, so something else is shown here, or a factor is missing.***

AR: Thank you for the reviewer's valuable comment. The color scale represents the probability density of the data points, not the probability itself. Since probability density is a density function over a continuous domain, its values can exceed 1. The integral (area under the curve) of the probability density over the entire plot sums to 1, consistent with probability theory.

EC: ***19) Page 26, Line 16: as above. AMF and a priori differences have already been corrected.***

AR: We thank the reviewer for this comment. We have completely removed the inaccurate statement regarding the impact of AMF and vertical-profile differences from the revised manuscript.

Text in manuscript

4 Summary and conclusion

...

By contrast, notable discrepancies emerge in June-September—dominated by July-August—likely linked to strong convection, high humidity, and elevated cloud/aerosol loading that differentially affect retrievals. ~~compounded by AMF and vertical-profile differences.~~

RC: ***I appreciate the authors' efforts to expand the literature review on VOC inversion studies. However, the current revision still lacks a clear methodological and chronological structure. The discussion of Wells et al. (2020, 2022) seems to misunderstand their relevance; these papers introduce new CrIS isoprene retrievals that represent an important measurement wise extension for future VOC inversions. My suggestion to include these studies was to highlight the expanding observational basis for VOC emission constraints, not to present them as inversion studies. While*** CH_2O ***has served as the baseline constraint, CHOCHO has provided an additional constraint,*** NO_2 ***has offered information on chemical interactions, and isoprene is the latest extension. The description of Choi et al. (2022) is also inaccurate because this study did not use TROPOMI observations and covered only a short May to June period, so it did not provide seasonal variations. In addition, the overall structure should include some methodological context, since Palmer et al. (2003) was the pioneering VOC satellite inversion using a Bayesian framework and there were important developments between that study and more recent studies, including this study using a 4DEnVar system. For this reason, the current flow from Palmer et al. (2003) directly to Wells et al. (2020, 2022) is not reasonable from either a methodological or a measurement wise perspective. Reorganizing this section accordingly would improve accuracy and readability.***

AR: We sincerely thank the reviewer for the careful reading and for these highly constructive suggestions. Following the reviewer's guidance, we have completely reorganized and rewritten the literature review in the Introduction (now structured as three clearly separated parts):
(1) A chronological overview of satellite formaldehyde (and related VOC tracer) observations, from early instruments (GOME/SCIAMACHY/GOME-2) to the current generation (OMI, OMPS, TROPOMI). The work of Wells et al. (2020, 2022) on CrIS direct isoprene retrievals is now presented at the end of this part as an important recent extension of the observational basis for future VOC emission inversions.
(2) The historical evolution of top-down approaches using formaldehyde observations to constrain NMVOC emissions. This part starts with the pioneering Bayesian inversion of Palmer et al. (2003), traces subsequent methodological developments, thereby providing clear methodological context for the present study.
(3) A focused summary of previous top-down NMVOC emission inversions over China that relied on formaldehyde assimilation. We explicitly highlight the limitations of these earlier studies and the gaps that motivate the innovations introduced here. Regarding Choi et al. (2022), we sincerely apologize for the earlier inaccurate description: that study used OMI and OMPS (not TROPOMI) observations and covered only a short May–June period. We have carefully reread the paper and corrected all related statements.

Text in manuscript

1 Introduction

...

Remote sensing observations of these compounds typically rely on spectral channels in the ultraviolet-visible (UV-Vis) range, with their primary absorption features occurring between 330 and 460 nm (Platt, 1979; Lerot et al., 2010; De Smedt et al., 2012).

~~Compared to glyoxal, satellite products for formaldehyde are better established.. Satellite observations of glyoxal began later with initial identifications made using Aura Ozone Monitoring Instrument (OMI), and its retrieval is more challenging than that of formaldehyde (Kurosu et al., 2005; Kurosu et al., 2014). Recent years have seen further advancements in satellite observational instruments and algorithms for formaldehyde (Abad, 2022; Abad, 2017), leading to new satellite observation products with improvements in both accuracy and resolution. In contrast, while the glyoxal retrieval algorithm has been updated for OMI product (Alvarado et al., 2014), its satellite products continue to face substantial uncertainties. Notably, the latest version of Sentinel-5 Precursor Tropospheric Monitoring Instrument (TROPOMI) glyoxal products is no longer publicly available on the official website.~~

~~Formaldehyde measurements from instruments such as the Global Ozone Monitoring Experiment-2 (GOME-2) (De Smedt et al., 2012), OMI (González Abad et al., 2015), Ozone Mapping and Profiler Suite (OMPS) (Li et al., 2015) and TROPOMI (De Smedt et al., 2018) have been used widely for estimating NMVOC emissions through data assimilation. The core of the methodology is to calculate the most likely NMVOC emissions given the formaldehyde observations and the prior information. For instance, Fu et al. (2007) used six years of continuous satellite measurements of formaldehyde columns from GOME (1996-2001) to improve regional emission estimates of reactive NMVOCs, including isoprene, olefins, formaldehyde, and xylene, for East Asia and South Asia. Similarly, formaldehyde data from the GOME-2A satellite were used to constrain NMVOC emissions in India for 2009 (Chaliyakunnel et al., 2019). Souri et al. (2020) used observations from OMPS satellites during the KORUS-AQ campaign to estimate NO_x and NMVOC emissions in East Asia from May to June 2016. Kaiser et al. (2018) also utilized high-resolution formaldehyde retrieval data from OMI instrument to quantify isoprene emissions at the ecosystem scale in the southeastern United States during August-September 2013. Those promising results have demonstrated that formaldehyde measurements could be utilized to optimize the existing NMVOC emissions that were established in a bottom-up manner.~~

Satellite remote sensing of formaldehyde has made substantial progress since the atmospheric formaldehyde abundance was first retrieved in 1997 (Burrows et al., 1999). The earliest retrievals of formaldehyde vertical column densities were based on the Global Ozone Monitoring Experiment (GOME) (Thomas et al., 1998; Chance et al., 2000). Subsequently, the Scanning Imaging Absorption Spectrometer for Atmospheric Chartography (SCIAMACHY) served as an important transitional instrument between GOME and GOME-2, offering significantly improved spatial resolution compared to GOME (De Smedt et al., 2008). In 2004, the launch of NASA's Aura satellite carrying the Ozone Monitoring Instrument (OMI) provided high signal-to-noise-ratio ultraviolet–visible (UV–Vis) spectra that greatly advanced trace-gas retrieval studies (De Smedt et al., 2015; González Abad et al., 2015). Approximately four years after GOME effectively ceased operational observations, its successor, GOME-2, began routine operations in 2007 and started delivering formaldehyde data (De Smedt et al., 2012). In recent years, high-resolution formaldehyde observations have continued to emerge, including those from the Ozone Mapping and Profiler Suite (OMPS) onboard the Suomi National Polar-orbiting Partnership (Suomi NPP) and NOAA-20 satellites (Li et al., 2015; González Abad et al., 2015, 2016; Nowlan et al., 2023), as well as from the Tropospheric Monitoring Instrument (TROPOMI) aboard the Sentinel-5 Precursor (Sentinel-5P) launched in 2017. TROPOMI's exceptional spatial resolution and near-daily global coverage have marked a new era in satellite formaldehyde monitoring (De Smedt et al., 2018, 2021). Furthermore, geostationary satellites now provide formaldehyde observations with high temporal resolution, including the Geostationary Environment Monitoring Spectrometer (GEMS) over East Asia (Kwon et al., 2019; Kim et al., 2020), the Tropospheric Emissions: Monitoring of Pollution (TEMPO)

instrument over North America (Chance et al., 2019), and Sentinel-4, successfully launched on 1 July 2025, which is conducting geostationary formaldehyde observations over Europe (Gulde et al., 2017).

Glyoxal retrieval product from satellite platform began relatively late, with the first global differential optical absorption spectroscopy (DOAS) retrievals reported by Wittrock et al. (2006) using SCIAMACHY, followed by their application to constrain NMVOC emissions by Stavrakou et al. (2009a). Because glyoxal is retrieved in a longer wavelength range (∼435–460 nm) than formaldehyde (∼330–360 nm), it exhibits markedly lower sensitivity to molecular scattering, which in turn reduces the sensitivity of the measurement to the lower troposphere (Palmer et al., 2001; Chan Miller et al., 2014). Moreover, glyoxal optical depths are very weak (order of 10^{-4}–10^{-3}), rendering the retrieval highly susceptible to fitting residuals from stronger absorbers, uncertainties in absolute radiometric calibration, and spectral features in surface reflectivity (Sinreich et al., 2013; Alvarado et al., 2014). For instruments with comparatively modest spectral resolution and signal-to-noise ratios, such as OMI, these interference effects are further amplified, leading to larger retrieval uncertainties for glyoxal columns than for formaldehyde (Chan Miller et al., 2014; Cao et al., 2018). Consequently, glyoxal satellite observations remain considerably less suitable than formaldehyde for high-spatiotemporal-resolution assimilation studies. Beyond glyoxal and formaldehyde, retrievals of other VOCs are also progressing, as exemplified by Fu et al. (2019) and Wells et al. (2020, 2022), who derived isoprene columns from Cross-track Infrared Sounder (CrIS) observations, representing an important step toward next-generation satellite constraints on volatile organic compounds.

~~Compared with formaldehyde retrieval, satellite remote sensing of glyoxal began much later and remains considerably more challenging. Shim et al. (2005) assimilated formaldehyde observations from the GOME using a global Bayesian inversion to constrain isoprene emissions. Although China was included within their East Asia region, the analysis lacked region-specific focus and did not provide detailed characterization of emission patterns over China, and the coarse spatial resolution (4° × 5°) in that study further limited the ability to resolve subregional emission features. Stavrakou et al. (2016) conducted a regional inversion in Eastern China using multi-year satellite formaldehyde data from GOME and OMI to constrain VOC emissions during the post-harvest burning period, and they indicated that the crop burning fluxes of VOCs in June exceeded by a factor of two the combined emissions from other anthropogenic activities in the NCP region from 2005 to 2012. Cao et al. (2018) conducted a relatively systematic satellite-based emission inversion study over China, using a 4DVar method and assimilating OMI and GOME-2A formaldehyde products to estimate monthly NMVOC emissions in 2007, though the spatial resolution (4° × 5°) was still too coarse. Choi et al. (2022) applied a 4DVar system to assimilate TROPOMI formaldehyde over East Asia, demonstrating the capability of high-resolution satellite data to capture regional and seasonal variability in VOC emissions, but the analysis was conducted only for May-June. Beyond China, a number of important studies have advanced top-down VOC inversion methodologies: Palmer et al. (2003) pioneered the use of GOME formaldehyde observations in a Bayesian framework to constrain global isoprene emissions, laying the foundation for subsequent satellite-based VOC studies; Wells et al. (2020, 2022) further advanced this field by retrieving isoprene emissions from CrIS measurements and providing high-resolution constraints on VOC oxidation chemistry at the global scale; and Oomen et al. (2024) derived weekly top-down VOC fluxes over Europe from TROPOMI formaldehyde data using the MAGRITTEv1.1 model, providing improved constraints on isoprene, biomass burning, and anthropogenic VOC emissions. Considering the increasingly stringent air pollution control in China (Wu et al., 2024), there is an urgent need for high-resolution top-down NMVOC emission optimization.~~

Top-down approaches, mainly assimilation techniques, with satellite formaldehyde columns have

become the primary method for constraining NMVOC emissions. Palmer et al. (2003) pioneered to apply a Bayesian inversion framework with GOME formaldehyde observations for constraining isoprene emissions over North America. The approach was subsequently extended to global and European domains by Shim et al. (2005) and Dufour et al. (2009), respectively. With the availability of OMI and GOME-2 formaldehyde products, inversion algorithms were further refined. Stavrakou et al. (2009b) first introduced an adjoint-based inversion to optimize biogenic emissions and, in a companion study the same year, revealed substantial underestimation of continental glyoxal sources worldwide (Stavrakou et al., 2009a). Concurrently, Millet et al. (2008) used OMI formaldehyde and identified the underestimated isoprene emissions over the north-central United States, while Zhu et al. (2014) reported that anthropogenic emissions of highly reactive VOCs (HRVOCs) in the Houston area were underestimated by a factor of 4.8 ± 2.7 compared to the US Environmental Protection Agency inventory. Formaldehyde product with much higher spatial resolution were then available since the launch of TROPOMI and OMPS, and made the city-scale emission optimizations possible (González Abad et al., 2016; De Smedt et al., 2018, 2021). In recent years, studies leveraging these new-generation instruments have proliferated. Choi et al. (2022) assimilated OMPS and OMI observations into an updated 4DVar system for East Asia during May–June 2016. Their inversion revealed a 47% increase in VOC emissions across Northeast Asia relative to the prior inventory, indicating that isoprene emissions over South Korea and anthropogenic NMVOC emissions over eastern China were underestimated in the bottom-up inventory. Oomen et al. (2024) used weekly-averaged TROPOMI formaldehyde observations from 2018–2021 with the MAGRITTEv1.1 adjoint model to derive top-down biogenic, pyrogenic, and anthropogenic VOC fluxes over Europe, substantially correcting previous underestimates of isoprene emissions. Feng et al. (2024) applied an Ensemble Kalman Filter (EnKF) approach to optimize August 2022 NMVOC emissions over China, revealing overestimation of biogenic emissions during an extreme heatwave and demonstrating consequential impacts on summertime ozone simulations. The advent of geostationary satellites (e.g., GEMS, TEMPO) with high-frequency observations has enabled the incorporation of diurnal cycle information into algorithm frameworks, making daily-scale top-down emission optimization feasible (Kwon et al., 2019, 2021). Meanwhile, multi-species constraint is gaining traction; Opacka et al. (2025) developed a novel inversion technique that simultaneously optimizes monthly-mean VOCs and NOx emissions from 2019 TROPOMI observations, uncovering severe underestimation of both NOx and VOCs in prior inventories over Africa.

Although substantial progress has been made globally in satellite-based top-down constraints on NMVOC emissions, high-resolution top-down emission optimization studies specifically over China remain relatively scarce in recent years. Shim et al. (2005) first used GOME formaldehyde observations in a global Bayesian inversion framework to constrain isoprene emissions. While their domain encompassed East Asia including China, the study lacked a dedicated focus on China and was limited by coarse model resolution (4° × 5°). Stavrakou et al. (2016) performed a regional inversion over eastern China using multi-year GOME and OMI formaldehyde columns, revealing that post-harvest agricultural burning in June contributed more than twice the VOC emissions of all other anthropogenic sources combined over the North China Plain during 2005–2012. Cao et al. (2018) conducted one of the most systematic satellite-constrained inversions for China to date, applying a 4DVar assimilation of OMI and GOME-2A formaldehyde products to estimate monthly NMVOC emissions in 2007, yet the analysis was still constrained by the same coarse 4° × 5° resolution. Choi et al. (2022) assimilated OMPS and OMI formaldehyde columns into a regional 4DVar system over East Asia but only for May–June; similarly, the top-down optimization of Chinese NMVOC emissions by Feng et al. (2024) was limited to a single month (August 2022). Given the increasingly stringent air pollution control policies in China (Wu et al., 2024), there is an urgent need for high spatial- and temporal-resolution

top-down NMVOC emission optimization to support more accurate air quality forecasting and effective regulatory strategies. In terms of the observation sources for assimilation, the OMI formaldehyde product remains one of the most widely used datasets in related studies to date. However, it has been affected globally by the persistent row anomaly since 2007 Kroon et al. (2011); Zhu et al. (2017), which degrades data quality in certain across-track positions and may reduce assimilation accuracy, particularly in high-resolution configurations. Although rigorous quality filtering and row anomaly flagging can mitigate this problem, the number of valid grid cells remaining after such screening is often severely limited, rendering OMI data insufficient for nationwide high-resolution emission inversion. In contrast, newer-generation instruments such as TROPOMI and OMPS provide formaldehyde products that are unaffected by the row anomaly, offer significantly higher spatial resolution and better data coverage (Section 2.3.3), and are therefore considerably more suitable for high-resolution top-down optimization of NMVOC emissions over China.

~~In terms of the assimilation observation sources, OMI product now remains as one of the most widely used formaldehyde products in related studies. However, the retrievals over China are reported to suffer from the so-called "row anomaly" issue (González Abad et al., 2015), which may reduce the assimilation accuracy especially in the high-resolution configuration. To ensure the quality of OMI data, rigorous filtering is usually required; however, the number of valid grid cells remaining after such screening is severely limited, making it insufficient for high-resolution assimilation. In contrast, the TROPOMI and OMPS formaldehyde products do not suffer from this issue and are therefore more suitable for high-resolution emission optimization at the national scale. Subsequently, the monthly NMVOC emission optimization in China is conducted. This is achieved by independently assimilating formaldehyde observations either from OMPS or from TROPOMI, based on the emission inversion system that couples the four-dimensional ensemble variational (4DEnVar) data assimilation algorithm and GEOS-Chem model. The effectiveness of this emission inversion system has been evaluated in our recent studies (Jin et al., 2023; Jin et al., 2025). In this study, we focus on the year 2020 for the main analysis, while results for 2019 are also presented in the Supplementary Information to provide additional context and support.~~

In this study, we conduct monthly optimization of anthropogenic NMVOC emissions over China at 0.5° latitude $\times$ 0.625° longitude horizontal resolution. It is achieved based on an emission inversion system that couples the four-dimensional ensemble variational (4DEnVar) data assimilation algorithm with the nested version of the GEOS-Chem model. The effectiveness of this emission inversion system has been evaluated in our recent studies of ammonia (Jin et al., 2023; Xia et al., 2025). Two independent assimilation experiments are performed: one assimilating OMPS total formaldehyde columns and the other assimilating TROPOMI tropospheric formaldehyde columns. In both cases, the satellite retrievals have been harmonized with the model by replacing the original shape profiles with GEOS-Chem profiles before assimilation. We focus on the year 2020 for the main analysis, while results for 2019 are also presented in the Supplementary Information to provide additional context and support. This paper is organized as follows: Section 2 describes the dataset and methodology, focusing on GEOS-Chem model, input emission sources (anthropogenic, biogenic, and biomass burning), and the satellite and ground-based observations utilized. Section 3 provides an analysis of the assimilation results, including the estimation of posterior NMVOC emissions and the validation of both formaldehyde columns and ground-level ozone simulations. Section 4 summarizes the key findings and concludes the study.

References

Abad, G. G.: OMPS-N20 L2 NM Formaldehyde (HCHO) Total Column swath orbital V1, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC), , accessed: [Data Access Date], 2022.

Alvarado, L., Richter, A., Vrekoussis, M., Wittrock, F., Hilboll, A., Schreier, S., and Burrows, J.: An improved glyoxal retrieval from OMI measurements, Atmospheric Measurement Techniques, 7, 4133–4150, 2014.

Bates, K. H. and Jacob, D. J.: A new model mechanism for atmospheric oxidation of isoprene: global effects on oxidants, nitrogen oxides, organic products, and secondary organic aerosol, Atmospheric Chemistry and Physics, 19, 9613–9640, 2019.

Bates, K. H., Jacob, D. J., Wang, S., Hornbrook, R. S., Apel, E. C., Kim, M. J., Millet, D. B., Wells, K. C., Chen, X., Brewer, J. F., et al.: The global budget of atmospheric methanol: new constraints on secondary, oceanic, and terrestrial sources, Journal of Geophysical Research: Atmospheres, 126, e2020JD033 439, 2021.

Boersma, K., Eskes, H., and Brinksma, E.: Error analysis for tropospheric NO2 retrieval from space, Journal of Geophysical Research: Atmospheres, 109, 2004.

Burrows, J. P., Weber, M., Buchwitz, M., Rozanov, V., Ladstätter-Weißenmayer, A., Richter, A., DeBeek, R., Hoogen, R., Bramstedt, K., Eichmann, K.-U., et al.: The global ozone monitoring experiment (GOME): Mission concept and first scientific results, Journal of the Atmospheric Sciences, 56, 151–175, 1999.

Cao, H., Fu, T.-M., Zhang, L., Henze, D. K., Miller, C. C., Lerot, C., Abad, G. G., De Smedt, I., Zhang, Q., van Roozendael, M., et al.: Adjoint inversion of Chinese non-methane volatile organic compound emissions using space-based observations of formaldehyde and glyoxal, Atmospheric Chemistry and Physics, 18, 15 017–15 046, 2018.

Chaliyakunnel, S., Millet, D. B., and Chen, X.: Constraining emissions of volatile organic compounds over the Indian subcontinent using space-based formaldehyde measurements, Journal of Geophysical Research: Atmospheres, 124, 10 525–10 545, 2019.

Chan Miller, C., Gonzalez Abad, G., Wang, H., Liu, X., Kurosu, T., Jacob, D., and Chance, K.: Glyoxal retrieval from the ozone monitoring instrument, Atmospheric Measurement Techniques, 7, 3891–3907, 2014.

Chance, K., Palmer, P. I., Spurr, R. J., Martin, R. V., Kurosu, T. P., and Jacob, D. J.: Satellite observations of formaldehyde over North America from GOME, Geophysical Research Letters, 27, 3461–3464, 2000.

Chance, K., Liu, X., Miller, C. C., Abad, G. G., Huang, G., Nowlan, C., Souri, A., Suleiman, R., Sun, K., Wang, H., et al.: TEMPO green paper: Chemistry, physics, and meteorology experiments with the tropospheric emissions: Monitoring of pollution instrument, in: Sensors, systems, and next-generation satellites XXIII, vol. 11151, pp. 56–67, SPIE, 2019.

Choi, J., Henze, D. K., Cao, H., Nowlan, C. R., González Abad, G., Kwon, H.-A., Lee, H.-M., Oak, Y. J., Park, R. J., Bates, K. H., et al.: An inversion framework for optimizing non-methane VOC emissions using remote sensing and airborne observations in Northeast Asia during the KORUS-AQ field campaign, Journal of Geophysical Research: Atmospheres, 127, e2021JD035 844, 2022.

Clarisse, L., Franco, B., Van Damme, M., Di Gioacchino, T., Hadji-Lazaro, J., Whitburn, S., Noppen, L., Hurtmans, D., Clerbaux, C., and Coheur, P.: The IASI NH 3 version 4 product: averaging kernels and improved consistency, Atmospheric Measurement Techniques Discussions, 2023, 1–31, 2023.

Cooper, M. J., Martin, R. V., Henze, D. K., and Jones, D.: Effects of a priori profile shape assumptions on comparisons between satellite NO 2 columns and model simulations, Atmospheric Chemistry and Physics, 20, 7231–7241, 2020.

De Smedt, I., Müller, J.-F., Stavrakou, T., Van Der A, R., Eskes, H., and Van Roozendael, M.: Twelve years of global observations of formaldehyde in the troposphere using GOME and SCIAMACHY sensors, Atmospheric Chemistry and Physics, 8, 4947–4963, 2008.

De Smedt, I., Van Roozendael, M., Stavrakou, T., Müller, J.-F., Lerot, C., Theys, N., Valks, P., Hao, N., and Van Der A, R.: Improved retrieval of global tropospheric formaldehyde columns from GOME-2/MetOp-A addressing noise reduction and instrumental degradation issues, Atmospheric Measurement Techniques, 5, 2933–2949, 2012.

De Smedt, I., Stavrakou, T., Hendrick, F., Danckaert, T., Vlemmix, T., Pinardi, G., Theys, N., Lerot, C., Gielen, C., Vigouroux, C., et al.: Diurnal, seasonal and long-term variations of global formaldehyde columns inferred from combined OMI and GOME-2 observations, Atmospheric Chemistry and Physics, 15, 12 519–12 545, 2015.

De Smedt, I., Yu, H., Richter, A., Beirle, S., Eskes, H., Boersma, K., Van Roozendael, M., Van Geffen, J., Lorente, A., and Peters, E.: QA4ECV HCHO tropospheric column data from OMI (Version 1.1)[Data set], Royal Belgian Institute for Space Aeronomy, 2017.

De Smedt, I., Theys, N., Yu, H., Danckaert, T., Lerot, C., Compernolle, S., Van Roozendael, M., Richter, A., Hilboll, A., Peters, E., et al.: Algorithm theoretical baseline for formaldehyde retrievals from S5P TROPOMI and from the QA4ECV project, Atmospheric Measurement Techniques, 11, 2395–2426, 2018.

De Smedt, I., Pinardi, G., Vigouroux, C., Compernolle, S., Bais, A., Benavent, N., Boersma, F., Chan, K.-L., Donner, S., Eichmann, K.-U., et al.: Comparative assessment of TROPOMI and OMI formaldehyde observations and validation against MAX-DOAS network column measurements, Atmospheric Chemistry and Physics, 21, 12 561–12 593, 2021.

Dufour, G., Wittrock, F., Camredon, M., Beekmann, M., Richter, A., Aumont, B., and Burrows, J.: SCIAMACHY formaldehyde observations: constraint for isoprene emission estimates over Europe?, Atmospheric Chemistry and Physics, 9, 1647–1664, 2009.

Eskes, H. and Boersma, K.: Averaging kernels for DOAS total-column satellite retrievals, Atmospheric Chemistry and Physics, 3, 1285–1291, 2003.

Feng, S., Jiang, F., Qian, T., Wang, N., Jia, M., Zheng, S., Chen, J., Ying, F., and Ju, W.: Constraining non-methane VOC emissions with TROPOMI HCHO observations: impact on summertime ozone simulation in August 2022 in China, Atmospheric Chemistry and Physics, 24, 7481–7498, 2024.

Fu, D., Millet, D. B., Wells, K. C., Payne, V. H., Yu, S., Guenther, A., and Eldering, A.: Direct retrieval of isoprene from satellite-based infrared measurements, Nature communications, 10, 3811, 2019.

Fu, T.-M., Jacob, D. J., Palmer, P. I., Chance, K., Wang, Y. X., Barletta, B., Blake, D. R., Stanton, J. C., and Pilling, M. J.: Space-based formaldehyde measurements as constraints on volatile organic compound emissions in east and south Asia and implications for ozone, Journal of Geophysical Research: Atmospheres, 112, 2007.

González Abad, G., Liu, X., Chance, K., Wang, H., Kurosu, T., and Suleiman, R.: Updated Smithsonian astrophysical observatory ozone monitoring instrument (SAO OMI) formaldehyde retrieval, Atmospheric Measurement Techniques, 8, 19–32, 2015.

González Abad, G., Vasilkov, A., Seftor, C., Liu, X., and Chance, K.: Smithsonian astrophysical observatory ozone mapping and profiler suite (SAO OMPS) formaldehyde retrieval, Atmospheric Measurement Techniques, 9, 2797–2812, 2016.

Gulde, S., Kolm, M., Smith, D., Maurer, R., Courrèges-Lacoste, G. B., Sallusti, M., and Bagnasco, G.: Sentinel 4: a geostationary imaging UVN spectrometer for air quality monitoring: status of design, performance and development, in: International Conference on Space Optics—ICSO 2014, vol. 10563, pp. 1158–1166, SPIE, 2017.

He, M., Zheng, J., Yin, S., and Zhang, Y.: Trends, temporal and spatial characteristics, and uncertainties in biomass burning emissions in the Pearl River Delta, China, Atmospheric Environment, 45, 4051–4059, 2011.

Hua, W., Lou, S., Huang, X., Xue, L., Ding, K., Wang, Z., and Ding, A.: Diagnosing uncertainties in global biomass burning emission inventories and their impact on modeled air pollutants, Atmospheric Chemistry and Physics, 24, 6787–6807, 2024.

Jin, J., Fang, L., Li, B., Liao, H., Wang, Y., Han, W., Li, K., Pang, M., Wu, X., and Lin, H. X.: 4DEnVar-based inversion system for ammonia emission estimation in China through assimilating IASI ammonia retrievals, Environmental Research Letters, 18, 034 005, 2023.

Kaiser, J., Heil, A., Andreae, M., Benedetti, A., Chubarova, N., Jones, L., Morcrette, J.-J., Razinger, M., Schultz, M., Suttie, M., et al.: Biomass burning emissions estimated with a global fire assimilation system based on observed fire radiative power, Biogeosciences, 9, 527–554, 2012.

Kaiser, J., Jacob, D. J., Zhu, L., Travis, K. R., Fisher, J. A., González Abad, G., Zhang, L., Zhang, X., Fried, A., Crounse, J. D., et al.: High-resolution inversion of OMI formaldehyde columns to quantify isoprene emission on ecosystem-relevant scales: application to the southeast US, Atmospheric Chemistry and Physics, 18, 5483–5497, 2018.

Kim, J., Jeong, U., Ahn, M.-H., Kim, J. H., Park, R. J., Lee, H., Song, C. H., Choi, Y.-S., Lee, K.-H., Yoo, J.-M., et al.: New era of air quality monitoring from space: Geostationary Environment Monitoring Spectrometer (GEMS), Bulletin of the American Meteorological Society, 101, E1–E22, 2020.

Kroon, M., De Haan, J., Veefkind, J., Froidevaux, L., Wang, R., Kivi, R., and Hakkarainen, J.: Validation of operational ozone profiles from the Ozone Monitoring Instrument, Journal of Geophysical Research: Atmospheres, 116, 2011.

Kurosu, T., Chance, K., and Volkamer, R.: Global measurements of BrO, HCHO, and CHOCHO from the Ozone Monitoring Instrument on EOS Aura, in: AGU Fall Meeting Abstracts, vol. 2005, pp. A54B–01, 2005.

Kwon, H.-A., Park, R. J., González Abad, G., Chance, K., Kurosu, T. P., Kim, J., De Smedt, I., Van Roozendael, M., Peters, E., and Burrows, J.: Description of a formaldehyde retrieval algorithm for the Geostationary Environment Monitoring Spectrometer (GEMS), Atmospheric Measurement Techniques, 12, 3551–3571, 2019.

Kwon, H.-A., Park, R. J., Oak, Y. J., Nowlan, C. R., Janz, S. J., Kowalewski, M. G., Fried, A., Walega, J., Bates, K. H., Choi, J., et al.: Top-down estimates of anthropogenic VOC emissions in South Korea using formaldehyde vertical column densities from aircraft during the KORUS-AQ campaign, Elem Sci Anth, 9, 00 109, 2021.

Lerot, C., Stavrakou, T., De Smedt, I., Müller, J.-F., and Van Roozendael, M.: Glyoxal vertical columns from GOME-2 backscattered light measurements and comparisons with a global model, Atmospheric Chemistry and Physics, 10, 12 059–12 072, 2010.

Levelt, P. F., Van Den Oord, G. H., Dobber, M. R., Malkki, A., Visser, H., De Vries, J., Stammes, P., Lundell, J. O., and Saari, H.: The ozone monitoring instrument, IEEE Transactions on geoscience and remote sensing, 44, 1093–1101, 2006.

Levelt, P. F., Joiner, J., Tamminen, J., Veefkind, J. P., Bhartia, P. K., Stein Zweers, D. C., Duncan, B. N., Streets, D. G., Eskes, H., van der A, R., et al.: The Ozone Monitoring Instrument: overview of 14 years in space, Atmospheric Chemistry and Physics, 18, 5699–5745, 2018.

Li, C., Joiner, J., Krotkov, N. A., and Dunlap, L.: A new method for global retrievals of HCHO total columns from the Suomi National Polar-orbiting Partnership Ozone Mapping and Profiler Suite, Geophysical Research Letters, 42, 2515–2522, 2015.

Lin, H., Long, M. S., Sander, R., Sandu, A., Yantosca, R. M., Estrada, L. A., Shen, L., and Jacob, D. J.: An Adaptive Auto-Reduction Solver for Speeding Up Integration of Chemical Kinetics in Atmospheric Chemistry Models: Implementation and Evaluation in the Kinetic Pre-Processor (KPP) Version 3.0. 0, Journal of Advances in Modeling Earth Systems, 15, e2022MS003 293, 2023.

Lorente, A., Folkert Boersma, K., Yu, H., Dörner, S., Hilboll, A., Richter, A., Liu, M., Lamsal, L. N., Barkley, M., De Smedt, I., et al.: Structural uncertainty in air mass factor calculation for NO 2 and HCHO satellite retrievals, Atmospheric Measurement Techniques, 10, 759–782, 2017.

Millet, D. B., Jacob, D. J., Boersma, K. F., Fu, T.-M., Kurosu, T. P., Chance, K., Heald, C. L., and Guenther, A.: Spatial distribution of isoprene emissions from North America derived from formaldehyde column measurements by the OMI satellite sensor, Journal of Geophysical Research: Atmospheres, 113, 2008.

Nowlan, C. R., González Abad, G., Kwon, H.-A., Ayazpour, Z., Chan Miller, C., Chance, K., Chong, H., Liu, X., O'Sullivan, E., Wang, H., et al.: Global formaldehyde products from the Ozone Mapping and Profiler Suite (OMPS) nadir mappers on Suomi NPP and NOAA-20, Earth and Space Science, 10, e2022EA002 643, 2023.

Oomen, G.-M., Müller, J.-F., Stavrakou, T., De Smedt, I., Blumenstock, T., Kivi, R., Makarova, M., Palm, M., Röhling, A., Té, Y., et al.: Weekly derived top-down volatile-organic-compound fluxes over Europe from TROPOMI HCHO data from 2018 to 2021, Atmospheric Chemistry and Physics, 24, 449–474, 2024.

Opacka, B., Stavrakou, T., Müller, J.-F., De Smedt, I., van Geffen, J., Marais, E. A., Horner, R. P., Millet, D. B., Wells, K. C., and Guenther, A. B.: Natural emissions of VOC and NO x over Africa constrained by TROPOMI HCHO and NO2 data using the MAGRITTEv1. 1 model, Atmospheric Chemistry and Physics, 25, 2863–2894, 2025.

Palmer, P. I., Jacob, D. J., Chance, K., Martin, R. V., Spurr, R. J., Kurosu, T. P., Bey, I., Yantosca, R., Fiore, A., and Li, Q.: Air mass factor formulation for spectroscopic measurements from satellites: Application to formaldehyde retrievals from the Global Ozone Monitoring Experiment, Journal of Geophysical Research: Atmospheres, 106, 14 539–14 550, 2001.

Palmer, P. I., Jacob, D. J., Fiore, A. M., Martin, R. V., Chance, K., and Kurosu, T. P.: Mapping isoprene emissions over North America using formaldehyde column observations from space, Journal of Geophysical Research: Atmospheres, 108, 2003.

Platt, C.: Remote sounding of high clouds: I. Calculation of visible and infrared optical properties from lidar and radiometer measurements, Journal of Applied Meteorology and Climatology, 18, 1130–1143, 1979.

Shim, C., Wang, Y., Choi, Y., Palmer, P. I., Abbot, D. S., and Chance, K.: Constraining global isoprene emissions with Global Ozone Monitoring Experiment (GOME) formaldehyde column measurements, Journal of Geophysical Research: Atmospheres, 110, 2005.

Sinreich, R., Ortega, I., and Volkamer, R.: Sensitivity Study of Glyoxal Retrievals at Different Wavelength Ranges (Poster), in: 2013 International DOAS Workshop, 2013.

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., et al.: 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.

Stavrakou, T., Müller, J.-F., De Smedt, I., Van Roozendael, M., Kanakidou, M., Vrekoussis, M., Wittrock, F., Richter, A., and Burrows, J.: The continental source of glyoxal estimated by the synergistic use of spaceborne measurements and inverse modelling, Atmospheric Chemistry and Physics, 9, 8431–8446, 2009a.

Stavrakou, T., Müller, J.-F., De Smedt, I., Van Roozendael, M., Van Der Werf, G., Giglio, L., and Guenther, A.: Global emissions of non-methane hydrocarbons deduced from SCIAMACHY formaldehyde columns through 2003–2006, Atmospheric Chemistry and Physics, 9, 3663–3679, 2009b.

Stavrakou, T., Müller, J.-F., Bauwens, M., De Smedt, I., Lerot, C., Van Roozendael, M., Coheur, P.-F., Clerbaux, C., Boersma, K., Van Der A, R., et al.: Substantial underestimation of post-harvest burning emissions in the North China Plain revealed by multi-species space observations, Scientific Reports, 6, 32 307, 2016.

Thomas, W., Hegels, E., Meisner, R., Slijkhuis, S., Spurr, R., and Chance, K.: Detection of trace species in the troposphere using backscatter spectra obtained by the GOME spectrometer, in: IGARSS'98. Sensing and Managing the Environment. 1998 IEEE International Geoscience and Remote Sensing. Symposium Proceedings.(Cat. No. 98CH36174), vol. 5, pp. 2612–2614, IEEE, 1998.

Van Der Werf, G. R., Randerson, J. T., Giglio, L., Van Leeuwen, T. T., Chen, Y., Rogers, B. M., Mu, M., Van Marle, M. J., Morton, D. C., Collatz, G. J., et al.: Global fire emissions estimates during 1997–2016, Earth System Science Data, 9, 697–720, 2017.

Vigouroux, C., Langerock, B., Bauer Aquino, C. A., Blumenstock, T., Cheng, Z., De Mazière, M., De Smedt, I., Grutter, M., Hannigan, J. W., Jones, N., et al.: TROPOMI–Sentinel-5 Precursor formaldehyde validation using an extensive network of ground-based Fourier-transform infrared stations, Atmospheric Measurement Techniques, 13, 3751–3767, 2020.

Wells, K., Millet, D., Payne, V., Vigouroux, C., Aquino, C., De Mazière, M., de Gouw, J., Graus, M., Kurosu, T., Warneke, C., et al.: Next-generation isoprene measurements from space: Detecting daily variability at high resolution, Journal of Geophysical Research: Atmospheres, 127, e2021JD036 181, 2022.

Wells, K. C., Millet, D. B., Payne, V. H., Deventer, M. J., Bates, K. H., de Gouw, J. A., Graus, M., Warneke, C., Wisthaler, A., and Fuentes, J. D.: Satellite isoprene retrievals constrain emissions and atmospheric oxidation, Nature, 585, 225–233, 2020.

Wiedinmyer, C., Akagi, S., Yokelson, R. J., Emmons, L., Al-Saadi, J., Orlando, J., and Soja, A.: The Fire INventory from NCAR (FINN): A high resolution global model to estimate the emissions from open burning, Geoscientific Model Development, 4, 625–641, 2011.

Wittrock, F., Richter, A., Oetjen, H., Burrows, J. P., Kanakidou, M., Myriokefalitakis, S., Volkamer, R., Beirle, S., Platt, U., and Wagner, T.: Simultaneous global observations of glyoxal and formaldehyde from space, Geophysical Research Letters, 33, 2006.

Wu, N., Geng, G., Xu, R., Liu, S., Liu, X., Shi, Q., Zhou, Y., Zhao, Y., Liu, H., Song, Y., et al.: Development of a high-resolution integrated emission inventory of air pollutants for China, Earth System Science Data, 16, 2893–2915, 2024.

Xia, J., Zhou, Y., Fang, L., Qi, Y., Li, D., Liao, H., and Jin, J.: South Asia anthropogenic ammonia emission inversion through assimilating IASI observations, Atmospheric Chemistry and Physics, 25, 7071–7086, , URL https://acp.copernicus.org/articles/25/7071/2025/, 2025.

Zhang, C., Liu, C., Hu, Q., Cai, Z., Su, W., Xia, C., Zhu, Y., Wang, S., and Liu, J.: Satellite UV-Vis spectroscopy: implications for air quality trends and their driving forces in China during 2005–2017, Light: Science & Applications, 8, 100, 2019.

Zhu, L., Jacob, D. J., Mickley, L. J., Marais, E. A., Cohan, D. S., Yoshida, Y., Duncan, B. N., Abad, G. G., and Chance, K. V.: Anthropogenic emissions of highly reactive volatile organic compounds in eastern Texas inferred from oversampling of satellite (OMI) measurements of HCHO columns, Environmental Research Letters, 9, 114 004, 2014.

Zhu, L., Jacob, D. J., Kim, P. S., Fisher, J. A., Yu, K., Travis, K. R., Mickley, L. J., Yantosca, R. M., Sulprizio, M. P., De Smedt, I., et al.: Observing atmospheric formaldehyde (HCHO) from space: validation and intercomparison of six retrievals from four satellites (OMI, GOME2A, GOME2B, OMPS) with SEAC 4 RS aircraft observations over the southeast US, Atmospheric chemistry and physics, 16, 13 477–13 490, 2016.

Zhu, L., Jacob, D. J., Keutsch, F. N., Mickley, L. J., Scheffe, R., Strum, M., González Abad, G., Chance, K., Yang, K., Rappenglück, B., et al.: Formaldehyde (HCHO) as a hazardous air pollutant: Mapping surface air concentrations from satellite and inferring cancer risks in the United States, Environmental Science & Technology, 51, 5650–5657, 2017.

Zhu, L., González Abad, G., Nowlan, C. R., Chan Miller, C., Chance, K., Apel, E. C., DiGangi, J. P., Fried, A., Hanisco, T. F., Hornbrook, R. S., et al.: Validation of satellite formaldehyde (HCHO) retrievals using observations from 12 aircraft campaigns, Atmospheric Chemistry and Physics, 20, 12 329–12 345, 2020.