Explicit aerosol correction of OMI formaldehyde retrievals

Yeonjin Jung[1], Gonzalo González Abad[1], Caroline R. Nowlan[1], Kelly Chance[1], Xiong Liu[1], Omar Torres[2], and Changwoo Ahn[3]

[1]Atmoic and Molecular Physics Division, Harvard-Smithsonian Center for Astrophysics, Cambridge, MA, USA

[2]Atmospheric Chemistry Dynamics Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD, USA

[3]Science Systems and Applications, Inc., Lanham, MD, USA

Key Points:

- We develop a measurement-based methodology to perform explicit aerosol correction for OMI HCHO retrievals using OMI aerosol measurements
- We quantify the impact of aerosols on the OMI HCHO retrievals
- Global mean aerosol-induced HCHO VCD differences are 27 %, 6 %, and -0.3 % for smoke, dust, and sulfate aerosols, respectively.

Corresponding author: Yeonjin Jung, `yeonjin.jung@cfa.harvard.edu`

Abstract

Atmospheric aerosols are significant sources of uncertainty in air mass factor (AMF) calculations for trace gas retrievals using ultraviolet measurements from space. Current trace gas retrievals typically do not consider aerosols explicitly as cloud products partially account for aerosol effects. Here, we propose a new measurement-based approach to correct for aerosols explicitly in the AMF calculation, apply it to Ozone Monitoring Instrument (OMI) formaldehyde (HCHO) retrievals and quantify the aerosol-induced HCHO vertical column density (VCD) difference for three aerosol types (smoke, dust, and sulfate) during 2006–2007. We use OMI aerosol retrievals for aerosol optical properties and vertical profiles to construct look-up-tables of scattering weights as functions of geometry, surface pressure, surface albedo, and aerosol information. The average difference between the NASA operational OMI HCHO product (not considering aerosols) and the results obtained in this study on a global scale are 27 %, 6 %, and -0.3 % for smoke, dust, and sulfate aerosols, respectively. The region with the largest aerosol effects is East China, where the explicit smoke aerosol correction enhances mean HCHO VCDs by 35 %, with corrections to individual observations sometimes larger than 100 %. The quantified aerosol effects are applicable under clear-sky conditions. This study highlights the need to implement aerosol corrections in the AMF calculation for HCHO retrievals. This is particularly relevant in regions with high levels of pollution where aerosols interfere the most with formaldehyde satellite observations.

1 Introduction

Formaldehyde (HCHO) is one of the most abundant hydrocarbons in the troposphere, and is mainly produced by the oxidation of volatile organic compounds (VOCs) from biogenic and anthropogenic sources (Houweling et al., 1998). As the oxidation of short-lived non-methane VOCs (NMVOCs) results in enhancements of total HCHO column over the continents, satellite measurements of HCHO can provide useful constraints on VOC emissions (Barkley et al., 2013; Marais et al., 2012; Palmer et al., 2003).

Space-based HCHO observations began with the launch of the Global Ozone Monitoring Experiment (GOME) on-board the European Remote sensing Satellite (ERS)-2, using backscattered solar radiation in the ultraviolet spectral region (Chance et al., 2000; Palmer et al., 2001), and have been continued by follow-up missions such as the SCanning Imaging Absorption SpectroMeter for Atmospheric CHartographY (SCIAMACHY) (De Smedt et al., 2008; Wittrock et al., 2006), the Ozone Monitoring Instrument (OMI) (De Smedt et al., 2015; González Abad et al., 2015), GOME-2 (De Smedt et al., 2008, 2012), the Ozone Mapping Profiler Suite (OMPS) (González Abad et al., 2016; Li et al., 2015), and the TROPOspheric Monitoring Instrument (TROPOMI) (De Smedt et al., 2018). These instruments on sun-synchronous satellites measure a given location once per day or less. To extend the number of observations per day, a new constellation of geostationary satellites is planned for the near future including the Tropospheric Emissions: Monitoring of Pollution (TEMPO) mission over North America (Zoogman et al., 2017), the Sentinel-4 mission over Europe (Ingmann et al., 2012) and the Geostationary Environment Monitoring Spectrometer (GEMS) over Asia (Kim et al., n.d.).

The air mass factor (AMF) is one of the major error sources in the retrieval of trace gas columns (Boersma et al., 2004; Martin et al., 2002). It depends primarily on the satellite viewing geometry, surface reflectance, vertical distribution of the trace gas of interest and aerosol characteristics (Palmer et al., 2001). Martin et al. (2002) found that the total error (53 %) of tropospheric NO_2 column retrievals has contributions from uncertainties in surface reflectivity (28 %), NO_2 profiles (15 %), aerosol properties (30 %), clouds (16 %) and radiative transfer assumptions (10 %).

Aerosols are a significant source of uncertainty in AMF calculations because they modify the path of the photons reaching the satellite and therefore the observed radiances. Current HCHO retrieval algorithms consider aerosol effects implicitly in their cloud correc-

tion strategies as cloud retrieval algorithms account for a large part of the aerosol effects by retrieving modified cloud fraction and pressure (Boersma et al., 2004, 2011). However, the implicit inclusion of aerosol effects in the cloud correction has several limitations, including the inability to account for strongly absorbing aerosols (Castellanos et al., 2015; Chimot et al., 2016). Martin et al. (2003) found that strongly absorbing aerosols reduced the AMF by up to 40 % over biomass-burning regions, while scattering aerosols increased the AMF by 5-10 % in the retrieval of tropospheric NO_2 columns from the GOME instrument by using monthly aerosol properties derived from the Global Ozone Chemistry Aerosol Radiation and Transport (GOCART) model and aerosol vertical distribution from GEOS-Chem simulations. An accurate characterization of aerosol information is needed to estimate the impact of aerosols on trace gas retrievals, as the effect of aerosols on AMF calculations depends on aerosol optical properties, and the amount and vertical distribution of aerosols (Leitão et al., 2010).

Recently, there has been an effort to examine the effects of aerosols explicitly on the retrievals. Using aerosol information from nested GEOS-Chem simulations over China, Lin et al. (2014) found that excluding aerosols from the retrieval process can change NO_2 vertical column density (VCD) by -70 to 90 % with an average of 14 % when aerosol optical depth (AOD) is greater than 0.8. The AODs in the simulations were constrained by the Aerosol Robotic Network (AERONET), the MODerate resolution Imaging Spectroradiometer (MODIS) and Multi-Axis Differential Optical Absorption Spectroscopy (MAX-DOAS) observations. The Cloud-Aerosol LIdar with Orthogonal Polarization (CALIOP) was used to constrain aerosol vertical profiles. In addition, Lin et al. (2015) showed that the implicit aerosol correction enhances annual mean NO_2 VCDs by 15–40 % over eastern China. They also found that the implicit aerosol correction tends to exclude highly polluted conditions mistakenly classified as cloudy pixels by cloud algorithms. M. Liu et al. (2019) improved the explicit aerosol correction in a second version of their algorithm (Lin et al., 2015, 2014) by constructing a monthly climatology of vertical aerosol profiles from CALIOP to constraint GEOS-Chem aerosol profiles, showing that retrieved NO_2 VCDs increase by 4–16 % over China with respect to the original algorithm.

Castellanos et al. (2015) accounted for the effects of biomass burning aerosols in the NO_2 AMF calculation during South American biomass burning seasons by including co-located aerosol vertical profile observations from CALIOP, and AOD and single scattering albedo (SSA) retrieved by the operational OMI near-UV aerosol retrieval product (OMAERUV).

They found the AMFs of observations in this region can increase by up to a factor of 2 when aerosols are considered explicitly. Focusing on HCHO retrievals, Kwon et al. (2017) used GEOS-Chem model simulations to show that including aerosols over East Asia changed the HCHO VCDs by -25 to 32 %. Lorente et al. (2017) found that the use of different aerosol corrections from different retrieval groups can result in AMF differences as large as 50 % under high aerosol loading conditions.

In this study, we propose a new methodology that applies explicit aerosol corrections to trace gas retrievals based on satellite aerosol measurements from the same instrument. To describe and evaluate aerosol impacts on HCHO AMF calculations globally, we use aerosol loading and optical properties from the OMI OMAERUV aerosol product. First, we analyze the sensitivity of OMI HCHO AMFs to AOD, SSA and aerosol layer height (ALH). Then, we apply the results of the sensitivity tests to construct look-up tables (LUTs) of scattering weights, compute aerosol correction factors for OMI HCHO products, and evaluate the impact of aerosols on OMI HCHO AMFs and VCDs quantitatively.

Section 2 describes the OMI instrument and the OMI HCHO and aerosol data products. Section 3 explains the methodology of the explicit aerosol correction on trace gas retrievals. Section 4 examines the sensitivity of OMI HCHO AMFs to aerosol information. Section 5 evaluates the aerosol-induced HCHO AMF and VCD differences by applying the aerosol correction factor to OMI measurements and discusses the uncertainties in the explicit aerosol correction. Section 6 summarizes and concludes this study.

2 Data

2.1 The OMI instrument

OMI is a nadir-viewing hyper-spectral spectrometer on board the National Aeronautics and Space Administration (NASA) Earth Observing System (NASA EOS) Aura satellite, which was launched on 15 July 2004 (Levelt et al., 2006). OMI measures solar back-scattered radiation in the ultraviolet and visible region in three channels (UV1: 270–310 nm, UV2: 310–365 nm, visible: 350–500 nm). Aura has an equator-crossing time of 13:42 in the ascending node. Sixty observations are collected across a 2600-km wide swath with a nominal spatial resolution on the ground of 13 $\times$ 24 km^2 at nadir, increasing to 26 $\times$ 128 km^2 at the edges of the swath. Until 2008, OMI provided daily global coverage. After 2008, OMI has been affected by the row anomaly which reduced its spatial coverage so that it achieves global coverage every 2 days (Torres et al., 2018).

2.2 OMI operational HCHO retrieval (OMHCHO)

The operational OMI HCHO retrieval algorithm follows a two-step approach to determine HCHO VCDs. The algorithm is described in detail in González Abad et al. (2015). The OMI radiances are first fitted to retrieve slant column density (SCD) in a window extending from 328.5 to 356.5 nm using a direct spectral fitting method (Chance, 1998). The SCDs are then converted to VCDs by applying AMFs (Palmer et al., 2001) calculated using pre-computed LUTs of scattering weights determined as a function of the viewing geometry, surface albedo, surface pressure, and cloud properties. The OMHCHO retrievals use the Lambertian equivalent reflectance (LER) from the OMI surface reflectance climatology (Kleipool et al., 2008) and cloud information derived from the OMI cloud products (OMCLDO2) (Acarreta et al., 2004) to derive the scattering weights. The a priori HCHO profile is obtained from a monthly climatology created using simulations performed with the GEOS-Chem chemical transport model (Bey et al., 2001) at a spatial resolution of 2°-latitude $\times$2.5°-longitude for 2007. The model results for each month between 13–14 local time (LT) are averaged to derive the climatology values at 47 vertical levels extending from the surface up to 0.01 hPa (González Abad et al., 2015). The uncertainties in the spectral fitting are estimated to be 45–100 %, and are usually reduced to 30 % over high HCHO regions. Uncertainties in the AMF calculation, not considering aerosols, are approximately 35 %, mainly due to uncertainties in the cloud parameters, surface albedo, and vertical

HCHO profiles. This results in typical total uncertainties in the HCHO VCD of 45–105 % (González Abad et al., 2015). In this study, OMHCHO retrievals are used with SZA less than 70°.

2.3 OMI operational aerosol retrieval (OMAERUV)

The operational OMAERUV product uses radiances at 354 and 388 nm to retrieve AOD and SSA for a selected aerosol type and aerosol vertical distribution over cloud-free scenes (Torres et al., 2013, 2007). The AOD and SSA retrievals are also provided at 354 nm, 388 nm, and 500 nm.

The OMAERUV retrieval algorithm assumes three major aerosol types: 1) carbonaceous aerosols associated with biomass burning, 2) desert dust, and 3) weakly absorbing sulfate-based aerosols (hereafter smoke, dust, and sulfate). Each aerosol type has an assumed particle size distribution (PSD) determined from AERONET inversion products and a real part of the refractive index obtained from the Optical Properties of Aerosols and Clouds (OPAC) database (Hess et al., 1998; Jethva & Torres, 2012).

Final retrieved AOD and SSA are determined by interpolation at the best-guess ALH from the LUT radiative transfer calculation at the ALH nodal points. For dust and smoke aerosol types, the best-guess ALH is obtained from a CALIOP-based climatology calculated as the attenuated-backscatter-weighted mean height during July 2006 to December 2008 (Torres et al., 2013). When climatological ALH information is not available, the height of the absorbing aerosol layer is prescribed as follows. For smoke aerosol layers between 30°N and 30°S are assumed to be centered at 3 km, while mid- and high- latitude ($\geq$45°) aerosols are assumed to be centered at 6 km. Over latitudes of 30–45°, smoke ALHs are interpolated with latitude between 3 and 6 km. For dust aerosol layers ALH is obtained from a multi-year climatology of GOCART simulations (Ginoux et al., 2001). Sulfate aerosol concentrations are assumed to decrease exponentially with increasing height, with the largest values at the surface. Additional details of the OMAERUV retrieval algorithm can be found in Torres et al. (2013, 2007).

The accuracy of the OMAERUV AODs is estimated to be the larger value of 30 % or 0.1 with 65 % of the retrievals within the expected uncertainty (Ahn et al., 2014). Approximately 46 % (69 %) of the retrieved OMAERUV SSAs agree with AERONET SSAs within OMAERUV expected uncertainties of 0.03 (0.05) (Jethva et al., 2014).

The OMAERUV product provides a simplified quality flag for each observation. In this study, we examine data flagged as 0 and 1. A flag of 0 indicates the pixel has a reliable retrieval for AOD, absorption aerosol optical depth (AAOD), and SSA with minimal cloud contamination. A flag of 1 means that only AAOD is reliable, with possible cloud contamination. Coincident OMAERUV and OMHCHO retrievals are available at 10% of the total number of OMI observations, of which approximately 35 % have quality flag of 0. Of these, 13 % are smoke, 22 % are dust, and 65 % are sulfate aerosols.

Figure 1 shows a histogram of AODs and SSAs at 340 nm, and ALHs from the OMAERUV products for the period of 2006–2007 with box-whisker plots showing the range for each aerosol type. The data provided in this figure are calculated at 340 nm (see Section 3.2) corresponding to the HCHO wavelength fitting window. As seen in this figure, AODs are mostly in the range from 0.5 to 2.5 with a maximum of about 6.0. The ranges of SSAs are dependent on the aerosol types. Smoke aerosol has the lowest median value followed by dust and then sulfate aerosols. The distribution of ALHs shows three distinct behaviors. The high frequency near 0 km is driven by sulfate aerosols, a second peak at 1 km is linked to dust and the broad distribution between 2 km and 4 km is due to both dust and smoke aerosols.

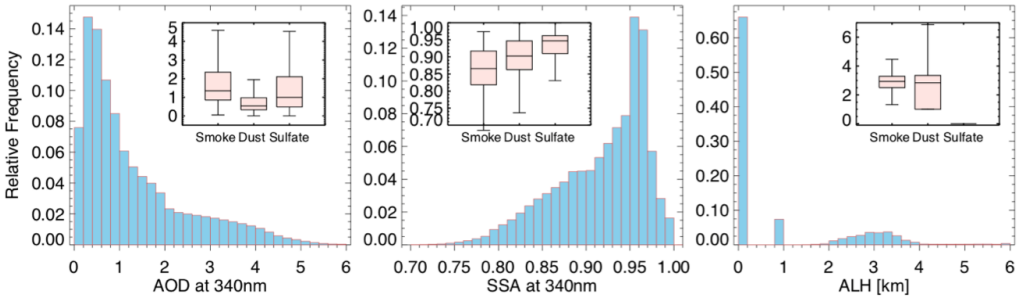

Figure 1. Histogram of AOD (left), SSA (middle) at 340 nm and ALH (right) from OMAERUV products during 2006–2007. The box-whisker plots include minimum and maximum, interquartile range (IQR; 25 to 75 percentiles) and median values. Minimum and maximum values exclude outliers which are less than 25 percentiles or greater than 75 percentiles by more than 1.5 times the IQR.

3 Aerosol-corrected HCHO AMF and VCD calculation

Typical HCHO retrievals in the UV spectral range follow a two-step approach. First, the HCHO SCD is derived by fitting a modeled spectrum to a measured spectrum of back-scattered solar radiation. The SCD is then converted to the VCD using the AMF as follows:

$$VCD = \frac{SCD}{AMF}. \tag{1}$$

AMFs are calculated using a radiative transfer model, assuming an atmospheric state and a vertical distribution of HCHO, typically derived from an atmospheric chemistry model (Palmer et al., 2001).

In this study, in order to investigate the aerosol impacts on HCHO retrievals, we calculate the aerosol-corrected HCHO AMF ($\mathrm{AMF_{aer}}$) with aerosol correction and HCHO AMF without aerosol correction, and produce the aerosol-corrected HCHO VCD ($\mathrm{VCD_{aer}}$) with aerosol correction and HCHO VCD without aerosol correction using OMHCHO SCDs at the same observation geometry, surface reflectance and HCHO profile used in the OMHCHO products. Figure 2 shows the flowchart of this approach. We construct LUTs of scattering weights using the linearized pseudo-spherical vector discrete ordinate radiative transfer (VLIDORT) model (R. J. Spurr, 2006) for various observing and aerosol conditions. A detailed description of the LUT is provided in Section 3.2. Using the LUTs of scattering weights, we calculate the aerosol-corrected scattering weights by applying aerosol information derived from OMAERUV retrievals for each pixel. Finally, $\mathrm{AMF_{aer}}$ and $\mathrm{VCD_{aer}}$ are calculated using the HCHO profile and SCD provided by the OMHCHO retrievals.

3.1 AMF formulation

According to Palmer et al. (2001), the AMF can be expressed by the integral of the product of atmospheric scattering weights and the vertical shape factor of a gas profile as follows:

$$AMF = \int_z w(z)s(z)dz. \tag{2}$$

Scattering weights (w) account for the sensitivity of the satellite measurement to each atmospheric layer, as determined using a radiative transfer model as functions of viewing

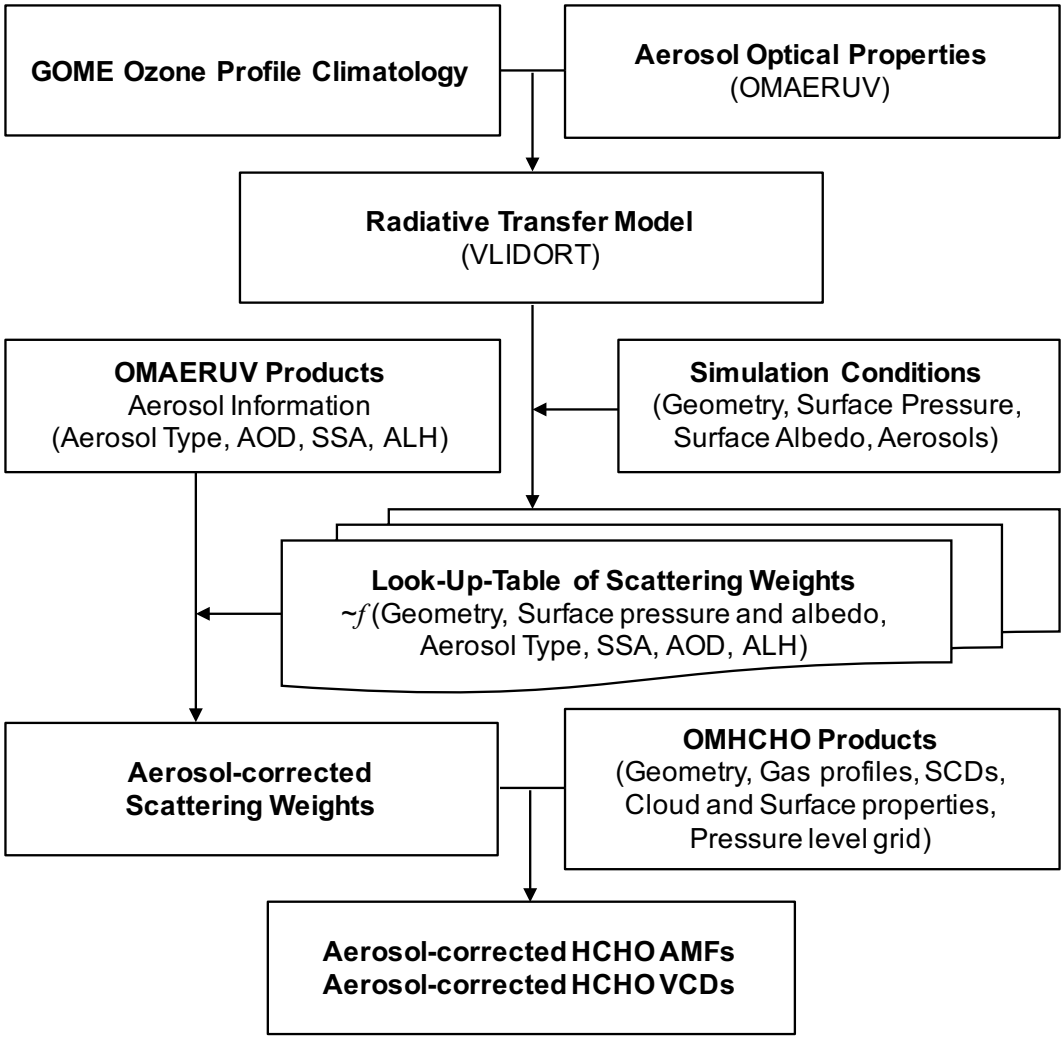

Figure 2. Flow diagram of explicit aerosol correction on trace gas retrievals

geometry, ozone profile, surface albedo, and aerosol information. The shape factor (s) represents the normalized vertical distribution of the trace gas profile, provided by correlative observations or model simulations. The decoupling of scattering weights and the shape factor makes the AMF calculation more flexible as different trace gas profiles may be used under the prescribed condition of scattering weights.

3.2 Look-up tables of scattering weights

To increase flexibility and computational efficiency, we construct LUTs of scattering weights, rather than of AMFs. A separate LUT is created for each aerosol type of dust, smoke, and sulfate. The GOME ozone profile climatology is used for meteorological data and ozone profiles. To build the climatology, the Smithsonian Astrophysical Observatory

(SAO) GOME ozone profile retrievals (X. Liu et al., 2007, 2005) are averaged within a latitude range (0–20° for tropics, 20–50° for mid-latitudes, and 60–90° for high latitudes) during March 1996–June 2006. Each profile consists of 46 layers from the surface up to 0.01 hPa. We use the mean mid-latitude profile with 350 DU as a standard profile to calculate scattering weights. We use the aerosol optical information (PSD and refractive index) used in the OMAERUV retrieval (*OMAERUV Readme File*, 2017) to calculate the scattering phase function (Mishchenko et al., 1997) so that we simulate the scattering weights at 340 nm by performing radiative transfer calculations with the VLIDORT version 2.7 radiative transfer model (R. J. Spurr, 2006). In general, scattering weights depend on wavelength. However, since ozone absorption has relatively weak spectral features in the HCHO fitting window (328.5–356.5 nm), the variability of scattering weights with respect to wavelength is less than 5 % over solar zenith angles (SZAs) smaller than 70° (González Abad et al., 2015). Therefore, we consider only the scattering weights at 340 nm in this study.

Since the scattering weights depend on the viewing geometry, surface albedo and aerosol information, the scattering weights in the LUTs are produced as functions of SZA, viewing zenith angle (VZA), relative azimuth angle (RAA), surface albedo, aerosol type, AOD, SSA and ALH, as well as surface pressure. Table 1 shows the parameters of the LUTs and the corresponding nodal points used in the LUTs. We use the same nodal points for viewing geometry and surface albedo as the operational OMI SAO HCHO algorithm LUT (González Abad et al., 2015).

The nodal points related to aerosol information (AOD, SSA, and ALH) are based on the values of OMAERUV products. As seen in Figure 1, AOD at 340 nm from the OMAERUV product is usually less than 6. The pixels with a quality flag of 0, which indicates the most reliable retrievals, only have an AOD greater than 3 in less than 1 % of the retrievals. As a result, the maximum nodal point of AOD in the LUT is set to 3 to minimize table size. The AOD nodal points of each aerosol type are determined by ensuring that the AMF interpolation error under all conditions is less than 5 % for the three HCHO representative profiles derived from GEOS-Chem (see Section 3.3 for more details on the representative profiles). The AOD nodal points of smoke and sulfate aerosol types are the same, but the dust type requires a greater number of nodal points to meet the 5 % AMF error criterion.

The OMAERUV algorithm uses a LUT with seven nodal points of imaginary refractive indices at 354 and 388 nm for each aerosol type. According to Kirchstetter et al. (2004)

and Russell et al. (2010), aerosol absorption can be expressed as a continuous function of wavelength and can be approximated using a power law with the value of absorption Ångström exponent (AAE) as follows:

$$k_\lambda = k_{\lambda_0} \cdot (\frac{\lambda}{\lambda_0})^{-AAE} \tag{3}$$

where k_λ and k_{λ_0} are absorption efficiency at the wavelength (λ) and reference wavelength (λ_0), respectively. In this study, the imaginary refractive index at 340 nm was calculated by using the imaginary refractive index at 354 and 388 nm using the OMAERUV assumption of AAE between 354-388 nm. The imaginary refractive indices at the seven nodal points are then converted into SSA at 340 nm by applying Mie theory (Mishchenko et al., 1997).

As seen in Figure 1, OMAERUV aerosols are mostly located near the surface (primarily sulfate) or at altitudes centered around 3 km. The aerosol vertical distribution is controlled by the aerosol layer shape and the altitude of the lower and upper boundary layer. The vertical distribution of smoke and dust aerosol types is based on the Gaussian distribution function with ALH and half width defined by R. Spurr & Christi (2014). For dust and smoke LUTs, we set the nodal points of ALH to 1 km intervals between 1–6 km. As the vertical distribution of sulfate aerosol is based on an exponential distribution peaking at the surface, the sulfate LUT does not contain an ALH variable. The nodal points of all parameters used in the LUTs are listed in Table 1.

For cloudy pixels, the scattering weights are calculated as follows:

$$w = (1 - \Phi) \cdot w_{\mathrm{clr}}(a_{\mathrm{s}}) + \Phi \cdot w_{\mathrm{cld}}(a_{\mathrm{c}}). \tag{4}$$

where w_{clr} and w_{cld} are the scattering weights for clear-sky and cloudy-sky conditions, respectively. We assume clouds to be opaque Lambertian surfaces with a cloud albedo (a_{c}) of 0.8. a_{s} is the surface albedo at each pixel from the OMHCHO retrievals, and Φ is the intensity-weighted cloud fraction (radiative cloud fraction) which is calculated as follows:

$$\Phi = \frac{C_{\mathrm{f}} \cdot I_{\mathrm{cld}}}{(1 - C_{\mathrm{f}}) \cdot I_{\mathrm{clr}} + C_{\mathrm{f}} \cdot I_{\mathrm{cld}}}. \tag{5}$$

where I_{clr} and I_{cld} are the radiance for clear-sky and cloudy-sky at the top of atmosphere considering aerosols. C_{f} is the cloud fraction extracted from the OMHCHO products, which is calculated by considering aerosol implicitly in the cloud algorithm. To accurately examine the effects of aerosols on the trace gas retrievals, aerosol must be explicitly considered in the cloud algorithm to calculate cloud fraction under aerosol free conditions. Even though cloudy pixels can be identified in the OMAERUV retrieval products, they are typically carried out under cloud free conditions. This study evaluates only pixels with a quality flag of 0, representing cloud free or minimally contaminated pixels. We find some pixels flagged as 0 in the OMAERUV product may include significant cloud fractions. Therefore, we analyze the dependency of our results with respect to cloud fraction in Section 5.4.

Table 1. Parameters used in the LUTs of scattering weights

Variables	Nodal points	Number of variables
Surface Pressure [hPa]	500, 600, 700, 800, 900, 1013, 1050	7
Aerosol Type	Smoke, Dust, Sulfate	3
Single Scattering Albedo (SSA)	1.00, 0.96, 0.93, 0.87, 0.82, 0.78, 0.75 (Smoke) 1.00, 0.96, 0.92, 0.87, 0.83, 0.79, 0.73 (Dust) 1.00, 0.95, 0.91, 0.88, 0.85, 0.82, 0.79 (Sulfate)	7
Aerosol Optical Depth (AOD)	0.0, 0.1, 0.3, 0.6, 1.0, 1.5, 2.0, 2.5, 3.0 (Smoke, Sulfate) 0.0, 0.1, 0.4, 0.7, 1.1, 1.2, 1.3, 1.7, 2.1, 2.5, 3.0 (Dust)	9 11
Aerosol Layer Height (ALH) [km]	Surface, 1, 2, 3, 4, 5, 6 (Smoke, Dust) Surface (Sulfate)	7 1
Surface Albedo	0, 0.01, 0.05, 0.1, 0.2, 0.4, 0.6, 0.8, 1.0	9
Solar Zenith Angle (SZA) [°]	0, 15, 30, 45, 50, 60, 70, 75, 80, 85, 89.9	11
Viewing Zenith Angle (VZA) [°]	0, 15, 30, 45, 55, 65, 70, 75	8
Relative Azimuth Angle (RAA) [°]	0, 30, 60, 90, 120, 150, 180	7

3.3 HCHO profiles

As stated above, the shape factor represents the normalized vertical distribution of a trace gas profile and generally depends on the measured time and region. The operational OMHCHO retrievals use monthly mean climatologies of vertical profiles derived from the GEOS-Chem chemical transport model (Bey et al., 2001) on 47 vertical levels (González Abad et al., 2015). For Section 5, where we apply an explicit aerosol correction to the OMI HCHO products, we use the GEOS-Chem HCHO profile climatologies used in the OMHCHO operational product. However, in order to closely examine the AMF sensitivity to aerosols in Section 4.1, we select three representative HCHO profiles derived from the GEOS-Chem climatology. Figure 3 shows the GEOS-Chem HCHO climatology at surface

level for OMI orbit 47225 on June 1, 2013, and the three profiles that are chosen to represent background, urban, and biogenic regions in our study.

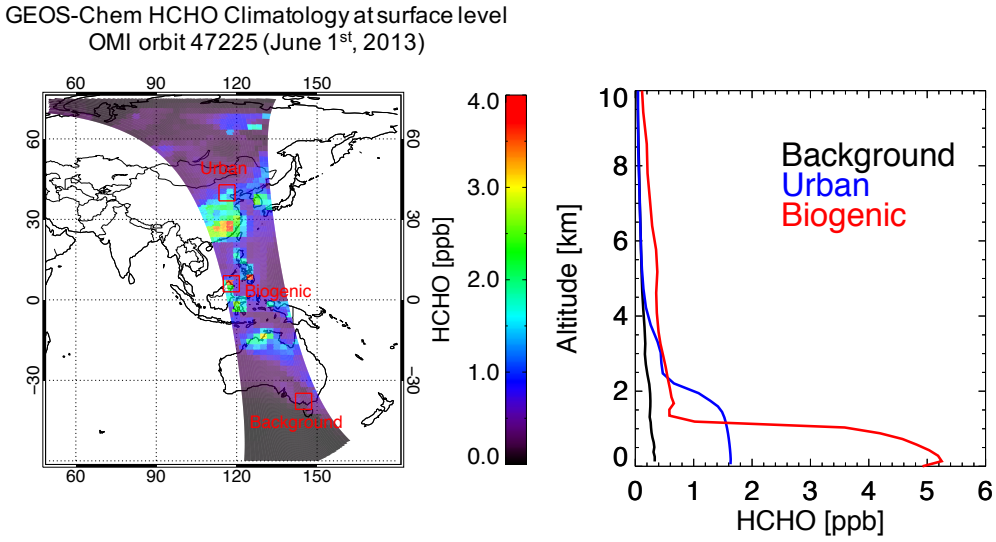

Figure 3. (left) The GEOS-Chem HCHO climatology at the surface level for OMI orbit 47225 and (right) derived HCHO profiles over background, urban, and biogenic regions on June 1, 2013.

.

4 Analysis of aerosol effects on AMF calculation

The presence of aerosols can change the light path of the detected photons and the contributions of each layer to the total radiance observed by the satellite. The total effect of aerosols strongly depends on the relative relation between the aerosol and gas vertical distributions as well as other parameters, including geometry, surface albedo, and aerosol optical properties. The presence of aerosols generally can have the following two impacts according to previous studies (Chimot et al., 2016; Leitão et al., 2010; Lin et al., 2015):

1. Sensitivity decreases below the aerosol layer as the fraction of detected photons at the top of the atmosphere that reaches the lowest part of the atmosphere is reduced compared to those in an aerosol-free scene (shielding effect).
2. Sensitivity increases within and above the aerosol layer since the fraction of all detected photons that sample the higher part of the atmosphere is increased as more photons are scattered back to the satellite (albedo effect). However, in the case of strongly absorbing aerosols, the sensitivity decreases even above the aerosol layer as the number of returning photons is reduced.

4.1 The sensitivity of the AMF to aerosol information

We have performed a sensitivity analysis to assess the influence of aerosol characteristics (optical properties, loading, and vertical distribution) on AMF calculations. The sensitivity analysis is carried out using three representative HCHO vertical profiles (background, urban, and biogenic) derived from GEOS-Chem simulations (see Figure 3). We perform the sensitivity study using a reference condition of surface pressure = 1013 hPa, SZA = 30°, VZA = 15°, RAA = 60°, surface albedo = 0.05, and ALH = 1 km for smoke and dust aerosols or ALH = surface for sulfate aerosols. The reference SSAs for smoke, dust, and sulfate aerosols are 0.87, 0.92, and 0.95, respectively, which are the median SSA values for each aerosol type from the OMAERUV products considering only retrievals with quality flag equal to 0. We carry out the sensitivity analysis of aerosol correction factors (ACFs) with respect to SSAs considering the range of nodal points in the LUTs. Here, we define the ACF as the ratio of the HCHO AMF_{aer} to AMF as follows:

$$ACF = \frac{AMF_{\text{aer}}}{AMF}. \tag{6}$$

where AMF is calculated from scattering weights by applying observing geometry, surface reflectance and HCHO profile consistent with AMF_{aer} except for aerosol information. Figure 4 shows the change of ACFs with respect to AOD for different surface albedos, ALHs, and SSAs for each aerosol type. The shaded regions indicate the variability in the ACFs due to using the three representative HCHO profiles. As expected, the effects of aerosols are amplified with increasing AOD.

Sulfate aerosols, with a median SSA of 0.95, are located near the surface. Being weakly absorbing aerosols, they have the effect of increasing AMFs at low albedos (<0.2). Under the sulfate aerosol reference conditions, AMFs steadily increase with AOD, up to 24 % for AOD = 3.0. At high aerosol loadings (AOD = 3.0), the increase of AMFs is enhanced for small albedos (ACF = 1.4 at an albedo of 0.01). Depending on the SSA, sulfate AMFs can increase up to 58 % (SSA = 1.00) or only 5 % (SSA = 0.91) at these high aerosol loadings.

Dust aerosols have the effect of increasing AMFs under the reference conditions, showing an increase of 10 % at an AOD of 3.0, as can be seen in the middle row panels. However, the effect of dust aerosols can change from increasing to decreasing the AMF as albedo or ALH increases. For example, as the ALH of dust aerosol increases, the ACF transitions from values above 1 to values as low as 0.51 for ALH = 6 km, which is the result of a strong shielding effect.

Smoke aerosols are typically strongly absorbing aerosols with small SSAs and in consequence they usually have the effect of decreasing AMFs with increasing AOD. At the reference conditions, the AMF decreases by 7 % at an AOD of 3.0, as can be seen in the top panels of Figure 4. The effect of a reduction in the AMFs is amplified up to 30 % when the surface albedo is increased to 0.2. As the ALH increases from 1 to 6 km, this effect is amplified even more and is as large as 57 % for ALH = 6 km (top-middle panel). Even though smoke aerosols have a median SSA of 0.87 in the OMAERUV retrievals, they can have increasing effects on the AMFs in the case of some smoke aerosols which have high retrieved SSA values (top-right panel). It can be seen that the variability due to the HCHO profiles (background, urban, and biogenic) of the ACFs for smoke and dust aerosols increases with increasing ALH.

Increasing SSA means that the aerosols relatively have more scattering and therefore they have enhancement effects on AMF calculations (see right column panels of Figure 4). This change from shielding effect to albedo effect occurs at SSAs around 0.88. This im-

plies that the absorptive characteristics of the aerosol is one of the most important factors in determining the effect of an increasing or decreasing AMF. Similarly, a decreasing surface albedo means that the relative scattering contribution of the aerosol becomes more significant, resulting in an increase in the AMF.

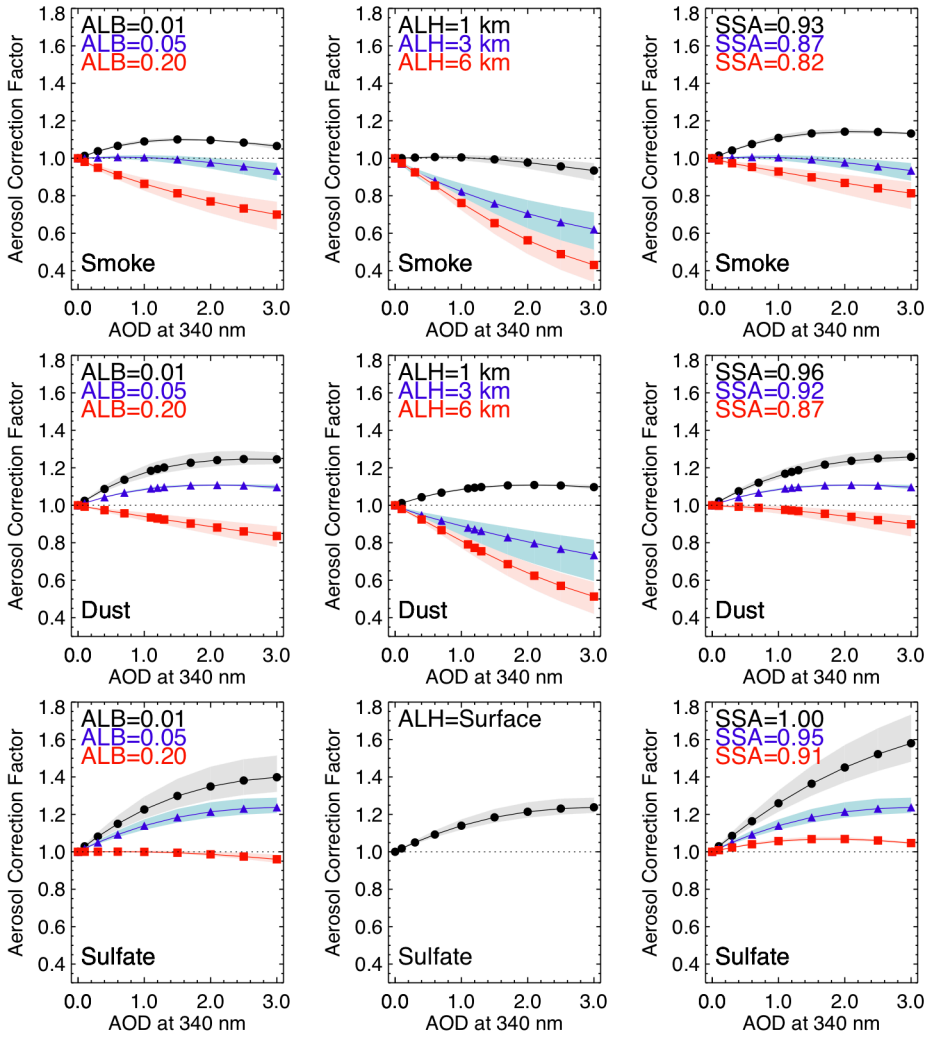

Figure 4. Aerosol correction factor (ACF) with respect to AOD at 340 nm depending on surface albedos (left column), ALHs (middle column), and SSAs (right column) for smoke (upper), dust (middle), and sulfate (lower), respectively. The shaded region shows the variability of the ACF using the three representative HCHO profiles.

4.2 Uncertainties due to OMAERUV errors

The OMAERUV product does not provide the uncertainties for each observation. Ahn et al. (2014) and Jethva et al. (2014) evaluated the statistical distribution of uncertainties

by comparing OMAERUV with AERONET measurements. They showed the expected uncertainty to be the largest of 30 % or 0.1, and 0.05 for OMAERUV AODs and SSAs, with 65 % and 69 % of OMAERUV AODs and SSAs falling within the expected uncertainty, respectively. In this section, we examine the uncertainties of ACFs due to uncertainties in the aerosol information in OMAERUV retrievals.

We can use our LUTs to assess the sensitivity of the ACF to uncertainties in the aerosol characteristics. At the reference conditions in Section 4.1, the uncertainties in the OMAERUV AOD product result in uncertainties in the ACF of 0.5 %, 1.0 %, and 1.4 % for smoke, dust, and sulfate aerosols, respectively. The expected uncertainty in the retrieved OMAERUV SSA of 0.05 (Jethva et al., 2014) results in ACF uncertainties of 10.3 %, 11.4 %, and 10.5 % for smoke, dust, and sulfate aerosols, respectively. OMAERUV uses different assumptions and climatologies to assign ALHs. We have evaluated the impact of the poor knowledge of the ALHs by perturbing OMAERUV ALHs by 1 km. We find that ACFs change, for the reference conditions, by 11.7 % and 12.2 % for smoke and dust aerosols, respectively.

It is worth noting that we quantify the aerosol effects on HCHO VCDs assuming spherical particles. However, the assumption of the particle shape is an important factor that influences the accuracy of OMAERUV retrievals (Gassó & Torres, 2016). In this study, we have assumed spherical particles. However, it is known that dust aerosols are irregularly shaped which may induce significant differences in retrieved AOD at scattering angles greater than 155° (Torres et al., 2018). These angles correspond to approximately 20 % of pixels identified as dust aerosols. The SSA errors due to the consideration of spherical particles in the case of dust aerosols are smaller than the expected uncertainty of OMAERUV SSA retrievals (Torres et al., 2018). In the subsequent analysis, we exclude pixels identified as dust aerosols with scattering angles greater than 155°. As part of a future study, we plan to improve the quantified aerosol effects by considering newly modeled dust aerosols that assume non-spherical particles in updated OMAERUV products (Torres et al., 2018).

5 Explicit aerosol correction applied to OMI products

In this section, we evaluate the effects of the explicit aerosol correction on OMHCHO retrievals. We quantify the aerosol-induced VCD difference, which is the difference between VCD_{aer} and VCD calculated by applying the AMF_{aer} and AMF determined from the LUTs to the OMHCHO SCD. The AMF_{aer} uses aerosol information from the OMAERUV at each ground pixel, while the AMF (aerosol-free case) is calculated using the same LUT but with AOD=0. The large value of aerosol-induced VCD difference implies that the effects of aerosols need to be considered on HCHO retrievals. We only consider pixels where the OMHCHO and OMAERUV retrievals are available for the period of 2006–2007, prior to the onset of the OMI row anomaly.

The OMAERUV product provides data quality flags. A quality flag of 0 indicates that the retrieved AOD and SSA from the OMAERUV product are reliable, while a quality flag of 1 indicates that only the retrieved AAOD is reliable. Except for in Section 5.4, only the retrievals with a quality flag of 0 from the OMAERUV retrievals are used.

5.1 Evaluation of the explicit aerosol correction: Case studies

We show the results for explicit aerosol correction applied to the OMHCHO retrievals for specific cases. Figure 5 shows the results of the explicit aerosol correction as well as the aerosol information for several OMI orbits over Africa and Europe on August 27, 2007, as also seen in Torres et al. (2013). Pixels with a quality flag of 1 or regions without OMI observations are shown as a gray color in the figures. This figure includes the aerosol information (aerosol type, AOD, SSA, and ALH), ACF, OMHCHO VCD and VCD_{aer}, and the aerosol-induced HCHO VCD difference, which is expressed in percentage change between VCD_{aer} and VCD. Figure 5 shows a large smoke plume from a fire in Greece that has moved southward across the Mediterranean Sea, reaching northern Africa. Over northern Africa, the OMAERUV product identifies the smoke plume as absorbing aerosols with AODs up to 3.4 at an altitude of approximately 3 km. Over the region, the ACFs range from 0.51 to 1.09, with values less than 1.0 in most areas, resulting in increased HCHO VCDs. Over the smoke plume, the AMFs decrease when aerosols are considered, resulting in an increase in the HCHO VCDs of up to 97 %.

In addition, we investigate a case of very high aerosol conditions over Asia on March 26, 2006, as shown in Figure 6. On this day, dust and smoke aerosols over certain areas

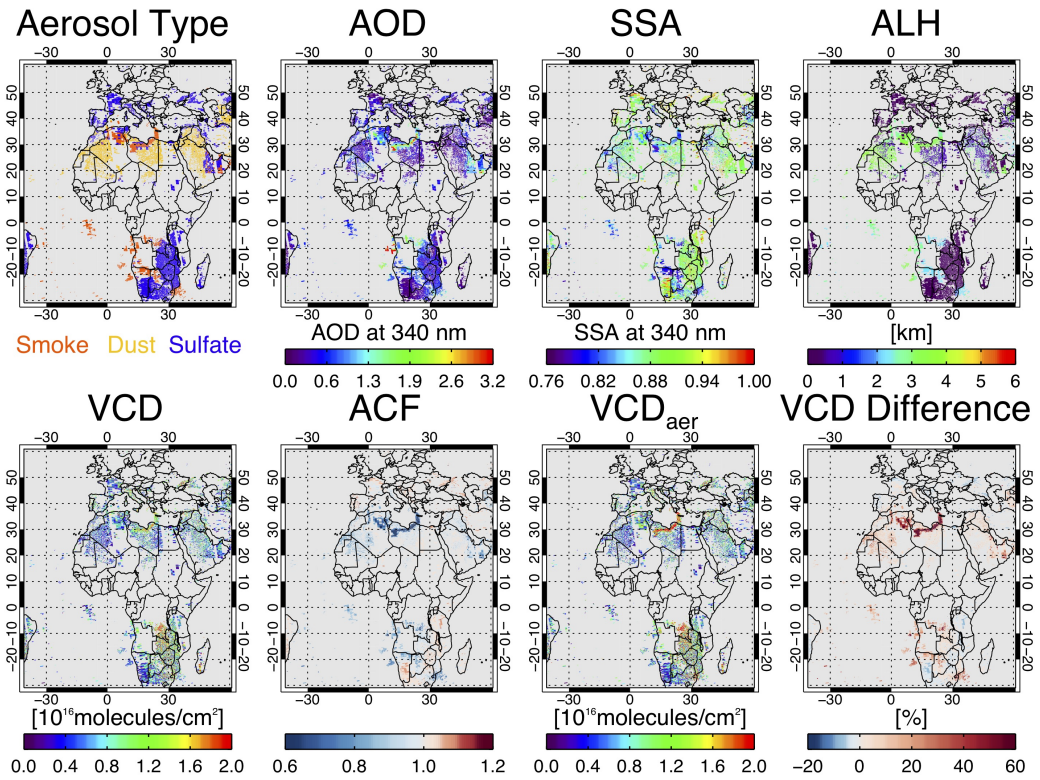

Figure 5. Aerosol correction results including aerosol type, AOD, SSA, ALH, HCHO VCD, aerosol correction factor (ACF) which is defined as the ratio of aerosol-corrected AMF (AMF_{aer}) to AMF, aerosol-corrected HCHO VCD (VCD_{aer}) and aerosol-induced HCHO VCD difference (VCD_{aer} - VCD) over Africa on August 27th, 2007. The HCHO VCD and aerosol-corrected HCHO VCD plots include only those observations where high quality OMI aerosol retrievals are available.

resulted in high AOD values, up to 3.5, and SSAs as low as 0.74. The AMF_{aer} is decreased relative to the OMHCHO AMF over most regions in the figure except for south India where relatively high SSA aerosols are present, showing ACFs between 0.52–1.09. The shielding effects dominate in many areas, given the absorbing mixture of dust and smoke aerosols. These effects are noticeable over central and northeastern China and also extend to South Korea. The aerosol-induced VCD differences in individual measurements over the entire region range from -9 % to 92 %, and are particularly strong over Myanmar due to smoke from biomass burning.

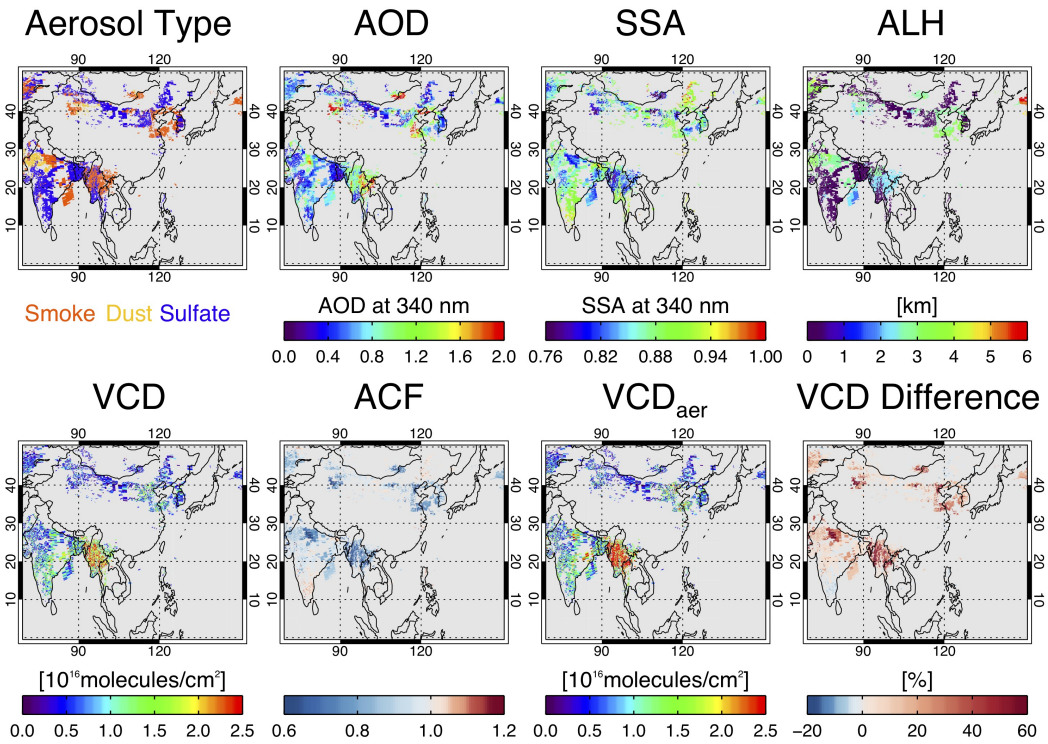

Figure 6. Same as Figure 5, but over East Asia on March 26th, 2006

5.2 Climatology of aerosol correction factors

In this section, we analyze the climatological results of the ACFs due to the explicit aerosol correction for OMHCHO retrievals during 2006–2007. We examine how the ACF changes with aerosol information, and show the values as functions of AOD, SSA, and ALH (Figures 7 and 8). The number of data points used to calculate the mean and standard deviation of ACFs is also shown in these figures. Smoke aerosols are mainly absorbing aerosols showing SSA <0.87, and mainly exist at ALHs of 2–4 km with AOD values mostly smaller than 1.5 (see the left column of each figure). As a result, as seen in the left-top panel of each figure, the ACFs typically have values between 0.7–0.9, indicating that AMF decreases for smoke aerosols. The standard deviation of ACFs for smoke aerosols increases with increasing AOD in both figures, showing that the variability of ACFs increases with increasing AOD. The majority of dust aerosols have the effect of decreasing AMFs, giving ACFs of less than 1. However, some dust aerosols have relatively high SSAs (>0.84) and low ALHs (<2.4 km). These aerosols have enhancement effects on the AMF for AOD<0.4 (top-middle panel in Figures 7 and 8). Sulfate aerosols, with higher SSA than smoke and dust aerosols, show a minimum ACF of 0.79. However, most sulfate aerosols are symmetrically distributed around a SSA of 0.9 with AODs less than 0.5 where the ACF is closer to 1. The variability of sulfate-induced ACFs for AOD<1.5 is also small.

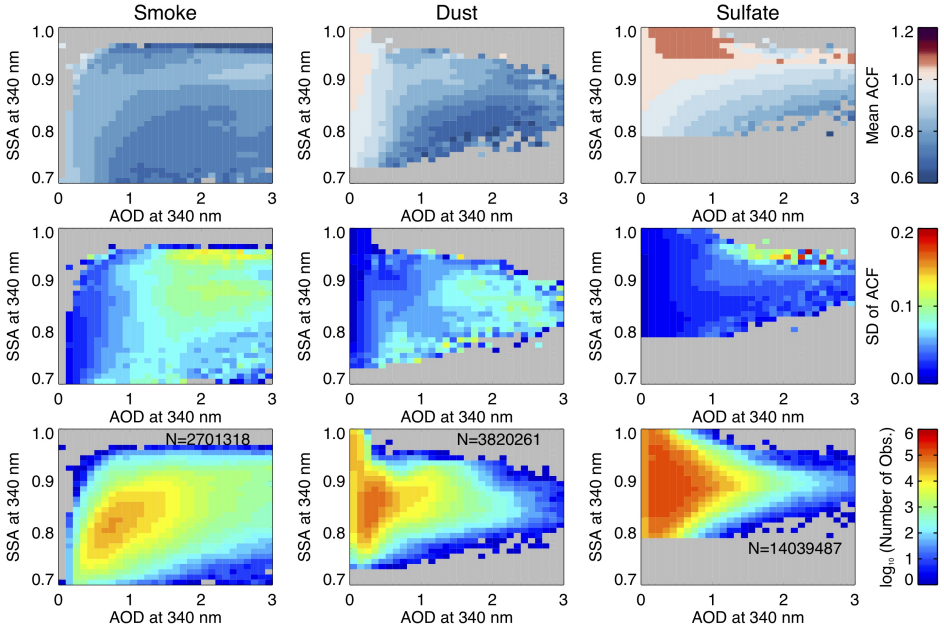

Figure 7. Mean value (upper) and standard deviation (SD) (middle) of the ACF, and the number of observations (lower) as functions of AOD and SSA at 340 nm for smoke (left column), dust (middle column) and sulfate (right column) aerosols during 2006–2007.

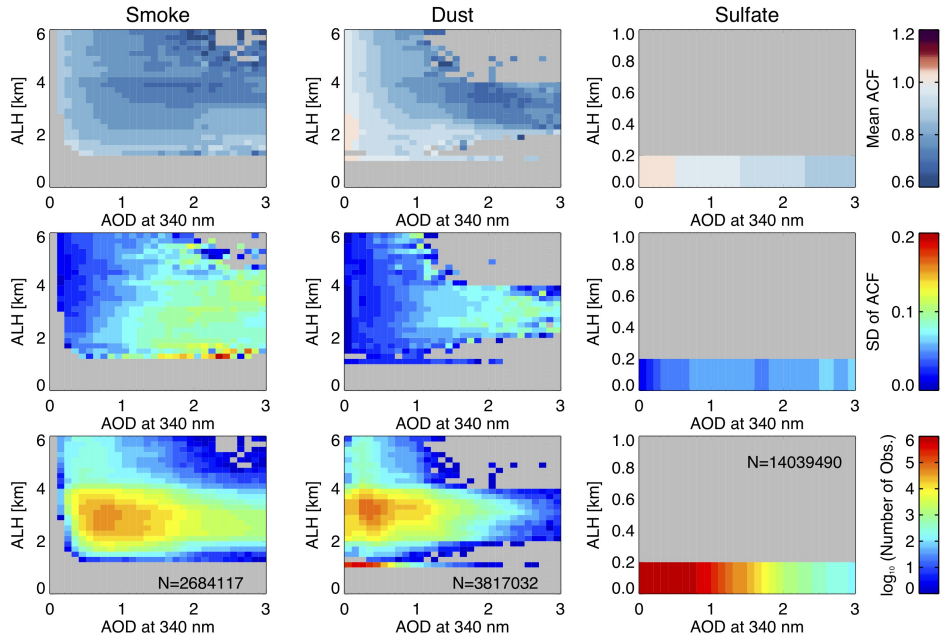

Figure 8. Same as Figure 7, but as functions of AOD at 340 nm and ALH.

5.3 Climatology of aerosol-corrected HCHO VCDs

By applying the ACF at each pixel, we can also analyze the seasonal mean aerosol-induced HCHO VCD difference during 2006–2007. Figure 9 shows the seasonal mean HCHO VCD, VCD_{aer} and the aerosol-induced HCHO VCD difference (VCD_{aer}-VCD) between the operational OMHCHO VCDs and our aerosol-corrected VCDs for the MAM (March to May), JJA (June to August), SON (September to November), and DJF (December to February) seasons. Considering the values between 25 and 75 percentiles, the global mean aerosol-induced VCD differences for the MAM and DJF seasons range from -2 % to 15 % and -1 % to 16 % with a minimum and maximum of -26 % and 41 %, respectively. Similarly, the values for the JJA and SON seasons range from -3 % to 11 % and -2 % to 10 % with a minimum and maximum of -26 % and 52 %, respectively.

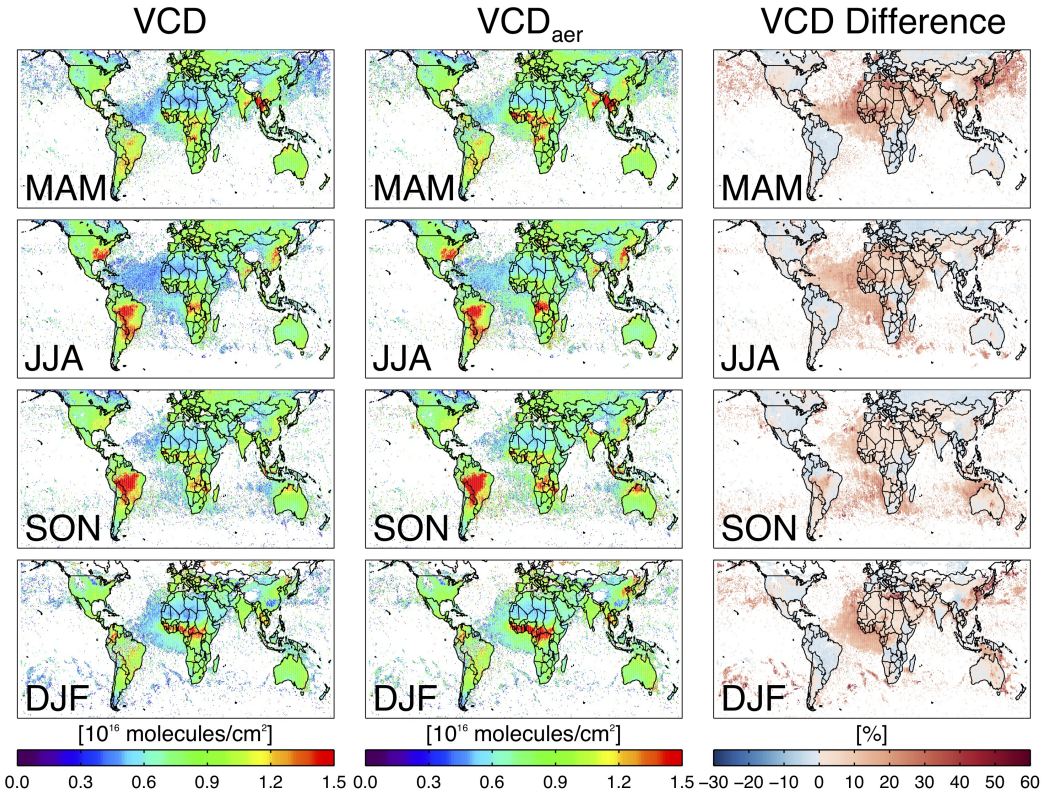

Figure 9. Seasonal mean global distribution of the HCHO VCD (left), the aerosol-corrected HCHO VCD (VCD_{aer}, middle), and the aerosol-induced HCHO VCD difference (VCD_{aer}-VCD, right) for the MAM, JJA, SON, and DJF seasons during 2006–2007. The black boxes show the different selected regions used to explore the regional aerosol-induced VCD differences.

We analyze the monthly aerosol-induced VCD difference for six regions. The geographic locations of these regions are indicated as black boxes in the right-hand column of Figure 9 and listed in Table 2. Figure 10 is a box-whisker plot of the monthly mean aerosol-induced VCD difference and relative frequency of occurrence for each aerosol type in the selected regions. In the box-whisker plot, the box extends to the interquartile range (IQR) with the upper and lower quartiles of the box edge. A horizontal line inside the box shows the median and the whiskers extend to the highest and lowest values, which are calculated after removing extreme values less than lower quartiles or upper quartiles by 1.5 times the IQR. A box-only plot without whiskers indicates that there is only one data point available for that particular month. As shown in the right-hand columns of Figure 10, sulfate aerosols show the highest monthly occurrence in most regions among all aerosol types. The aerosol-induced VCD differences of sulfate aerosols are mostly negative due to the fact that they are mostly located near the surface and have high SSA (a median SSA of 0.95). The monthly mean aerosol-induced VCD differences of sulfate aerosols show similar values in all regions. On the other hand, the monthly mean aerosol-induced VCD differences from smoke and dust aerosols show greater variability than sulfate aerosols.

HCHO VCDs can be very large over the continental tropical regions, in particular over the Amazon basin and Africa, as was seen in Figure 9. Biomass burning contributes significantly to high HCHO columns during the dry season, and also biogenic isoprene emissions are expected to be at a peak during the season (Müller et al., 2008). The southeastern US and southeastern Asian regions are also regions with significant HCHO VCD during certain seasons. Over the southeastern US, HCHO VCD shows a strong seasonal variation related to the oxidation of biogenic VOCs during the summer, with several studies reporting values about two times higher in summer than in winter (Abbot et al., 2003; Palmer et al., 2006). Over the southeastern US and Europe, as can be seen in Figure 10, sulfate aerosols dominate OMAERUV retrievals, and dust and smoke aerosols rarely occur. For example, dust aerosol is only once detected during February–May over the southeastern US, and the aerosol-induced VCD difference of dust aerosols cannot be regarded as characteristic of this region. We find dust aerosols contribute very weakly to increasing or decreasing AMF here, with most ACFs close to 1.0 with insignificant variability. The VCD differences in the southeastern US are typically minimal, although the VCD difference in summer shows a slightly positive value compared to other seasons. Similar to the southeastern US, HCHO VCDs over Asia have high seasonal variability. The HCHO VCDs over southern China show a

pronounced seasonal variation related to increased biogenic emission during the JJA season. Over Myanmar and Thailand, the maximum HCHO VCDs occur in the MAM season and are a factor of two higher as compared to the other seasons.

Over the Amazon and southeastern Asia, sulfate aerosols also appear with high frequency, but are detected less often than smoke aerosols in September or October. In addition, the AOD and SSA of smoke aerosols are higher in September and October than those during other seasons (not shown here). This results in a decrease of aerosol-corrected AMF and an increase of aerosol-corrected VCD. Over the Amazon basin, the monthly mean aerosol-induced VCD differences for smoke, dust, and sulfate aerosols are 24 %, 69 %, and -3 %, respectively. The aerosol-induced VCD difference is particularly variable over the Amazon region in August, showing a maximum value of 97 %. Over East China and tropical Africa, the frequencies of occurrence of each aerosol type are similar, with higher aerosol loading in winter. In East China, AODs are highest in June and September. Except for the period of July–September, monthly mean SSAs shows high variability so that aerosol-induced VCD differences also show the largest variability. The monthly mean aerosol-induced VCD difference over East China is 35 %, 13 %, and -0.1 % for smoke, dust, and sulfate aerosol, respectively.

Table 3 lists the mean, standard variation, minimum, and maximum values of the aerosol-induced VCD difference for each aerosol type and region. The aerosol-induced VCD difference for smoke aerosols is large and positive in all regions. Smoke aerosols cause an aerosol-induced VCD difference of 27 % in the global average. The largest mean values are seen over East China with individual values as high as 142 %. The aerosol-induced VCD difference caused by dust aerosols shows a global average of 6 %, and usually the exclusion of dust in the AMF calculations has a smaller influence on the mean VCDs than excluding smoke aerosols for the studied regions. The mean dust aerosol-induced VCD difference in each region ranges from 6–16 %. Biases as high as 83 % in individual VCDs can occur when dust is not considered in the trace gas retrievals. Sulfate aerosols have relatively weak effects on the VCD difference as compared with dust and smoke aerosols. The mean sulfate aerosol-induced VCD difference shows slightly negative values, except for East China where the effect in some months is positive. The global mean aerosol-induced VCD difference for sulfate on average is -0.3 % with a minimum of -13 % and a maximum of 13 %.

Table 2. Geographic limits of the regions shown in Table 9

Regions	Geographic limits
Southeastern USA	30°–41°N, 77°–95°W
Amazon basin	15°S–0°, 50°–70°W
Europe	40°–52°N, 0°–25°E
Southeastern Asia	5°S–5°N, 95°–120°E
Tropical Africa	1.5°S–11°N, 8°W –30°E
East China	28°–39°N, 111°–120°E

Table 3. Mean, standard deviation (STD), minimum (Min), and maximum (Max) of the aerosol-induced HCHO VCD difference (VCD_{aer}-VCD) of each aerosol type (smoke, dust, and sulfate) in different regions as listed in Table 2. Geographic limits of these regions shown in Figure 9.

.

VCD difference [%]	Smoke			Dust			Sulfate		
	Min	Mean ± STD	Max	Min	Mean ± STD	Max	Min	Mean ± STD	Max
Southeastern US	4	25 ± 8	81	-7	10 ± 6	44	-13	-3 ± 3	9
Amazon basin	-19	24 ± 9	97	-12	6 ± 6	55	-14	-3 ± 3	9
Europe	-4	23 ± 8	62	-5	8 ± 4	32	-15	-2 ± 3	11
Tropical Africa	-10	22 ± 8	67	-1	16 ± 4	39	-16	-0.3 ± 5	17
East China	-61	35 ± 17	142	-8	13 ± 3	83	-18	-0.1 ± 5	23
Southeastern Asia	-5	12 ± 6	41	-2	10 ± 1	19	-11	-2 ± 3	8
Global Mean	-10	27 ± 11	72	-21	6 ± 6	39	-13	-0.3 ± 4	13

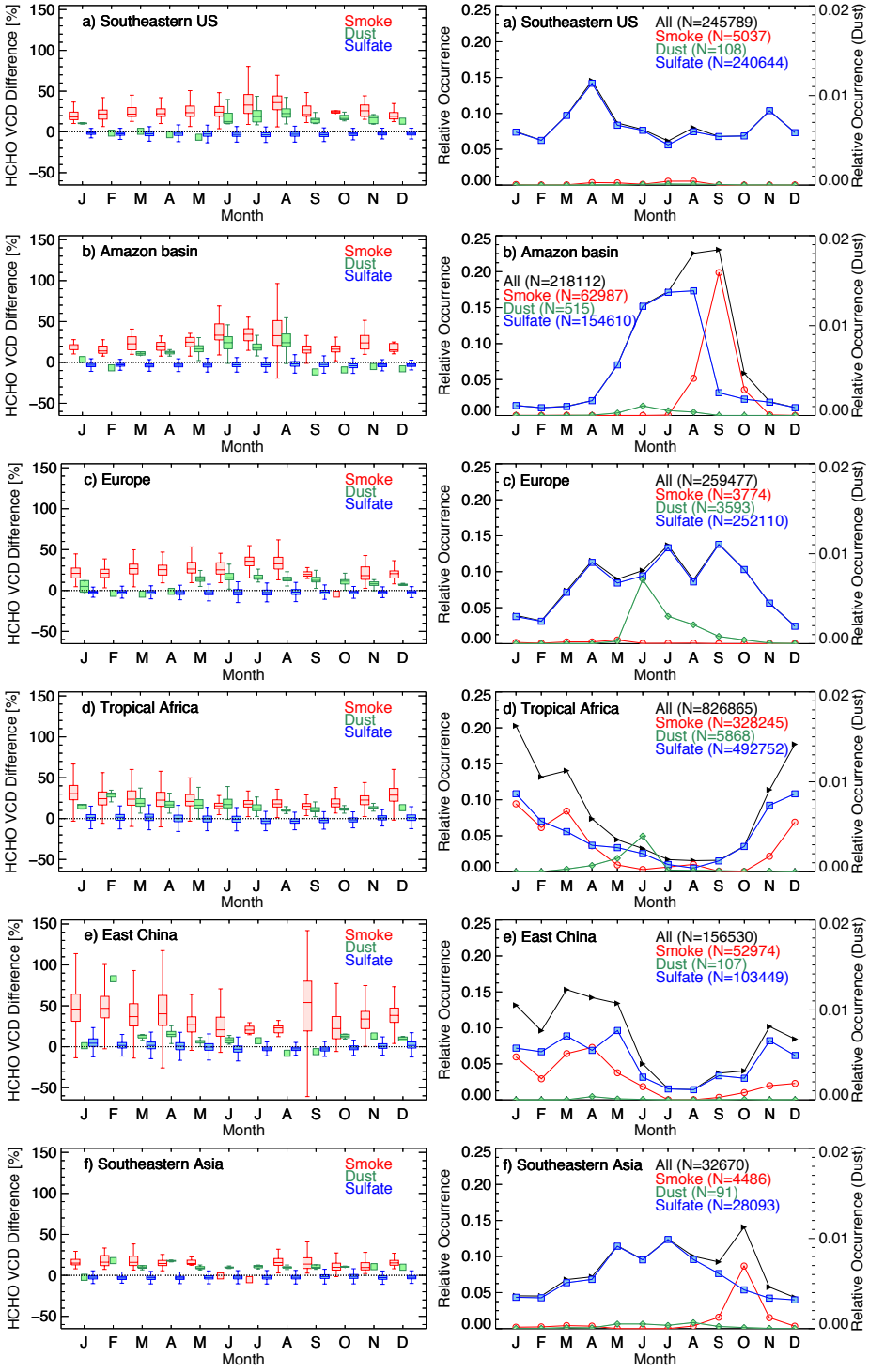

Figure 10. (left) Box-whisker plot of monthly mean aerosol-induced VCD percent difference (VCD_{aer}-VCD) of smoke, dust, and sulfate aerosols over (a) southeastern US, (b) Amazon basin, (c) Europe, (d) tropical Africa, (e) east China and (f) southeastern Asia during 2006–2007 as listed in Table 2 . The box-whisker plots include the minimum, lower quartile, median, upper quartile and maximum values. The box-only symbol means that there is only one detected pixel in that month. (right) Monthly mean relative frequency of each aerosol type over the selected regions during 2006–2007. The relative frequency of dust aerosols is shown on the right-side y-axis as they show very small occurrence relative to smoke and sulfate.

5.4 Cloud-aerosol interaction

We derived the results shown in Section 5.3 using OMAERUV retrievals with a quality flag of 0. Since the OMCLDO2 cloud products already contain aerosol information (Acarreta et al., 2004), there exists a correlation between the retrieved cloud fraction and the AOD. We wish to identify the range of cloud fractions at which the explicit aerosol correction is stable and suffers limited interference from clouds. Figure 11 shows the aerosol-induced HCHO VCD difference of each aerosol type with respect to radiative cloud fraction, as well as the histogram of the radiative cloud fraction for the OMAERUV retrievals with quality flags of 0 and 1 during the period of 2006–2007.

Retrievals with a quality flag of 0 offer highly reliable AOD and SSA values (top row of Figure 11). Most retrievals classified as dust or sulfate aerosols show low cloud fractions, less than 0.15, with cloud fractions greater than 0.15 being no more than 3 % of the total retrievals. The aerosol-induced VCD differences of sulfate aerosols are consistent for cloud fractions below 0.15, mostly showing near-zero aerosol-induced VCD differences. The histogram distribution of cloud fractions for dust aerosols is similar to that of sulfate aerosols, but the aerosol-induced VCD differences do not show a consistent behavior with cloud fraction. Given the dependence on cloud fraction, we recommend using the dust aerosol correction only for cloud fractions below 0.05. On the other hand, retrievals of smoke aerosols can provide some information at cloud fractions up to 0.1. Interestingly, despite the presence of high cloud fraction, the aerosol-induced VCD differences of smoke aerosols are consistent at cloud fractions less than 0.1. In the presence of clouds, the effects of aerosols can be compensated depending on cloud altitude and cloud fraction (Lin et al., 2014). In order to ensure the quality of the aerosol correction, a detailed analysis that considers the interaction between aerosols and clouds is planned for the future.

In the case of OMAERUV retrievals with a quality flag of 1 (bottom row of Figure 11), there are relatively significant number of retrievals with high cloud fractions for all aerosol types, compared with the case of a quality flag of 0. In addition, the inconsistency with the behaviours seen for the case of a quality flag of 0 indicates the retrievals with a quality flag of 1 are not suitable for a reliable aerosol correction.

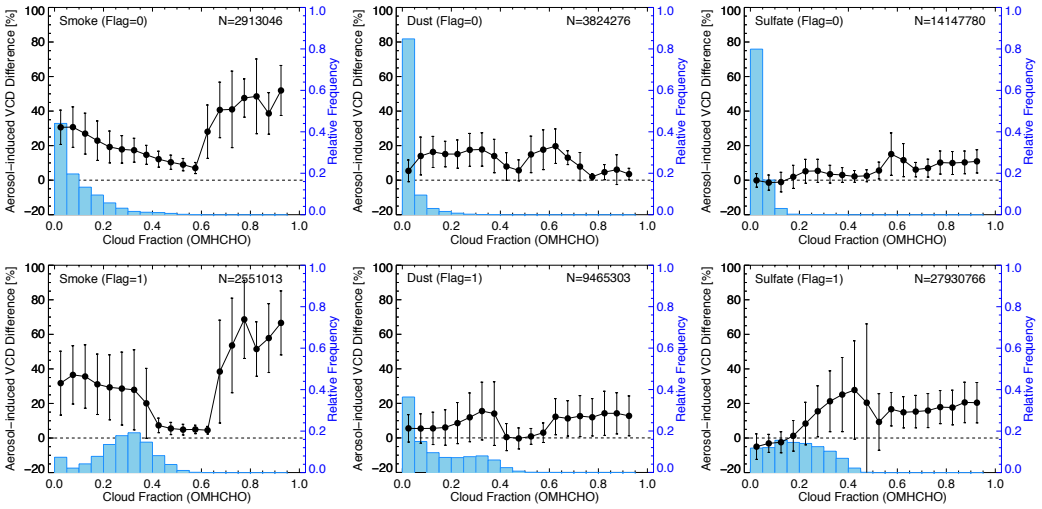

Figure 11. Annual mean aerosol-induced HCHO VCD difference (VCD_{aer}-VCD) and relative frequency with respect to radiative cloud fraction for smoke (left), dust (middle), and sulfate (right) aerosols, during the period of 2006–2007 for the cases for retrievals with a quality flag of 0 (top) and 1 (bottom).

6 Summary and conclusions

A new measurement-based explicit aerosol correction approach for OMI HCHO retrievals using OMI aerosol observations is presented. To perform the aerosol correction, we recalculate the aerosol-corrected scattering weights considering the geometry, surface albedo and aerosol information (aerosol type, AOD, SSA, and ALH) from the OMAERUV retrievals using the VLIDORT radiative transfer model. We examine the sensitivity of AMFs to aerosol information using representative HCHO profiles from GEOS-Chem model simulations. Applying the approach to OMI HCHO retrievals, we calculate the aerosol-corrected HCHO VCDs and quantify the aerosol-induced VCD differences of each aerosol type for the time period 2006–2007.

The effects of aerosols shown in this study are consistent with previous investigations (e.g. Chimot et al., 2016; Leitão et al., 2010). As SSA increases, aerosols are more scattering and thus have enhancement effects on AMF calculations. In addition, if the aerosol layer is below the gas layer, the aerosol-corrected HCHO AMFs are increased by the albedo effect, except for strongly absorbing aerosols. In this study, we have seen from the climatology of ACFs that of the three OMAERUV aerosol types, smoke aerosols (strongly absorbing) tend to decrease AMFs, while dust and sulfate aerosols may increase or decrease AMFs depending on the aerosol vertical distribution and the SSA.

We have examined the effects of aerosols on HCHO retrievals and quantified the aerosol-induced HCHO VCD difference of each aerosol type in six regions of the world. Even though sulfate aerosols are the dominant type in OMAERUV retrievals, they have small effects on HCHO measurements. More substantial aerosol effects appear over regions with smoke or dust aerosols. Excluding the aerosol effect underestimates HCHO VCDs by an average of 27 % over regions dominated by smoke aerosols with differences found in individual retrievals as high as 142 % over East China. Globally, dust aerosols increase HCHO VCDs on average by 6 %. The impact of dust aerosols on individual observations can be large, with impacts as high as 83 % over East China. Sulfate aerosols generally have the smallest effect, decreasing HCHO VCDs by 0.3 % on average globally.

We quantified the aerosol effects using the OMAERUV retrievals with quality flags of 0. However, even retrievals including supposedly reliable AODs and SSAs are sometimes affected by cloud contamination. The quantified aerosol effects are most valid under clear-sky conditions, but we find they can be informative at cloud fractions up to 0.1, 0.05, and 0.15

for smoke, dust, and sulfate aerosols, respectively. We recommend using these criteria when quantifying aerosol effects on HCHO retrievals. The interaction between aerosols and clouds is complicated, requiring more detailed analysis to characterize their effects independently.

As a part of a future study, we plan to improve the aerosol properties by implementing non-spherical dust aerosols. In addition, given the low frequency availability of the OMAERUV retrievals, additional aerosol measurements or models, such as the Modern-Era Retrospective analysis for Research and Applications (MERRA) (Rienecker et al., 2011), could be utilized in future research for a more complete global analysis of the aerosol effect on OMI HCHO retrievals. Finally, we plan to implement an error budget into OMI HCHO operational retrievals using the aerosol correction results from this study. This explicit aerosol correction can also be applied to quantify the aerosol effects on other OMI products such as NO_2 and water vapor (H_2O).

Acknowledgments

This study is supported by the NASA Atmospheric Composition Program/Aura Science Team awards NNX11AE58G and NNX17AH47G. We are grateful to the OMI science teams for their efforts in providing the satellite data. We would like to thank Christopher Chan Miller at Harvard-Smithsonian Center for Astrophysics and Robert J.D. Spurr at RT Solutions, Inc. for supporting the radiative transfer model. Data are available online from https://doi.org/10.5281/zenodo.3361858.

References

Abbot, D. S., Palmer, P. I., Martin, R. V., Chance, K. V., Jacob, D. J., & Guenther, A. (2003). Seasonal and interannual variability of north american isoprene emissions as determined by formaldehyde column measurements from space. *Geophysical Research Letters*, *30*(17).

Acarreta, J., De Haan, J., & Stammes, P. (2004). Cloud pressure retrieval using the O_2-O_2 absorption band at 477 nm. *Journal of Geophysical Research: Atmospheres*, *109*(D5).

Ahn, C., Torres, O., & Jethva, H. (2014). Assessment of omi near-uv aerosol optical depth over land. *Journal of Geophysical Research: Atmospheres*, *119*(5), 2457–2473.

Barkley, M. P., De Smedt, I., Van Roozendael, M., Kurosu, T. P., Chance, K., Arneth, A., ... Mao, J. (2013). Top-down isoprene emissions over tropical south america inferred from SCIAMACHY and OMI formaldehyde columns. *Journal of Geophysical Research:*

Atmospheres, *118*(12), 6849–6868.

Bey, I., Jacob, D. J., Yantosca, R. M., Logan, J. A., Field, B. D., Fiore, A. M., ... Schultz, M. G. (2001). Global modeling of tropospheric chemistry with assimilated meteorology: Model description and evaluation. *Journal of Geophysical Research: Atmospheres*, *106*(D19), 23073–23095.

Boersma, K., Eskes, H., & Brinksma, E. (2004). Error analysis for tropospheric NO_2 retrieval from space. *Journal of Geophysical Research: Atmospheres*, *109*(D4).

Boersma, K., Eskes, H., Dirksen, R., Van Der A, R., Veefkind, J., Stammes, P., ... Zhou, Y. (2011). An improved tropospheric NO_2 column retrieval algorithm for the Ozone Monitoring Instrument. *Atmospheric Measurement Techniques*, *4*(9), 1905–1928.

Castellanos, P., Boersma, K., Torres, O., & De Haan, J. (2015). OMI tropospheric NO_2 air mass factors over south america: effects of biomass burning aerosols. *Atmospheric Measurement Techniques*, *8*(9), 3831–3849.

Chance, K. (1998). Analysis of BrO measurements from the global ozone monitoring experiment. *Geophysical Research Letters*, *25*(17), 3335–3338.

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

Chimot, J., Vlemmix, T., Veefkind, J., De Haan, J., & Levelt, P. (2016). Impact of aerosols on the OMI tropospheric NO_2 retrievals over industrialized regions: how accurate is the aerosol correction of cloud-free scenes via a simple cloud model? *Atmospheric Measurement Techniques*, *9*(2), 359.

De Smedt, I., Müller, J.-F., Stavrakou, T., Van Der A, R., Eskes, H., & Van Roozendael, M. (2008). Twelve years of global observation of formaldehyde in the troposphere using GOME and SCIAMACHY sensors. *Atmospheric Chemistry and Physics Discussions*, *8*(2), 4947–4063.

De Smedt, I., Stavrakou, T., Hendrick, F., Danckaert, T., Vlemmix, T., Pinardi, G., ... Van Roozendael, M. (2015). Diurnal, seasonal and long-term variations of global formaldehyde columns inferred from combined OMI and GOME-2 observations. *Atmospheric Chemistry and Physics*, *15*(21), 12519–12545.

De Smedt, I., Theys, N., Yu, H., Danckaert, T., Lerot, C., Compernolle, S., ... others (2018). Algorithm theoretical baseline for formaldehyde retrievals from S5P TROPOMI and from the QA4ECV project.

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

Gassó, S., & Torres, O. (2016). The role of cloud contamination, aerosol layer height and aerosol model in the assessment of the omi near-uv retrievals over the ocean. *Atmospheric Measurement Techniques*, *9*(7).

Ginoux, P., Chin, M., Tegen, I., Prospero, J. M., Holben, B., Dubovik, O., & Lin, S.-J. (2001). Sources and distributions of dust aerosols simulated with the GOCART model. *Journal of Geophysical Research: Atmospheres*, *106*(D17), 20255–20273.

González Abad, G., Liu, X., Chance, K., Wang, H., Kurosu, T., & Suleiman, R. (2015). Updated smithsonian Astrophysical Observatory Ozone Monitoring Instrument (SAO OMI) formaldehyde retrieval. *Atmospheric Measurement Techniques*, *8*(1), 19–32.

González Abad, G., Vasilkov, A., Seftor, C., Liu, X., & Chance, K. (2016). Smithsonian Astrophysical Observatory Ozone Mapping and Profiler Suite (SAO OMPS) formaldehyde retrieval. *Atmospheric Measurement Techniques*, *9*(7), 2797–2812.

Hess, M., Koepke, P., & Schult, I. (1998). Optical properties of aerosols and clouds: The software package (OPAC). *Bulletin of the American meteorological society*, *79*(5), 831–844.

Houweling, S., Dentener, F., & Lelieveld, J. (1998). The impact of nonmethane hydrocarbon compounds on tropospheric photochemistry. *Journal of Geophysical Research: Atmospheres*, *103*(D9), 10673–10696.

Ingmann, P., Veihelmann, B., Langen, J., Lamarre, D., Stark, H., & Courrèges-Lacoste, G. B. (2012). Requirements for the GMES atmosphere service and ESA's implementation concept: Sentinels-4/-5 and-5p. *Remote Sensing of Environment*, *120*, 58–69.

Jethva, H., & Torres, O. (2012). Satellite-based evidence of wavelength-dependent aerosol absorption in biomass burning smoke inferred from ozone monitoring instrument.

Jethva, H., Torres, O., & Ahn, C. (2014). Global assessment of omi aerosol single-scattering albedo using ground-based aeronet inversion. *Journal of Geophysical Research: Atmospheres*, *119*(14), 9020–9040.

Kim, J., Jeong, U., Ahn, M.-H., Kim, J. H., Park, R., Lee, H., ... Dobber, M. (n.d.). New Era of Air Quality Monitoring from Space: Geostationary Environment Monitoring Spectrometer (GEMS). *Bulletin of the American Meteorological Society*, *8*(00), 00.

Kirchstetter, T. W., Novakov, T., & Hobbs, P. V. (2004). Evidence that the spectral dependence of light absorption by aerosols is affected by organic carbon. *Journal of Geophysical Research: Atmospheres*, *109*(D21).

Kleipool, Q., Dobber, M., de Haan, J., & Levelt, P. (2008). Earth surface reflectance climatology from 3 years of OMI data. *Journal of Geophysical Research: Atmospheres*, *113*(D18).

Kwon, H.-A., Park, R. J., Jeong, J. I., Lee, S., González Abad, G., Kurosu, T. P., ... Chance, K. (2017). Sensitivity of formaldehyde (HCHO) column measurements from a geostationary satellite to temporal variation of the air mass factor in East Asia. *Atmospheric Chemistry and Physics*, *17*(7), 4673–4686.

Leitão, J., Richter, A., Vrekoussis, M., Kokhanovsky, A., Zhang, Q., Beekmann, M., & Burrows, J. (2010). On the improvement of NO_2 satellite retrievals–aerosol impact on the airmass factors. *Atmospheric Measurement Techniques*, *3*(2), 475–493.

Levelt, P. F., van den Oord, G. H., Dobber, M. R., Malkki, A., Visser, H., de Vries, J., ... Saari, H. (2006). The ozone monitoring instrument. *IEEE Transactions on geoscience and remote sensing*, *44*(5), 1093–1101.

Li, C., Joiner, J., Krotkov, N. A., & Dunlap, L. (2015). A new method for global retrievals of HCHO total columns from the Suomi National Polar-orbiting Partnership Ozone Mapping and Profiler Suite. *Geophysical Research Letters*, *42*(7), 2515–2522.

Lin, J.-T., Liu, M., Xin, J., Boersma, K., Spurr, R., Martin, R., & Zhang, Q. (2015). Influence of aerosols and surface reflectance on satellite NO_2 retrieval: seasonal and spatial characteristics and implications for NO_x emission constraints. *Atmospheric Chemistry and Physics*, *15*(19), 11217–11241.

Lin, J.-T., Martin, R., Boersma, K., Sneep, M., Stammes, P., Spurr, R., ... Irie, H. (2014). Retrieving tropospheric nitrogen dioxide from the Ozone Monitoring Instrument: effects of aerosols, surface reflectance anisotropy, and vertical profile of nitrogen dioxide. *Atmospheric Chemistry and Physics*, *14*(3), 1441–1461.

Liu, M., Lin, J., Boersma, K. F., Pinardi, G., Wang, Y., Chimot, J., ... others (2019). *Atmospheric Measurement Techniques*, *12*(1), 1–21.

Liu, X., Chance, K., & Kurosu, T. (2007). Improved ozone profile retrievals from GOME data with degradation correction in reflectance. *Atmospheric Chemistry and Physics*, *7*(6), 1575–1583.

Liu, X., Chance, K., Sioris, C., Spurr, R., Kurosu, T., Martin, R., & Newchurch, M. (2005). Ozone profile and tropospheric ozone retrievals from the Global Ozone Monitoring Experiment: Algorithm description and validation. *Journal of Geophysical Research: Atmospheres*, *110*(D20).

Lorente, A., Boersma, K. F., Yu, H., Dörner, S., Hilboll, A., Richter, A., ... Krol, M. (2017). Structural uncertainty in air mass factor calculation for NO_2 and HCHO satellite retrievals. *Atmospheric Measurement Techniques*, *10*, 759–782.

Marais, E., Jacob, D., Kurosu, T., Chance, K., Murphy, J., Reeves, C., ... Mao, J. (2012). Isoprene emissions in Africa inferred from OMI observations of formaldehyde columns. *Atmospheric Chemistry and Physics*, *12*(14), 6219–6235.

Martin, R. V., Chance, K., Jacob, D. J., Kurosu, T. P., Spurr, R. J., Bucsela, E., ... Koelemeijer, R. B. (2002). An improved retrieval of tropospheric nitrogen dioxide from GOME. *Journal of Geophysical Research: Atmospheres*, *107*(D20), ACH–9.

Martin, R. V., Jacob, D. J., Chance, K., Kurosu, T. P., Palmer, P. I., & Evans, M. J. (2003). Global inventory of nitrogen oxide emissions constrained by space-based observations of NO_2 columns. *Journal of Geophysical Research: Atmospheres*, *108*(D17).

Mishchenko, M. I., Travis, L. D., Kahn, R. A., & West, R. A. (1997). Modeling phase functions for dustlike tropospheric aerosols using a shape mixture of randomly oriented polydisperse spheroids. *Journal of Geophysical Research: Atmospheres*, *102*(D14), 16831–16847.

Müller, J.-F., Stavrakou, T., Wallens, S., Smedt, I. D., Roozendael, M. V., Potosnak, M., ... Guenther, A. (2008). Global isoprene emissions estimated using MEGAN, ECMWF analyses and a detailed canopy environment model. *Atmospheric Chemistry and Physics*, *8*(5), 1329–1341.

Omaeruv readme file. (2017). Retrieved from `https://aura.gesdisc.eosdis.nasa.gov/data//Aura\_OMI\_Level2/OMAERUV.003/doc/README.OMAERUV.pdf`

Palmer, P. I., Abbot, D. S., Fu, T.-M., Jacob, D. J., Chance, K., Kurosu, T. P., ... others (2006). Quantifying the seasonal and interannual variability of North American isoprene emissions using satellite observations of the formaldehyde column. *Journal of Geophysical Research: Atmospheres*, *111*(D12).

Palmer, P. I., Jacob, D. J., Chance, K., Martin, R. V., Spurr, R. J., Kurosu, T. P., ... Li, Q. (2001). Air mass factor formulation for spectroscopic measurements from satellites: Application to formaldehyde retrievals from the Global Ozone Monitoring Experiment.

Journal of Geophysical Research: Atmospheres, *106*(D13), 14539–14550.

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

Rienecker, M. M., Suarez, M. J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., ... others (2011). MERRA: NASA's modern-era retrospective analysis for research and applications. *Journal of climate*, *24*(14), 3624–3648.

Russell, P., Bergstrom, R., Shinozuka, Y., Clarke, A., DeCarlo, P., Jimenez, J., ... Strawa, A. (2010). Absorption Angstrom Exponent in AERONET and related data as an indicator of aerosol composition. *Atmospheric Chemistry and Physics*, *10*(3), 1155–1169.

Spurr, R., & Christi, M. (2014). On the generation of atmospheric property jacobians from the (v) lidort linearized radiative transfer models. *Journal of Quantitative Spectroscopy and Radiative Transfer*, *142*, 109–115.

Spurr, R. J. (2006). VLIDORT: A linearized pseudo-spherical vector discrete ordinate radiative transfer code for forward model and retrieval studies in multilayer multiple scattering media. *Journal of Quantitative Spectroscopy and Radiative Transfer*, *102*(2), 316–342.

Torres, O., Ahn, C., & Chen, Z. (2013). Improvements to the OMI near-UV aerosol algorithm using A-train CALIOP and AIRS observations. *Atmospheric Measurement Techniques*, *6*(11), 3257–3270.

Torres, O., Bhartia, P. K., Jethva, H., & Ahn, C. (2018). Impact of the ozone monitoring instrument row anomaly on the long-term record of aerosol products. *Atmospheric Measurement Techniques*, *11*(5), 2701–2715.

Torres, O., Tanskanen, A., Veihelmann, B., Ahn, C., Braak, R., Bhartia, P. K., ... Levelt, P. (2007). Aerosols and surface UV products from Ozone Monitoring Instrument observations: An overview. *Journal of Geophysical Research: Atmospheres*, *112*(D24S47).

Wittrock, F., Richter, A., Oetjen, H., Burrows, J. P., Kanakidou, M., Myriokefalitakis, S., ... Wagner, T. (2006). Simultaneous global observations of glyoxal and formaldehyde from space. *Geophysical Research Letters*, *33*(L16804).

Zoogman, P., Liu, X., Suleiman, R., Pennington, W., Flittner, D., Al-Saadi, J., ... Chance, K. (2017). Tropospheric missions: Monitoring of pollution (TEMPO). *Journal of Quantitative Spectroscopy and Radiative Transfer*, *186*, 17–39.