Author Contributions:

Conceptualization: Weitao Fu, Lei Zhu
Data curation: Weitao Fu, Yali Li
Formal analysis: Weitao Fu
Funding acquisition: Lei Zhu, Hyeong-Ahn Kwon
Investigation: Weitao Fu, Lei Zhu
Methodology: Weitao Fu, Lei Zhu, Hyeong-Ahn Kwon, Rokjin J. Park, Gitaek T. Lee, Song Liu, Xicheng Li, Yuyang Chen, Xiaoxing Zuo, Peng Zhang, Yali Li, Zhuoxian Yan
Project administration: Lei Zhu
Resources: Lei Zhu
Software: Weitao Fu, Lei Zhu
Supervision: Lei Zhu, Isabelle De Smedt, Song Liu, Huizhong Shen, Jianhuai Ye, Chen Wang, Tzung-May Fu, Xin Yang
Validation: Weitao Fu

Evaluating GEMS HCHO Retrievals With TROPOMI Product, Pandora Observations, and GEOS-Chem Simulations

Weitao Fu[1] , Lei Zhu[1,2,3] , Hyeong-Ahn Kwon[4] , Rokjin J. Park[5] , Gitaek T. Lee[5] , Isabelle De Smedt[6] , Song Liu[1] , Xicheng Li[1] , Yuyang Chen[1] , Dongchuan Pu[1] , Juan Li[1] , Xiaoxing Zuo[1] , Peng Zhang[1] , Yali Li[1] , Zhuoxian Yan[1] , Xue Zhang[1] , Jiaming Zhang[1] , Xingyi Wu[1] , Huizhong Shen[1,2,3] , Jianhuai Ye[1,2,3] , Chen Wang[1,2,3] , Tzung-May Fu[1,2,3] , and Xin Yang[1,2,3]

[1]School of Environmental Science and Engineering, Southern University of Science and Technology, Shenzhen, China, [2]Coastal Atmosphere and Climate of the Greater Bay Area Observation and Research Station of Guangdong Province, Southern University of Science and Technology, Shenzhen, China, [3]Shenzhen Key Laboratory of Precision Measurement and Early Warning Technology for Urban Environmental Health Risks, School of Environmental Science and Engineering, Southern University of Science and Technology, Shenzhen, China, [4]Department of Environmental & Energy Engineering, University of Suwon, Hwaseong-si, Republic of Korea, [5]School of Earth and Environmental Sciences, Seoul National University, Seoul, Republic of Korea, [6]Royal Belgian Institute for Space Aeronomy (BIRA-IASB), Brussels, Belgium

Abstract Satellite column formaldehyde (HCHO) is an indicator of regional volatile organic compounds (VOC) emissions as HCHO is a short-lived intermediate oxidation product. The Geostationary Environment Monitoring Spectrometer (GEMS), launched in 2020, is the first geostationary satellite to monitor hourly HCHO. GEMS offers unprecedented potential to reveal the diurnal variations of VOC emissions in Asia. Here, we present the first study to evaluate year-round GEMS HCHO retrievals using TROPOMI satellite and ground-based Pandora spectrometers. Our study shows that GEMS HCHO aligns with TROPOMI ($r = 0.59$–0.85; differences within 20% for most areas). Moreover, GEMS captures monthly and diurnal HCHO variations observed by Pandora spectrometers across Asia with differences overall within 15% ($r \sim 0.85$). Diurnally, we find strong HCHO variations over urban areas but not in forests. During the fire season of mainland Southeast Asia, GEMS HCHO increases in the afternoon, in line with diurnal emission estimates from the Global Fire Emissions Database Version 4 with small fires (GFED4s) and GEOS-Chem simulations. GEMS also captures the spatial patterns of fire emissions in GFED4s. GEMS HCHO shows negative bias when observing with a high ($>60°$) viewing zenith angle (VZA) and overly relies on model correction for observations to the north of 30°N.

Plain Language Summary Formaldehyde (HCHO) is an intermediate oxidation product of most VOCs. Because of its short lifetime, studies have used satellite HCHO observations as a proxy of regional VOC emissions. Previous HCHO-observing satellites are all onboard low-earth-orbit satellites, which can only provide daily observations. Here, we evaluate hourly HCHO observations from the first geostationary air quality satellite, GEMS, based on ground-based observations and previous satellite observations. Overall, we find diurnal HCHO observations from GEMS align with validation data sets in China, Korea, and mainland Southeast Asia. HCHO diurnal variations are strong over megacities, indicating dynamic VOC emissions.

1. Introduction

VOCs pose threats to human health directly and contribute to the formation of tropospheric ozone and secondary organic aerosols (Ehn et al., 2014; Gu et al., 2021; Sarigiannis et al., 2011; Shao et al., 2009). Understanding diurnal VOC emissions is critical for developing effective pollution control strategies (Coggon et al., 2018; Liu et al., 2015; Menchaca-Torre et al., 2015). However, diverse VOC species hinder us from observing them comprehensively, especially on a regional scale. HCHO, the short-lived intermediate oxidation product of VOCs, is used as a proxy of local VOC emissions (Green et al., 2021; Hong et al., 2021). Here, we examine HCHO observations from the first geostationary satellite for trace gases, the Geostationary Environment Monitoring Spectrometer (GEMS).

The past two decades have witnessed the development of space-based HCHO retrievals and their broad presence in air quality studies. Palmer et al. (2003) presented the first study to map isoprene emissions based on HCHO

Visualization: Weitao Fu
Writing – original draft: Weitao Fu
Writing – review & editing: Weitao Fu,
Lei Zhu, Hyeong-Ahn Kwon, Rokjin
J. Park, Gitaek T. Lee, Isabelle De Smedt,
Song Liu, Xicheng Li, Yuyang Chen,
Dongchuan Pu, Juan Li, Xiaoxing Zuo,
Peng Zhang, Yali Li, Zhuoxian Yan,
Xue Zhang, Jiaming Zhang, Xingyi Wu,
Huizhong Shen, Jianhuai Ye, Chen Wang,
Tzung-May Fu, Xin Yang

observations from the Global Ozone Monitoring Experiment (GOME; Burrows et al., 1999) satellite instrument. Since then, HCHO observations have been retrieved from a series of satellite instruments, including SCanning Imaging Absorption spectroMeter for Atmospheric CHartographY (SCIAMACHY; De Smedt et al., 2008), Global Ozone Monitoring Experiment-2 (GOME-2; De Smedt et al., 2012), Ozone Monitoring Instrument (OMI; González Abad et al., 2015), Ozone Mapping and Profiler Suite (OMPS; Nowlan et al., 2023), TROPOspheric Monitoring Instrument (TROPOMI; De Smedt et al., 2018), and Environmental Trace Gases Monitoring Instrument (EMI; Su et al., 2022). Previous studies have used HCHO observations from these low-earth-orbit (LEO) satellites to investigate global VOC emissions (Cao et al., 2018; Choi et al., 2022; Sun et al., 2021), surface ozone formation sensitivity (Duncan et al., 2010; Ren et al., 2022; Travis et al., 2022), and the abundance of hydroxyl radical and secondary organic aerosols (Anderson et al., 2023; Liao et al., 2019; Wolfe et al., 2019). However, one limitation of these LEO satellites is that they provide only 1–2 observations every day over a specific location.

As the first air quality geostationary satellite, GEMS addresses such a limitation with hourly HCHO observations in the daytime. Here, we present the first study to characterize year-round diurnal variations of GEMS HCHO over different emission sources, including urban areas, forests, and fires. We evaluate GEMS HCHO against TROPOMI retrievals, ground-based observations from Pandora spectrometers, and state-of-the-art GEOS-Chem simulations.

2. Data and Methods

2.1. GEMS HCHO Retrievals

GEMS is a UV-visible hyperspectral spectrometer launched in 2020 onboard the GK-2B satellite (Choi et al., 2018; Kim et al., 2020). Sitting on a geostationary orbit, it provides ∼8 hourly observations each day across Asia with stable VZA (Figure S1 in Supporting Information S1). Each hour, it operates with one out of four different observation domains, scanning Asia from east to west and lasting around 30 min (*e.g.*, 9:45 to 10:15) (Figure S2 in Supporting Information S1). GEMS mostly operates with full central or full west modes, both of which cover China, Korea, and Southeast Asia.

There are 3 main steps in the operational GEMS HCHO Version 2 retrieval algorithm (Kwon et al., 2019; Lee et al., 2024), namely spectral fitting, background correction, and conversion from slant column density (SCD) to vertical column density (VCD). First, it retrieves the differential slant column density (DSCD) of HCHO relative to its reference area (120–150°E) of the same latitude to reduce fitting uncertainties, based on a nonlinear solar radiance fitting method named basic optical absorption spectroscopy (BOAS; Chance et al., 2000). Then, GEOS-Chem is run to simulate the VCD over the reference area (VCD_{ref}), which is then converted to SCD_{ref} with air mass factor (AMF) over the reference area (AMF_{ref}), by accounting for the observing geometries, atmospheric scattering, and clouds. Finally, the total VCD is computed with regional AMF, as shown in Equation 1.

$$VCD = \frac{DSCD + VCD_{ref} \times AMF_{ref}}{AMF} \quad (1)$$

Therefore, the contribution from spectral fitting (DVCD) and background correction (VCD_0) is as follows:

$$DVCD = \frac{DSCD}{AMF} \quad (2)$$

$$VCD_0 = \frac{VCD_{ref} \times AMF_{ref}}{AMF} \quad (3)$$

We obtain the GEMS HCHO Version 2 data set from the Korean Environmental Satellite Center (https://nesc.nier.go.kr, last access: 2/1/2024). Lee et al. (2024) find that GEMS HCHO columns show strong correlations with HCHO satellite retrieval from TROPOMI during the first year of GEMS operation, including the in-orbit test period (August-October 2020), both spatially ($r = 0.62/0.9$ for the entire domain/north-eastern Asia) and temporally (monthly $r = 0.58–0.82$ in major cities). GEMS also reproduces diurnal HCHO variations from ground-based observations from multi-axis differential optical absorption spectroscopy (MAX-DOAS; $r = 0.79$) and Fourier transform infrared (FTIR; $r = 0.85$) instruments in Xianghe (116.96°E, 39.75°N). This study only

uses observations that pass its official quality control (*FinalAlgorithmFlags* = 0) with a low cloud fraction (<30%) and solar zenith angle (SZA; <60°).

2.2. TROPOMI HCHO Retrievals

TROPOMI is onboard a low-earth-orbit Sentinel-5P satellite (Veefkind et al., 2012). Launched in 2017, its operational HCHO product (De Smedt et al., 2018) has gone through rigorous validation against ground-based HCHO observations from MAX-DOAS and FTIR instruments, outperforming previous satellites (De Smedt et al., 2021; Vigouroux et al., 2020), and have been widely used for spatiotemporal analysis (Levelt et al., 2022; Pu et al., 2022; Sun et al., 2021). We also apply the official quality control (*qa_value* > 0.5) and same cloud and SZA filters. Considering TROPOMI has an overpassing time of ∼13:30 local time, GEMS observations around either 13:00 or 14:00 are used for intercomparison. Here we note that GEMS HCHO is the total VCD, while TROPOMI HCHO only includes tropospheric HCHO. Nevertheless, GEOS-Chem simulations show that tropospheric HCHO constitutes ∼97% of the total VCD on average (Figure S3 in Supporting Information S1).

2.3. Ground-Based Pandora HCHO Observations

We obtain HCHO observations of 19 stations from the Pandonia Global Network (https://www.pandonia-global-network.org/, last access: 9/22/2024). Pandora spectrometers operate with various modes, including direct-sun and multi-axis (Kreher et al., 2020; Pinardi et al., 2013; Zhao et al., 2019). Spinei et al. (2018) and Herman et al. (2018) reported that the direct-sun Pandora HCHO product is overall larger (+16%) than aircraft in-situ measurements and two times of OMI satellite product. Spinei et al. (2021) later found that those early instruments used polyoxymethylene (POM-H Delrin) that releases HCHO, while Pandora spectrometers manufactured after summer 2019 solved this issue; they also suggested that Pandora HCHO columns retrieved with multi-axis measurements are less impacted by this internal HCHO releases compared to direct-sun measurements. We note that all Pandora spectrometers used here do not contain POM-H Delrin. Previous studies have used Pandora to evaluate both OMI (HCHO and NO_2) and GEMS (NO_2) (Chang et al., 2022; Herman & Mao, 2024; Kim et al., 2023; Lange et al., 2024).

This study uses both Pandora direct-sun total (fus5) and multi-axis tropospheric (fuh5) HCHO products (Luft-Blick, 2023), with the following flags: both *L1_data* and *L2_fit* = 0 or 10 (high quality); *L2_data* = 0 or 1 or 10 or 11, which are "high/medium quality" data. We use all useable data, including those whose *L2_data* are labeled with "unassured", or "medium" quality, to maximize data coverage, as a recent study by Rawat et al. (2024) suggests that some Pandora HCHO observations flagged with "low" quality are also valid.

To obtain hourly observations, we calculate median Pandora HCHO observations within ±15 min of hourly GEMS observations similar to Judd et al. (2020). All GEMS observations falling into 0.1° × 0.1° pixel centered at each Pandora station are then used to calculate the averaged value based on inverse distance weighting (Bartier & Keller, 1996). Finally, for each site, we filter out GEMS HCHO outliers that are larger than the third quartile plus 1.5 interquartile range (IQR) or less than the first quartile minus 1.5 IQR. Their locations and total numbers of coincident hourly observations are in Table S1 and Figure S1 in Supporting Information S1.

2.4. GEOS-Chem Simulations

We simulate hourly HCHO profiles using GEOS-Chem 14.1.1 (Yantosca et al., 2023). GEOS-Chem is a widely used global three-dimensional chemical transport model (Bey et al., 2001), with state-of-the-art meteorological reanalysis, emission inventories, and chemical mechanisms to simulate global distributions of chemical species. In this study, we first run 4° × 5° classic GEOS-Chem using the MERRA-2 data (Gelaro et al., 2017) as meteorological inputs to obtain boundary conditions. Then, these boundary conditions are used in the finer 0.5° × 0.625° nested simulations for Asia (11°S–55°N, 60°E−150°E). We use biogenic, anthropogenic, and pyrogenic emissions from the Model of Emissions of Gases and Aerosols from Nature (MEGAN), Community Emissions Data System (CEDS), and GFED4s, respectively (Guenther et al., 2012; Hoesly et al., 2018; McDuffie et al., 2020; Randerson et al., 2017), to be consistent with GEOS-Chem simulations used for producing GEMS *a-priori* HCHO profiles (Lee et al., 2024). Note that the CEDS used in this study is based on 2019, because the CEDS inventory for 2021 was released in April 2024 and hasn't been processed to be used for GEOS-Chem simulations as of 1 June 2024. According to its release note, the change in global non-methane VOC emission is modest (https://github.com/JGCRI/CEDS/wiki/Release-Notes).

To better intercompare GEMS and GEOS-Chem, we apply GEMS averaging kernels (*AK*) on GEOS-Chem simulated HCHO. *AK* accounts for the sensitivities of GEMS HCHO retrieval to the vertical distributions of HCHO (Kwon et al., 2019; Rodgers & Connor, 2003). The first step is to calculate the *AK* of GEMS at different vertical layers (*i*) based on its scattering weights (*w*) and AMF (Equation 4). Then, we apply *AK* as weights to the GEOS-Chem simulated partial column (ΔVCD) at different vertical layers (Equation 5). The final AK-smoothed GEOS-Chem VCD (VCD$_{GC}$) is the sum of AK-smoothed partial columns (ΔVCD$_{AK}$) (Equation 6).

$$AK(i) = \frac{w(i)}{AMF} \tag{4}$$

$$\Delta VCD_{AK}(i) = \Delta VCD(i) \times AK(i) \tag{5}$$

$$VCD_{GC} = \sum \Delta VCD_{AK}(i) \tag{6}$$

2.5. Sampling GEMS HCHO From Different Areas

First, we focus on 20 major cities with high HCHO columns (shown in Figure 7). City boundaries are extracted from Food and Agriculture Organization Global Administrative Unit Layers (FAO, 2015). Because the administrative boundaries of some cities (*e.g.*, Chongqing) contain a lot of rural areas, we further remove areas where the annual mean NO$_2$ VCD is less than 3.0×10^{15} molecules cm^{-2} based on TROPOMI observations (van Geffen et al., 2022), as NO$_2$ VCD in major cities often exceeds this threshold (Fioletov et al., 2022).

Then, we investigate four $2° \times 2°$ forest areas shown in Global Land Cover and Land Use Change data set (Potapov et al., 2022a), centered at Borneo Islands (2°N, 115°E), Malaysia (5°N, 102°E), the Philippines (18°N, 121°E), and North Korea (41°N, 128°E), where year-round fire emissions (Figure S4 in Supporting Information S1; <1 g C m^{-2} month^{-1}) and population density (Figure S5 in Supporting Information S1; <100 persons km^{-2}) are relatively low, according to GFED4s (Randerson et al., 2017; van der Werf et al., 2017) and Gridded Population of the World Version 4 (Center for International Earth Science Information Network, 2018). Here we examine temporal variations of biogenic VOC emission estimates and temperatures based on CAMS-GLOB-BIO v3.1 (Sindelarova et al., 2022) and ERA5 reanalysis (Hersbach et al., 2020), as biogenic VOC emissions are dependent on temperature (Huang et al., 2018; Saunier et al., 2017).

For fire-affected areas, we focus on mainland Southeast Asia from January to April, because of the strong fire emissions according to GFED4s. First, we conduct several sensitivity simulations with $4° \times 5°$ GEOS-Chem to investigate the impacts of fire. Then, we calculate the mass of monthly total fire-induced HCHO (m_{HCHO}; Equation 7) based on the mass (*m*) of burned dry matter from different land cover types (*i*), which is provided by GFED4s inventory, and emission factors (*k*) to HCHO documented in Akagi et al. (2011). These emission factors are measured after smoke cooled down to ambient temperature without undergoing significant photochemical processes. Finally, we compare the spatial patterns of change (2022–2021) of GEMS HCHO at different hours with GFED4s HCHO change in March (as the simulations show fire emissions dominate interannual HCHO variations).

$$m_{HCHO} = \sum_i m_i \times k_i \tag{7}$$

3. Intercomparison and Validation of GEMS HCHO Columns

3.1. Intercomparison With TROPOMI

Figure 1 shows the comparison between GEMS and TROPOMI HCHO. Overall, GEMS HCHO VCD aligns with TROPOMI in Korea, mainland Southeast Asia, and major parts of China (spatial $r = 0.59$–0.85). Although both products retrieve DSCD relative to their reference area, the reference area of GEMS (120°–150°E, including some highly industrialized regions in Korea, Japan, eastern parts of China, and nearby seas) is not as clean as that of TROPOMI (Pacific Ocean; De Smedt et al., 2018; Kwon et al., 2019; Lee et al., 2024). Consequently, GEMS has lower DVCD but higher VCD$_0$ (Figure S6 and S7 in Supporting Information S1). GEMS VCD$_0$ also shows stronger latitudinal variations, which leads to systematic latitudinal differences in Figure 1 and S8 in Supporting Information S1 (GEMS minus TROPOMI). Regardless, differences in total VCD are within 20% in most polluted

areas. In some low-HCHO regions (*e.g.*, Tibet Plateau), the absolute differences (Figure S8 in Supporting Information S1) are within $\pm 2.5 \times 10^{15}$ molecules cm^{-2} despite relative differences deviating from 100% (Figure 1).

Major discrepancies exist in the west and south of the GEMS domain. The underestimation in the west domain shows a similar pattern with GEMS VZA (Figure S1 in Supporting Information S1), which is stable because GEMS is a geostationary satellite. GEMS starts to show lower HCHO VCD when VZA >50° (light purple line in Figure 1) and is less than half of TROPOMI when VZA >60° (dark purple line in Figure 1). This indicates that current retrieval algorithm (spectral fitting and AMF) under high VZA needs improvements; specifically, it should better account for the impacts of atmospheric scattering because of the long light path (Lee et al., 2024). As for the positive bias near the equator during summer, we find it is due to both DSCD fitting and background correction (Figure S6 in Supporting Information S1).

3.2. Validation With Pandora Observations

This section shows the comparison between coincident GEMS and Pandora direct-sun/multi-axis HCHO observations. Here we consider Pandora multi-axis observations as the benchmark, as the multi-axis method is a well-established method to retrieve HCHO and is widely used by most previous studies to evaluate satellite HCHO (Chan et al., 2020; Chong et al., 2024; Kumar et al., 2020; Ryan et al., 2023). In addition, Frieß et al. (2019) intercompared different multi-axis algorithms and suggested that the one used in Pandora HCHO retrieval is accurate.

As shown in Figure 2, GEMS total HCHO VCD is overall only slightly lower (−6%) than Pandora multi-axis tropospheric VCD during summer and spring, and 8% higher in winter. Considering GEOS-Chem simulations show that tropospheric HCHO constitutes ~97% of the total VCD in these Pandora sites (Figure S3 in Supporting Information S1), GEMS may slightly underestimate high HCHO VCD.

However, Pandora direct-sun HCHO is overall ~45% higher than GEMS (Figure S9 in Supporting Information S1). This aligns with the second Cabauw Intercomparison of Nitrogen Dioxide measuring Instruments (CINDI-2), which suggested that direct-sun HCHO products are less consistent and higher compared to multi-axis HCHO products (Tirpitz et al., 2021). A recent study also reported that Pandora direct-sun HCHO product is about twice of OMI satellite observations (Herman & Mao, 2024). Since HCHO above the tropopause is not sufficient to explain the differences between Pandora multi-axis and direct-sun HCHO, the Pandora direct-sun product likely overestimates HCHO.

Individual scatterplots of all coincident hourly (methods in Section 2.3) GEMS and Pandora multi-axis/direct-sun observations can be found in Figure S10 and S11 in Supporting Information S1. For Pandora multi-axis HCHO, clear linear relationships with GEMS ($r \sim 0.5$, slope ~1) are found in most sites, except for Dalanzadgad in Mongolia, a site located in the northwestern edge of the GEMS domain (VZA ~ 55°). The Pandora direct-sun HCHO also shows similar linear relationships with GEMS but is generally higher.

In Figure 3, monthly HCHO variations of GEMS align with coincident Pandora multi-axis observations ($r > 0.8$ for most stations). GEMS shows stronger underestimations in more polluted regions (*e.g.*, ~−16% in Bangkok and Beijing). Similarly, Vigouroux et al. (2020) and De Smedt et al. (2021) also found TROPOMI HCHO is systematically ~25% lower than ground-based observations in the polluted area (HCHO VCD is higher than 8.0×10^{15} molecules cm^{-2}). The negative bias may be partly attributed to its weaker sensitivity to near-surface HCHO compared to Pandora. It is worth noting that, the bias of GEMS relative to Pandora multi-axis (direct-sun) HCHO is calculated with coincident multi-axis (direct-sun) observations. For visualization purposes, GEMS monthly (and diurnal) variations shown are calculated with all coincident (either with multi-axis or direct-sun) observations.

We find GEMS captures summertime diurnal HCHO variations ($r = 0.67$–0.98; Figure 4) in cities such as (b) Bangkok and (c) Beijing. Two Pandora HCHO products show close diurnal patterns but differ in magnitude. Figure S12–S14 in Supporting Information S1 also show their comparisons in other seasons during which the diurnal variations are less pronounced; overall they show a r of ~0.8 where the diurnal variations exist. Meanwhile, in less polluted cities, both GEMS and Pandora show weak diurnal HCHO variations, such as (e) Dalanzadgad and (n) Singapore. The discrepancies in (o) Tokyo and (p, q, and r) Tsukuba are likely linked to different observation domains of each hour, which impacts the reference area. In addition, large differences
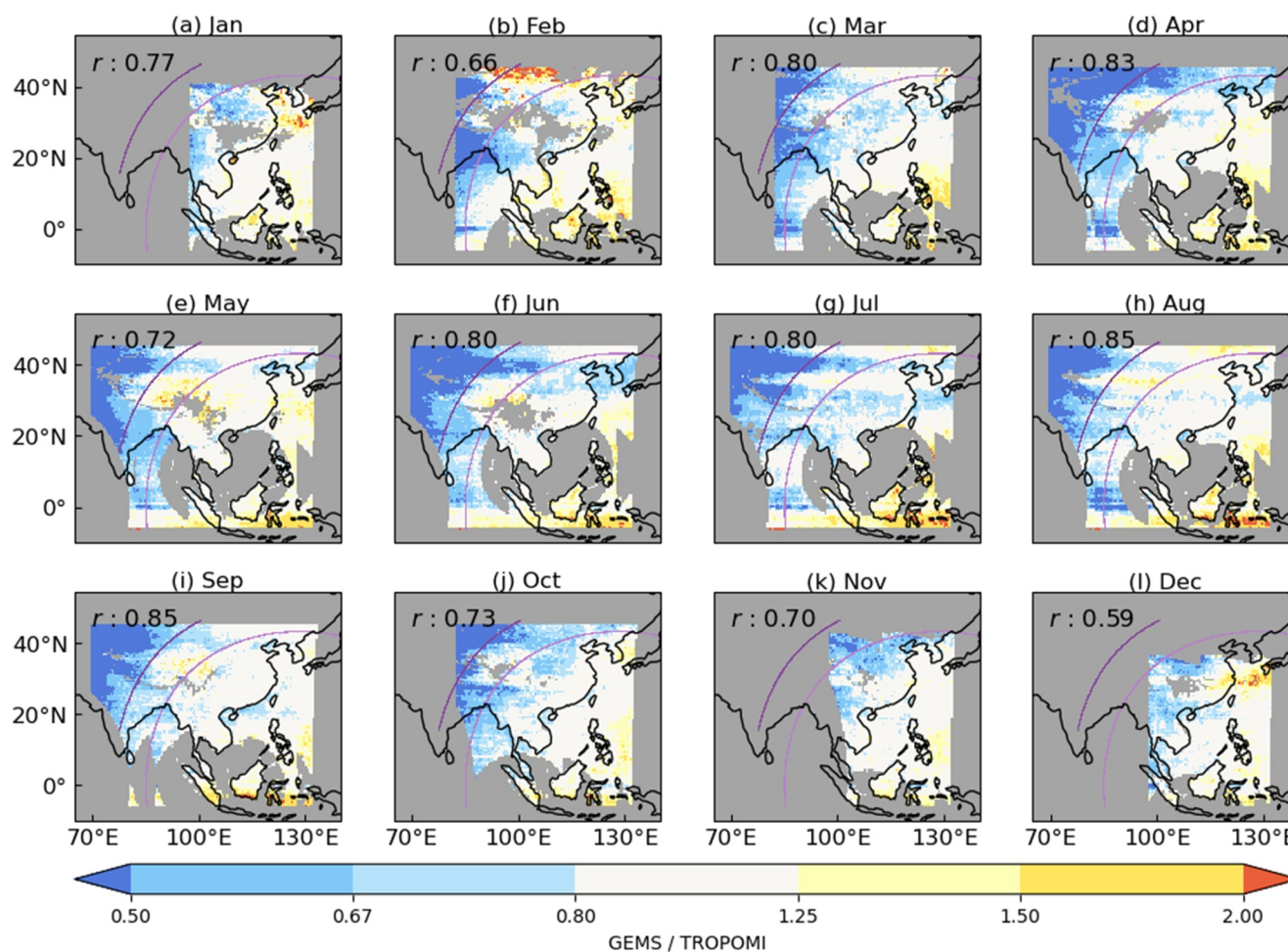
Figure 1. Ratios of monthly mean GEMS HCHO vertical column density (observed during 13:45–14:15 local time) to TROPOMI (~13:30) in 2022. Dark and light purple lines indicate regions where GEMS VZA is 60° and 50° respectively. Both GEMS and TROPOMI HCHO observations are gridded to 0.5° × 0.5° using the oversampled method described in Zhu et al. (2017). Correlation coefficients (*r*) are calculated with colocated GEMS-TROPOMI observations where the GEMS VZA is less than 60°.

between the two Pandora sites in Seoul are likely due to ground-based instruments being much more sensitive to their surroundings than satellite observations. Biogenic emissions from nearby mountains may be one of the potential explanations for strong afternoon enhancement in Seoul-SNU. It also should be noted that the diurnal variations shown here are not as representative as Section 4 because only coincident observations are used for validation purposes.

3.3. Comparisons With GEOS-Chem HCHO Columns

Figure 5 shows an example of the effect of averaging kernels near Hong Kong in summer; applying averaging kernels slightly improves its agreement with GEMS (Figure 5b; smaller RMSE). We choose Hong Kong as it is a southern metropolitan city, where the model contribution (VCD_0) to the total GEMS VCD is weak (details in Section 4.1). Notably, diurnal HCHO simulated by GEOS-Chem in all 3 months always follow a similar pattern; it peaks in the noon during summer; while GEMS observations show much larger variations across 3 months. This indicates that the current model simulation may not be able to fully characterize sophisticated anthropogenic environments. Potential reasons include but are not limited to its inventories, characterizations of diurnal variations of atmospheric oxidants, and planetary boundary layer dynamics (Li et al., 2023; Yang et al., 2024). For instance, the latest CEDS anthropogenic emission inventory available for our GEOS-Chem simulations is based on 2019; its default description of the diurnal emission variation may also not be fully representative in Asia.
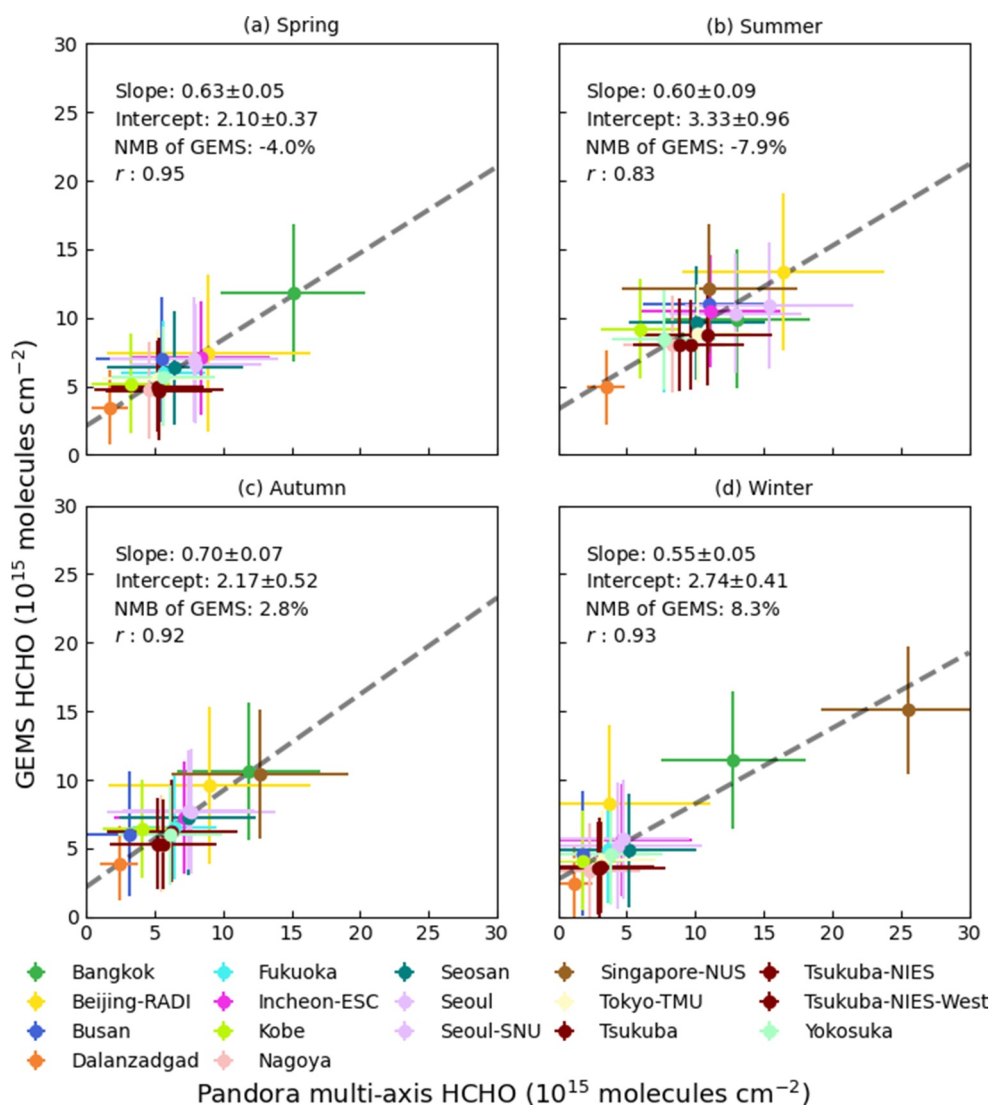
Figure 2. Scatterplots of mean GEMS and Pandora multi-axis HCHO VCD in each season during 2021–2023. Each point represents a Pandora site. The error bar shows the standard deviations of coincident observations. The dashed line is based on the reduced major axis regression. Key statistics are also shown in each plot.

Likewise, we calculate these statistics for each pixel in each of the four seasons (Figure 6). First, their correlations, as shown in the first column, are above 0.8 in most regions. However, GEOS-Chem does not correlate with GEMS in areas below 30°N, especially in summer and winter, which can be attributed to the background correction of GEMS (details in Section 4.1). Second, the slopes (second column) from their linear regressions are near one in major polluted regions; on the one hand, HCHO VCD from both GEOS-Chem and TROPOMI is higher than GEMS in the southwestern domain; on the other hand, GEOS-Chem is much lower than GEMS over Tibet Plateau, where GEMS and TROPOMI observations are close (Figure 1; S8 in Supporting Information S1). Third, applying averaging kernels has different impacts on the agreements; for example, it caused a lower fitted RMSE in India but a higher RMSE in eastern Chinese cities.

4. GEMS HCHO Variations

4.1. Urban Areas

As shown in Figure 7, GEMS effectively captures anthropogenic enhancements of HCHO across Asia. Each city has its distinct HCHO diurnal variations in different seasons. It should be aware that even with strict criteria to
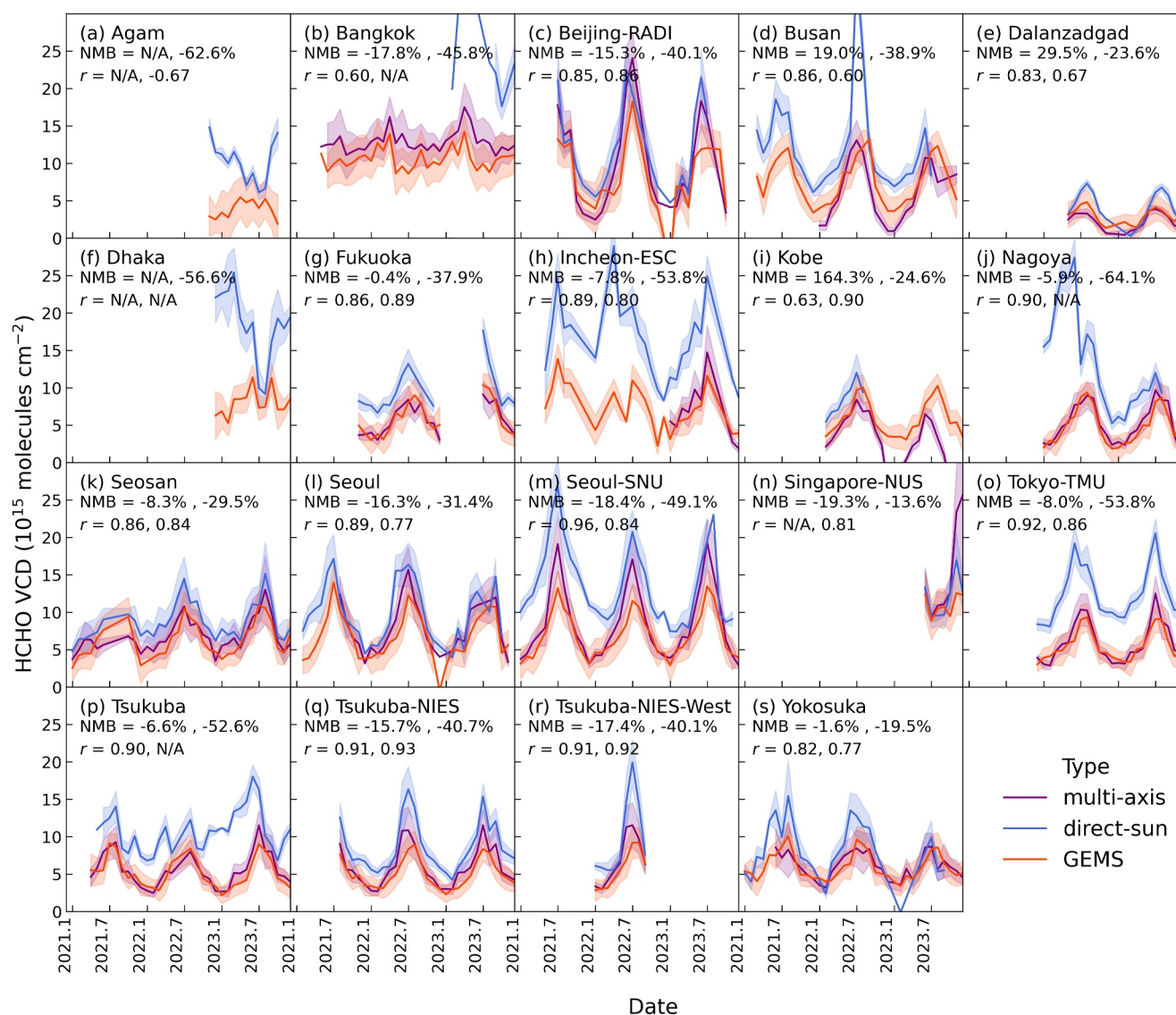
Figure 3. Monthly variations of coincident HCHO VCD from GEMS and Pandora in each station during 2021–2023. The error bands show ±0.5 standard deviations of all observations for that specific month. NMB refers to the normalized mean bias of GEMS HCHO relative to Pandora multi-axis (the first number) and direct-sun (the second number) observations. Only statistically significant (*p*-value <0.05) Pearson correlation coefficient (*r*) is shown.

confine the study area to the urban area (Section 2.5), HCHO in these cities is jointly contributed by many factors, including oxidations of background long-lived VOCs (*e.g.*, methane), biogenic emissions from urban vegetation, anthropogenic emissions, biomass burning and transport from surrounding areas.

Additionally, to quantify the impacts of background corrections on diurnal GEMS HCHO variations, we calculate the model contributed VCD_0 (Equation 3; $VCD = DVCD + VCD_0$; dashed lines in Figure 7). We find that VCD_0 in northern cities shows stronger diurnal variation and is overall larger ($>7 \times 10^{15}$ molecules cm^{-2}) than that in southern cities ($\sim4 \times 10^{15}$ molecules cm^{-2}). Such differences are due to the latitudinal background correction discussed in Section 3.1. Because the reference areas GEMS uses are more polluted in the northern areas, the model-simulated diurnal variations are much stronger. It also means that during spectral fitting, the radiances of nearby coastal cities (*e.g.*, Busan, Seoul) are largely offset by the reference area, leading to flat diurnal DVCD. However, in northern megacities (*e.g.*, Beijing), the difference between VCD_0 and total VCD is clear in the afternoon during summer due to much stronger emissions.
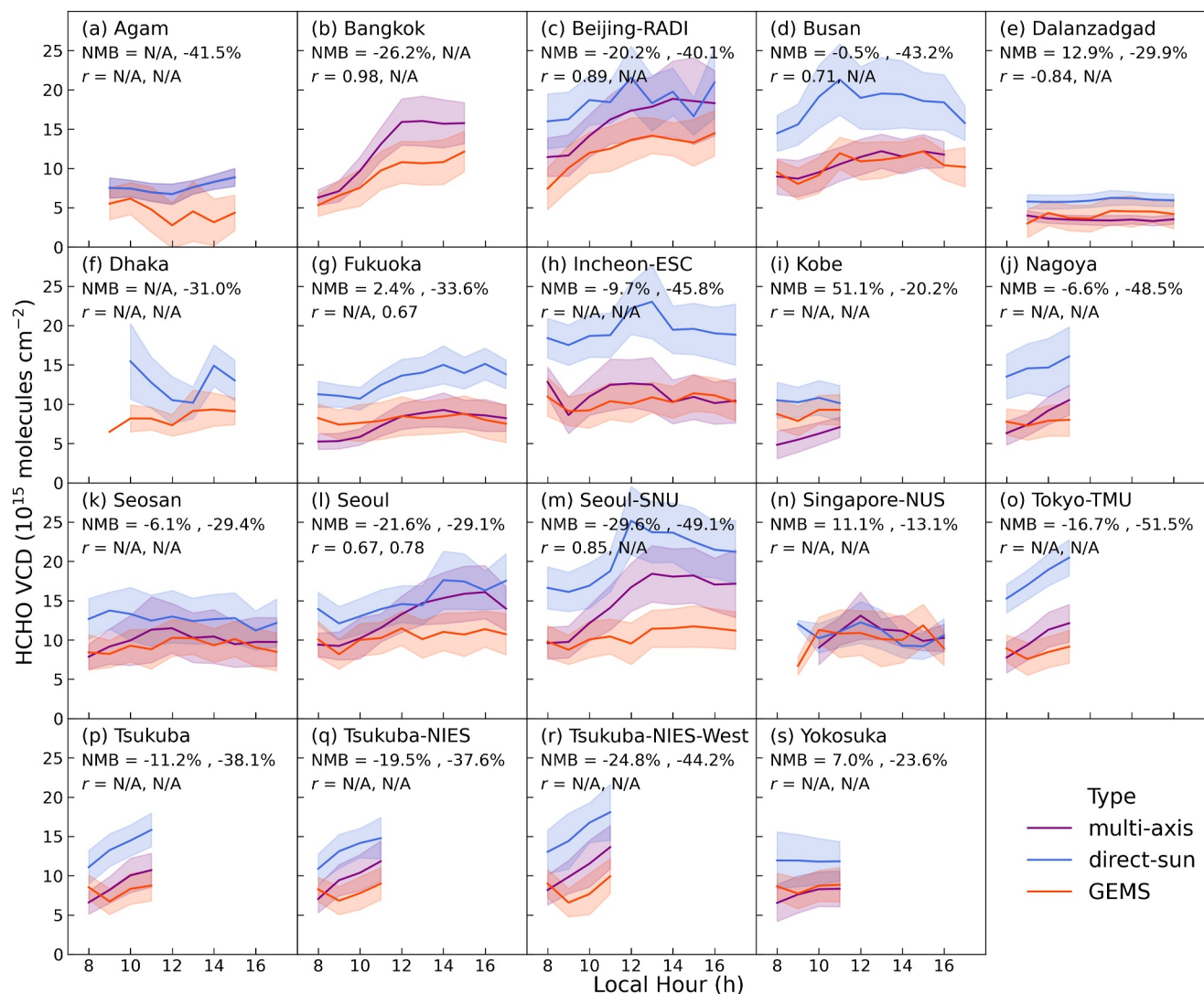
Figure 4. Diurnal variations of HCHO VCD summer average (June, July, and August) from GEMS and Pandora in each station during 2021–2023. The X-axis shows the hour of observation converted to local time based on its longitude. The error bands show ±0.5 standard deviations of all observations for that specific hour. The NMB refers to normalized mean bias of GEMS HCHO relative to Pandora multi-axis (the first number) and direct-sun (the second number) observations. Only statistically significant (p-value <0.05) Pearson correlation coefficient (r) is shown.

We also find GEMS diurnal HCHO variations agree well with previous ground-based MAX-DOAS measurements. In Beijing (Figure 7c) HCHO increases during the morning rush hour, with a mild decrease around 11:00 a.m., then continues to increase in the afternoon (Javed et al., 2019). In Shenyang (Figure 7b), summertime HCHO continues to increase during the day while the variations in other seasons are mild (Xue et al., 2022). In Shanghai (Figure S15h in Supporting Information S1) during autumn, HCHO VCD continues to increase after 9:00 a.m. (Zhang et al., 2021).

Higher (compared to midday) HCHO VCD in the morning is likely caused by early commuting traffic in some of the most densely populated cities (*e.g.*, Shanghai, Seoul). For other cities, there is an overall increasing trend during the day, which may result from strong industrial activities and relatively weak removal of HCHO. Seasonal HCHO variations are partly due to temperature impacting the oxidation of VOCs which forms HCHO (Atkinson & Arey, 2003). Considering Bangkok and Ho Chi Minh City are close to fire emission sources discussed in Section 4.3, surrounding fires likely contribute to their wintertime HCHO.
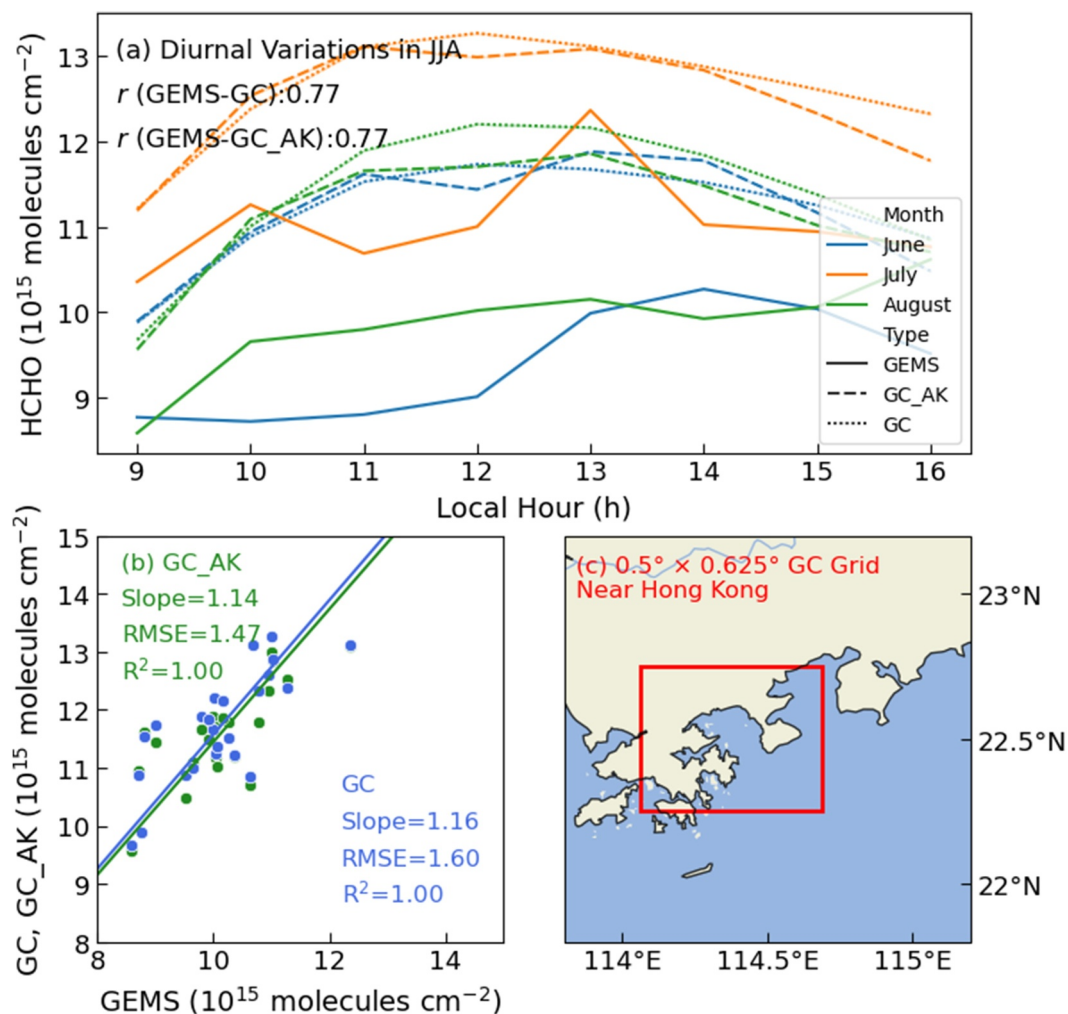
Figure 5. Summertime diurnal variations of HCHO from GEMS and GEOS-Chem near Hong Kong in 2021. Panel (a) shows the diurnal HCHO variations from GEMS, GEOS-Chem (GC), and GEOS-Chem with GEMS averaging kernels (GC_AK); correlation coefficients (*r*) are calculated from hourly mean HCHO of each month (24 pairs in total). Panel (b) shows the scatterplots between GEMS and GC HCHO; the regression lines are fitted without intercepts. Panel (c) shows the location of the GC grid used in this figure.

4.2. Forests

In forest areas, GEMS HCHO does not show significant diurnal HCHO variations, but monthly HCHO shows similar patterns with temperature and isoprene emissions (Figure 8). For instance, in Malaysia, Borneo Island, and the Philippines, where the year-round temperature is above 20° Celsius, the monthly trends of HCHO and isoprene emissions are similar. In North Korea, temperature, isoprene emissions, and HCHO all peaked in summer. Because the oxidation of methane and other long-lived VOCs contributed to the background HCHO, the total HCHO VCD in these different regions is close.

4.3. Fire-Affected Regions

Fires in mainland Southeast Asia from January to April severely threaten regional air quality and human health (Huang et al., 2016; Lee et al., 2018; Reddington et al., 2021). To evaluate the impacts unrelated (*e.g.*, meteorological) and related to fires on HCHO, we conduct a series of GEOS-Chem sensitivity simulations. We note that GFED4s fire emission inventory used in our GEOS-Chem simulations accounts for small fires, which is necessary yet largely increases (~2 times depending on the date) the total fire emissions and uncertainties in Southeast Asia
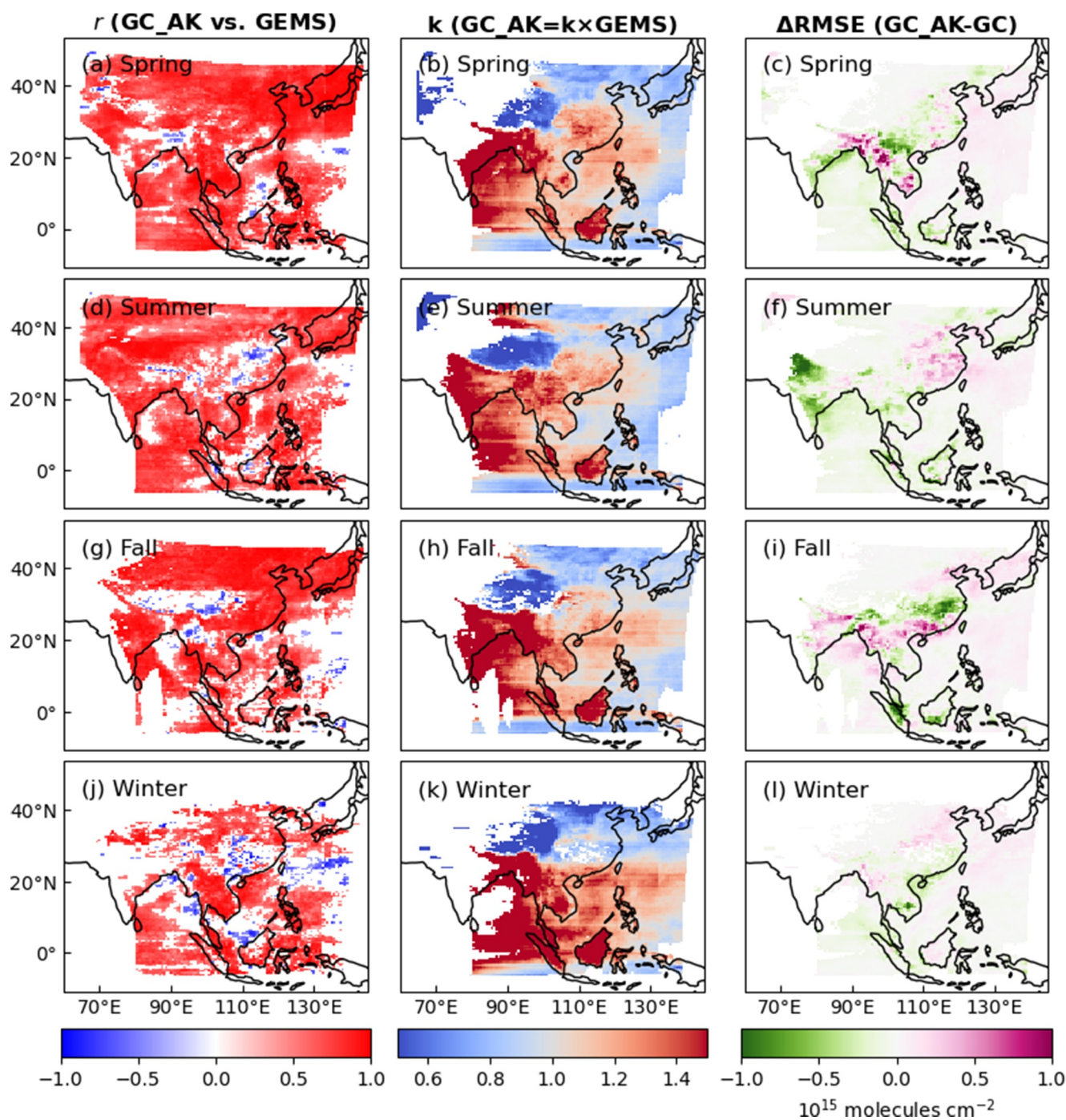
Figure 6. Pixelwise statistics of diurnal HCHO between GEOS-Chem and GEMS. The first column shows correlation coefficients (*r*; same as Figure 5a) between hourly mean GEMS and GC HCHO of each month (minimum 15 pairs except 12 pairs for winter. The second column shows the pixelwise slope of linear regression (without an intercept), based on pairs of diurnal GC_AK and GEMS HCHO within 3 months (same as Figure 5b). The third column shows changes in the root mean square error (RMSE) from linear regressions after applying averaging kernels (same as Figure 5b). Note that only statistically significant pixels are shown (*p-value* < 0.05); for the second and third columns, an additional requirement is that the R^2 of the linear regression model needs to be larger than 0.9.

(van der Werf et al., 2017). Its 3-hr diurnal emission factors in Southeast Asia are estimated from geostationary active fire observations in Southern Hemisphere South America (Mu et al., 2011). Nevertheless, it remains one of the most reliable fire emission inventories (Chen et al., 2023; Pan et al., 2020; Xu et al., 2023; Zheng et al., 2023).
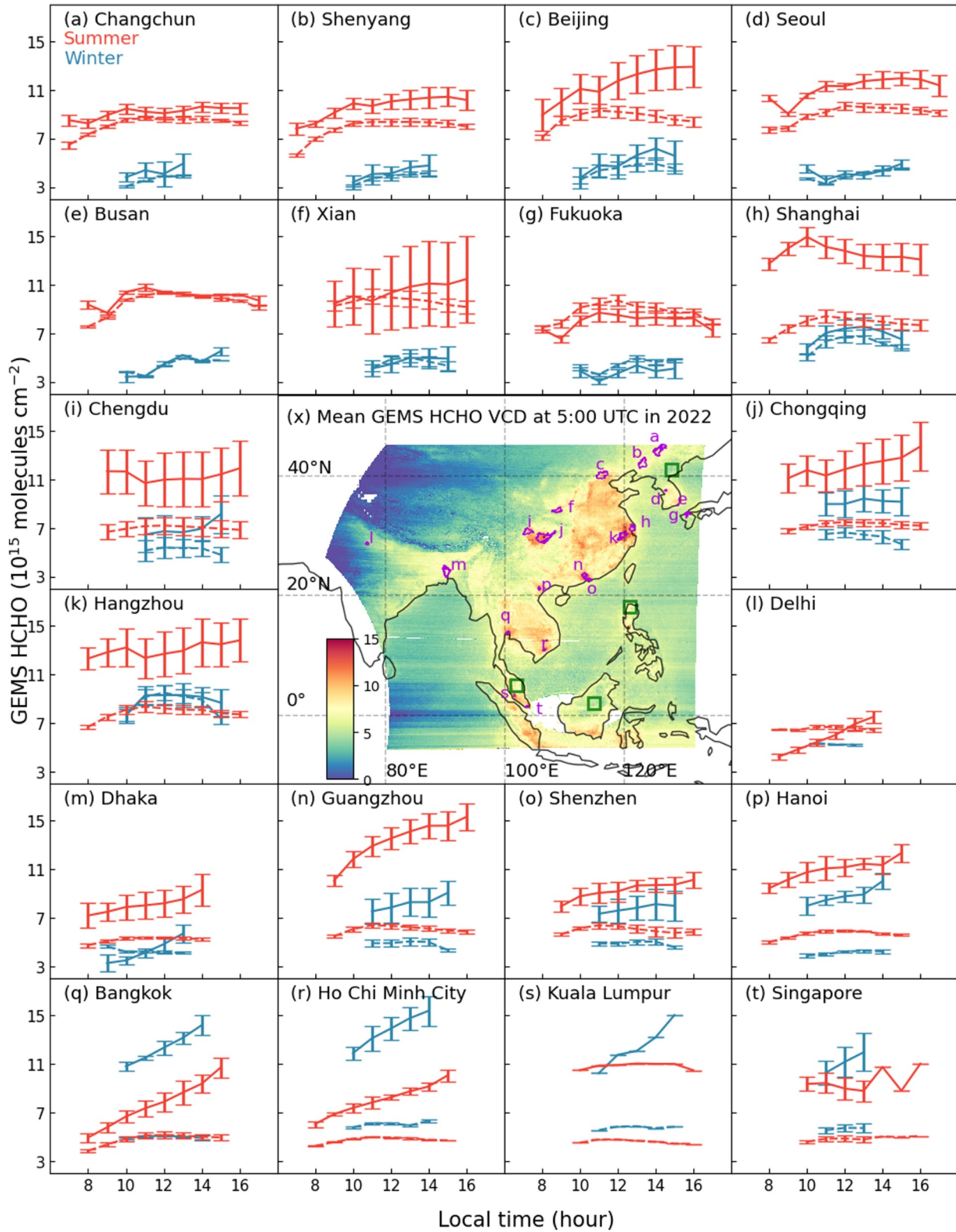
Figure 7. Diurnal variations of summer (red) and winter (blue) average GEMS HCHO VCD (solid line) and model-contributed VCD_0 (dashed line) over 20 major Asian cities. The central panel shows the annual mean GEMS HCHO at 5:00 UTC in 2022; each city in the surrounding panels (a–t) is drawn with purple in the central panel; four green rectangles represent forest areas discussed in Section 2.5 and upcoming Section 4.2. The surrounding panels show the diurnal variations of GEMS HCHO in different seasons for each city; error bars show regional standard deviations.
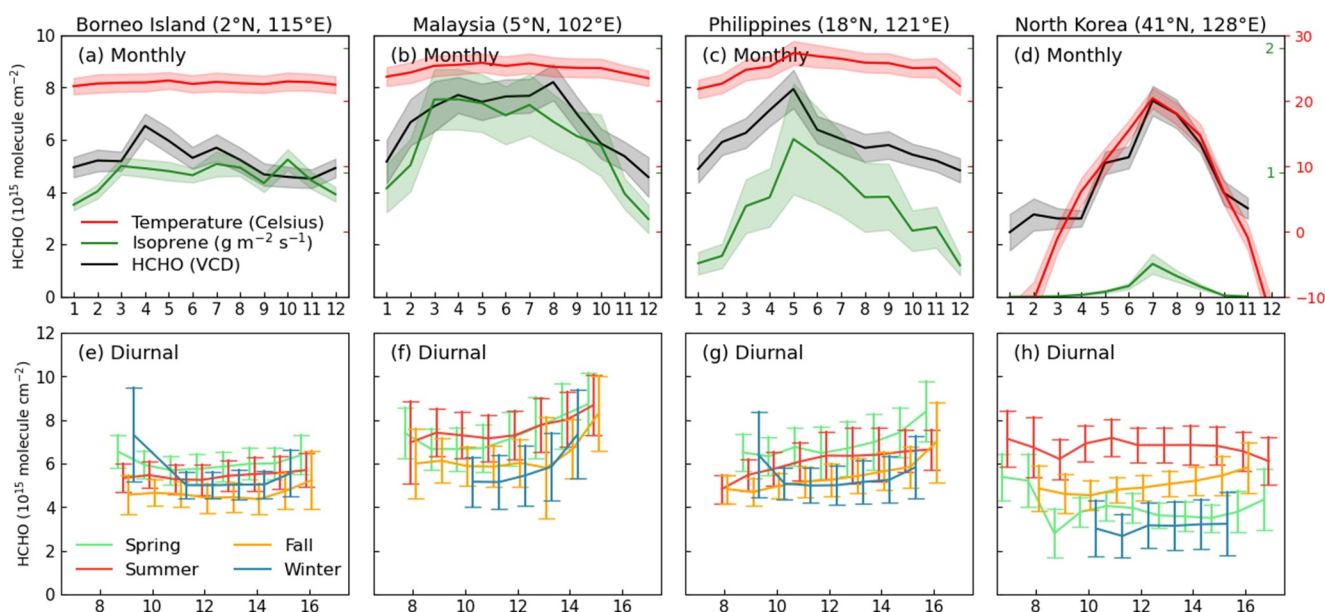
Figure 8. Monthly (a–d) and diurnal (e–h) variations of GEMS HCHO VCD in four $2° \times 2°$ forest areas. Their locations are shown by the green rectangles in the central panel (x) of Figure 7. Error bars (e–h) show standard deviations regionally. Error bands (a–d) show ± 0.5 standard deviations.

We do not aim to address their absolute differences here, as observations and simulations are inherently different. Instead, we focus on whether GEMS HCHO can capture the temporal and spatial characteristics of fire emissions.

GEOS-Chem sensitivity simulations indicate that, in March, fire dominates interannual (2022–2021) HCHO changes in mainland Southeast Asia. The size of the study area (12–28°N, 92.5–107.5°E) is chosen based on GEOS-Chem grids. First, we turn off GFED4s fire emissions in both 2021 and 2022. The results (Figure S16 in Supporting Information S1) show that the interannual HCHO variations are within $\pm 1 \times 10^{15}$ molecules cm^{-2} in January and March, and $\sim 2 \times 10^{15}$ molecules cm^{-2} in February and April. Then, we compare the HCHO VCD with/without fire emissions (Figure 9). We find fire emissions cause increases of $\sim 6 \times 10^{15}$ molecules cm^{-2} HCHO (from $\sim 12 \times 10^{15}$) in March 2021 and 2022×10^{15} in February and April, while little differences are found in January.

In addition, improved agreements are found between GEOS-Chem and GEMS when using GFED4s diurnal fire emission factors ($\sim 80\%$ of fire emissions are released after midday), which indicates increasing HCHO after 10 a. m. can be partly attributed to fires. Thus, we further compare the spatial patterns of changes in GFED4s and GEMS HCHO in March (Figure 10), during which the interannual GEMS HCHO variability should roughly represent the changes in fire intensity. As expected, afternoon HCHO changes coincide the best with GFED4s 2022–2021 changes.

To investigate morning HCHO increases near Myanmar (20–28°N, 92.5–97.5°E), we compare interannual diurnal HCHO variations based on GEOS-Chem simulations. The results (Figure 11) show that diurnal HCHO near Myanmar is always higher in 2022 compared to 2021, regardless of whether fire emission is turned on, which is consistent with GEMS morning HCHO changes (Figure 10a). Notably, the gap between diurnal HCHO in 2021 and 2022 reduces from morning ($\sim 1 \times 10^{15}$ molecules cm^{-2}) to afternoon (close to 0). This is consistent with GEMS afternoon observations (Figure 10c), as GEMS also only observed a modest decrease in a small region.

In Figure S17 in Supporting Information S1, similar agreements are found near Vietnam (20–24°N, 102.5–107.5°E). Fire-induced HCHO was substantially higher in 2021 (up to 5×10^{15} molecules cm^{-2}) compared to 2022 ($\sim 2 \times 10^{15}$). Interannual HCHO differences peaked before dawn ($\sim 4 \times 10^{15}$) and continuously decreased during the day, in line with GEMS HCHO changes shown in Figure 10. In summary, the spatiotemporal characteristics of GEMS HCHO align with GFED4s and corresponding GEOS-Chem simulations.
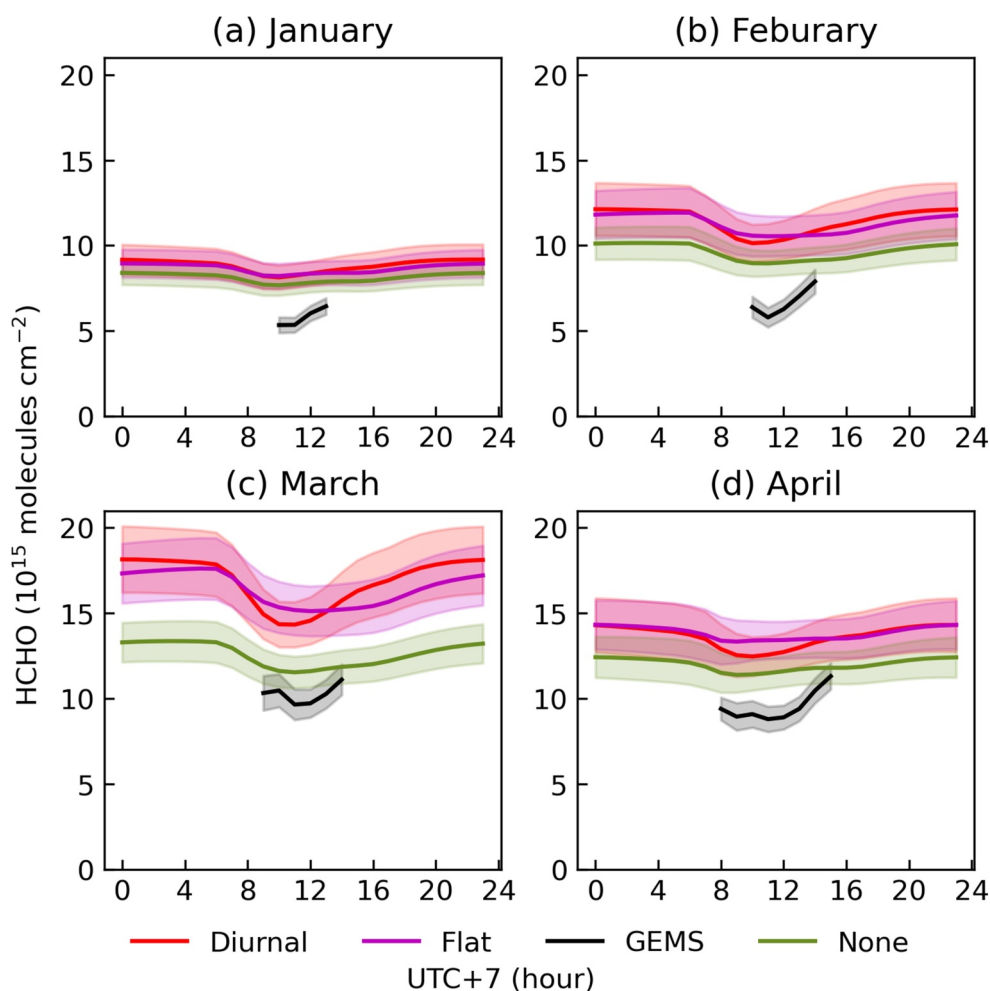
Figure 9. Diurnal HCHO variations in mainland Southeast Asia from GEOS-Chem. Type "Diurnal" refers to GEOS-Chem simulated HCHO using GFED4s fire emission and its diurnal emission factors; "Flat"—using GFED4s fire emission without diurnal emission factors; "None"—without fire emissions. The error bands show ±0.25 regional standard deviations.

5. Discussion and Conclusion

In this paper, we present the first study to evaluate the year-round operational GEMS HCHO Version 2 product and characterize diurnal HCHO variations over different areas. We find that GEMS HCHO is in close agreement (differences less than 20%) with the TROPOMI satellite in most areas in China, Korea, and mainland Southeast Asia. GEMS diurnal HCHO variations in major cities are also close to Pandora multi-axis HCHO ($r \sim 0.85$; bias $\sim -15\%$).

GEMS reveals distinct diurnal HCHO variations over different areas. Over urban areas, the diurnal variations are pronounced. For instance, HCHO continues to increase in Guangzhou during summer, starting from 0.9×10^{16} molecules cm^{-2} in the morning to 1.6×10^{16} molecules cm^{-2} in the afternoon. Over forests, HCHO does not show diurnal variations. Instead, it is mainly influenced by seasonal temperature variations. During the fire season in mainland Southeast Asia, GEMS HCHO continues to increase during the day, in line with GEOS-Chem simulations and the underlying GFED4s, which estimates that only $\sim 10\%$ of fire emissions were released between night and morning (22:00 to 10:00). We also find GEMS afternoon HCHO observations show similar spatial patterns with GFED4s in March.

Despite showing exceptional agreements with TROPOMI, Pandora, and GEOS-Chem, GEMS HCHO retrievals may need further updates. Our study shows GEMS Version 2 HCHO retrieval over 30°N is dominated by model corrections. However, the model may not always be able to characterize complex anthropogenic environments
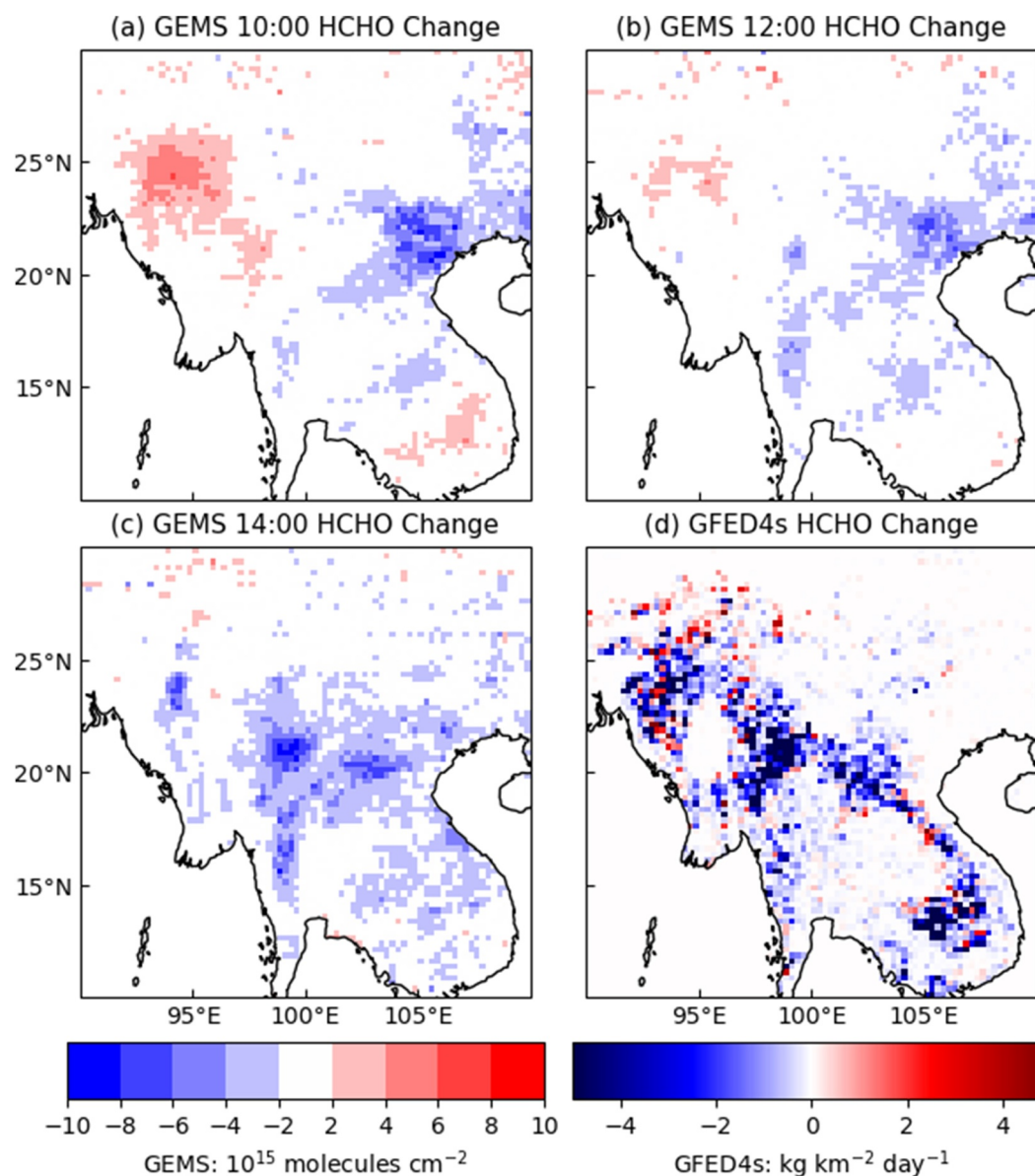
Figure 10. Changes (2022–2021) of mean HCHO in March from GEMS (a, b, c) and GFED4s (d). GFED4s HCHO is calculated using the method described in Section 2.5.

(Section 3.3), thus it is crucial to avoid being overly dependent on the model. One potential solution could be fitting DSCD over clean reference (below 30°N) rather than strictly using reference areas of the same latitude. GEMS HCHO under high VZA (>60°) is less than half of TROPOMI, indicating improvements in air mass factors and/or spectral fittings are needed. A tentative solution to correct these systematic underestimations is to use scaling factors (Zhu et al., 2016, 2017) based on coincident data sets (*e.g.*, TROPOMI or ground-based observations). Here, we note whether GEMS can still capture diurnal HCHO trends under extreme viewing geometries (*e.g.*, India) remains to be investigated by future studies using in-situ measurements.

To conclude, this study compares the first geostationary GEMS satellite HCHO observations with the TROPOMI HCHO product, Pandora observations, and GEOS-Chem simulations. Their broad agreements show how hourly HCHO observations from geostationary satellites can help to improve our understanding of diurnal anthropogenic VOC emissions and fire activity. We also present one of the first studies to analyze both multi-axis and direct-sun HCHO products from the fast-developing Pandonia network. In almost every site, both products show similar trends, demonstrating how this centralized-managed open data can benefit the evaluation of geostationary
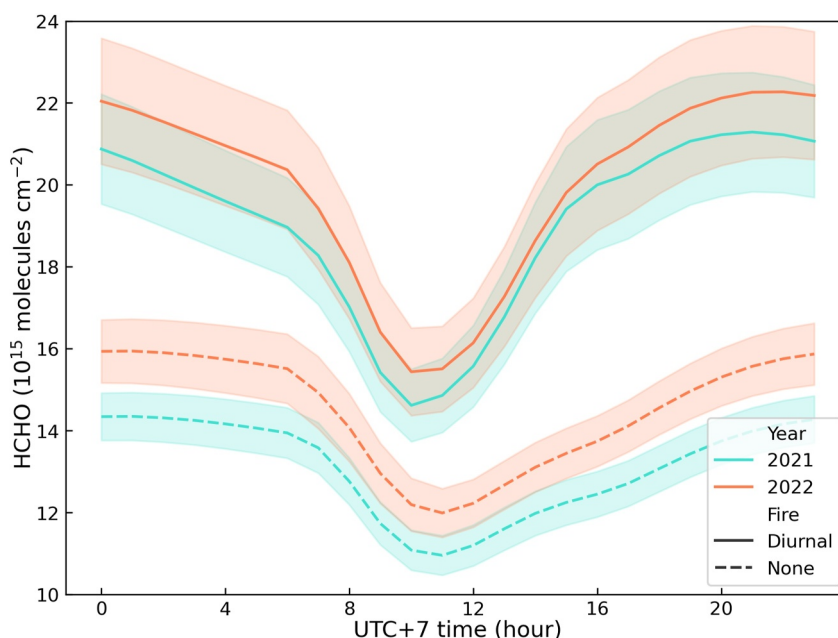
Figure 11. Diurnal HCHO variations near Myanmar (20–28°N, 92.5–97.5°E) simulated by GEOS-Chem using GFED4s "diurnal" fire emission factors or without fire emissions. The error bands show ±0.25 regional standard deviations.

satellites, including GEMS, recently launched TEMPO (Zoogman et al., 2017), and upcoming Sentinel-4 (Courrèges-Lacoste et al., 2017).

Data Availability Statement

GEMS HCHO Version 2 data can be requested from the Korean Environmental Satellite Center (Park & Kwon, 2020). TROPOMI HCHO is available at Copernicus Data Space (European Space Agency, 2020). Pandora HCHO observations can be accessed from the Pandonia Global Network (LuftBlick, 2023). GEOS-Chem model codes are available at Zenodo (Yantosca et al., 2023). CAMS-GLOB-BIO v3.1 biogenic isoprene inventory is downloaded from the Global Emission Initiative data portal (Granier et al., 2019). The Gridded Population of the World Version 4 for 2020 is hosted by Columbia University (Center for International Earth Science Information Network, 2018). The Global Land Cover and Land Use Change data set is accessed from Google Earth Engine (Potapov et al., 2022b). The GFED4s fire emission inventory is from Randerson et al. (2024). Python codes used in this study are available at GitHub (https://github.com/Fu-Weitao/Evaluate-GEMS-HCHO) and Zenodo (Fu, 2024).

References

Akagi, S. K., Yokelson, R. J., Wiedinmyer, C., Alvarado, M. J., Reid, J. S., Karl, T., et al. (2011). Emission factors for open and domestic biomass burning for use in atmospheric models. *Atmospheric Chemistry and Physics*, *11*(9), 4039–4072. https://doi.org/10.5194/acp-11-4039-2011

Anderson, D. C., Duncan, B. N., Nicely, J. M., Liu, J., Strode, S. A., & Follette-Cook, M. B. (2023). Technical note: Constraining the hydroxyl (OH) radical in the tropics with satellite observations of its drivers – First steps toward assessing the feasibility of a global observation strategy. *Atmospheric Chemistry and Physics*, *23*(11), 6319–6338. https://doi.org/10.5194/acp-23-6319-2023

Atkinson, R., & Arey, J. (2003). Atmospheric degradation of volatile organic compounds. *Chemical Reviews*, *103*(12), 4605–4638. https://doi.org/10.1021/cr0206420

Bartier, P. M., & Keller, C. P. (1996). Multivariate interpolation to incorporate thematic surface data using inverse distance weighting (IDW). *Computers and Geosciences*, *22*(7), 795–799. https://doi.org/10.1016/0098-3004(96)00021-0

Bey, I., Jacob, D. J., Yantosca, R. M., Logan, J. A., Field, B. D., Fiore, A. M., et al. (2001). Global modeling of tropospheric chemistry with assimilated meteorology: Model description and evaluation. *Journal of Geophysical Research*, *106*(D19), 23073–23095. https://doi.org/10.1029/2001JD000807

Burrows, J. P., Weber, M., Buchwitz, M., Rozanov, V., Ladstätter-Weißenmayer, A., Richter, A., et al. (1999). The global ozone monitoring experiment (GOME): Mission concept and first scientific results. *Journal of the Atmospheric Sciences*, *56*(2), 151–175. https://doi.org/10.1175/1520-0469(1999)056<0151:TGOMEG>2.0.CO;2

Cao, H., Fu, T. M., Zhang, L., Henze, D. K., Miller, C. C., Lerot, C., et al. (2018). Adjoint inversion of Chinese non-methane volatile organic compound emissions using space-based observations of formaldehyde and glyoxal. *Atmospheric Chemistry and Physics*, *18*(20), 15017–15046. https://doi.org/10.5194/acp-18-15017-2018

Center for International Earth Science Information Network. (2018). Gridded population of the World, version 4 (GPWv4): Population density, revision 11 [Dataset]. https://doi.org/10.7927/H49C6VHW

Chan, K. L., Wiegner, M., van Geffen, J., De Smedt, I., Alberti, C., Cheng, Z., et al. (2020). MAX-DOAS measurements of tropospheric NO2 and HCHO in Munich and the comparison to OMI and TROPOMI satellite observations. *Atmospheric Measurement Techniques*, *13*(8), 4499–4520. https://doi.org/10.5194/amt-13-4499-2020

Chance, K., Palmer, P. I., Spurr, R. J., Martin, R. V., Kurosu, T. P., & Jacob, D. J. (2000). Satellite observations of formaldehyde over North America from GOME. *Geophysical Research Letters*, *27*(21), 3461–3464. https://doi.org/10.1029/2000GL011857

Chang, L. S., Kim, D., Hong, H., Kim, D. R., Yu, J. A., Lee, K., et al. (2022). Evaluation of correlated Pandora column NO2 and in situ surface NO2 measurements during GMAP campaign. *Atmospheric Chemistry and Physics*, *22*(16), 10703–10720. https://doi.org/10.5194/acp-22-10703-2022

Chen, Y., Hall, J., van Wees, D., Andela, N., Hantson, S., Giglio, L., et al. (2023). Multi-decadal trends and variability in burned area from the fifth version of the Global Fire Emissions Database (GFED5). *Earth System Science Data*, *15*(11), 5227–5259. https://doi.org/10.5194/essd-15-5227-2023

Choi, J., Henze, D. K., Cao, H., Nowlan, C. R., González Abad, G., Kwon, H.-A., et al. (2022). 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*(7), e2021JD035844. https://doi.org/10.1029/2021JD035844

Choi, W. J., Moon, K.-J., Yoon, J., Cho, A., Kim, S.-k., Lee, S., et al. (2018). Introducing the geostationary environment monitoring spectrometer. *Journal of Applied Remote Sensing*, *12*(4), 044005. https://doi.org/10.1117/1.JRS.12.044005

Chong, K., Wang, Y., Liu, C., Gao, Y., Boersma, K. F., Tang, J., & Wang, X. (2024). Remote sensing measurements at a rural site in China: Implications for satellite NO2 and HCHO measurement uncertainty and emissions from fires. *Journal of Geophysical Research: Atmospheres*, *129*(2), e2023JD039310. https://doi.org/10.1029/2023JD039310

Coggon, M. M., McDonald, B. C., Vlasenko, A., Veres, P. R., Bernard, F., Koss, A. R., et al. (2018). Diurnal variability and emission pattern of decamethylcyclopentasiloxane (D5) from the application of personal care products in two north American cities. *Environmental Science and Technology*, *52*(10), 5610–5618. https://doi.org/10.1021/acs.est.8b00506

Courrèges-Lacoste, G. B., Sallusti, M., Bulsa, G., Bagnasco, G., Veihelmann, B., Riedl, S., et al. (2017). The Copernicus Sentinel 4 mission: A geostationary imaging UVN spectrometer for air quality monitoring. *Sensors, Systems, and Next-Generation Satellites XXI*. https://doi.org/10.1117/12.2282158

De Smedt, I., Müller, J.-F., Stavrakou, T., Van Der A, R., Eskes, H., & Van Roozendael, M. (2008). Twelve years of global observations of formaldehyde in the troposphere using GOME and SCIAMACHY sensors. *Atmospheric Chemistry and Physics*, *8*(16), 4947–4963. https://doi.org/10.5194/acp-8-4947-2008

De Smedt, I., Pinardi, G., Vigouroux, C., Compernolle, S., Bais, A., Benavent, N., et al. (2021). Comparative assessment of TROPOMI and OMI formaldehyde observations and validation against MAX-DOAS network column measurements. *Atmospheric Chemistry and Physics*, *21*(16), 12561–12593. https://doi.org/10.5194/acp-21-12561-2021

De Smedt, I., Theys, N., Yu, H., Danckaert, T., Lerot, C., Compernolle, S., et al. (2018). Algorithm theoretical baseline for formaldehyde retrievals from S5P TROPOMI and from the QA4ECV project. *Atmospheric Measurement Techniques*, *11*(4), 2395–2426. https://doi.org/10.5194/amt-11-2395-2018

De Smedt, I., Van Roozendael, M., Stavrakou, T., Müller, J. F., Lerot, C., Theys, N., et al. (2012). Improved retrieval of global tropospheric formaldehyde columns from GOME-2/MetOp-A addressing noise reduction and instrumental degradation issues. *Atmospheric Measurement Techniques*, *5*(11), 2933–2949. https://doi.org/10.5194/amt-5-2933-2012

Duncan, B. N., Yoshida, Y., Olson, J. R., Sillman, S., Martin, R. V., Lamsal, L., et al. (2010). Application of OMI observations to a space-based indicator of NOx and VOC controls on surface ozone formation. *Atmospheric Environment*, *44*(18), 2213–2223. https://doi.org/10.1016/j.atmosenv.2010.03.010

Ehn, M., Thornton, J. A., Kleist, E., Sipilä, M., Junninen, H., Pullinen, I., et al. (2014). A large source of low-volatility secondary organic aerosol. *Nature*, *506*(7489), 476–479. https://doi.org/10.1038/nature13032

European Space Agency. (2020). Copernicus Sentinel-5P (processed by ESA), 2020, TROPOMI level 2 formaldehyde total column products [Dataset]. *Copernicus Data Space Ecosystem*. https://doi.org/10.5270/S5P-vg1i7t0

FAO. (2015). Food and agriculture organization global administrative unit layers (GAUL) [Dataset]. Retrieved from https://data.apps.fao.org/map/catalog/srv/api/records/9c35ba10-5649-41c8-bdfc-eb78e9e65654/attachments/GAUL2015_Documentation.zip

Fioletov, V., McLinden, C. A., Griffin, D., Krotkov, N., Liu, F., & Eskes, H. (2022). Quantifying urban, industrial, and background changes in NO2 during the COVID-19 lockdown period based on TROPOMI satellite observations. *Atmospheric Chemistry and Physics*, *22*(6), 4201–4236. https://doi.org/10.5194/acp-22-4201-2022

Frieß, U., Beirle, S., Alvarado Bonilla, L., Bösch, T., Friedrich, M. M., Hendrick, F., et al. (2019). Intercomparison of MAX-DOAS vertical profile retrieval algorithms: Studies using synthetic data. *Atmospheric Measurement Techniques*, *12*(4), 2155–2181. https://doi.org/10.5194/amt-12-2155-2019

Fu, W. (2024). Supplementary codes of manuscript submitted to AGU earth and space science [Software]. *Zenodo*. https://doi.org/10.5281/zenodo.13954404

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., et al. (2017). The modern-era retrospective analysis for Research and applications, version 2 (MERRA-2). *Journal of Climate*, *30*(14), 5419–5454. https://doi.org/10.1175/JCLI-D-16-0758.1

González Abad, G., Liu, X., Chance, K., Wang, H., Kurosu, T. P., & Suleiman, R. (2015). Updated Smithsonian Astrophysical Observatory Ozone Monitoring Instrument (SAO OMI) formaldehyde retrieval. *Atmospheric Measurement Techniques*, *8*(1), 19–32. https://doi.org/10.5194/amt-8-19-2015

Granier, C., Darras, S., Denier van der Gon, H., Doubalova, J., Elguindi, N., Galle, B., et al. (2019). The Copernicus atmosphere monitoring service global and regional emissions (April 2019 version) report April 2019 version [Dataset]. *Copernicus Atmosphere Monitoring Service*. https://doi.org/10.24380/d0bn-kx16

Green, J. R., Fiddler, M. N., Fibiger, D. L., McDuffie, E. E., Aquino, J., Campos, T., et al. (2021). Wintertime formaldehyde: Airborne observations and source apportionment over the eastern United States. *Journal of Geophysical Research: Atmospheres*, *126*(5), e2020JD033518. https://doi.org/10.1029/2020JD033518

Gu, S., Guenther, A., & Faiola, C. (2021). Effects of anthropogenic and biogenic volatile organic compounds on Los Angeles air quality. *Environmental Science and Technology*, *55*(18), 12191–12201. https://doi.org/10.1021/acs.est.1c01481

Guenther, A. B., Jiang, X., Heald, C. L., Sakulyanontvittaya, T., Duhl, T., Emmons, L. K., & Wang, X. (2012). The model of emissions of gases and aerosols from nature version 2.1 (MEGAN2.1): An extended and updated framework for modeling biogenic emissions. *Geoscientific Model Development*, *5*(6), 1471–1492. https://doi.org/10.5194/gmd-5-1471-2012

Herman, J., & Mao, J. (2024). Seasonal variation of total column formaldehyde, nitrogen dioxide, and ozone over various Pandora spectrometer sites with a comparison of OMI and diurnally varying DSCOVR-EPIC satellite data. *EGUsphere*, *2024*, 1–20. https://doi.org/10.5194/egusphere-2024-1216

Herman, J., Spinei, E., Fried, A., Kim, J., Kim, J., Kim, W., et al. (2018). NO2 and HCHO measurements in Korea from 2012 to 2016 from Pandora spectrometer instruments compared with OMI retrievals and with aircraft measurements during the KORUS-AQ campaign. *Atmospheric Measurement Techniques*, *11*(8), 4583–4603. https://doi.org/10.5194/amt-11-4583-2018

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., et al. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049. https://doi.org/10.1002/qj.3803

Hoesly, R. M., Smith, S. J., Feng, L., Klimont, Z., Janssens-Maenhout, G., Pitkanen, T., et al. (2018). Historical (1750–2014) anthropogenic emissions of reactive gases and aerosols from the Community Emissions Data System (CEDS). *Geoscientific Model Development*, *11*(1), 369–408. https://doi.org/10.5194/gmd-11-369-2018

Hong, Q., Liu, C., Hu, Q., Zhang, Y., Xing, C., Su, W., et al. (2021). Evaluating the feasibility of formaldehyde derived from hyperspectral remote sensing as a proxy for volatile organic compounds. *Atmospheric Research*, *264*, 105777. https://doi.org/10.1016/j.atmosres.2021.105777

Huang, J., Hartmann, H., Hellén, H., Wisthaler, A., Perreca, E., Weinhold, A., et al. (2018). New perspectives on CO2, temperature, and light effects on BVOC emissions using online measurements by PTR-MS and cavity ring-down spectroscopy. *Environmental Science and Technology*, *52*(23), 13811–13823. https://doi.org/10.1021/acs.est.8b01435

Huang, W.-R., Wang, S.-H., Yen, M.-C., Lin, N.-H., & Promchote, P. (2016). Interannual variation of springtime biomass burning in Indochina: Regional differences, associated atmospheric dynamical changes, and downwind impacts. *Journal of Geophysical Research: Atmospheres*, *121*(17), 10016–10028. https://doi.org/10.1002/2016JD025286

Javed, Z., Liu, C., Khokhar, M. F., Tan, W., Liu, H., Xing, C., et al. (2019). Ground-based MAX-DOAS observations of CHOCHO and HCHO in Beijing and baoding, China. *Remote Sensing*, *11*(13), 1524. https://doi.org/10.3390/rs11131524

Judd, L. M., Al-Saadi, J. A., Szykman, J. J., Valin, L. C., Janz, S. J., Kowalewski, M. G., et al. (2020). Evaluating sentinel-5P TROPOMI tropospheric NO2 column densities with airborne and Pandora spectrometers near New York city and long Island sound. *Atmospheric Measurement Techniques*, *13*(11), 6113–6140. https://doi.org/10.5194/amt-13-6113-2020

Kim, J., Jeong, U., Ahn, M.-H., Kim, J. H., Park, R. J., Lee, H., et al. (2020). New Era of air quality monitoring from space: Geostationary environment monitoring spectrometer (GEMS). *Bulletin of the American Meteorological Society*, *101*(1), E1–E22. https://doi.org/10.1175/bams-d-18-0013.1

Kim, S., Kim, D., Hong, H., Chang, L. S., Lee, H., Kim, D. R., et al. (2023). First-time comparison between NO2 vertical columns from geostationary environmental monitoring spectrometer (GEMS) and Pandora measurements. *Atmospheric Measurement Techniques*, *16*(16), 3959–3972. https://doi.org/10.5194/amt-16-3959-2023

Kreher, K., Van Roozendael, M., Hendrick, F., Apituley, A., Dimitropoulou, E., Frieß, U., et al. (2020). Intercomparison of NO2, O4, O3 and HCHO slant column measurements by MAX-DOAS and zenith-sky UV–visible spectrometers during CINDI-2. *Atmospheric Measurement Techniques*, *13*(5), 2169–2208. https://doi.org/10.5194/amt-13-2169-2020

Kumar, V., Beirle, S., Dörner, S., Mishra, A. K., Donner, S., Wang, Y., et al. (2020). Long-term MAX-DOAS measurements of NO2, HCHO, and aerosols and evaluation of corresponding satellite data products over Mohali in the Indo-Gangetic Plain. *Atmospheric Chemistry and Physics*, *20*(22), 14183–14235. https://doi.org/10.5194/acp-20-14183-2020

Kwon, H. A., Park, R. J., González Abad, G., Chance, K., Kurosu, T. P., Kim, J., et al. (2019). Description of a formaldehyde retrieval algorithm for the geostationary environment monitoring spectrometer (GEMS). *Atmospheric Measurement Techniques*, *12*(7), 3551–3571. https://doi.org/10.5194/amt-12-3551-2019

Lange, K., Richter, A., Bösch, T., Zilker, B., Latsch, M., Behrens, L. K., et al. (2024). Validation of GEMS tropospheric NO2 columns and their diurnal variation with ground-based DOAS measurements. *Atmospheric Measurement Techniques*, *17*(21), 6315–6344. https://doi.org/10.5194/amt-17-6315-2024

Lee, G. T., Park, R. J., Kwon, H. A., Ha, E. S., Lee, S. D., Shin, S., et al. (2024). First evaluation of the GEMS formaldehyde product against TROPOMI and ground-based column measurements during the in-orbit test period. *Atmospheric Chemistry and Physics*, *24*(8), 4733–4749. https://doi.org/10.5194/acp-24-4733-2024

Lee, H. H., Iraqui, O., Gu, Y., Yim, S. H. L., Chulakadabba, A., Tonks, A. Y. M., et al. (2018). Impacts of air pollutants from fire and non-fire emissions on the regional air quality in Southeast Asia. *Atmospheric Chemistry and Physics*, *18*(9), 6141–6156. https://doi.org/10.5194/acp-18-6141-2018

Levelt, P. F., Stein Zweers, D. C., Aben, I., Bauwens, M., Borsdorff, T., De Smedt, I., et al. (2022). Air quality impacts of COVID-19 lockdown measures detected from space using high spatial resolution observations of multiple trace gases from Sentinel-5P/TROPOMI. *Atmospheric Chemistry and Physics*, *22*(15), 10319–10351. https://doi.org/10.5194/acp-22-10319-2022

Li, Y., Martin, R. V., Li, C., Boys, B. L., van Donkelaar, A., Meng, J., & Pierce, J. R. (2023). Development and evaluation of processes affecting simulation of diel fine particulate matter variation in the GEOS-Chem model. *Atmospheric Chemistry and Physics*, *23*(19), 12525–12543. https://doi.org/10.5194/acp-23-12525-2023

Liao, J., Hanisco, T. F., Wolfe, G. M., St. Clair, J., Jimenez, J. L., Campuzano-Jost, P., et al. (2019). Towards a satellite formaldehyde – In situ hybrid estimate for organic aerosol abundance. *Atmospheric Chemistry and Physics*, *19*(5), 2765–2785. https://doi.org/10.5194/acp-19-2765-2019

Liu, H., Man, H., Tschantz, M., Wu, Y., He, K., & Hao, J. (2015). VOC from vehicular evaporation emissions: Status and control strategy. *Environmental Science and Technology*, *49*(24), 14424–14431. https://doi.org/10.1021/acs.est.5b04064

LuftBlick. (2023). Data products readme document [Dataset]. *Pandonia Global Network*. Retrieved from https://www.pandonia-global-network.org/wp-content/uploads/2023/11/PGN_DataProducts_Readme_v1-8-8.pdf

McDuffie, E. E., Smith, S. J., O'Rourke, P., Tibrewal, K., Venkataraman, C., Marais, E. A., et al. (2020). A global anthropogenic emission inventory of atmospheric pollutants from sector- and fuel-specific sources (1970–2017): An application of the community emissions data System (CEDS). *Earth System Science Data*, *12*(4), 3413–3442. https://doi.org/10.5194/essd-12-3413-2020

Menchaca-Torre, H. L., Mercado-Hernández, R., & Mendoza-Domínguez, A. (2015). Diurnal and seasonal variation of volatile organic compounds in the atmosphere of Monterrey, Mexico. *Atmospheric Pollution Research*, *6*(6), 1073–1081. https://doi.org/10.1016/j.apr.2015.06.004

Mu, M., Randerson, J. T., van der Werf, G. R., Giglio, L., Kasibhatla, P., Morton, D., et al. (2011). Daily and 3-hourly variability in global fire emissions and consequences for atmospheric model predictions of carbon monoxide. *Journal of Geophysical Research*, *116*(D24). https://doi.org/10.1029/2011JD016245

Nowlan, C. R., González Abad, G., Kwon, H.-A., Ayazpour, Z., Chan Miller, C., Chance, K., et al. (2023). Global formaldehyde products from the Ozone Mapping and Profiler Suite (OMPS) nadir mappers on Suomi NPP and NOAA-20. *Earth and Space Science*, *10*, e2022EA002643. https://doi.org/10.1029/2022EA002643

Palmer, P. I., Jacob, D. J., Fiore, A. M., Martin, R. V., Chance, K., & Kurosu, T. P. (2003). Mapping isoprene emissions over North America using formaldehyde column observations from space. *Journal of Geophysical Research*, *108*(D6). https://doi.org/10.1029/2002JD002153

Pan, X., Ichoku, C., Chin, M., Bian, H., Darmenov, A., Colarco, P., et al. (2020). Six global biomass burning emission datasets: Intercomparison and application in one global aerosol model. *Atmospheric Chemistry and Physics*, *20*(2), 969–994. https://doi.org/10.5194/acp-20-969-2020

Park, R. J., & Kwon, H.-A. (2020). Geostationary environment monitoring spectrometer (GEMS) algorithm theoretical basis document VOC (HCHO/CHOCHO) retrieval algorithm [Dataset]. *National Institute of Environmental Research*. Retrieved from https://nesc.nier.go.kr/en/html/datasvc/index.do

Pinardi, G., Van Roozendael, M., Abuhassan, N., Adams, C., Cede, A., Clémer, K., et al. (2013). MAX-DOAS formaldehyde slant column measurements during CINDI: Intercomparison and analysis improvement. *Atmospheric Measurement Techniques*, *6*(1), 167–185. https://doi.org/10.5194/amt-6-167-2013

Potapov, P., Hansen, M. C., Pickens, A., Hernandez-Serna, A., Tyukavina, A., Turubanova, S., et al. (2022a). The global 2000-2020 land cover and land use change dataset derived from the landsat archive: First results [original Research]. *Frontiers in Remote Sensing*, *3*. https://doi.org/10.3389/frsen.2022.856903

Potapov, P., Hansen, M. C., Pickens, A., Hernandez-Serna, A., Tyukavina, A., Turubanova, S., et al. (2022b). Global land cover and land use dataset in 2020 [Dataset]. *Google Earth Engine*. Retrieved from https://code.earthengine.google.com/57162b69fd81ebd97bac80957eade6c6

Pu, D., Zhu, L., De Smedt, I., Li, X., Sun, W., Wang, D., et al. (2022). Response of anthropogenic volatile organic compound emissions to urbanization in Asia probed with TROPOMI and VIIRS satellite observations. *Geophysical Research Letters*, *49*(18), e2022GL099470. https://doi.org/10.1029/2022GL099470

Randerson, J. T., Van Der Werf, G. R., Giglio, L., Collatz, G. J., & Kasibhatla, P. S. (2017). Global Fire Emissions Database, Version 4.1 (GFEDv4) [Dataset]. https://doi.org/10.3334/ORNLDAAC/1293

Randerson, J. T., Van Der Werf, G. R., Giglio, L., Collatz, G. J., & Kasibhatla, P. S. (2024). Global fire emissions Database version 4 with small fires [Dataset] https://www.geo.vu.nl/~gwerf/GFED/GFED4

Rawat, P., Crawford, J. H., Travis, K. R., Judd, L. M., Demetillo, M. A. G., Valin, L. C., et al. (2024). Maximizing the use of Pandora data for scientific applications. *Atmos. Meas. Tech. Discuss.*, *2024*, 1–36. https://doi.org/10.5194/amt-2024-114

Reddington, C. L., Conibear, L., Robinson, S., Knote, C., Arnold, S. R., & Spracklen, D. V. (2021). Air pollution from forest and vegetation fires in Southeast Asia disproportionately impacts the poor. *GeoHealth*, *5*(9), e2021GH000418. https://doi.org/10.1029/2021GH000418

Ren, J., Guo, F., & Xie, S. (2022). Diagnosing ozone–NOx–VOC sensitivity and revealing causes of ozone increases in China based on 2013–2021 satellite retrievals. *Atmospheric Chemistry and Physics*, *22*(22), 15035–15047. https://doi.org/10.5194/acp-22-15035-2022

Rodgers, C. D., & Connor, B. J. (2003). Intercomparison of remote sounding instruments. *Journal of Geophysical Research*, *108*(D3). https://doi.org/10.1029/2002JD002299

Ryan, R. G., Marais, E. A., Gershenson-Smith, E., Ramsay, R., Muller, J. P., Tirpitz, J. L., & Frieß, U. (2023). Measurement report: MAX-DOAS measurements characterise central London ozone pollution episodes during 2022 heatwaves. *Atmospheric Chemistry and Physics*, *23*(12), 7121–7139. https://doi.org/10.5194/acp-23-7121-2023

Sarigiannis, D. A., Karakitsios, S. P., Gotti, A., Liakos, I. L., & Katsoyiannis, A. (2011). Exposure to major volatile organic compounds and carbonyls in European indoor environments and associated health risk. *Environment International*, *37*(4), 743–765. https://doi.org/10.1016/j.envint.2011.01.005

Saunier, A., Ormeño, E., Boissard, C., Wortham, H., Temime-Roussel, B., Lecareux, C., et al. (2017). Effect of mid-term drought on Quercus pubescens BVOCs' emission seasonality and their dependency on light and/or temperature. *Atmospheric Chemistry and Physics*, *17*(12), 7555–7566. https://doi.org/10.5194/acp-17-7555-2017

Shao, M., Lu, S., Liu, Y., Xie, X., Chang, C., Huang, S., & Chen, Z. (2009). Volatile organic compounds measured in summer in Beijing and their role in ground-level ozone formation. *Journal of Geophysical Research*, *114*(D2). https://doi.org/10.1029/2008JD010863

Sindelarova, K., Markova, J., Simpson, D., Huszar, P., Karlicky, J., Darras, S., & Granier, C. (2022). High-resolution biogenic global emission inventory for the time period 2000–2019 for air quality modelling. *Earth System Science Data*, *14*(1), 251–270. https://doi.org/10.5194/essd-14-251-2022

Spinei, E., Tiefengraber, M., Müller, M., Gebetsberger, M., Cede, A., Valin, L., et al. (2021). Effect of polyoxymethylene (POM-H Delrin) off-gassing within the Pandora head sensor on direct-sun and multi-axis formaldehyde column measurements in 2016–2019. *Atmospheric Measurement Techniques*, *14*(1), 647–663. https://doi.org/10.5194/amt-14-647-2021

Spinei, E., Whitehill, A., Fried, A., Tiefengraber, M., Knepp, T. N., Herndon, S., et al. (2018). The first evaluation of formaldehyde column observations by improved Pandora spectrometers during the KORUS-AQ field study. *Atmospheric Measurement Techniques*, *11*(9), 4943–4961. https://doi.org/10.5194/amt-11-4943-2018

Su, W., Liu, C., Hu, Q., Zhang, C., Liu, H., Xia, C., et al. (2022). First global observation of tropospheric formaldehyde from Chinese GaoFen-5 satellite: Locating source of volatile organic compounds. *Environmental Pollution*, *297*, 118691. https://doi.org/10.1016/j.envpol.2021.118691

Sun, W., Zhu, L., De Smedt, I., Bai, B., Pu, D., Chen, Y., et al. (2021). Global significant changes in formaldehyde (HCHO) columns observed from space at the early stage of the COVID-19 pandemic. *Geophysical Research Letters*, *48*(4), 2e020GL091265. https://doi.org/10.1029/2020GL091265

Tirpitz, J. L., Frieß, U., Hendrick, F., Alberti, C., Allaart, M., Apituley, A., et al. (2021). Intercomparison of MAX-DOAS vertical profile retrieval algorithms: Studies on field data from the CINDI-2 campaign. *Atmospheric Measurement Techniques*, *14*(1), 1–35. https://doi.org/10.5194/amt-14-1-2021

Travis, K. R., Judd, L. M., Crawford, J. H., Chen, G., Szykman, J., Whitehill, A., et al. (2022). Can column formaldehyde observations inform air quality monitoring strategies for ozone and related photochemical oxidants? *Journal of Geophysical Research: Atmospheres*, *127*(13), e2022JD036638. https://doi.org/10.1029/2022JD036638

van der Werf, G. R., Randerson, J. T., Giglio, L., van Leeuwen, T. T., Chen, Y., Rogers, B. M., et al. (2017). Global fire emissions estimates during 1997–2016. *Earth System Science Data*, *9*(2), 697–720. https://doi.org/10.5194/essd-9-697-2017

van Geffen, J., Eskes, H., Compernolle, S., Pinardi, G., Verhoelst, T., Lambert, J. C., et al. (2022). Sentinel-5P TROPOMI NO2 retrieval: Impact of version v2.2 improvements and comparisons with OMI and ground-based data. *Atmospheric Measurement Techniques*, *15*(7), 2037–2060. https://doi.org/10.5194/amt-15-2037-2022

Veefkind, J. P., Aben, I., McMullan, K., Förster, H., de Vries, J., Otter, G., et al. (2012). TROPOMI on the ESA Sentinel-5 precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications. *Remote Sensing of Environment*, *120*, 70–83. https://doi.org/10.1016/j.rse.2011.09.027

Vigouroux, C., Langerock, B., Bauer Aquino, C. A., Blumenstock, T., Cheng, Z., De Mazière, M., et al. (2020). TROPOMI–Sentinel-5 Precursor formaldehyde validation using an extensive network of ground-based Fourier-transform infrared stations. *Atmospheric Measurement Techniques*, *13*(7), 3751–3767. https://doi.org/10.5194/amt-13-3751-2020

Wolfe, G. M., Nicely, J. M., St. Clair, J. M., Hanisco, T. F., Liao, J., Oman, L. D., et al. (2019). Mapping hydroxyl variability throughout the global remote troposphere via synthesis of airborne and satellite formaldehyde observations. *Proceedings of the National Academy of Sciences*, *116*(23), 11171–11180. https://doi.org/10.1073/pnas.1821661116

Xu, R., Ye, T., Yue, X., Yang, Z., Yu, W., Zhang, Y., et al. (2023). Global population exposure to landscape fire air pollution from 2000 to 2019. *Nature*, *621*(7979), 521–529. https://doi.org/10.1038/s41586-023-06398-6

Xue, J., Zhao, T., Luo, Y., Miao, C., Su, P., Liu, F., et al. (2022). Identification of ozone sensitivity for NO2 and secondary HCHO based on MAX-DOAS measurements in northeast China. *Environment International*, *160*, 107048. https://doi.org/10.1016/j.envint.2021.107048

Yang, L. H., Jacob, D. J., Dang, R., Oak, Y. J., Lin, H., Kim, J., et al. (2024). Interpreting Geostationary Environment Monitoring Spectrometer (GEMS) geostationary satellite observations of the diurnal variation in nitrogen dioxide (NO2) over East Asia. *Atmospheric Chemistry and Physics*, *24*(12), 7027–7039. https://doi.org/10.5194/acp-24-7027-2024

Yantosca, B., Sulprizio, M., & Lundgren, L. (2023). GEOS-Chem 14.1.1 [Software]. *Zenodo*. https://doi.org/10.5281/zenodo.7696632

Zhang, S., Wang, S., Zhang, R., Guo, Y., Yan, Y., Ding, Z., & Zhou, B. (2021). Investigating the sources of formaldehyde and corresponding photochemical indications at a suburb site in Shanghai from MAX-DOAS measurements. *Journal of Geophysical Research: Atmospheres*, *126*(6), e2020JD033351. https://doi.org/10.1029/2020JD033351

Zhao, X., Griffin, D., Fioletov, V., McLinden, C., Davies, J., Ogyu, A., et al. (2019). Retrieval of total column and surface NO2 from Pandora zenith-sky measurements. *Atmospheric Chemistry and Physics*, *19*(16), 10619–10642. https://doi.org/10.5194/acp-19-10619-2019

Zheng, B., Ciais, P., Chevallier, F., Yang, H., Canadell, J. G., Chen, Y., et al. (2023). Record-high CO_2 emissions from boreal fires in 2021. *Science*, *379*(6635), 912–917. https://doi.org/10.1126/science.ade0805

Zhu, L., Jacob, D. J., Keutsch, F. N., Mickley, L. J., Scheffe, R., Strum, M., et al. (2017). Formaldehyde (HCHO) as a hazardous air pollutant: Mapping surface air concentrations from satellite and inferring cancer Risks in the United States. *Environmental Science and Technology*, *51*(10), 5650–5657. https://doi.org/10.1021/acs.est.7b01356

Zhu, L., Jacob, D. J., Kim, P. S., Fisher, J. A., Yu, K., Travis, K. R., et al. (2016). Observing atmospheric formaldehyde (HCHO) from space: Validation and intercomparison of six retrievals from four satellites (OMI, GOME2A, GOME2B, OMPS) with SEAC4RS aircraft observations over the Southeast US. *Atmospheric Chemistry and Physics*, *16*(21), 13477–13490. https://doi.org/10.5194/acp-16-13477-2016

Zoogman, P., Liu, X., Suleiman, R. M., Pennington, W. F., Flittner, D. E., Al-Saadi, J. A., et al. (2017). Tropospheric emissions: Monitoring of pollution (TEMPO). *Journal of Quantitative Spectroscopy and Radiative Transfer*, *186*, 17–39. https://doi.org/10.1016/j.jqsrt.2016.05.008