https://doi.org/10.5194/egusphere-2026-716
Preprint. Discussion started: 23 February 2026

Direct-sun versus Sky-Scan Pandora Formaldehyde Retrievals: Implications for OMI Validation in Tropical Southeast Asia

Santanasawry A. L. David Arul[1,6], Jackson Hian-Wui Chang[1,2,*], Yong Jie Wong[3], Maggie Chel-Gee Ooi[4], Juneng Liew[5], Fuei Pien Chee[6], Jedol Dayou[6], Justin Sentian[6], Putu Aryastana[7], Neng-Huei Lin[2]

[1]Preparatory Center for Science and Technology, University Malaysia Sabah (UMS), Jalan UMS, 88400, Kota Kinabalu, Sabah, Malaysia
[2]Department of Atmospheric Sciences, National Central University (NCU), Taoyuan, 32001, Taiwan
[3]Department of Bioenvironmental Design, Faculty of Bioenvironmental Sciences, Kyoto University of Advanced Science, Kyoto, 606-8501, Japan
[4]Institute of Climate Change, National University of Malaysia (UKM), Bangi, 43600, Malaysia
[5]Department of Earth Sciences and Environment, Faculty of Sciences and Technology, National University of Malaysia (UKM), Bangi, 43600, Malaysia
[6]Faculty of Science and Technology, University Malaysia Sabah (UMS), Jalan UMS, 88400, Kota Kinabalu, Sabah, Malaysia
[7]Department of Civil Engineering, Faculty of Engineering and Planning, Warmadewa University, Denpasar, 80239, Indonesia

Correspondence to: Jackson Hian-Wui Chang (jacksonchw@ums.edu.my)

Abstract. Ground-based Pandora spectrometers are widely used for validating satellite formaldehyde (HCHO) retrievals; however, the influence of scanning geometry and spatiotemporal representativeness remains insufficiently quantified in tropical environments. This study evaluates Pandora Level-2 HCHO total vertical columns from five Southeast Asian stations (Bangkok, Bandung, Agam, Pontianak, and Singapore-NUS) over 2021–2025, comparing Direct-sun and Sky-scan retrievals and assessing their consistency with OMI Aura observations. HCHO distributions exhibit strong inter-site variability and pronounced skewness, with Direct-sun retrievals showing higher medians and substantially larger variance than Sky-scan observations. Mean Direct-sun HCHO columns are strongly influenced by episodic enhancements at biomass-burning-affected sites, particularly Agam, whereas Sky-scan retrievals display lower central values and reduced variability, consistent with broader atmospheric sampling and diminished sensitivity to localized plumes. Satellite–ground comparisons are conducted using nine spatiotemporal averaging configurations that vary OMI spatial footprints (nearest grid, 3×3, and 5×5) and Pandora temporal averaging. Direct-sun comparisons generally yield weak or unstable correlations ($R \approx -0.1$ to 0.3) and large errors (RMSE ≈ 8–14 × 10^{15} molecules cm^{-2}). In contrast, Sky-scan retrievals show systematically improved agreement, with optimized configurations achieving RMSE values of ~5 × 10^{15} molecules cm^{-2}, MAE of ~4–7 × 10^{15} molecules cm^{-2}, and moderate positive correlations ($R \approx 0.4$–0.6) at several sites. Solar zenith angle–dependent analysis reveals persistent positive biases in Direct-sun retrievals (~10–20 × 10^{15} molecules cm^{-2}), while Sky-scan retrievals exhibit near-zero bias at low to moderate SZAs and substantially reduced extremes. Overall, the results demonstrate that scanning

https://doi.org/10.5194/egusphere-2026-716
Preprint. Discussion started: 23 February 2026

geometry exerts a first-order control on Pandora–OMI consistency in the tropics, with Sky-scan observations providing a more spatially representative reference for satellite validation, although optimal configurations remain site dependent.

Keywords. Formaldehyde (HCHO); Pandora; Direct-sun; Sky-scan; Ozone Monitoring Instrument (OMI); Southeast Asia; Spatiotemporal averaging

1. Introduction

Formaldehyde (HCHO) is a key intermediate in tropospheric photochemistry and one of the most important carbonyl compounds in the atmosphere. It is produced primarily through the oxidation of volatile organic compounds (VOCs) from both biogenic and anthropogenic sources and serves as an effective proxy for VOC emissions at local to regional scales (Nascimento et al., 2022; Zhu et al., 2017). Through photolysis, HCHO represents a major source of hydroperoxy (HO_2) radicals, thereby enhancing ozone (O_3) production in the presence of nitrogen oxides (NO_x) and contributing to secondary organic aerosol formation. Owing to its short atmospheric lifetime, typically on the order of hours (Lim et al., 2019), HCHO exhibits strong spatial and temporal variability and is highly sensitive to changes in emissions, meteorology, and photochemical activity (Fang et al., 2017; Liao et al., 2021; Lim et al., 2019). Accurate characterization of HCHO is therefore essential for understanding air quality, constraining chemical transport models, and evaluating emission control strategies, particularly in regions with intense photochemistry and episodic pollution.

A range of techniques exists for measuring atmospheric HCHO, including in situ sensors and ground-based remote sensing methods such as differential optical absorption spectroscopy (DOAS) (Liu et al., 2020; Pinardi et al., 2013), Fourier transform infrared spectroscopy (FTIR) (Jones et al., 2009; Vigouroux et al., 2018), and cavity-enhanced absorption spectrometers (Glowania et al., 2021). While these instruments can provide high-precision observations, their deployment is limited by cost, logistical complexity, and maintenance requirements (Lee et al., 2024; Tian et al., 2019). As a result, routine HCHO measurements are rarely included in national air quality monitoring networks, leading to substantial observational gaps, especially in rapidly developing regions such as Southeast Asia. Satellite remote sensing has therefore become a critical tool for monitoring HCHO, offering consistent spatial coverage and long-term observations that enable the identification of emission hotspots, seasonal variability, and regional trends.

Among satellite instruments, the Ozone Monitoring Instrument (OMI) onboard NASA's Aura satellite has provided global HCHO observations since 2004 (Tanskanen et al., 2006), with near-daily coverage and a nadir footprint of approximately 13 × 24 km² (Ahn et al., 2008). OMI HCHO products have been widely used for air quality and atmospheric chemistry studies (Nascimento et al., 2022; Zhu et al., 2017b); however, their accuracy is affected by cloud contamination, aerosol loading, surface reflectance, and viewing geometry. Consequently, robust validation using independent ground-based measurements remains essential (Harkey et al., 2021). In recent years, the Pandonia Global Network (PGN) of Pandora spectrometers has

https://doi.org/10.5194/egusphere-2026-716
Preprint. Discussion started: 23 February 2026

emerged as a key resource for satellite validation. Pandora instruments retrieve total vertical columns of trace gases using high-resolution UV–visible spectroscopy and offer standardized, long-term observations across a growing global network.

For HCHO, Pandora spectrometers provide two physically distinct retrieval modes—Direct-sun and Sky-scan—yet their differing sensitivities and implications for satellite validation remain insufficiently quantified (Herman et al., 2015). Direct-sun retrievals, based on direct solar viewing, exhibit high precision and strong sensitivity to near-surface HCHO but sample a narrow atmospheric column and are therefore prone to localized plume influences and solar-geometry-dependent biases (Ghahremanloo et al., 2025). In contrast, Sky-scan retrievals sample scattered skylight across multiple viewing angles, yielding a more spatially integrated column that may better approximate the effective footprint of satellite instruments (Herman et al., 2009). Despite these fundamental differences, most previous validation studies have implicitly treated Pandora HCHO as a single product, without explicitly assessing how scan geometry controls representativeness, bias behavior, and agreement with satellite retrievals (Tzortziou et al., 2012). This lack of distinction is particularly consequential in tropical environments, where strong emission heterogeneity, biomass burning, and rapid photochemical production amplify sub-pixel variability (Herman et al., 2015). A systematic evaluation of Direct-sun versus Sky-scan Pandora HCHO, and their respective consistency with satellite observations, therefore represents a critical but largely unexplored gap in current validation frameworks.

Validation efforts for satellite HCHO products have largely focused on mid-latitude regions in North America, Europe, and East Asia (Palmer et al., 2003; Spinei et al., 2018; Zhu et al., 2016), often leveraging intensive field campaigns. Comparatively few studies have examined tropical environments, where high solar irradiance, frequent convection, complex cloud fields, and recurrent biomass burning introduce additional challenges for both satellite and ground-based retrievals(Hansen et al., 2019). Southeast Asia is a particularly critical yet underexplored region, characterized by dense urban emissions, seasonal agricultural burning, and persistent transboundary haze, all of which drive strong variability in HCHO (Cheong et al., 2019; Fu et al., 2007). These conditions amplify the importance of understanding how ground-based sampling geometry interacts with satellite spatial resolution.

Recent satellite instruments such as the TROPOspheric Monitoring Instrument (TROPOMI) provide substantially higher spatial resolution than OMI and enable improved detection of localized HCHO enhancements under favorable conditions (Lee et al., 2024; Su et al., 2020). However, higher spatial resolution alone does not eliminate representativeness errors when comparing satellite and ground-based observations, particularly in heterogeneous tropical environments (Boersma et al., 2016). TROPOMI HCHO retrievals remain sensitive to cloud fraction, aerosol loading, and surface reflectance, and the smaller pixel size can increase susceptibility to localized plumes that are not spatially representative of satellite-scale columns (Boersma et al., 2016; De Smedt et al., 2018). In persistently cloudy regions such as Southeast Asia, these sensitivities often lead to fragmented temporal coverage, limiting the robustness of daily-scale and climatological validation

https://doi.org/10.5194/egusphere-2026-716
Preprint. Discussion started: 23 February 2026

analyses. In contrast, OMI's coarser footprint and more continuous data availability provide a stable reference for isolating first-order effects related to measurement geometry and spatiotemporal representativeness (Herman et al., 2019; Vasilkov et al., 2017). Importantly, the effective atmospheric sampling of Pandora Sky-scan retrievals more closely approximates the OMI footprint than that of Direct-sun observations, making OMI particularly well suited for evaluating how Pandora scanning geometry influences satellite–ground consistency (Herman et al., 2009; Tzortziou et al., 2012). As such, OMI–Pandora intercomparisons offer a physically grounded framework for diagnosing retrieval biases that arise from sampling geometry rather than spatial resolution alone. Insights gained from this analysis therefore provide a critical foundation for the interpretation and future validation of higher-resolution HCHO products from TROPOMI, TEMPO, and Sentinel-4. and upcoming geostationary missions, which are very lacking in tropical Southeast Asia environments.

In this study, we present a comprehensive evaluation of Pandora HCHO measurements from five Southeast Asian sites—Bangkok, Bandung, Agam, Pontianak, and Singapore—explicitly distinguishing between Direct-sun and Sky-scan retrievals and assessing their consistency with OMI satellite observations. The objectives are to (i) quantify systematic differences between Direct-sun and Sky-scan HCHO retrievals, (ii) evaluate how spatiotemporal averaging strategies influence Pandora–OMI agreement for each retrieval mode, and (iii) assess the role of solar zenith angle and atmospheric representativeness in driving retrieval biases. By explicitly addressing scan geometry as a first-order factor in satellite validation, this work fills a critical methodological gap and provides guidance for the optimal use of Pandora HCHO data in tropical environments. The findings offer important implications for current and future satellite missions and establish a robust framework for HCHO validation in one of the most chemically active yet observationally sparse regions of the world.

2. Method

2.1 Ground-Based Pandonia Global Network of HCHO Observations

This study uses formaldehyde (HCHO) observations from five Pandora spectrometer systems located across Southeast Asia (Fig. 1), a region characterized by persistent high solar angles, frequent cloudiness, and strong spatiotemporal variability. Table 1 summarizes the Pandora stations included in this analysis, including geographic location, altitude, product status, and data availability. While previous satellite validation studies have largely focused on Pandora sites in mid-latitude regions (e.g., North America, Europe, and East Asia; Lee et al., 2024; Spinei et al., 2018; Tzortziou et al., 2012), relatively few have examined low-latitude and equatorial stations. This distinction is important, as tropical observing conditions introduce challenges related to solar geometry, cloud cover, and atmospheric heterogeneity that may influence retrieval performance and representativeness.

https://doi.org/10.5194/egusphere-2026-716
Preprint. Discussion started: 23 February 2026

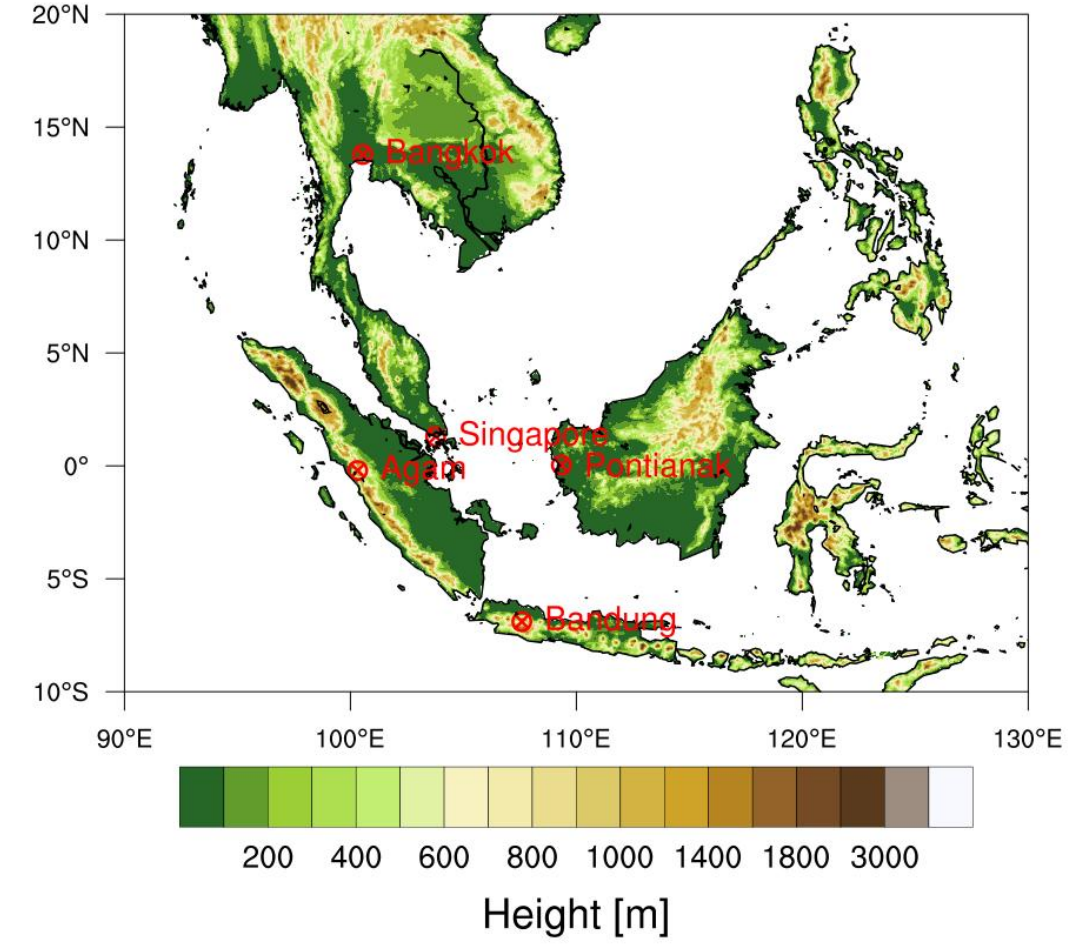


Figure 1. Geographic distribution of the Pandora observation sites used in this study, including Bangkok (a), Bandung (b), Agam (c), Pontianak (d), and Singapore-NUS (e). All sites utilize Pandora data version rfus5p1-8 (Direct-sun) and rfuh5p1-8 (Sky-scan).

All Pandora instruments analyzed here operate using both Direct-sun and Sky-scan viewing geometries, allowing a systematic evaluation of geometry-dependent retrieval behavior (Herman et al., 2015). We use Level-2 HCHO products from the rfus5p1-8 and rfuh5p1-8 processing streams (last accessed: 27 February 2025), which provide total vertical column HCHO retrievals derived from direct-Sun and diffuse-sky measurements. The rfus5p1-8 product is emphasized in this study owing to its improved numerical stability and reduced noise relative to rfuh5p1-8, which incorporates horizon scans and exhibits greater sensitivity to heterogeneous cloud and aerosol conditions. The availability of both Direct-sun and Sky-scan products enables a consistent assessment of how viewing geometry influences HCHO retrieval characteristics under tropical conditions.

https://doi.org/10.5194/egusphere-2026-716
Preprint. Discussion started: 23 February 2026

The selected stations include Bangkok (190s1, 13.78°N, 100.54°E, 60 m a.s.l.), an urban megacity with heavy traffic and industrial emissions; Bandung (210s1, −6.89°S, 107.59°E, 752 m a.s.l.), a highland city in Indonesia surrounded by volcanic mountains and agricultural activity; Agam (211s1, −0.20°S, 100.32°E, 865 m a.s.l.), a remote and elevated background site in West Sumatra with limited anthropogenic influence; Pontianak (212s1, 0.04°N, 109.34°E, 1 m a.s.l.), a coastal equatorial station in West Kalimantan, Indonesia, known for frequent cloud cover and convective activity; and Singapore-NUS (77s1, 1.30°N, 103.77°E, 77 m a.s.l.), an urban tropical island site with a dense population and significant marine and urban air interactions. All stations provide "Official" quality data under the Pandora global network, with observational periods ranging from May 2021 to February 2025. Collectively, this network offers a valuable opportunity to evaluate satellite HCHO products in complex tropical environments that are typically underrepresented in validation studies.

2.2 OMI/Aura HCHO Total Column Retrievals

The Ozone Monitoring Instrument (OMI), onboard NASA's Aura satellite, is a nadir-viewing hyperspectral sensor launched in 2004. OMI formaldehyde (HCHO) retrievals have undergone several algorithm updates, with the current OMI SIPS HCHO Version 003 product widely used in long-term and regional analyses (Herman et al., 2018; Lamsal et al., 2014). The retrieval is based on differential optical absorption spectroscopy (DOAS) and provides tropospheric vertical column densities at a nadir spatial resolution of approximately 13 × 24 km². Standard quality filtering was applied following established recommendations, including qa_flag = 0 (good quality), solar zenith angle (SZA) < 60°, and cloud radiance fraction < 0.3 (Johnson et al., 2024). OMI has a fixed local overpass time of approximately 13:30 LT, enabling direct comparison with Pandora observations filtered to the same midday period. Although OMI retrievals are subject to pixel-level noise and cloud-related uncertainties, their long-term continuity and global coverage make them a key reference dataset for satellite–ground intercomparisons, particularly in regions where in situ and ground-based measurements are limited (Harkey et al., 2021).

3. Results

3.1 Pandora HCHO Total Vertical Column: Direct-sun vs Sky-scan

Pandora Level-2 formaldehyde (HCHO) total vertical column retrievals (rfus5p1-8 and rfuh5p1-8) from five Southeast Asian stations were analyzed for the period 2021–2025. The dataset includes both Direct-sun (rfus5p1-8) and Sky-scan (rfuh5p1-8) observations, providing complementary sampling of atmospheric HCHO under differing viewing geometries. Station characteristics and data availability are summarized in Table 1, while Table 2 reports the number of valid observations retained after filtering negative, missing, or otherwise unusable retrievals.

https://doi.org/10.5194/egusphere-2026-716
Preprint. Discussion started: 23 February 2026

Table 1. Summary of Pandora monitoring stations used in this study, including location, altitude, product status, and data availability. Data description: Formaldehyde (HCHO) Level 2, Version: rfus5p1-8 and rfuh5p1-8 (Last accessed: 27 Feb 2025).

Station ID	Station Name	Lat	Lon	Altitude (m)	rfus5p1-8 and rfuh5p1-8		
					Product Status	Data Start	Last Updated
190s1	Bangkok	13.7847	100.5400	60	Official	20210520	20250221
210s1	Bandung	-6.8948	107.5865	752	Official	20230611	20240920
211s1	Agam	-0.2046	100.3195	865	Official	20220913	20240521
212s1	Pontianak	0.0415	109.3366	1	Official	20240309	20250226
77s1	Singapore	1.2990	103.7710	77	Official	20230621	20250226

Table 2. Summary of Pandora Formaldehyde (HCHO) Level 2 data Direct-sun (Version rfus5p1-8) and sky-scan (Version rfuh5p1-8) used in this study. Data flagged as unusable, missing, or negative were excluded from the analysis (last accessed 27 February 2025).

Station ID	Station Name	L2 QC Flag			Missing Value	Negative Value	Total Used
		Assured	Not-Assured	Unusable			
Direct-sun rfus5p1-8							
190s1	Bangkok	0	81169	0	0	833	80336
210s1	Bandung	0	35671	0	339	1027	34305
211s1	Agam	0	44450	10171	9881	3594	32188
212s1	Pontianak	0	27224	0	0	1526	25698
77s1	Singapore	7646	33660	0	0	1506	39800
Sky-scan rfuh5p1-8							
190s1	Bangkok	0	138343	0	1	2657	135685
210s1	Bandung	0	49816	0	48	2422	47346
211s1	Agam	0	49546	0	8	9023	40515
212s1	Pontianak	0	38395	0	0	2098	36297
77s1	Singapore	0	64021	0	0	2536	61485

In total, 80,336 Direct-sun measurements were available from Bangkok (190s1), the majority of which were classified as "Not-Assured" owing to the absence of "Assured" flags during the study period. Bandung (210s1) contributed 34,305 valid Direct-sun retrievals, with approximately 1 % excluded based on quality flags. Agam (211s1) yielded 32,188 Direct-sun observations, but also exhibited the highest fraction of excluded data, with a substantial number of negative and unusable values removed during preprocessing. Pontianak (212s1), a recently established site, provided 25,698 Direct-sun retrievals, all classified as "Not-Assured." Singapore–NUS (77s1) contributed 39,800 Direct-sun observations, including 7,646 "Assured" retrievals, representing the only station with a notable fraction of high-confidence data. Sky-scan observations were more numerous overall, with 135,685 valid retrievals at Bangkok, 47,346 at Bandung, 40,515 at Agam, 36,297 at Pontianak, and 61,485 at Singapore–NUS. As with Direct-sun data, unusable and negative Sky-scan retrievals were removed prior to analysis.

https://doi.org/10.5194/egusphere-2026-716
Preprint. Discussion started: 23 February 2026

Table 3. Descriptive statistics of Pandora Level-2 total column formaldehyde (HCHO) observations obtained from Direct-sun (Version rfus5p1-8) and sky-scan (Version rfuh5p1-8) retrievals at selected Southeast Asian stations. Statistics are presented for hourly-averaged and daily-averaged data, which include the mean ± standard deviation (SD), median with interquartile range (IQR; Q1–Q3), minimum–maximum values, and the number of valid observations (n). All HCHO columns are scaled to 1×10^{15} molecules cm^{-2}. Data were last accessed on 27 February 2025.

(a) Direct-sun HCHO

Station ID	Station	Time scale	Mean ± SD	Median (IQR)	Min–Max	n
190s1	Bangkok	Hourly	20.53 ± 7.64	19.95 (15.37–25.01)	0.09–50.99	746
		Daily	20.30 ± 7.20	19.77 (15.44–24.41)	0.09–42.54	571
210s1	Bandung	Hourly	19.83 ± 34.56	14.84 (9.86–21.08)	0.28–428.18	339
		Daily	21.20 ± 40.42	15.24 (10.11–21.12)	0.28–428.18	231
211s1	Agam	Hourly	407.52 ± 985.19	9.87 (7.17–56.11)	0.03–8103.20	310
		Daily	237.89 ± 720.51	9.05 (6.36–12.55)	0.09–5844.17	196
212s1	Pontianak	Hourly	11.89 ± 4.88	11.81 (9.00–14.70)	0.53–28.64	237
		Daily	11.92 ± 4.73	11.81 (8.95–14.27)	2.63–28.64	160
77s1	Singapore	Hourly	10.62 ± 5.95	9.34 (6.91–12.93)	0.32–45.79	344
		Daily	10.81 ± 5.80	9.41 (7.19–13.16)	0.32–45.79	249

(b) Sky-scan HCHO

Station ID	Station	Time scale	Mean ± SD	Median (IQR)	Min–Max	n
190s1	Bangkok	Hourly	13.54 ± 35.77	11.90 (8.07–15.99)	0.00–1241.50	1218
		Daily	14.11 ± 43.85	12.01 (8.54–15.90)	0.00–1241.50	799
210s1	Bandung	Hourly	13.14 ± 10.88	11.35 (7.16–16.17)	0.17–151.22	436
		Daily	13.11 ± 8.69	11.80 (7.64–16.55)	0.22–84.76	288
211s1	Agam	Hourly	$2.26\times10^{7} \pm 4.28\times10^{8}$	4.21 (2.50–6.87)	$0.06–8.1\times10^{9}$	359
		Daily	$3.31\times10^{7} \pm 5.19\times10^{8}$	4.39 (2.79–6.78)	$0.06–8.1\times10^{9}$	245
212s1	Pontianak	Hourly	7.93 ± 5.31	7.04 (4.57–9.96)	0.15–55.68	306
		Daily	7.92 ± 5.47	6.98 (4.69–9.54)	0.63–55.68	190
77s1	Singapore	Hourly	10.96 ± 7.78	9.13 (6.17–12.99)	0.42–63.73	554
		Daily	11.17 ± 7.13	9.41 (6.60–13.10)	0.90–56.64	362

Descriptive statistics for hourly and daily averaged HCHO columns are summarized in Table 3. At all stations, arithmetic means substantially exceed medians, indicating strongly positively skewed distributions dominated by episodic high-HCHO events rather than persistent background levels. This skewness is most pronounced in Direct-sun retrievals. Bangkok exhibits consistently elevated HCHO columns, with relatively small differences between mean and median values, reflecting stable and sustained enhancements. Bandung shows a larger separation between mean and median and a broad distribution with extreme maxima, indicative of strong intermittency. The most extreme skewness is observed at Agam, where median Direct-sun HCHO columns remain low (≈10×10^{15} molecules cm^{-2}), while mean values are strongly inflated by rare but very large enhancements, resulting in maxima exceeding $10^{3}–10^{4} \times 10^{15}$ molecules cm^{-2}. This divergence highlights the sensitivity of mean statistics to extreme events and underscores the importance of robust measures such as medians for characterizing typical conditions at this site. In contrast, Pontianak and Singapore–NUS display comparatively compact distributions, with smaller mean–median differences and narrower ranges.

Sky-scan retrievals exhibit systematically lower median HCHO columns and reduced variance relative to Direct-sun observations at all stations. While extreme values are still present, particularly at Agam, their influence on summary statistics is substantially reduced, and median Sky-scan values (≈$4–5 \times 10^{15}$ molecules cm^{-2} at Agam) are comparable to those at other

https://doi.org/10.5194/egusphere-2026-716
Preprint. Discussion started: 23 February 2026

regional sites. This behavior is consistent with the broader effective sampling of Sky-scan observations, which reduces sensitivity to localized enhancements and short-lived extremes.

Frequency distributions of hourly and daily HCHO columns (Figs. 2 and 3) further illustrate these characteristics. Hourly Direct-sun distributions are generally broad and right-skewed, with pronounced high-value tails at Bandung and Agam, whereas daily averages show narrower distributions with damped extremes due to temporal averaging. Despite these differences, the relative ranking of stations is preserved across temporal resolutions and retrieval modes, indicating robust spatial contrasts captured by the Pandora network. Sky-scan distributions are consistently smoother and less skewed than their Direct-sun counterparts, although the overall inter-station patterns remain similar.

Time series of hourly and daily HCHO columns derived from Direct-sun and Sky-scan observations (Fig. 4) show substantially greater scatter at hourly resolution across all sites, reflecting short-term variability superimposed on longer-term background levels. Daily averaging effectively suppresses high-frequency fluctuations while preserving persistent inter-site differences, confirming that observed spatial contrasts are not driven solely by isolated extreme events. Together, these results demonstrate that while Direct-sun retrievals are highly sensitive to episodic enhancements, Sky-scan observations provide a more stable representation of background HCHO conditions across diverse tropical environments. Time series analysis reveals clear station-dependent behaviour. Bangkok shows persistently elevated and temporally stable HCHO columns, while Bandung and Agam exhibit stronger intermittency characterized by sporadic enhancements superimposed on lower background levels. Pontianak and Singapore–NUS display comparatively stable and lower HCHO columns with limited temporal dispersion.

https://doi.org/10.5194/egusphere-2026-716
Preprint. Discussion started: 23 February 2026

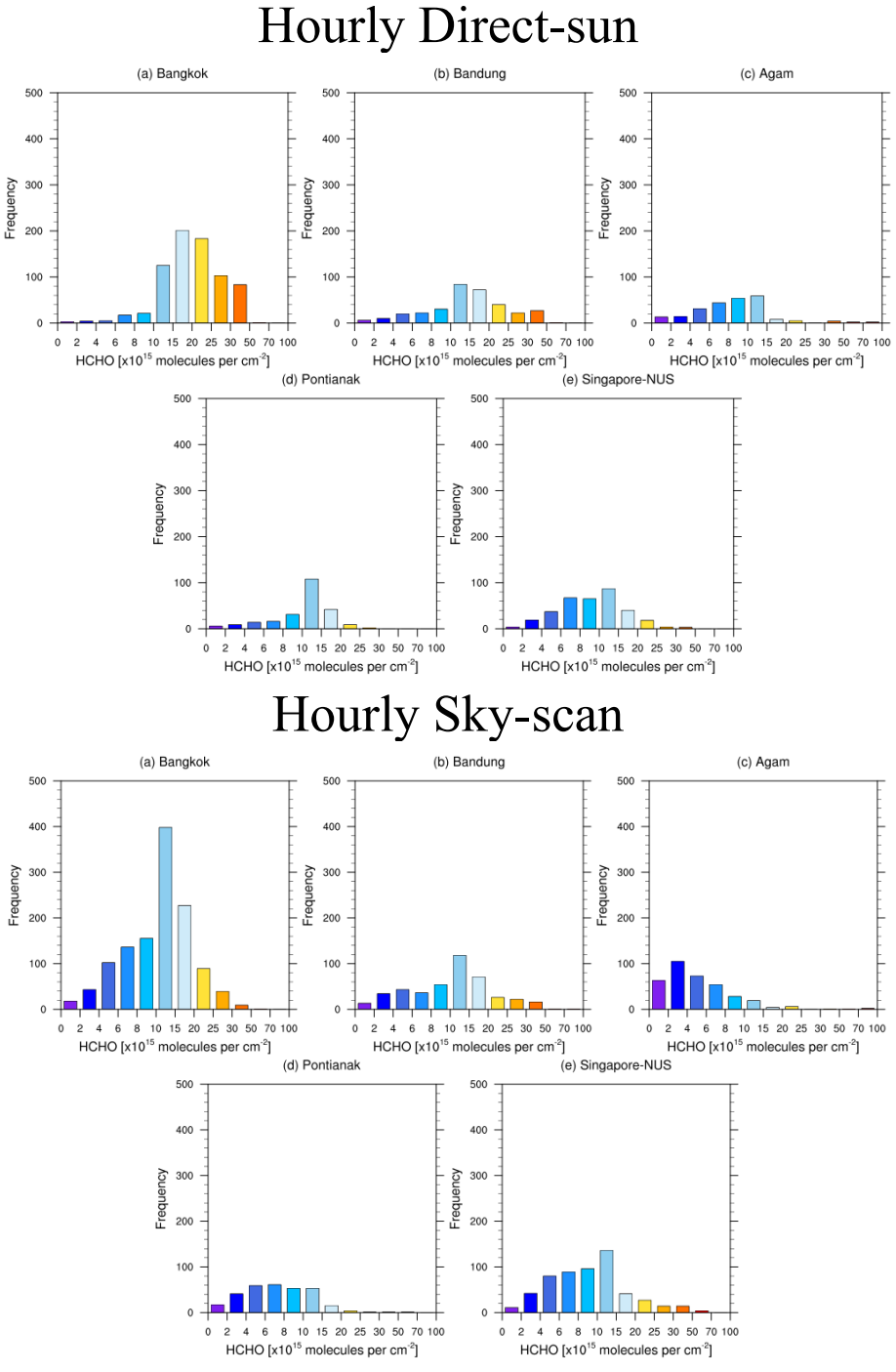


Figure 2: Frequency distribution of hourly average HCHO column densities measured by Pandora monitoring stations in Bangkok (a), Bandung (b), Agam (c), Pontianak (d), and Singapore-NUS (e) from 2021 to 2024, using data version rfus5p1-8.

https://doi.org/10.5194/egusphere-2026-716
Preprint. Discussion started: 23 February 2026

265

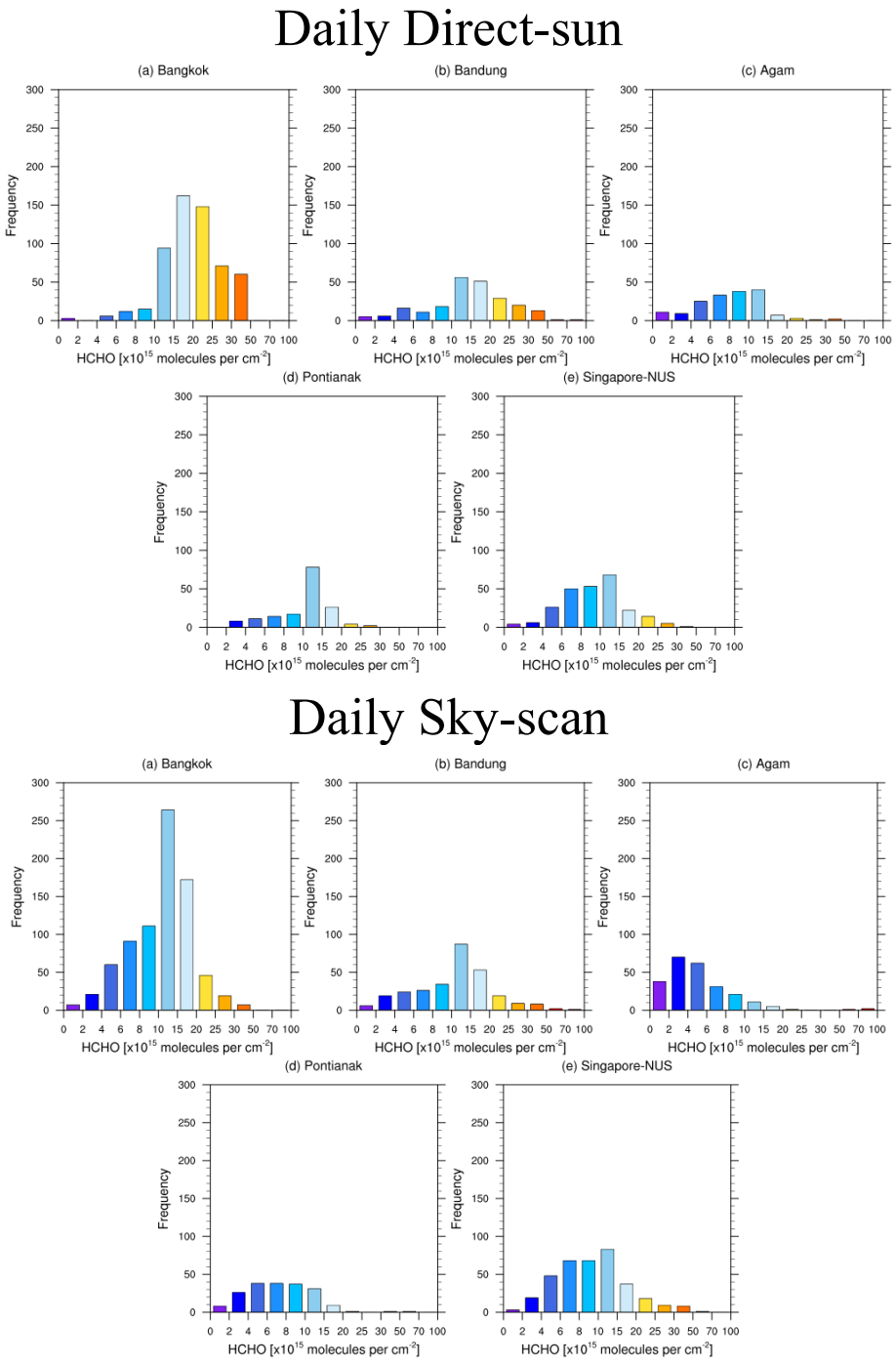


Figure 3: Frequency distribution of daily average HCHO column densities measured by Pandora monitoring stations in Bangkok (a), Bandung (b), Agam (c), Pontianak (d), and Singapore-NUS (e) from 2021 to 2024, using data version rfus5p1-8 (Direct-sun) and rfuh5p1-8 (Sky-scan).

https://doi.org/10.5194/egusphere-2026-716
Preprint. Discussion started: 23 February 2026

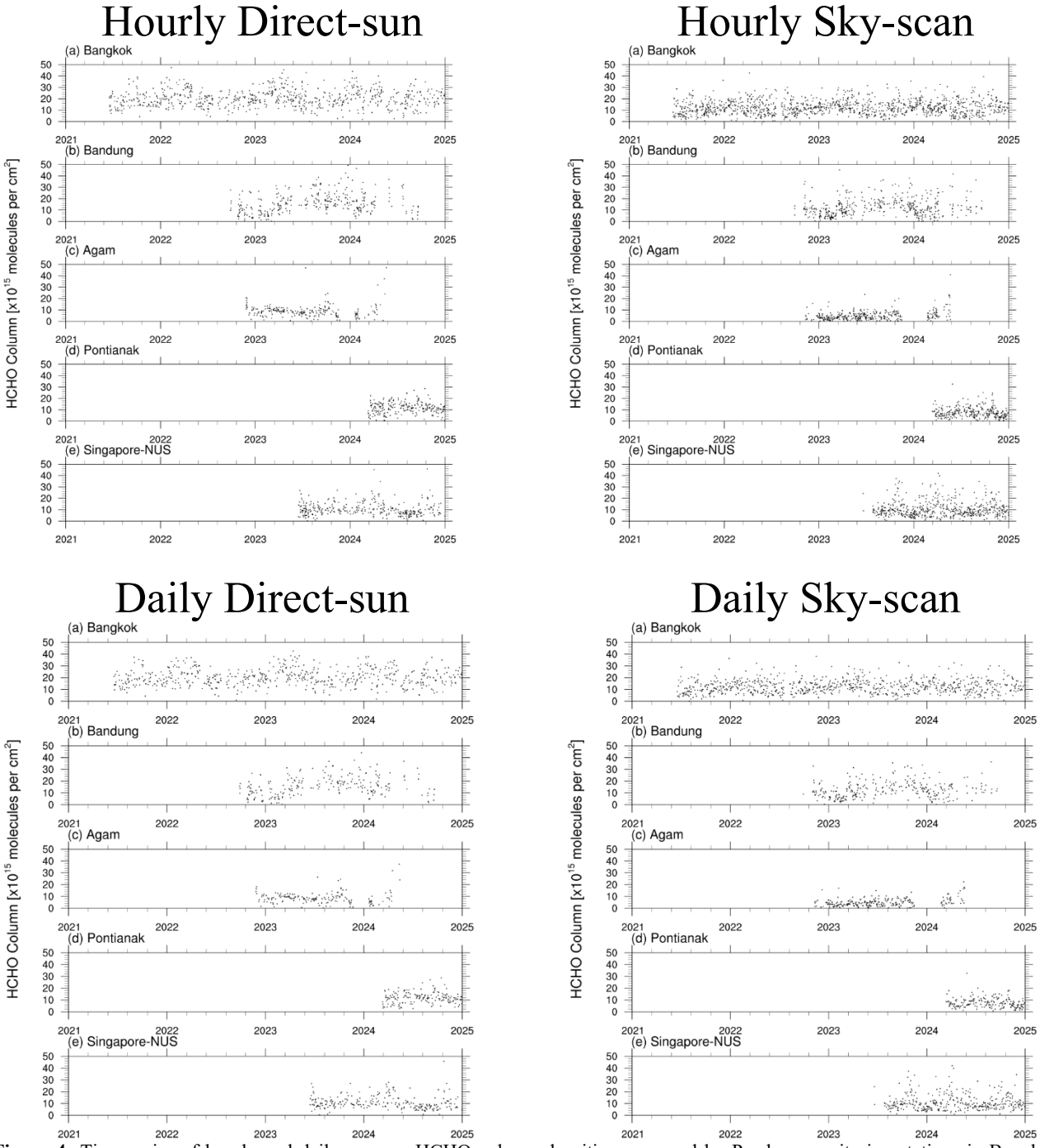


Figure 4: Time series of hourly and daily average HCHO column densities measured by Pandora monitoring stations in Bangkok (a), Bandung (b), Agam (c), Pontianak (d), and Singapore-NUS (e) from 2021 to 2024, using data version rfus5p1-8 (Direct-sun) and rfuh5p1-8 (sky-scan). Data greater than 50 x 10^{15} molecules per cm^2 are filtered.

https://doi.org/10.5194/egusphere-2026-716
Preprint. Discussion started: 23 February 2026

Table 4: Experimental design for the comparative assessment of HCHO column densities retrieved from Pandora and OMI. Each experiment (E1–E9) varies in OMI spatial averaging (nearest pixel, 3×3, 5×5) and Pandora temporal averaging (daily average, daytime, noontime) to evaluate the impact of spatial and temporal matching on dataset agreement.

Exp	OMI spatial averaging	Pandora temporal averaging	Remarks
E1	Nearest lat/lon	00:00 – 23:00 LT	Daily average
E2		07:00 – 09:00 LT	Daytime
E3		13:00 – 15:00 LT	Noontime
E4	3 x 3 (~50 km)	00:00 – 23:00 LT	Daily average ~50 km
E5		07:00 – 09:00 LT	Daytime ~50 km
E6		13:00 – 15:00 LT	Noontime ~50 km
E7	5 x 5 (~100 km)	00:00 – 23:00 LT	Daily average ~100 km
E8		07:00 – 09:00 LT	Daytime ~100 km
E9		13:00 – 15:00 LT	Noontime ~100 km

3.2 Comparative Assessment between Pandora and OMI HCHO

To quantify the influence of spatiotemporal averaging on satellite–ground consistency, nine experimental configurations (E1–E9; Table 4) were applied separately to Direct-sun and Sky-scan Pandora HCHO retrievals. The configurations systematically vary OMI spatial averaging (nearest grid cell, 3×3, and 5×5 domains centered on each station) and Pandora temporal averaging (daily, daytime, and noontime). This framework enables an explicit evaluation of representativeness effects in satellite–ground comparisons, particularly at daily timescales that remain underexplored in tropical validation studies.

Statistical performance metrics for all configurations and stations are summarized in Table 5. Overall, satellite–ground agreement is highly sensitive to both spatial and temporal filtering, with clear contrasts between Direct-sun and Sky-scan retrieval modes. Direct-sun comparisons generally exhibit weak or unstable correlations, with RMSE values ranging from ~5 to >14 × 10^{15} molecules cm^{-2} and correlation coefficients frequently near zero or negative. Moderate positive correlations ($R \approx 0.4–0.5$) are obtained only under a limited subset of configurations (e.g., E2, E5, and E9) and primarily at Bangkok and Singapore–NUS.

In contrast, Sky-scan retrievals show systematically improved agreement with OMI across most stations. RMSE and MAE values are consistently lower than for Direct-sun retrievals, and positive correlations are more robust, particularly for configurations combining intermediate OMI spatial averaging (3×3 or 5×5) with temporally constrained Pandora sampling (e.g., E7–E8). The best-performing configurations achieve RMSE values as low as ~5 × 10^{15} molecules cm^{-2} and correlation coefficients up to ~0.4–0.6 at Bangkok, Bandung, and Agam. Despite these improvements, no single configuration optimizes all metrics across all sites, highlighting persistent site-dependent representativeness differences.

https://doi.org/10.5194/egusphere-2026-716
Preprint. Discussion started: 23 February 2026

Table 5: Statistical evaluation (RMSE, MAE, and R) of OMI and Pandora HCHO column comparisons across five Southeast Asian stations (Bangkok, Bandung, Agam, Pontianak, and Singapore-NUS) under nine experimental configurations (E1–E9). Each experiment uses daily configuration and varies in OMI spatial averaging (nearest grid, 3×3, or 5×5) and Pandora temporal averaging (daily, daytime, noontime) to assess the influence of spatiotemporal filtering on satellite-ground agreement. Bold numbers (**in red color**) represent the best metric among all configurations (E1-E9).

(a) Direct-sun

	Bangkok			Bandung			Agam			Pontianak			Singapore-NUS		
Exp	**RMSE**	**MAE**	**R**	**RMSE**	**MAE**	**R**	**RMSE**	**MAE**	**R**	**RMSE**	**MAE**	**R**	**RMSE**	**MAE**	**R**
E1	11.70	9.85	-0.08	13.22	10.65	-0.08	13.22	10.65	0.05	7.32	5.86	-0.04	8.90	6.40	0.09
E2	8.91	6.58	**0.50**	**5.40**	**4.47**	**0.42**	19.25	15.21	-0.74	**6.02**	5.07	**0.14**	5.85	4.41	0.09
E3	13.21	11.49	0.04	13.91	9.48	0.25	10.97	7.67	-0.30	8.86	7.32	-0.26	7.57	5.77	0.11
E4	10.44	8.56	0.12	10.10	7.84	-0.16	9.45	7.53	0.03	6.61	5.24	-0.02	8.33	6.02	0.00
E5	**7.62**	**6.09**	0.48	7.19	6.08	-0.40	5.22	4.68	-0.16	6.17	**4.95**	0.13	5.57	4.37	0.13
E6	11.32	9.92	0.16	14.56	11.05	-0.32	8.36	7.42	0.37	7.66	6.31	-0.20	6.50	4.99	0.20
E7	10.46	8.63	0.09	9.46	7.39	0.01	11.47	8.37	-0.09	6.57	5.25	-0.02	7.44	5.26	0.05
E8	7.97	6.40	0.38	7.13	6.42	0.17	**4.95**	**4.40**	-0.23	7.33	5.78	-0.11	5.95	**4.20**	0.12
E9	11.30	9.73	0.21	14.09	10.15	0.05	7.81	7.15	**0.46**	7.50	6.06	-0.09	**5.53**	4.23	**0.32**

(b) Sky-scan

	Bangkok			Bandung			Agam			Pontianak			Singapore-NUS		
Exp	**RMSE**	**MAE**	**R**	**RMSE**	**MAE**	**R**	**RMSE**	**MAE**	**R**	**RMSE**	**MAE**	**R**	**RMSE**	**MAE**	**R**
E1	8.34	6.45	0.16	8.68	6.80	0.30	11.29	9.49	0.25	7.92	5.63	0.00	9.34	6.91	0.06
E2	8.46	6.93	0.34	8.71	6.44	0.28	9.18	6.97	0.35	10.56	8.59	-0.15	11.36	9.84	-0.14
E3	9.66	6.96	-0.02	10.09	8.19	-0.07	11.72	10.17	**0.60**	6.84	4.88	**0.26**	9.72	6.82	0.05
E4	7.05	5.53	0.11	7.29	5.54	0.18	8.74	7.21	-0.06	5.72	4.35	-0.10	8.60	6.32	-0.01
E5	7.84	6.70	0.36	7.66	5.55	**0.32**	10.21	8.15	-0.00	5.88	4.88	-0.02	9.80	7.94	-0.06
E6	7.82	5.62	-0.04	8.55	6.50	-0.14	12.04	9.96	-0.28	5.86	4.70	-0.07	9.06	6.62	0.02
E7	**6.90**	**5.46**	0.12	6.88	5.33	0.13	**7.74**	**6.49**	-0.17	5.59	4.21	-0.07	7.80	**5.74**	0.01
E8	7.99	6.88	**0.44**	**5.09**	**4.12**	0.22	7.81	6.53	0.46	**5.00**	**4.16**	0.14	**7.45**	6.59	**0.11**
E9	7.53	5.48	0.05	7.81	6.00	-0.03	11.41	8.89	-0.55	5.60	4.48	-0.04	8.46	6.27	0.08

Figure 5 illustrates the day-to-day correspondence between OMI and Pandora HCHO columns for selected configurations, including the baseline E1 and the optimized configurations for Direct-sun (E2) and Sky-scan (E8). Direct-sun time series exhibit substantial scatter relative to OMI across all stations, with intermittent alignment punctuated by large deviations. In contrast, Sky-scan observations show reduced dispersion and clearer co-variability with OMI under the optimized configuration, particularly at Bangkok and Bandung. This indicates that Sky-scan retrievals more closely capture variability at the spatial scale sampled by the satellite.

Frequency distributions of daily mean HCHO columns (Fig. 6) further support these findings. While both Pandora and OMI exhibit log-normal behavior across all stations, Sky-scan distributions show closer alignment with OMI than Direct-sun retrievals, particularly in the central range of HCHO values. The improvement is most pronounced at Bangkok, where the Sky-scan configuration substantially reduces discrepancies in modal values and distribution shape relative to OMI.

Scatter plot analyses (Figs. 7 and 8) quantitatively confirm the benefits of optimized configurations. Under the baseline E1 setup, both retrieval modes show weak correspondence with OMI, with near-zero correlations at several sites. Substantial improvements are achieved under the optimized Direct-sun (E2) and Sky-scan (E8) configurations, with reduced error metrics and strengthened correlations across all stations. The strongest improvements are observed for Sky-scan E8, where data cluster more closely around the 1:1 line and both RMSE and MAE are minimized, confirming the effectiveness of combined spatial and temporal averaging in improving satellite–ground HCHO consistency.

https://doi.org/10.5194/egusphere-2026-716
Preprint. Discussion started: 23 February 2026

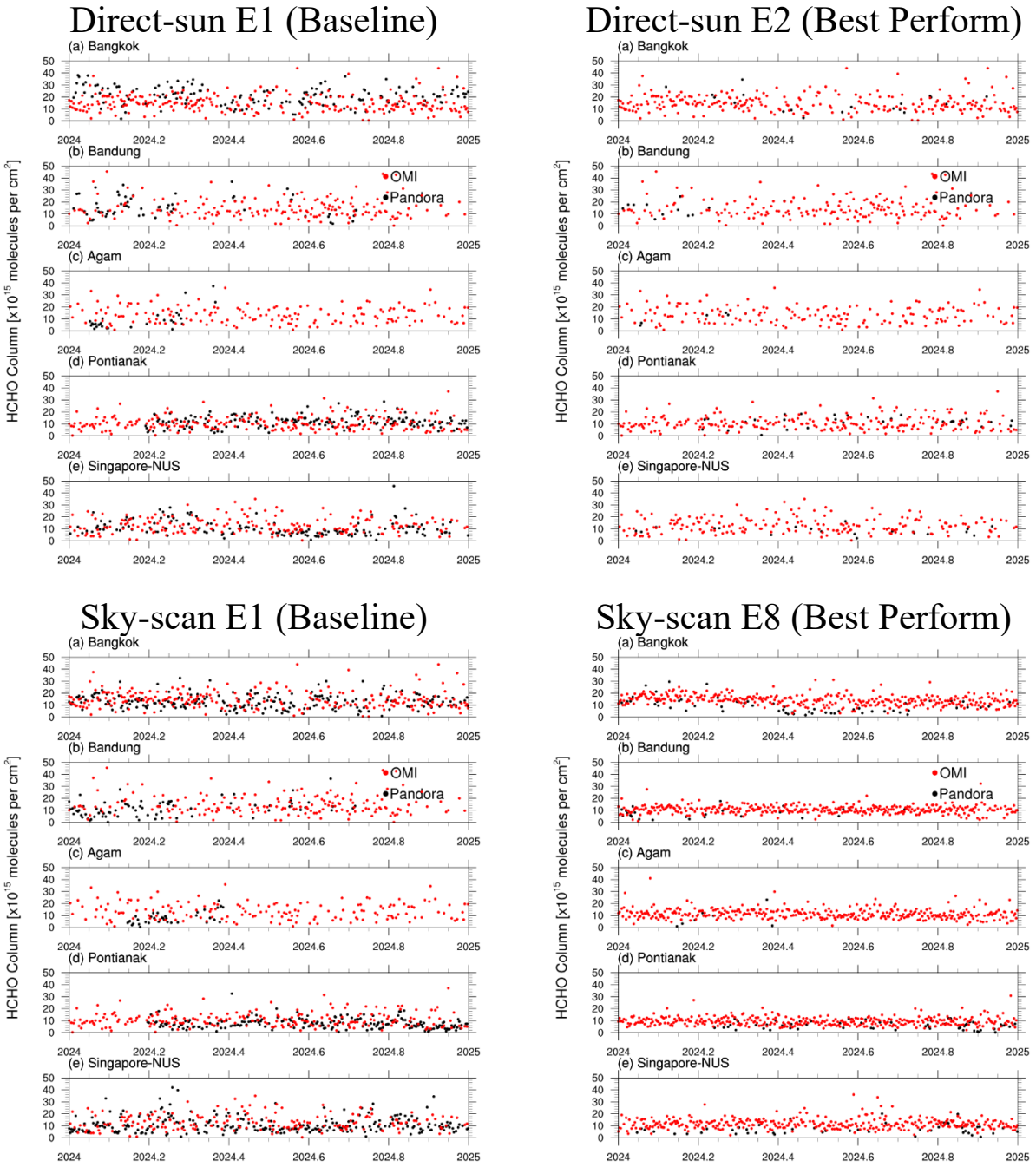


Figure 5: Time series of Direct-sun and Sky-scan E1 (baseline), Direct-sun E2 and Sky-scan E8 daily average HCHO column densities measured by Pandora monitoring stations (black marker) and OMI Aura (red marker) in Bangkok (a), Bandung (b), Agam (c), Pontianak (d), and Singapore-NUS (e) from 2024 January to 2024 December. Pandora HCHO is retrieved using data version rfus5p1-8; OMI HCHO is retrieved using data version 003 Level 2 Global Gridded 0.25-degree x 0.25-degree (OMHCHOG). Data greater than 50 x 10^{15} molecules per cm^2 are filtered.

https://doi.org/10.5194/egusphere-2026-716
Preprint. Discussion started: 23 February 2026

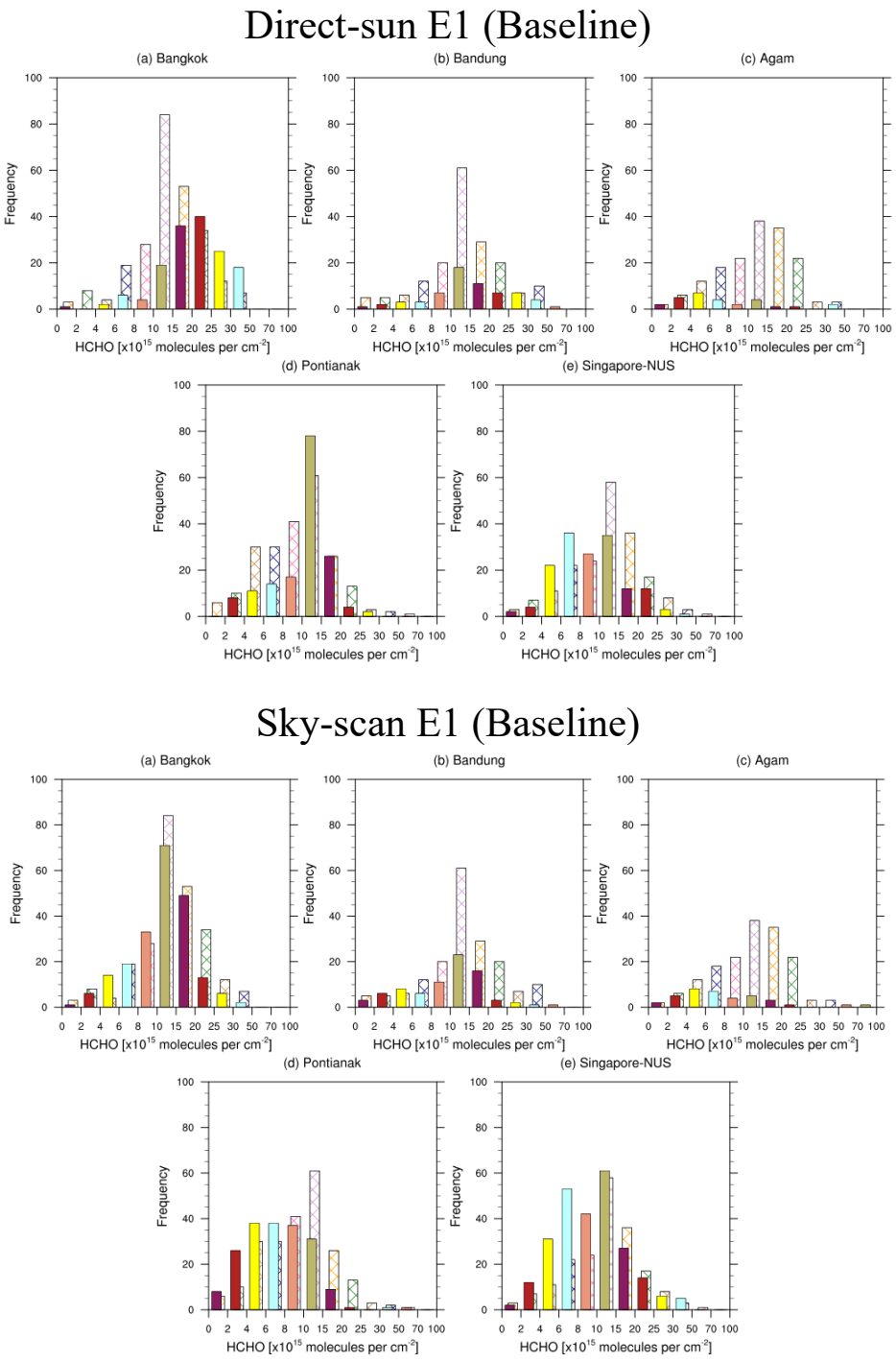
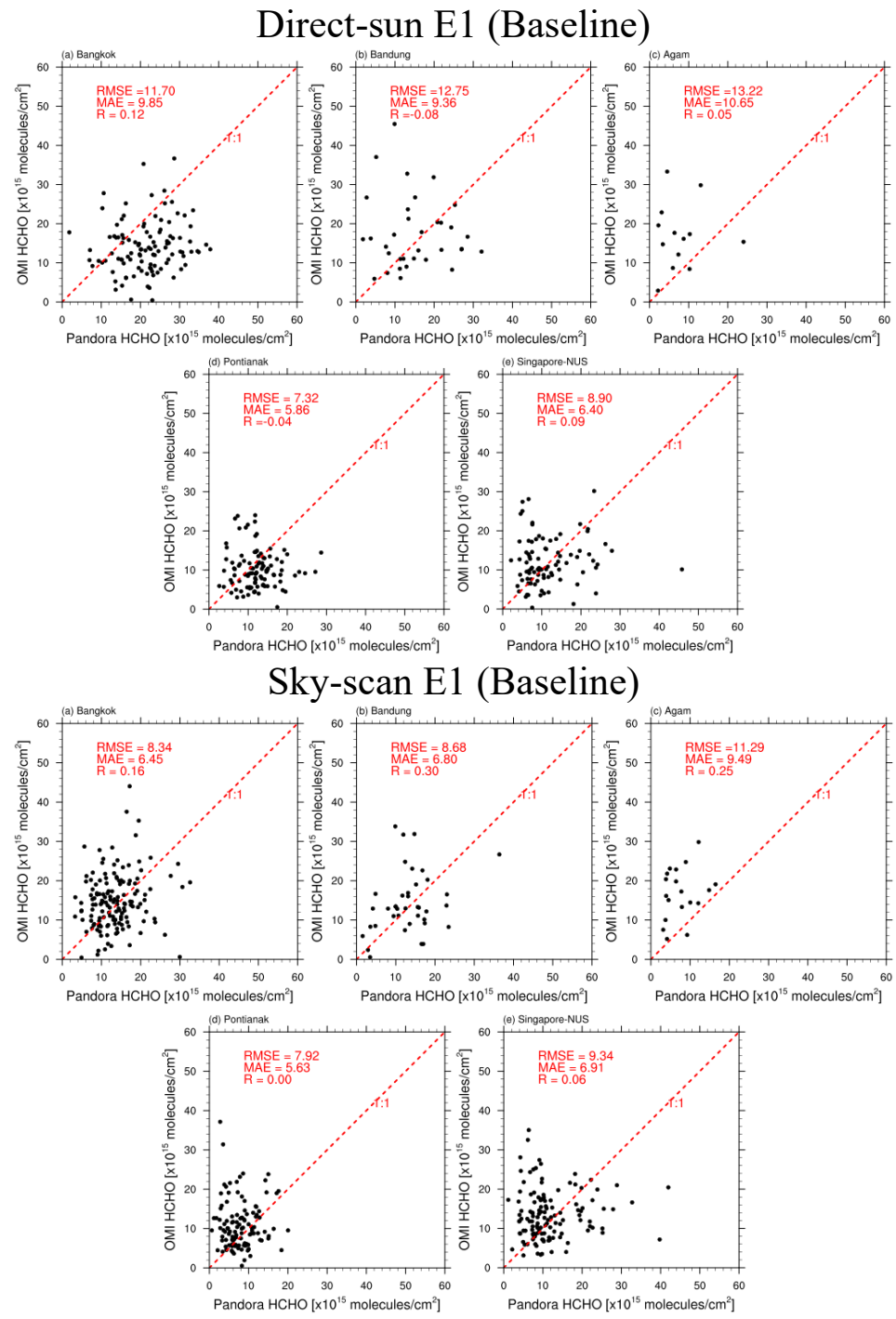


Figure 6: Frequency distribution of Direct-sun and Sky-scan E1 (baseline) daily average HCHO column densities measured by Pandora monitoring stations (solid-fill bars) and OMI Aura (pattern-filled bars) in Bangkok (a), Bandung (b), Agam (c), Pontianak (d), and Singapore-NUS (e) from 2024 January to 2024 December. Pandora HCHO is retrieved using data version rfus5p1-8; OMI HCHO is retrieved using data version 003 Level 2 Global Gridded 0.25-degree x 0.25-degree (OMHCHOG).

https://doi.org/10.5194/egusphere-2026-716
Preprint. Discussion started: 23 February 2026



335 **Figure 7:** Scatter plot of Direct-sun and Sky-scan E1 (baseline) daily average HCHO column densities measured by Pandora monitoring stations and OMI Aura in Bangkok (a), Bandung (b), Agam (c), Pontianak (d), and Singapore-NUS (e) from 2024 January to 2024 December. Pandora HCHO is retrieved using data version rfus5p1-8; OMI HCHO is retrieved using data version 003 Level 2 Global Gridded 0.25-degree x 0.25-degree (OMHCHOG). Red dotted line represents 1:1 regression.

https://doi.org/10.5194/egusphere-2026-716
Preprint. Discussion started: 23 February 2026

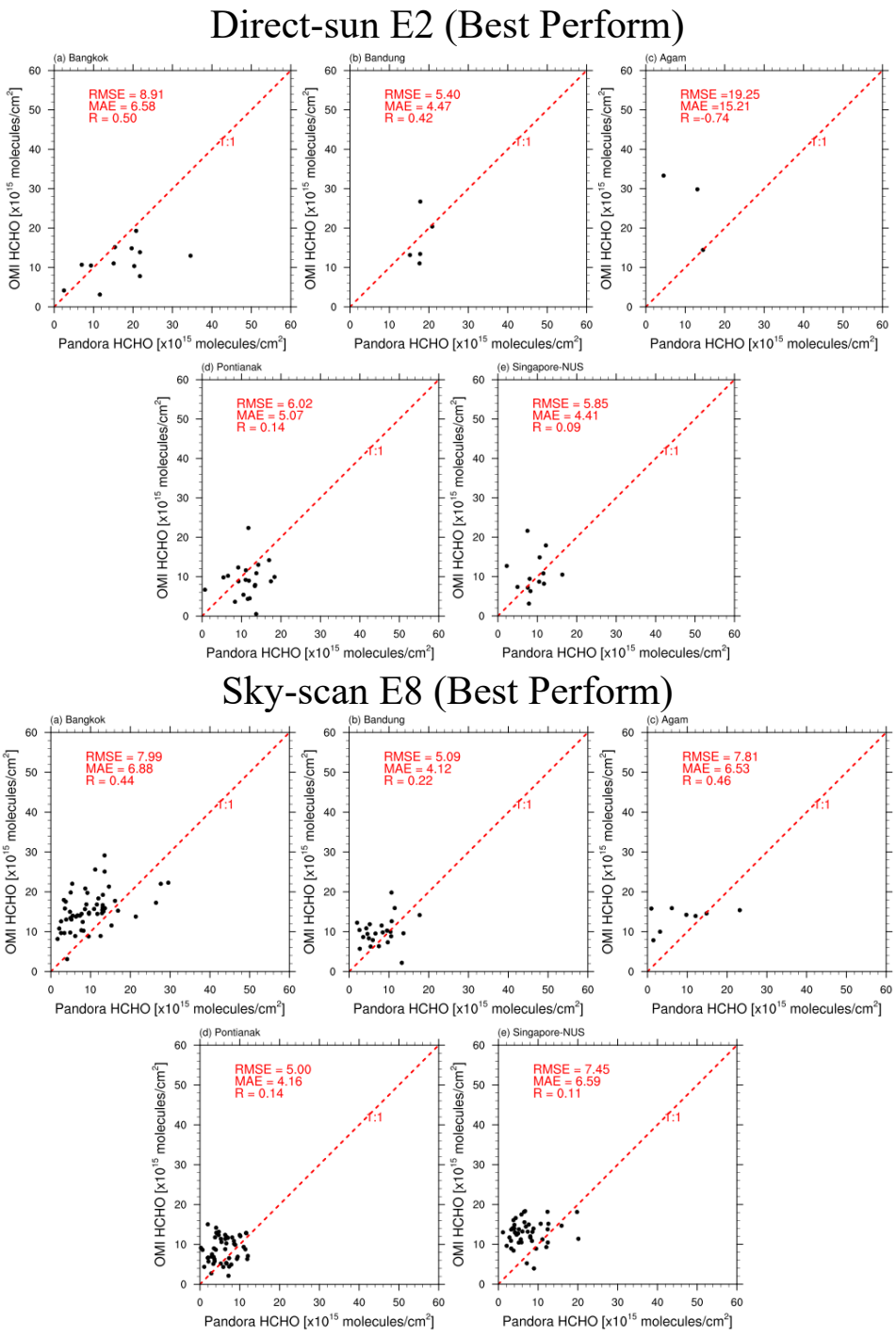


Figure 8: Scatter plot of Direct-sun E2 and Sky-scan E8 daily average HCHO column densities measured by Pandora monitoring stations and OMI Aura in Bangkok (a), Bandung (b), Agam (c), Pontianak (d), and Singapore-NUS (e) from 2024 January to 2024 December. Pandora HCHO is retrieved using data version rfus5p1-8; OMI HCHO is retrieved using data version 003 Level 2 Global Gridded 0.25-degree x 0.25-degree (OMHCHOG). Red dotted line represents 1:1 regression.

https://doi.org/10.5194/egusphere-2026-716
Preprint. Discussion started: 23 February 2026

4. Discussion

4.1 Effect of Scan Geometry on Pandora HCHO

The solar zenith angle (SZA)–dependent bias analysis (Fig. 9) demonstrates that scanning geometry exerts a first-order control on the consistency between Pandora and OMI HCHO column retrievals across Southeast Asia. This finding is consistent with previous studies showing pronounced SZA-dependent uncertainties in OMI trace-gas products, arising from enhanced multiple scattering, changes in air-mass factors, and residual geometric effects at large SZAs (Balis et al., 2007; Bak et al., 2024). Under the Direct-sun E1 baseline configuration, Pandora retrievals generally exhibit a positive bias relative to OMI across most sites and SZA bins, indicating a systematic overestimation that persists across varying illumination conditions. This behaviour is particularly evident at Bangkok, where the Direct-sun bias remains approximately constant ($\sim$10–20 $\times$ 10^{15} molecules cm^{-2}) across the full SZA range.

In contrast, Sky-scan Pandora retrievals exhibit substantially reduced SZA-dependent bias. By sampling diffuse skylight from multiple viewing directions, Sky-scan observations effectively integrate HCHO over a larger atmospheric volume and horizontal footprint. At Bangkok, Sky-scan biases remain close to zero for SZAs below ~40°, effectively removing the systematic overestimation observed in the Direct-sun mode. This behaviour reflects the reduced sensitivity of Sky-scan measurements to fine-scale spatial heterogeneity and localized enhancements that can disproportionately influence narrow line-of-sight observations. Given OMI's coarse spatial resolution (~13 km × 24 km at nadir), the more spatially integrated nature of Sky-scan Pandora observations provides a closer representational match to the satellite sampling, resulting in improved consistency and reduced systematic bias (Herman et al., 2018).

https://doi.org/10.5194/egusphere-2026-716
Preprint. Discussion started: 23 February 2026

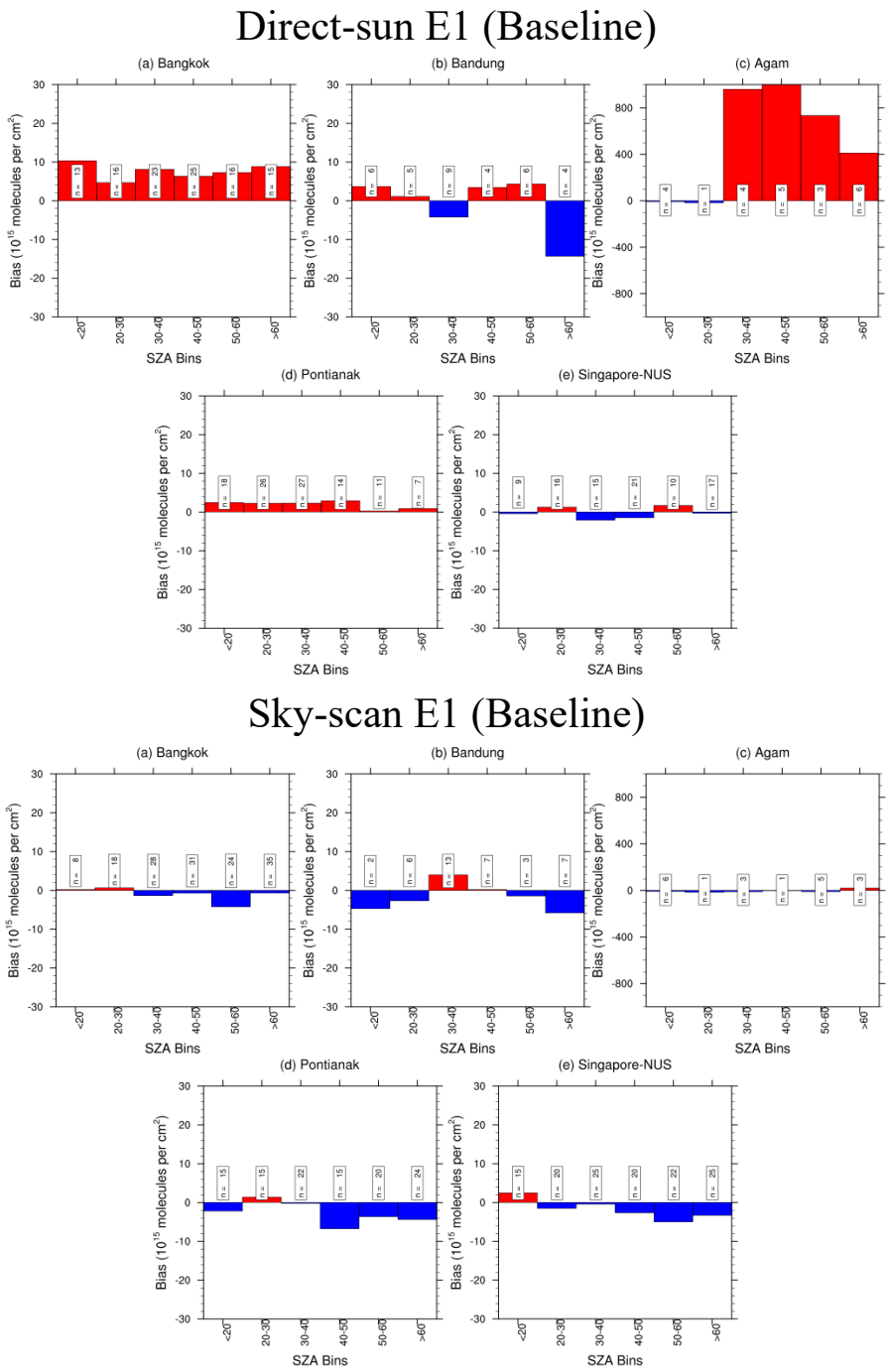


Figure 9. Bar chart showing the bias between daily average HCHO column densities from Direct-sun and Sky-scan Pandora retrievals (E1 Baseline) and OMI Aura across five locations: Bangkok (a), Bandung (b), Agam (c), Pontianak (d), and Singapore-NUS (e), covering the period from January to December 2024. The bias is calculated as Pandora HCHO minus OMI HCHO and is grouped by solar zenith angle (SZA) bins. Pandora HCHO retrievals use data version *rfus5p1-8*, while OMI HCHO data is from version 003 Level 2 Global Gridded 0.25° × 0.25° (OMHCHOG). *n* indicates the number of data points in each AOD bin. Note: the y-axis range in panel (c) for Agam is set to [800, –800] molecules/cm².

https://doi.org/10.5194/egusphere-2026-716
Preprint. Discussion started: 23 February 2026

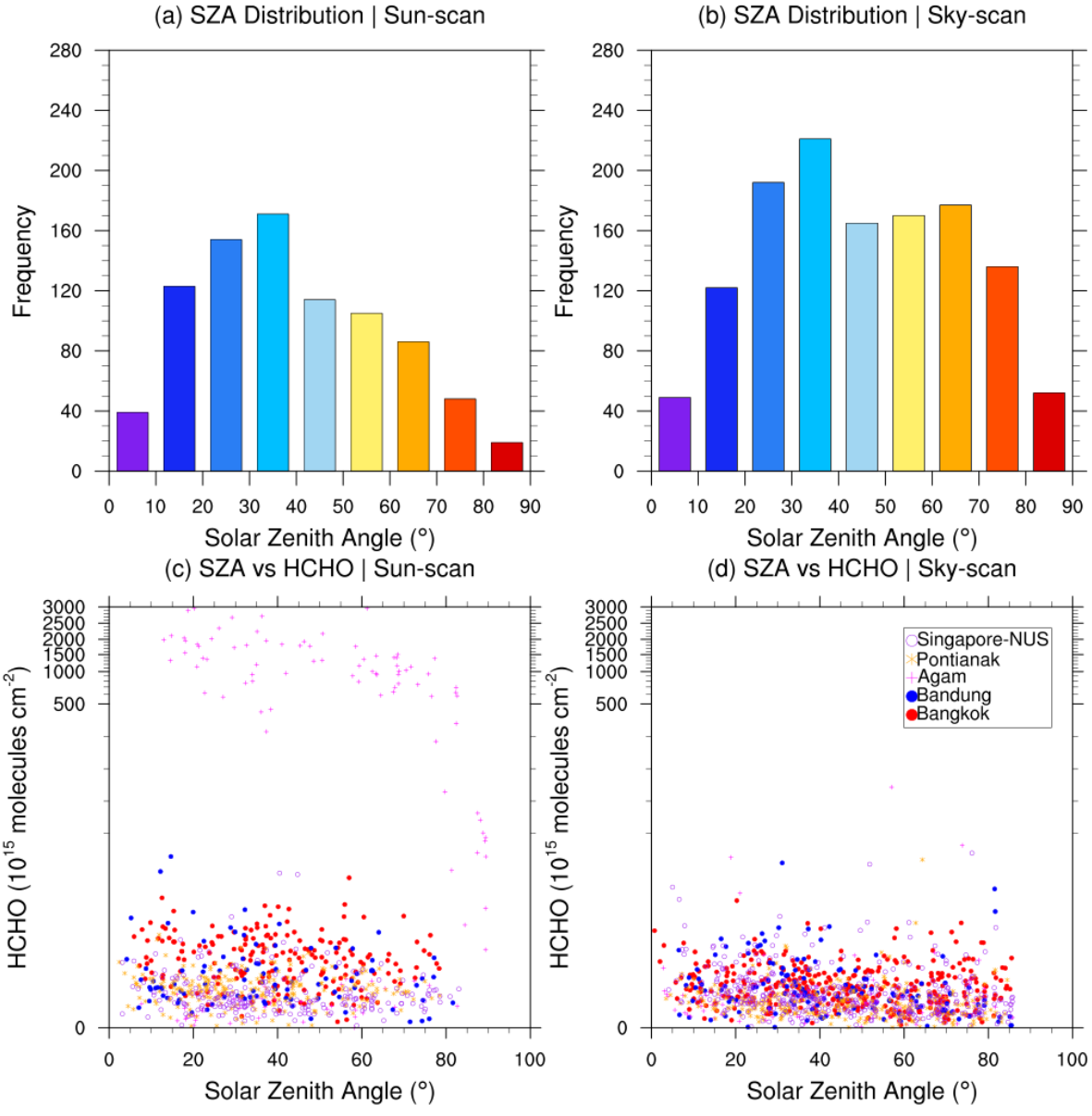


Figure 10. Distribution of solar zenith angle (SZA) and its influence on Pandora HCHO column retrievals for Direct-sun and Sky-scan geometries in 2024. Panels (a) and (b) show the frequency distributions of SZA for Direct-sun and Sky-scan observations, respectively. Panels (c) and (d) present scatter plots of HCHO column densities as a function of SZA for Direct-sun and Sky-scan measurements. Extreme HCHO outliers (>3000 × 10^{15} molecules cm^{-2}) have been excluded to reduce the influence of rare retrieval artefacts. Note that the y-axis uses explicitly defined, irregular tick intervals to visualize both low and high HCHO variability on a linear scale.

Figure 10 further illustrates systematic differences in SZA sampling and retrieval behaviour between Direct-sun and Sky-scan observations. Sky-scan measurements span a broader and more uniformly populated SZA range, particularly at moderate to high angles (~40–80°), whereas Direct-sun observations are concentrated at lower SZAs, reflecting geometric

https://doi.org/10.5194/egusphere-2026-716
Preprint. Discussion started: 23 February 2026

constraints associated with direct solar viewing (Fig. 10a–b). Station-resolved SZA–HCHO relationships (Fig. 10c–d) show that Direct-sun retrievals exhibit larger variability and stronger site dependence across the full SZA range, including episodic high-column events. These features are consistent with increased air-mass factors at large SZAs and enhanced sensitivity to localized atmospheric structures. The most extreme HCHO values are primarily observed at Agam, indicating the influence of site-specific factors such as viewing geometry, local heterogeneity, or retrieval sensitivity under certain illumination conditions. In contrast, Sky-scan retrievals display a more compact and overlapping distribution across stations, with fewer extreme values and weaker apparent SZA dependence, consistent with the spatial averaging inherent to scattered-light measurements. Overall, these results highlight the contrasting sensitivities imposed by viewing geometry, with Sky-scan retrievals yielding more geometrically stable HCHO columns, while Direct-sun measurements retain heightened sensitivity to localized variability.

4.2 Regional Bias Behavior and Extremes

Beyond Bangkok, Sky-scan retrievals demonstrate a stabilizing influence across sites characterized by diverse atmospheric regimes. Agam represents the most extreme case, where Direct-sun E1 biases exceed 400 × 10^{15} molecules cm^{-2} at SZAs greater than ~30°. These large anomalies are effectively suppressed under the Sky-scan E1 configuration (Fig. 9c), yielding near-zero bias across the full SZA range and indicating robust performance under challenging observational conditions. At Bandung and Singapore–NUS, the transition from positive Direct-sun bias to slightly negative Sky-scan bias at higher SZAs (>40°) suggests a mild underestimation at increased optical path lengths; however, the magnitude of this bias remains substantially smaller than the corresponding Direct-sun overestimation. Pontianak exhibits the most stable bias behaviour, with Sky-scan retrievals maintaining a consistently low-magnitude bias across all SZAs, reflecting reduced sensitivity to solar geometry in a comparatively homogeneous observational environment.

4.3 Interaction Between Bias Correction and Retrieval Optimization

The combined analysis of bias, correlation, and error metrics demonstrates that optimal satellite–ground agreement is inherently site dependent. While the Sky-scan E1 configuration is highly effective in minimizing SZA-dependent bias—particularly at Bangkok and Agam—additional optimization through spatiotemporal averaging further improves overall statistical performance. At Bandung and Singapore–NUS, the optimized Direct-sun E2 configuration yields the largest reductions in RMSE and the strongest improvements in correlation, despite Sky-scan providing superior bias stability. Conversely, at Agam and Pontianak, the Sky-scan E8 configuration achieves the most consistent alignment with OMI, delivering both reduced error and improved correlation. These results indicate that bias minimization alone does not guarantee optimal agreement and that the interplay between viewing geometry, local heterogeneity, and spatiotemporal representativeness governs the effectiveness of individual retrieval configurations.

https://doi.org/10.5194/egusphere-2026-716
Preprint. Discussion started: 23 February 2026

5. Conclusions

This study demonstrates that scanning geometry and retrieval optimization critically influence the reliability of Pandora HCHO column measurements for satellite validation in tropical environments. The SZA-dependent bias analysis confirms that the Sky-scan configuration provides a more robust and satellite-consistent baseline than Direct-sun, effectively minimizing path-length-related and local-scale biases. Bangkok emerges as the clearest example of Sky-scan superiority, where near-zero SZA-dependent bias and strong agreement with OMI highlight the effectiveness of Sky-scan geometry in complex urban environments with high aerosol and emission variability. Sky-scan retrievals also play a crucial role in stabilizing extreme anomalies, as demonstrated at Agam. However, the optimal configuration varies by site. Direct-sun E2 proves highly effective for reducing errors in Bandung and Singapore-NUS, while Sky-scan E8 delivers the best overall performance in Agam and Pontianak. These findings support a site-specific validation strategy, where the selection between optimized Direct-sun and Sky-scan configurations is guided by local SZA-dependent bias behavior and error characteristics. Overall, this work underscores the necessity of adaptive retrieval strategies when using ground-based spectrometers for satellite validation in the tropics and provides a framework for selecting optimal Pandora configurations to support robust HCHO satellite product evaluation. Sky-scan retrievals are recommended as the baseline for satellite validation in heterogeneous tropical environments.

Acknowledgements

The authors gratefully acknowledge the Pandonia Global Network (PGN) and the teams responsible for maintaining the Pandora spectrometer systems in Bangkok, Bandung, Agam, Pontianak, and Singapore. We also thank the NASA Goddard Earth Sciences Data and Information Services Center (GES DISC) for providing OMI formaldehyde (OMHCHOG v003) data. Their efforts in ensuring the quality, accessibility, and continuity of these datasets were essential to the success of this study.

Author contributions

Conception of the work: Jackson HW Chang. Acquisition of the data: Putu Aryastana. Interpretation of the data: Yong Jie Wong, Chee Fuei Pien. Supervision: Justin Sentian, Juneng Liew, Neng-Huei Lin. Writing—original draft: Jackson HW Chang, Maggie Chel-Gee Ooi. Writing—review and editing: Justin Sentian, Juneng Liew, Neng-Huei Lin.

Funding

The authors did not receive support from any organization for the submitted work.

Availability of data and materials

The datasets used and/or analyzed in the current study are available from the corresponding author on reasonable request.

https://doi.org/10.5194/egusphere-2026-716
Preprint. Discussion started: 23 February 2026

Declarations

The authors declare that they have no potential conflicts of interest. This research did not involve human participants and/or animals, and informed consent was not required for this study.

Competing interests

The authors declare that they have no competing interest.

References

Ahn, C., Torres, O., and Bhartia, P. K.: Comparison of Ozone Monitoring Instrument UV Aerosol Products with Aqua/Moderate Resolution Imaging Spectroradiometer and Multiangle Imaging Spectroradiometer observations in 2006, J. Geophys. Res., 113, https://doi.org/10.1029/2007jd008832, 2008.

Bak, J., Liu, X., Yang, K., Gonzalez Abad, G., O'Sullivan, E., Chance, K., and Kim, C.-H.: An improved OMI ozone profile research product version 2.0 with collection 4 L1b data and algorithm updates, Atmos. Meas. Tech., 17, 1891–1911, https://doi.org/10.5194/amt-17-1891-2024, 2024.

Balis, D., Kroon, M., Koukouli, M. E., Brinksma, E. J., Labow, G., Veefkind, J. P., and Mcpeters, R. D.: Validation of Ozone Monitoring Instrument total ozone column measurements using Brewer and Dobson spectrophotometer ground-based observations, J. Geophys. Res., 112, https://doi.org/10.1029/2007jd008796, 2007.

Boersma, K. F., Vinken, G. C. M., and Eskes, H. J.: Representativeness errors in comparing chemistry transport and chemistry climate models with satellite UV–Vis tropospheric column retrievals, Geosci. Model Dev., 9, 875–898, https://doi.org/10.5194/gmd-9-875-2016, 2016.

Cheong, K. H., Ngiam, N. J., Morgan, G. G., Pek, P. P., Tan, B. Y.-Q., Lai, J. W., Koh, J. M., Ong, M. E. H., and Ho, A. F. W.: Acute Health Impacts of the Southeast Asian Transboundary Haze Problem-A Review., IJERPH, 16, 3286, https://doi.org/10.3390/ijerph16183286, 2019.

De Smedt, I., Theys, N., Yu, H., Danckaert, T., Lerot, C., Compernolle, S., Van Roozendael, M., Richter, A., Hilboll, A., Peters, E., Pedergnana, M., Loyola, D., Beirle, S., Wagner, T., Eskes, H., Van Geffen, J., Boersma, K. F., and Veefkind, P.: Algorithm theoretical baseline for formaldehyde retrievals from S5P TROPOMI and from the QA4ECV project, Atmos. Meas. Tech., 11, 2395–2426, https://doi.org/10.5194/amt-11-2395-2018, 2018.

https://doi.org/10.5194/egusphere-2026-716
Preprint. Discussion started: 23 February 2026

Fang, Z., Deng, W., Zhang, Y., Ding, X., Tang, M., Liu, T., Hu, Q., Zhu, M., Wang, Z., Yang, W., Huang, Z., Song, W., Bi, X., Chen, J., Sun, Y., George, C., and Wang, X.: Open burning of rice, corn and wheat straws: primary emissions, photochemical aging, and secondary organic aerosol formation, Atmos. Chem. Phys., 17, 14821–14839, https://doi.org/10.5194/acp-17-14821-2017, 2017.

Fu, T., 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, J. Geophys. Res., 112, https://doi.org/10.1029/2006jd007853, 2007.

Ghahremanloo, M., Nowlan, C., Abad, G. G., Garraffo, C., Liu, X., Wang, H., Henderson, B., Baumann, E., Valin, L., Geddes, J. A., and Zhao, X.: Comprehensive Analysis of Bias in TEMPO NO2 Column Densities Through Pandora Observations., JGR Atmospheres, 130, https://doi.org/10.1029/2025jd044150, 2025.

Glowania, M., Rohrer, F., Dorn, H.-P., Hofzumahaus, A., Holland, F., Kiendler-Scharr, A., Wahner, A., and Fuchs, H.: Comparison of formaldehyde measurements by Hantzsch, CRDS and DOAS in the SAPHIR chamber, Atmos. Meas. Tech., 14, 4239–4253, https://doi.org/10.5194/amt-14-4239-2021, 2021.

Hansen, A. B., Witham, C. S., Chong, W. M., Kendall, E., Chew, B. N., Gan, C., Hort, M. C., and Lee, S.-Y.: Haze in Singapore – source attribution of biomass burning PM10 from Southeast Asia, Atmos. Chem. Phys., 19, 5363–5385, https://doi.org/10.5194/acp-19-5363-2019, 2019.

Harkey, M., Holloway, T., Kim, E. J., Baker, K. R., and Henderson, B.: Satellite Formaldehyde to Support Model Evaluation., JGR Atmospheres, 126, https://doi.org/10.1029/2020jd032881, 2021.

Herman, J., Evans, R., Cede, A., Abuhassan, N., Petropavlovskikh, I., and Mcconville, G.: Comparison of ozone retrievals from the Pandora spectrometer system and Dobson spectrophotometer in Boulder, Colorado, Atmos. Meas. Tech., 8, 3407–3418, https://doi.org/10.5194/amt-8-3407-2015, 2015.

Herman, J., Spinei, E., Fried, A., Kim, J., Kim, J., Kim, W., Cede, A., Abuhassan, N., and Segal-Rozenhaimer, M.: 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, Atmos. Meas. Tech., 11, 4583–4603, https://doi.org/10.5194/amt-11-4583-2018, 2018.

https://doi.org/10.5194/egusphere-2026-716
Preprint. Discussion started: 23 February 2026

Jones, N. B., Riedel, K., Allan, W., Wood, S., Palmer, P. I., Chance, K., and Notholt, J.: Long-term tropospheric formaldehyde concentrations deduced from ground-based fourier transform solar infrared measurements, Atmos. Chem. Phys., 9, 7131–7142, https://doi.org/10.5194/acp-9-7131-2009, 2009.

Lamsal, L. N., Krotkov, N. A., Celarier, E. A., Swartz, W. H., Pickering, K. E., Bucsela, E. J., Gleason, J. F., Martin, R. V., Philip, S., Irie, H., Cede, A., Herman, J., Weinheimer, A., Szykman, J. J., and Knepp, T. N.: Evaluation of OMI operational standard NO 2 column retrievals using in situ and surface-based NO 2 observations, Atmos. Chem. Phys., 14, 11587–11609, https://doi.org/10.5194/acp-14-11587-2014, 2014.

Lee, G. T., Park, R. J., Kwon, H.-A., Ha, E. S., Lee, S. D., Shin, S., Ahn, M.-H., Kang, M., Choi, Y.-S., Kim, G., Lee, D.-W., Kim, D.-R., Hong, H., Langerock, B., Vigouroux, C., Lerot, C., Hendrick, F., Pinardi, G., De Smedt, I., Van Roozendael, M., Wang, P., Chong, H., Cho, Y., and Kim, J.: First evaluation of the GEMS formaldehyde product against TROPOMI and ground-based column measurements during the in-orbit test period, Atmos. Chem. Phys., 24, 4733–4749, https://doi.org/10.5194/acp-24-4733-2024, 2024.

Liao, K., Chen, Q., Liu, Y., Li, Y. J., Lambe, A. T., Zhu, T., Huang, R.-J., Zheng, Y., Cheng, X., Miao, R., Huang, G., Khuzestani, R. B., and Jia, T.: Secondary Organic Aerosol Formation of Fleet Vehicle Emissions in China: Potential Seasonality of Spatial Distributions., Environ. Sci. Technol., 55, 7276–7286, https://doi.org/10.1021/acs.est.0c08591, 2021.

Lim, C. Y., Hagan, D. H., Coggon, M. M., Koss, A. R., Sekimoto, K., De Gouw, J., Warneke, C., Cappa, C. D., and Kroll, J. H.: Secondary organic aerosol formation from the laboratory oxidation of biomass burning emissions, Atmos. Chem. Phys., 19, 12797–12809, https://doi.org/10.5194/acp-19-12797-2019, 2019.

Liu, J., Li, X., Yang, Y., Wang, H., Kuang, C., Zhu, Y., Chen, M., Hu, J., Zeng, L., and Zhang, Y.: Sensitive Detection of Ambient Formaldehyde by Incoherent Broadband Cavity Enhanced Absorption Spectroscopy., Anal. Chem., 92, 2697–2705, https://doi.org/10.1021/acs.analchem.9b04821, 2020.

Pinardi, G., Van Roozendael, M., Abuhassan, N., Adams, C., Cede, A., Clémer, K., Fayt, C., Frieß, U., Gil, M., Herman, J., Hermans, C., Hendrick, F., Irie, H., Merlaud, A., Navarro Comas, M., Peters, E., Piters, A. J. M., Puentedura, O., Richter, A., Schönhardt, A., Shaiganfar, R., Spinei, E., Strong, K., Takashima, H., Vrekoussis, M., Wagner, T., Wittrock, F., and Yilmaz, S.: MAX-DOAS formaldehyde slant column measurements during CINDI: intercomparison and analysis improvement, Atmos. Meas. Tech., 6, 167–185, https://doi.org/10.5194/amt-6-167-2013, 2013.

https://doi.org/10.5194/egusphere-2026-716
Preprint. Discussion started: 23 February 2026

Spinei, E., Whitehill, A., Fried, A., Tiefengraber, M., Knepp, T. N., Herndon, S., Herman, J. R., Müller, M., Abuhassan, N., Cede, A., Richter, D., Walega, J., Crawford, J., Szykman, J., Valin, L., Williams, D. J., Long, R., Swap, R. J., Lee, Y., Nowak, N., and Poche, B.: The first evaluation of formaldehyde column observations by improved Pandora spectrometers during the KORUS-AQ field study., Atmos. Meas. Tech., 11, 4943–4961, https://doi.org/10.5194/amt-11-4943-2018, 2018.

Su, W., Liu, C., Chan, K. L., Hu, Q., Liu, H., Ji, X., Zhu, Y., Liu, T., Zhang, C., Chen, Y., and Liu, J.: An improved TROPOMI tropospheric HCHO retrieval over China, Atmos. Meas. Tech., 13, 6271–6292, https://doi.org/10.5194/amt-13-6271-2020, 2020.

Tanskanen, A., Krotkov, N. A., Herman, J. R., and Arola, A.: Surface ultraviolet irradiance from OMI, IEEE Trans. Geosci. Remote Sensing, 44, 1267–1271, https://doi.org/10.1109/tgrs.2005.862203, 2006.

Tzortziou, M., Herman, J. R., Cede, A., and Abuhassan, N.: High precision, absolute total column ozone measurements from the Pandora spectrometer system: Comparisons with data from a Brewer double monochromator and Aura OMI, J. Geophys. Res., 117, n/a, https://doi.org/10.1029/2012jd017814, 2012.

Vigouroux, C., Bauer Aquino, C. A., Bauwens, M., Becker, C., Blumenstock, T., De Mazière, M., García, O., Grutter, M., Guarin, C., Hannigan, J., Hase, F., Jones, N., Kivi, R., Koshelev, D., Langerock, B., Lutsch, E., Makarova, M., Metzger, J.-M., Müller, J.-F., Notholt, J., Ortega, I., Palm, M., Paton-Walsh, C., Poberovskii, A., Rettinger, M., Robinson, J., Smale, D., Stavrakou, T., Stremme, W., Strong, K., Sussmann, R., Té, Y., and Toon, G.: NDACC harmonized formaldehyde time series from 21 FTIR stations covering a wide range of column abundances, Atmos. Meas. Tech., 11, 5049–5073, https://doi.org/10.5194/amt-11-5049-2018, 2018.

Zhu, L., Jacob, D. J., Keutsch, F. N., Mickley, L. J., Scheffe, R., Strum, M., González Abad, G., Chance, K., Yang, K., Rappenglück, B., Millet, D. B., Baasandorj, M., Jaeglé, L., and Shah, V.: Formaldehyde (HCHO) As a Hazardous Air Pollutant: Mapping Surface Air Concentrations from Satellite and Inferring Cancer Risks in the United States., Environ. Sci. Technol., 51, 5650–5657, https://doi.org/10.1021/acs.est.7b01356, 2017