



A Bayesian Maximum Entropy Framework Using Vertical Profiles to Improve Surface Ozone Estimation From IASI+GOME2, OMI/MLS, And Cris Satellite Ozone Observations

5 Hantao Wang¹, Marc L. Serre¹, Kazuyuki Miyazaki², Juan Cuesta³, Jerry R. Ziemke⁴, J. Jason West¹

¹Department of Environmental Sciences and Engineering, University of North Carolina at Chapel Hill, Chapel Hill, 27599, USA

²Jet Propulsion Laboratory, California Institute of Technology, Pasadena, 91125, USA

10 ³Univ. Paris Est Créteil and Université de Paris Cité, CNRS, LISA, 94010 Créteil, France

⁴NASA Goddard Space Flight Center, Greenbelt, 20771, USA

Correspondence to: J. Jason West (jasonwest@unc.edu)

Abstract

Satellite observations are essential for global tropospheric ozone monitoring, but their ability to estimate ground-level
15 ozone remains limited because of weak sensitivity and substantial uncertainty in retrieved ozone near the surface. In
this study, we develop new methods for adjusting satellite ozone observations (IASI+GOME2, OMI/MLS, and CrIS)
through chemistry-transport reanalysis and in situ ozone vertical profile measurements. Using these methods, we
create global maps of ground-level ozone concentrations based on satellite observations. We use the Bayesian
Maximum Entropy framework to horizontally interpolate the vertical profiles from ozonesondes and IAGOS and
20 improve the accuracy of both the satellite column measurements and the surface-to-column ratios from a chemical
reanalysis. This is done for monthly average maximum daily 8-hr ozone concentrations over several years. For the
three satellites, surface ozone estimated from the BME-adjusted column-to-surface conversion showed improved
agreement with TOAR-II observations. For IASI+GOME2 (2017-2020), global R^2 increased from 0.25 to 0.51, and
RMSE was reduced from 10.74 ppb to 9.44 ppb. For OMI/MLS tropospheric column (2005-2022), global R^2 increased
25 from 0.26 to 0.57, and RMSE decreased from 22.21 ppb to 7.79 ppb. For the CrIS 0-3km ozone (2022), global R^2
increased from 0.30 to 0.56, and RMSE decreased from 16.48 ppb to 9.45 ppb. The method's efficacy was found to
be highest within 6° of a vertical profile station and most impactful when the original satellite data quality was low.
The resulting satellite-based monthly ground-level ozone estimates can be used further as an independent input to data
fusion methods.

30 1. Introduction

Ozone at ground level is a secondary air pollutant with harmful impacts on human health (Balme, 2022) and plants
(Mills et al., 2018). Ozone is also a major greenhouse gas, ranking third in its contribution to anthropogenic climate
change behind carbon dioxide and methane (Masson-Delmotte et al., 2021). According to the Global Burden of
Disease 2021 (GBD) study, surface ozone contributed to the deaths of approximately 490,000 (95% UI: 107,000–



35 837,000) people globally in 2021, accounting for 0.72% (95% UI: 0.16% – 1.18%) of all deaths (Brauer et al., 2024). Recognizing its importance, the Tropospheric Ozone Assessment Report (TOAR) provides the most extensive compilation of ground-level ozone data by aggregating thousands of monitoring stations worldwide, and its second phase (TOAR-II) is further expanding coverage in regions like China and India (Schröder et al., 2021; Schultz et al., 2017). The TOAR dataset enables analysis of ozone distribution and trends (Gaudel et al., 2018). For example, 40 research using TOAR-II data reveals that in the United States, although extreme ozone concentrations have declined, winter concentrations have actually increased in the Southwest and Midwest regions since the 1990s (Chang et al., 2025). In contrast, East Asia and Southeast Asia have emerged as current hotspots for ozone pollution, with ozone concentration significantly exceeding those currently observed in the United States and Europe (Lu et al., 2025). In densely populated regions of China, such as the North China Plain and the Yangtze River Delta, the upward trend in 45 ozone concentrations has intensified markedly, with summer seasonal averages in these areas now exceed 75 ppb (Li et al., 2025a). In South America, based on TOAR-II data, trends in ground-level ozone concentrations vary by geographic location, with high-altitude background monitoring stations exhibiting an upward trend, whereas the southernmost regions demonstrate a downward trend (Seguel et al., 2024).

50 Despite significant progress, the limited availability of ground-based monitoring, particularly in developing countries and remote areas, makes it challenging to accurately track surface ozone and fully understand its variations over time and space. To bridge gaps in regions lacking ground-based ozone monitors, researchers have employed various data fusion methods, including geostatistical approaches, machine learning, and chemical reanalysis. Geostatistical datasets fuse ozone observations with composite outputs from multiple chemistry models, leveraging spatial and temporal 55 autocorrelations (Becker et al., 2023). Machine learning datasets use neural networks to fuse various data sources, including ground-based observations, chemical model outputs, and meteorology (Liu et al., 2022; Sun et al., 2022). Chemical reanalysis products utilize atmospheric composition from satellite data and have not included ground-level ozone observations in their data assimilation process (Miyazaki, 2024; Inness et al., 2019). A recent comparison of six global ground-level ozone datasets reveals significant discrepancies in ozone magnitude, spatial patterns, and 60 trends, with average estimates differing by as much as 5–10 parts per billion (ppb), while most datasets tend to overestimate ozone at lower concentration (Wang et al., 2025).

Satellite observations have the potential to improve our understanding of the global distribution of ground-level ozone. Satellite remote sensing offers a distinct advantage in this regard, providing daily, global coverage of tropospheric 65 ozone overcoming the sparse distribution of other monitoring networks (Yang et al., 2025). Satellite retrievals have been used as a key input for chemical reanalysis products, offering a way to understand ozone column densities, as well as some machine learning datasets (Hickman et al., 2025; Sekiya et al., 2025). These continuous global records are important for quantifying long-term trends and understanding the transport of tropospheric ozone (Van Malderen et al., 2025). However, satellite retrievals can be compromised by biases between different products, instrument drift, 70 and cloud interference (Pennington et al., 2025). More importantly, these satellite retrievals lack the sensitivity necessary to accurately measure ozone concentrations near the surface, which is the layer most relevant to human



health (Ziemke et al., 2019). To overcome the limited sensitivity of satellite retrievals to near-surface ozone, multispectral retrieval techniques that synergistically combine measurements from different spectral ranges have been developed over past years (Mettig et al., 2022). The OMI/MLS product provides tropospheric column ozone using the
75 Tropospheric Ozone Residual (TOR) technique, which subtracts the stratospheric column ozone measured by the Microwave Limb Sounder (MLS) from the total column ozone measured by the Ozone Monitoring Instrument (OMI) (Ziemke et al., 2006). IASI+GOME2, a satellite product that couples Thermal Infrared (TIR) radiances from the Infrared Atmospheric Sounding Interferometer (IASI) and Ultraviolet (UV) reflectance from the Global Ozone Monitoring Experiment 2 (GOME2), offers insights into the lowermost troposphere (LMT, 0–3 km) and effectively
80 lowers the altitude of peak sensitivity to around 2.2 km over land (Cuesta et al., 2022; Cuesta et al., 2013). Similarly, the AIRS/OMI product integrates TIR radiances from the Atmospheric Infrared Sounder (AIRS) and UV reflectance from the OMI to generate vertical profiles capable of distinguishing between upper and lower tropospheric ozone (Fu et al., 2018). Recently, satellite data from single high-resolution TIR instruments like the Cross-track Infrared Sounder (CrIS) have been processed using the advanced algorithm MUSES (MULTi-SpEctra, MULTi-SpEcies, MULTi-SEnsors)
85 to provide a reliable ozone retrieval (Pennington et al., 2025; Fu et al., 2018).

To adjust and validate satellite ozone retrievals, researchers rely on high-quality, in-situ vertical profile observations. Ozonesondes are considered a reference standard due to their high accuracy and fine vertical resolution, making them essential for correcting satellite data (Gaudel et al., 2024). Similarly, the In-service Aircraft for a Global Observing
90 System (IAGOS) program utilizes instruments on commercial aircraft to provide high-quality ozone profiles, offering a denser sampling frequency than ozonesondes and valuable data over regions like Asia and Africa that often lack long-term ozonesonde sites (Wang et al., 2022). While vertical profile observations are valuable for assessing the performance of satellite products, their utility is constrained by sparse global coverage and low sampling frequency (typically weekly). These limitations introduce significant uncertainty into long-term records, complicating the
95 detection of ozone trends (Chang et al., 2020; Chang et al., 2024). Consequently, these vertical profiles cannot be directly applied to correct satellite retrievals of ground-level ozone on a global scale.

Previous studies have integrated ground-based measurements with satellite data to constrain estimates of ground-level ozone. Xue et al. (2020) utilized random forest models to estimate daily surface ozone across China by combining in
100 situ observations with chemical transport model simulations (CMAQ) and satellite-derived ozone profiles from the Ozone Monitoring Instrument (OMI), achieving R^2 of around 0.70 against ground observations for 2013–2017. However, machine learning methods usually do not utilize the vertical profile structure to physically derive surface ozone from the satellite column; instead, they incorporate satellite column ozone within a data-driven model to estimate ground-level observations (Li and Cheng, 2021). Also, the reliance of machine learning on ground-based
105 monitors limits its ability to estimate surface ozone before 2013 in China, when the monitoring network was established (Wei et al., 2022). Zhang et al. (2021) developed a deep learning Long Short-Term Memory (LSTM) model to estimate ozone concentrations from the lower stratosphere to near surface by training it on a combination of ozonesonde profiles, meteorological data (ERA5), and satellite radiances (AIRS/OMI). While the LSTM model



utilized ozonesonde vertical profiles to derive surface ozone from satellite observations, its ability to capture vertical
110 ozone distributions was limited to the locations of the ozonesonde sites, and their validation of ground-level ozone in
China was aggregated into three broad regions (Sichuan Basin, North China Plain, and Yangtze River Delta), without
reporting detailed performance metrics for more locations. Colombi et al. (2021) developed a method to infer surface
ozone from multispectral satellite data (TES/OMI) by integrating surface-to-column ratios from a chemical reanalysis,
and correcting those ratios based on ozonesondes and ground-based observations. While this approach demonstrated
115 that utilizing vertical information from a chemical reanalysis allows for the effective inference of ground-level ozone
from satellite columns, its application is limited to the United States and China during summer (2005–2009) with a 2°
 $\times 2^\circ$ resolution (Colombi et al., 2021). Furthermore, the method relied on spatially uniform bias corrections over broad
regions, which did not account for how satellite ozone and vertical ozone profile estimates are biased in nonlinear
and spatially heterogeneous ways. While recent methods have increasingly attempted to leverage vertical information,
120 they remain constrained by their reliance on ground-based observations for training or bias correction, which restricts
their application to specific regions and time periods.

In this study, we attempt to improve global surface ozone estimates from satellites by developing a geostatistical
framework for adjusting the column to surface conversion that effectively propagates vertical profile information
125 across the horizontal domain. Unlike the spatially uniform corrections employed by Colombi et al. (2021), our method
utilizes the Bayesian Maximum Entropy (BME) method to correct nonlinear and regionally varying biases in both the
satellite ozone column concentrations and the chemical reanalysis surface-to-column ratios. By applying a BME
adjustment, our study propagates these corrections smoothly in space based on the spatial autocorrelation of monthly
fitted ozonesonde and IAGOS observations. We apply this BME-adjusted column to surface conversion framework
130 to three different satellite datasets to estimate monthly global ground-level ozone, over the years for which these
satellite data are available from 2005 to 2022. The three satellite datasets are IASI+GOME2 0-3 km ozone (Cuesta et
al., 2013; Cuesta et al., 2022), OMI/MLS tropospheric column ozone (Ziemke et al., 2006), and CrIS 0-3 km or
tropospheric column ozone (Bowman, 2022). The BME-adjusted surface ozone estimates allow us to infer global
surface monthly mean Daily Maximum 8-Hour Average (DMA8) ozone at a $1^\circ \times 1^\circ$ resolution over multiple years.
135 We selected the monthly mean DMA8 for this study as it is widely used in health-related research, effectively
capturing both the ozone seasonal cycle and the burden of ground-level exposure (Fleming et al., 2018). Our estimates
are evaluated against independent ground observations from TOAR-II and benchmarked against the original satellite
measurements and the Tropospheric Chemistry Reanalysis Version 2 (TCR-2) surface ozone. By intentionally
excluding ground-based observations from the adjustment process, we ensure that our satellite-based product serves
140 as an independent input for future data fusion studies.

2. Data

As shown in Table 1, this study uses three different satellite ozone products, the TCR-2 chemical reanalysis, and
vertical profiles from ozonesonde and IAGOS observations. While we focus on the monthly average DMA8 because
it is a critical health metric, methodologically, using a monthly metric also helps reduce systematic biases stemming



145 from the low sampling frequency of ozonesonde and IAGOS data (Chang et al., 2024). However, the monthly metric
may hide extreme short-term events that would be found in daily or hourly data (Li et al., 2025b).

2.1 Satellite data

IASI+GOME2 uses multispectral retrieval techniques derived from the combined radiances in the TIR detected by
IASI and UV measured by GOME2 (synergism of level 1 radiance measurements). Both instruments are aboard on
150 the MetOp satellites, providing daily global coverage at a ground resolution of approximately 12 km diameter pixels
at the satellite overpass time around 9:30 AM local time (Cuesta et al., 2013). The IASI+GOME2 multispectral
approach is specifically designed to enhance sensitivity to ozone from the surface up to 3 km altitude, where single
retrievals from either TIR or UV instruments have limited sensitivity. For this, it combines information from the two
spectral domains and optimizes it through altitude dependent Phillips-Tikhonov type constraints. Degrees of Freedom
155 (DOF) represent the number of independent vertical structures that can be resolved by the satellite measurement,
which is calculated as the trace of the averaging kernel matrix (Colombi et al., 2021). While IASI and GOME2 each
typically provide 3.3 DOF for the entire atmosphere, IASI+GOME2 improves to 5 DOF for the entire atmosphere
(accounting for approximately 1.6 DOF for tropospheric ozone) in average (Cuesta et al., 2013). IASI+GOME2 has
been validated against ozonesondes at multiple locations (46 stations distributed worldwide) with a total of 1035
160 records, demonstrating good agreement in the lowermost troposphere with a high correlation ($R = 0.85$) (Cuesta et al.,
2018). In this study, we use monthly average 0-3 km ozone obtained from daily IASI+GOME2 observations at
approximately 9:30 AM local time overpass from 2017 to 2020 at $1^\circ \times 1^\circ$ resolution.

The OMI/MLS tropospheric column ozone product is derived using the TOR technique, which subtracts the
165 stratospheric column ozone measured by MLS from the total column ozone measured by OMI. A key advantage of
this product is that both instruments are on the NASA Aura satellite (1:45 PM overpass), resulting in a more reliable
estimate of tropospheric ozone (Ziemke et al., 2006). Characterized by a standard deviation (random uncertainty) of
approximately 1.3 Dobson Units (DU) for monthly observations, the OMI/MLS dataset has proven robust for
assessing seasonal and long-term atmospheric changes (Elshorbany et al., 2024; Fadnavis et al., 2025). In this study,
170 we use OMI/MLS monthly tropospheric column ozone at $1^\circ \times 1^\circ$ resolution from 2005 to 2022 (Ziemke et al., 2006),
with the tropopause defined by the thermal tropopause from National Centers for Environmental Prediction (NCEP)
reanalysis data.

Unlike multispectral products such as IASI+GOME2 and OMI/MLS, CrIS is a lower noise TIR instrument aboard the
175 Suomi-NPP satellite launched in 2011 (1:30 PM overpass) whose data are processed using the MUSES algorithm with
a spatial resolution of 14 km (Fu et al., 2018; Mettig et al., 2022). While CrIS provides high spectral resolution, its
vertical sensitivity is physically constrained by limited thermal contrast, resulting in lower sensitivity in the lower
troposphere compared to the upper troposphere. Validation indicates that the Tropospheric Ozone and its Precursors
from Earth System Sounding (TROPESS) CrIS product generally exhibits a positive bias in the tropospheric column
180 (4.8%) relative to ozonesondes (Pennington et al., 2025). This study uses the CrIS L2 Ozone forward stream product
for 2022, which provides 26 vertical layers (from the surface to 0.1 hPa) and approximately 23,569 targets per day



(Bowman, 2022). We use monthly 0-3 km and tropospheric column ozone products as predictors of ground-level ozone. For any satellite retrieval product, when the sensitivity to the lower troposphere is limited, the retrieved estimates are strongly influenced by the priori profiles.

185 2.2 Ozone vertical profile observations

Ozonesondes provide high quality vertical profile observations of ozone from the surface into the stratosphere. The instruments are launched on balloons and measure ozone partial pressure with a high vertical resolution of approximately 100 m. However, these datasets are limited by a low sampling frequency (often weekly or less). Furthermore, their long-term data records often require complex homogenization to correct for instrumental changes over years and may exhibit biases at high altitudes (Ancellet et al., 2022). To address the latter issue, this study uses the homogenized ozonesonde time series from the TOAR-II Harmonization and Evaluation of Ground-based Instruments for Free-Tropospheric Ozone Measurements (HEGIFTOM) working group, which corrects for instrumental and processing biases to reduce uncertainty for sites across global networks like Southern Hemisphere Additional Ozonesondes (SHADOZ) and the World Ozone and Ultraviolet Radiation Data Centre (WOUDC) (Smit and Thompson, 2021; Van Malderen et al., 2025). For this study, we utilize 36 ozonesonde locations from HEGIFTOM and 8 from WOUDC, selecting only those with a minimum of six qualifying months (defined as months having at least four measurements) over 2005 to 2022 (list of sites can be found in Table S1).

IAGOS provides vertical ozone profiles from instruments on commercial passenger aircraft as they take off and land, which has greater sampling frequency than ozonesondes, but its observations are unevenly distributed, with dense coverage over Europe and North America but sparse sampling in the Southern Hemisphere and other remote regions (Gaudel et al., 2020). In addition, since the IAGOS program relies on commercial flight routes, near-surface data are largely restricted to airport vicinities where local airport and urban emissions may limit their spatial representativeness and relevance for general population exposure assessments (Petetin et al., 2016; Gaudel et al., 2018). Therefore, we only use the IAGOS data above 200 m to limit the influence of local airport emissions. For this study, we include 48 IAGOS locations where vertical profiles were measured during aircraft ascent and descent near major airports. These airports were selected using the same standards as the ozonesondes (list of airports can be found in Table S1).

Figure 1(a) shows the network for vertical ozone profiles from both ozonesondes and IAGOS from 2005 to 2022. We incorporated IAGOS data to supplement the ozonesonde network, particularly in North America, Asia, and Africa. The IAGOS network expanded significantly between 2007 and 2017, while the ozonesonde network remained relatively stable. However, the combined network is heavily concentrated in the Northern Hemisphere, with notable coverage gaps in South America, Australia, and Africa. Temporal changes in data availability (Fig. 1(b)) show that the total number of monthly ozone profiles increased steadily from 2005 to 2018, peaking around 2016. The Day Count Category is defined as the total number of days per month containing valid data records. We show thresholds of 4 and 7 days; a 4-day threshold represents an average sampling frequency of once per week, while a 7-day threshold indicates that data coverage exceeds 25% of days in the month. From Fig. 1(b), we found data robustness improved in terms of the absolute number of monthly records, with a growing number of profiles meeting the >4 days per month



sampling frequency required for accurate long-term trend analysis (Gaudel et al., 2024; Chang et al., 2024). The sharp decline in profiles after 2018 is due mainly to reduce IAGOS measurements which may result from fewer commercial flights during COVID-19 (Cuesta et al., 2022).

2.3 Chemical reanalysis

TCR-2 is generated by assimilating satellite observations of ozone, CO, NO₂, HNO₃ and SO₂ from satellite instruments such as OMI, MLS and MOPITT into the MIROC-Chem model (Miyazaki et al., 2020a). In situ surface observations were not included in the data assimilation. When compared to worldwide ozonesonde measurements in the lower troposphere (up to 500 hPa), TCR-2 shows a low bias among world regions of -0.4 to 4.2 ppb (Miyazaki et al., 2020b). This study uses TCR-2 2-hourly ozone profiles from 9-11 am and 12-2 pm at a 1.125° x 1.125° resolution, where 9-11 am is intended to correspond to the overpass time of IASI+GOME2, and 12-2 pm to the overpass time of OMI/MLS and CrIS. Details on the input satellite data used in data assimilation of TCR-2 are summarized in Table S2.

2.4 Ground-level ozone observations

For evaluation, we use the latest TOAR-II ground-based observation dataset from 2005 to 2022, focusing on monthly mean DMA8 ozone. Building on TOAR-I (Schultz et al., 2017), TOAR-II represents the most extensive collection of global tropospheric ozone measurements, incorporating an expanded dataset that notably includes monitoring data from approximately 1,400 stations across China after 2015 (<https://toar-data.fz-juelich.de/gui/v2/dashboard/>, last access: 20 May 2025) (Schröder et al., 2021). As TOAR-II observations for the United States are only available through spring 2022, the observation dataset is supplemented with observations from the US EPA Air Quality System (AQS) for the remainder of 2022. We process the ground-based observations using a 75% data recording threshold, calculating a monthly mean only if a station reported valid DMA8 ozone values for at least 75% of the days that month. Due to the detection of possible incorrect units, we have removed all France data in 2014 (Wang et al., 2025). The total number of stations used in our evaluation that met this standard varied from a minimum of 3,355 in January 2005 to a peak of 7,073 in June 2022. The spatial distribution of monitoring stations and their proximity to the nearest vertical profiles is shown in Fig. 1(c), with more yearly maps available in Figure S1.

3. Methodology

As illustrated in Fig. 2, this study leverages vertical information from a chemical reanalysis to define a surface-to-column ratio, which is then used to convert satellite column ozone to ground-level ozone concentrations. We apply bias adjustment to both the satellite ozone observations (0–3 km or tropospheric column) and the chemical reanalysis surface-to-column ratios, utilizing global vertical profile measurements from ozonesondes and IAGOS. These bias adjustments are spatially propagated using the BME framework, which ensures a smooth correction field based on the spatial covariance structure. The final ground-level ozone concentration is estimated by multiplying the BME-adjusted satellite column ozone by the BME-adjusted surface-to-column ratio, applying the fundamental equation from Colombi et al. (2021) as follows:

$$\text{Surface ozone} = \text{Column ozone (satellite)} \times \frac{\text{Surface ozone (reanalysis)}}{\text{Column ozone (reanalysis)}} \quad (1)$$



3.1 Monthly trend of vertical profiles

To improve the reliability of our analysis, we calculate the monthly fitted value of vertical ozone profiles from daily observations of 44 ozonesonde and 48 IAGOS sites globally. While ozonesondes offer highly accurate data, their low sampling frequency (typically once a week) can create a significant bias when calculating monthly averages. Approximating a monthly mean from only four measurements can lead to an error of about 4 ppb (10%) (Chang et al., 2024). Following Chang et al. (2025), we applied a Generalized Additive Model (GAM) to calculate the monthly fitted value for both the 0-3 km and the tropospheric column ozone with respect to its seasonal and interannual changes.

The specific equation used to derive these values is:

$$y(t, h) = f_{season}(m, h) + f_{index}(t, h) + error \quad (2)$$

where h represents altitude, grouped into 100 m vertical bins from the surface to the top of the troposphere (for IAGOS data, only measurements above 200 m are used to avoid unusual ground-level measurements near airports); m represents the month (1-12); t is a continuous monthly index for all years with valid observations (N_{years}) at each site where $t = 1, \dots, 12 \times N_{years}$; and f is the basic function, for which we use Duchon splines. After filtering the data, we aggregate the ozone concentrations by altitude, month and year. The GAM above is then used to fit the aggregated data, allowing us to generate a continuous, monthly fitted vertical ozone profile at a 100 m resolution over years covered for each site. Examples of the resulting fitted time series are shown in Figure S2.

3.2 Filtering and Preprocessing Input Data

Data preprocessing aimed to standardize datasets including satellite products, TCR-2 and TOAR-II observations to a common $1^\circ \times 1^\circ$ resolution and monthly scale. This began with a filtering stage to ensure input data quality. For IASI+GOME2 and CrIS, we retained 0–3 km column ozone data only if the DOF exceeded 0.2. For IASI+GOME2 and OMI/MLS, we removed the data if cloud reflectivity was more than 0.30. For the CrIS dataset, we filtered for grid cells containing at least 10 valid daily records per month (see Figure S3 for data density details). We also excluded grid cells from CrIS and TCR-2 where the surface pressure was below 700 hPa or surface height is over 3000 m before proceeding with 0-3 km column ozone calculations.

Next, we defined vertical boundaries for integration using surface pressure from TCR-2 and tropopause data from NCEP. For OMI/MLS, we calculated monthly average tropospheric column ozone concentrations, using the defined vertical boundaries to convert the original Dobson Units (DU) into ppb in Equation 3:

$$O_3(\text{ppbv}) = 1270 \times \left[\frac{\text{Tr}O_3(\text{DU})}{P_{\text{surface}}(\text{hPa}) - P_{\text{tropopause}}(\text{hPa})} \right] \quad (3)$$

where $\text{Tr}O_3$ is the tropospheric ozone column concentration in DU, P_{surface} is the surface pressure obtained from TCR-2, and $P_{\text{tropopause}}$ is the tropopause pressure from NCEP.



In this study, tropospheric column ozone refers to the mean mixing ratio of the column rather than the total column density. Then, we aggregated the daily values from TCR-2, CrIS and IASI+GOME2 to produce monthly means. For TCR-2 and CrIS, we calculated the average concentrations of 0-3 km and tropospheric column ozone by integrating the multiple vertical layers using the trapezoidal rule. Finally, data from TOAR-II, TCR-2, and CrIS are regridded to the same $1^\circ \times 1^\circ$ grid as IASI+GOME2 and OMI/MLS. For sparse point data from TOAR-II, we used an inverse distance weighted method in regridding; when fewer than three data points were available within a 2° radius, a grid cell's value is derived from an average of its internal points. For TCR-2 and CrIS, we applied a bilinear interpolation to resample the data to the $1^\circ \times 1^\circ$ resolution.

3.3 BME data fusion and column to surface conversion

We employ the Bayesian Maximum Entropy (BME) framework, a geostatistical data fusion method that produces estimates of a homogeneous, stationary, space random field by integrating multiple forms of knowledge (Serre and Christakos, 1999; Christakos et al., 2004; Christakos, 1990). Prior to applying the BME framework, we evaluated the statistical distributions of all 0-3 km ozone, tropospheric column ozone and ratio datasets to verify the Gaussian assumption inherent to geostatistical estimation. As illustrated in Figure S4, the 0-3 km ozone concentrations, surface-to-column ratios (0-3 km), and tropospheric column ozone concentrations exhibited approximate normality and were modeled in their original data form. However, the surface-to-column ratios (tropospheric column) displayed a significant positive skew with a long tail and a high frequency of values near zero. To address this behavior, we applied a logarithmic transformation to the surface-to-column ratios (tropospheric column) prior to applying the BME framework. In this study, we employ a spatial-only BME framework, as the temporal component was already addressed by using a GAM to fit monthly values from the daily vertical profile observations.

We use the spatial-only BME for two primary estimation tasks. First, we estimate column ozone concentrations (0-3 km and troposphere) by integrating biased satellite observations (as the global offset) with observational data from ozonesondes and IAGOS. Second, we similarly estimate the surface-to-column ratio by integrating biased ratios derived from TCR-2 with the same ozonesonde and IAGOS data. The key of the BME method involves blending general knowledge (representing large-scale spatial patterns and correlation model derived from the data) with site-specific information (the knowledge of in-situ measurements) to produce robust estimates (Delang et al., 2021). In this study, the BME framework focuses exclusively on the spatial autocorrelation, because seasonal and interannual changes were explicitly used to fit the ozonesonde and IAGOS observations in Section 3.1. The BME estimation is implemented numerically using the BMElib library in MATLAB (Christakos et al., 2002). The final outputs for both the ozone columns and surface-to-column ratios are spatially continuous maps that keep the specific ozonesonde and IAGOS measurements at their locations. In regions distant from these vertical profiles, the maps smoothly transition to the background fields of TCR-2, guided by the derived spatial correlation models.

The general knowledge base of $X(s)$ at location s includes the mean function $m(s) = E[X]$, which is assumed to be zero in this study, and the correlation model $c_X(s, s') = E[(X(s) - m(s))(X(s') - m(s'))]$, which is determined



325 by the experimental covariance of data. For the homogeneous space random field $X(s)$, we use an exponential covariance model to best fit the experimental covariance as follows:

$$c_X(r) = c_0[\gamma \exp \frac{-3r}{a_{r1}} + \lambda \exp \frac{-3r}{a_{r2}}] \quad (4)$$

330 where r is the distance between location s and s' ; c_0 is the sill (variance); a_{r1} , a_{r2} are the spatial correlation ranges; γ and λ are the weights of each spatial correlation range. Numerical values for all parameters are provided in the Supporting Information Section S1. We observe that the surface-to-column ratios exhibit relatively short spatial correlation ranges (70% of the weight within 1° for 0-3 km ratio; 70% of the weight within 2° for the log-transformed tropospheric column ozone ratio), whereas ozone concentration fields maintain correlation over much larger distances
335 (40% of the weight within 5° for 0-3 km ozone; 40% of the weight within 10° for the tropospheric column ozone). This spatial correction range implies that the mean mixing ratio of column ozone remains informative up to a 10° radius, with the most significant spatial influence concentrated within 5° . This spatial autocorrelation range is larger than that of ground-level ozone, which generally exhibits a spatial correlation range of less than 1° (for around 50% weight) (Becker et al., 2023).

340

The BME framework is applied by the following methods, separately for the ozone concentration based on satellite observations and the surface-to-column ratio from TCR-2. We use a global offset ($o(s)$), representing the mean trend. Residuals (x_h) are calculated at each ozonesonde and IAGOS location (s_h) by subtracting the global offset value from the in-situ observation (z_h). Then we calculate general knowledge based on the spatial structure of these residuals,
345 assuming a zero mean and using a spatial covariance model ($c_X(r)$) described above fitted to the experimental covariance of the residuals to describe how their correlation decays with distance. We then define the site-specific information consisting of these residuals (x_h) at the in-situ observation locations, treated as hard data (with no uncertainty). The BME method integrates the general knowledge (zero mean, spatial covariance $c_X(r)$) and the site-specific information (hard data z_h and residuals x_h) using Bayesian principles to derive a posterior probability density
350 function (PDF) for the residual value (x_k) at any desired estimation location (s_k) (Abbott et al., 2025). Because only hard data are used for the site-specific information and the residual mean is assumed zero, this estimation step mathematically simplifies to Simple Kriging (SK) performed on the residuals (Delang et al., 2021). The SK estimate for the residual ($\tilde{x}_k$) at location (s_k) is essentially a weighted linear combination of the nearby hard data residuals (x_h), where the weights (λ) depend on the spatial covariance structure ($c_X(r)$). Specifically, the weights are calculated to
355 minimize the estimation error variance (σ_k^2). The SK equations for the estimated residual mean ($\tilde{x}_k$) and the estimation variance (σ_k^2) are:

$$\tilde{x}_k = m_X + c_{kh}c_{hh}^{-1}(x_h - m_h) \quad (5)$$

$$\sigma_k^2 = c_{kk} - c_{kh}c_{hh}^{-1}c_{hk} \quad (6)$$

360



where m_x is the assumed mean of the residuals (zero in this study), x_h is the vector of nearby hard data residuals, m_h is the vector of means at the hard data locations (zero in this study), c_{kk} is the variance of the residual at the estimation point (from the covariance model), c_{hh} is the covariance matrix between the hard data points, and c_{kh} (and its transpose c_{hk}) is the covariance vector between the estimation point and the hard data points.

365

Finally, the estimated ozone concentration or surface-to-column ratio ($\tilde{z}_k$) at location s_k is obtained by adding the global offset value $o(s_k)$ back to the SK estimate of the residual $\tilde{x}_k$:

$$\tilde{z}_k = \tilde{x}_k + o(s_k) \quad (7)$$

370

For the surface-to-column ratio (tropospheric column), which was modeled in the logarithmic form, we subsequently convert the estimates from the logarithmic form back to their original form. Further details on the specific definitions and parameters for BME framework can be found in Supporting Information Section S1.

375

As illustrated in Fig. 2, for application to IASI+GOME2, we use the satellite 0-3 km ozone measurements from IASI+GOME2, and surface-to-column ratios derived from the TCR-2 as our global offset column ozone and ratio fields. We employ the BME method to: (i) adjust the satellite 0-3 km (or tropospheric column) measurements using the fitted monthly ozonesondes and IAGOS 0-3 km observations as hard data, and (ii) adjust the TCR-2 surface-to-column ratio using ratios derived from the fitted monthly ozonesonde and IAGOS ratios as hard data. The BME method produces bias-adjusted monthly estimates of 0-3 km ozone and surface-to-column ratios at a $1^\circ \times 1^\circ$ resolution.

380

As shown in the first column of Fig. 2, the BME bias adjustment for 0-3 km ozone spans a much larger radius than for the vertical ratio, a direct outcome of the spatial covariance range (Equation 4). The covariance function indicates that the spatial autocorrelation for the surface-column ratio (0-3 km) decays rapidly beyond 1° , whereas the column-average ozone concentration maintains significant correlation over 5° . A similar structural pattern is observed for the tropospheric column ozone products, where the ozone concentration also exhibits a longer spatial correlation range than the surface-to-column ratio. However, the spatial correlation ranges for the tropospheric column ozone products are generally larger than those of the 0-3 km products, extending to approximately 10° for the concentration and 2° for the ratio. Finally, we calculate the global monthly DMA8 ground-level ozone estimates by multiplying the bias-adjusted satellite 0-3 km ozone by the bias-adjusted TCR-2 ratio at each grid cell (Equation 1). This produces a consistent set of monthly ground-level ozone maps ($1^\circ \times 1^\circ$). We adapt this same methodology to calculate ground-level ozone based on the tropospheric column ozone from CrIS and OMI/MLS.

390

Our approach offers several key improvements compared to the methodology of Colombi et al. (2021). Colombi et al. applied a uniform, regionally averaged bias correction to the chemical reanalysis based on seasonal means from ozonesondes and ground-based observations within each large region. This approach corrects the overall regional mean but may not capture nonlinear spatial patterns in bias within those regions. In contrast, our BME method adjusts the bias based on spatial covariance models, allowing for more localized and nonlinear bias correction. While Colombi

395



et al. focused their analysis on inferring ground-level ozone in only three specific regions (western US, eastern US, and China) and restricted their analysis to summertime data (June-August) when satellite retrieval performance is typically best (Cuesta et al., 2013), our BME method provides global coverage and is applied to all months from 2005 to 2022. Also, our method yields estimates at a finer spatial resolution, potentially offering more detail than the previous study. Throughout this paper, we refer to our method as the "BME-adjusted column to surface conversion" or the "BME-adjusted conversion method." We define the terms "BME-adjusted column ozone" and "BME-adjusted model surface to column ratio" to represent the column ozone concentrations and surface-to-column ratios as adjusted by the BME framework using vertical profiles. The term "BME-adjusted surface ozone" or "BME conversion satellite data" are used to denote the final surface ozone concentrations derived from applying BME-adjusted column to surface conversion method to the satellite data.

3.4 Evaluation of derived surface ozone with TOAR-II ground observations

We evaluate the derived surface ozone fields with respect to monthly average DMA8 ozone TOAR-II observations (2005–2022). The primary metrics for evaluation are the coefficient of determination (R^2) and root mean square error (RMSE), which are analyzed across different seasons and regions. For comparison, we evaluate the performance of different datasets against TOAR-II observations over a common temporal range: (1) the original satellite observations (monthly average), (2) TCR-2 (monthly average DMA8), and (3) our final BME-adjusted surface ozone (monthly average DMA8). We quantify the BME-adjusted conversion method against the conversion method without the BME adjustment process (only using the TCR2 ratio directly in Equation 1), focusing on the distance from a given TOAR-II grid cell to the nearest vertical profile location. Furthermore, we conduct an analysis to investigate how indicators of satellite observation quality for IASI+GOME2 impact the final derived surface ozone estimates.

For the surface ozone derived from OMI/MLS, we also perform a long-term trend evaluation against TOAR-II trends from 2005 to 2022, using area-weighted ozone means. Prior to trend estimation, we deseasonalize the time series to remove the influence of the seasonal cycle, consistent with TOAR-II recommendations. The seasonal signal is modeled using four harmonic terms based on monthly data from 2005 to 2022 (Supporting Information Section S2). After removing the seasonal signal for each region, trends are calculated using median quantile regression, and the confidence in each trend was classified by its p-value: very high ($p \leq 0.01$), high ($0.01 < p \leq 0.05$), medium ($0.05 < p \leq 0.1$), low ($0.1 < p \leq 0.33$), or no evidence ($p > 0.33$). Additionally, the 95% confidence interval (CI) for each dataset was determined through a linear regression against the gridded TOAR-II observations. We use regional groupings defined by Hemispheric Transport Air Pollution 2 (HTAP2) (Koffi et al., 2016), as detailed in Table S3.

4. Results

4.1 Evaluation of BME-adjusted conversion for IASI+GOME2, OMI/MLS and CrIS in different regions

Based on Table 2, the BME-adjusted conversion method substantially improves the performance of all satellite products compared to their original performance. The adjustment and conversion produce higher R^2 values which indicates a stronger correlation with TOAR-II observations and lower RMSE values across different regional scales



which indicates a lower bias against TOAR-II observations. The high RMSE for the original tropospheric column data results from the fact that ozone is typically higher in the middle and upper troposphere, and is expected (Van Malderen et al., 2025). Consequently, the improvement in RSME achieved by the adjustment and conversion is particularly obvious for the tropospheric column datasets (OMI/MLS, CrIS). For the OMI/MLS product, the BME-adjusted conversion increases global R^2 from 0.26 to 0.57 and its global RMSE reduces from 22.21 ppb to 7.79 ppb. The BME-adjusted surface ozone from IASI+GOME2 (0-3 km) shows a global R^2 increase from 0.25 to 0.51 and global RMSE reduce from 10.74 ppb to 9.44 ppb. The CrIS tropospheric column also shows a great improvement, with its R^2 increasing from 0.31 to 0.57 and its RMSE being cut by more than half, from 29.52 ppb to 10.81 ppb. Notably, the CrIS 0-3 km product achieves higher precision, with R^2 increasing from 0.30 to 0.56, and RMSE dropping to 9.45 ppb.

For tropospheric column products (OMI/MLS and CrIS), the BME-adjusted conversion proved universally beneficial across all metrics and regions in Table 2. In every region, the BME-adjusted surface ozone shows a much lower RMSE compared to the original data. For the “other” region, which typically lacks dense ground monitoring, we observed small improvements in both OMI/MLS and CrIS tropospheric column products. While increases in spatial correlation were modest, the reduction in bias was substantial as expected. Although the BME correction method provides a substantial improvement on tropospheric column products, its efficacy for satellite 0-3 km ozone observation is not uniform among regions. Performance improved most notably in the EU-27 and the United Kingdom. In China and the United States, improvements were positive but less pronounced. Furthermore, in Japan where the original 0-3 km satellite data already demonstrated strong performance, the BME-adjusted conversion yielded no improvement. This suggests that while the adjustment and conversion are highly effective at fixing large bias, it may not provide a large benefit for specific regions where the original satellite LMT (0-3 km) ozone observations already perform well.

The BME-adjusted surface ozone from the CrIS 0–3 km dataset exhibits better performance for global and regional bias compared to the CrIS tropospheric column product. Consequently, we focus exclusively on the CrIS 0–3 km data for the remainder of this study. Figure 3 presents the spatial distribution of the BME-adjusted surface ozone and the difference between BME-adjusted surface ozone and original data. All three BME-adjusted surface ozone datasets show consistent spatial patterns, such as high ozone concentrations over Asia during the summer, and low ozone concentrations over remote ocean areas. For IASI+GOME2 (Fig. 3a), the surface ozone map reveals a significant positive adjustment over polluted land regions, including North Africa, the Middle East, and parts of East Asia, particularly during the summer. In contrast, the OMI/MLS map (Fig. 3b) shows a widespread negative adjustment, meaning that the correction consistently reduces the ozone estimates, as the original OMI/MLS data represent tropospheric column averages. The adjustment and conversion of the OMI/MLS tropospheric column ozone is strongest over the tropical oceans, while the adjustment and conversion is minimal over areas like Central Africa and East Asia. The CrIS 0-3 km ozone (Fig. 3c) has a negative adjustment and conversion over most ocean areas, North America, and Europe, but a strongly positive adjustment over land area in North Africa and Australia, as well as in polluted regions of East Asia and Southeast Asia during the summer. Notably, a comparison of the seasonal maps (left



470 columns) shows that the 2022 CrIS data (Fig. 3c) has higher concentrations than the 2017-2020 IASI+GOME2 data (Fig. 3a) in North Africa, East Asia and Southeast Asia.

Figure 4 provides a regional analysis of the bias after adjustment and conversion, and the impact of BME-adjusted conversion method for the three satellite datasets in Europe (EUR), North America (NAM), and East Asia (EAS). The maps for IASI+GOME2 (Fig. 4a) show a relatively low bias after BME correction. In Europe and North America, there is a slight underestimation in winter and an overestimation in summer, with most bias remaining within ± 10 ppb. The low bias, particularly in the European winter, may be related to the denser distribution of nearby vertical ozone profiles. OMI/MLS (Fig. 4b) shows a very similar bias spatial pattern to IASI+GOME2. The impact maps show a dramatic reduction in absolute bias after the adjustment and conversion, especially in East Asia and Europe, but the final product still tends to overestimate. Similarly, CrIS 0-3 km (Fig. 4c) still suffers from a clear overestimation in winter across all three regions. Although the impact maps show a decrease in absolute bias in winter, a significant positive bias remains. All three satellite products exhibit a consistent seasonal pattern, with the adjustment and conversion having a more pronounced impact during the winter than the summer. Even after BME-adjusted conversion, the satellite products continue to overestimate ground-level ozone in some regions, although the magnitude of the bias is strongly reduced.

4.2 Assessment of improvement over different months and years

Table 2 indicates that the BME-adjusted conversion method improves the performance of satellite products over different regions. To investigate the stability of this method over time, we conducted validation across different months and years to assess temporal consistency. In Fig. 5, the R^2 of the derived surface ozone is superior to the original satellite ozone across all months for all products. Its performance exhibits a seasonal cycle consistent with surface ozone concentration changes. The R^2 of derived surface ozone against TOAR-II observations peaks in the summer months and declines in the winter. For the CrIS 0-3 km dataset, the BME-adjusted conversion method significantly improves performance, especially in winter when the original satellite data exhibit extremely low R^2 and very high RMSE. The BME conversion for CrIS 0-3 km ozone yields better R^2 during the winter compared to the original data, when surface ozone concentration is low. In terms of RMSE, the derived surface ozone from CrIS 0-3 km ozone performs better during the spring and fall. The monthly performance in Fig. 5(a) and 5(b) show that the OMI/MLS based BME-adjusted conversion consistently outperforms the original satellite data, with BME-adjusted surface ozone exhibiting similar seasonal trends in R^2 (peaking in late summer, decline in winter). The BME conversion for OMI/MLS maintains a lower RMSE over almost all months in 2017 to 2020 and 2022. As indicated in Figure S5, the improvement in R^2 is primarily between 60°S and 60°N , with less enhancement at higher latitudes. The improvement in bias exhibits greater temporal stability than the effect on R^2 across months, demonstrating that the BME method is robust in reducing systematic errors across different seasons. The improvements are particularly better in winter yielding notably larger bias reduction, when the original satellite data performs poorly.



Figure 6 evaluates the yearly performance of both the BME-adjusted surface ozone estimates and the original satellite retrievals against TOAR-II surface observations. The results clearly demonstrate that over the 18-year period, the BME conversion method provides a stable improvement across multiple satellite data products, continuously enhancing the accuracy of surface ozone estimates. For the long-term OMI/MLS record (2005–2022), this conversion method increases the R^2 from 0.2–0.4 to 0.5–0.7, indicating a much stronger correlation with surface observations. Concurrently, the RMSE decreases from 20–25 ppb to 6–9 ppb, substantially reducing the system bias. The BME-adjusted surface ozone estimates also show greater stability with less fluctuation, and enhanced performance when applied to different satellite products.

When compared to the TCR-2 chemical reanalysis (Fig. 7), the BME-adjusted conversions demonstrate notable improvements in minimizing bias. For 2017–2020, both IASI+GOME2 and OMI/MLS yield lower RMSE across all TOAR-II sites (9.44 ppb and 8.30 ppb, respectively) than TCR-2 (11.15 ppb). The derived surface ozone estimates align more closely with the 1:1 reference line, whereas TCR-2 overestimates surface ozone. Although the satellite-derived products exhibit slightly lower overall spatial correlations than TCR-2, they achieve higher R^2 values specifically in rural areas. This trend largely persists in 2022 (Fig. 7), with OMI/MLS continuing to show a reduced RMSE despite a slightly lower R^2 , though the surface estimates converted from CrIS do not outperform TCR-2 in RMSE. Also, from Figure S6, derived surface ozone estimates from OMI/MLS over 2005–2022 maintains a lower bias than TCR-2. In general, satellite-based estimates offer reduced bias compared to chemical reanalysis, and they consistently perform better in rural regions than in urban ones. This persistent performance gap between urban and rural areas may underscore the inherent difficulties in resolving the high spatial heterogeneity characteristic of urban environments, as well as the complex and localized dynamics of emissions.

4.3 Analysis of Factors Influencing the BME-adjusted conversion for IASI+GOME2

Figure 8 evaluates the BME-adjusted surface ozone estimates from IASI+GOME2 0–3 km ozone against TOAR-II ground observations (2017–2020), analyzing how the method's accuracy changes based on the distance to vertical profiles, and the DOF and height of maximum sensitivity reported by IASI+GOME2. This analysis explicitly distinguishes the respective impacts of the vertical ratio conversion and the BME bias adjustment. First, Fig. 8(a) directly comparing the raw IASI+GOME2 0–3 km column to ground observations yields a weak spatial correlation (R^2) and high bias (RMSE). This is expected, as the lower-tropospheric column does not directly reflect actual ground-level ozone. If only the TCR-2 vertical ratio is applied to downscale this column data to surface concentrations, the surface spatial patterns are successfully captured, significantly improving the spatial correlation (R^2 increases from 0.2 to 0.5). However, this ratio conversion alone is insufficient to resolve the inherent biases. In fact, relying solely on the ratio conversion leads to an increase in RMSE compared to the raw satellite data.

Conversely, applying the BME joint adjustment before ratio conversion drastically reduces this absolute error and further improves R^2 , though exhibiting different spatial behavior. Within a 6° radius of the vertical profile location, the BME-adjusted conversion increases R^2 compared to the conversion without BME adjustment. Beyond this range, the advantage gradually diminishes, eventually converging with the unadjusted ratio conversion method. In contrast,



the BME adjustment's ability to reduce RMSE is not limited by this spatial range. Across all evaluated distances, the BME conversion method consistently maintains a lower bias than the unadjusted conversion method, effectively dropping the RMSE from 10 ppb to 7.5 ppb. This bias reduction pattern is also seen in the BME conversion results for OMI/MLS data over 2005 to 2022 (Figure S7). These results clearly illustrate the distinct roles of the two processes: the ratio conversion primarily improves spatial correlation (R^2) but introduces additional error in the bias. The joint BME adjustment (column ozone and surface-to-column ratio), on the other hand, is essential for correcting the bias and, within its effective range (up to 6°), further enhances spatial correlation. By comparing the surface ozone estimates across these three scenarios, it is evident that vertical profile data from IAGOS and ozonesondes can effectively constrain and adjust satellite data over a broad spatial range. This effective range far exceeds the typical $<1^\circ$ limit of the spatial autocorrelation of ground-level ozone (Becker et al., 2023; Delang et al., 2021). This is likely because ozone mixing is more thorough in the troposphere than at the surface (Gaudel et al., 2024).

The efficacy of our BME method is closely linked to the quality of the original satellite observations, as shown in Fig. 8(b) and Fig. 8(c). Figure 8(b) examines the satellite's sensitivity to the lower troposphere using the DOF (0–3 km) metric defined in the methods. A higher DOF indicates a greater amount of independent information retrieved from this layer, which generally corresponds to lower uncertainty in satellite observation. As seen in the original satellite data, increased DOF correlates with improved performance (higher R^2 and lower RMSE) and higher retrieved surface ozone concentrations. Compared to the original satellite data, the BME-adjust conversion yields the most significant gains in R^2 and RMSE when the DOF is low (the raw satellite data quality is poor). Our method consistently achieves a higher R^2 than the ratio conversion (without the BME adjustment), though this performance gap decreases as DOF increases. Furthermore, while the ratio conversion tends to increase RMSE, the BME-adjusted conversion effectively mitigates bias. Similarly, Fig. 8(c) analyzes the impact of the height of maximum sensitivity. The original IASI+GOME2 0–3 km data exhibits a negative correlation between the height of maximum sensitivity and performance, meaning that 0–3 km satellite data quality worsens as the height of maximum sensitivity increases, likely because the satellite's peak signal is shifted into the free troposphere and away from the near surface layer. Consistent with the DOF findings, our BME-adjusted conversion proves most beneficial when the height of maximum sensitivity is high (data quality is low). While the BME method outperforms the ratio conversion in terms of R^2 , this advantage reduces when the height of maximum sensitivity is near the surface. Notably, regarding RMSE, the ratio correction (without the BME adjustment) increases bias, whereas the BME-adjusted conversion effectively reduces it.

4.4 Long Term Trend of OMI/MLS BME-adjusted Surface Ozone

Figure 9 displays time-series data for the OMI/MLS tropospheric column, BME-adjusted surface ozone estimates, and TOAR-II observations from 2005 to 2022 at the grid cells of TOAR-II stations across Europe (EUR), North America (NAM) and East Asia (EAS). The original OMI/MLS satellite data consistently higher than surface concentrations as it measures the entire troposphere. In Europe, the discrepancy between TOAR-II and BME-adjusted surface ozone is very small, although the estimate reveals an increasing trend with very high confidence (0.23 [0.20, 0.26] ppb/year, $p < 0.01$) after removing seasonal effects. The effect is particularly highlighted in North America, where our BME-adjusted surface ozone successfully captured a downward trend with very high confidence (-0.11 [-0.16, -0.04]



580 ppb/year, $p < 0.01$), while the original OMI/MLS data shows an increasing trend with very high confidence. In East Asia, TOAR-II observations show a significant increasing trend with very high confidence (0.32 [0.23, 0.48] ppb/year, $p < 0.01$), which may be related to rapid industrialization and economic growth (Zhan et al., 2023), and the discrepancy between BME-adjusted surface ozone and TOAR-II observations narrowed during this period. This time series plot demonstrates the effectiveness of the BME-adjusted conversion method. Although discrepancies in long term trend
585 directions persist, alignment with the magnitude of TOAR-II observational data has improved.

5. Discussion

The validation results in Table 2 demonstrate that the BME-adjusted conversion method significantly improves surface ozone estimation from satellite data. This performance enhancement is driven by the use of observed vertical profiles to adjust the surface-to-column ratio from TCR-2 and column ozone observations from the satellites. The surface-to-
590 column ratio effectively characterizes key physical processes such as boundary layer mixing and photochemical production (Colombi et al., 2021). Specifically, a high ratio implies strong near-surface photochemical production coupled with a stable boundary layer that traps ozone near the ground. Conversely, a low ratio indicates vertical mixing or convective transport that moves ozone concentrations into the free troposphere. The vertical profile observations of column ozone also play a critical role by reducing systematic errors in the satellite column observations (0–3 km
595 and tropospheric). By fusing these fitted monthly vertical profiles from ozonesondes and IAGOS, the BME framework removes spatially varying retrieval biases, which arise from the sensor or a priori assumptions. Thus, the final derived surface ozone estimates benefit from the surface-to-column ratio conversion, which constrains vertical distribution physics, and the column ozone adjustment, which calibrates the magnitude of the ozone concentration.

600 The spatial maps (Fig. 3 and Fig. 4) and validation metrics (Table 2) reveal crucial insights in how the BME adjustment interacts with satellite products. For tropospheric column products, which exhibit a weak correlation with surface ozone due to the integration of the full tropospheric column, the BME adjustments yields significant improvements globally. However, for the 0–3 km ozone products, which have better sensitivity to the lower atmosphere, the efficacy of the adjustment is not uniform spatially. The pronounced regional heterogeneity in adjustment performance
605 highlights the BME method's dependence on three types of input data: the quality of the original satellite retrieval, the density of the vertical profile network, and the surface-to-column ratio accuracy of the chemical reanalysis. In the United States and Europe, the method achieves its highest performance gains despite the large biases in the original satellite data. This success is driven by the dense network of vertical profile observations, which allows the BME framework to effectively model the local spatial covariance structure and constrain the satellite retrieval. In contrast,
610 regions with sparse vertical profiles observations (like China) depend heavily on the TCR-2 ratio, since the spatial correlation of the surface-to-column ratio decays rapidly. However, when this reanalysis uncertainty interacts with original satellite data, the BME-adjusted conversion effectively introduces bias, explaining the performance degradation observed in some regions (like Japan, South Korean).



615 As noted by Colombi et al. (2021), the combination of high surface temperatures (critical for TIR sensors) and high
solar elevation angles (critical for UV sensors) maximize the DOF for signal in the retrieval. However, summer also
has active photochemical production, which introduces high ozone concentrations and complex vertical mixing
profiles that are challenging to model. Consequently, even though satellite datasets perform well in R^2 , they still have
larger bias during this season. In contrast, winter observations are often constrained by low thermal contrast, causing
620 the satellite retrieval to rely heavily on a priori information rather than the actual ozone concentration (Pennington et
al., 2025). While the signal is naturally lower in winter due to lower background ozone concentrations, the relative
improvement provided by the BME-adjusted conversion is more pronounced than in summer. Another consistent
pattern observed in Fig. 7 is that derived surface ozone performs better in rural areas than in urban areas for all satellite
products. This disparity is likely driven by representativeness errors associated with the original satellite data. Rural
625 ozone concentrations are more spatially uniform, meaning that a single satellite pixel can better represent a broad area.,
while urban environments tend to exhibit sharper ozone gradients over short distances.

The uncertainty maps (in Figure S8) of BME-adjusted surface ozone reflect uncertainties from three primary inputs:
the original satellite retrieval data, the chemical reanalysis (TCR-2), and the vertical ozone profile. The satellite data
630 quality is linked to the information content of the satellite retrieval. In low information scenarios, when characterized
by low DOF for the signal, satellite retrievals largely revert to the a priori climatology, resulting in a lack of
independent sensitivity to real surface ozone variability. Although the LMT (0-3 km) product is theoretically better
able to measure near-surface ozone, performance disparities suggest this advantage is often compromised by the
critical influence of sampling time. In Fig. 5, we found that data fusion-based estimate of surface ozone using
635 OMI/MLS tropospheric columns and IASI+GOME2 LMT show similar results for both correlation and bias (it is
slightly better using the OMI/MLS, although the period of comparison is not the same). It is worth noting that
OMI/MLS measurements around 1:45 PM local time likely capture the daily photochemical peak, and would be
expected to correspond well with daily DMA8 ozone. On the other hand, IASI+GOME2 measurements around 9:30
AM capture the rapid morning transition period, which may be less correlated with DMA8. In Fig. 6, our BME-
640 adjusted conversion method robustly enhances performance amidst year-to-year variations, effectively smoothing out
the interannual fluctuations present in the original satellite datasets.

Although the chemical reanalysis (TCR-2) provides crucial surface-to-column ratios in our method, it still contains
errors. As shown in Fig.7, TCR-2 overestimates surface ozone concentrations. However, our methodology mitigates
645 this by utilizing the surface-to-column ratio rather than the absolute concentration as the input. This strategy rests on
the premise that even if the reanalysis model possesses a bias in magnitude, its simulation of the atmospheric vertical
structure may remain relatively reliable (Colombi et al., 2021). The ratio-based conversion is also applicable to the
retrieval of other air pollutants (Malings et al., 2024). Another source of uncertainty arises from the temporal
representativeness of the vertical profiles. Ozone sonde observations are typically taken at irregular and sparse intervals
650 throughout the week. Although daily homogenized ozone sonde observations are precise, calculating monthly means
from sparse daily observations introduces under-coverage bias due to the high inter-daily variability of free



655 tropospheric ozone (Chang et al., 2024; Chang et al., 2025). However, the limited sampling density of historical
ozonesondes makes them inadequate for robust trend detection. Recognizing this limitation, our primary objective in
utilizing these vertical profiles is not to calculate temporal trends, but rather to adjust the baseline magnitude of the
satellite column ozone concentrations. We applied a filter standard, keeping only locations with at least 4 profiles per
month. This limited sampling size successfully improves the overall magnitude, which is evidenced by our validation
results after applying the BME-adjusted column-to-surface conversion. However, it does not overcome the inability
to capture the precise temporal trend. Indeed, our derived trend estimates, which are further complicated by the
inherent uncertainties and limitations of the satellite data itself, remain insufficiently accurate, as explicitly reflected
660 in Figure 9.

6. Conclusion

This study represents the first global application of a geostatistical framework to estimate ground-level ozone based
on multiple satellite ozone datasets, effectively leveraging vertical profile observations from ozonesondes and IAGOS,
alongside the surface-to-column ratio from a chemical reanalysis (TCR-2). Our global validation against TOAR-II
665 observations demonstrates that the BME-adjusted conversion method significantly improves the accuracy of surface
ozone estimates from OMI/MLS, IASI+GOME2, and CrIS satellite products. The success of this approach is driven
by dual adjustments to the surface-to-column ratio to constrain vertical distribution physics, and to the satellite
observations to reduce systematic biases in the satellite column magnitude. In contrast to purely data-driven machine
learning approaches (e.g., Xue et al., 2020; Wei et al., 2022), which often treat the atmosphere as a 2D surface and
670 rely heavily on dense ground monitoring networks for training, our method integrates a physics-based vertical
constraint via the surface-to-column ratio conversion. Another key advantage of our approach is the use of spatial
autocorrelation with vertical profile observations to weight corrections; this allows the BME framework to capture
non-linearities and heterogeneous bias structures that uniform regional corrections may miss, and to smoothly apply
these corrections spatially. Furthermore, unlike previous studies restricted to specific regions, temporal scales, or data
675 from a single satellite, we established a robust framework for correcting multiple satellite datasets over long-term
periods on a global scale.

Our analysis shows that vertical profile observations exert the most significant corrective influence on spatial
correlation (R^2) when satellite measurements are within 6° radius. In contrast, for systematic bias (RMSE), the
680 adjustment range is considerably larger. Even where vertical profile data are sparse, the method remains highly
effective at reducing the satellite retrieval bias, particularly evident in the widespread improvement of RMSE. By
demonstrating this capacity to correct surface ozone estimates, this study extends the utility of these vertical profile
datasets beyond their traditional applications in tropospheric ozone trend analysis. Another innovation is the complete
exclusion of ground monitoring station data from the adjustment process. While previous approaches often assimilated
685 ground observations to improve local accuracy, this can bias models toward data-rich regions. By relying exclusively
on sparse vertical profiles for reducing biases, our method generates an independent satellite-based surface ozone
product.



690 Despite these achievements, several limitations persist. The method's adjustment and conversion are limited by the spatiotemporal density of the vertical profile observations and the quality of chemical reanalysis. Therefore, it offers minimal benefits in regions where vertical profiles are few and chemical reanalysis performance is poor. Furthermore, temporal variations in the magnitude of bias correction across months and years can mathematically mask underlying physical trends. Even after adjustment, persistent biases remain in satellite-derived products. These residuals stem from inherent satellite data uncertainties, the inability of chemical reanalysis surface-to-column ratios to fully capture vertical distributions, and the lack of vertical profile observations in many regions. Consequently, these results underscore the need for increasing observations of ozone vertical profiles. In summary, the BME-adjusted surface ozone offers a unique, independent estimate derived from satellite observations. We show that it performs comparably with a chemical reanalysis, but not as good as ground-level ozone datasets that are currently available. Its primary value therefore lies not as a global 'best estimate,' but as a background field based on satellite data for use in future data fusion frameworks and global air quality assessments particularly in monitor-sparse regions.

7. Code and data availability

Ground observational data are publicly available from the TOAR-II data portal (last accessed on 20 May 2025, <https://toar-data.fz-juelich.de/gui/v2/dashboard/>). The homogenized ozonesonde data (Van Malderen et al., 2025) can be obtained by contacting the HEGIFTOM working group. IAGOS data are publicly available from the data portal (last accessed on 20 May 2025, iagos.org). The TCR-2 reanalyses data (Miyazaki, 2024) are publicly available from https://disc.gsfc.nasa.gov/datasets/TRPSCRO32H2D_1/summary. IASI+GOME2 data (Cuesta et al., 2013) are publicly available from https://iasi.aeris-data.fr/O3_IAGO2/ and can be obtained by contacting the creators. OMI/MLS data (Ziemke et al., 2006) are publicly available from https://acd-ext.gsfc.nasa.gov/Data_services/cloud_slice. CrIS data (Bowman, 2022) are publicly available from https://daac.gsfc.nasa.gov/datasets/TRPSYL2O3CRS1FS_1/summary. The BME-adjusted dataset produced in this study will be made publicly available on a public website (Zenodo) prior to publication.

8. Author Contributions

This research was conceived by HW, JJW, MLS and KM. KM, JC, and JZ provided satellite ozone observations. KM provided chemical reanalysis dataset. HW collected the TOAR-II observational data and vertical profiles. Data analyses and numerical results were generated by HW with input from MLS and JJW. HW and JJW wrote the paper and all authors provided edits and comments on drafts.

9. Competing Interests

Some authors are members of the editorial board for *Atmospheric Chemistry & Physics*. The authors declare that they have no other conflicts of interest.



720 10. Financial Support

We gratefully acknowledge support for this work through National Aeronautics and Space Administration (NASA) grants #NNX16AQ30G and #80NSSC23K0930. Part of this work was conducted at the Jet Propulsion Laboratory, California Institute of Technology, under contract with NASA. We also acknowledge the support of the NASA Atmospheric Composition: Aura Science Team Program (19-AURAST19-0044), Atmospheric Composition
725 Modeling and Analysis Program (22-ACMAP22-0013), NASA Earth Science U.S. Participating Investigator program (22-EUSPI22-0005).

Table 1: Overview of input data. Each dataset is converted from its original frequency to monthly averages.

Dataset	Frequency	Record local time	Spatial resolution	Year	Type
IASI-GOME2	Daily	9:30 am	1° x 1°	2017-2020	Satellite
OMI/MLS	Daily	1:45 pm	1° x 1°	2005-2022	Satellite
CrIS	Daily	1:30 pm	14 km x 14 km	2022	Satellite
TCR-2	2-Hourly	9-11 am/12-2 pm	1.125° x 1.125°	2005-2022	Chemical reanalysis
Ozonesonde	Daily	Variable	44 locations	2005-2022	Vertical profile
IAGOS	Daily	Variable	48 locations	2005-2022	Vertical profile

730

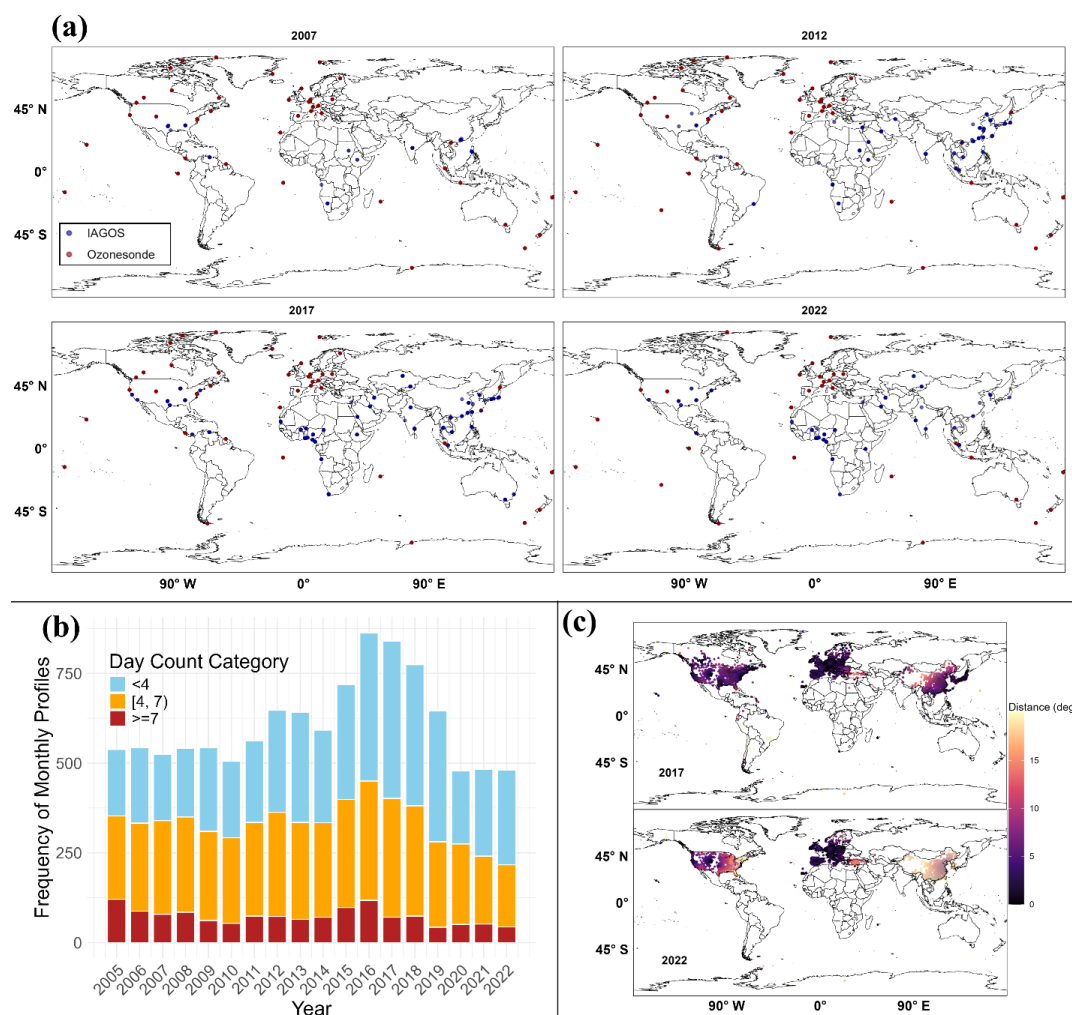


Figure 1: (a) Maps of ozonesonde and IAGOS locations in 2007, 2012, 2017, 2022. (b) Annual distribution of daily vertical profile records per month. (c) Maps illustrating the spatial proximity of TOAR-II stations to the nearest vertical profile for the years 2017 and 2022. Day Count Category means total number of days with data records in a given month.

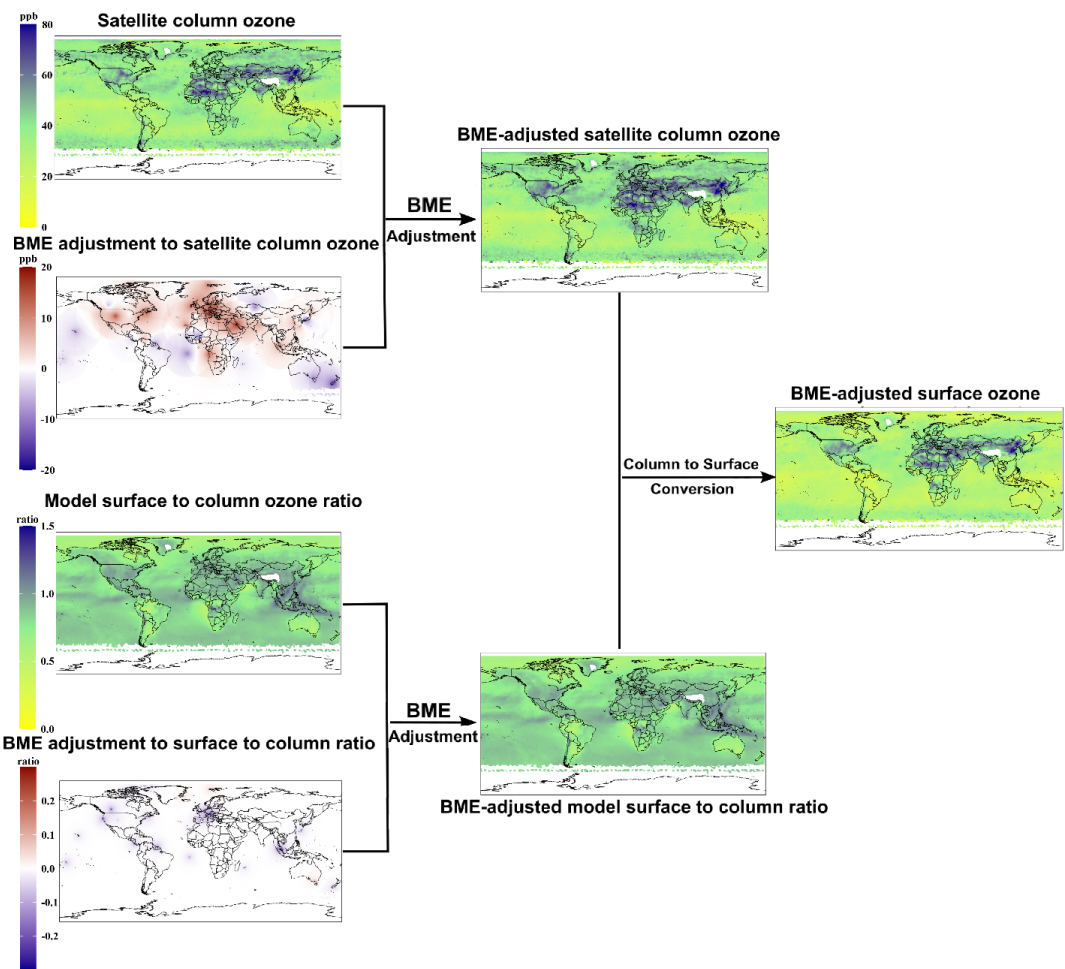


Figure 2: Method workflow illustrated for June 2017 IASI+GOME2.

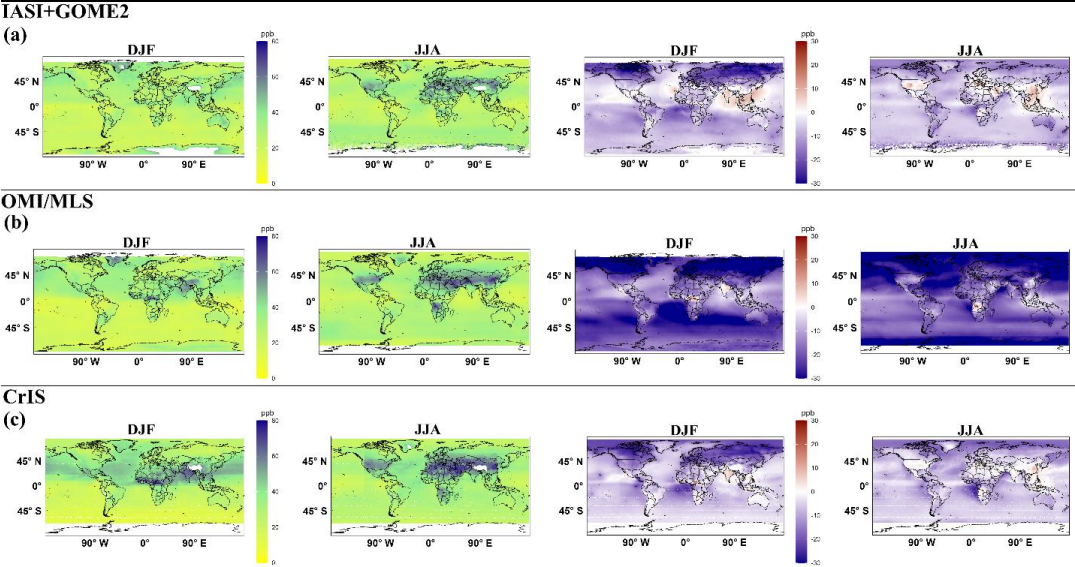
Table 2: Comparison of satellite product performance for countries (and the EU) with the most monitors against TOAR-II observations of monthly DMA8 ozone. As satellite estimates are not direct measurements of surface ozone and thus may not align perfectly with TOAR-II observations, this table is provided to demonstrate the magnitude of improvement across different regions after applying the BME-adjusted conversion method. Estimate is the spatial area-weighted average of the grid estimates for each satellite product at the 1°*1° resolution re-gridded TOAR-II locations in each country. The Lower and Upper Bound represent the 95% confidence interval for the estimate, calculated from the linear regression of each dataset against TOAR-II observations. Country names are China (CHN), Japan (JPN), South Korea (KOR), United Kingdom (UK), United States of America (USA). EU-27 includes Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden. Others are all other countries in TOAR-II apart from those listed. GLO means global, which includes all stations from TOAR-II.



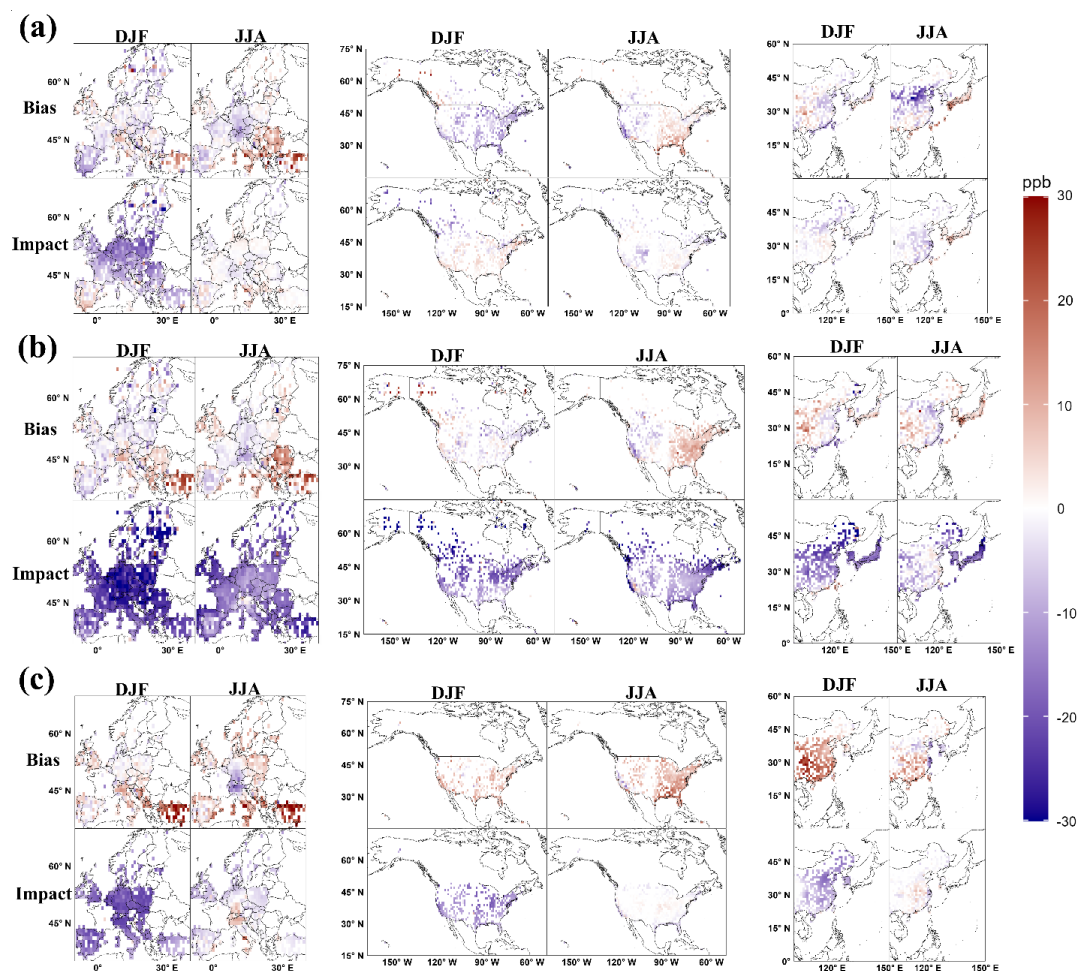
dataset	Country	GLO	CHN	EU-27	JPN	KOR	UK	USA	Others
TOAR 2017-2020	Mean	39.07	45.80	36.40	37.87	44.97	32.45	40.17	32.01
IASI+GOME2 0-3 km 2017-2020	R ²	0.25	0.41	0.16	0.46	0.57	0.04	0.28	0.08
	RMSE	10.74	12.55	10.78	7.02	8.43	9.65	8.13	15.29
	Estimate	39.93	42.07	40.10	38.93	41.17	38.05	38.79	39.35
	Lower	25.66	27.80	25.83	24.66	26.90	23.78	24.52	25.08
	Upper	54.20	56.34	54.37	53.20	55.44	52.32	53.06	53.62
IASI+GOME2 0-3 km BME-adjusted surface ozone 2017-2022	R ²	0.51	0.57	0.61	0.28	0.45	0.43	0.47	0.22
	RMSE	9.44	11.74	7.03	8.41	10.45	5.29	8.14	12.94
	Estimate	36.38	40.05	34.82	38.39	39.30	32.47	36.86	35.17
	Lower	21.90	25.57	20.34	23.91	24.82	17.99	22.38	20.69
	Upper	50.86	54.53	49.30	52.87	53.78	46.95	51.34	49.65
TOAR 2005-2022	Mean	38.32	44.73	35.82	38.29	42.66	31.52	41.26	32.50
OMI/MLS tropospheric column 2005-2022	R ²	0.26	0.28	0.31	0.23	0.38	0.21	0.29	0.16
	RMSE	22.21	20.65	23.79	21.17	20.25	24.67	18.77	26.41
	Estimate	57.83	60.63	57.58	57.13	60.52	54.98	57.79	55.00
	Lower	41.63	44.43	41.38	40.93	44.32	38.78	41.59	38.80
	Upper	74.03	76.83	73.78	73.33	76.72	71.18	73.99	71.20
OMI/MLS BME-adjusted surface ozone 2005-2022	R ²	0.57	0.56	0.63	0.50	0.47	0.50	0.52	0.26
	RMSE	7.79	10.18	7.05	7.62	8.75	5.30	6.86	11.77
	Estimate	37.41	43.45	34.51	40.69	43.07	31.37	39.61	35.15
	Lower	24.21	30.25	21.31	27.49	29.87	18.17	26.41	21.95
	Upper	50.61	56.65	47.71	53.89	56.27	44.57	52.81	48.35
TOAR 2022	Mean	40.33	46.42	37.31	49.83	50.30	33.12	40.34	32.18
CrIS 0-3 km 2022	R ²	0.30	0.25	0.52	0.71	0.23	0.19	0.27	0.15
	RMSE	16.48	16.41	16.41	4.64	9.05	16.57	13.57	25.07
	Estimate	53.22	56.77	51.69	49.96	50.37	48.68	52.04	53.41
	Lower	43.13	46.68	41.60	39.87	40.28	38.59	41.95	43.32
	Upper	63.31	66.86	61.78	60.05	60.46	58.77	62.13	63.50
CrIS 0-3 km BME-adjusted surface ozone 2022	R ²	0.56	0.64	0.68	0.49	0.5	0.54	0.48	0.22
	RMSE	9.45	11.88	8.77	6.21	8.69	8.26	9.24	20.39
	Estimate	47.79	54.4	42.94	49.31	48.49	39.44	48.17	46.51
	Lower	34.46	41.07	29.61	35.98	35.16	26.11	34.84	33.18
	Upper	61.12	67.73	56.27	62.64	61.82	52.77	61.5	59.84
CrIS tropospheric	R ²	0.31	0.26	0.50	0.33	0.17	0.38	0.28	0.15
	RMSE	29.52	28.52	29.11	21.37	24.56	30.26	27.38	39.88



column 2022	Estimate	67.90	71.73	65.45	69.86	72.89	62.96	66.79	68.98
	Lower	52.42	56.25	49.97	54.38	57.41	47.48	51.31	53.50
	Upper	83.38	87.21	80.93	85.34	88.37	78.44	82.27	84.46
CrIS tropospheric column BME-adjusted surface ozone 2022	R ²	0.57	0.60	0.66	0.54	0.54	0.52	0.50	0.22
	RMSE	10.81	11.93	10.05	6.19	9.36	7.08	9.17	21.11
	Estimate	48.11	53.79	43.78	51.89	54.6	38.24	47.72	46.49
	Lower	32.16	37.64	27.63	35.74	38.45	22.09	31.57	30.34
	Upper	64.46	69.94	59.93	68.04	70.75	54.39	63.87	62.64



750 **Figure 3: Global distribution of BME-adjusted surface ozone (two left columns) and the post adjusted-conversion difference (two right columns) shown for these satellites: (a) IASI+GOME2 2017-2020, (b) OMI/MLS 2005-2022. (c) CrIS 0-3km 2022. Results are shown as seasonal means for Northern Hemisphere winter (DJF) and summer (JJA).**



755 Figure 4: Bias of BME-adjusted surface ozone and the reduction in absolute bias after adjustment and conversion
 ("impact") against 1°×1° resolution re-gridded TOAR-II observations in the EUR, NAM, and EAS regions, for these
 satellites: (a) IASI+GOME2 2017-2020, (b) OMI/MLS 2005-2022, (c) CrIS 0-3km 2022. Red indicates positive values (i.e.,
 an overestimation in the "Bias" map, or an increased absolute bias in the "Impact" map), while blue indicates negative
 values.

760

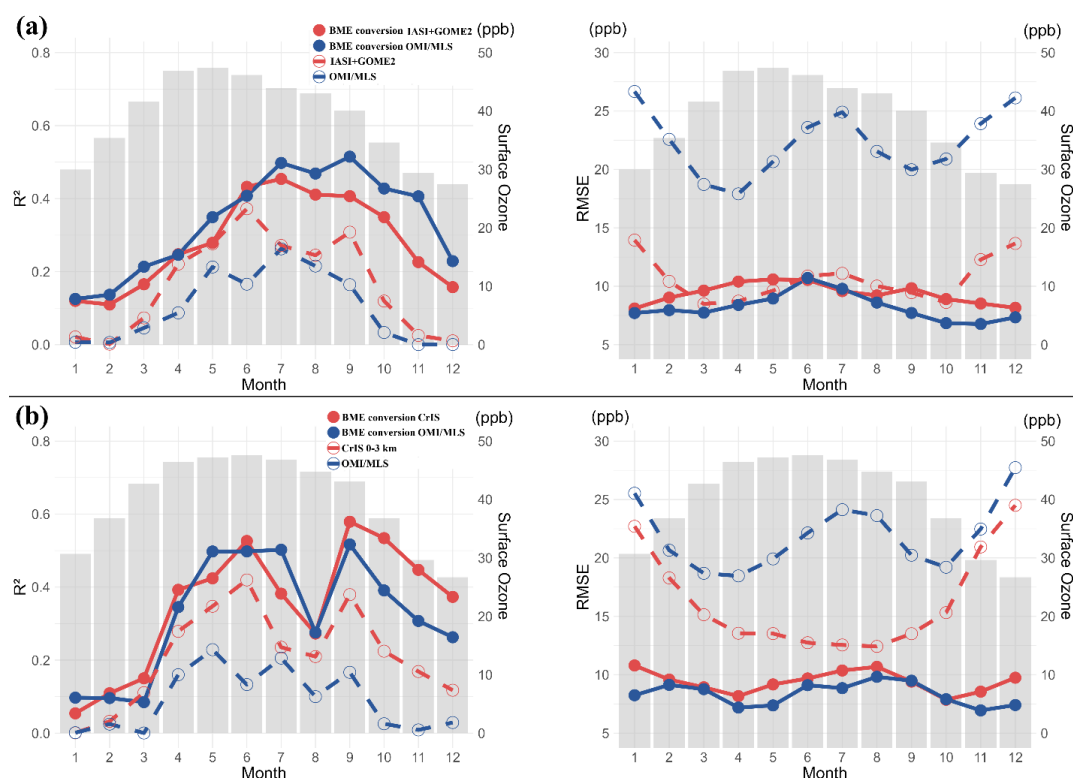
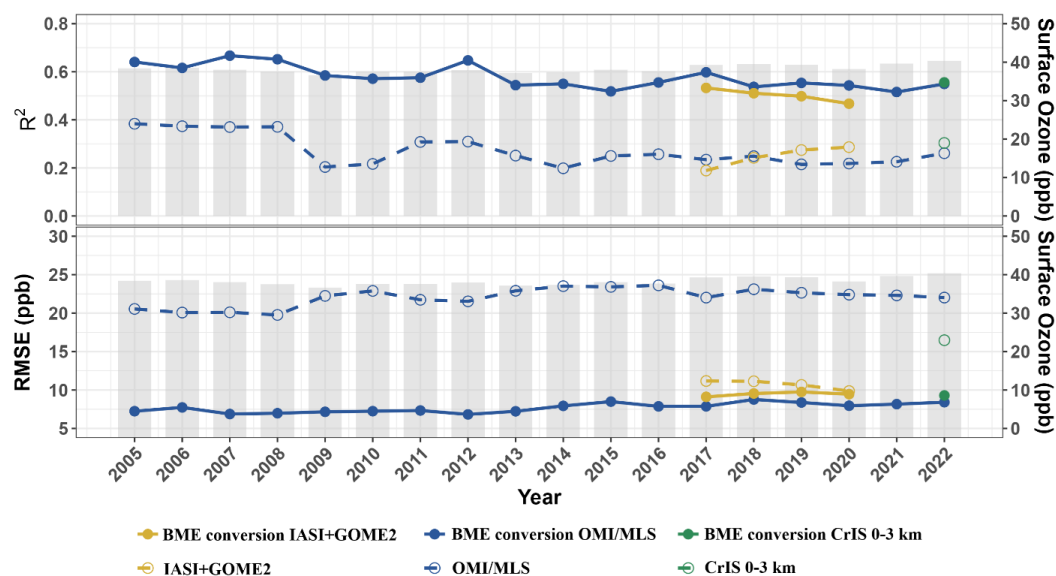
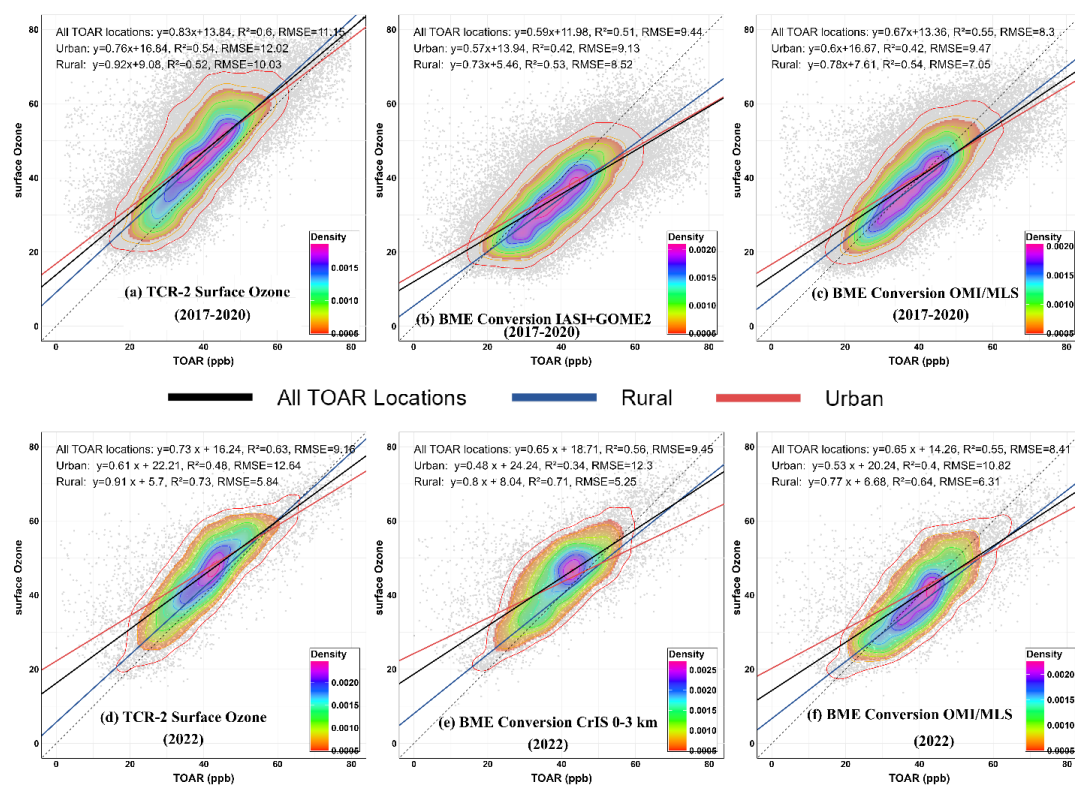


Figure 5: Monthly R^2 and RMSE for monthly average DMA8 surface ozone for original satellites and after BME-adjusted conversion, for all TOAR-II observations globally: (a) IASI+GOME2 and OMI/MLS 2017-2020, (b) CrIS 0-3km and OMI/MLS 2022. Background grey histograms represent the average surface ozone concentration over all TOAR-II observations.

765



770 **Figure 6: Time series of yearly R^2 (top) and RMSE (bottom) for BME-adjusted ozone estimates and original satellite data against TOAR-II DMA8 surface ozone (2005-2022). Dashed lines and open circles indicate baseline satellite data (OMI/MLS, IASI+GOME2, and CrIS), while solid lines and filled circles show the improved performance following the BME-adjusted conversion. Background grey histograms represent the average surface ozone concentration over all TOAR-II observations.**



775

Figure 7: Evaluation BME-adjusted surface ozone and TCR-2 surface ozone against TOAR-II DMA8 surface ozone: (a) TCR-2 2017-2020, (b) IASI+GOME2 2017-2020, (c) OMI/MLS 2017-2020, (d) TCR-2 2022, (e) CrIS 0-3km 2022, (f) OMI/MLS 2022. Stations are categorized into urban and rural sites based on the TOAR-II metadata.

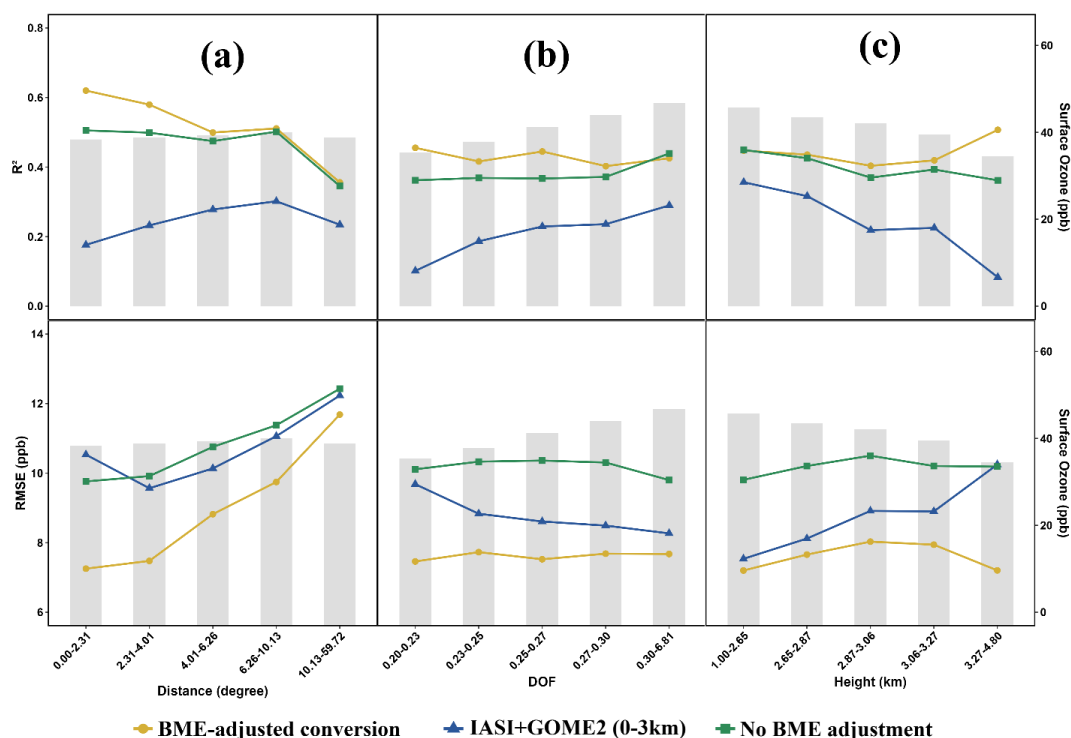


Figure 8: Evaluation of the BME-adjusted conversion method for IASI+GOME2 against TOAR-II monthly average DMA8 ozone observations (2017-2020) (a) model accuracy versus distance to nearest vertical profile location, (b) model accuracy versus DOF (surface to 3 km) for TOAR-II sites within 6° of the nearest vertical profile location, (c) model accuracy versus the height of maximum sensitivity for TOAR-II sites within 6° of the nearest vertical profile location. Yellow line with circle represents BME-adjusted surface ozone (final output), green line with square represents only using the TCR-2 ratio to derive surface ozone without the BME Joint-Adjustment, blue line with triangle represents original satellite observation from IASI+GOME2 0-3km. Background grey histograms represent the average surface ozone concentration from TOAR-II.

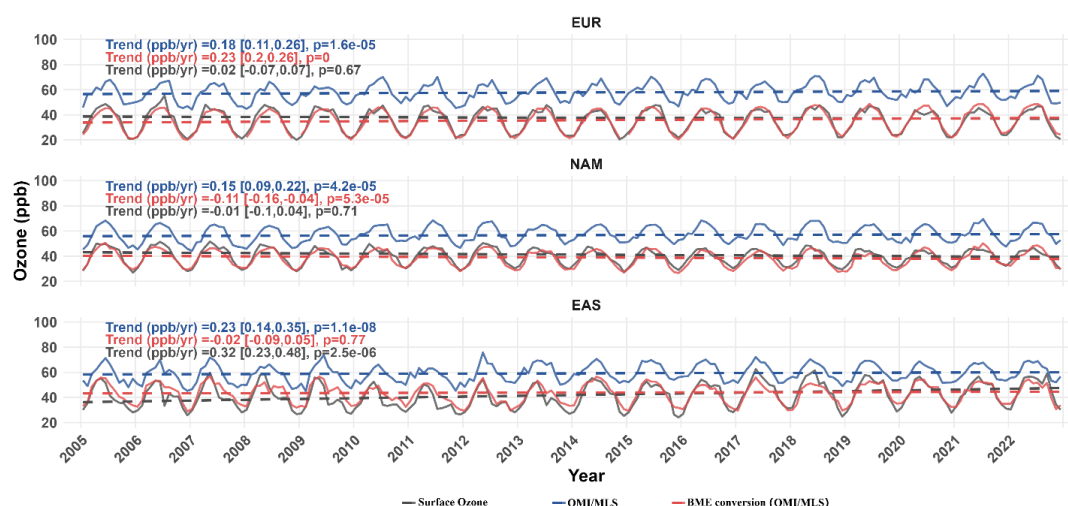


Figure 9: Deseasonalized trends in monthly average surface MDA8 ozone averaged over TOAR-II station locations in the EUR, NAM, and EAS regions from BME-adjusted conversion OMI/MLS data, original OMI/MLS column ozone, and TOAR-II observations (2005-2022). The deseasonalized trends (in ppb yr⁻¹) are calculated by median quantile regression after removing the seasonal effect, details are in Section S2.

Reference

Abbott, N. A., Semone, L., Strott, R., Dodda, P., Wang, C.-T., Green, J., Baek, B. H., Engel, L. S., Vizuete, W., and Serre, M. L.: Evaluation of Bayesian Maximum Entropy Data Fusion Approaches to Estimate Styrene, Benzene, Toluene, Ethylbenzene, and Xylenes and to Inform Epidemiological Analyses in the US Gulf States, *Environmental Science & Technology*, 59, 45-55, 10.1021/acs.est.4c05094, 2025.

Ancellet, G., Godin-Beekmann, S., Smit, H. G., Stauffer, R. M., Van Malderen, R., Bodichon, R., and Pazmiño, A.: Homogenization of the Observatoire de Haute Provence electrochemical concentration cell (ECC) ozonesonde data record: comparison with lidar and satellite observations, *Atmospheric Measurement Techniques*, 15, 3105-3120, 10.5194/amt-15-3105-2022, 2022.

Balmes, J. R.: Long-Term Exposure to Ozone and Small Airways: A Large Impact?, *American Journal of Respiratory and Critical Care Medicine*, 205, 384-385, 10.1164/rccm.202112-2733ED, 2022.

Becker, J. S., DeLang, M. N., Chang, K.-L., Serre, M. L., Cooper, O. R., Wang, H., Schultz, M. G., Schröder, S., Lu, X., Zhang, L., Deushi, M., Josse, B., Keller, C. A., Lamarque, J.-F., Lin, M., Liu, J., Marécal, V., Strode, S. A., Sudo, K., Tilmes, S., Zhang, L., Brauer, M., and West, J. J.: Using Regionalized Air Quality Model Performance and Bayesian Maximum Entropy data fusion to map global surface ozone concentration, *Elementa: Science of the Anthropocene*, 11, 10.1525/elementa.2022.00025, 2023.

Bowman, K. W.: TROPES CrIS-JPSS1 L2 Ozone for Forward Stream, Summary Product V1 [dataset], 10.5067/MHH8R0UZ5BMJ, 2022.

Brauer, M., Roth, G. A., Aravkin, A. Y., Zheng, P., Abate, K. H., Abate, Y. H., Abbafati, C., Abbasgholizadeh, R., Abbasi, M. A., and Abbasian, M.: Global burden and strength of evidence for 88 risk factors in 204 countries and 811 subnational locations, 1990–2021: a systematic analysis for the Global Burden of Disease Study 2021, *The Lancet*, 403, 2162-2203, 10.1016/S0140-6736(24)00933-4, 2024.



- 820 Chang, K.-L., Cooper, O. R., Gaudel, A., Petropavlovskikh, I., Effertz, P., Morris, G., and McDonald, B. C.: Challenges in detecting free tropospheric ozone trends in a sparsely sampled environment, *Atmospheric Chemistry and Physics*, 24, 6197-6218, 10.5194/acp-24-6197-2024, 2024.
- Chang, K. L., McDonald, B. C., Harkins, C., and Cooper, O. R.: Surface ozone trend variability across the United States and the impact of heat waves (1990–2023), *Atmos. Chem. Phys.*, 25, 5101-5132, 10.5194/acp-25-5101-2025, 2025.
- 825 Chang, K. L., Cooper, O. R., Gaudel, A., Petropavlovskikh, I., and Thouret, V.: Statistical regularization for trend detection: an integrated approach for detecting long-term trends from sparse tropospheric ozone profiles, *Atmos. Chem. Phys.*, 20, 9915-9938, 10.5194/acp-20-9915-2020, 2020.
- Christakos, G.: A Bayesian/maximum-entropy view to the spatial estimation problem, *Mathematical Geology*, 22, 763-777, 10.1007/BF00890661, 1990.
- 830 Christakos, G., Bogaert, P., and Serre, M.: *Temporal GIS: advanced functions for field-based applications*, Springer Science & Business Media 2002.
- Christakos, G., Kolovos, A., Serre, M. L., and Vukovich, F.: Total ozone mapping by integrating databases from remote sensing instruments and empirical models, *IEEE Transactions on Geoscience and Remote Sensing*, 42, 991-1008, 10.1109/TGRS.2003.822751, 2004.
- 835 Colombi, N., Miyazaki, K., Bowman, K. W., Neu, J. L., and Jacob, D. J.: A new methodology for inferring surface ozone from multispectral satellite measurements, *Environmental Research Letters*, 16, 105005, 10.1088/1748-9326/ac243d, 2021.
- Cuesta, J., Kanaya, Y., Takigawa, M., Dufour, G., Eremenko, M., Foret, G., Miyazaki, K., and Beekmann, M.: Transboundary ozone pollution across East Asia: daily evolution and photochemical production analysed by IASI+ GOME2 multispectral satellite observations and models, *Atmospheric Chemistry and Physics*, 18, 9499-9525, 10.5194/acp-18-9499-2018, 2018.
- 840 Cuesta, J., Costantino, L., Beekmann, M., Siour, G., Menut, L., Bessagnet, B., Landi, T. C., Dufour, G., and Eremenko, M.: Ozone pollution during the COVID-19 lockdown in the spring of 2020 over Europe, analysed from satellite observations, in situ measurements, and models, *Atmospheric Chemistry and Physics*, 22, 4471-4489, 10.5194/acp-22-4471-2022, 2022.
- 845 Cuesta, J., Eremenko, M., Liu, X., Dufour, G., Cai, Z., Höpfner, M., Von Clarmann, T., Sellitto, P., Foret, G., Gaubert, B., Beekmann, M., Orphal, J., Chance, K., Spurr, R., and Flaud, J. M.: Satellite observation of lowermost tropospheric ozone by multispectral synergism of IASI thermal infrared and GOME-2 ultraviolet measurements over Europe, *Atmospheric Chemistry and Physics*, 13, 9675-9693, 10.5194/acp-13-9675-2013, 2013.
- 850 DeLang, M. N., Becker, J. S., Chang, K. L., Serre, M. L., Cooper, O. R., Schultz, M. G., Schroder, S., Lu, X., Zhang, L., Deushi, M., Josse, B., Keller, C. A., Lamarque, J. F., Lin, M., Liu, J., Marecal, V., Strode, S. A., Sudo, K., Tilmes, S., Zhang, L., Cleland, S. E., Collins, E. L., Brauer, M., and West, J. J.: Mapping Yearly Fine Resolution Global Surface Ozone through the Bayesian Maximum Entropy Data Fusion of Observations and Model Output for 1990-2017, *Environ Sci Technol*, 55, 4389-4398, 10.1021/acs.est.0c07742, 2021.
- 855 Elshorbany, Y., Ziemke, J. R., Strode, S., Petetin, H., Miyazaki, K., De Smedt, I., Pickering, K., Seguel, R. J., Worden, H., and Emmerichs, T.: Tropospheric ozone precursors: global and regional distributions, trends, and variability, *Atmospheric Chemistry and Physics*, 24, 12225-12257, 10.5194/acp-24-12225-2024, 2024.
- 860 Fadnavis, S., Elshorbany, Y., Ziemke, J., Barret, B., Rap, A., Chandran, P. S., Pope, R. J., Sagar, V., Taraborrelli, D., and Le Flochmoen, E.: Influence of nitrogen oxides and volatile organic compounds emission changes on tropospheric ozone variability, trends and radiative effect, *Atmospheric Chemistry and Physics*, 25, 8229-8254, 10.5194/acp-25-8229-2025, 2025.



- Fleming, Z. L., Doherty, R. M., von Schneidmesser, E., Malley, C. S., Cooper, O. R., Pinto, J. P., Colette, A., Xu, X., Simpson, D., Schultz, M. G., Lefohn, A. S., Hamad, S., Moolla, R., Solberg, S., and Feng, Z.: Tropospheric Ozone Assessment Report: Present-day ozone distribution and trends relevant to human health, *Elementa: Science of the Anthropocene*, 6, 10.1525/elementa.273, 2018.
- 865 Fu, D., Kulawik, S. S., Miyazaki, K., Bowman, K. W., Worden, J. R., Eldering, A., Livesey, N. J., Teixeira, J., Irion, F. W., and Herman, R. L.: Retrievals of tropospheric ozone profiles from the synergism of AIRS and OMI: methodology and validation, *Atmospheric Measurement Techniques*, 11, 5587-5605, 10.5194/amt-11-5587-2018, 2018.
- 870 Gaudel, A., Cooper, O. R., Ancellet, G., Barret, B., Boynard, A., Burrows, J. P., Clerbaux, C., Coheur, P.-F., Cuesta, J., and Cuevas, E.: Tropospheric Ozone Assessment Report: Present-day distribution and trends of tropospheric ozone relevant to climate and global atmospheric chemistry model evaluation, *Elem Sci Anth*, 6, 39, 2018.
- Gaudel, A., Cooper, O. R., Chang, K.-L., Bourgeois, I., Ziemke, J. R., Strode, S. A., Oman, L. D., Sellitto, P., Nédélec, P., Blot, R., Thouret, V., and Granier, C.: Aircraft observations since the 1990s reveal increases of tropospheric ozone at multiple locations across the Northern Hemisphere, *Science Advances*, 6, eaba8272, doi:10.1126/sciadv.aba8272, 875 2020.
- Gaudel, A., Bourgeois, I., Li, M., Chang, K. L., Ziemke, J., Sauvage, B., Stauffer, R. M., Thompson, A. M., Kollonige, D. E., Smith, N., Hubert, D., Keppens, A., Cuesta, J., Heue, K. P., Veeffkind, P., Aikin, K., Peischl, J., Thompson, C. R., Ryerson, T. B., Frost, G. J., McDonald, B. C., and Cooper, O. R.: Tropical tropospheric ozone distribution and trends from in situ and satellite data, *Atmos. Chem. Phys.*, 24, 9975-10000, 10.5194/acp-24-9975-2024, 2024.
- 880 Hickman, S. H., Kelp, M. M., Griffiths, P. T., Doerksen, K., Miyazaki, K., Pennington, E. A., Koren, G., Iglesias-Suarez, F., Schultz, M. G., and Chang, K.-L.: Applications of machine learning and artificial intelligence in tropospheric ozone research, *Geoscientific Model Development*, 18, 8777-8800, 10.5194/gmd-18-8777-2025, 2025.
- Inness, A., Ades, M., Agustí-Panareda, A., Barré, J., Benedictow, A., Blechschmidt, A. M., Dominguez, J. J., Engelen, R., Eskes, H., Flemming, J., Huijnen, V., Jones, L., Kipling, Z., Massart, S., Parrington, M., Peuch, V. H., Razinger, M., Remy, S., Schulz, M., and Suttie, M.: The CAMS reanalysis of atmospheric composition, *Atmos. Chem. Phys.*, 885 19, 3515-3556, 10.5194/acp-19-3515-2019, 2019.
- Koffi, L. B., Dentener, F., Janssens-Maenhout, G., Guizzardi, D., Crippa, M., Diehl, T., Galmarini, S., and Solazzo, E.: Hemispheric Transport of Air Pollution (HTAP): Specification of the HTAP2 experiments: Ensuring harmonized modelling, Publications Office of the European Union, 10.2788/725244, 2016.
- 890 Li, K., Tan, R., Qiao, W., Lee, T., Wang, Y., Zhang, D., Tang, M., Zhao, W., Gu, Y., and Fan, S.: Surface and tropospheric ozone over East Asia and Southeast Asia from observations: distributions, trends, and variability, *Atmospheric Chemistry and Physics*, 25, 11575-11596, 10.5194/acp-25-11575-2025, 2025a.
- Li, S., Song, G., Xing, J., Dong, J., Zhang, M., Fan, C., Meng, S., Yang, J., Dong, L., and Gong, W.: Unraveling overestimated exposure risks through hourly ozone retrievals from next-generation geostationary satellites, *Nature Communications*, 16, 3364, 10.1038/s41467-025-58652-2, 2025b. 895
- Li, T. and Cheng, X.: Estimating daily full-coverage surface ozone concentration using satellite observations and a spatiotemporally embedded deep learning approach, *International Journal of Applied Earth Observation and Geoinformation*, 101, 102356, 10.1016/j.jag.2021.102356, 2021.
- 900 Liu, X., Zhu, Y., Xue, L., Desai, A. R., and Wang, H.: Cluster-Enhanced Ensemble Learning for Mapping Global Monthly Surface Ozone From 2003 to 2019, *Geophysical Research Letters*, 49, e2022GL097947, 10.1029/2022GL097947, 2022.
- Lu, X., Liu, Y., Su, J., Weng, X., Ansari, T., Zhang, Y., He, G., Zhu, Y., Wang, H., and Zeng, G.: Tropospheric ozone trends and attributions over East and Southeast Asia in 1995–2019: an integrated assessment using statistical methods,



- 905 machine learning models, and multiple chemical transport models, *Atmospheric Chemistry and Physics*, 25, 7991-8028, 10.5194/acp-25-7991-2025, 2025.
- Malings, C., Knowland, K. E., Pavlovic, N., Coughlin, J. G., King, D., Keller, C., Cohn, S., and Martin, R. V.: Air quality estimation and forecasting via data fusion with uncertainty quantification: theoretical framework and preliminary results, *Journal of Geophysical Research: Machine Learning and Computation*, 1, e2024JH000183, 10.1029/2024JH000183, 2024.
- 910 Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S. L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., and Gomis, M.: Climate change 2021: the physical science basis, Contribution of working group I to the sixth assessment report of the intergovernmental panel on climate change, 2, 2391, 2021.
- Mettig, N., Weber, M., Rozanov, A., Burrows, J. P., Veefkind, P., Thompson, A. M., Stauffer, R. M., Leblanc, T., Ancellet, G., and Newchurch, M. J.: Combined UV and IR ozone profile retrieval from TROPOMI and CrIS measurements, *Atmospheric Measurement Techniques*, 15, 2955-2978, 10.5194/amt-15-2955-2022, 2022.
- 915 Mills, G., Pleijel, H., Malley, C. S., Sinha, B., Cooper, O. R., Schultz, M. G., Neufeld, H. S., Simpson, D., Sharps, K., Feng, Z., Gerosa, G., Harmens, H., Kobayashi, K., Saxena, P., Paoletti, E., Sinha, V., and Xu, X.: Tropospheric Ozone Assessment Report: Present-day tropospheric ozone distribution and trends relevant to vegetation, *Elementa: Science of the Anthropocene*, 6, 10.1525/elementa.302, 2018.
- 920 Miyazaki, K.: TROPESS Chemical Reanalysis Surface O₃ 2-Hourly 2-dimensional Product V1, Goddard Earth Sciences Data and Information Services Center (GES DISC) [dataset], 10.5067/NN87W53OVGUS, 2024.
- Miyazaki, K., Bowman, K. W., Yumimoto, K., Walker, T., and Sudo, K.: Evaluation of a multi-model, multi-constituent assimilation framework for tropospheric chemical reanalysis, *Atmospheric Chemistry and Physics*, 20, 931-967, 10.5194/acp-20-931-2020, 2020a.
- 925 Miyazaki, K., Bowman, K., Sekiya, T., Eskes, H., Boersma, F., Worden, H., Livesey, N., Payne, V. H., Sudo, K., Kanaya, Y., Takigawa, M., and Ogochi, K.: Updated tropospheric chemistry reanalysis and emission estimates, TCR-2, for 2005–2018, *Earth Syst. Sci. Data*, 12, 2223-2259, 10.5194/essd-12-2223-2020, 2020b.
- Pennington, E. A., Osterman, G. B., Payne, V. H., Miyazaki, K., Bowman, K. W., and Neu, J. L.: Quantifying biases in TROPESS AIRS, CrIS, and joint AIRS+ OMI tropospheric ozone products using ozonesondes, *Atmospheric Chemistry and Physics*, 25, 8533-8552, 10.5194/acp-25-8533-2025, 2025.
- 930 Petetin, H., Thouret, V., Fontaine, A., Sauvage, B., Athier, G., Blot, R., Boulanger, D., Cousin, J.-M., and Nédélec, P.: Characterising tropospheric O₃ and CO around Frankfurt over the period 1994–2012 based on MOZAIC–IAGOS aircraft measurements, *Atmospheric Chemistry and Physics*, 16, 15147-15163, 10.5194/acp-16-15147-2016, 2016.
- Schröder, S., Schultz, M. G., Selke, N., Sun, J., Ahring, J., Mozaffari, A., Romberg, M., Epp, E., Lensing, M., and Apweiler, S.: TOAR Data Infrastructure, Version v1. 0, FZ-Juelich B2SHARE [data set], 10, 10.34730/4D9A287DEC0B42F1AA6D244DE8F19EB3, 2021.
- 935 Schultz, M. G., Schröder, S., Lyapina, O., Cooper, O. R., Galbally, I., Petropavlovskikh, I., Von Schneidmesser, E., Tanimoto, H., Elshorbany, Y., and Naja, M.: Tropospheric Ozone Assessment Report: Database and metrics data of global surface ozone observations, *Elem Sci Anth*, 5, 58, 10.1525/elementa.244, 2017.
- 940 Seguel, R. J., Castillo, L., Opazo, C., Rojas, N. Y., Nogueira, T., Cazorla, M., Gavidia-Calderón, M., Gallardo, L., Garreaud, R., and Carrasco-Escaff, T.: Changes in South American surface ozone trends: exploring the influences of precursors and extreme events, *Atmospheric Chemistry and Physics*, 24, 8225-8242, 10.5194/acp-24-8225-2024, 2024.
- 945 Sekiya, T., Emili, E., Miyazaki, K., Inness, A., Qu, Z., Pierce, R. B., Jones, D., Worden, H., Cheng, W. Y., and Huijnen, V.: Assessing the relative impacts of satellite ozone and its precursor observations to improve global



- tropospheric ozone analysis using multiple chemical reanalysis systems, *Atmospheric Chemistry and Physics*, 25, 2243–2268, 10.5194/acp-25-2243-2025, 2025.
- Serre, M. L. and Christakos, G.: Modern geostatistics: Computational BME analysis in the light of uncertain physical knowledge—the Equus Beds study, *Stochastic Environmental Research and Risk Assessment*, 13, 1–26, 10.1007/s004770050029, 1999.
- Smit, H. and Thompson, A.: GAW Report No. 268. Ozonesonde Measurement Principles and Best Operational Practices ASOPOS 2.0 (Assessment of Standard Operating Procedures for Ozonesondes), 2021.
- Sun, H., Shin, Y. M., Xia, M., Ke, S., Wan, M., Yuan, L., Guo, Y., and Archibald, A. T.: Spatial Resolved Surface Ozone with Urban and Rural Differentiation during 1990–2019: A Space–Time Bayesian Neural Network Downscaler, *Environmental Science & Technology*, 56, 7337–7349, 10.1021/acs.est.1c04797, 2022.
- Van Malderen, R., Thompson, A. M., Kollonige, D. E., Stauffer, R. M., Smit, H. G., Maillard Barras, E., Vigouroux, C., Petropavlovskikh, I., Leblanc, T., and Thouret, V.: Global ground-based tropospheric ozone measurements: reference data and individual site trends (2000–2022) from the TOAR-II/HEGIFTOM project, *Atmospheric Chemistry and Physics*, 25, 7187–7225, 10.5194/acp-25-7187-2025, 2025.
- Wang, H., Lu, X., Jacob, D. J., Cooper, O. R., Chang, K.-L., Li, K., Gao, M., Liu, Y., Sheng, B., and Wu, K.: Global tropospheric ozone trends, attributions, and radiative impacts in 1995–2017: an integrated analysis using aircraft (IAGOS) observations, ozonesonde, and multi-decadal chemical model simulations, *Atmospheric Chemistry and Physics Discussions*, 2022, 1–55, 10.5194/acp-22-13753-2022, 2022.
- Wang, H., Miyazaki, K., Sun, H. Z., Qu, Z., Liu, X., Inness, A., Schultz, M., Schröder, S., Serre, M., and West, J. J.: Intercomparison of global ground-level ozone datasets for health-relevant metrics, *Atmos. Chem. Phys.*, 25, 15969–15990, 10.5194/acp-25-15969-2025, 2025.
- Wei, J., Li, Z., Li, K., Dickerson, R. R., Pinker, R. T., Wang, J., Liu, X., Sun, L., Xue, W., and Cribb, M.: Full-coverage mapping and spatiotemporal variations of ground-level ozone (O₃) pollution from 2013 to 2020 across China, *Remote Sensing of Environment*, 270, 112775, 10.1016/j.rse.2021.112775, 2022.
- Xue, T., Zheng, Y., Geng, G., Xiao, Q., Meng, X., Wang, M., Li, X., Wu, N., Zhang, Q., and Zhu, T.: Estimating Spatiotemporal Variation in Ambient Ozone Exposure during 2013–2017 Using a Data-Fusion Model, *Environmental Science & Technology*, 54, 14877–14888, 10.1021/acs.est.0c03098, 2020.
- Yang, Z., Li, Z., Cheng, F., Lv, Q., Li, K., Zhang, T., Zhou, Y., Zhao, B., Xue, W., and Wei, J.: Two-decade surface ozone (O₃) pollution in China: Enhanced fine-scale estimations and environmental health implications, *Remote Sensing of Environment*, 317, 114459, 10.1016/j.rse.2024.114459, 2025.
- Zhan, C., Xie, M., Lu, H., Liu, B., Wu, Z., Wang, T., Zhuang, B., Li, M., and Li, S.: Impacts of urbanization on air quality and the related health risks in a city with complex terrain, *Atmos. Chem. Phys.*, 23, 771–788, 10.5194/acp-23-771-2023, 2023.
- Zhang, X., Zhang, Y., Lu, X., Bai, L., Chen, L., Tao, J., Wang, Z., and Zhu, L.: Estimation of lower-stratosphere-to-troposphere ozone profile using long short-term memory (LSTM), *Remote Sensing*, 13, 1374, 10.3390/rs13071374, 2021.
- Ziemke, J., Chandra, S., Duncan, B., Froidevaux, L., Bhartia, P., Levelt, P., and Waters, J.: Tropospheric ozone determined from Aura OMI and MLS: Evaluation of measurements and comparison with the Global Modeling Initiative's Chemical Transport Model, *Journal of Geophysical Research: Atmospheres*, 111, 10.1029/2006JD007089, 2006.
- Ziemke, J. R., Oman, L. D., Strode, S. A., Douglass, A. R., Olsen, M. A., McPeters, R. D., Bhartia, P. K., Froidevaux, L., Labow, G. J., Witte, J. C., Thompson, A. M., Haffner, D. P., Kramarova, N. A., Frith, S. M., Huang, L. K., Jaross,

<https://doi.org/10.5194/egusphere-2026-2812>

Preprint. Discussion started: 12 June 2026

© Author(s) 2026. CC BY 4.0 License.



990 G. R., Seftor, C. J., Deland, M. T., and Taylor, S. L.: Trends in global tropospheric ozone inferred from a composite record of TOMS/OMI/MLS/OMPS satellite measurements and the MERRA-2 GMI simulation, Atmos. Chem. Phys., 19, 3257-3269, 10.5194/acp-19-3257-2019, 2019.