ABSTRACT

Title of Dissertation: REMOTE SENSING OF
ATMOSPHERIC TRACE GASES FROM
SPACEBORNE UV MEASUREMENTS

Xinzhou Huang
Doctor of Philosophy, 2023

Dissertation Directed by: Research Scientist Kai Yang
Department of Atmospheric and Oceanic Science
Professor Russell R. Dickerson
Department of Atmospheric and Oceanic Science

Satellite measurements of atmospheric trace gases provide continuous long-term information for monitoring the atmospheric chemical environment and air quality at local, regional, and global scales. Trace gas retrievals play a critical role in chemical data assimilation, air quality modeling and forecast, and regulatory decision-making. In this dissertation, I present retrievals of three trace gases species (O_3, SO_2, and NO_2) from measurements of ultraviolet (UV) radiation made from the imaging spectrometers onboard operational satellites, including the Earth Polychromatic Imaging Camera (EPIC) onboard the Deep Space Climate Observatory (DSCOVR), the Ozone Mapping and Profiler Suite - Nadir Mapper (OMPS-NM) onboard Suomi-NPP (SNPP), and the OMPS-NM onboard NOAA-20 satellite. The retrievals of the trace gas vertical columns are achieved through the Direct Vertical Column Fitting (DVCF) algorithm, which is designed to maximize the absorption signature from the Earth's atmosphere in the UV spectral range.

This dissertation first demonstrates the theoretical basis and mathematical procedures of the DVCF algorithm used for retrieving total vertical columns of ozone (O_3) and sulfur dioxide (SO_2) from DSCOVR EPIC. We describe algorithm advances, including an improved O_3 profile representation that enables profile adjustments from multiple spectral measurements and the spatial optimal estimation (SOE) scheme that reduces O_3 artifacts resulted from EPIC's band-to-band misregistrations. Furthermore, we present detailed error analyses to quantify retrieval uncertainties from various sources, assess EPIC observed volcanic plumes, and validate O_3 and SO_2 retrievals with correlative data.

The second part of this dissertation presents a suite of efforts to retrieve the tropospheric and stratospheric NO_2 vertical columns from the new NOAA-20 OMPS hyperspectral ultraviolet-visible (UV-Vis) instrument, covering retrieval algorithm, Stratosphere-Troposphere Separation (STS) scheme, measurement sensitivity assessment, inter-comparison with the Ozone Monitoring Instrument (OMI), evaluation with ground-based Pandora spectrometers, as well as a case study of drastic NO_2 changes during COVID-19 pandemic.

The third part of my dissertation focuses on validation and algorithm improvements for the tropospheric NO_2 retrievals from SNPP OMPS UV measurements. OMPS column NO_2 was validated against coincidence measurements from two ground-based MAX-DOAS spectrometers deployed in eastern China. To achieve higher retrieval accuracy, we developed and implemented a series of algorithm improvements, including an explicit aerosol correction scheme to account for changes in measurement sensitivity caused by aerosol scattering and absorption, the replacement of climatological a priori NO_2 profile with more accurate NO_2 vertical distribution from high-resolution CMAQ model simulations, and the application of model-derived spatial weighting kernel to account for the effect of heterogeneous subpixel distribution. These improvements yield

more accurate OMPS NO_2 retrievals in better agreement with MAX-DOAS NO_2 measurements. The analysis concluded that explicit aerosol correction and a priori profile adjustment are critical for improving satellite NO_2 observations in highly polluted regions and spatial downscaling is helpful in resolving NO_2 subpixel variations.

REMOTE SENSING OF ATMOSPHERIC TRACE GASES FROM SPACEBORNE UV MEASUREMENTS

by

Xinzhou Huang

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2023

Advisory Committee:
  Professor Russell R. Dickerson, Chair/Co-Advisor
  Research Scientist Dr. Kai Yang, Co-Chair/Advisor
  Professor Zhanqing Li
  Research Scientist Dr. Juying X. Warner
  Associate Professor Dongdong Wang

Preface

Materials presented in Chapters 1 to 6 of this dissertation are adapted from two published articles I lead. Materials presented in Chapters 7 and 8 are expected to be submitted soon. Please refer to the list of publications with corresponding Chapters below.

- Chapters 1 to 4

  Huang, X., Yang, K., 2022. Algorithm Theoretical Basis for Ozone and Sulfur Dioxide Retrievals from DSCOVR EPIC. Atmos. Meas. Tech., 15, 5877–5915, 2022. https://doi.org/10.5194/amt-15-5877-2022.

- Chapters 5 and 6

  Huang, X., Yang, K., Kondragunta, S., Wei, Z., Valin, L., Szykman, J., Goldberg, M., 2022. NO_2 retrievals from NOAA-20 OMPS: Algorithm, evaluation, and observations of drastic changes during COVID-19. Atmos. Environ. 290, 119367. https://doi.org/10.1016/j.atmosenv.2022.119367.

- Chapters 7 and 8

  Huang, X., Yang, K., He, H., Wang, Y., Wagner, T., Chang, X., Zhao B., Wang, S., Dickerson, R. R., 2023. Tropospheric NO_2 from Suomi-NPP OMPS: validation and algorithm improvements. Atmos. Meas. Tech., to be submitted.

I have conducted an extensive amount of original investigations to complete my Ph.D. research. In the following, I will highlight some of the new discoveries that were presented in the dissertation.

For many remote sensing retrieval algorithms, radiance matching between forward modeling and measurement is needed for accurate retrievals. In Chapter 1, we demonstrated that forward modeling needs to properly simulate the atmosphere's photon sampling (or photon mean paths), in addition to radiance matching. This discovery suggests that many radiance correction schemes attempted before are bound to fail because they did not consider improving the photon mean path representation in the correction scheme. This finding points out the direction of future improvements in remote sensing retrieval algorithms. In Chapter 2, we presented a new ozone profile representation scheme that allows for first-time profile retrieval with a small number of spectral measurements. Furthermore, we developed a novel spatial optimal estimation technique, which was applied to the EPIC measurements for artifact correction.

The other Chapters focus on the actual retrieval procedure (including algorithm & algorithm improvements we have developed), the error analysis, as well as the validation of the retrieval products. These results elucidate the new knowledge and skills that my dissertation has developed to advance trace gas retrievals from spaceborne UV measurements.

Dedication

My dissertation is dedicated to my father, Guangyin Huang, and my mother, Longlin Zhou, for being such an incredible and supportive family and for always believing in me.

I owe a debt of gratitude to my father, who has instilled in me the value of hard work and dedication through his exemplary work ethic. His determination and perseverance inspired me to always keep motivated and achieve every step of my goals.

I am extremely thankful to my mother. Her adventurous spirit has inspired me to take on new challenges and never be self-limiting. My mom is always an optimist. She has taught me to maintain a positive outlook and approach life with a can-do attitude even when things get tough. Her favorite quote, “If others can do it, you can too,” has been a guiding force behind many of the first times in my life. I find that adopting this “yes I can” attitude empowers me to explore the unknowns of research and overcome obstacles throughout my graduate school experience. Thank you Mom for your words of wisdom and encouragement.

My dedication also goes out to my husband and best friend, Shuo Sun, for always standing by my side and cheering me up during those down moments of my graduate school journey. Meeting and getting to know you is the best thing that happened to me at UMD. The past few years haven’t been easy as we navigated through three years of a cross-country (East-to-West coast) long-distance relationship, and followed by over one year of a cross-continent (US-China) ultra-long-distance relationship due to COVID-19 travel bans. Despite these challenges, I am grateful that our love has not faded and we have been supporting each other in every possible way. This dissertation would not have been possible without your unwavering love and support.

Acknowledgments

First of all, I would like to express my gratitude to my research advisor Dr. Kai Yang, who has supported me with his knowledge, guidance, trust, and encouragement since day one I started doing research with him. I do not have a typical graduate-school trajectory like most of my peers. I switched research direction at a late stage of my first program at UMD. Coming from a background as an atmospheric modeler, steering my way in a new field can be exciting and difficult at times. I have learned a lot from Kai throughout the past five years, from his solid knowledge of atmospheric physical processes to his hands-on skills for problem-solving. His positive attitude, critical thinking, and his great passion for research have inspired me to go deeper into this research direction. I would also like to thank Prof. Russell Dickerson for his contribution as my academic advisor and his guidance that ultimately helped me to complete my dissertation. I am very grateful for the amazing advisors I have at UMD.

I would like to take this opportunity to acknowledge our collaborators, Dr. Zigang Wei and Dr. Shobha Kondragunta at NOAA, and Dr. Lukas Valin and Dr. James Szykman at EPA for providing input and feedback to my NOAA-20 research project. I would like to thank Dr. Yang Wang and Prof. Thomas Wagner at the German Aerospace Center for providing the MAX-DOAS observations used as reference data in my SNPP project. I would also like to thank our fellow modeler collaborators, Dr. Hao He at UMD, Prof. Bin Zhao, Dr. Xing Chang, and Prof. Shu-Xiao Wang at Tsinghua University for providing the high-resolution CMAQ model simulations to aid in the retrieval algorithm development for the SNPP data product.

I would then like to thank my committee members, Prof. Russell Dickerson, Prof. Zhanqing Li, Dr. Juying Warner, and Prof. Dongdong Wang for taking the time to serve on my dissertation committee and providing valuable feedback.

My endless appreciation also goes out to the village of people who have helped me along the way. Thank you to all the teachers, professors, mentors, colleagues, and my dear friends. I would not be where I am today without the support from all of you. I also want to express my gratitude to the AOSC front office staff, particularly Tammy Hendershot, Bernadette Gatewood, and June Sherer, for their assistance whenever I need it.

Lastly, I would like to extend my deepest gratitude to my most incredible family, my mom, dad, and husband, as well as the rest of my family. They are my rock. I could not have done this without their unconditional love and encouragement throughout my entire life.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AE	Aerosol Extinction
AI	Aerosol Index
AK	Averaging Kernel
AMF	Air Mass Factor
AOD	Aerosol Optical Depth
BDM	Brion-Daumont-Malicet Ozone Cross-Sections
BIO	carbonaceous aerosols from BIOmass combustion
BRDF	Bidirectional Reflectance Distribution Function
BW	Birk and Wagner
CCD	Charged-Coupled Device
CMAQ	Community Multiscale Air Quality
CNY	Chinese New Year holiday
CONUS	Continental United States
DF	Direct Fitting
DOAS	Differential Optical Absorption Spectroscopy
DOY	Day Of Year
dSCD	the differential slant column density
DSCOVR	The Deep Space Climate Observatory
DST	Mineral Dust
DVCF	Direct Vertical Column Fitting
DU	Dobson Units
EOF	Empirical Orthogonal Function
EPA	Environmental Protection Agency
EPIC	The Earth Polychromatic Imaging Camera
ENVISAT	ESA's ENVIronmental SATellite
FOV	Field Of View
GDF	Generalized Distribution Function
GEO	Geostationary Earth Orbit

GLER	Geometry-dependent Lambertian-equivalent reflectivity
GOME	Global Ozone Monitoring Experiment
GOME-2	Global Ozone Monitoring Experiment–2
IFOV	Instantaneous Field Of View
ISRF	Instrument Spectral Response Function
IR	Infrared
JPSS	Joint Polar Satellite System
LEO	Low Earth Orbit
LER	Lambertian-equivalent reflectivity
LIDORT-RRS	Linearized Discrete Ordinate Radiative Transfer model with RRS
LLM	Labow–Logan–McPeters Ozone Profile climatology
LT	Local Time
M2TCO3	the total-column-dependent climatology
MAX-DOAS	Multi-AXis Differential Optical Absorption Spectroscopy
MERRA-2	the Modern-Era Retrospective Analysis for Research and Applications version 2
MLER	Mixed Lambertian-equivalent reflectivity
Mt	Million Tons
NASA	National Aeronautics and Space Administration
NOAA	National Oceanic and Atmospheric Administration
OCP	Optical Centroid Pressure
OE	Optimal Estimation
OLS	Ordinary Least Square
OMI	Ozone Monitoring Instrument
OMPS	Ozone Mapping and Profiler Suite
OMPS-NM	Ozone Mapping and Profiler Suite - Nadir Mapper
PGN	the Pandonia Global Network
PM	Particulate Matter
ppbv	parts per billion by volume
RAA	Relative Azimuthal Angle
RCF	Radiance Contribution Function
RRS	Rotational Raman Scattering
RT	Radiative Transfer

S5P	Sentinel-5 Precursor
SCIAMACHY	Scanning Imaging Absorption Spectrometer for Atmospheric Cartography
SLF	sulfate
SNR	Signal-to-Noise Ratio
SMOKE	the Sparse Matrix Operator Kernel Emissions
SOE	Spatial Optimal Estimation
SP	Standard Product
SSA	Single Scattering Albedo
STS	Stratosphere-Troposphere Separation
SVCDs	Stratospheric Vertical Column Densities
SZA	Solar Zenith Angle
TOA	Top Of the Atmosphere
TOMRAD	Radiative Transfer Code based Dave’s successive iteration method
TOMS	Total Ozone Mapping Spectrometer
TOMS-V8	TOMS Version 8
TROPOMI	the TROPOspheric Monitoring Instrument
TVCDs	Tropospheric Vertical Column densities
UTLS	Upper Troposphere and Lower Stratosphere
UV	Ultraviolet
UV-Vis	Ultraviolet-visible
VLIDORT	Vector Linearized Discrete Ordinate Radiative Transfer model
VOC	Volatile Organic Compounds
VZA	Viewing Zenith Angle
WRF	Weather Research and Forecasting

Chapter 1: Introduction

1.1 Background

Ozone (O_3), sulfur dioxide (SO_2), and nitrogen dioxide (NO_2) are important trace gases in the Earth's atmosphere affecting human health and climate (Lelieveld et al., 2015; Seinfeld and Pandis, 2016).

O_3 is a highly reactive trace gas that occurs in both Earth's stratosphere and troposphere. In the stratosphere, the ozone layer is vitally important to life because it absorbs biologically harmful UV radiation coming from the Sun. Ozone received much public attention when in the 1980s its enormous reduction was observed during Antarctic spring (Farman et al., 1985). This so-called ozone hole is the result of human-made chlorofluorocarbon compounds and the recovery of the ozone hole is monitored continuously. In the troposphere, O_3 is a secondary air pollutant formed primarily from the photochemical reactions between the volatile organic compounds (VOC) and nitrogen oxides (NO_x). High ambient ozone concentrations are typically observed in the summer months when sunlight and heat are the most intense. Tropospheric ozone causes premature aging of the lungs (Bell et al., 2004) and stunts the growth of plants (Sandermann Jr, 1996) under high concentrations. To protect human health and agriculture, the US Environmental Protection Agency (EPA) limited ambient ozone to an 8 hr daily maximum mixing ratio of 70 parts per billion by volume (ppbv) in 2015. The tropospheric ozone levels are determined by emissions of ozone precursors, atmospheric photochemistry, and transport (Jacob et al., 1993). Understanding the nonlinear relationship between ozone production and its precursors is critical for the development of an effective ozone control strategy (Sillman, 1999).

SO_2 is emitted by both volcanoes and anthropogenic activities. It is a major air pollutant and a precursor to sulfate aerosols. Sulfates have long residence time (up to a few years), depending mostly on their altitudes in the atmosphere. They can be deposited back to the ground where they have an adverse impact on the environment or reside in the atmosphere as aerosols and affect radiative forcing, either directly by reflecting the incoming solar radiation or indirectly by changing the cloud albedo and lifetime (Bréon et al., 2002; Gassó, 2008). Anthropogenic sources of SO_2 come from the burning of sulfur-rich fossil fuels such as coal and petroleum, and they can also be produced from the smelting of ores. Thanks to new technologies such as flue gas desulfurization and sulfur scrubbers implemented in coal-fired power plants, the level of sulfur pollution can be greatly reduced. While anthropogenic SO_2 emissions have gone down in recent years (Li et al., 2017a; Zheng et al., 2018), natural sources of SO_2 from volcanic eruptions are known to influence global or hemispheric climate (Robock, 2000). The strongest eruption in the twentieth century (Pinatubo, June 1991), released about 20 million tons (Mt) of SO_2 directly into the stratosphere and caused an average global cooling of 0.3 - 0.5 °C that lasted for about 2 years (Robock, 2000; Stenchikov et al., 2021). Although eruptions of this magnitude rarely happen, volcanic degassing is continuous and smaller eruptions occur sporadically (a few times in a year) in various regions. The cumulative contributions from volcanic emissions (large and small) over the globe play a major role in regulating the burden of sulfate aerosols in the troposphere and stratosphere (Vernier et al., 2009). In order to understand their impacts on atmospheric chemistry and climate, it is important to measure both the abundance and altitude of these volcanic SO_2 emissions (Yang et al., 2010). In many cases volcanic SO_2 and ash are collocated, thus making SO_2 a useful proxy for the location of the ash plume. This makes routinely volcanic SO_2 monitoring very helpful for aviation hazard mitigation and volcanic plume forecasting (Carn et al., 2009).

Nitrogen oxides ($NO_x = NO_2 + NO$) relate strongly to ozone destruction and halogen compound reactions in the stratosphere (Wennberg et al., 1994) and work as important precursors of ozone and nitrate aerosols in the troposphere (Sillman, 1999). NO is quickly oxidized to NO_2 in the atmosphere via ozone or in the presence of hydroperoxy (HO_2) or organic peroxy radicals (RO_2), on a timescale of seconds, and the photolysis of NO_2 converts NO_2 back into NO. Thus, the NO and NO_2 species are often grouped into a single species called NO_x. NO_2 is a regulated air pollutant toxic to both human health and crop growth (Chauhan et al., 2003; Lobell et al., 2022). There are some natural sources of nitrogen oxides, such as from soil microbial processes (Conrad, 1996), lightning (Ridley et al., 1996), and natural wildfires (Val Martín et al., 2006), but the majority of the NO_2 in the atmosphere today originates from anthropogenic sources related with fossil fuels combustion (van Vuuren et al., 2011). The ultimate sink of tropospheric NO_2 is often nitric acid (HNO_3), a chemical species easily dissolved in the water and responsible for acid rain. During the day, the formation of HNO_3 occurs via the reaction of NO_2 and hydroxyl radical (OH). During the night, heterogeneous hydrolysis of N_2O_5 on the surface of aqueous aerosol particles represents another possibility for NO_x to be removed from the atmosphere and leads to HNO_3 formation (Riemer et al., 2003). The photochemical lifetime of NO_2 is short, which varies from $\sim$2 - 6 hr in summer to $\sim$12 - 27 hr in winter (Beirle et al., 2011; Laughner and Cohen, 2019; Shah et al., 2020). Due to its short lifetime, tropospheric NO_2 concentrations are spatially correlated with local NO_x emissions at spatial scales of $\sim$10 km (Beirle et al., 2019).

Concentrations of atmospheric trace gases can be measured locally by in-situ monitors and detected remotely in an atmospheric column by ground-based or airborne instruments. However, these measurements often only cover a limited region and or during a short time period such as deployed in designated field campaigns. The satellite-based remote sensing technique offers continuous long-term observations of the atmospheric chemical environment with spatial coverage over the entire globe, enabling a wide range of research applications including many not feasible from in-situ, ground-based, and campaign-based airborne measurements.

In Chapter 1, we describe the theoretical basis for satellite remote sensing of these three gaseous light-absorbers and provide detailed mathematical procedures of the direct vertical column fitting (DVCF) algorithm that our retrievals are built upon. We outline in Section 1.4 several applications of the DVCF algorithm as well as the algorithm improvements we developed in this thesis to provide critical data products for three NASA/NOAA satellites.

1.2 Theoretical basis of satellite remote sensing: algorithm physics

In this section, we describe the algorithm theoretical basis of satellite remote sensing using EPIC retrievals of O_3 and SO_2 as an example. These algorithmic basis are general principles that apply to the retrievals of trace gas and other geophysical quantities from any satellite instrument, including the OMPS-NMs NO_2 retrievals elucidated later in this dissertation.

Algorithm physics is a term first used by Chance (2006) to denote the physical processes contributing to the spaceborne measurement of radiance spectra. A measured radiance L_m (in units of $W \cdot sr^{-1} \cdot m^{-2} \cdot nm^{-1}$) from space consists of sunlight photons within a narrow spectral range (typically $< \sim$2 nm), specified by the instrument spectral response function S (ISRF, e.g.,

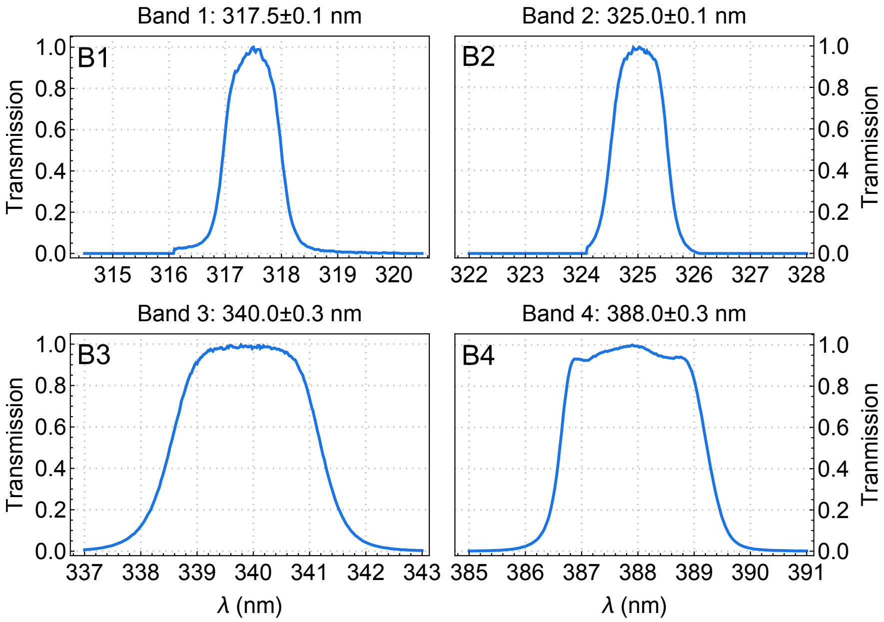


Figure 1.1: Filter transmission functions for the four EPIC UV channels. The widths are ∼1 nm for EPIC bands 1 and 2, similar to those for TOMS and OMPS-NM. Note that the filter transmissions as functions of wavelength are measured in the air (see Figure 1 in Herman et al. 2017). Here we have converted the wavelength in the air to wavelength in vacuum using the formula of Edlén (1966). The filter values are normalized to 1 at band centers (noted on top of each panel with uncertainty).

EPIC UV filter transmissions shown in Fig. 1.1), and is modeled as

$$L_M = \frac{\int S(\lambda) I_{TOA}(\lambda) F(\lambda) d\lambda}{\int S(\lambda) d\lambda}, \tag{1.1}$$

where $F(\lambda)$ (in units of $W \cdot m^{-2} \cdot nm^{-1}$) is the monochromatic spectral solar irradiance, and $I_{TOA}(\lambda)$ the sun-normalized monochromatic top-of-the-atmosphere (TOA) radiance (in units of sr^{-1}) for a wavelength λ (in units of nm). The sun-normalized measured radiance I_M for a spectral band is defined as $I_M = L_M/F_M$, where $F_M = \int S(\lambda)F(\lambda)d\lambda / \int S(\lambda)d\lambda$, and the λ integrations in these equations are performed over the valid range of the ISRF S for the spectral band. Hereafter we drop 'sun-normalized' when referring to I_M, which is simply called measured radiance. Quantities for a spectral band are flux-weighted bandpass averages to account for the differential contributions from individual wavelengths within the bandpass. Without loss of generality, $I_{TOA}(\lambda)$ and other spectral-dependent quantities are hereafter used to denote flux-weighted bandpass averages, with λ representing the characterized wavelength of the spectral band.

To reach a sensor at TOA, sunlight photons are either back-scattered by air molecules or particles or reflected by the underlying Earth surface. As these photons traverse through the atmosphere along many possible optical paths connecting the Sun to the sensor, they may be absorbed by the underlying surface or by some atmospheric constituents, such as trace gases (e.g. O_3 and SO_2) and light-absorbing particles (e.g. dust and smoke). The photons that complete the journey carry information about atmospheric absorbers along their paths. The accumulation of photons from each contributing path yields the TOA radiance, which may be modeled with radiative transfer (RT) simulation if the properties of surface reflection and atmospheric absorption and scattering are known explicitly. The ability to model the TOA radiance accurately is the prerequisite for interpreting the observations and relating the gas absorptions with TOA radiance measurements.

We describe next the characteristics of UV photon sampling of the atmosphere, and the construction of surface and atmospheric models to enable proper simulation of the photon sampling of the atmosphere. Dividing the atmosphere into infinitesimal thin layers, the quantity that specifies the photon sampling is the mean path length of photons traversing through a layer. This mean path length normalized by the geometric thickness of the layer is the local or altitude-resolved air mass factor (AMF, m_z). The proper simulation of photon sampling requires that the modeled mean path length through each layer closely matches that in the actual observing condition.

In theory, a TOA radiance, I_{TOA}, depends on the viewing-illumination geometry, the optical properties of the atmospheric constituents (both absorbers and non-absorbers), and their amounts and vertical distributions, as well as on the reflective properties of the underlying surface. For a wavelength λ, I_{TOA} can be expressed as the sum of two contributions,

$$I_{TOA} = I_a + I_s, \tag{1.2}$$

where I_a consists of solar photons scattered once or more by molecules and particles in the atmosphere without interacting with the underlying surface, and I_s are solar photons reflected at least once or multiple times by the underlying surface.

1.2.1 Path radiance

I_a is also known as the atmospheric path radiance, i.e., photons backscattered to the sensor along a path without any intersection with the underlying surface. Conceptually it is the accumulation of TOA photons that are last backscattered toward the sensor along the line of sight from atmospheric layers at different levels of extinction optical depths. Algebraically it is expressed as the path integration of virtual emission $J(t)$ (Dave 1964) in the direction specified by the view zenith angle (θ_v), attenuated ($e^{-t/\mu}$, where $\mu = \cos\theta_v$) by atmospheric scattering and absorption, over the extinction optical depth t along the path of line of sight from the top ($t = 0$) to the bottom ($t = \tau$) of the atmosphere:

$$I_a = \int_0^{\tau} J(t)\ e^{-t/\mu}\,\omega(t)\ dt/\mu. \tag{1.3}$$

The source of virtual emission, $J(t)$, consists of all the photons scattered towards to the sensor, including photons of the direct solar radiation being scattered once only and photons of diffuse radiation (i.e. photons scattered to level t) being scattered once more at t. The strength of the virtual emission of a thin layer at t is proportional to its scattering optical thickness, which is equal to the product of the layer total optical thickness (dt) and the single scattering albedo $\omega(t)$ (defined as the ratio of layer scattering optical thickness over the layer total optical thickness). Here we use $\Psi(t) = J(t)e^{-t/\mu}\omega(t)/\mu$ to represent the radiance contribution per unit optical thickness to I_a from a layer at t. Eq. (1.3) describes how the solar photons sample the atmosphere from top to bottom and how atmospheric absorption is directly imprinted (via the attenuation $e^{-t/\mu}$) on the path radiance.

A path radiance I_a for a molecular (i.e., an aerosol- and cloud-free) atmosphere with

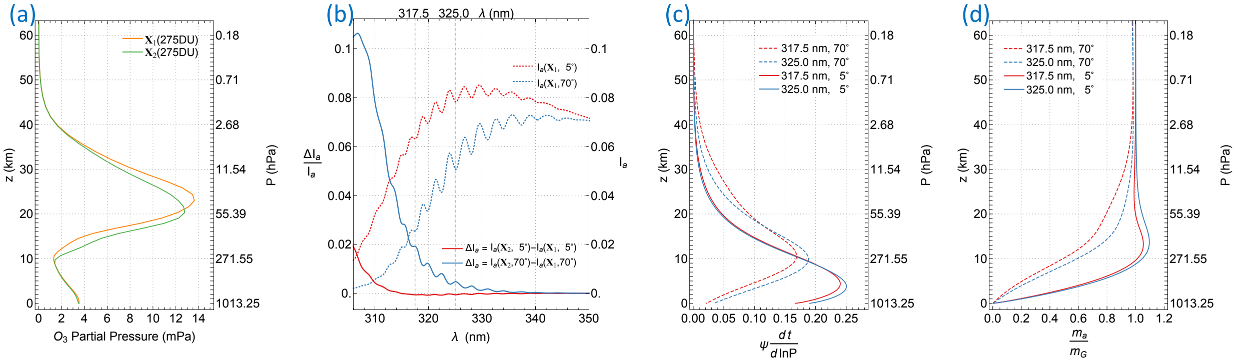


Figure 1.2: Sample results from RT simulations for a molecular atmosphere with O_3 profiles $\mathbf{X}_1$ and $\mathbf{X}_2$ in panel (a). Both $\mathbf{X}_1$ and $\mathbf{X}_2$ are mid-latitude zone ($30^\circ \leq$ latitude $\leq 60^\circ$) climatological O_3 profiles with the same total vertical column of 275 Dobson units, where 1 DU = 2.69×10^{16} molecules/cm^2. RT simulations are performed for two viewing-illumination geometries: 1) low zenith angles, $\theta_s = \theta_v = 5^\circ$ and relative azimuthal angle (RAA), $\phi = 45^\circ$ and 2) high zenith angles, $\theta_s = \theta_v = 70^\circ$ and $\phi = 45^\circ$. (b) Path radiances $I_a(\mathbf{X}_1)$ for the low and high zenith geometries, and their fractional changes ($\Delta I_a/I_a$) when O_3 profile is changed to $\mathbf{X}_2$. (c) Normalized RCFs, ψ for EPIC bands 1 and 2. Here $\psi(t)$ is converted into $\psi(\ln\mathrm{P})$ by the multiplication of factor $dt/d\ln\mathrm{P}$. (d) Mean photon path lengths (m_a) of EPIC bands 1 and 2 as functions of altitude z for the low and high zenith geometries, normalized by the respective geometric air mass factors, m_G.

absorption from trace gases can be accurately determined with RT simulations. For example, the path radiances for the low and high zenith angle geometries (see Fig. 1.2b) are calculated with a vector RT code (e.g., TOMRAD, Dave 1964, or VLIDORT, Spurr 2006) as a function of wavelength for a molecular atmosphere with the O_3 profile $\mathbf{X}_1$ in Fig. 1.2a, and the corresponding radiance contributions to the path radiances at EPIC bands 1 and 2 are shown in Fig. 1.2c. The radiance contribution function (RCF) for a wavelength in the UV range (300 – 400 nm) is determined by Rayleigh scattering and absorption by trace gases (primarily O_3). O_3 is ubiquitous in the atmosphere, with the bulk of it located in the stratosphere (e.g., Fig. 1.2a or Fig. 2.2), and its absorption cross-sections $\sigma(O_3)$ increase rapidly with shorter wavelengths in the UV range (see Fig. 1.10). Rayleigh scattering, whose cross-sections are proportional to $\frac{1}{\lambda^4}$, also increase with shorter wavelength. The strong O_3 absorption and large Rayleigh cross-sections at short wavelengths greatly reduce the number of solar photons reaching the lower atmosphere. Conversely, at longer wavelengths, weaker O_3 absorption and smaller Rayleigh cross-sections allow more solar photons to reach the lower atmosphere where higher air density increases the intensity of backscattering. Similar to the effect of reducing wavelength, lengthening the slant path (by increasing solar or viewing or both zenith angles) would enhance ozone absorption and Rayleigh scattering along the slant path, raising the altitude profile of RCF. These spectral and angular characteristics of RCF are illustrated in Fig. 1.2c, which shows the normalized RCFs ($\psi = \Psi/I_a$) of EPIC bands 1 and 2 for two different observation geometries and a mid-latitude O_3 profile labeled as $\mathbf{X}_1$ in Fig. 1.2a. The results in Fig. 1.2c show that at longer wavelengths and lower zenith angles, path radiance contains more photons that are backscattered from the lower atmosphere. The RCF peak reaches $\sim$4 km altitude for band 2 at 5° zenith angle, while at shorter wavelength and higher zenith angle, the RCF peak moves to the higher altitude, and it rises to $\sim$10 km for band 1 at 70° zenith angle. The shifting shapes of RCF shown in Fig. 1.2c illustrate the changes in the photon sampling of the atmosphere with different wavelengths and zenith angles. The rising RCF peak position signifies diminishing sensitivity to absorptions below the peak while favoring those above it.

The measurement sensitivity to a thin molecular absorber layer is equal to the product of the absorption cross-sections (σ) and the mean path length (m_a) of photons passing through the layer, where $m_a = -\partial \ln I_a / \partial \tau_z$ and τ_z is the absorption optical depth at the layer center altitude z. Note that the photon path length is equal to the geometric AMF, $m_G = 1/\cos(\theta_s) + 1/\cos(\theta_v)$, for a plane-parallel atmosphere if there is no scattering. Figure 1.2d shows the mean optical path lengths of EPIC bands 1 and 2 as a function of altitude for the low and high zenith viewing-illumination geometries, showing that m_a decreases rapidly as the layer descends nearing the surface due to fewer photons reaching the lower atmosphere while m_a approaches m_G as the layer rises towards TOA due to fewer path altering scatterings resulted from lower air density. In the upper troposphere and lower stratosphere (UTLS), m_a of the low zenith geometry usually exceeds m_G due to a significant fraction of photons undergo multiple scattering below and within UTLS, while m_a of the high zenith geometry drops continuously from TOA down to the surface in the case when the RCF peak is sufficiently high that fewer multiple scatterings contribute to the path radiance. In general, the mean path length m_a is shorter for a wavelength with stronger O_3 absorption, which reduces the number of photons reaching the lower atmosphere. The variation of m_a with a changing altitude signifies the path radiance dependence on the absorber profile. The path radiance fractional change due to profile change ($\Delta\mathbf{X} = \mathbf{X}_2 - \mathbf{X}_1$) can be expressed as

$$\frac{\Delta I_a}{I_a} = \frac{I_a(\mathbf{X}_2) - I_a(\mathbf{X}_1)}{I_a(\mathbf{X}_1)} = -\int_0^\infty \Delta X(z)\, \sigma(T_z)\, m_a\, dz, \tag{1.4}$$

where T_z is the atmospheric temperature and $X_1(z)$ and $X_2(z)$ are absorber concentration at altitude z. Figure 1.2b illustrates the change in path radiance caused by a O_3 profile change while keeping its total vertical column the same: lowering the O_3 profile (e.g., $\mathbf{X}_1$ to $\mathbf{X}_2$ in Fig. 1.2a) tend to increase the path radiance. Path radiance changes more with shorter wavelengths at higher zenith angles, thus becoming more sensitive to the shape of the O_3 profile. At low zenith angles, the change may have the opposite sign of the change at large zenith angle for certain wavelengths (e.g., the changes plotted as red solid lines for $\lambda > 316$ nm in Fig. 1.2b), but the magnitude of change is much smaller, indicating the path radiances under these conditions are primarily functions of total columns, since they are less sensitive to the profile shapes. The differential responses of the spectral path radiance to profile changes imply that more than one piece of information about O_3 may be contained in the multi-spectral measurements. Retrieval constrained by multi-spectral radiances instead of a single spectral band may achieve a more accurate O_3 measurement.

1.2.2 Surface reflection

The path radiance I_a includes backscattered photons that are independent of the underlying surface, while the surface contribution to TOA radiance, I_s (referred to as surface radiance hereafter), consists of photons reflected once or more from the surface. For a molecular atmosphere bounded by a surface with well-characterized optical reflection properties, the surface radiance I_s can be accurately predicted with RT modeling. For a Lambertian surface, which reflects radiation isotropically independent of the incident direction, the surface radiance I_s can be expressed as (Dave, 1964)

$$I_s = \frac{T_{\downarrow} r_s T_{\uparrow}}{1 - r_s S_b}, \tag{1.5}$$

where r_s is the reflectance or albedo of the Lambertian surface, $T_{\downarrow}$ is the total (direct and diffuse) transmittance from the Sun to the surface along the direction of incoming solar irradiation and $T_{\uparrow}$ from the surface to the TOA along the viewing direction, and S_b is the atmospheric spherical albedo, which is the fraction of the reflected radiation backscattered from the overlaying atmosphere to the surface. The surface contribution from the Lambertian surface, I_s, may be described as the once-reflected radiance ($T_{\downarrow} r_s T_{\uparrow}$), enhanced by the series of interactions: backscattering from the overlaying atmosphere and reflection from the underlying surface, which are accumulated to produce the amplification factor $1/(1 - r_s S_b)$.

The reflection property of a surface is represented by a bidirectional reflectance distribution function (BRDF), which specifies the angular distribution of reflected radiance as a fraction of directional incident spectral irradiance. Field measurements (Brennan and Bandeen, 1970) demonstrate that the reflection from natural surfaces (such as cloud, water, and land surfaces) are anisotropic in the UV, exhibiting different apparent reflectances when viewed from different directions. For instance, a water surface looks bright when viewed from the direction near the specular reflection, but is much darker outside the glitter (e.g., see Fig. 1.3a). Here the apparent reflectance is the Lambertian-Equivalent reflectivity (LER), i.e., the isotropic reflectance r_s that reproduces the radiance I_s from a surface with an anisotropic BRDF at a viewing-illumination geometry. This LER is also referred to as geometry-dependent surface LER (GLER) to indicate its dependence on the viewing-illumination geometry.

Reflection of UV sunlight from natural surfaces has long been measured by instruments onboard satellites in sun-synchronous polar orbits (e.g. Eck et al., 1987). Since BRDFs for most natural surfaces (except for water surfaces) have not been adequately characterized in the UV, satellite measurements provide scene reflectivities that are quantified with LERs at wavelengths in the range of weak gaseous absorption. To derive LER r_s from a measured radiance I_M, the atmospheric path radiance I_a, transmissions $T_\downarrow$ and $T_\uparrow$, and reflectance S_b for a spectral band are calculated for a molecular atmosphere and the inversion of Eq. (1.5) yields

$$r_s = \frac{I_s}{T_\downarrow T_\uparrow + S_b I_s}, \tag{1.6}$$

where $I_s = I_M - I_a$. A vast majority of scene LERs derived from satellite observations contain contributions from scattering from clouds or aerosols or both (see section 1.2.3 for their treatment). To characterize reflective properties of natural surfaces, many investigations have devoted to creating global LER climatologies by selecting gridded LERs that are minimally affected by clouds or aerosols from the repeated observations over a period of time (typically a calendar month). These climatologies include spectral surface LER databases constructed from the TOMS radiance measurements between 340–380 nm from 1978–1993 (Herman and Celarier, 1997), GOME-1 between 335–772 nm from 1995–2000 (Koelemeijer, 2003), SCIAMACHY between 335–1670 nm from 2002–2012 (Tilstra et al., 2017), OMI between 328–499 nm from 2005–2009 (Kleipool et al., 2008), and GOME-2 between 335–772 nm from 2007–2013 (Tilstra et al., 2017). Inter-comparisons of these spectral LERs from different satellite missions show good agreement among corresponding measurements (Tilstra et al., 2017) despite differences in observation time periods, viewing-illumination geometry, and footprint size. For a location on Earth, its surface is usually observed at nearly the same local solar time from a sun-synchronous orbit, thus the sampling of its surface BRDF is limited to a small range of SZAs. Furthermore, the selection of cloud- and aerosol-free LERs tends to favor low LER values, thus likely excluding the LERs at high VZAs. LER values of natural surfaces tend to be quite close when SZAs fall within a small range and large VZAs are excluded, hence these LER climatologies are presented as independent of viewing-illumination geometry. The low LER sensitivity to varying viewing-illumination geometry (within limited ranges of SZA and VZA) indicates that natural surfaces (excluding glittering water surface) have weak anisotropy and can be treated as Lambertian surfaces. These climatological data reveal that the surface LER in the UV for snow- and ice-free areas vary within the range of 0.02–0.1 for most land and (off-glint) water surfaces, except for a few places on Earth, such as the Saharan desert and the salt flat in Bolivia, where surface LERs may exceed 0.1. These low surface LER values derived from satellite observations have been validated in field experiments (Coulson and Reynolds 1971; Doda and Green 1980, 1981; Feister and Grewe

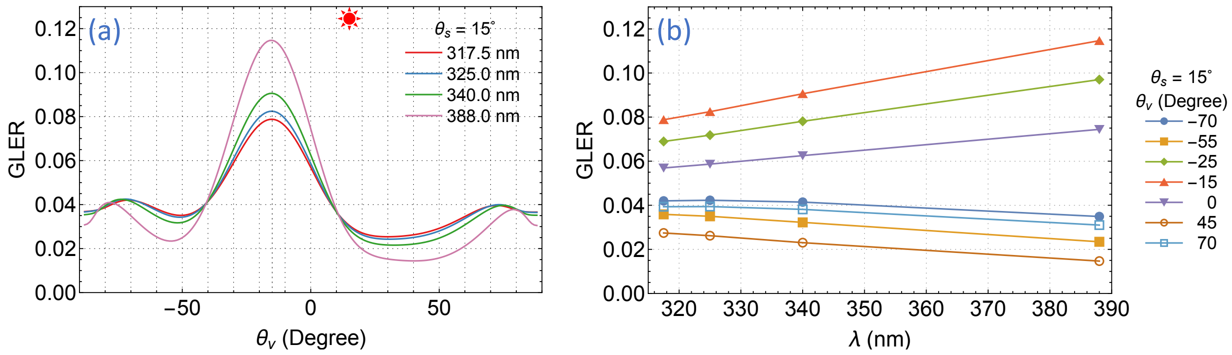


Figure 1.3: Apparent reflectances of an ocean surface, described by a Cox-Munk BRDF (Cox and Munk 1954a,b) for a wind speed of 6 m/s, viewed along the plane of incidence with the Sun at a zenith angle of $\theta_s = 15°$. (a) GLER at four EPIC UV bands vs viewing zenith angle θ_v. Here positive θ_v denotes $\phi = 0°$ and negative θ_v for $\phi = 180°$. (b) GLER at several viewing zenith angles vs. wavelength λ.

1995), which have found that the spectral reflectances of natural surfaces, such as the open ocean, forest, grassland, and desert, fall within the same range of satellite LER measurements. These field experiments have also demonstrated that the spectral reflectances of natural surfaces vary slowly and smoothly with changing wavelengths. The spectrally smooth GLER of natural surfaces permits accurate estimation of GLER within the UV range with measurements at two or more wavelengths, and specifically, the extrapolation of GLERs determined at the long (weak O_3 absorption) wavelengths to estimate the GLERs at short (strong O_3 absorption) wavelengths.

Based on the reflective characteristics of natural surfaces described above, the forward model for retrieval treats the reflections from a surface as Lambertian, whose reflectance is determined from the radiance measurement of the spectral band with weak gaseous absorption or is extrapolated from the weak to the strong absorption band. We use the reflection from an ocean surface as an example to illustrate the success and deficiency of the isotropic surface treatment and the GLER extrapolation, since a water surface is likely the most anisotropic surface encountered in satellite remote sensing. Figure 1.3a displays the GLERs of an ocean surface at the four EPIC UV bands as a function of VZA along the incident plane with the Sun at $\theta_s = 15°$. Viewing in the specular direction ($\theta_v = 15°$ and $\phi = 180°$), the GLER decreases with longer wavelengths but the reverse is true when viewing in directions ~25° or greater away on either side of it. In other words, the reflection appears to be less anisotropic at shorter wavelengths. This is due to less direct beam, thus more diffuse radiation (resulted from more photons are Rayleigh scattered by air molecules) at the shorter wavelengths. While the reflection of a direct beam yields anisotropic outgoing radiation according to the BRDF, the diffuse radiation impinges on the surface from every possible direction of the hemisphere above, usually resulting in a much less anisotropic reflected radiation, which follows the angular distribution specified by the hemispherically averaged BRDF. Figure 1.3b shows the spectral dependence of GLER on wavelength, illustrating that linear extrapolation of GLER at longer wavelengths (340.0 nm and 388.0 nm) yields highly accurate GLER estimations at shorter wavelengths (317.5 nm and 325.0 nm), usually with errors much less than 1%.

The Lambertian surface treatment enables an accurate estimation of the surface radiance I_s without the knowledge of the actual BRDF, provided that the GLERs estimated at some (usually the weak absorbing) wavelengths can be extended (linearly extrapolated) to other wavelengths accurately. However, the paths traversed by photons reflected from a Lambertian surface differ from those from an anisotropic one, as illustrated in Fig. 1.4, which displays the mean optical path lengths, $m_s = -\partial \ln I_s / \partial \tau_z$, of EPIC band 1 as a function of altitude for two viewing-illumination geometries. As shown in Fig. 1.4, the path lengths differ the most just above the surface, but the difference decreases with higher altitudes due to less course-altering atmospheric scattering resulting from lower air density and vanishes around 25 km above the surface. Thus the Lambertian treatment of an anisotropically reflective surface may introduce an error, called

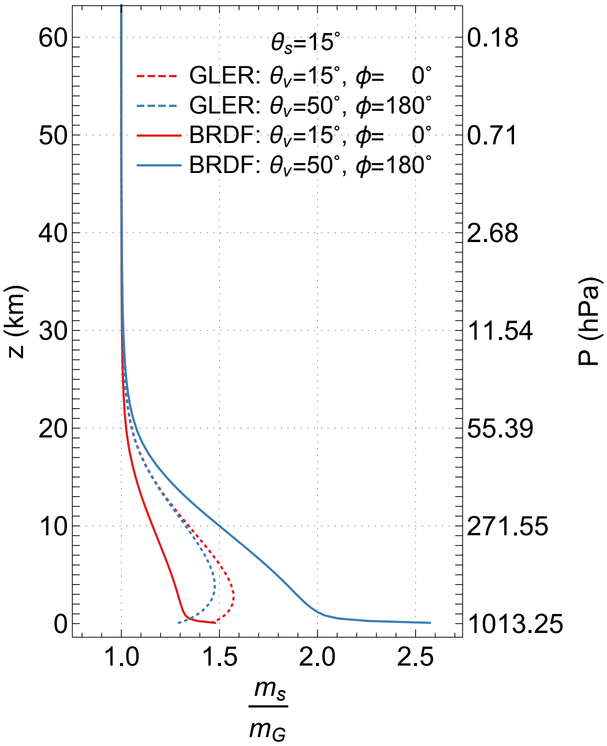


Figure 1.4: Mean path lengths (m_s) of EPIC band 1 reflective photons from an ocean surface (with the same BRDF described in Fig. 1.3) and its Lambertian equivalent surfaces. Here the mean path lengths m_s, normalized by the respective geometric air mass factors (m_G), are plotted as functions of altitude z for two viewing-illumination geometries: one view from the direction of specular reflection, $\theta_v = 15°, \phi = 180°$, and the other at $\theta_v = 50°, \phi = 0°$, while the Sun at $\theta_s = 15°$ for both geometries.

the AMF error, in accounting for atmospheric absorption due to the difference in the photon sampling of the atmosphere. Since this difference is larger in the lower troposphere, but becomes negligible in the stratosphere, implying that the effect of anisotropic reflection, i.e., the BRDF effect, has a larger impact on the quantification of trace gas absorption in the troposphere, but a smaller one for trace gases in the stratosphere. Because the bulk O_3 ($\sim$ 90%) is located in the stratosphere, the Lambertian treatment does not introduce a significant AMF error in total O_3 absorption.

As described above, UV reflectivities for most natural surfaces are quite low (GLER < 0.1), therefore the surface contributions I_s are typically much smaller than (¡ 10% at 317.5 nm) the path radiance I_a (see Fig. 1.5). In modeling a measured radiance I_M, an error in surface radiance I_s is compensated for with the path radiance I_a. The uncertainty of extrapolated GLER is usually less than 1%, corresponding to a less than 1% error in I_s, hence less than 0.1% error in the path radiance I_a. Furthermore, the AMF error due to the Lambertian treatment of an anisotropic surface is insignificant, since the combined mean photon path lengths,

$$m_z = -\partial \ln I_{TOA}/\partial \tau_z = (I_a m_a + I_s m_s)/I_{TOA}, \tag{1.7}$$

contain minor contributions from surface radiance I_s.

Natural surfaces with high UV reflectivities (GLER >0.2) are surfaces covered with snow or ice or both. The highest GLER values are found over Antarctica and Greenland, where typical GLER values are higher than 0.9, as shown in Fig. 1.6). Figure 1.6 shows sample results of a climatological GLER database for Antarctic ice constructed from the observations of polar-orbiting instruments, including Aura OMI and SNPP OMPS, and it reveals a sizeable dependence

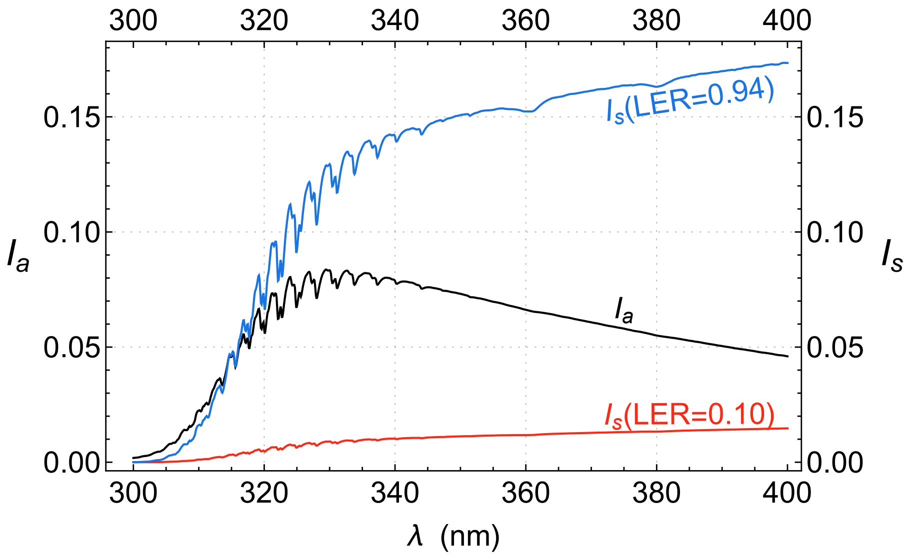


Figure 1.5: Example path radiance, I_a, and surface radiance I_s for $\theta_s = 45°$, $\theta_v = 40°$, and $\phi = 135°$. I_a is the middle line in black, and I_s for LER = 0.1 and LER =0.94 are the lower (red) and upper (blue) lines, respectively.

of ice GLER values on the viewing-illumination geometry, indicating that the reflection from ice is significantly anisotropic. Since the much higher surface radiance I_s (e.g., Fig. 1.5 blue line), the Lambertian treatment of ice surface can lead to large AMF errors. However, the ice GLER varies within a small range (0.94 to 0.98) and hence ice reflection has weak anisotropy for low SZA and VZA ($< 70°$). Because the stronger O_3 absorption and Rayleigh scattering at shorter wavelengths reduce the fraction of direct solar beam but increase that of the diffuse radiation reaching the surface, further weakening the BRDF effect, the error of Lambertian treatment of ice surface in the sampling of atmospheric O_3 absorption is suppressed for the low SZA and VZA observations.

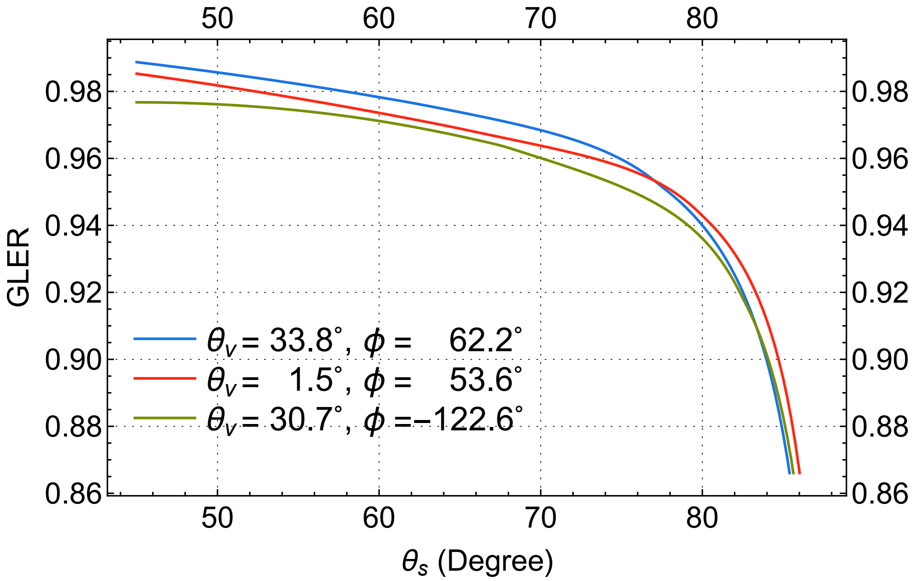


Figure 1.6: Climatological Antarctic GLER values at 331 nm as functions of SZA (θ_s) for three viewing geometries, revealing a significant dependence of ice GLER on the viewing-illumination geometry.

1.2.3 Particle scattering and absorption

Atmospheric particles, including clouds and aerosols, reside mostly in the troposphere and cover a large portion (∼67% by clouds alone, King et al. 2013) of the Earth's surface. Radiative transfer modeling of sunlight through a particle-laden atmosphere can be performed to quantify the TOA contributions from possible light paths, provided that the optical (scattering and absorption) properties of these particles, their amounts, and vertical distributions are specified. However, for UV remote sensing observations, the quantitative information about particles needed for radiative transfer modeling is in general not known sufficiently, precluding their explicit treatment. In this section, we describe an implicit treatment of atmospheric particles for the simulation of measured radiances with the mean photon path approximately matching that through the particle-laden atmosphere.

Atmospheric particles scatter and possibly absorb UV photons, thus can significantly alter their paths through layers from closely above the particles down to the ground surface, usually shortening the path lengths below while lengthening those above the particles. Observing from space, the apparent effect of atmospheric particles is the enhancement of the TOA radiance contributed by backscattering from them. Since this effect is very similar to the consequence of an increased surface albedo, it is often referred to as the albedo effect. The albedo effect can be modeled by placing in a molecular atmosphere an elevated bright surface that partially covers an IFOV. This treatment is called the mixed Lambertian-equivalent reflectivity (MLER) model, which is frequently employed by many algorithms for trace gas retrievals. Based on the MLER model, the TOA radiance for an IFOV is expressed as

$$I_{TOA} = I_g(R_g, p_g)(1 - f_c) + I_c(R_c, p_c)f_c, \tag{1.8}$$

the weighted sum of two independent contributions I_g and I_c. Here I_g is the radiance from the cloud-free portion of the IFOV, containing a Lambertian surface of reflectivity R_g at pressure p_g. Similarly, I_c is from the cloudy portion, and f_c is the cloud fraction and R_c the reflectivity of the Lambertian surface at pressure p_c.

The MLER model can reproduce measured radiances I_m through the determination of cloud fraction f_c. First, the scene LER r_s at surface pressure p_g is estimated using Eq. (1.6). If r_s is less than or equal to the climatological LER value R_g (e.g. Kleipool et al., 2008), this IFOV is treated as particle-free scene ($f_c = 0$). If r_s is greater than or equal to the LER value for cloud $R_c = 0.8$ (Ahmad et al., 2004; Koelemeijer and Stammes, 1999), this IFOV is treated as fully cloud covered ($f_c = 1$). When r_s is in between R_g and R_c, the cloud fraction is inverted

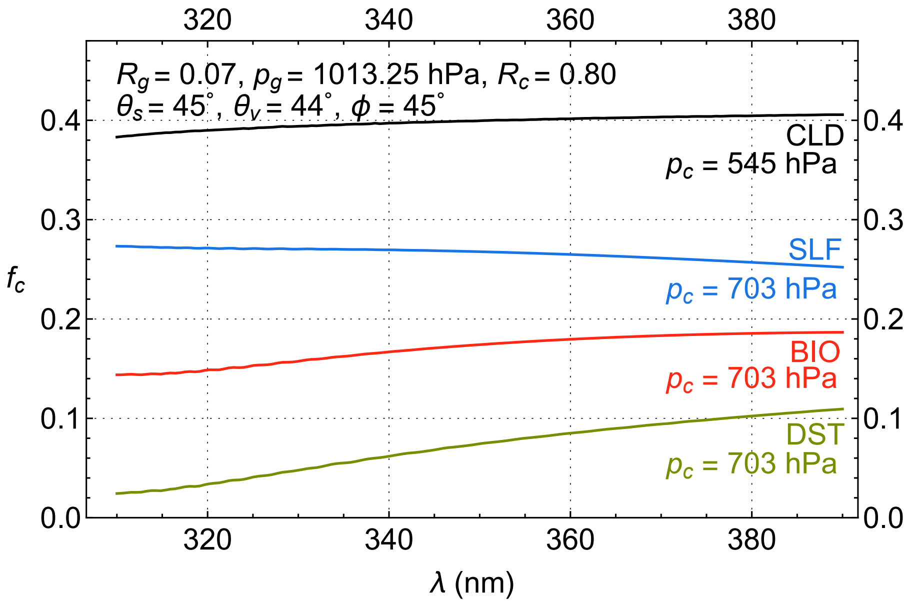


Figure 1.7: Four examples of cloud fractions (f_c) derived from explicitly modeled TOA radiances for particle-laden atmospheres. The first of these is the atmosphere with a 1.5-km-thick layer of C1 cloud (CLD, Deirmendjian 1969) with a single scattering albedo $\omega = 1$ and an optical thickness $\tau = 5$ at 340 nm centered at 5 km altitude (or pressure level of 545 hPa). The others are atmospheres with a 1-km-thick layer of aerosols, including SLF ($\omega = 0.996$), BIO ($\omega = 0.921$), and DST ($\omega = 0.900$) aerosols (SLF, BIO, and DST models are taken from Torres et al. 2007), with an optical thickness $\tau = 1.5$ at 340 nm centered at 3 km altitude (or pressure level of 703 hPa). The insets list the MLER parameters, R_g, p_g, R_c, and p_c, as well as the angles (θ_s, θ_v, and ϕ) that specify the viewing-illumination geometry.

from Eq. (1.8), which yields

$$f_c = \frac{I_M - I_g}{I_c - I_g}. \tag{1.9}$$

In case of $f_c = 0$ or 1, surface LER r_g or cloud LER r_c is determined using Eq. (1.6) to ensure that modeled radiance I_{TOA} is equal to the measurement I_M. Figure 1.7 shows cloud fractions (f_c) as a function of wavelength for several examples of particle-laden atmospheres.

The radiance intensity scattered from atmospheric particles varies with wavelength smoothly without high-frequency spectral structures. For instance, the contributions to TOA radiances (I_{TOA}) from backscattering by meteorological clouds change smoothly and slowly with wavelength (see Fig. 1.7, the CLD curve). The selection of $R_c = 0.8$ facilitates the MLER model to closely simulate the spectral variation of clouds observed from space (Ahmad et al., 2004), such that retrieved f_c has a small spectral variation (i.e., f_c nearly the same for different wavelengths) for most cloudy observations. The small and smooth change of f_c with wavelengths allows its extrapolation to provide a reliable estimate of f_c at shorter wavelengths from those determined at longer wavelengths.

Certain types of aerosols, such as continental aerosols containing soot, smoke from fires, mineral dust from deserts, and ash from volcanic eruptions, both scatter and absorb UV photons passing through them. Usually, aerosol absorptions cause the underlying surface (including clouds) to appear darker, more so at shorter wavelengths. The change in I_{TOA} due to the addition of aerosols and hence the cloud fraction (f_c or the surface LER, r_g) are smooth in wavelength (e.g., see Fig. 1.7, smooth curves for weakly absorbing sulfate-based aerosols (SLF), carbonaceous aerosols from biomass burning (BIO), and mineral dust (DST), Torres et al. 2007). Therefore f_c (when $f_c > 0$, from Eq. 1.9) or r_g (when $f_c = 0$, from Eq. 1.6) determined at longer wavelengths where atmospheric absorption is weak, maybe linearly extrapolated to O_3 sensitive wavelengths for estimation of contributions to TOA radiance from surface reflection and particle backscattering (referred to as the $r_g f_c$ extrapolation method hereafter).

The UV aerosol index (AI) (Herman et al., 1997; Torres et al., 1998), which measures the deviation of spectral variation of TOA radiance from that of a pure molecular atmosphere, is proportional to the spectral slope c_l used in the $r_g f_c$ extrapolation scheme. Algebraically, AI is calculated as the N-value (defined as $-100\ \log_{10} I$) difference between the modeled (I_{TOA}) and the measured (I_M) radiances at a wavelength λ

$$\begin{aligned} AI =& 100 \ \log_{10} \frac{I_M(\lambda)}{I_{TOA}(\lambda, R_e)} && (1.10) \\ =& 100 \ c_l \ \Delta\lambda \ \frac{\partial \log_{10} I_{TOA}(\lambda, R)}{\partial R}\bigg|_{R=R_e}. && (1.11) \end{aligned}$$

Here, the modeled radiance $I_{TOA}(\lambda, R_e)$ is calculated for a molecular atmosphere with an estimated reflectivity parameter R_e, which may be the LER value r_s or the MLER cloud fraction f_c determined at a well-separated wavelength ($\lambda + \Delta\lambda$). The pair of wavelengths used for the AI calculation are in the UV spectral range with weak molecular absorption and their separation $\Delta\lambda$ should be sufficiently large (> 10 nm) to capture the spectral contrast of Rayleigh scattering. Using $I_M(\lambda) = I_{TOA}(\lambda, R_m) = I_{TOA}(\lambda, R_e + \Delta R)$, since the reflectivity parameter R_m is derived from $I_M(\lambda)$ and $\Delta R = R_m - R_e = c_l \ \Delta\lambda$, we arrive at Eq. (1.11) from the definition of AI, Eq. (1.10). In short, the spectral slope c_l is equivalent to the AI, which is significantly positive for particles (such as smoke, dust, and volcanic ash) with large absorption and slightly positive to negative for non-absorbing and weakly absorbing particles (such as clouds and sulfate aerosols). Note that for the conventional AI (a.k.a. LER AI) calculation, radiance I_{TOA} is modeled for a Rayleigh scattering-only atmosphere over a Lambertian surface. To capture the spectral slope of the $r_s f_c$ extrapolation scheme, we switch the LER treatment with the MLER modeling of I_{TOA} for AI calculation. The resulting MLER AI is usually higher than the corresponding LER AI when $f_c > 0$, but otherwise can be similarly used to indicate the presence of UV-absorbing aerosol.

The MLER treatment enables the modeling of measured radiances without the knowledge of the optical properties or the full vertical distributions of atmospheric particles. The accuracy of the modeled radiances at the extrapolated wavelengths depends on how close the MLER

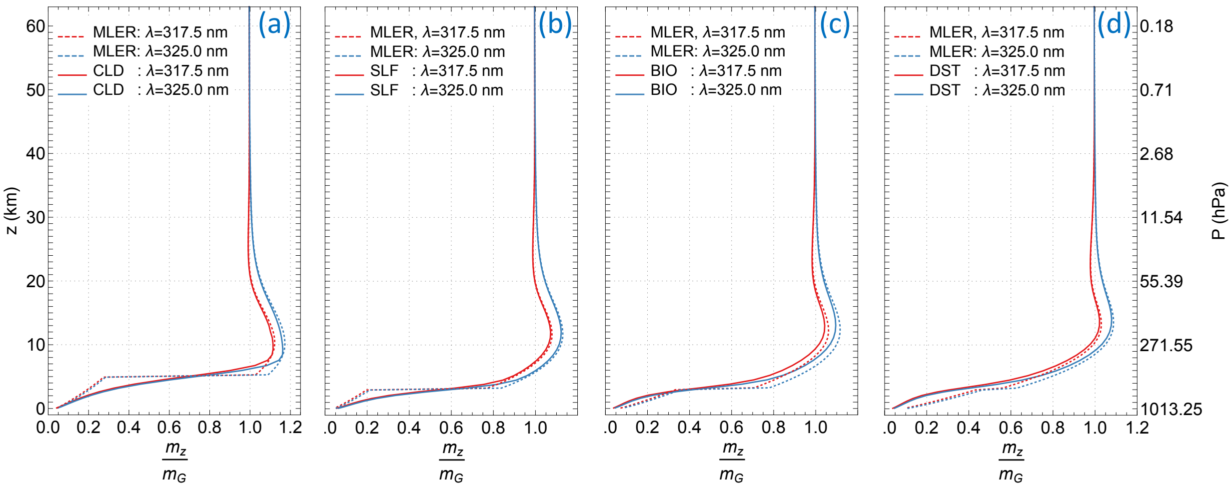


Figure 1.8: Mean photon path lengths m_z, normalized by the geometric AMF m_G, of EPIC bands 1 and 2 as functions of altitude z for particle-laden atmospheres and their MLER treatments. See the caption of Fig. 1.7 for the description of aerosol characterizations, MLER treatments, and the viewing-illumination geometry.

parameter (r_g or f_c) follows the linear relationship among different wavelengths. In reality, the spectral dependence of natural surface reflection (r_g) or particle scattering and absorption (f_c) are nonlinear, though moderately as exemplified in Figs. 1.3b and 1.7, therefore $r_g f_c$ extrapolation yields small errors in r_g or f_c at the extrapolated wavelengths. The radiance uncertainties associated with the $r_g f_c$ extrapolation error are below 1% for the vast majority of remote sensing observations. Higher radiance uncertainties usually occur in the presence of highly elevated or strongly absorbing aerosols. These observations may be flagged with high AI values.

In addition to the mostly small radiance errors at the extrapolated wavelengths, the MLER treatment can simulate the photon sampling of particle-laden atmospheres with a diverse range of particle types and vertical distributions. Figure 1.8 shows comparisons of mean photon path lengths of particle-laden atmospheres with those from the corresponding MLER treatments. These comparisons illustrate that the layer mean photon paths based on the MLER model deviate from those of the particle-laden atmospheres, mostly in the region immediately above the particles down to the underlying surface. These deviations diminish with higher altitudes where lower air density reduces the chance of photons being scattered. Since the vast majority of clouds and aerosols are in the lower troposphere ($<\sim$10 km), the MLER treatment does not introduce significant AMF errors in accounting for O_3 absorption, which occurs mostly in the stratosphere. This is similar to how the Lambertian treatment of surface reflection works for the estimation of total O_3 absorption (see section 1.2.2).

The MLER treatment relies on a few adjustable parameters, including the cloud fraction f_c and cloud pressure p_c, to model a vast range of conditions encountered in remote sensing of Earth's atmosphere. The cloud fraction f_c, obtained directly from radiance measurements using Eq. (1.9), provides an estimate of the cloud amount in an IFOV. The pressure p_c of the elevated Lambertian surface needs to be set at a proper level to best approximate the layer mean photon paths of a particle-laden atmosphere. As seen in Fig. 1.8, the optimal placement of the elevated Lambertian surface is within the particle layer, as p_c locates too high or too low from the optical centroid pressure (OCP) (Joiner and Vasilkov, 2006; Vasilkov et al., 2008) would make layer mean photon paths deviate further from those of the particle-laden atmosphere. The effective cloud pressures retrieved from the EPIC measurements of O_2 A-band (Y.Y.Yang et al. 2019) are usually located within the particle vertical distributions and therefore used to set the cloud pressures p_c for processing EPIC observations.

The use of OCP for p_c enables the MLER model to account for the measurement sensitivity change when a layer of particles is introduced into the atmosphere: enhancing the photon attenuation by absorbers inside and above the layer, while reducing them below, as the mean photon paths or AMFs from the MLER model lengthen above p_c, but shorten below it, as illustrated in Fig. 1.8. Since the MLER model captures the enhancement and shielding effects on trace gas absorption by

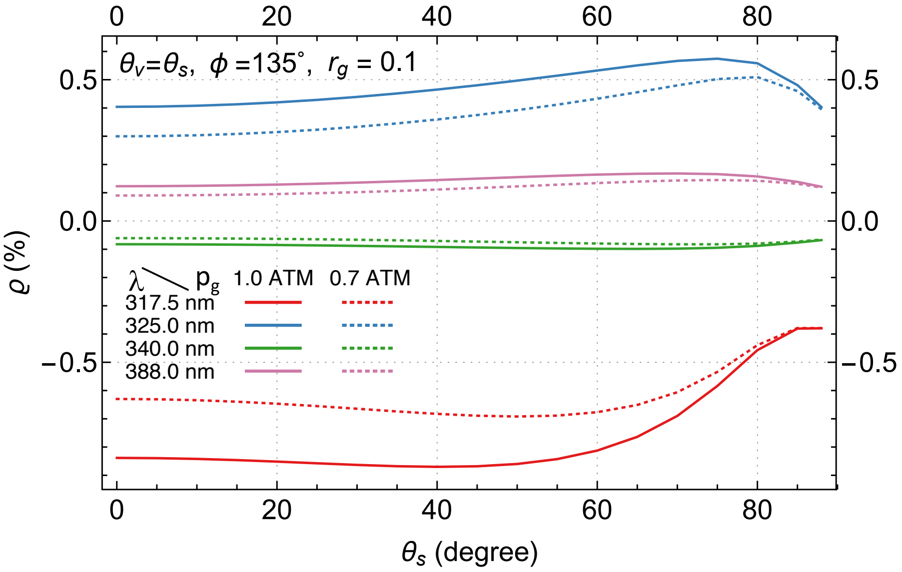


Figure 1.9: Filling-in factors ($\varrho = \frac{I_{RRS}-I_{ELA}}{I_{ELA}} \times 100$) for EPIC UV bands as a function zenith angle at two surface pressures $p_g = 0.7$ ATM and $p_g = 1.0$ ATM, with a surface albedo of $r_g = 0.1$.

atmospheric particles, it is widely adopted due to its simplicity for retrievals of trace gases besides O_3, such as NO_2 and SO_2 in the troposphere. However, sizeable AMF errors are prevalent for modeling tropospheric absorptions based on the MLER treatment, which usually yields significantly different mean photon paths from those of explicit treatment in the troposphere.

1.2.4 Inelastic molecular scattering

The scattering of sunlight with atmospheric constituents is mostly elastic, i.e., the energy and thus the wavelength of a photon remain the same before and after the interaction. But a small portion (~4%) of molecular scattering is inelastic, resulting in energy gain or loss of the scattered photons. Specifically, the rotational Raman scattering (RRS) from air molecules (such as nitrogen and oxygen) can alter the wavelengths of scattered photons, with UV wavelength shifts $\Delta\lambda < \pm$ 2 nm (Chance and Spurr, 1997; Joiner et al., 1995; Vountas et al., 1998). These inelastic scatterings cause the filling-in of telluric lines (i.e., trace gas absorption features) and solar Fraunhofer lines (also known as the Ring effect, which was first noticed by Grainger and Ring 1962).

The filling-in effect is a function of wavelength and depends on the optical properties of the atmosphere, the viewing and illumination geometry, and the surface reflectivity and pressure. The filling-in effect also depends on the ISRF, especially on the instrument spectral resolution, which is the width of its ISRF, since the measured radiance of a band is a convolution of spectral radiance and the ISRF (see Eq. 1.1). This effect is quantified with the filling-in factors, defined as $\varrho = (I_{RRS} - I_{ELA})/I_{ELA}$, where I_{ELA} is the TOA radiance calculated assuming all molecular scattering is elastic, while I_{RRS} includes the inelastic (RRS) contributions. To illustrate the significance of RRS, we show in Fig. 1.9 examples of the filling-in factors, calculated for EPIC bands using the scalar LIDORT-RRS radiative transfer code (Spurr et al., 2008). Since RRS is weakly dependent on polarization, a scalar radiative transfer model, from which both I_{ELA} and I_{RRS} are calculated without including radiation polarization, can accurately provide filling-in factors (Landgraf et al., 2004; Wagner et al., 2010).

The filling-in factors provide estimates of the modeling errors in I_{TOA} when RRS contributions are neglected, and results in Fig. 1.9 show variations of modeling errors with different observing conditions. These errors are usually systematic for a spectral band and are between half to one percent for measurements of EPIC bands 1 and 2. These errors are sufficiently large that corrections are required for achieving high ($\sim$1%) O_3 retrieval accuracy. The filling-in factors (ϱ), modeled using a scaler code (like LIDORT-RRS), may be used to correct the results (I_{TOA}) from vector radiative transfer codes (e.g. Dave, 1964; Spurr, 2006) that perform elastic modeling only, i.e., the RRS corrected TOA radiance $= I_{TOA}(\varrho + 1)$.

1.3 Inversion technique

Section 1.2 describes algorithm physics treatments of interactions of solar radiation with atmospheric particles and surfaces to enable RT modeling of photons traversing through a molecular atmosphere to reproduce the measured TOA radiances with photons that follow the paths similar to those through the actual atmosphere and therefore establish the relationship between spectral measurements and the atmospheric state, as well as surface reflectivity and instrumental parameters. At its core, the RT modeling sets up a forward mapping from the vertical distributions of gaseous absorbers and the surface reflectivity parameters to measured TOA radiances. The retrieval of gas absorbers, such as O_3 and SO_2, is the inverse of this mapping, i.e., to find their vertical distributions and the surface reflectivity parameters for which forward modeling closely reproduces the measured TOA radiances. However, this inverse mapping is inherently an ill-posed problem, as the solution is not unique, i.e., more than one set of profiles and surface parameters can yield the same measurements. This problem is made worse with measurement uncertainties, which expand the profile and surface combinations that can reproduce, within error bars, the measured spectra.

For successful inversion, analytical constraints are placed on the profiles of gas absorbers and the spectral variations of ground reflectivity and atmospheric particle (aerosol and cloud) back-scattering. For O_3 retrieval, Eq. (2.1) embodies the profile constraint, while the MLER model with $r_s f_c$ extrapolation regulates the surface reflection and particle back-scattering. These constraints control the dimension of the inverse mapping space and manifest themselves as the retrieval (i.e., adjustable) parameters, which, in the case of O_3 retrieval, consist of total O_3 column Ω, a number (p) of modal expansion coefficients $\{\gamma_k, k = 1 \ldots p\}$, surface LER (r_s)

or cloud fraction (f_c), and a number (q) of polynomial coefficients $\{c_l, l = 1 \dots q\}$ of the $r_s f_c$ extrapolation. The set of adjustable parameters forms the state vector ($\mathbf{x}$) whose length (n) is the dimension of inverse mapping space. Proper selection of adjustable parameters by limiting the number of the modal coefficients ($p \geq 0$) and the polynomial coefficients ($q \leq 1$) ensures the inverse problem is well-posed and simultaneously maximizes the amount of information collected from the spectral measurements. Here $p = 0$ indicates no modal expansion, equivalent to restricting the profile to a climatological column-dependent O_3 profile, and $q = 0$ for the spectral invariant reflectivity parameter.

1.3.1 Exact solution

Conceptually, the inversion is to find the state vector ($\mathbf{x}$) that satisfies a set of m simultaneous equations, $\{\Delta y_i = 0, i = 1 \dots m\}$, one for each spectral band difference, $\Delta y_i = \ln I_M(\lambda_i) - \ln I_{TOA}(\mathbf{x}, \lambda_i)$, between the radiance measurement I_M and the forward modeling I_{TOA}. Here λ_i the wavelength that characterizes the i^{th} ($1 \leq i \leq m$) spectral band and Δy_i the residual of this band. In matrix form, the m simultaneous equations can be expressed as

$$\Delta \mathbf{y} = 0, \tag{1.12}$$

where $\Delta \mathbf{y}$ is residual column vector $\{\Delta y_i, i = 1 \dots m\}$. Since the forward mapping $I_{TOA}(\mathbf{x})$ is a nonlinear function of the state vector $\mathbf{x}$ and has no analytical inverse, the solution to Eq. (1.12) is usually obtained iteratively. The iteration is started with an initial (i.e., iteration number $L = 0$)

state vector $\mathbf{x}_L$ to linearize the equation between residuals and the state vector

$$\Delta y_i = \ln I_M(\lambda_i) - \ln I_{TOA}(\mathbf{x}_L, \lambda_i) - \sum_{j=1}^{n} (x_j - x_{Lj}) \frac{\partial \ln I_{TOA}(\mathbf{x}, \lambda_i)}{\partial x_j} \Bigg|_{\mathbf{x}=\mathbf{x}_L}, \tag{1.13}$$

where x_j and x_{Lj} are the j^{th} components of $\mathbf{x}$ and $\mathbf{x}_L$ respectively, $\Delta x_j = x_j - x_{Lj}$ the j^{th} components of state adjustment vector, and $K_{ij} = \frac{\partial \ln I_{TOA}(\mathbf{x},\lambda_i)}{\partial x_j}\big|_{\mathbf{x}=\mathbf{x}_L}$ the Jacobian, also known as the weighting function for the retrieval parameter x_j at spectral band λ_i. The m residual elements, each written in Eq. (1.13), can be expressed in matrix form as

$$\Delta \mathbf{y} = \Delta \mathbf{y}_L - \mathbf{K} \Delta \mathbf{x}, \tag{1.14}$$

where $\Delta \mathbf{y}_L$ is the column vector $\{\ln I_M(\lambda_i) - \ln I_{TOA}(\mathbf{x}_L, \lambda_i), i = 1 \ldots m\}$, $\Delta \mathbf{x} = \mathbf{x} - \mathbf{x}_L$ the state adjustment vector, and $\mathbf{K}$ the $m \times n$ Jacobian matrix with the $\{K_{ij}, i = 1 \ldots m, j = 1 \ldots n\}$ as its elements. Putting Eq. (1.14) into Eq. (1.12) yields

$$\Delta \mathbf{y}_L = \mathbf{K} \Delta \mathbf{x}, \tag{1.15}$$

which may be solved exactly (under strict conditions) to determine state adjustment vector $\Delta \mathbf{x}$. After each iteration, the linearization state vector is updated to

$$\mathbf{x}_{L+1} = \mathbf{x}_L + \Delta \mathbf{x}. \tag{1.16}$$

The final state $\mathbf{x}$ is found when the iteration converges, i.e., when the absolute change of state vector $\Delta \mathbf{x}$ is below a threshold.

The linear equation (Eq. 1.15) may be solved exactly only when the number of measurements is equal to the number of retrieval parameters (i.e. $m = n$) and the Jacobian matrix $\mathbf{K}$ is invertible (i.e., non-singular matrix), as exemplified in the well-known TOMS-V8 total O_3 algorithm (Bhartia and Wellemeyer, 2002). The TOMS-V8 algorithm determines the two-component state vector, $\mathbf{x} = \{\Omega, r_s \text{ or } f_c\}$, from radiance measurements of two spectral bands: one with low O_3 sensitivity to estimate the MLER parameter (r_s or f_c), and the other with high O_3 sensitivity to derive total O_3 column Ω. However, few other algorithms adopt this inversion method, since it requires $m = n$ and $\mathbf{K}$ being a nonsingular matrix. Even if both these conditions are met, inverting Eq. (1.15) to obtain exact solutions tends to enhance the impact of measurement uncertainties (noises) on the retrieved results, as in cases that $\mathbf{K}$ matrices are nearly but not quite singular. These cases occur when the spectral variation of a Jacobian has some similarity or a high degree of correlation with that of another retrieval parameter, leading to algorithm difficulty in distinguishing two retrieval parameters corresponding to the two Jacobians, thus yielding unstable retrieval results, such as in the case of simultaneous retrieval of total O_3 and SO_2 columns from EPIC UV measurements.

1.3.2 Direct fitting

Since spectral measurements have errors and $m \neq n$ in general, the inversion is achieved by finding the solution $\mathbf{x}$ that minimizes the cost function

$$\Upsilon(\mathbf{x}) = \left| \mathbf{S}_\epsilon^{-\frac{1}{2}} \Delta\mathbf{y} \right|_2^2 = \Delta\mathbf{y}^T \mathbf{S}_\epsilon^{-1} \Delta\mathbf{y} \tag{1.17}$$

$$= \sum_{i=1}^{m} \left(\frac{\Delta y_i}{\mu_i} \right)^2, \tag{1.18}$$

where $\mathbf{S}_\epsilon$ is the measurement error covariance matrix, with its i^{th} diagonal element equal to μ_i^2. Here μ_i is the fractional standard deviation of radiance error of the i^{th} band. In case of independent measurement error, i.e., no error correlation between different spectral bands, Eq. (1.17) can then be explicitly written as Eq. (1.18), which is the formulation of the least-squares method.

The minimization of the cost function Υ can be started by linearizing the residuals with an initial (i.e., iteration number $L = 0$) state vector $\mathbf{x}_L$. Substituting $\Delta\mathbf{y}$ (Eq. 1.14) into Eq. (1.17), we minimize this cost function to obtain the state adjustment vector

$$\Delta\mathbf{x} = (\mathbf{K}^T\mathbf{S}_\epsilon^{-1}\mathbf{K})^{-1}\mathbf{K}^T\mathbf{S}_\epsilon^{-1}\Delta\mathbf{y}_L = \mathbf{G}_{\mathrm{DF}}\Delta\mathbf{y}_L, \tag{1.19}$$

which is the solution of linear weighted least-square regression. Here, $\mathbf{G}_{\mathrm{DF}} = (\mathbf{K}^T\mathbf{S}_\epsilon^{-1}\mathbf{K})^{-1}\mathbf{K}^T\mathbf{S}_\epsilon^{-1}$ is the direct fitting (DF) gain matrix.

This procedure of iterative minimization of the difference between measurements and modelings to determine the bulk parameters is called the direct vertical column fitting (DVCF) algorithm. The DVCF algorithm is quite general and valid for both discrete-wavelength and hyperspectral measurements, as well as for different types of retrieval parameters, such as MLER parameters, layer partial columns of various absorbing trace gases, and their total vertical columns. This algorithm has been applied to retrievals of total O_3 vertical column (Joiner and Bhartia, 1997; Lerot et al., 2014; Yang et al., 2004), combo of total O_3 and SO_2 vertical columns (Yang et al., 2007, 2009a, 2013), combo of O_3 and altitude-resolved SO_2 vertical columns (Yang et al., 2009b, 2010), and stratospheric and tropospheric NO_2 vertical columns (Yang et al., 2014). This algorithm is named DVCF to contrast with the DOAS (the Differential Optical Absorption

Spectroscopy) method (Platt, 2017), which derives a slant column and then uses an air mass factor (AMF) at a single wavelength (λ_0) to converts it to a vertical column.

In general, the DVCF algorithm works well when the changes in radiance measurements responding to changes in the state vectors are significantly different between any two retrieval parameters, i.e., that columns of $\mathbf{K}$, which are the Jacobians of a retrieval parameter at different wavelengths, exhibit significantly different spectral dependence from one another. This is usually true for any two bulk retrieval parameters over a sufficiently broad spectral range, such as total O_3 column (Ω) and an expansion coefficient (γ_k) of differential profile $\mathbf{e}_k$ (see Eq. 2.1), or the SO_2 vertical column and its layer altitude. With measurements from a broad spectral range, the DVCF algorithm can discriminate subtle spectral features contained in hyperspectral measurements to enhance the retrieval accuracy (e.g., Yang et al. 2009b, 2010). Besides contrasting with the DOAS method, the name DVCF emphasizes vertical column because this algorithm is usually not suitable for traditional profile retrieval, due to the high similarity of partial column Jacobians between adjacent layers and hence the difficulty in distinguishing their partial columns.

1.3.3 Optimal estimation

In many cases, such as sparse spectral sampling or narrow spectral range, the performance of the direct fitting inversion method may decline as a result of limited information contained in the spectral measurements. For stabilizing the retrieved results, the inversion process can be regulated with an additional constraint, which is frequently based on the *a priori* knowledge of the retrieval parameters. Algebraically, adding an *a priori* constraint to Eq. (1.17) yields a new cost function

$$\Upsilon(\mathbf{x}) = \Delta\mathbf{y}^T\mathbf{S}_\epsilon^{-1}\Delta\mathbf{y} + (\mathbf{x}-\mathbf{x}_a)^T\mathbf{S}_a^{-1}(\mathbf{x}-\mathbf{x}_a), \tag{1.20}$$

where $\mathbf{x}_a$ is the *a priori* state vector and $\mathbf{S}_a$ the *a priori* state vector covariance matrix. The first term on the right-hand side (r.h.s) of Eq. (1.20) strives to diminish the difference between measured and modeled radiances, performing the same function as the direct fitting retrieval, while the second r.h.s term seeks to reduce the deviation of retrieved $\mathbf{x}$ from the *a priori* $\mathbf{x}_a$. This *a priori* constraint effectively stabilizes the retrieval by guiding the state vector adjustment when the measurements contain little information to differentiate the contributions from different components of the state vector. Using the optimal estimation (OE) technique (Rodgers, 2000) to minimize the cost function Eq. (1.20) yields a posterior state adjustment vector at the Lth iteration

$$\Delta\mathbf{x} = (\mathbf{S}_a^{-1} + \mathbf{K}^T\mathbf{S}_\epsilon^{-1}\mathbf{K})^{-1}(\mathbf{K}^T\mathbf{S}_\epsilon^{-1}\Delta\mathbf{y}_L + \mathbf{S}_a^{-1}\Delta\mathbf{x}_{aL}) \tag{1.21}$$

$$= (\mathbf{S}_a^{-1} + \mathbf{K}^T\mathbf{S}_\epsilon^{-1}\mathbf{K})^{-1}\left(\mathbf{K}^T\mathbf{S}_\epsilon^{-1}\mathbf{K}\;\mathbf{G}_{\mathrm{DF}}\Delta\mathbf{y}_L + \mathbf{S}_a^{-1}\Delta\mathbf{x}_{aL}\right) \tag{1.22}$$

$$= \Delta\mathbf{x}_{aL} + \mathbf{S}_a\mathbf{K}_T(\mathbf{K}\mathbf{S}_a\mathbf{K}_T + \mathbf{S}_\epsilon)^{-1}(\Delta\mathbf{y}_L - \mathbf{K}\Delta\mathbf{x}_{aL}), \tag{1.23}$$

where $\Delta\mathbf{x}_{aL} = \mathbf{x}_a - \mathbf{x}_L$ and the primed quantities are defined in section 1.3.2. Inserting $\mathbf{I}_n = (\mathbf{K}^T\mathbf{S}_\epsilon^{-1}\mathbf{K})(\mathbf{K}^T\mathbf{S}_\epsilon^{-1}\mathbf{K})^{-1}$, an $n \times n$ identity matrix in front of the term $\mathbf{K}^T\mathbf{S}_\epsilon^{-1}\Delta\mathbf{y}_L$ in Eq. (1.21) yields Eq. (1.22). At iteration $L = 0$, a state vector close to the actual one is sought to be the initial state vector $\mathbf{x}_0$, and a frequent selection is the *a priori* state vector: $\mathbf{x}_0 = \mathbf{x}_a$. This is a more robust inversion scheme that works for $m > n$, $m = n$, and $m < n$. Eq. (1.23) is often used in the case of $m < n$, as the inversion deals with an $m \times m$ (i.e., a smaller) matrix.

Eq. (1.22) describes the difference between the current and previous state vectors $\Delta\mathbf{x}$ as a combination of the direct fitting solution $\mathbf{G}_{\mathrm{DF}}\Delta\mathbf{y}_L$ (see Eq. 1.19, which is derived without any *a priori* constraint) and the difference between the *a priori* and the previous state vectors $\Delta\mathbf{x}_{aL}$ weighted by matrices $\mathbf{K}^T\mathbf{S}_\epsilon^{-1}\mathbf{K}$ and $\mathbf{S}_a^{-1}$ respectively. For the state vector component with a strong *a priori* constraint, i.e., a small variance in $\mathbf{S}_a$, the retrieved result gravitates towards the value of the *a priori* state vector, while for the one with a weak constraint, i.e., a high variance in $\mathbf{S}_a$, its retrieved value is primarily determined from the measurements.

The variance of a retrieved parameter is equal to the corresponding diagonal element of the covariance matrix $(\mathbf{S}_a^{-1} + \mathbf{K}^T\mathbf{S}_\epsilon^{-1}\mathbf{K})^{-1}$ (see Eq. 1.22), thus less or equal to the corresponding *a priori* variance in the *a priori* $\mathbf{S}_a$ matrix. In other words, the change magnitude of a retrieval parameter at each iteration is usually smaller than its *a priori* standard deviation. Consequently the OE method can be used as an inversion scheme to ensure retrieval stability and preserve the dependence of the retrieved results on the measurements, through a careful construction of the *a priori* covariance matrix $\mathbf{S}_a$. To further reduce the dependence on the *a priori* state vector, it is updated at each iteration with the linearization point, setting $\mathbf{x}_a = \mathbf{x}_L$, and hence Eq. (1.21) becomes

$$\Delta\mathbf{x} = (\mathbf{S}_a^{-1} + \mathbf{K}^T\mathbf{S}_\epsilon^{-1}\mathbf{K})^{-1}\mathbf{K}^T\mathbf{S}_\epsilon^{-1}\Delta\mathbf{y}_L = \mathbf{G}\Delta\mathbf{y}_L, \tag{1.24}$$

where $\mathbf{G} = (\mathbf{S}_a^{-1} + \mathbf{K}^T\mathbf{S}_\epsilon^{-1}\mathbf{K})^{-1}\mathbf{K}^T\mathbf{S}_\epsilon^{-1}$ is the optimal estimation gain matrix. This setting floats the anchor point of the retrieval, allowing the measurements to drive the iteration to its final state, with the *a priori* covariance to limit the deviation from the anchor.

By relaxing the *a priori* constraints through increasing the diagonal terms (i.e., the variances) of $\mathbf{S}_a$ such that $\mathbf{S}_a^{-1} \rightarrow \mathbf{0}$, $\mathbf{G}$ becomes $\mathbf{G}_{\mathrm{DF}}$ and Eq. (1.21), as well as Eq. (1.24), becomes Eq. (1.19). In other words, the direct fitting inversion is a special case of the OE inversion scheme,

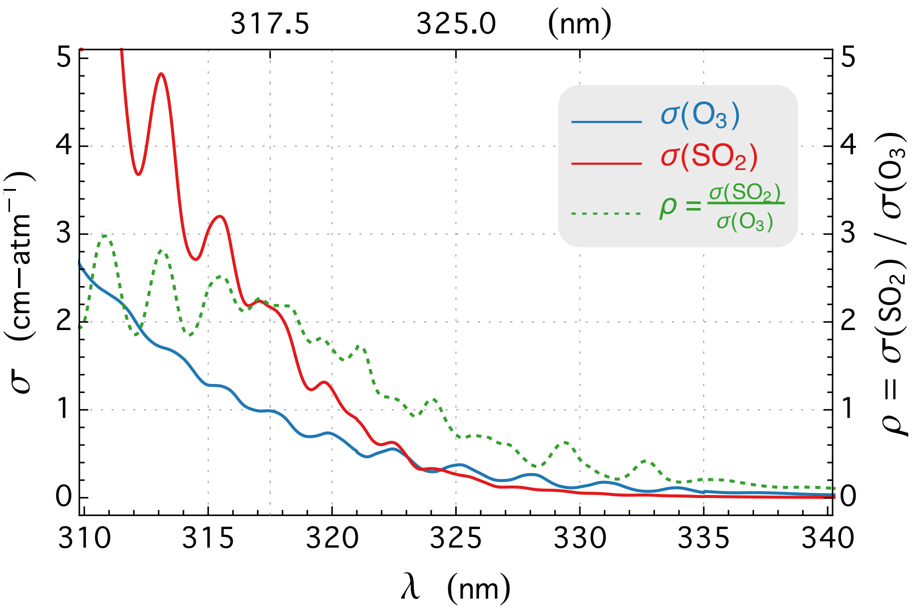


Figure 1.10: EPIC bandpass-averaged cross sections σ for O_3 and SO_2 at 280 Kelvin and their ratio, ρ=σ(SO_2)/σ (O_3).

which is more appropriately called the regulated direct fitting inversion. Using the knowledge of their variances ($\mathbf{S}_a$) to limit some of their ranges while allowing others to change freely, the DVCF algorithm with regulated inversion scheme is suitable for retrieving multiple parameters from discrete measurements . It is applied to EPIC UV observations for simultaneous O_3 and SO_2 retrievals.

1.4 Outline of the thesis

This dissertation addresses the current status & recent progress regarding trace gas retrievals from spaceborne UV measurements via the DVCF algorithm, including retrieving O_3 and SO_2 from DISCOVR EPIC, and retrieving NO_2 from NOAA-20 OMPS and SNPP OMPS instruments.

- Chapter 1 describes the general background and theoretical principle for satellite remote sensing of trace gases, including the physical processes contributing to the satellite spectral measurements, mathematical procedures, and retrieval inversion techniques.

- Chapter 2 to 4 discuss the retrieval processes of O_3 and SO_2 from EPIC UV bands. We present two innovative algorithm improvements (improved O_3 profile representation and the SOE scheme), a thorough error analysis to quantify retrieval uncertainties due to various sources and simplified algorithm physics treatments, and validation of EPIC O_3 and SO_2 products with correlative data (Brewer spectrometers, MERRA-2 reanalysis, and OMPS-NM on SNPP).

- Chapter 5 to 6 present a suite of efforts behind the new NOAA-20 OMPS NO_2 product development, including retrieval algorithm, instrument measurement sensitivity assessment, inter-comparison with OMI tropospheric and stratospheric columns, validation against Pandora measurements, and an application during COVID-19 pandemic.

- Chapter 7 to 8 focus on validation and algorithm improvements for the tropospheric NO_2 column retrievals from SNPP OMPS UV measurements. OMPS column NO_2 is validated against MAX-DOAS measurements acquired from two stations in eastern China. We proposed three algorithm improvements with demonstrated success in improving the accuracy of OMPS NO_2 retrievals.

- Chapter 9 presents a summary of this thesis and future research directions.

Chapter 2: O_3 and SO_2 retrievals from EPIC UV channels

2.1 The DSCOVR EPIC instrument

The Deep Space Climate Observatory (DSCOVR) was launched on the 11^{th} of February 2015, and after a 116-day journey maneuvered successfully into its Lissajous orbit around the first Earth-Sun system Lagrangian (L1) point, which is about 1.5×10^6 km from the Earth and located between the Sun and the Earth on the ecliptic plane. At the L1 point, where the net gravitational forces equal the centrifugal force, DSCOVR orbits the Sun at the same rate as the Earth, staying close in line along the Sun and the Earth and thus allowing the Earth-pointing EPIC to monitor the entire sunlit planet continuously.

The Earth Polychromatic Imaging Camera (EPIC) measures the solar backscattered and reflected radiances from the Earth using a 2-dimensional (2048×2048) charged-coupled device (CCD), recording a set of ten spectral images using different narrowband filters successively. While EPIC may observe the Earth continuously from the vicinity of the L1 point, only a number of spectral image sets are taken in a day, limited by accessible contact windows of the two ground stations located in Wallops island (Virginia) and Fairbanks (Alaska). Currently, between 13 and 22 spectral image sets, recorded at a sampling rate of one set in every 110 minutes during boreal winter and every 65 minutes during boreal summer, are transmitted back to the ground stations in a day.

EPIC takes about six and a half minutes to complete an image set. The first in the set is the blue band (centered at 443 nm), which takes ∼2 minutes to complete the imaging at native resolution (2048×2048 pixels). The images of the nine remaining bands are recorded sequentially at a reduced resolution (1024×1024 pixels, achieved through an onboard average of 2×2 pixels), separating by a time cadence of ∼30 seconds between adjacent bands. Due to the Earth rotation and spacecraft jitter, each spectral image records a slightly different (i.e., rotated) sunlit hemisphere. As a result, the images of two different channels appear to be displaced from each other, usually by a distance of about one to a few native pixels, depending on their observation time difference.

Each native pixel has a ∼1 arc second or 2.778×10^{-4} degree angular instantaneous field of view (IFOV), yielding a geometric ground footprint size of ∼8×8 km^2 at the image center of the sunlit disk. The effective footprint size is about 10×10 km^2, which is larger than the geometric one due to the effect of the optical point-spread function of the EPIC imaging system. For a reduced resolution image (1024×1024 pixels), the effective central ground IFOV size is about 18×18 km^2, which is significantly smaller than the nadir footprints of some past and present satellite instruments that provided global ozone mapping from the low Earth orbit (LEO), such as the Total Ozone Mapping Spectrometer (TOMS, nadir pixel size 50×50 km^2) on a series of satellites, the Scanning Imaging Absorption Spectrometer for Atmospheric Cartography (SCIAMACHY, 60×30 km^2, Bovensmann et al. 1999) on ESA's ENVIronmental SATellite (ENVISAT), the Ozone Mapping and Profiler Suite Nadir Mapper (OMPS-NM, 50×50 km^2, Flynn et al. 2014) on Suomi National Polar Partnership (SNPP), the Global Ozone Monitoring Experiment–2 (GOME-2, Callies et al. 2000; Munro et al. 2016) on Metop-A (40×40 km^2), Metop-B (80×40 km^2), and Metop-C (80×40 km^2). Though it is slightly larger than the nadir footprint of the

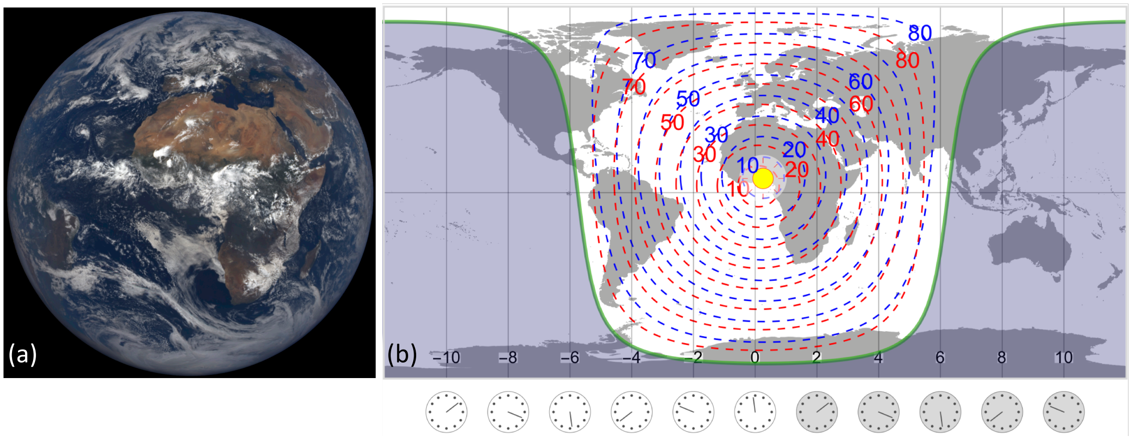


Figure 2.1: (a) Example of EPIC Field of View (FOV): EPIC earth image at 11:40:31 UTC on 4 September 2015. Image source: NASA EPIC Team, https://epic.gsfc.nasa.gov. (b) Viewing and illumination angles are taken from FOV on the left. The subsolar point is marked on the map with a yellow dot. The area shaded with midnight blue is in the dark, i.e., without direct sunlight, while the unshaded area is the sunlit hemisphere, with sunrise on the left (west of subsolar point) and sunset on the right (east of subsolar point). Contours of solar zenith angles (SZAs, blue dashed lines) and viewing zenith angles (VZAs, red dashed lines), going from 10° to 80° with a step 10°, are shown in the sunlit area. Note that the SZA (θ_s) and VZA (θ_v) of an EPIC IFOV have similar values and both angles increase as the IFOV moves from the center towards the edge of the sunlit disk.

Ozone Monitoring Instrument (OMI, 13×24 km^2, Levelt et al. 2006) on Aura and the OMPS-NM (17×13 km^2, Flynn et al. 2016) on NOAA-20, and much bigger than that of the TROPOspheric Monitoring Instrument (TROPOMI, 5.5×3.5 km^2, Veefkind et al. 2012) on the ESA Sentinel-5 Precursor (S5P), EPIC's spatial resolution are sufficiently high to map small-scale O_3 natural variations and observe many volcanic emissions, from degassing to eruption.

EPIC, combining moderate spatial resolution with high temporal cadences from the unique vantage point of L1, provides unprecedented Earth observations, from sunrise to sunset simultaneously (see Fig. 2.1). This synoptic (i.e., concurrent, globally unified, and spatially resolved) perspective is quite distinctive from satellite observations from a LEO or a geostationary Earth orbit (GEO): LEO observations are often made within a narrow range of local time with a small number of samplings at a location per day, while GEO observations have limited spatial coverage, constrained to roughly 60° away from its position. The EPIC observations can have simultaneous co-located observations with measurements from any contemporaneous LEO and GEO platforms, allowing direct comparisons and synergistic use of data acquired from different perspectives. This overlapping feature has been exploited to calibrate some EPIC channels by matching its measured albedo values to those of OMPS-NM on SNPP (Herman et al., 2017).

The ten narrow bands of EPIC, spanning ultra-violet (UV), visible, and near-infrared wavelengths, are selected to yield diverse information about the Earth, from atmospheric compositions to surface reflectivity and vegetation. Four of the ten bands measure UV spectral radiances, which are used primarily for total ozone (O_3) retrievals. These UV bands also provide sensitive detection of sulfur dioxide (SO_2) and volcanic ash, both of which may be episodically injected into the atmosphere during explosive volcanic eruptions.

2.2 The improved O_3 and temperature vertical profiles

As shown in section 1.2.1, the O_3 vertical distribution or profile directly affects the magnitude of a measured radiance in the spectral region with significant O_3 absorption. Hence the interpretation of radiance change due to O_3 absorption requires some knowledge of its profile. In general, the retrieval of quantitative information about a gaseous absorber (such as O_3 and SO_2) requires a model to prescribe its vertical distribution. The skill of this model in representing the actual vertical distribution of the absorber contributes significantly to the quantification accuracy. In this section, we describe a recently developed O_3 profile model for remote sensing retrieval algorithms and its improvements over the model commonly used by other total O_3 algorithms.

Ozone is naturally present throughout the atmosphere and its spatial and temporal distribution controlled by atmospheric processes of O_3 production, destruction, and transport. The O_3 distribution exhibits a high abundance of O_3 in the stratosphere and a minor portion ($\sim$10%) in the troposphere, with the peak O_3 concentration occurring at a lower altitude as the latitude increases towards the poles. These characteristics are well captured by O_3 profile climatologies (e.g. Fortuin and Kelder, 1998; McPeters and Labow, 2012; McPeters et al., 2007), which provide the mean and variance of O_3 vertical distribution as a function of latitude and calendar month. These climatologies also reveal that O_3 profile has the highest variability in the upper troposphere and lower stratosphere (UTLS), contributing the most to the natural variations in total O_3. This high O_3 variability is the consequence of atmospheric movements that blend air masses with different O_3 concentrations, such as uplifting of O_3 poor air in the troposphere or lowering of O_3 rich air in the stratosphere resulting from the rise and fall of the tropopause. Predictors of O_3 profile shape, including tropopause pressure and total O_3 columns, are developed to capture the dynamical influences on O_3 vertical distributions, resulting in the construction of tropopause-sensitive (Bak et al., 2013; Sofieva et al., 2014; Wei et al., 2010) and total-column-dependent (Bhartia and Wellemeyer, 2002; Labow et al., 2015; Lamsal et al., 2004; Wellemeyer et al., 1997) O_3 profile climatologies.

The O_3 profile model for the Total Ozone Mapping Spectrometer Version 8 (TOMS-V8) total O_3 algorithm combines the latitude-dependent monthly mean Labow-Logan-McPeters (LLM) climatology (McPeters et al., 2007) with the latitude- and total-column-dependent annual mean climatology (Bhartia and Wellemeyer, 2002) to determine the O_3 profile as a function of latitude, time (day of year, DOY), and total O_3 column. This model has been adopted by nearly all the contemporary total O_3 algorithms (e.g. Bhartia and Wellemeyer, 2002; Eskes et al., 2005; Lerot et al., 2010, 2014; Loyola et al., 2011; Van Roozendael et al., 2006, 2012; Veefkind et al., 2006;

Wassmann et al., 2015), owing to its capability of characterizing O_3 profile variation with the total column.

To improve the representation of O_3 profile, we construct both tropopause-dependent and total-column-dependent climatologies using the Modern-Era Retrospective Analysis for Research and Applications version 2 (MERRA-2, Bosilovich et al. 2015; Gelaro et al. 2017) O_3 record between 2005 and 2016. The total-column-dependent climatology, named M2TCO3, is more appropriate for use as the O_3 profile model needed by a total O_3 algorithm, as it is generally more reliable than the tropopause-dependent version in prescribing realistic O_3 profiles (Yang and Liu, 2019).

Figure 2.2 compares daytime M2TCO3 (Yang and Liu 2019, referred to as M2TCO3 hereafter) and TOMS-V8 profiles for two months and four latitude zones, illustrating the similarities and differences between the two O_3 models. Both show north-south asymmetry, i.e., profiles in the northern hemisphere differ from those in the southern hemisphere for the corresponding months and latitude zones (e.g., September and 60°S–50°S vs. March and 50°N–60°N in Fig. 2.2), substantial seasonal variations (e.g., 60°S–50°S, March vs. September in Fig. 2.2), strong dependence on latitude, exhibiting lower altitudes of O_3 concentration peaks at higher latitudes for similar total columns, and characteristic dependence on the total column, which gets smaller with a higher O_3 peak altitude (e.g., March and 50°N–60°N in Fig. 2.2). Figure 2.2 shows good agreements of zonal mean profiles (e.g., close matches between solid black and dotted black curves in each panel of this figure), but significant differences between M2TCO3 and TOMS-V8 profiles for similar total columns. These differences are due to TOMS-V8's use of annual mean column-dependent climatology to account for profile variations with the total column throughout the year (Bhartia and Wellemeyer, 2002), thus ignoring the significant seasonal dependence. An

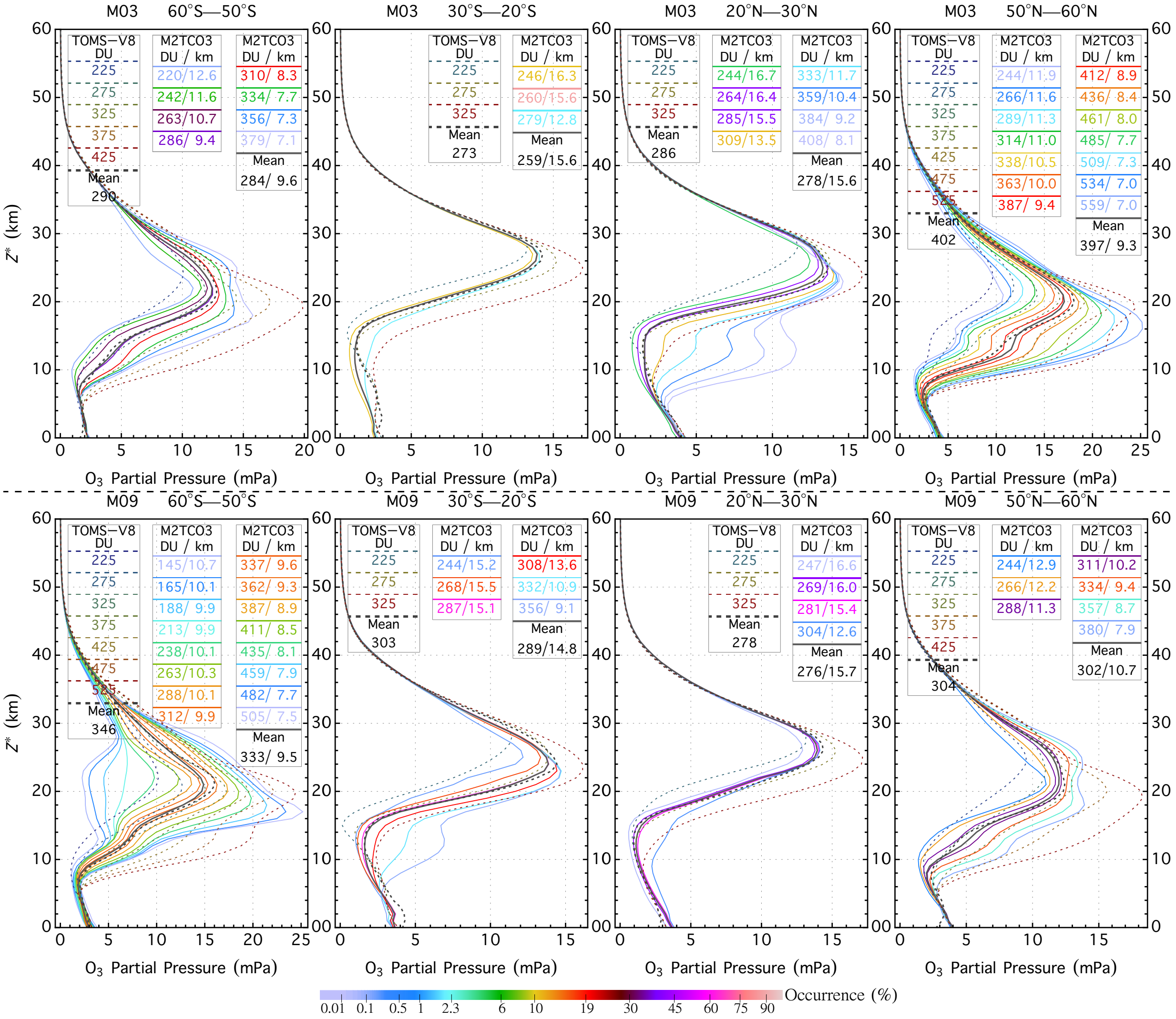


Figure 2.2: Profile comparisons between M2TCO3 and TOMS-V8 for two months (March and September) and four latitude zones: 60°S–50°S, 30°S–20°S, 20°N–30°N, and 50°N–60°N. Colored solid lines represent M2TCO3 profiles, while the dotted ones for TOMS-V8 profiles. The color of a solid line indicates the percentage occurrence of the climatological profile, and its line legend displays the mean tropopause altitude and the mean total column O_3 of the profile. The solid black lines represent the downgraded M2TCO3 (i.e., the monthly zonal mean) profiles and dotted lines are TOMS-V8 monthly zonal mean (i.e., the LLM climatological) profiles. Here pressure altitude is defined as $Z^* = 16\ \log_{10}[\frac{p_s}{p}]$, where p is pressure level (in hPa) and $p_s = 1013.25$ hPa.

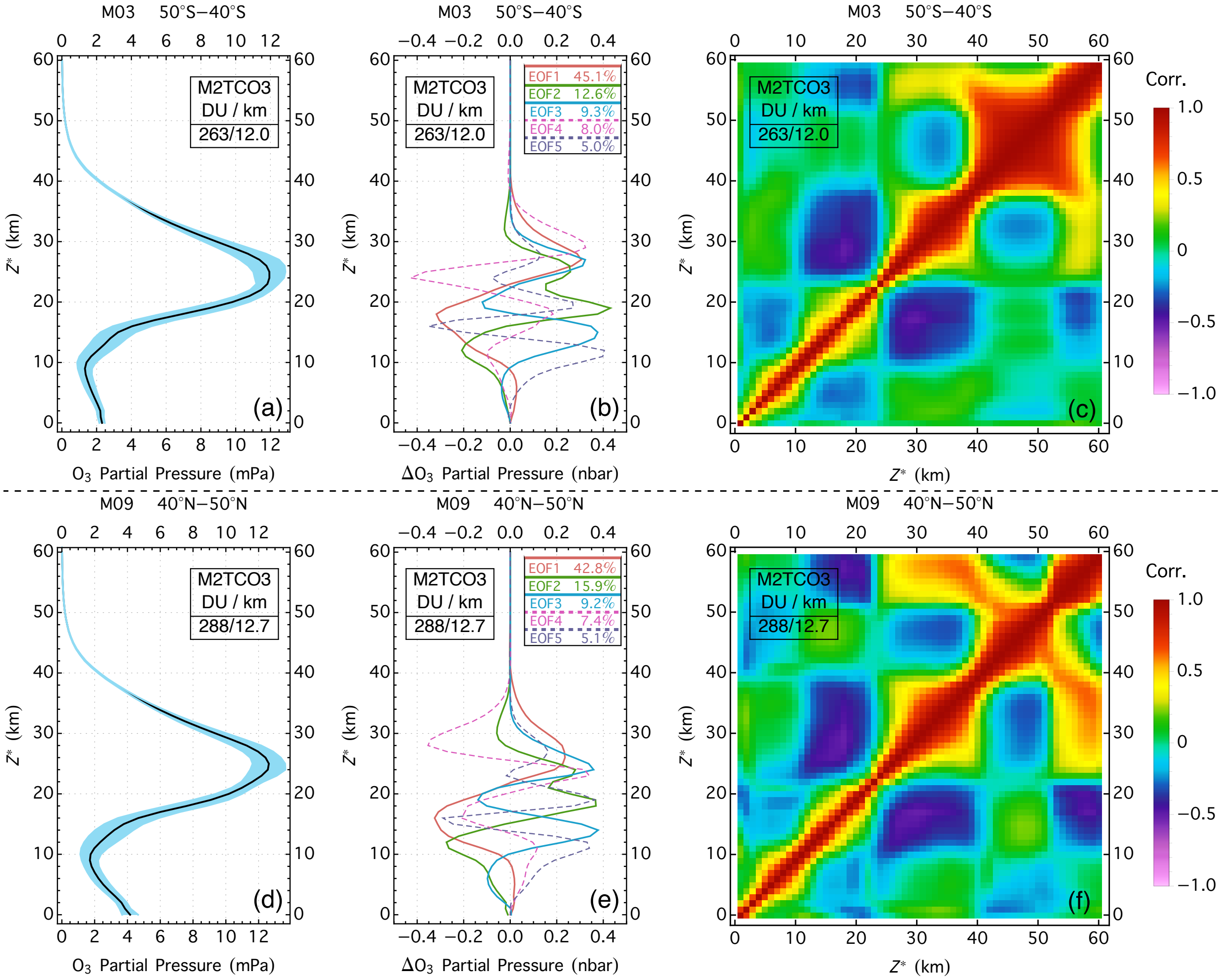


Figure 2.3: Examples of M2TCO3 climatological profiles for the southern midlatitude zone in March (panel a) and the northern midlatitude zone in September (panel b), the associated correlation matrices (panels c and f), and the corresponding modal O_3 profiles (panels b and e). The blue shaded areas in panels a and d are within one standard deviation of the mean. The correlation matrices in panels c and f are standardized (i.e., diagonal element normalized to 1) covariance matrices. The five modal profiles in panels b and e are the first five ordered eigenvectors (also known as empirical orthogonal functions or EOFs) of the corresponding covariance matrices, with percentages of the profile variance explained by the EOFs displayed in the line legends. The text box in each panel displays the average tropopause altitude (in km) and the average total O_3 column (in DU) for the climatological profile.

additional deficiency of TOMS-V8 contributing to the differences is its inadequate representation of latitude-dependent O_3 profile variation with the total column, including broad (30°) latitude zones and omission of north-south asymmetry. These deficiencies are eliminated with M2TCO3, which improves the realism of O_3 profile representation.

In short, M2TCO3 better captures the dynamical changes and spatiotemporal variations in O_3 profiles with higher resolutions in total O_3 column (25 DU), latitude (10°), and time (monthly). Taking into account the substantial change of atmospheric O_3 over a long time, M2TCO3 is more accurate to represent atmospheric O_3 vertical distribution from recent past to near future than the TOMS-V8 model, which was compiled from earlier satellite and ozonesonde data (mostly from the 1980s and 1990s, McPeters et al. 2007; Wellemeyer et al. 1997). Hence we use the M2TCO3 climatology as the O_3 profile model for total O_3 retrieval from EPIC.

The M2TCO3 climatology contains not only mean profiles that represent the likely O_3 vertical distributions, but also the modal O_3 adjustment profiles that specify the probable deviations from the means. These modal profiles are determined from the O_3 profile covariance statistics, as illustrated in Fig. 2.3, showing examples of M2TCO3 climatological O_3 profiles and the associated modal profiles, which are the eigenvectors (also known as the Empirical Orthogonal Functions or EOFs) of the profile covariance matrices. Algebraically the representation of an O_3 profile $\mathbf{X}$ is expressed as

$$\mathbf{X} = \mathbf{X_m}(\mathbf{v}) + \sum_{k=1}^{p} \gamma_k \, \mathbf{e}_k(\mathbf{v}), \tag{2.1}$$

where $\mathbf{X_m}(\mathbf{v})$ is a climatological profile that depends on a set of variables $\mathbf{v}$, which for M2TCO3 consists of the total column (Ω_0), time, and location. $\mathbf{e}_k(\mathbf{v})$ is the k^{th} modal profile, γ_k the k^{th} coefficient, and p the number of $\mathbf{e}_k(\mathbf{v})$, with a maximum equal to the number of levels used to represent an O_3 profile in the climatology. Usually, a few modal profiles are sufficient to account for the majority of profile variance. For example in Fig. 2.3, the first five EOFs (panels b and e) of the covariance matrices (panels c and f) account for 80% profile variances (blue shaded area in panels a and d). An actual O_3 profile $\mathbf{X}$, which deviates invariably from the mean $\mathbf{X_m}$, can be accurately represented using Eq. (2.1) with a small number of expansion coefficients γ_k. Much like the mean the profile $\mathbf{X_m}$ represents the most probable vertical distribution of O_3, the modal profiles, $\{\mathbf{e}_k, k = 1 \ldots\}$, describe the most, the second most, and so on, likely vertical patterns of deviations from the mean profile. Each modal profile describes a rearrangement, like shifting, shrinking, or broadening, of the mean profile without substantially changing the total column. With these modal profiles constraining how a profile can be adjusted, the retrieval algorithm can exploit the O_3 profile information contained in multi-spectral measurements to improve the O_3 profile representation by determining one or more linear expansion coefficients $\{\gamma_k, k = 1 \ldots\}$. Note that for most total ozone algorithms, the O_3 profile representation is limited to the climatological mean only, equivalent to restricting $\gamma_k = 0$ for all k in Eq. (2.1).

The total column is a good predictor of an O_3 profile, especially accurate for the shape in the stratosphere, but less so in the troposphere. Tropospheric O_3 exhibits characteristic spatiotemporal distribution, which is captured in the MERRA-2 tropospheric O_3 climatology (Yang and Liu, 2019). To better represent the O_3 profile, the tropospheric part of a column-dependent M2TCO3 profile, $\mathbf{X_m}$, is scaled with the ratio of the MERRA-2 climatological tropospheric column to the tropospheric column integrated from the downgraded M2TCO3 profile (see Fig. 2.2 for sample M2TCO3 and downgraded M2TCO3 profiles). In other words, the profile $\mathbf{X_m}$ in Eq. (2.1) has its tropospheric part tied to the spatiotemporally varying climatological tropospheric column, to which the tropospheric column of the mean $\mathbf{X_m}$ profile (obtained by averaging over the different column amounts) is matched.

In addition to knowledge of profiles of light-absorbing trace gases, such as O_3 and SO_2, radiative transfer modeling of measured radiance requires knowledge of the atmospheric temperature profile because the absorption cross-sections of these trace gases depend on temperature significantly. For total O_3 retrieval from EPIC, this knowledge is taken from the temperature profile climatology created from MERRA-2 data together with the ozone profile climatlogy (Yang and Liu, 2019). This temperature climatology provides mean temperature profiles corresponding to the climatological O_3 profiles, capturing the dependence of temperature profile on season and location, as well as the variation of temperature with O_3 profile. It is an improvement over the TOMS-V8 temperature profile climatology, which provides latitude-month dependent temperature profiles, but without accounting for the strong correlation between temperature and O_3 profiles.

2.3 Retrieval procedures using the DVCF algorithm

EPIC have four UV channels (see Fig. 1.1), referred to as B1, B2, B3, and B4 and characterized by wavelengths $\lambda_1 = 317.5$ nm, $\lambda_2 = 325.0$ nm, $\lambda_3 = 340.0$ nm, and $\lambda_4 = 388.0$ nm, respectively. The radiance measurements from shorter UV channels, EPIC B1 and B2, are sensitive to both O_3 and SO_2 absorptions (see Fig. 1.10), containing information that allow the retrieval of total O_3 and SO_2 vertical columns, provided that the reflectivity of the underlying surface is known. This knowledge is obtained from the radiance measurements of EPIC B3 and B4, the longer wavelength channels. These channels provide information about the surface reflection and particle back-scattering and have very low sensitivities to O_3 and SO_2 absorption, such that changes in O_3 and SO_2 amounts result in little difference in the radiance measurements of these two bands. The reflectivity determined from B3 and B4 is used to estimate the reflectivity at the shorter wavelength (O_3 sensitive) channels, accomplished with the $r_s f_c$ extrapolation scheme (see section 1.2.3). The reflectivity spectral slope c_l of this extrapolation is proportional to the AI (see Eq. 1.11). The reflectivity parameter (R) is either the LER value r_s estimated from Eq. (1.6) or the MLER cloud fraction f_c from Eq. (1.9) depending on the value of f_c: $R = r_s$ when $f_c = 0$, $R = f_c$ when $f_c > 0$, and its spectral slope is calculated as $c_l = (R_4 - R_3)/(\lambda_4 - \lambda_3)$.

In this section, we describe the application of the DVCF algorithm to EPIC UV measurements, the scheme to solve the difficulty arisen from the non-coincidence among the different EPIC spectral observations, and examples to illustrate the success of this application.

2.3.1 Reflectivity correction by spatial optimal estimation (SOE)

The estimation of O_3 column from EPIC radiance measurements requires accurate reflectivity information of the underlying surface, which is extrapolated from the reflectivity determined at the longer wavelength bands (B3 and B4), but the uncertainty of this extrapolation becomes large due to EPIC's asynchronous spectral measurements. Unlike most space-borne UV instruments which provide coincident measurements from different spectral bands, EPIC takes the spectral images sequentially, separating by a time delay of $\sim$30 seconds between adjacent UV bands. Due to the Earth's self-rotation and spacecraft jitter, different spectral images record slightly different (i.e., rotated) sunlit hemispheres. The geolocation procedure of EPIC (Blank, 2019) aligns different spectral images and further refines the band-to-band registration using the image correlation technique (Yang et al., 2000), which is estimated to provide better than 0.1-pixel (a pixel refers to an IFOV) registration accuracy for EPIC bands. Despite this high registration

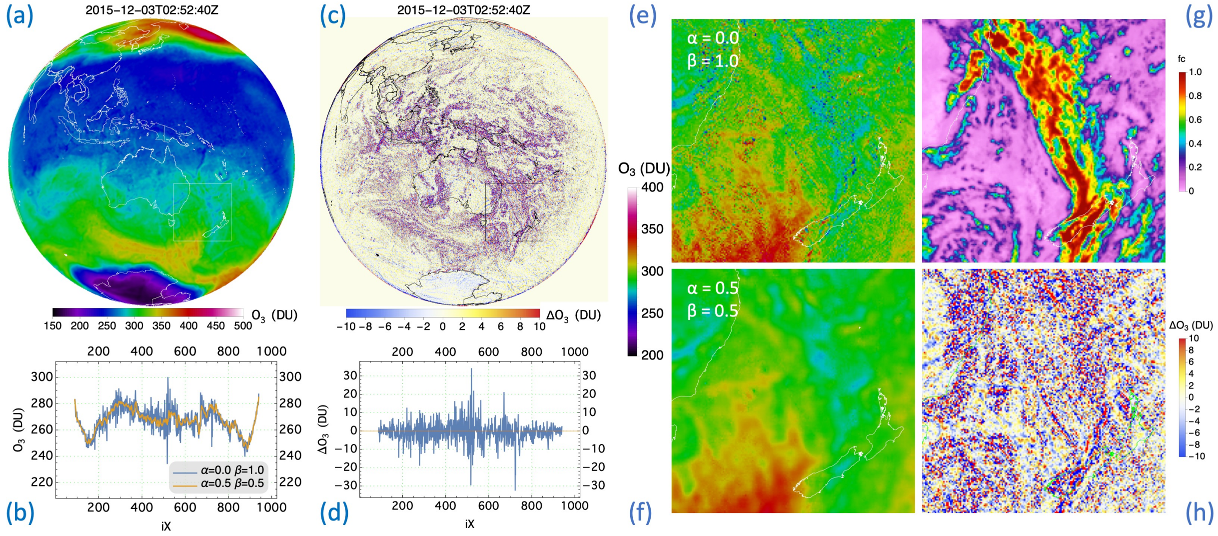


Figure 2.4: Retrieved O_3 from EPIC measurements of bands B1, B3, and B4 on December 3, 2015. (a) Optimized (i.e., $\alpha = 0.5, \beta = 0.5$) O_3 map based on SOE method; (b) a comparison of optimized (orange) and independent-pixel (blue, $\alpha = 0, \beta = 1$) O_3 along the horizontal line (left-to-right) across the middle of the O_3 map in (a); (c) the O_3 difference map: $\Delta O_3 = O_3(\text{Optimized}) - O_3(\text{IndependentPixel})$; (d) the O_3 difference along the horizontal line across the middle of the map in (c); (e) a zoom-in of the independent-pixel O_3 map; (f) the optimized O_3 corresponding to the rectangle in (a); (g) cloud fraction f_c corresponding to (e) and (f); (h) O_3 difference (Optimized – Independent-pixel), a zoom-in corresponding to the rectangle in (c).

accuracy, $r_s f_c$ extrapolation (see section 1.2.3) becomes less accurate for a significant fraction of EPIC IFOVs as substantial reflectivity changes may occur with small shifts in viewing and solar zenith angles since near the direct backscattering direction the particle scattering phase functions have a high angular sensitivity and the shadow areas of structured scenes change non-linearly with viewing-illumination geometry. This difficulty is unlikely to improve even with better alignment and requires a new approach to correct the extrapolated reflectivity.

The basic idea to obtain a more accurate reflectivity at an O_3 sensitive band is to derive it from the radiance measurement of this band with an optimally estimated total O_3 column from the nearby O_3 distribution. This O_3 estimation is attainable because an actual spatial distribution of total O_3 column is a smooth function of geolocation and exhibits a high degree of close-range correlation (Liu et al., 2009). Algebraically, the Spatial Optimal Estimation (SOE) method finds the reflectivity (R_B) at EPIC band B by minimizing the cost function that embodies the *a priori* knowledge of R_B and O_3 spatial distribution. The first part of cost function supports a smooth (i.e., homogeneous) O_3 distribution, while the second part penalizes the difference between R_B and its *a priori* value, which is the extrapolated reflectivity (R_E) from the longer wavelength EPIC bands. Hence the cost function is written as

$$\begin{aligned} \Upsilon &= \alpha \sum_{\substack{i=1\\j=1}}^{n,n} wt(i,j) \left[\frac{\Omega(i)-\Omega(j)}{\langle\Omega\rangle}\right]^2 + \beta \sum_{i=1}^{n} \left[\frac{R_B(i)-R_E(i)}{R_E(i)}\right]^2 & (2.2) \\ &= \sum_{i=1}^{n} \left(\alpha \sum_{j=1}^{n} wt(i,j) \left[\frac{\Omega(i)-\Omega(j)}{\langle\Omega\rangle}\right]^2 + \beta \left[\frac{R_B(i)-R_E(i)}{R_E(i)}\right]^2\right) = \sum_{i=1}^{n} \Upsilon_i, & (2.3) \end{aligned}$$

subject to the measurement constraints $\{I_M(\lambda_B, i) = I_{TOA}(\Omega(i), R_B(i), \lambda_B)$, IOFV index $i =$

$1 \ldots n$, the size of the IFOV group$\}$, which is linearized to become

$$\begin{aligned}\Omega &= \Omega(R_E) + \left(\frac{\partial \ln I_{TOA}}{\partial R_B}\Big|_{R_B=R_E} \Big/ \frac{\partial \ln I_{TOA}}{\partial \Omega}\Big|_{\Omega=\Omega(R_E)} \right)(R_B - R_E) \\ &= \Omega(R_E) + \frac{\partial \Omega}{\partial R_B}\bigg|_{R_B=R_E} (R_B - R_E) = \ \Omega(R_E) + S(R_B - R_E) \end{aligned} \tag{2.4}$$

where $S = \frac{\partial \Omega}{\partial R_B}\Big|_{R_B=R_E}$. The IFOV index i is dropped in Eq. (2.4) without losing clarity. Here j is also an index, labeling the pairing (or other) IFOV in the group and $wt(i, j)$ is the weighting factor that depends on the distance between the i, j pair. $\langle \Omega \rangle$ is the average O_3 column for the group. Given R_E, which is the band B reflectivity extrapolated from the longer wavelength bands, the total O_3 column $\Omega(R_E)$ is retrieved from band B radiance measurement using the exact solution method (see section 1.3.1), and the associated O_3 profile is the column-dependent M2TCO3 climatological profile $\mathbf{X}_m(\Omega)$. The equation of measurement constraint (Eq. 2.4) describes a positive (since $S > 0$ usually) linear relationship between total O_3 column Ω and the surface reflectivity (in the neighborhood of R_E), increasing R_B requires more O_3 absorption to maintain $I_M = I_{TOA}$.

Minimizing only the first r.h.s term of Eq. (2.2) leads to the same O_3 column for all the IFOVs (i.e., $\{\Omega(i) = \langle \Omega \rangle, i = 1 \ldots n\}$), while minimizing only the second term makes $R_B = R_E$ for each IFOVs. The constants α and β are weights to emphasize respectively the smoothness of O_3 spatial distribution and the closeness of reflectivity between extrapolation and estimation. In the SOE scheme, weights are $\alpha = 0$ and $\beta = 1$ for the traditional O_3 retrieval, also referred to as independent-pixel retrieval, while for optimized retrieval, equal weights $\alpha = \beta = 0.5$ are used.

For optimized retrieval, the minimization of the cost function Υ (Eq. 2.2) can be accomplished by iteratively finding $R_B(i)$ to minimizing each component Υ_i. The solution $R_B(i)$ that minimizes

Υ_i is found by solving this equation

$$\frac{\partial \Upsilon_i}{\partial R_B(i)} = \beta \frac{R_B(i) - R_E(i)}{R_E^2(i)} + \alpha \sum_{j=1}^{n} wt(i,j) \frac{\big(\Omega(i) - \Omega(j)\big) S_i}{\langle \Omega \rangle^2} = 0, \tag{2.5}$$

which yields

$$R_B(i) = R_E(i) - \frac{\alpha\, R_E^2(i)\, S_i\, \Big(n'\Omega(i, R_E(i)) - \sum_{j=1}^{n'} \Omega(j)\Big)}{\beta\, \langle \Omega \rangle^2 + \alpha\, n'\, R_E^2(i)\, S_i^2}. \tag{2.6}$$

From Eq. (2.5) to Eq. (2.6), only the n' nearby IFOVs are included, i.e., $wt(i,j) = 1$ for i-j separation within a few (< 4) adjacent IFOVs, otherwise $wt(i,j) = 0$. At the start of iteration, $\{\Omega(j) = \Omega(j, R_E), 1 \ldots n\}$, and they are then updated using Eq. (2.4) with $R_B(i)$ from Eq. (2.6) for the next iteration, which stops until changes in $\{R_B(i), i = 1 \ldots n\}$ becomes sufficiently small. In practice, no more than a couple of iterations are needed to reach convergence.

Figure 2.4 shows an example of simultaneous retrieval from the IFOVs of an EPIC hemispheric view using the SOE method. The high variability O_3 map (Fig. 2.4e) from the independent pixel retrieval contains many artifacts (high spikes and low dips in O_3 columns), which are substantially reduced using the SOE method, resulting in a much more realistic (smooth) O_3 map (Fig. 2.4f). The O_3 differences ($\Delta\Omega$) between optimized and independent-pixel retrievals (see Fig. 2.4c, d, and h) illustrate the quantitative improvements, with a small mean O_3 difference (mean $\Delta\Omega$ within ± 0.5 DU) and a sizeable reduction in O_3 noise level (standard deviation of $\Delta\Omega \approx 7$ DU). The corresponding reflectivity corrections are quite significant ~ 0.02 on average, with a maximum of ~ 0.1 deviation from the $r_s f_c$ extrapolations.

In summary, the SOE method performs single band (B1 or B2) multiple IFOVs (or image-based) retrieval, yielding reflectivity (R) and total O_3 column (Ω), with the associated profile determined by the O_3 model (Eq. 2.1) that retains only the column-dependent M2TCO3 climatological profile $\mathbf{X}_m$. The *a priori* knowledge of the O_3 distribution, which is spatially smooth, provides the extra information to correct the initial reflectivity estimation extrapolated from the characterization based on the longer wavelength bands.

2.3.2 Total O_3 column retrieval

Radiance measurements of EPIC B1 and B2 radiances have O_3 profile sensitivity, which is higher at B1 than at B2, especially at high zenith (SZA or VZA or both) angles (as illustrated in Fig. 1.2). Compared to the measurement of a single O_3 sensitive band, both bands jointly provide more information that allows the refinement of climatological representation of the O_3 profile. This refinement is performed by adjusting the climatological profile with the most probable modal profile ($\mathbf{e}_1$, see Eq. 2.1) so that both B1 and B2 yield the same total O_3 column.

For retrieval from EPIC, the full state vector to be inverted is $\mathbf{x} = \{\Omega_0, \gamma_1, \Xi, R_1, R_2\}$, where Ω_0 is the total O_3 column, γ_1 the O_3 profile adjustment factor, Ξ the total vertical SO_2 column, R_1, R_2 the MLER parameters at EPIC B1 and B2. The regulated direct fitting of EPIC B1 and B2 radiances is applied to obtain retrieved full state vector $\mathbf{x}$.

For each IFOV of EPIC, the O_3 vertical column is estimated first assuming there is no SO_2. The iteration starts with an initial state vector $\mathbf{x}_0 = \{\Omega_0 = \Omega_c, \gamma_1 = 0, \Xi = 0, R_1 = R_1^S, R_2 = R_2^S\}$, where Ω_c is the climatological total column selected from the M2TCO3 climatology based on time and location. R_1^S and R_2^S are the corrected MLER parameters at B1 and B2 respectively using the SOE method (see section 2.3.1).

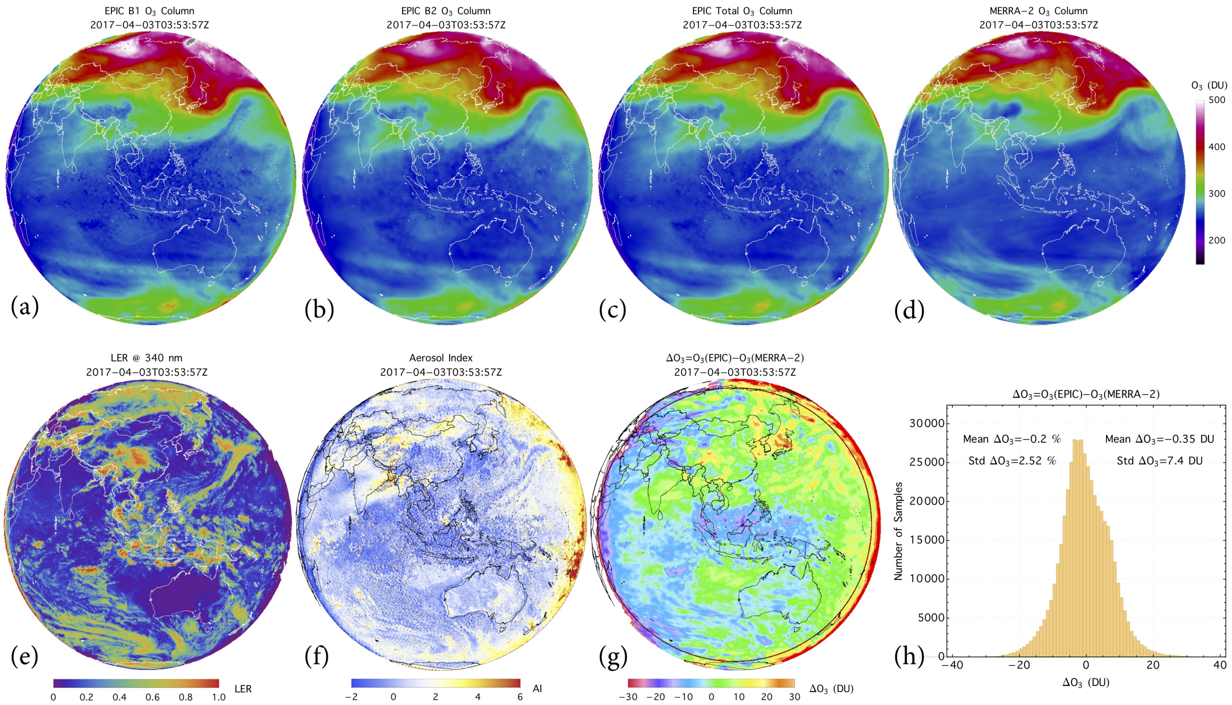


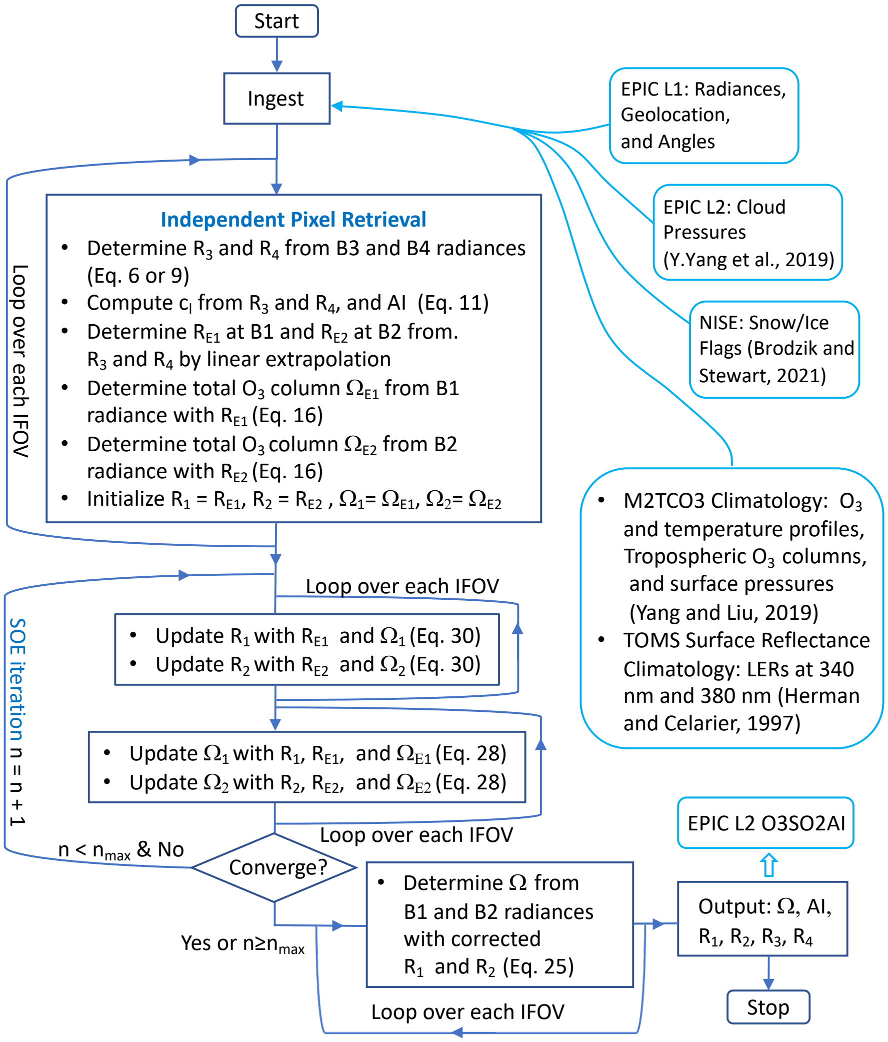
Figure 2.5: An L2 O3SO2AI granule contains the total O_3 vertical columns (c), LER at 340 nm (e), and AI (f), retrieved from EPIC UV measurements at 03:53:57 UTC on 04/03/2017. (a) Total O_3 column (referred to as B1 total O_3 column) retrieved from EPIC B1, B3, and B4. (b) Total O_3 column (referred to as B2 total O_3 column) retrieved from EPIC B2, B3, and B4. (c) Total O_3 from all four bands. (d) Coincident MERRA-2 total O_3 columns. (g) The total O_3 difference: $\mathrm{O_3(EPIC)} - \mathrm{O_3}$(MERRA-2). (h) The histogram of the O_3 differences with SZA $\leq 70°$, i.e., samples within the circle in g, with a mean difference $\mu(\mathrm{EPIC}) = -0.20\%$ (or -0.35 DU) and a standard deviation $\sigma(\mathrm{EPIC}) = 2.52\%$ (or 7.4 DU). Similarly the O_3 difference, $\mathrm{O_3(EPIC\,B1)} - \mathrm{O_3}$(MERRA-2), has a mean of $\mu(\mathrm{EPIC\,B1}) = 0.25\%$ (or 1.03 DU) and a standard deviation $\sigma(\mathrm{EPIC\,B1}) = 2.68\%$ (or 7.9 DU), and $\mathrm{O_3(EPIC\,B2)} - \mathrm{O_3}$(MERRA-2) has a mean $\mu(\mathrm{EPIC\,B2}) = -0.41\%$ (or -1.08 DU) and a standard deviation $\sigma(\mathrm{EPIC\,B2}) = 2.68\%$ (or 7.8 DU).

Since EPIC radiance measurement errors between any two bands are not correlated, the measurement error covariance matrix is diagonal: $\mathbf{S}_\epsilon = \mathrm{diag}(\sigma_{B1}^2 = 0.00345^2, \sigma_{B2}^2 = 0.00345^2)$, estimated from the random errors of the radiance (I_M) measurements (see section 3.2).

There is no correlation among retrieval parameters: total O_3 column (Ω_0), the deviation (ω_1) of O_3 profile from the mean, SO_2 column (Ξ), and the MLER parameters R, except between R_1 and R_2. The diagonal elements of the *a priori* covariance matrix are $\mathbf{S}_a = \mathrm{diag}(\varepsilon_{\Omega_0}^2 = 10^2\ \mathrm{DU}^2, \varepsilon_{\gamma_1}^2 = 2^2\ \mathrm{DU}^2, \varepsilon_{\Xi}^2 = 0.0001^2\ \mathrm{DU}^2,\ \varepsilon_{R1}^2 = 0.001^2, \varepsilon_{R2}^2 = \ 0.001^2)$. The off-diagonal elements are equal to zero, $\{S_a(i,j) = 0, \text{when } i \neq j\}$, except for the elements associated with R_1 and R_2, which may be set at $S_a(4,5) = S_a(5,4) = 0.98\ \varepsilon_{R1}\ \varepsilon_{R2} = 0.0099^2$, representing a high degree of correlation (0.99) between R_1 and R_2. This $\mathbf{S}_a$ essentially limits the adjustments at each iteration: $|\Delta\Omega_0|\varepsilon_{\Omega_0}(10\ \mathrm{DU})$, $|\Delta\omega_1|\varepsilon_{\gamma_1}(2\ \mathrm{DU})$, $|\Delta\Xi|\varepsilon_{\Xi}(10^{-4}\ \mathrm{DU})$, $|\Delta R_1|\varepsilon_{R1}(0.001)$, and $|\Delta R_2|\varepsilon_{R2}(0.001)$. The strong constraint on SO_2 ensures that its column Ξ never deviates far (> 0.01 DU) from its initial value 0 during the iteration, essentially enforcing an SO_2-free retrieval. The strong constraints on R_1 and R_2 also ensures that they remain nearly the same as their initial values R_1^S and R_2^S. The constraints on O_3 parameters are quite loose. Especially towards the convergence of the iteration, the absolute adjustment of each component is much smaller than the corresponding standard deviation, i.e., the square root of the corresponding diagonal element of $\mathbf{S}_a$.

With the setup of error and *a priori* covariance matrices $\mathbf{S}_\epsilon$ and $\mathbf{S}_a$, the initial state vector $\mathbf{x}_0$ is updated (Eq. 1.16) iteratively using $\Delta\mathbf{x}$ from Eq. (1.24), until the exit of the iteration when $|\Delta\Omega_0| < 0.5$ DU and $|\Delta\gamma_1| < 0.5$ DU. The retrieved total O_3 column (Ω) is obtained by integrating the profile $\mathbf{X} = \mathbf{X_m}(\Omega_0) + \gamma_1\mathbf{e}_1(\Omega_0)$. In processing EPIC data, the initial O_3 column Ω_c of $\mathbf{x}_0$ for an IFOV may be set to the column Ω of a previous (or nearby) IFOV to improve the speed of convergence of the iteration.

Algorithm 1 EPIC Total Ozone Retrieval: starts from the ingest of EPIC L1B (Blank, 2019; Blank et al., 2021; Cede et al., 2021), EPIC L2 Cloud (Y.Yang et al. 2019), ancillary data about snow/ice coverage (Brodzik and Stewart, 2021), and climatological databases (including O_3 and temperature profiles (Yang and Liu, 2019) and surface reflectance (Herman and Celarier, 1997)), followed by independent pixel retrievals, SOE corrections, then DVCF inversion, and end with writing the O3SO2AI output (NASA/LARC/SD/ASDC, 2018).



Here we list the algorithmic procedure (see flowchart Algorithm 1), titled EPIC Total Ozone Retrieval, that is applied to each EPIC level-1b (L1B) granule, which contains spectral measurements, as well as geolocation and angular information of all the IFOVs of a snapshot of the sunlit side of the Earth, to produce the level-2 (L2) O3SO2AI product. The contents of a sample O3SO2AI granule are displayed in Fig. 2.5, including total O_3, LER, and AI, respectively shown in panels (c), (e), and (f). For comparison, the MERRA-2 assimilated O_3 total columns, interpolated to the time and location of EPIC IFOVs, are included in Fig. 2.5(d), their differences $O_3(\mathrm{EPIC}) - O_3$(MERRA-2) in (g) and the histogram in (h). This comparison reveals excellent agreement between MERRA-2 and EPIC total O_3, showing near identical O_3 spatial distributions with similar highs and lows. Quantitatively, the differences for samples with VZA$\leq 70^\circ$ are characterized by a low mean offset ($\mu(\mathrm{EPIC}) = -0.2\%$) and a narrow spread (standard deviation $\sigma(\mathrm{EPIC}) = 2.52\%$). Figure 2.5 includes the intermediate results of the EPIC total O_3 processing (see the procedure in Algorithm 1), showing the total O_3 columns retrieved from B1 in panel (a) and from B2 in panel (b) using the SOE method. Both B1 and B2 total columns closely resemble the MERRA-2 (d) and EPIC (c) total O_3 fields, with difference statistics showing slightly worse offsets ($\mu(\mathrm{EPIC\ B1}) = 0.25\%$ and $\mu(\mathrm{EPIC\ B2}) = -0.41\%$) and higher standard deviations ($\sigma(\mathrm{EPIC\ B1}) = 2.68\%$ and $\sigma(\mathrm{EPIC\ B2}) = 2.68\%$). The improved agreement with MERRA-2 is significant, reducing the B1 O_3 spread by $\sqrt{\sigma^2(\mathrm{EPIC\ B1}) - \sigma^2(\mathrm{EPIC})} = 0.9\%$ (or 2.8 DU) and the B2 O_3 spread by a similar amount. These better agreements are consistent over time and location, substantiating the improved retrieval with both O_3-sensitive bands over a single one, which is adopted by the TOMS-V8 algorithm. Since MERRA-2 O_3 field, from the assimilation of independent measurements of the Aura OMI and Aura Microwave Limb Sounder (MLS), provides highly realistic spatiotemporal O_3 representation, the smaller spread between the two-band (B1 and B2) EPIC and MERRA-2 total O_3 columns indicates that the inclusion of more O_3 sensitive bands enables more accurate retrievals.

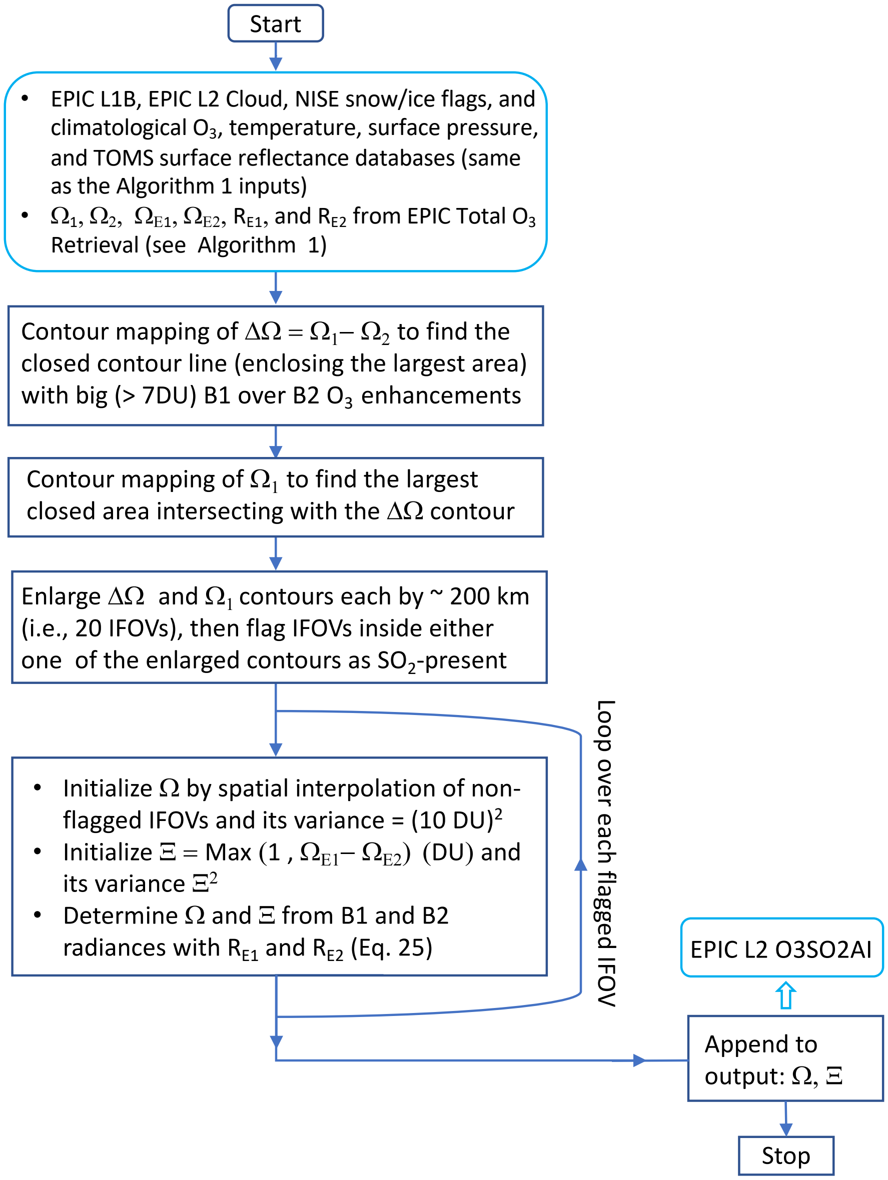
2.3.3 Volcanic SO_2 retrieval

EPIC B1 and B2 radiances respond to both O_3 and SO_2 absorptions, but with very different (see Fig. 1.10) sensitivities: SO_2 is more than twice as UV-absorbent as O_3 at B1, in contrast, it is significantly less at B2, about 70% as absorbent as O_3. Consequently, the estimate of O_3 absorption signals at these two bands would result in an error due to the presence of SO_2 in the atmosphere: 1 DU of SO_2 would usually yield more than 2 DU O_3 error at B1, but only about 0.7 DU error at B2. This big difference in absorption sensitivities facilitates the detection of SO_2 in the atmosphere. Given a radiance SNR of 285:1, the theoretical minimum detectable level of SO_2 enhancement is ~ 0.5 DU in the upper troposphere and above. However, it is difficult to distinguish SO_2 at this minimum level from other changes, such as O_3 profile or surface spectral reflectance, since they can induce similar changes in the measured radiances. This difficulty is increased by the EPIC's asynchronous spectral measurements, which may yield spectral variation similar to the response to adding SO_2 in the atmosphere. Consequently, low levels of SO_2 elevation can not be reliably detected in EPIC observations. For significant SO_2 elevations, typically those from volcanic eruptions, B1 O_3 is much higher than B2 O_3 (i.e., $\Omega_1 > \Omega_2$) from the total O_3 retrieval (described in section 2.3.2). Adjusting the O_3 profile shape or changing the spectral reflectance of the underlying surface usually cannot eliminate this large O_3 discrepancy between the two bands. Therefore a high positive value of $\Delta\Omega$ can be used to flag the presence of SO_2. Furthermore, a volcanic plume usually occupies a contiguous area with a limited spatial extent. Thus, $\Delta\Omega$ and Ω_1 enhancements resulted from volcanic SO_2 plume occur over a large group of connected IFOVs instead of isolated or a small group of disconnected IFOVs. Based on these characteristics of volcanic SO_2 plumes, we describe next an algorithmic procedure to flag IFOVs with SO_2 enhancements.

For reliable SO_2 detection, the following procedure is applied to identify the presence of SO_2 in an IFOV. First, IFOVs of likely SO_2 elevations are flagged through spatial analysis of the differential O_3 field (i.e., $\Delta\Omega = \Omega_1 - \Omega_2$, EPIC B1 and B2 O_3 difference), accomplished through contour mapping to find closed areas of local $\Delta\Omega$ enhancements, i.e., areas within closed contours with Ω_1 considerably higher ($\geq$ 7 DU) than the Ω_2 values (e.g., see Fig.2.6b). The IFOVs within this $\Delta\Omega$ contour likely have a SO_2 elevation around 5 DU or above. Next, contour mapping of Ω_1 is performed to find the longest closed contour line in the area that extends 150 pixels off the extrema of the $\Delta\Omega$ contour (i.e., an image rectangle with a minimum of 300 $\times$ 300 IFOVs or 3000 $\times$ 3000 km^2 that covers the $\Delta\Omega$ contour). Within this closed Ω_1 contour, IFOVs with likely SO_2 enhancements are flagged when $\Omega_1 > \Omega_2$. This flagging is then extended to the adjacent areas outside of the two contours to identify IFOVs with possible SO_2 contamination. For most volcanic plumes, these two contours overlap each other greatly. Including area within the Ω_1 contour and the adjacent outside regions are designed to capture plumes with lower SO_2 elevations.

Once detected, the SO_2 quantification follows the DVCF retrieval with the initial state and *a priori* covariance setting described next. For the IFOV identified with a SO_2 contamination, the initial O_3 values are spatially interpolated from background Ω_1 field, $\gamma_1 = 0$, and initial SO_2 column $\Xi_0 = \Omega_1 - \Omega_2$, integrated from a vertical profile specified by a generalized distribution function (GDF, Yang et al. 2010) with a width and a center altitude appropriate for the plume. The corresponding elements of the *a priori* covariance matrix are $\mathbf{S}_a = \mathrm{diag}(\varepsilon^2_{\Omega_0} = 10^2\ \mathrm{DU}^2, \varepsilon^2_{\gamma_1} = 2^2\ \mathrm{DU}^2, \varepsilon^2_{\Xi} = {\Xi_0}^2\ \mathrm{DU}^2)$, i.e., the variances associated with O_3 are the same as those for total O_3 retrieval, while the SO_2 column variance is equal to the square of the initial SO_2 estimate

Algorithm 2 EPIC Total SO_2 Retrieval: starts from the intermediate results from EPIC total O_3 retrieval, followed by spatial and spectral analysis to mask SO_2-present IFOVs, then DVCF inversion to retrieve O_3 (Ω) and SO_2 (Ξ), and end with appending retrieved results to the O3SO2AI output.



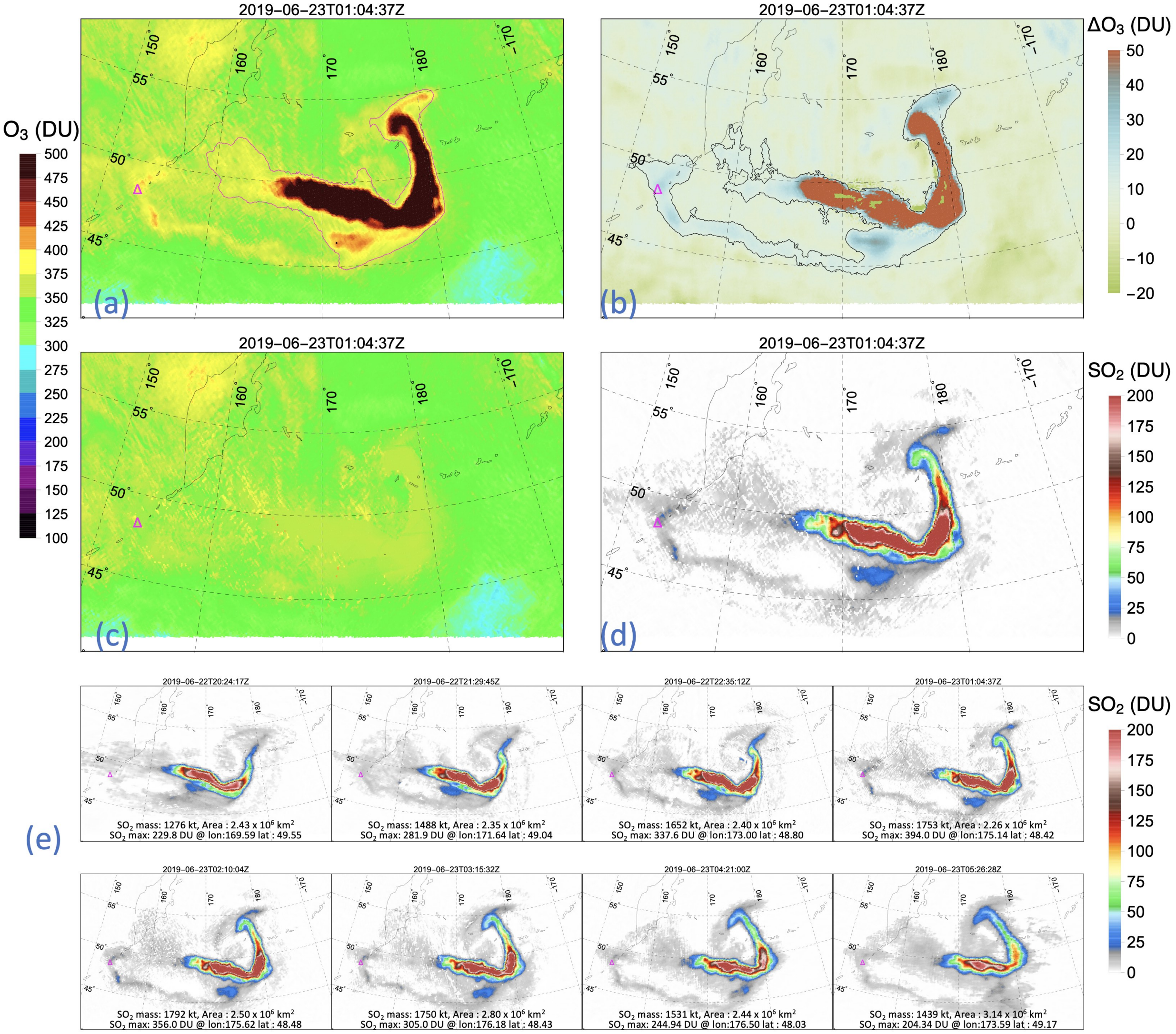


Figure 2.6: EPIC observation of the volcanic plume on 23 June 2019 from the previous day's eruption of Raikoke volcano (represented by Δ in each panel) in the central Kuril Islands of Russia. (a) B1 O_3 column (Ω_1) from EPIC total ozone retrieval and the elevated O_3 contour. (b) B1 and B2 O_3 column difference ($\Delta\Omega = \Omega_1 - \Omega_2$) and elevated $\Delta\Omega$ contour. (c) Vertical O_3 column from EPIC total SO_2 retrieval (see Algorithm 2). (d) Vertical SO_2 column from EPIC total SO_2 retrieval. (e) SO_2 vertical column retrieved from a series of eight consecutive EPIC observations of the Raikoke plume, represented by a 1.5 km thick GDF layer centered at an altitude of 13 km above sea level.

(Ξ_0), which is a weak constraint to allow Ξ change freely responding to the measurement . Other retrieval settings are kept the same as in the total O_3 retrieval described in the previous section. We list the complete algorithmic procedure in flowchart Algorithm 2, titled EPIC Total SO_2 Retrieval, for SO_2 detection and quantification from EPIC UV observations.

The algorithmic procedure (listed in Algorithm 2) applies to regions where EPIC observes volcanic plumes to produce the EPIC volcanic SO_2 product. Figure 2.6 illustrates the detection and quantification of volcanic SO_2 from EPIC observations. Spatial analyses (i.e., contour mappings) of the intermediate results (Figs. 2.6a and b) of the total O_3 processing (see Algorithm 1) provide reliable detection of SO_2 elevations. The SO_2 flagged IFOVs are then processed with the DVCF algorithm to retrieve total vertical O_3 and SO_2 columns simultaneously, with results showing in Figs. 2.6 (c) and (d) respectively. Comparison of the two O_3 fields in Figs. 2.6 shows that the initial O_3 elevations (Figs. 2.6a) due to the presence of SO_2 are nearly entirely removed in the final O_3 field (Figs. 2.6c), demonstrating that the combo retrieval of O_3 and SO_2 achieves consistent O_3 values inside and outside of the plumes. The achieved internal consistency indirectly validates the SO_2 columns. In Fig. 2.6e, we show maps of DVCF retrieved SO_2 columns from a series of eight consecutive EPIC observations of the Raikoke plume in Fig. 2.6 (e), with the maximum SO_2 value, the total SO_2 mass, and the total area covered by elevated SO_2 displayed in each snapshot. These results illustrate the high-cadence observing capability and high-quality SO_2 measurements of EPIC.

Chapter 3: Error analysis and uncertainty estimates of EPIC retrievals

We describe in this section how algorithm physics treatments and various sources contribute to the retrieval uncertainties and provide error estimates of the EPIC O_3 and SO_2 products.

3.1 General expression of retrieval errors

The spectral measurements, represented by a column vector $\mathbf{y}$ of length m (the number of wavelength bands), are written explicitly with all the dependent parameters and possible errors and then expanded with respect to the linearization point ($\mathbf{x}_L$)

$$\begin{aligned}\mathbf{y} = \ln \mathbf{I}_m &= \ln \mathbf{I}_{TOA}(\boldsymbol{\omega}, \boldsymbol{\xi}, \mathbf{b}) + \boldsymbol{\epsilon}_m \\ &= \ln \mathbf{I}_{TOA}(\boldsymbol{\omega}_L, \boldsymbol{\xi}_L, \mathbf{b}_L) + \mathbf{k}_{\boldsymbol{\omega}}(\boldsymbol{\omega} - \boldsymbol{\omega}_L) + \mathbf{k}_{\boldsymbol{\xi}}(\boldsymbol{\xi} - \boldsymbol{\xi}_L) + \mathbf{k}_{\mathbf{b}}(\mathbf{b} - \mathbf{b}_L) + \boldsymbol{\epsilon}_f + \boldsymbol{\epsilon}_m,\end{aligned} \tag{3.1}$$

where the column vectors of length n_l (the number of atmospheric layers), $\boldsymbol{\omega}$ and $\boldsymbol{\xi}$, represent respectively the actual vertical profiles for O_3 and SO_2, while $\boldsymbol{\omega}_L$ is the climatological O_3 profile equal to $\mathbf{X_m}(\Omega_L) + \gamma_1 \mathbf{e}_1(\Omega_L)$ and $\boldsymbol{\xi}_L$ the prescribed SO_2 profile specified by a GDF layer with an integrated vertical column equal to Ξ_L. The profile weighting functions, $\mathbf{k}_{\boldsymbol{\omega}} = \frac{-\partial \ln \mathbf{I}_{TOA}}{\partial \boldsymbol{\omega}}|_{\boldsymbol{\omega}=\boldsymbol{\omega}_L}$ and $\mathbf{k}_{\boldsymbol{\xi}} = \frac{-\partial \ln \mathbf{I}_{TOA}}{\partial \boldsymbol{\xi}}|_{\boldsymbol{\xi}=\boldsymbol{\xi}_L}$, are $m \times n_l$ matrices, with each of its rows equal to the product of absorber cross-sections at one spectral band and the layer AMFs (i.e., the mean photon path lengths through the atmospheric layers). Likewise, $\mathbf{b}$ and $\mathbf{b}_L$ are respectively a set of the exact forward model parameters and those used in the linearization, with the corresponding sensitivity matrix $\mathbf{k_b} = \frac{-\partial \ln \mathbf{I}_{TOA}}{\partial \mathbf{b}}|_{\mathbf{b}=\mathbf{b}_L}$. These forward model parameters may include the spectral-dependent MLER parameters, the ground surface and the OCP cloud pressures, the atmospheric temperature profile, the absorption cross-sections of O_3 and SO_2, and the parameters that specify the ISRFs of the spectral bands. The column vector $\boldsymbol{\epsilon}_f$ is the forward modeling errors of the spectral bands, such as the approximate radiative transfer through Earth's spherical atmosphere and the incomplete accounting for RRS contributions. The last term $\boldsymbol{\epsilon}_m$ is a column vector representing the spectral radiance errors of the instrument, including random noises and radiometric calibration biases.

Using the definition $\Delta\mathbf{x} = \mathbf{x} - \mathbf{x}_L$ and putting Eq. (3.1) into residual $\Delta\mathbf{y}_L$, Eq. (1.24) is re-written as

$$\begin{aligned}\mathbf{x} - \mathbf{x}_L &= \mathbf{G}\left[\mathbf{k}_{\boldsymbol{\omega}}(\boldsymbol{\omega} - \boldsymbol{\omega}_L) + \mathbf{k}_{\boldsymbol{\xi}}(\boldsymbol{\xi} - \boldsymbol{\xi}_L) + \mathbf{k_b}(\mathbf{b} - \mathbf{b}_L) + \boldsymbol{\epsilon}_f + \boldsymbol{\epsilon}_m\right] \\ &= \mathbf{A}_{\boldsymbol{\omega}}(\boldsymbol{\omega} - \boldsymbol{\omega}_L) + \mathbf{A}_{\boldsymbol{\xi}}(\boldsymbol{\xi} - \boldsymbol{\xi}_L) + \mathbf{G}\mathbf{k_b}(\mathbf{b} - \mathbf{b}_L) + \mathbf{G}\boldsymbol{\epsilon}_f + \mathbf{G}\boldsymbol{\epsilon}_m, \end{aligned} \tag{3.2}$$

where $\mathbf{A}_{\boldsymbol{\omega}} = \mathbf{G}\mathbf{k}_{\boldsymbol{\omega}}$ and $\mathbf{A}_{\boldsymbol{\xi}} = \mathbf{G}\mathbf{k}_{\boldsymbol{\xi}}$ are the averaging kernels (AKs) for O_3 and SO_2, respectively. Eq. (3.2) describes how various error sources, from mismatches in absorber profiles to errors in model and measurement, propagate into the final result ($\mathbf{x}$). The rows associated with the O_3 and SO_2 columns can be extracted from the vector equation (Eq. 3.2) and written as

$$\Omega - \Omega_T = (\mathbf{A}_\Omega - \mathbf{1})(\boldsymbol{\omega} - \boldsymbol{\omega}_L) + \mathbf{G}_\Omega\boldsymbol{\epsilon}_\Omega, \tag{3.3}$$

$$\Xi - \Xi_T = (\mathbf{A}_\Xi - \mathbf{1})(\ \boldsymbol{\xi} -\ \boldsymbol{\xi}_L) + \mathbf{G}_\Xi\boldsymbol{\epsilon}_\Xi, \tag{3.4}$$

after subtracting $\Omega_T - \Omega_L$ and $\Xi_T - \Xi_L$ from the row equations respectively. Here Ω_T and Ξ_T are the true O_3 and SO_2 columns, integrated from the corresponding true O_3 ($\boldsymbol{\omega}$) and SO_2 ($\boldsymbol{\xi}$) profiles. $\mathbf{A}_\Omega$ and $\mathbf{G}_\Omega$ are the row vectors associated with the retrieved O_3 column Ω from the corresponding matrices $\mathbf{A}_\omega$ and $\mathbf{G}$. Analogously, $\mathbf{A}_\Xi$ and $\mathbf{G}_\Xi$ are the row vectors related to the retrieved SO_2 column Ξ taken from the matrices $\mathbf{A}_\xi$ and $\mathbf{G}$ respectively. The constant row vector $\mathbf{1}$ contains the value 1 for all its elements. Thus its dot product with a vertical profile (a column vector) is equivalent to the summation of all the individual layer amounts, yielding the total column. The column vector $\boldsymbol{\epsilon}_\Omega$ represents the total error combined from various sources impacting the total O_3 accuracy, including errors in model parameters $\mathbf{k_b}(\mathbf{b} - \mathbf{b}_L)$, forward modeling $\boldsymbol{\epsilon}_f$, spectral measurements $\boldsymbol{\epsilon}_m$, and the other absorber $\mathbf{k}_{\boldsymbol{\xi}}(\boldsymbol{\xi} - \boldsymbol{\xi}_L)$. Similarly, $\boldsymbol{\epsilon}_\Xi$ represents the combined total error affecting total SO_2 accuracy with the other absorber term being replaced by $\mathbf{k}_{\boldsymbol{\omega}}(\boldsymbol{\omega} - \boldsymbol{\omega}_L)$.

Retrieval errors can be characterized using Eqs. (3.3) and (3.4), provided errors from various sources are sufficiently small that forward modeling responds linearly to these deviations. However, substantial retrieval errors usually are resulted from simplified physics treatments, which constrain the forward model to be radiative transfer in a molecular atmosphere over Lambertian surfaces. These errors may be called the AMF errors because the simplified physics treatments can not, in general, reproduce the paths of photons through the observed atmosphere, even though they enable radiance matching between measurement and modeling. The deviations of mean paths lead to retrieval errors in O_3 and SO_2 because the interpretation of measured radiance through radiance matching requires accurate modeling of mean photon paths (i.e., the AMFs). The retrieval errors from the simplified physics treatment can be estimated using closed-loop tests (i.e., realistic forward modelings and then inverse retrieval with simplified physics treatments). Next, we provide uncertainty estimates of O_3 and SO_2 retrievals contributed from various errors

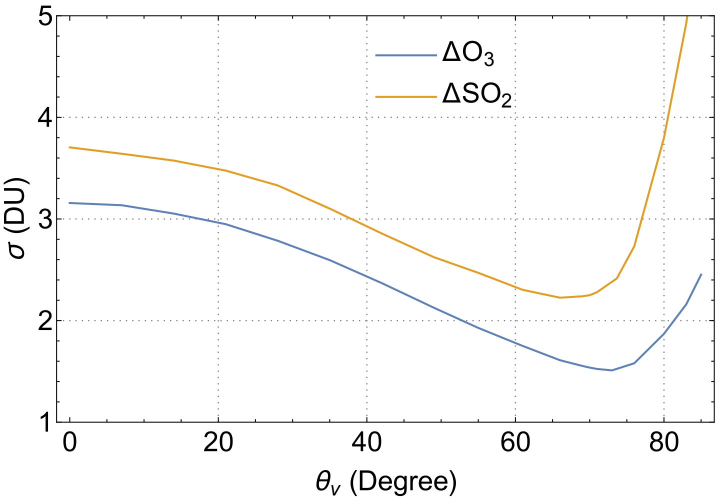


Figure 3.1: Noise levels, i.e., standard deviations (σ) of O_3 and SO_2 errors (Δ) contributed from the random noises on EPIC spectral measurements. The SO_2 noise estimate is for a layer at an altitude 11 km above sea level.

sources and simplified physics treatments.

3.2 Uncertainty estimates

3.2.1 Measurement errors

Errors in EPIC spectral measurements contribute to uncertainties in retrieved O_3 and SO_2 columns (Ω and Ξ). Taken the terms associated with radiance errors from Eqs. (3.3) and (3.4), retrieval errors are written as

$$\Delta\Omega = \mathbf{G}_\Omega(\boldsymbol{\epsilon}_m + \mathbf{k}_\mathbf{R}\Delta\mathbf{R}), \tag{3.5}$$

$$\Delta\Xi = \mathbf{G}_\Xi(\boldsymbol{\epsilon}_m + \mathbf{k}_\mathbf{R}\Delta\mathbf{R}). \tag{3.6}$$

These equations specify how measurement errors ($\boldsymbol{\epsilon}_m$) of the O_3 sensitive bands and the MLER parameter errors ($\Delta\mathbf{R}$) due to the measurement errors in the weak absorption bands propagate into retrieved vertical columns.

Biases in radiance measurements lead to systematic errors in retrieved vertical columns. While the actual radiance biases are unknown, they are likely less than 1% for EPIC UV bands. For radiance biases within $\pm 1\%$, the systematic O_3 and SO_2 column errors are within $\pm \sim$15 DU and $\pm \sim$8 DU, respectively estimated from Eqs. (3.5) and (3.6). These retrieval column errors are primarily controlled by the relative differences of spectral errors without significant dependence on the column amounts or surface reflectance. The retrieval biases vary with observing conditions given the same percentage radiance errors due to gain matrices ($\mathbf{G}_\Omega$ and $\mathbf{G}_\Xi$) depend significantly on viewing and illumination angles.

In addition to systematic errors, radiance measurement noises add random errors onto the retrieved columns. Retrieval errors due to random radiance noises (specified with normal distributions) are unbiased, with mean values, $\mu(\Delta\Omega)$ and $\mu(\Delta\Xi)$, close to zero and standard deviations, $\sigma(\Delta\Omega)$ and $\sigma(\Delta\Xi)$, proportional to standard deviations of radiance noises. The signal-to-noise ratios for EPIC UV bands are 290:1 (Herman et al., 2018), equivalent to a noise level (standard deviation) of 0.345% (= 1/290). This level is consistent with high-frequency radiance fluctuations (with standard deviations equal to 0.373%, 0.354%, 0.354%, and 0.368% for B1 to B4, respectively) within cloud-free scenes observed by EPIC. With a setting of equal standard deviations for the four UV bands (i.e., $\boldsymbol{\epsilon}_m = \{0.345\%, 0.345\%, 0.345\%, 0.345\%\}$), the estimated column O_3 noise level is $\sigma(\Delta\Omega)$ $\simeq$3.2 DU at low viewing zenith angles, decreases gradually with higher zenith angles, reaches a minimum of $\sim$1.5 DU at $\sim 75°$, then rebounds quickly with further increases in zenith angles (see Fig. 3.1). The noise level of SO_2 columns, $\sigma(\Delta\Xi)$, exhibits a similar angular dependence as shown in Fig. 3.1, primarily following the angular variation of the gain matrix $\mathbf{G}_\Xi$. These angular-dependent column noises are insensitive to the column amounts or the surface reflectance.

3.2.2 Model parameter errors

Retrieval errors due to uncertainties in model parameters, including molecular absorption cross-sections ($\boldsymbol{\sigma}$) and atmospheric temperature profiles ($\mathbf{T}$), are estimated as

$$\Delta\Omega = \mathbf{G}_\Omega \left((\mathbf{m}_z \Delta\boldsymbol{\sigma}_{O_3})\boldsymbol{\omega} + (\mathbf{m}_z \frac{\partial \boldsymbol{\sigma}_{O_3}}{\partial \mathbf{T}})(\boldsymbol{\omega}\Delta\mathbf{T}) \right), \tag{3.7}$$

$$\Delta\Xi = \mathbf{G}_\Xi \left((\mathbf{m}_z \Delta\boldsymbol{\sigma}_{SO_2})\boldsymbol{\xi} + (\mathbf{m}_z \frac{\partial \boldsymbol{\sigma}_{SO_2}}{\partial \mathbf{T}})(\boldsymbol{\xi}\Delta\mathbf{T}) \right), \tag{3.8}$$

where $(\mathbf{m}_z \Delta\boldsymbol{\sigma}_{O_3,SO_2})$ and $(\mathbf{m}_z \frac{\partial \boldsymbol{\sigma}_{O_3,SO_2}}{\partial \mathbf{T}})$ are $m \times n_l$ matrices with elements $\{m_{z_j}(\lambda_i)\Delta\sigma_{O_3,SO_2}(\lambda_i, T_j), i = 1..m, j = 1..n_l\}$ and $\{m_{z_j}(\lambda_i)\frac{\partial \sigma_{O_3,SO_2}(\lambda_i,T)}{\partial T}|_{T_j}, i = 1..m, j = 1..n_l\}$, respectively. Here m_z (see Eq. 1.7) is the mean photon path length through a layer at altitude z, $\Delta\boldsymbol{\sigma}_{O_3,SO_2}$ are errors in O_3 or SO_2 cross-sections, and $\Delta\mathbf{T}$ are errors in atmospheric temperature profiles.

The BDM O_3 cross-sections (Brion et al., 1993; Daumont et al., 1992; Malicet et al., 1995) and the BW SO_2 cross-sections (Birk and Wagner, 2018) are used in O_3 and SO_2 retrievals from EPIC. These baseline cross-sections contain errors, which are not known quantitatively but can be estimated by comparing with alternative cross-sections. Specifically, the BW O_3 cross-sections (Birk and Wagner, 2021) and the SO_2 absorption cross-sections of Bogumil et al. (2003) are the alternatives that can replace the baselines for EPIC retrievals. The cross-section errors, $\Delta\boldsymbol{\sigma}_{O_3}$ in Eq. (3.7) and $\Delta\boldsymbol{\sigma}_{SO_2}$ in Eq. (3.8), are estimated based on the differences between the alternatives and the baselines, showing that alternative O_3 and SO_2 cross-sections are slightly (about 0.1% to 1.1%) lower than the corresponding baselines at EPIC B1 and B2. These biases in cross-sections result in O_3 column biases between 0.5% and 2% and SO_2 column biases between

1% and 2% depending on the effective cross-section differences. The temperature-dependence of the BW O_3 cross-sections behaves quite differently from the BDM (Bak et al., 2020), especially at EPIC B2, contributing to the high ends of O_3 biases ($> 1.0\%$), which occur predominantly at high (viewing or solar) zenith angles when O_3 retrieval becomes more sensitive to EPIC B2 radiance.

Both O_3 and SO_2 cross-sections are significantly dependent on temperatures. Thus accurate temperature profiles are needed to determine atmospheric absorption properties for modeling of measured radiances. As mentioned in section 2.2, MERRA-2 climatological temperature profiles (Yang and Liu, 2019) are used for retrievals from EPIC. Actual temperature profiles differ from the climatological profiles. Over a short period (e.g., a day), the spatial distribution of these differences is not random, leading to retrieval errors that are unevenly distributed spatially. However, actual temperature profiles are normally distributed around the climatological mean over a long period (e.g., a month) for a location. Therefore temperature profile mismatches add random components, which average to zero over a long time, to the total errors. The variances of these random errors are proportional to the layer-column weighted temperature error variances. Estimated from the variances of temperature profiles (Yang and Liu, 2019), the random components, $\sigma(\Delta\Omega)$, are $\sim$0.3% in the tropics, increase to $\sim$0.7% in the mid-latitudes, and reach $\sim$1% at high latitudes. Similarly, random errors, $\sigma(\Delta\Xi)$, are $\sim$0.8% in the tropics, $\sim$1.7% in the mid-latitudes, and $\sim$3.5% at high latitudes.

3.2.3 Forward modeling errors

The MLER treatment adopted for the retrieval algorithm allows the use of the vector radiative transfer code, TOMRAD, as the forward model to simulate measured radiances and weighting functions. TOMRAD implements Dave's iterative solution (Dave, 1964) with pseudo-spherical approximation (Caudill et al., 1997) to the problem of the transfer of solar radiations through a molecular atmosphere over a Lambertian surface. The forward modeling with TOMRAD is accurate for EPIC observations around the center of its hemispheric view, with radiance errors (ϵ_f) of all EPIC UV bands less than $\pm$ 0.2% for VZA $< 50^\circ$. Note that for EPIC observations, each of its IFOVs has similar VZA and SZA (see Fig. 2.1) with differences VZA$-$SZA $< \pm 9^\circ$. As EPIC observations move towards the edge, the pseudo-spherical model atmosphere deviates more from Earth's spherical atmosphere in accounting for atmospheric attenuation and multiple scattering, resulting in more significant errors in modeled radiances, whose maximum errors increase to about $\pm 1\%$ at 75° VZA and about $\pm 2\%$ at 85° VZA (Caudill et al., 1997). RRS corrections are included in the forward modeling, and they are well within $\pm 1\%$ for EPIC UV bands (e.g., see Fig. 1.9). Incomplete RRS corrections are expected to add less than $\pm 0.1\%$ to the forward modeling errors.

Unlike the calibration biases being insensitive to observing conditions and having no correlation among different bands, the forward modeling errors vary with absorber amounts and surface reflection and over- or underestimate similarly for all the UV bands depending on the viewing and illumination geometry. How these radiance errors propagate into the retrieved columns can be estimated using Eqs. (3.5) and (3.6), with error source terms replaced by ϵ_f and $\Delta \mathbf{R}$ due to modeling errors in the long-wavelength bands. With the radiance errors estimated above, these equations yield retrieval errors up to $\pm \sim 0.6$ DU and $\pm \sim 0.3$ DU when VZA $< 50°$, increasing to $\pm \sim 1.5$ DU and $\pm \sim 1$ DU at VZA = 75°, and $\pm \sim 5$ DU and $\pm \sim 15$ DU at VZA = 85°, respectively for O_3 and SO_2 vertical column errors. These are systematic errors and vary between high and low biases spatially depending on observing conditions, especially the viewing and illumination geometry.

3.2.4 Profile errors

As described in section 2.2, a column-dependent O_3 profile, whose tropospheric integration matches the climatological tropospheric column, is used to specify the vertical distribution of a retrieved total O_3 column. This retrieved profile ($\boldsymbol{\omega}_L$), which represents likely vertical distribution of the retrieved O_3 vertical column, differs invariably from the actual profile ($\boldsymbol{\omega}$). The O_3 error ($\Delta\Omega$) due to a profile errors ($\boldsymbol{\omega} - \boldsymbol{\omega}_L$) can be quantified using the first term on the r.h.s of Eq. (3.3), which is regulated by the retrieval AK ($\mathbf{A}_\Omega$). Examples of AKs for EPIC total O_3 retrievals are shown in Fig. 3.2 (panels a and c), illustrating how $\mathbf{A}_\Omega$ changes with viewing geometry. For low VZAs ($< 55°$), O_3 AKs are close to 1 above the upper troposphere, and therefore, profile mismatches in this altitude region result in insignificant retrieval errors. However, profile mismatches produce sizeable retrieval errors for high VZAs. In the troposphere, O_3 AKs change with surface reflectance in addition to angular dependence. Under cloud-free conditions, O_3 AKs drop quickly towards low-reflectivity surfaces, more so at high zenith angles (e.g., see Fig. 3.2a). Above highly reflective surfaces (e.g., snow, ice, or bright clouds), O_3 AKs increase drastically (see Fig. 3.2c), indicating enhanced sensitivity to tropospheric profile, especially at low zenith angles. Evidently, O_3 errors due to profile mismatches primarily come from the troposphere for low zenith angles for both low and high reflectivity surfaces, though stratospheric contributions increase substantially with higher zenith angles, more significantly for a low reflectivity surface. In general, errors due to the profile shape are reduced for high reflectivity surfaces.

Over a short period (e.g., one day), O_3 errors due to profile mismatches are local biases (reductions or enhancements) that vary with location smoothly. However, they are random errors since mismatches ($\boldsymbol{\omega} - \boldsymbol{\omega}_L$) are normally distributed around their near-zero means over a long period (e.g., one month). The variances of O_3 errors ($\Delta\Omega$) can be written as

$$\begin{aligned}\mathrm{Var}(\Delta\Omega) &= \mathbf{E}\left[\left((\mathbf{A}_\Omega - \mathbf{1})(\boldsymbol{\omega} - \boldsymbol{\omega}_L)\right)^2\right] \\ &= (\mathbf{A}_\Omega - \mathbf{1})\mathbf{E}\left[(\boldsymbol{\omega} - \boldsymbol{\omega}_L)(\boldsymbol{\omega} - \boldsymbol{\omega}_L)^T\right](\mathbf{A}_\Omega - \mathbf{1})^T \\ &= (\mathbf{A}_\Omega - \mathbf{1})\mathbf{S}_{n_l}(\mathbf{A}_\Omega - \mathbf{1})^T \qquad (3.9)\end{aligned}$$

where the expected values (i.e., the statistical means), $\mathbf{E}\left[(\boldsymbol{\omega} - \boldsymbol{\omega}_L)(\boldsymbol{\omega} - \boldsymbol{\omega}_L)^T\right]$, are O_3 profile covariance matrices $\mathbf{S}_{n_l}$, which depend on total columns (Ω), season, and latitude. This random component is estimated as a function of VZA using the column-dependent $\mathbf{S}_{n_l}$ from the M2TCO3 climatology (Yang and Liu, 2019). Figure 3.3 shows that the standard deviation of this error component increases gradually with higher VZA, from 1% at nadir to 1.7% at 75°, then rapidly with further elevated VZA.

Retrieval of SO_2 requires knowledge of the altitude at the center of the volcanic plume, which can be represented by a narrow (e.g., a width of 1.5 km) GDF. Error in the plume altitude leads to SO_2 retrieval error, which can be estimated using the retrieval AK. Figure 3.2 (panels b and d) shows sample AKs and their variations with VZA for an SO_2 plume center at 11 km altitude. The values of these AKs are equal to 1 at 11 km, meaning no retrieval error when the

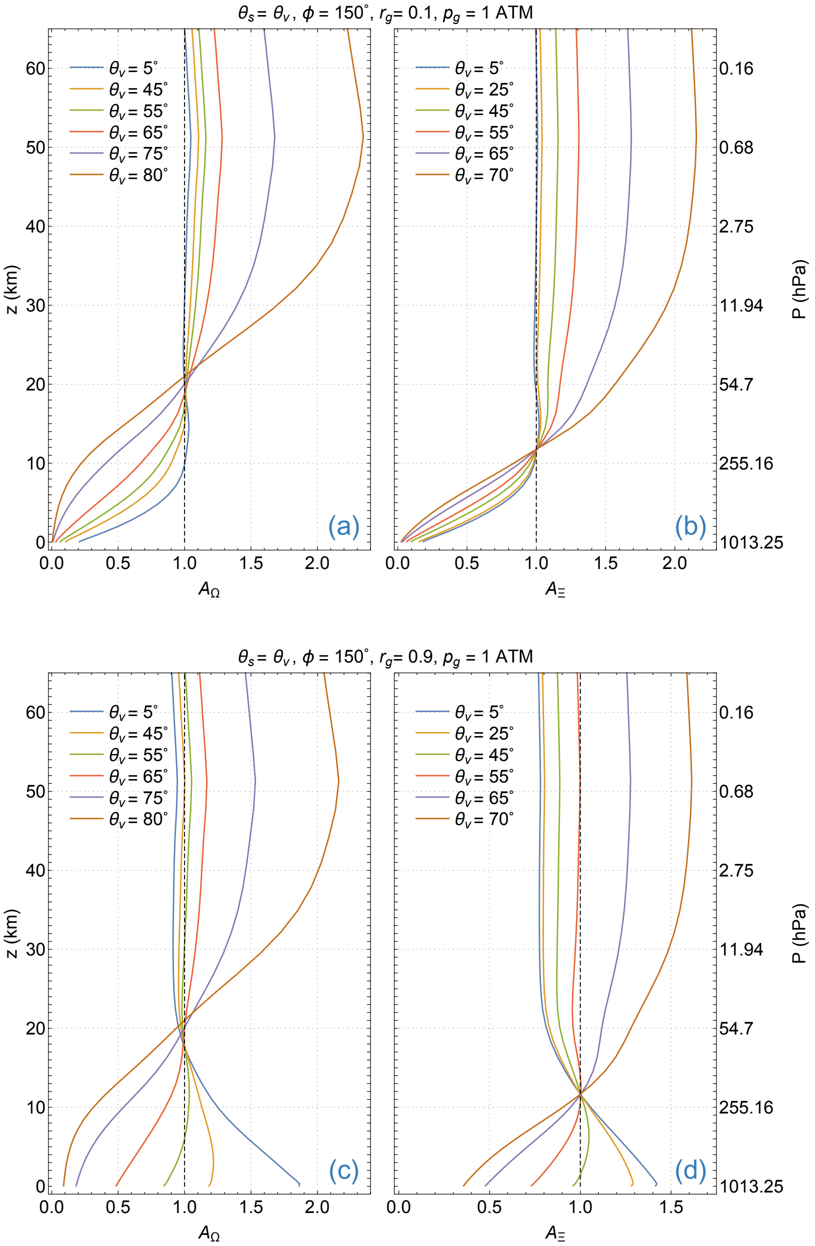


Figure 3.2: Examples of EPIC total O_3 AKs (a, c) and SO_2 AKs (b, d) as functions of geometric altitude (z) above seal level for several VZAs (θ_v). These AKs are calculated for a molecular atmosphere at mid-latitude with 275 DU total O_3 and 30 DU of SO_2 in a layer (1.5 km thick) at 11 km altitude over a low reflectance (a, b) and a high (c, d) reflectance Lambertian surfaces. Observing conditions are listed at the top of panels.

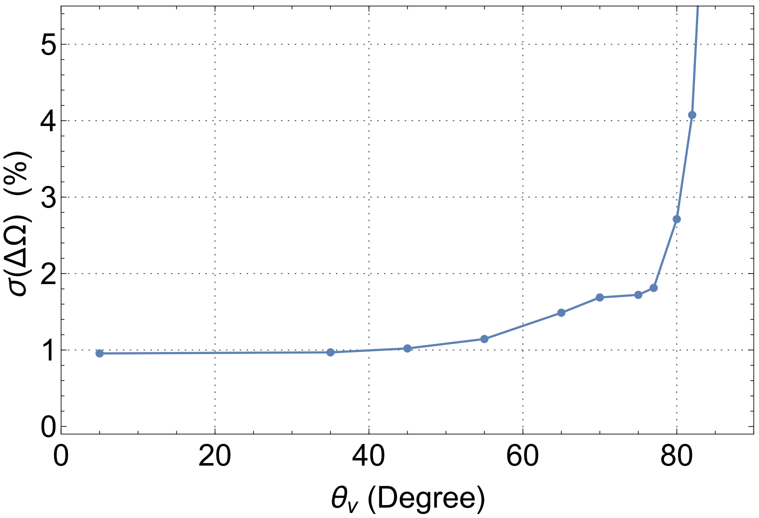


Figure 3.3: The standard deviation (σ) of O_3 errors due to profile mismatches as a function of the viewing zenith angle (θ_v), estimated using the M2TCO3 profile covariance matrix for the December mid-latitude zone (40°N–50°N).

altitude used in the retrieval is equal to the actual plume altitude. Here, we examine the case of a low reflectivity surface. The retrieved SO_2 column overestimates (underestimates) the actual column when the plume altitude is higher (lower) than the assumed altitude (11 km). At low VZAs ($< 35°$), AK values are close to 1 above the assumed altitude (11 km), indicating small (less than a few percents) errors for plumes at higher altitudes. Overestimation (up to 150%) increases quickly with larger VZAs when the plume is at higher altitudes. Underestimation is more severe with a lower altitude of the plume, by 10 to 20% per 1 km lower than the assumed altitude.

3.2.5 Errors from Lambertian treatment of natural surfaces

Reflections from surfaces are anisotropic but treated as isotropic. To estimate errors in O_3 and SO_2 columns due to this simplification, we performed DVCF retrieval from simulated radiances. First, TOA radiances of the four EPIC UV bands are modeled using a state-of-art

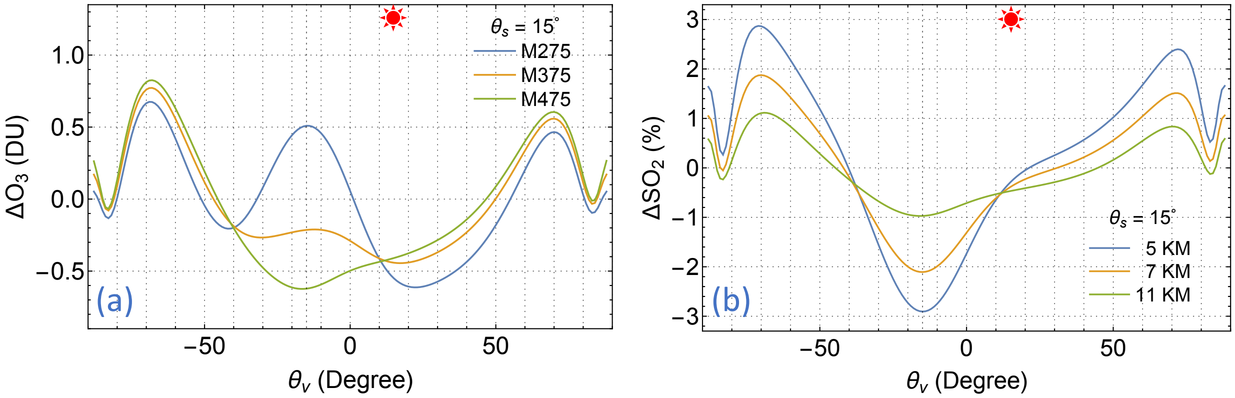


Figure 3.4: Errors in retrieved O_3 and SO_2 due to Lambertian surface treatment of an anisotropic surface. (a) O_3 errors in DU for midlatitude O_3 profile with total columns of 275 DU, 375 DU, and 475 DU. (b) SO_2 errors in percent for SO_2 layer at altitudes of 5 km, 7 km, and 11 km above sea level. See Fig. 1.3 caption for the specification of surface BRDF and viewing and illumination geometry.

radiative transfer model, VLIDORT (Spurr, 2006), for a molecular atmosphere with various O_3 and SO_2 profiles over a surface characterized by an anisotropic BRDF. Next, GLERs are determined at the long-wavelength bands (B3 and B4) and then linearly extrapolated to the short-wavelength bands (B1 and B2). Finally, retrieved O_3 and SO_2 columns from simulated B1 and B2 radiances using the extrapolated GLERs are compared with the column settings of the forward modeling to quantify retrieval errors. Examples of O_3 and SO_2 errors determined this way are shown in Fig. 3.4 for observing conditions described in Fig. 1.3 caption. In the closed-loop testing, surface reflection is specified by the Cox-Munk BRDF, which is highly anisotropic, more so than the land surface BRDFs that are well characterized by the combinations of Lambertian, Ross, and Li kernels (Lucht et al., 2000; Schaaf et al., 2011). Hence, the Cox-Munk BRDF selection provides error ranges due to the Lambertian treatment of surface reflections. Closed-loop tests are performed for a wide range of viewing-illumination geometries and vertical distributions of O_3 and SO_2. Test results (e.g., see Fig. 3.4) show that errors in total O_3 are mostly within ±1 DU, while SO_2 errors are within ±5% percent for SO_2 layers above 5 km, and decrease (increase) with higher (lower) layer altitudes. As shown in Fig. 1.4, the AMF errors due to Lambertian treatment occur below 20 km altitude. Consequently, a small fraction of the O_3 profile is affected by this approximation. Thus, O_3 errors are proportional to the tropospheric columns but are insensitive to the total column amounts. Since a vast majority of volcanic SO_2 clouds are located below 20 km altitude, SO_2 errors are proportional to the total SO_2 columns. Higher SO_2 clouds are not affected by this treatment.

3.2.6 Errors from MLER treatment of clouds and aerosols

The MLER model is adopted to treat atmospheric particles, including clouds and aerosols, which reside predominantly in the lower troposphere. The modeled light paths (especially in the troposphere) based on this treatment differ significantly from those for light transfer through the particle-laden atmosphere (e.g., see Fig. 1.8). The retrieval errors due to this simplification are again estimated using closed-loop testing. First, TOA radiances of the EPIC UV bands are simulated using VLIDORT for particle-laden atmospheres with various O_3 and SO_2 profiles over Lambertian surfaces of different reflectivities. Then inversion from the simulated radiances with the MLER treatment permits the identification of conditions under which retrieval errors are significant.

Clouds

The error in total O_3 due to the MLER treatment of a low-lying (below 10 km) cloud is mostly within ±2 DU (e.g., Fig. 3.5a). This O_3 error decreases slightly with a lower cloud altitude (or higher cloud pressure) but is insensitive to the cloud fraction or the total O_3 column.

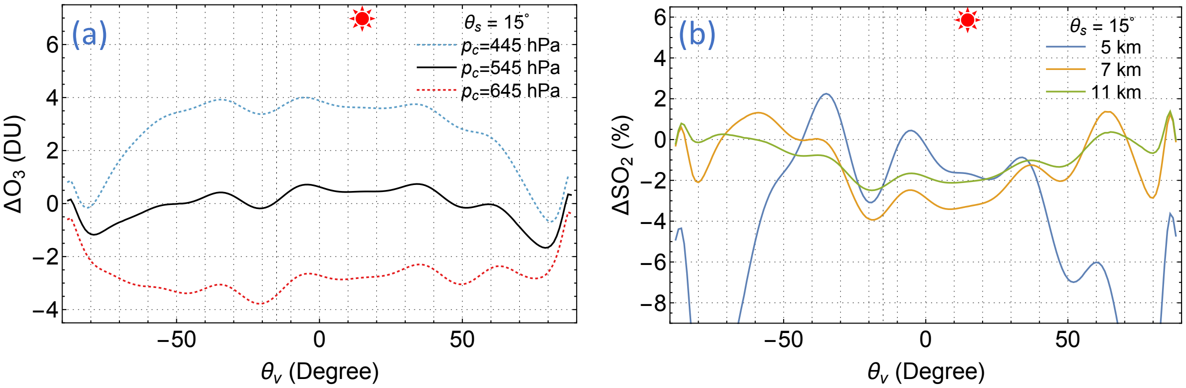


Figure 3.5: Errors in retrieved O_3 and SO_2 due to MLER treatment of clouds, which are represented by 1.5 km thick C1 particle layers with an optical thickness $\tau = 15$ at 340 nm. (a) O_3 errors in DU for correct ($p_c = 545$ hPa) and biased ($p_c = 545 \pm 100$ hPa) cloud OCPs. (b) SO_2 errors in percent for SO_2 layers at three altitudes (5, 7, and 11 km) above a layer of cloud (at 3 km altitude). See Fig. 1.3 caption for the specification of viewing and illumination geometry.

In other words, the MLER treatment does not contribute to large uncertainty in the retrieved total O_3 column, provided that an accurate OCP for the cloud is used for the MLER cloud surface. However, OCP has some uncertainty, contributing to additional uncertainty in the O_3 column: a low (high) bias in OCP results in a positive (negative) error in total O_3, quantitatively ±100 hPa causes about ∓4 DU (see Fig. 3.5a). The OCP uncertainty is estimated to be within ±50 hPa, thus contributing ∓2 DU to the total O_3 uncertainty. Combining O_3 uncertainties due to the OCP error and the MLER treatment yields ±4 DU uncertainty in total O_3 under cloudy conditions.

The error in the total SO_2 column due to the MLER cloud treatment is within ±2% when the SO_2 layer in the troposphere is well above the underlying cloud. This SO_2 error increases with a smaller separation between the SO_2 layer and the cloud, reaching ±15% when the SO_2 layer is just above the cloud. These characteristics of SO_2 error are illustrated in Fig. 3.5b. In contrast to the MLER-treatment O_3 error, which is insensitive to the total column, this SO_2 error is proportional to the total SO_2 column. When the SO_2 layer is below or within the cloud, the uncertainty of SO_2 quantification increases drastically. Depending on the relative distributions of

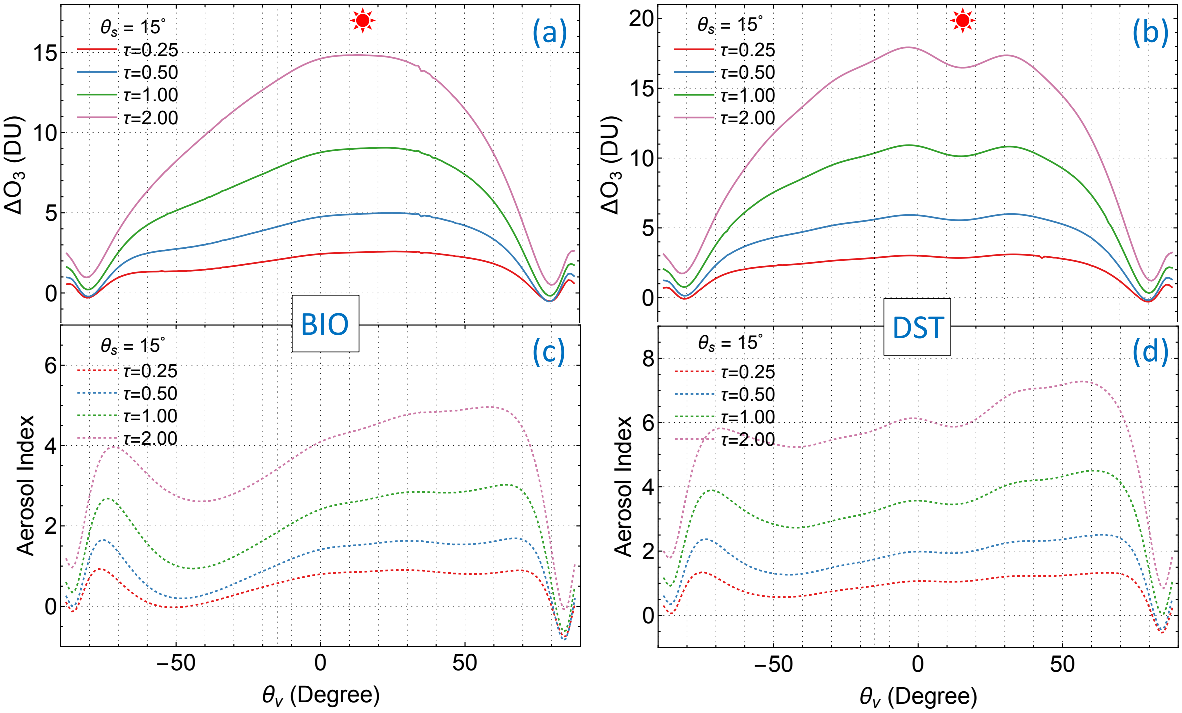


Figure 3.6: Errors in retrieved O_3 due to MLER treatment of two common UV-absorbing aerosols, (a) BIO ($\omega = 0.921$) and (b) DST ($\omega = 0.900$), with various optical thicknesses (τ = 0.25, 0.5, 1.0, and 2.0 at 340 nm) located at 5 km altitude. See Fig. 1.3 caption for the specification of viewing and illumination geometry. The AIs associated with each observation scenario are shown in panels (c) and (d).

SO_2 and the cloud particles, the retrieved SO_2 based on the MLER treatment can be a fraction of or a few times the actual column.

Aerosols

Besides clouds, the MLER treatment is applied to IFOVs contaminated by aerosols, which reside primarily in the troposphere and cover a significant portion of Earth's surface. These aerosols are suspended tiny (micron-scale) particles that scatter and possibly absorb sunlight. The frequently observed non-absorbing (or weakly absorbing) aerosols are sea salt and sulfate (SLF), and UV-absorbing aerosols are smoke (i.e., carbonaceous aerosols from biomass combustion, BIO), mineral dust (DST), and volcanic ash. Moderate and high positive AI values indicate the presence of UV-absorbing aerosols in an IFOV, while negative and slightly positive AI values the presence of non-absorbing or weakly absorbing aerosols.

Closed-loop testing shows that MLER treatment of non-absorbing and weakly absorbing aerosols reside in the lower troposphere ($<$ 7 km) result in small ($< \pm 2$ DU) errors in total O_3 retrievals, provided that the proper OCP for the elevated cloud surface is used. This error range is nearly independent of the total O_3 column or the aerosol loading.

The MLER treatment errors for UV-absorbing aerosols close to the surface ($<$ 1 km altitude) are mostly within ± 1 DU, similar to the error range associated with the LER treatment of BRDF surfaces. For elevated UV-absorbing aerosols, the MLER treatment and the linear $r_g f_c$ extrapolation scheme (see section 1.2.3) results in a positive bias in the retrieved total O_3 columns (e.g., see Fig. 3.6). This O_3 bias depends on the viewing-illumination geometry and generally increases with stronger aerosol absorption (i.e., lower single scattering albedo, ω), larger aerosol optical thickness, and higher altitude of aerosol layers. Regression analysis of results from closed-testing with many combinations of viewing-illumination geometries, particle-laden atmospheres (with various optical properties, optical thicknesses, and vertical distributions), surface reflectivities, and O_3 profiles reveals a positive correlation between column O_3 error and the UV AI. Quantitatively this relationship can be written as $\Delta O_3 = (1.5 \pm 1) \times$ AI DU for AI values greater than 0.5 and less than 8. This relationship provides a rough estimate of O_3 bias based on the observed AI. Typically, AI values fall between 1 and 4 with a median value of 1.5 for EPIC observations of UV-absorbing aerosols, corresponding to a mean O_3 bias of about 3 DU for IFOVs contaminated with UV-absorbing aerosols. The MLER treatment sometimes fails when aerosol absorption is strong such that the derived LER becomes negative. In this case, the explicit aerosol treatment may be needed to reduce the retrieval uncertainty.

EPIC's high-cadence observation has more chances to view volcanic clouds during or soon after eruptions. These young volcanic clouds contain mixtures of ash particles and water or ice clouds, as eruptions inject ash and gases (including SO_2) into the atmosphere. Since ash particles absorb UV strongly, the MLER treatment of volcanic plumes leads to huge uncertainties in the retrieved SO_2 columns, which are often over- or under-estimated greatly depending on the relative distributions between SO_2 and ash particles. An explicit treatment of volcanic ash is needed for accurate retrieval of SO_2 when ash particles are co-located with or slightly separated from the gas.

3.3 Error summary

Uncertainty estimates (section 3.2) have detailed various contributions to systematic and random errors in retrieved O_3 and SO_2 vertical columns. In this summary, these error types are separately combined to estimate their total systematic and random errors respectively.

First, random errors from various sources, including measurement noise and errors in temperature profiles and O_3 vertical profiles, are combined to estimate total random errors. The random O_3 error (characterized by its standard deviation) is 1.5% at low VZAs ($< 45°$), increasing to 2.0% at 75° for IFOVs without clouds and aerosols. Similarly for IFOVs with clouds or non-absorbing aerosols, the random error in total O_3 column is 1.8% for low VZAs ($< 45°$) and 2.5% at 75°. Random SO_2 error has two terms: one is independent of the SO_2 column (see section 3.2.1), but the other is proportional to this column (see section 3.2.2). Hence they are not combined.

Next, the radiometric biases, forward radiance modeling errors, and errors in molecular cross-sections are the primary contributions to the systematic errors in the retrieved vertical columns. The possible ranges due to these sources are provided earlier in this section, but their actual contributions are unknown. We obtain the total systematic error by combining the likely ranges (i.e., ~half of the possible ranges) of these contributions and estimate the bias in the total O_3 column to be $\pm2\%$ for VZA $\leq 70°$ and $\pm3.5\%$ for VZA $\leq 85°$, and the bias in the total SO_2 column high (> 10 km) in the atmosphere to be $\pm$ 1.5% for VZA $\leq 70°$ and $\pm$ 5% for VZA $\leq 85°$. Some biases depend on geophysical conditions. For instance, O_3 has a positive bias for IFOVs with UV absorbing aerosols, increasing with a higher AI (see section 3.2.6). Underestimation (overestimation) of a SO_2 column occurs when its layer altitude used in the retrieval is higher (lower) than the actual altitude (see section 3.2.3).

Chapter 4: Validation of EPIC O_3 and SO_2 products

4.1 O_3 Validation

We validate the DVCF O_3 retrievals from EPIC using ground-based Brewer spectrophotometer measurements and the assimilated O_3 product from MERRA-2, the Modern-Era Retrospective Analysis for Research and Applications, Version 2 (Gelaro et al., 2017).

We compare EPIC total O_3 columns with the Brewer O_3 data at ten selected ground stations with high-cadence measurements, distributed in five latitude zones. At each of these selected stations, a Brewer spectrometer makes a measurement every few ($\sim$ 10) minutes during the daylight hours each day, thus providing total vertical O_3 columns that are coincident (within $\pm$15 minutes) with EPIC observations at these stations. For inter-comparison, Brewer O_3 data are interpolated to the times when EPIC observes these locations. Coincident O_3 data from these two independent sources are displayed in the upper panels of Fig. 4.1(a-e), their differences in the lower panels of Fig. 4.1(a-e), and the EPIC vs. Brewer scatter plots in the right panels of Fig. 4.1(a-e). We include coincident data with VZA $\leq 70^\circ$ only for statistical analysis, due to EPIC-Brewer IFOV differences that usually increase with slant path lengths and EPIC's footprint sizes and due to calibration biases at large VZAs or SZAs. The mean difference and standard deviations in percent are displayed in the difference plots, while those in DU and the correlation coefficients are in the scatter plots. Time series of O_3 difference (see lower panels of Fig. 4.1(a-e)

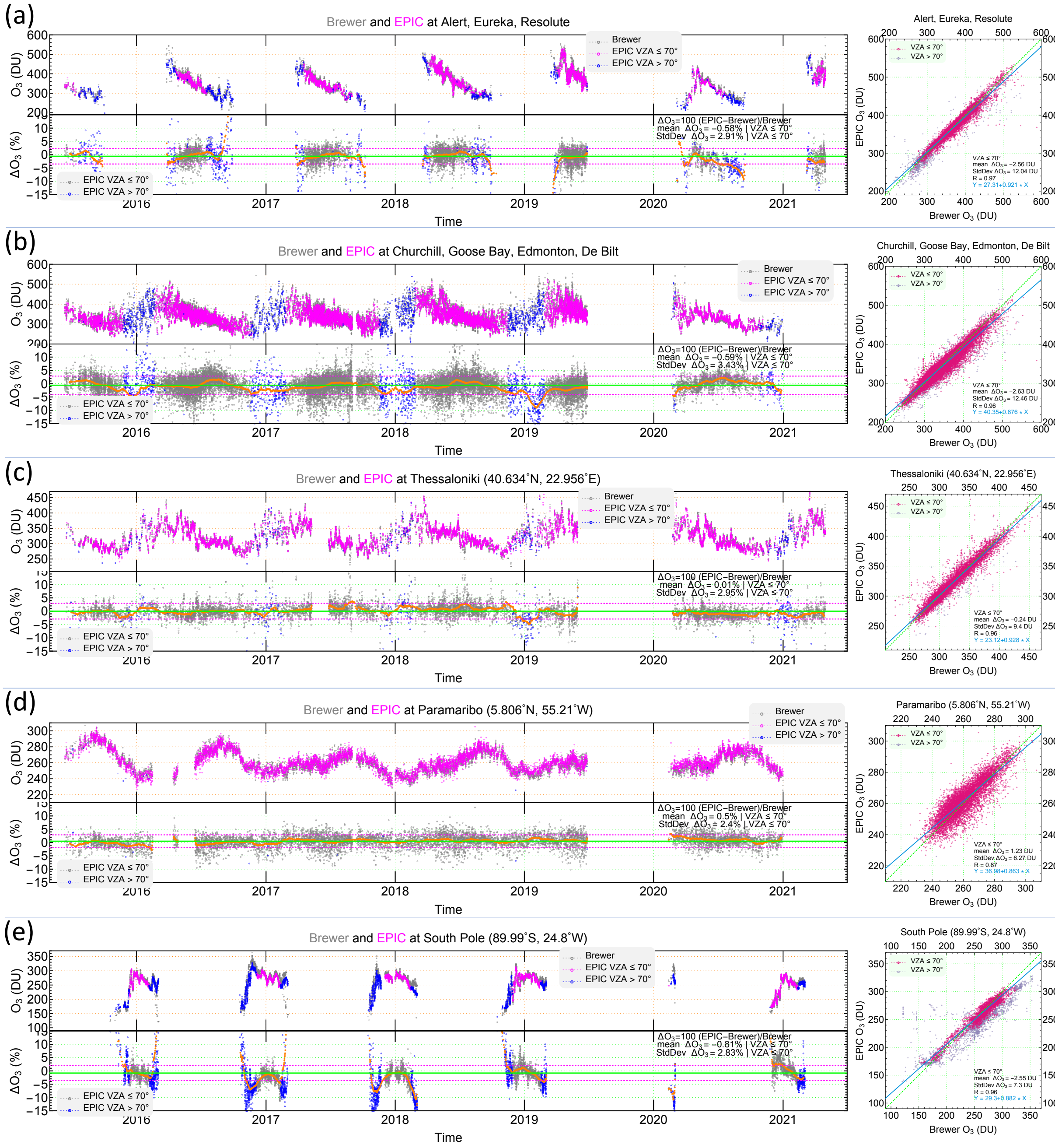


Figure 4.1: Inter-comparison of total O_3 from EPIC and the ground-based Brewer spectrophotometers at ten selected ground stations with high-cadence measurements: Alert (82.50°N), Eureka (79.99°N),), Resolute (74.72°N), Churchill (58.75°N), Edmonton (53.55°N), Goose Bay (53.31°N), De Bilt (52.10°N), Thessaloniki (40.63°N), Paramaribo (5.806°N), and South Pole (-89.99°N), from July 2015 – April 2021. EPIC and Brewer coincident pairs are used in the plots and data with VZA $\leq$ 70°N only are included in the difference statistics.

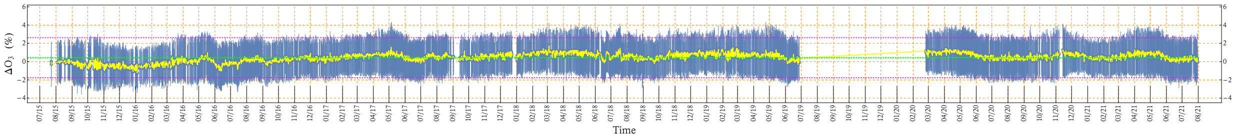


Figure 4.2: Comparison of synoptic EPIC O_3 with MERRA-2 assimilated O_3: time series of mean daily differences and standard deviations for EPIC observations with VZA $\leq 70°$.

between EPIC and Brewer are highly stable with similar moving averages and standard deviations from June 2015 to April 2021, showing that EPIC O_3 are consistent over time, without noticeable drift. The correlations between EPIC and Brewer are very high with the correlation coefficients $R \geq 0.96$ for most stations (except for the Paramaribo station near the equator, where R = 0.87), demonstrating that EPIC captures O_3 variability accurately. EPIC O_3 agrees with the Brewer measurements to better than 1% with standard deviations of differences less than 3.5% for all the ground stations, validating the high accuracy of EPIC total O_3.

From October 2004, MERRA-2 O_3 field is assimilated from Aura MLS and OMI and provides highly realistic global distributions of O_3 in the stratosphere and upper troposphere while inheriting the uncertainty characteristics of its sources (Davis et al., 2017; Stajner et al., 2008; Wargan et al., 2015). We compare the MERRA-2 synoptic O_3 field with the EPIC hemispheric view for the same observation time to access EPIC's capability in capturing the realistic O_3 distribution. For instance, strikingly similar O_3 spatial distributions are observed in EPIC measurements (Fig. 2.5c) and the MERRA-2 assimilation (Fig. 2.5d), with an agreement at -0.20 ± 2.52% (or -0.35 ± 5.6 DU, Fig. 2.5h). We expand this synoptic comparison to each EPIC hemispheric view obtained from July 2015 to August 2021 and plot in Fig. 4.2 the time series of daily statistics. This time series shows that nearly the same level of agreement is achieved for the entire period, with a mean bias and standard deviation of 0.44 ± 2.19% (or 1.16 ± 6.34 DU). Considering the mean bias (about –1.2%) of MERRA-2 total O_3 (Wargan et al., 2017), we estimate the accuracy of EPIC total O_3 to be –0.76 ± 2.19%.

4.2 SO_2 Validation

Volcanic eruptions occur sporadically and without warning but EPIC on DSCOVR, from the unique L1 vantage point, usually provides multiple daily observations of volcanic SO_2 and ash clouds once injected into the atmosphere. In contrast, ground-based instruments rarely detect volcanic clouds unless they drift over one in operation. We thus rely on polar-orbiting instruments, which may observe a volcanic cloud once (or more at high latitude) per day to provide validation measurements.

The OMPS-NM on SNPP provides high-quality hyperspectral measurements in the UV, from which highly accurate retrievals of O_3 and SO_2 are achieved using the DVCF algorithm. The DVCF algorithm can apply to both discrete spectral measurements (e.g., TOMS and EPIC) as well as hyper-spectral ones (e.g., OMI and OMPS-NM). The main difference is more information can be extracted from hyperspectral measurements to improve the accuracy and precision of the retrieved geophysical parameters. For instance, the altitude of a SO_2 layer can be determined in addition to its amount simultaneously using the DVCF algorithm (Yang et al., 2010). Having the altitude information significantly improves the accuracy of SO_2 quantification because the SO_2 measurement sensitivity varies strongly with its altitude. Thus DVCF height-resolved SO_2 retrievals from hyperspectral instruments, such as OMI and OMPS-NM, provide the most accurate quantification of SO_2 vertical columns. To validate DVCF SO_2 retrievals from EPIC, we compare the SO_2 mass loading of a volcanic plume integrated from EPIC observations with SNPP OMPS-NM for the same event.

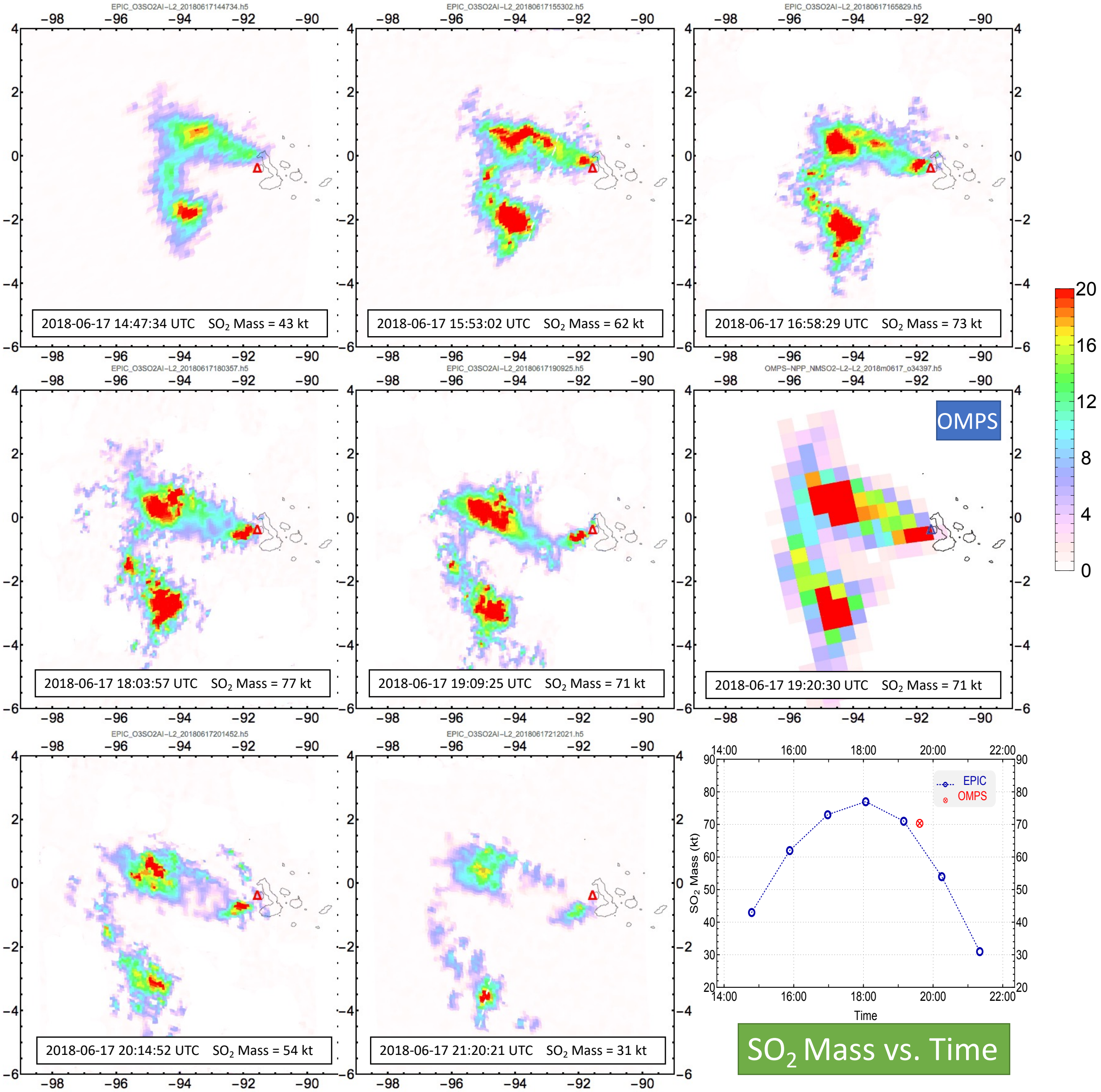


Figure 4.3: EPIC and OMPS observations of volcanic SO_2 plumes on 17 June 2018 from the eruption of Fernandina volcano (Δ) in the Galapagos Islands. This eruption injected significant amount on SO_2 into the troposphere at about 3.5 km above sea level. The mass loading of a SO_2 plume is obtained by summing the SO_2 masses of all IFOVs with SO_2 vertical columns $\geq$ 1 DU. The lower right panel plots the EPIC and OMPS SO_2 masses vs. the observation time (UTC).

Figure 4.3 compares the DVCF retrievals of the volcanic plume from the explosive eruption of Fernandina volcano in the Galapagos Islands on 17 June 2018. Seven plume exposures about 65 minutes apart are taken by EPIC on this day. Soon after the fifth EPIC exposure, the OMPS observed this plume for the first time. For the exposures at ten minutes apart, both instruments estimate the mass loading at 71 kt, validating the EPIC SO_2 result.

The lower right panel of Fig. 4.3 plots the plume mass vs. the observation time, showing the mass loading peaks near the local noontime. The observed mass change results from the continuing emission from the volcano, the conversion of SO_2 into sulfate, and the changing measurement sensitivity with viewing illumination conditions since low SO_2 columns may be missed at large (VZA or SZA or both) angles due to lower sensitivity. EPIC's high-cadence observations allow better identification of the peak loading of volcanic SO_2 plume, thus usually can provide more accurate estimates of the lower bound of SO_2 emission compared to polar-orbiting instruments.

We have conducted many mass loading comparisons between EPIC and OMPS and found that the agreements are usually within 20%. These findings indicate that DVCF SO_2 retrieval from EPIC provides better than 20% (an estimate of the upper error bound) accuracy in total mass for eruptions with greater than 50 kt emissions.

Chapter 5: NO_2 retrievals from NOAA-20 OMPS: instrument and algorithm

5.1 NOAA-20 OMPS instrument

OMPS Nadir Mapper (NM) is a nadir-viewing hyperspectral instrument that measures backscattered ultraviolet (UV) radiance spectra. The NOAA-20 OMPS spacecraft launched in November 2017, is the second of several OMPS missions planned for the next decade and beyond on the NOAA/NASA Joint Polar Satellite Systems (JPSS) spacecrafts, with the first OMPS mission launched in October 2011, aboard SNPP spacecraft. Similar to SNPP, NOAA-20 is in a Sun-synchronous orbit with a local ascending (northbound) equator crossing-time at 1:30 P.M. local time (LT), close in time to the Aura/OMI & TROPOMI overpasses at 1:45 P.M. LT (Table 5.1). NOAA-20 OMPS has a spatial resolution of 17×13 km^2 at the nadir, improved over the nadir resolution of 50×50 km^2 of SNPP OMPS, and OMPS resolution will be continually improved on the subsequent JPSS satellites.

NOAA-20 OMPS measures UV radiance in the 300-420 nm wavelength range at a spectral resolution of 1 nm and a sampling rate of 0.42 nm per pixel. Although NOAA-20 OMPS extends the spectral coverage to 420 nm (compared to SNPP OMPS in the 300-380 nm range), its radiance quality is poor for wavelength longer than 390 nm and thus not used for NO_2 retrieval, and the shorter wavelength spectra ($<$ 345 nm) are strongly affected by ozone absorption. Therefore, the 345-390 nm wavelength range was utilized for OMPS NO_2 retrieval, shorter in wavelength

Table 5.1: Comparison of satellite NO_2 instruments on Low Earth Orbit, including OMI, SNPP OMPS, NOAA-20 OMPS, and TROPOMI.

	OMI	SNPP OMPS	NOAA-20 OMPS	TROPOMI
Spectral window	405 – 465 nm	345 – 380 nm	345 – 390 nm	405 – 465 nm
Spectral resolution	0.63 nm	1 nm	1 nm	0.63 nm
Swath width	2600 km	2800 km	2800 km	2600 km
FOV	75°	110°	110°	75°
Signal-to-noise ratio	1200	2500	600-800[1]	1200
Nadir resolution	24×13 km^2	50×50 km^2	13×17 km^2	5.5×3.5 km^2
Overpassing time	13:45 LT	13:30 LT	13:30 LT	13:45 LT

[1] Note that the signal-to-noise ratio of NOAA-20 OMPS is estimated to be about $1/\sqrt{11}$ of that of SNPP OMPS.

than other legacy UV-Vis instruments (Table 5.1). We adopted the Direct Vertical Column Fitting (DVCF) technique to retrieve NO_2 from NOAA-20 OMPS-NM UV radiance, which is the algorithm currently implemented in the operational SNPP OMPS NO_2 product (Yang et al., 2014). Details about the DVCF algorithm and challenges for NO_2 retrievals in the UV spectra are elucidated in Section 5.2.

5.2 DVCF algorithm implemented for OMPS NO_2 retrievals

The Direct Vertical Column Fitting (DVCF) algorithm is applied to the NOAA-20 OMPS-NM spectral measurements to retrieve the atmospheric NO_2 vertical columns. The approach of this algorithm is to find retrieved parameters so that the modeled radiance spectra ($\mathbf{I_{TOA}}$) at the top of the atmosphere (TOA) match the satellite-measured spectra ($\mathbf{I_m}$). Algebraically, radiance matching is accomplished by minimizing the cost function $\left\|\Delta \mathbf{y} \mathbf{S_y}^{-\frac{1}{2}}\right\|^2$, where $\mathbf{S_y}$ is the measurement error covariance matrix and $\Delta \mathbf{y} = \{\ln \mathbf{I_m} - \ln \mathbf{I_{TOA}}\}$ is the residual vector for all wavelengths in a spectral window, one of which at wavelength λ can be written as:

$$\Delta\mathbf{y}(\lambda) = V\int_0^\infty \frac{\partial \ln I_{TOA}(\lambda)}{\partial \tau_z} S_z\sigma(\lambda, T_z)\,dz - \sum_i^m \xi_i\sigma_i(\lambda, T_i)$$
$$+\sum_{k=0}^n \frac{\partial \ln I_{TOA}(\lambda)}{\partial R}\Delta R_k(\lambda-\lambda_0)^k + \varepsilon \tag{5.1}$$

The least-square solution to the set of Eq. 5.1 described the retrieval of NO_2 vertical column (V) as a process of fitting the residuals with the vertical column weighting function (WF, i.e., $\int_0^\infty \frac{\partial \ln I_{TOA}(\lambda)}{\partial \tau_z} S_z\sigma(\lambda, T_z)\,dz$) and the slant columns $\{\xi_i, i = 1, ..., m\}$ of other trace gases (including O_3, HCHO, BrO, and OClO, thus m = 4) with their molecular absorption cross sections $\{\sigma_i(T_i), i = 1, ..., m\}$ at their respective temperature $\{T_i, i = 1, ..., m\}$. S_z is the shape factor, which is the normalized vertical profile; T_z is the atmospheric temperature, a function of altitude (z); and ε is the total error, which includes satellite measurement error and the forward modeling uncertainty. Here, τ_z is the optical thickness of an infinitesimally thin layer at z, and the total absorption optical thickness(τ) is the integration of τ_z: $\tau = \int_0^\infty \tau_z\,dz = V\int_0^\infty \sigma(T_z)S_z\,dz$. The radiance matching is primarily through adjusting the reflectivity parameters $\{R_k, k = 0, ..., n\}$, which specify the Mixed Lambert-Equivalent Reflectivity (MLER) model. Here n = 1, which describes the reflectivity change linearly with wavelength, a simplified treatment to account for aerosol effects. The spectral structures in the measured spectra are then reproduced by finding the correct vertical column (V) and other absorbers' slant columns (ξ_i).

After the direct retrieval of total vertical columns (V) as described in Eq. 5.1, OMPS stratospheric and tropospheric NO_2 vertical columns are separated using an orbit-based sliding median correction approach. The basic premise behind Stratosphere-Troposphere Separation (STS) is that the spatial distribution of stratospheric NO_2 is more homogeneous than that of tropospheric NO_2 due to the localized anthropogenic emission and short lifetime of the latter. The sliding median STS technique used in NOAA-20 OMPS retrieval was first developed for SO_2 retrieval in OMI (Yang et al., 2007, 2009a), and then applied in NO_2 retrieval in SNPP OMPS (Yang et al., 2014). It follows a simple procedure: first, retrieved total vertical columns are partitioned into stratospheric (V_s^i) and tropospheric components using tropopause inputs and the a priori shape factors. Second, the initial stratospheric columns get refined by locating and smoothing out the high-frequency structures that are attributed to the inaccuracies in a priori shape factors. Specifically, two empirical latitudinal bands (e.g., 2° and 20°, subject to modifications in certain conditions) are used to construct two smoothed stratospheric fields from the initial field along the orbital track for each cross-track position of a satellite orbit using the sliding median method, as detailed in (Yang et al., 2014). The smaller latitude band is used to generate a higher-frequency smoothed field (m_h) that retains possible tropospheric signals, while the larger band is used to construct a lower-frequency smoothed field (m_l) with minimal tropospheric contributions that is representative of background median values. Thus, the excesses (+) and deficits (-) of stratospheric NO_2 are obtained from the difference between the two smoothed fields ($m_h - m_l$). The corrected stratospheric NO_2 column is then adjusted as $V_s = V_s^i - (m_h - m_l)$. After the stratospheric vertical columns are consolidated, the corresponding tropospheric NO_2 columns (V_t) are then retrieved by solving a new set of linear equations:

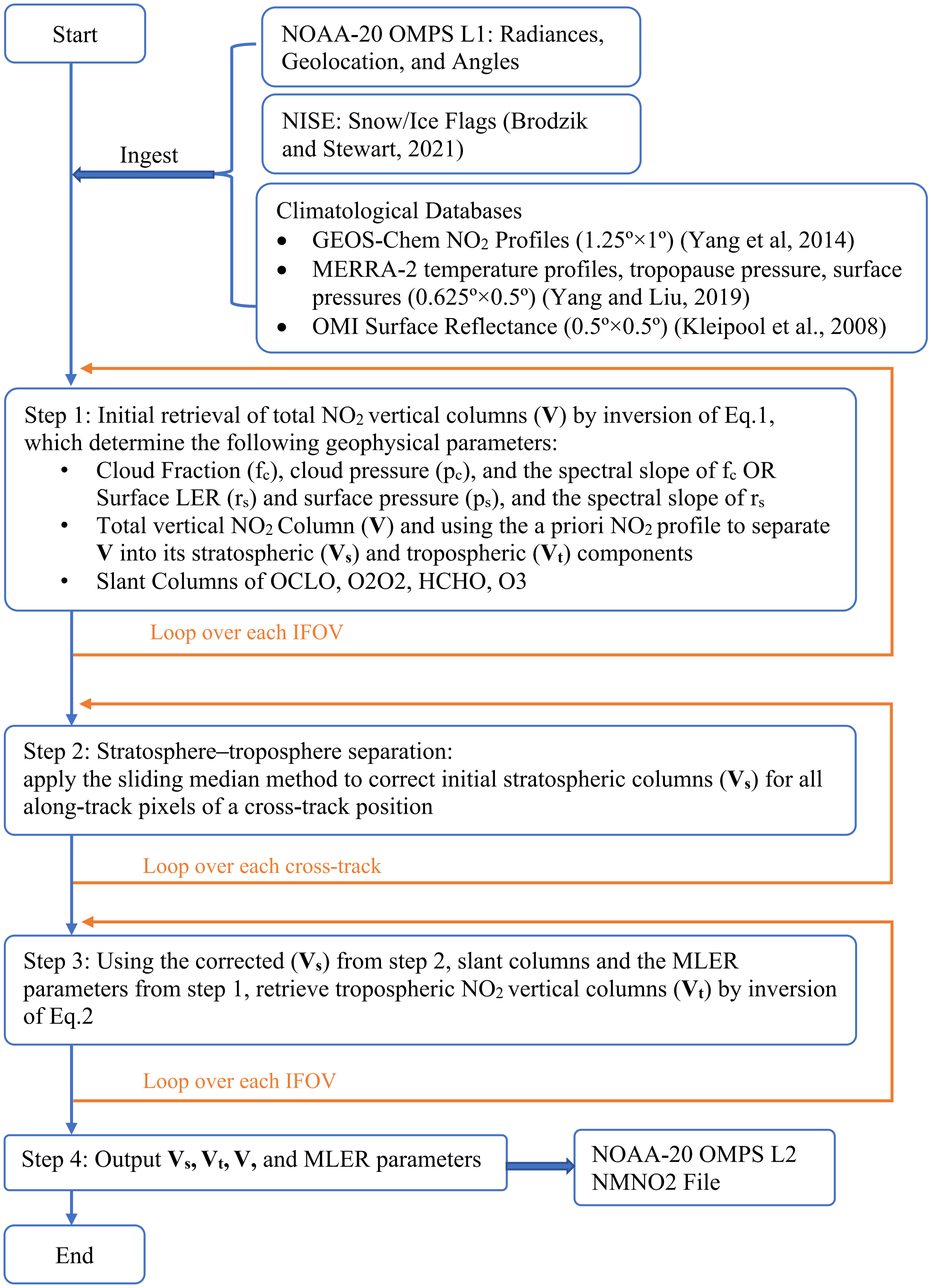
$$\begin{aligned}\Delta \mathbf{y}(\lambda) - V_s \int_{z_{tp}}^{\infty} \frac{\partial \ln I_{TOA}(\lambda)}{\partial \tau_z} S_z \sigma(\lambda, T_z) dz + \sum_i^m \xi_i \sigma_i(\lambda, T_i) \\ - \sum_{k=0}^{n} \frac{\partial \ln I_{TOA}(\lambda)}{\partial R} \Delta R_k (\lambda - \lambda_0)^k \\ = V_t \int_0^{z_{tp}} \frac{\partial \ln I_{TOA}(\lambda)}{\partial \tau_z} S_z \sigma(\lambda, T_z) dz + \varepsilon \qquad (5.2)\end{aligned}$$

, where Z_{tp} is the tropopause altitude. This completes the whole process of DVCF retrieval of OMPS tropospheric and stratospheric NO_2 vertical columns.

The key improvement of the DVCF algorithm over the traditional Differential Optical Absorption Spectroscopy (DOAS) approach lies in the more accurate representation of NO_2 measurement sensitivity, and thus more accurate NO_2 retrieval. In UV, the Rayleigh scattering from air molecules is quite strong and varies with wavelength drastically ($\sim 1/\lambda^4$). Consequently, the tropospheric air mass factors (AMFs) depend on the wavelength significantly. The spectrally dependent WF used in the DVCF captures the measurement sensitivity more accurately than the single-wavelength AMFs employed in the DOAS algorithm. Furthermore, retrieving surface reflectance or cloud fraction from the same spectral range, instead of taking it from ancillary inputs, such as climatological values or measurements from different spectra, improves the quantification of measurement sensitivity. Both improvements enable better spectral fits to the measured spectra and provide more accurate vertical column weighting functions, and thus allows more accurate and precise retrievals of NO_2 vertical columns than the traditional DOAS approach. Typically, the DOAS retrieval from UV spectra underestimates heavy NO_2 pollutions (> 2 DU) in the boundary layer by more than 10% compared to the corresponding DVCF retrieval.

With the theoretical background of the DVCF algorithm, here we summarize the algorithmic procedure applied to NOAA-20 Level-1 (L1) data to produce the Level-2 (L2) NO_2 product in the flowchart Algorithm 3, including references to the input ancillary (Brodzik and Stewart, 2021) and climatological data (Kleipool et al., 2008; Yang and Liu, 2019).

Algorithm 3 Flowchart that shows the processing of NOAA-20 OMPS by the DVCF algorithm.



NOAA-20 OMPS L1: Radiances, Geolocation, and Angles

NISE: Snow/Ice Flags (Brodzik and Stewart, 2021)

Climatological Databases

- GEOS-Chem NO$_2$ Profiles (1.25°×1°) (Yang et al, 2014)
- MERRA-2 temperature profiles, tropopause pressure, surface pressures (0.625°×0.5°) (Yang and Liu, 2019)
- OMI Surface Reflectance (0.5°×0.5°) (Kleipool et al., 2008)

Step 1: Initial retrieval of total NO$_2$ vertical columns (**V**) by inversion of Eq.1, which determine the following geophysical parameters:

- Cloud Fraction (f_c), cloud pressure (p_c), and the spectral slope of f_c OR Surface LER (r_s) and surface pressure (p_s), and the spectral slope of r_s
- Total vertical NO$_2$ Column (**V**) and using the a priori NO$_2$ profile to separate **V** into its stratospheric ($\mathbf{V_s}$) and tropospheric ($\mathbf{V_t}$) components
- Slant Columns of OCLO, O2O2, HCHO, O3

Step 2: Stratosphere–troposphere separation: apply the sliding median method to correct initial stratospheric columns ($\mathbf{V_s}$) for all along-track pixels of a cross-track position

Step 3: Using the corrected ($\mathbf{V_s}$) from step 2, slant columns and the MLER parameters from step 1, retrieve tropospheric NO$_2$ vertical columns ($\mathbf{V_t}$) by inversion of Eq.2

Step 4: Output $\mathbf{V_s}$, $\mathbf{V_t}$, **V**, and MLER parameters

NOAA-20 OMPS L2 NMNO2 File

5.3 Measurement sensitivity of NOAA-20 OMPS NO_2

The precision (sensitivity) of a satellite instrument is often assessed over remote areas, where the measurement variability is dominated by random errors originating from measurement noise. The measurement sensitivity of NOAA-20 OMPS NO_2 tropospheric vertical column densities (TVCDs) over the remote ocean (Indian Ocean) and remote desert (Arabian Peninsula) are 0.5×10^{15}, 0.7×10^{15} molecules/cm^2, respectively. The values are 1 σ (standard deviation) of the directly retrieved tropospheric NO_2 vertical columns (Figure 5.1). We adopted the same method used to quantify OMI NO_2 sensitivity as demonstrated in (Boersma et al., 2007; Valin et al., 2011). The areas selected to report OMPS sensitivity are 2° by 2° boxes between 56°E and 58°E, the box for the remote desert over Arabian Peninsula is between 21°N and 23°N, and the box for the remote Indian Ocean is between 15°N and 17°N (Figure 5.1). It is worth noting that we used vertical columns to report sensitivity instead of the slant columns as used by Valin et al. (2011), because the OMPS DVCF retrieval algorithm retrieves the vertical columns directly from a spectral fit to the Earth reflectance spectrum. In other words, the slant column is a derived quantity from the vertical column retrieval. Therefore, it makes more sense to use vertical columns to characterize OMPS sensitivity. For OMI, since slant columns are determined with spectral fit in the first step of the DOAS retrieval algorithm, it is better to use the slant column to quantify OMI sensitivity.

The sensitivity of SNPP OMPS NO_2 TVCDs is 0.4×10^{15} molecules/cm^2 (Yang et al., 2014), better than NOAA-20 OMPS. Although the two products are built on the same retrieval system, the NOAA-20 NO_2 is noisier than SNPP primarily because the NOAA-20 OMPS instrument has a smaller signal-to-noise ratio (SNR) than its predecessor SNPP OMPS (Table 5.1). Since

SNPP OMPS has a bigger pixel size (50 × 50 km^2) than NOAA-20 OMPS (17 × 13 km^2), if we were to estimate NOAA-20 SNR from SNPP, we can aggregate 11 NOAA-20 pixels into 1 SNPP pixel to make NOAA-20 equivalent to SNPP. This aggregation process cancels out noise but keeps the signal, which means that the NOAA-20 SNR is about $\sqrt{11}$ $\sim$3.32 times lower than SNPP. Therefore, NOAA-20 OMPS measurement sensitivity is intrinsically limited by its smaller signal-to-noise ratio and the DVCF retrieval algorithm is specially designed to amplify its measurement sensitivity as possible.

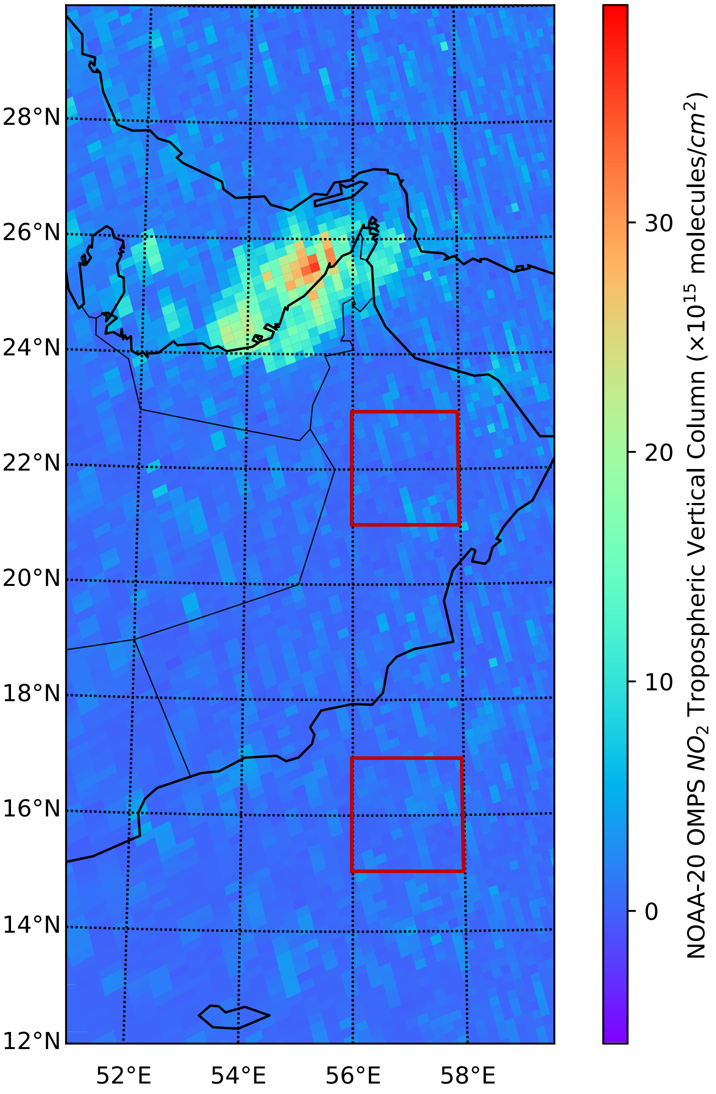


Figure 5.1: NOAA-20 OMPS NO_2 tropospheric vertical columns over the Eastern Arabian Peninsula on November 10, 2019. The scan time on the map is 09:00 to 09:05 UTC. The sensitivity of NOAA-20 tropospheric NO_2 columns is reported over the remote ocean and desert, where red boxes indicate.

Chapter 6: Evaluating NOAA-20 OMPS NO_2 retrievals: comparisons with OMI, Pandora, and a case study during COVID-19

6.1 Stratospheric NO_2: comparison with OMI

Before evaluating NOAA-20 tropospheric NO_2 retrievals, we first examine the stratospheric NO_2 observations from NOAA-20, since the stratospheric columns represent the clean background values over which tropospheric NO_2 enhancements are detected. We compared the seasonal averaged NO_2 stratospheric vertical column densities (SVCDs) observed from NOAA-20 OMPS and OMI in Figure 6.1. The daily NO_2 SVCDs (Level-2 data) collected from the two instruments were zonally averaged using 2° latitude bins for all cross-track iFOVs (OMI row anomaly affected pixels are excluded), and the seasonal averaged SVCDs were then plotted as a function of latitude. Since OMPS and OMI have similar overpassing time, the observed SVCDs are compared directly without photochemical corrections to compensate for NO_2 diurnal cycles (Rivas et al., 2014). In all seasons, the stratospheric NO_2 field is characterized by a tropical minimum over the equatorial NO_y (odd nitrogen) production zone, where total nitrogen is subject to upward and poleward transport. Outside the tropical regions, the stratospheric NO_2 field is characterized by a winter minimum and a summer maximum. The seasonal evolution of stratospheric NO_2 is explained by the sunlight-driven exchange between NO_x (nitrogen oxides) and other reservoir oxidized nitrogen species: N_2O_5 (primarily), HNO_3 and $ClONO_2$. As the amount of daily photolysis decreases over winter, NO_x begins to store into inactive N_2O_5 reservoirs, which results in a decrease of NO_x columns (Solomon and Garcia, 1983). Conversely, as the solar angle decreases in summer, the photolytic release of reservoir species increases NO_2 columns.

OMI and NOAA-20 OMPS retrievals of stratospheric NO_2 columns over the tropics and mid-latitude are very similar (Figure 6.1). In high latitudes, the differences are larger. This is primarily due to the sunlight-driven NO_2 diurnal variations at large solar zenith angles (SZA). The large SZA at higher latitude is more prone to the sharp NO_2 gradient at day-night transition, making direct column comparisons more difficult. In addition, large SZA increases the uncertainty in satellite retrieval of the NO_2 total columns due to stronger absorption in the stratosphere and lower signal-to-noise ratio. Studies found that the differences between satellite- and ground-based NO_2 measurements are generally larger for SZA above 45° (Ialongo et al., 2020). We have compared OMI cross-track positions that are not affected by row anomaly against the equivalent OMPS cross-track positions based on similar view zenith angle. We find that the row anomaly caused sampling mismatch contributes to some but is not the main reason for the large discrepancy at high latitudes (See Figure A.1 in Appendix Section A.1).

OMPS and OMI stratospheric NO_2 columns show an agreement with $r = 0.96$ and average relative difference = -3% for the region between 65°S and 65°N. The excellent agreement between NOAA-20 OMPS and OMI stratospheric columns is promising given that each relies on independent measurements and very different retrieval methodologies. Also, since stratospheric NO_2 is homogeneously distributed, this comparison is not subject to instrumental resolution difference.

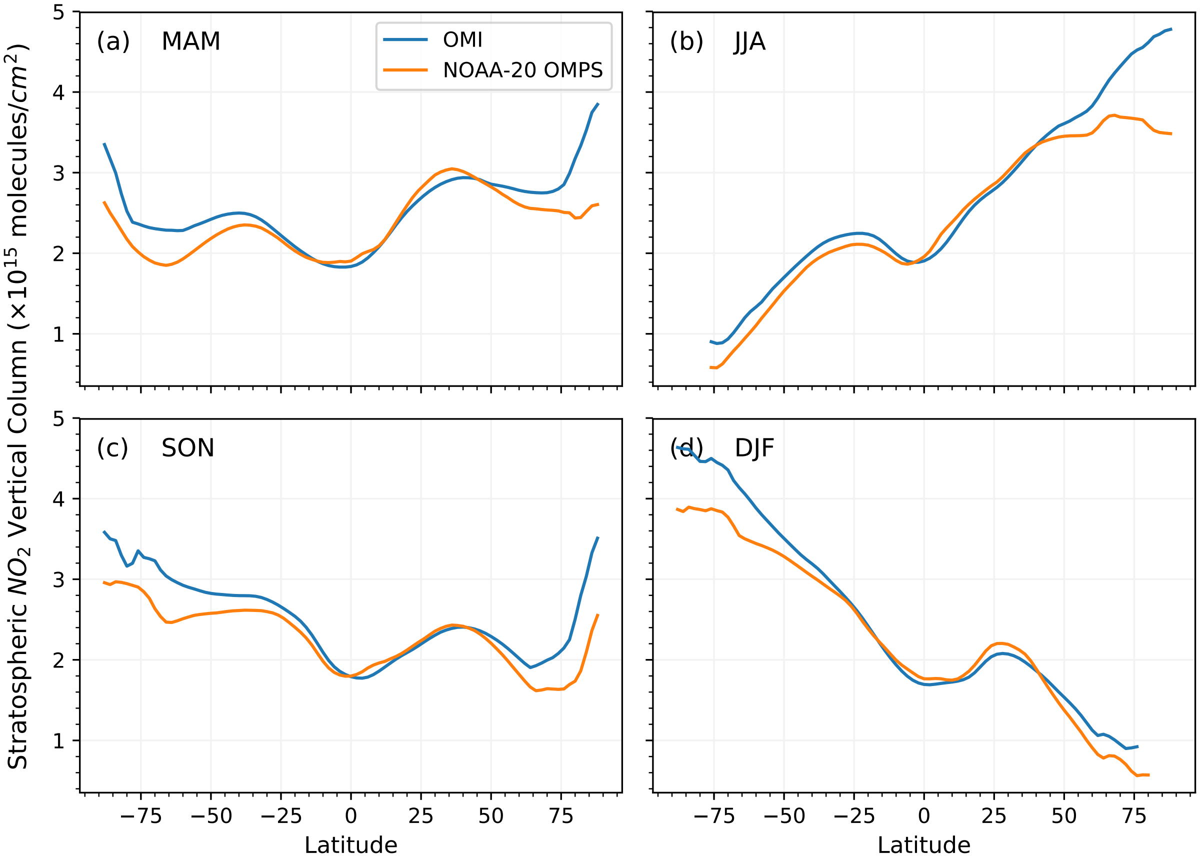


Figure 6.1: Seasonal averaged stratospheric NO_2 vertical columns observed from NOAA-20 OMPS (orange curve) and OMI (blue curve) as a function of latitude for **(a)** MAM, **(b)** JJA, **(c)** SON, **(d)** DJF, over the period from 2019-03-01 to 2020-04-30. OMPS and OMI show excellent agreement with r = 0.96 and mean relative difference = -3% for the region between 65°S and 65°N. OMI pixels affected by row anomaly are excluded.

6.2 Tropospheric NO_2: comparison with OMI

Figure 6.2 shows maps of the gridded monthly mean NO_2 tropospheric vertical column densities (TVCDs) derived from NOAA-20 OMPS and OMI for July and December 2019. OMPS monthly mean NO_2 TVCDs are derived from OMPS Level-2 data and are compared directly with OMI monthly mean columns derived from OMI Level-2 data using an identical gridding procedure. OMI and OMPS data are both gridded at 0.25° × 0.25° resolution, with the same cloud screening applied: iFOVs (pixels) with radiative cloud fraction $>30\%$ are excluded. OMI data affected by the row anomaly are also excluded. We computed OMPS and OMI monthly averages from respective Level-2 data in the following procedure: the value at each grid cell (0.25° × 0.25°) is determined by the weighted mean of the qualifying iFOVs that overlap with the grid cell over the month. The weight is an observation coverage, defined as the ratio of GridCell-iFOV overlapping area to the iFOV area. The gridding strategy is often called 'oversampling' over a long temporal window, and we use the same gridding method to generate OMPS Level-3 data and calculate mean NO_2 TVCDs over the designated periods in Section 6.4.

The monthly maps provide perspectives of where persistent tropospheric NO_2 enhancements are located. Places like the United States East Coast, Western Europe, East Asia, and northern India that exhibit elevated NO_2 pollution, are the world's major industrial and densely populated regions. Both OMPS and OMI observe these NO_2 enhancements. To highlight the similarities and differences between the two NO_2 products, we plot the longitudinal variations of OMPS and OMI measurements in July and December 2019 mean TVCDs across 38.625°N, where the highest OMI monthly mean value is found in December 2019 (Figure 6.2e, f). The NO_2 TVCDs from OMPS and OMI agree very well over China (between 100° and 140°) at this latitude, but

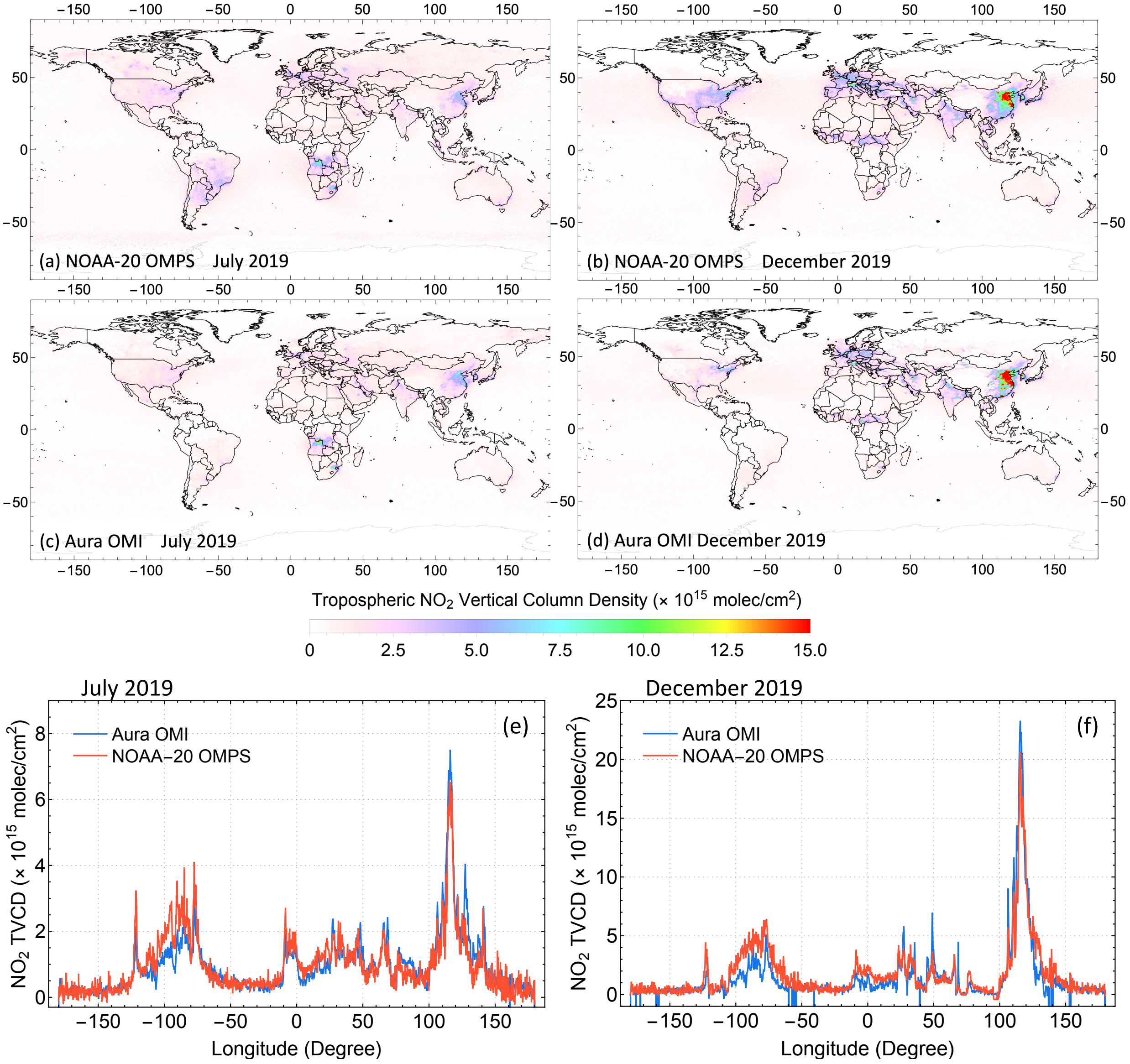


Figure 6.2: Monthly averages of NO_2 tropospheric vertical column densities (TVCDs) observed by **(a, c)** NOAA-20 OMPS for July 2019 and **(b, d)** OMI for December 2019, pixels with cloud fraction of 30% and above are excluded. **(e, f)** Quantitative comparison of NOAA-20 OMPS and OMI monthly averaged NO_2 TVCDs at 38.625°N, from 180°W to 180°E.

OMPS TVCDs are higher than OMI over the U.S. (between -100° and -60°) and Europe (between -10° and 20°). These differences are likely due to different a priori profile assumptions over these regions. The a priori NO_2 profiles used in the current NOAA-20 NO_2 product are taken from the monthly mean profiles of a 2012 GEOS-Chem global simulation at a coarse resolution (1° latitude × 1.25° longitude). These a priori profiles describe much higher boundary layer NO_2 concentrations than those of the more recent years. A higher boundary layer NO_2 in the a priori shape factors would result in higher NO_2 column retrievals. This potentially causes higher OMPS column NO_2 retrievals than OMI in the U.S. and Europe. On the other hand, for China, although more recent-year a priori profiles might reflect lower NO_2 concentrations benefited from environmental regulations, there is still a relatively large abundance of anthropogenic emissions near the surface compared to upper attitudes and thus the NO_2 vertical distributions (i.e., profile shapes) are not expected to change much. Therefore, the current agreement between OMPS and OMI in China would probably sustain in more recent-year a priori profiles. We are developing new a priori NO_2 profiles that are more appropriate for the current pollution levels to address the potential errors from inaccurate profile assumptions in the retrievals. Overall, the similar spatial patterns and good quantitative agreement demonstrate the high tropospheric NO_2 measurement sensitivity of NOAA-20 OMPS that is comparable to OMI.

6.3 Evaluating total NO_2 column with Pandora ground-based observations

The accuracy of NOAA-20 OMPS NO_2 columns measurements was preliminarily evaluated against Pandora ground-based observations over the continental United States (U.S.) during the period from 2019-02-14 to 2020-04-30 (Figure 6.3). Pandora instruments can retrieve NO_2

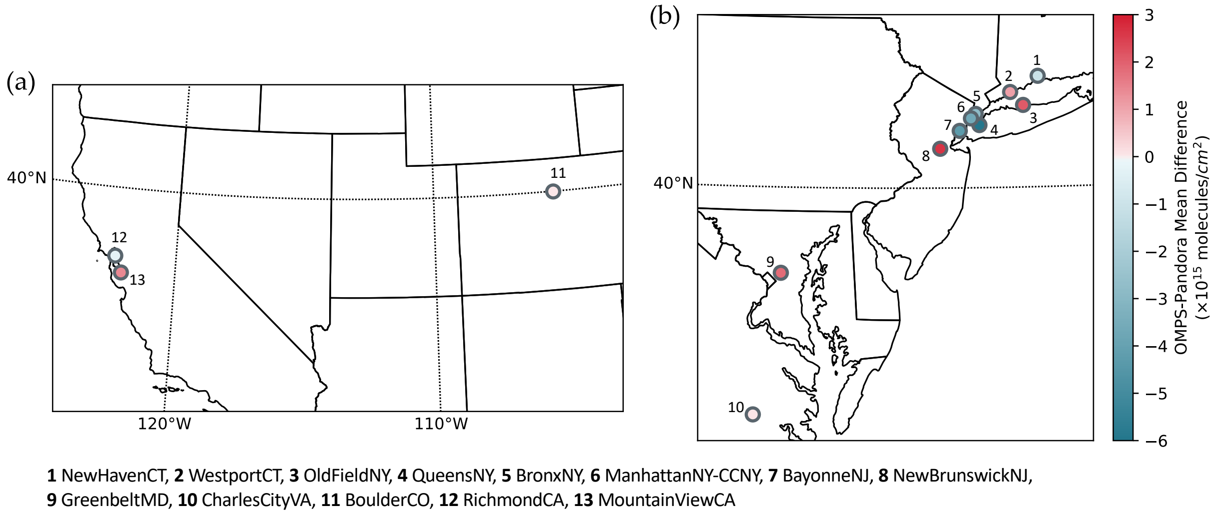


Figure 6.3: Locations of Pandora ground stations over **(a)** western U.S. (3 stations) and **(b)** eastern U.S. (10 stations), colored by the average difference between OMPS and Pandora measured total NO_2 columns.

vertical column densities (VCDs) through two viewing geometries, either direct-sun or zenith sky. For the time of interest, 13 Pandora instruments operated in direct-sun mode over the U.S. are compared to NOAA-20 OMPS column measurements. The direct-sun mode Pandora instruments provide high-quality reference measurements for evaluating trace gas retrievals from satellite sensors due to their low uncertainties in AMFs (Judd et al., 2020). The ground stations used in this analysis cover a variety of atmospheric environments, including 4 Pandoras located in the New York City (NYC) region: Manhattan NY-CCNY, Queens NY, Bronx NY, and Bayonne NJ (Figure 6.4b), and 9 other Pandoras located over mid-Atlantic and western U.S. states, representing urban/suburban/remote atmospheric conditions (Figure 6.3). All the sites are operated as part of the Pandonia Global Network (PGN; www.pandonia-global-network.org). Only high-quality Pandora measurements with a quality flag of 0 or 10 were included in this analysis.

For the comparison between OMPS and Pandora NO_2 total vertical columns, we adopted the following coincidence criteria: 1) the average Pandora total NO_2 VCDs are calculated within

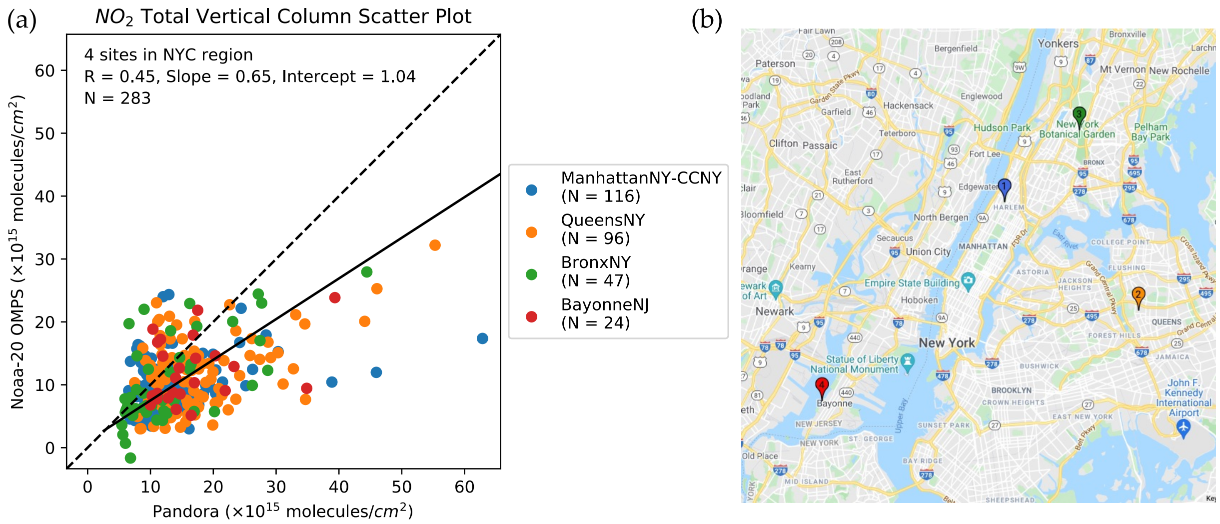


Figure 6.4: **(a)** Scatter plot of NOAA-20 OMPS and Pandora observed NO_2 total vertical columns over 4 ground stations in the New York metropolitan area, from 2019-02-14 to 2020-04-30. The statistics of linear regression fit are shown on the plot (N represents the number of coincidences). Note that different stations have different date spans and thus different number of coincidences with OMPS, coincidence by the station is shown in the legend; **(b)** The locations of 4 stations on Google Map, the color of each station on the map corresponds to the color used in the scatter plot.

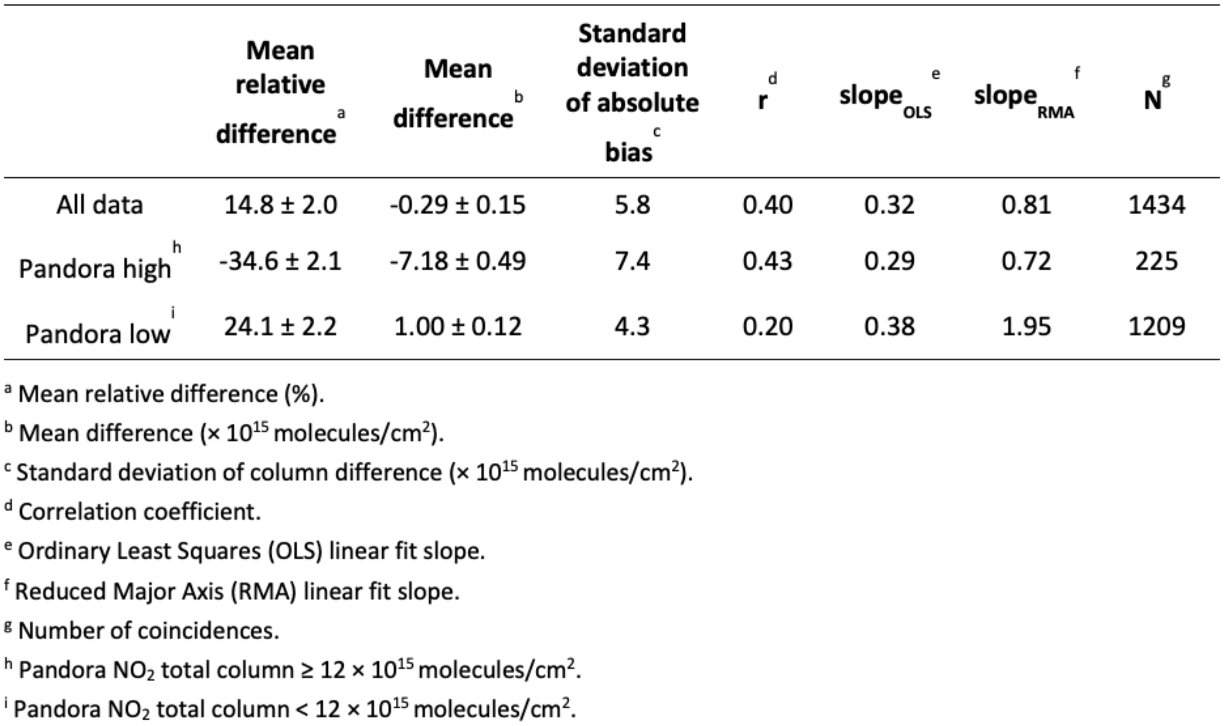
±30 min of OMPS overpass, and 2) all OMPS data have radiative cloud fractions less than 30%. The coincidence criteria are similar to those used in other validation studies (Ialongo et al., 2016; Judd et al., 2019). We calculated the linear regression statistics using Reduced Major Axis regression with correlation coefficient. This regression is chosen over Ordinary Least Square to recognize the potential errors/uncertainties in both evaluated and reference measurements. Note that the Ordinary Least Square statistics are also provided as a reference in Table 6.1. The difference and relative difference of the two column measurements are also calculated and analyzed, and are calculated in the following convention:

$$\text{column difference} = \text{OMPS measurement} - \text{Pandora measurement} \tag{6.1}$$

$$\text{relative difference } (\%) = \frac{\text{column difference}}{\text{Pandora measurement}} \times 100\% \tag{6.2}$$

Figure 6.4a shows the scatter plot and linear regression statistics of OMPS and Pandora NO_2 total columns coincidences from 4 sites over NYC area (N = 283). NOAA-20 OMPS has an average low bias of 28% (median relative difference, Figure A.2) and is moderately correlated (r = 0.45) with Pandora spectrometer measurements for the 4 NYC sites. The mean difference between OMPS and Pandora retrievals shows OMPS ubiquitously underestimates in the NYC region from -6.0×10^{15} (Queens NY) to -2.8×10^{15} (Bronx NY) molecules/cm^2 (Figure 6.3). Outside of the NYC metro area, the average OMPS column NO_2 is generally higher than or close to Pandoras, with the mean difference between -0.3×10^{15} (Richmond CA) and 2.7×10^{15} (New Brunswick NJ) molecules/cm^2, except for New Haven CT, which OMPS underestimates with an average difference of -1.1×10^{15} molecules/cm^2 from Pandora (Figure 6.3). To assess the statistical distribution of the OMPS biases, we plot the column NO_2 difference and percent difference as a function of pollution levels in Figure 6.5. For the least polluted columns ($< 3 \times 10^{15}$ molecules/cm^2), the inter-quantile range of column difference is 0.6 to 4.5×10^{15}, with a median of 3.3×10^{15} molecules/cm^2. When the pollution level increases, the median difference gradually shifts from positive towards negative. For the more polluted columns (12 - 15 and $> 15 \times 10^{15}$ molecules/cm^2), the inter-quantile range of column differences is both in the negative range, with a median difference of -4 and -10×10^{15} molecules/cm^2, respectively. Considering all data points from 13 sites during the 15-month validation span (N = 1434), the median difference and relative difference between NOAA-20 OMPS and Pandora are -0.1×10^{15} molecules/cm^2 and -1% respectively, with an inter-quantile range of -2.8 to 2.9×10^{15} molecules/cm^2 and -32% to 44% respectively (Figure 6.5). The overall linear correlation between NOAA-20 and Pandora total columns is 0.40 and the correlation is higher (r = 0.43) at higher pollution levels (Table 6.1). The quality of statistics of NOAA-20 OMPS is reasonably comparable to other

Table 6.1: Statistics of the comparison between NOAA-20 OMPS and Pandora NO_2 total columns, based on all data from 13 U.S. Pandora stations during 2019-02-14 to 2020-04-30. The uncertainties are the corresponding standard errors of the mean.



	Mean relative difference[a]	Mean difference[b]	Standard deviation of absolute bias[c]	r[d]	slope$_{\text{OLS}}$[e]	slope$_{\text{RMA}}$[f]	N[g]
All data	14.8 ± 2.0	-0.29 ± 0.15	5.8	0.40	0.32	0.81	1434
Pandora high[h]	-34.6 ± 2.1	-7.18 ± 0.49	7.4	0.43	0.29	0.72	225
Pandora low[i]	24.1 ± 2.2	1.00 ± 0.12	4.3	0.20	0.38	1.95	1209

[a] Mean relative difference (%).

[b] Mean difference (× 10^{15} molecules/cm^2).

[c] Standard deviation of column difference (× 10^{15} molecules/cm^2).

[d] Correlation coefficient.

[e] Ordinary Least Squares (OLS) linear fit slope.

[f] Reduced Major Axis (RMA) linear fit slope.

[g] Number of coincidences.

[h] Pandora NO_2 total column ≥ 12 × 10^{15} molecules/cm^2.

[i] Pandora NO_2 total column < 12 × 10^{15} molecules/cm^2.

satellite instrument bias with regard to Pandora measurements, see Text A.2 for details (Herman et al., 2019; Ialongo et al., 2016, 2020; Judd et al., 2019; Lamsal et al., 2014).

These results from multiple Pandora spectrometer instruments indicate that OMPS NO_2 total columns underestimate for relatively large Pandora NO_2 total columns, corresponding to polluted urban regions and episodes of elevated pollution, while overestimate for relatively small NO_2 total columns. The low bias (OMPS underestimation) can be partially attributed to the sampling mismatch in spatial representativity between a point measurement from the ground-based spectrometer and an area-averaged quantity from the satellite iFOV (instantaneous Field of View, i.e., pixel). As the more polluted NO_2 columns observed by Pandora are likely occurring over spatial scales much smaller than the satellite resolutions, the satellite-to-Pandora linear

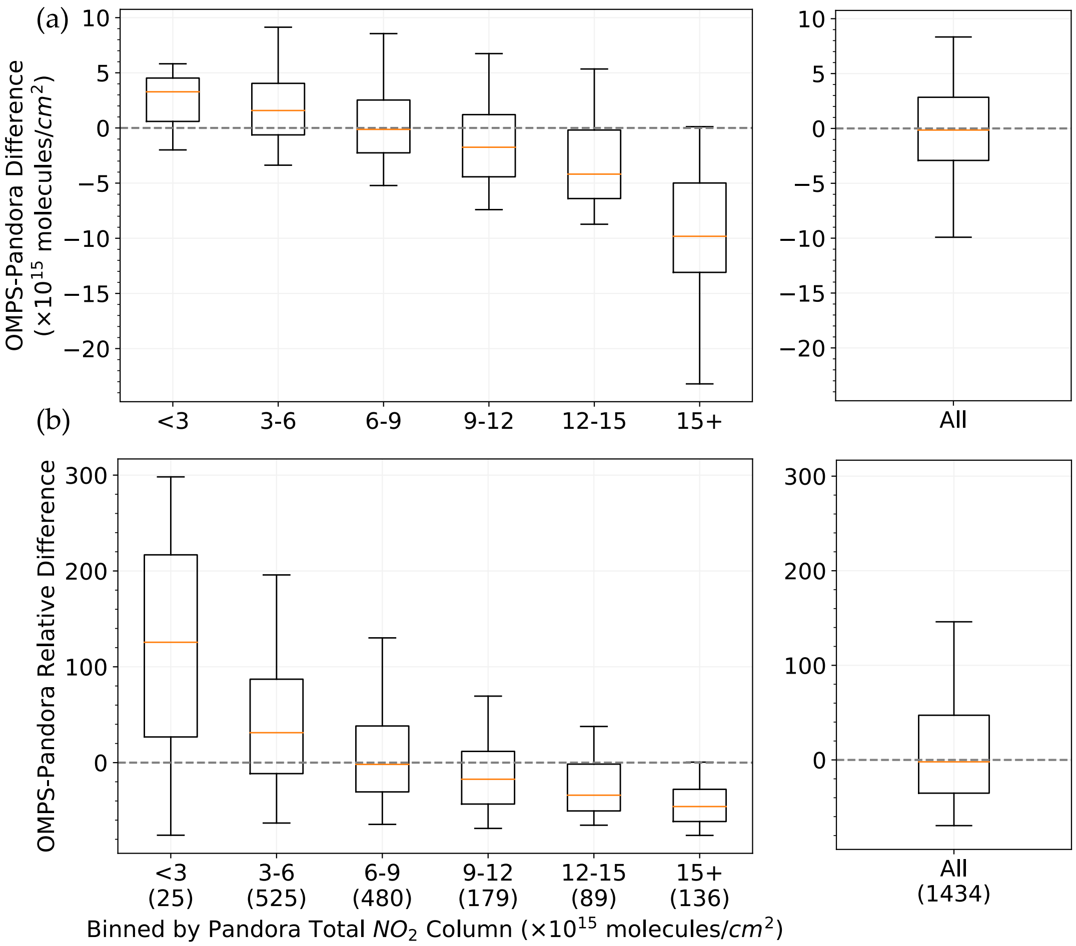


Figure 6.5: Box-whisker plots (95-75-50-25-5 percentiles) showing the **(a)** absolute difference and **(b)** relative difference between NOAA-20 OMPS and Pandora measured total NO_2 columns, binned by Pandora columns at the labeled thresholds (left), as well as all data points (right). The number of points in each bin and all data are indicated by the numbers in parentheses. The data used in the analysis are collected from the 13 U.S. Pandora stations as shown in Figure 6.3, over a period from 2019-02-14 to 2020-04-30.

relationship progressively worsens with increasing satellite pixel size, simply resulting from the flattening of higher NO_2 enhancement over larger spatial areas (Judd et al., 2019). Such behavior is more often associated with localized heterogeneous features rather than more well-mixed regional-scale enhancements. In addition, because of the relatively coarse resolution of the OMPS a priori profiles, OMPS tropospheric columns are expected to have a low bias over polluted areas where the actual peak in the NO_2 profiles is close to the surface, and the boundary layer column is underestimated in the a priori. Similarly, the less polluted columns could be overestimated due to a slight overestimate of boundary layer NO_2, resulting from the averaging effect of low-resolution a priori profiles in situations of large spatial heterogeneity. Replacing the coarse (1° × 1.25°) a priori NO_2 profiles with high-resolution profiles from chemical transport models can potentially improve the agreement between NOAA-20 OMPS and Pandora.

6.4 Case study: tropospheric NO_2 column reductions observed during COVID-19

In this section, we demonstrate the high sensitivity of NOAA-20 OMPS NO_2 observations with COVID-19 application and quantify the impact of COVID-19 outbreak on global NO_2 pollution. During the early half of 2020, many countries around the world enforced physical distancing measures in response to the outbreak of the COVID-19 crisis. China's policy interventions are among the most stringent. Figure 6.6 shows a visual comparison of OMPS observed tropospheric NO_2 columns over China before and after the lockdown in 2020 (a-e) and over the same period in 2021 (f-j), with indications of the Chinese New Year holiday (by red lantern, top left) and of the lockdown period (by padlock, bottom right). In 2021, OMPS observed large winter NO_2

abundances (Figure 6.6f and g) followed by a drop during the Chinese New Year holiday (CNY hereafter, Figure 6.6h). The NO_2 TVCDs decline during CNY is a typical phenomenon observed every year because most Chinese factories shut down for the holiday and the traffic volumes decrease, resulting in a decrease in fuel consumption and thus NOx emissions. A rebound of NO_2 TVCDs is usually observed right after CNY, marking the end of the 7-day CNY holiday and people get back to work (Figure 6.6i). Note that the NO_2 rebound after CNY is much lower than its January peak, due to seasonality caused by the shorter NO_2 lifetime in the warmer season. In 2020, since the initial phase lockdown is coincident with the CNY holiday, NOx emissions curtail significantly and NOAA-20 OMPS observations indicate a steep drop of NO_2 TVCDs, reaching a factor of 2 or more at most Chinese cities (Figure 6.6b). The average NO_2 reduction in 2020 over China is 35% from "before" (Figure 6.6a) to "after" (Figure 6.6b), while a reduction of 15% in 2021 is observed. This suggests that the observed reduction in 2020 far exceeds the typical holiday-related reduction. In addition, unlike the typical years that we see a clear NO_2 reduction during and a quick increase after CNY, NO_2 columns do not bounce back after the week of 2020 CNY holiday (Figure 6.6c). In fact, it remains low for several weeks during strict COVID-19 quarantine (Jan 31 – Feb 17, 2020), after which NO_2 columns gradually recover, reflecting the return of economic activities and NO_x emissions (Figure 6.6d and e).

A quantitative analysis of the impact of the COVID-19 measures on NO_2 in China as well as in other countries is given in Table 6.2. Note that the relatively large and not fully understood contribution of background NO_2 columns has a large impact on trend analyses as more background signal is incorporated into the analysis, whether by incorporating a large spatial area or by computing the analysis over less polluted cities (Qu et al., 2021; Silvern et al., 2019) (Qu et al., 2021; Silvern et al., 2019). We compare the observed NO_2 TVCDs during the

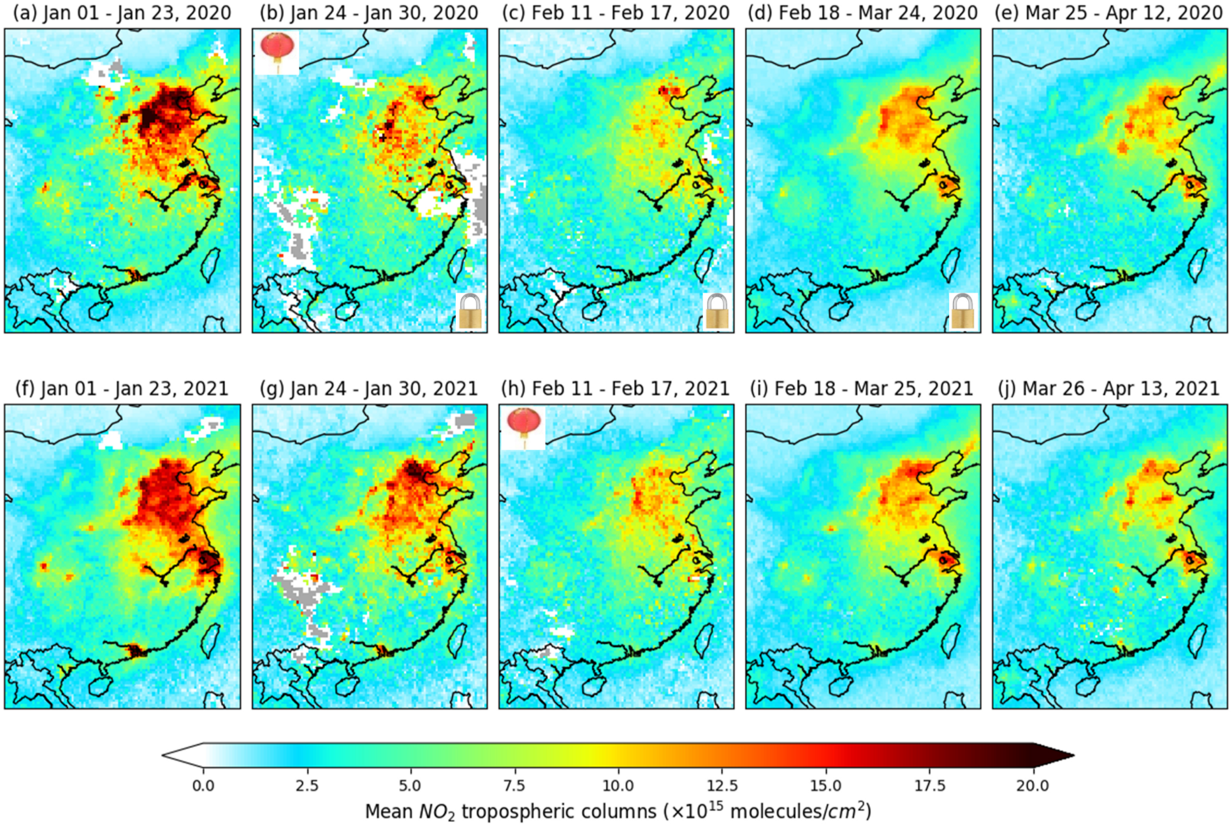
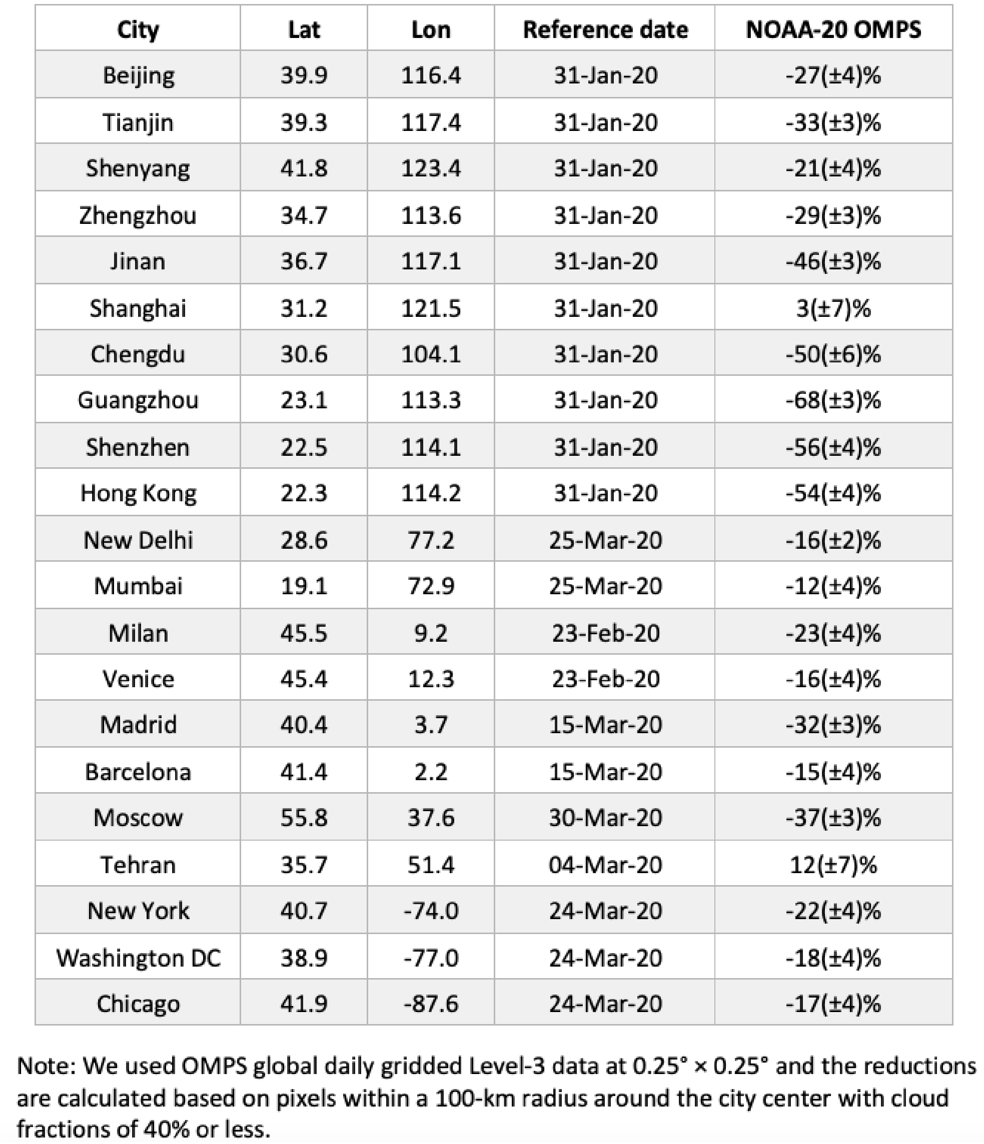


Figure 6.6: Mean tropospheric NO_2 columns over China as observed by NOAA-20 OMPS **(a)** before and **(b-e)** after the COVID-19 lockdowns. For the comparison, the same time periods are shown for 2021 **(f-j)**. The Chinese New Year holiday covers the weeks of Jan 24-30 in 2020 and Feb 11-17 in 2021, which are indicated by the red lanterns in panel **(b)** and **(h)**. The lockdown measures are initiated during and extended after the 2020 Chinese New Year holiday, shown by the padlock sign in panel **(b-d)**, and partial loosening of the restrictions starting Mar 25, 2020, shown in panel **(e)**. Grey areas on the maps indicate no valid data due to the 30% cloud fraction filter.

lockdown in 2020 versus a recovering year NO_2 in 2021. This year-over-year comparison calculates NO_2 column averages starting on the same reference date and last for 21 days, to exclude seasonality-caused NO_2 changes. For the Chinese cities in Table 6.2, we averaged NO_2 TVCDs between January 31 and February 10, 2020 (11 days) compared to the same period in 2021, in order to eliminate the interference of CNY holidays. Similarly, the lockdown period for Iran was chosen between March 4 and March 19 (16 days) to eliminate the interference of the Nowruz holiday. Substantial NO_2 column reductions in 2020 (relative to 2021) are evident in many cities around the world where strict COVID-19 precautions were enforced. The observed column decreases are largely due to the decline of traffic emissions, by far the dominant NO_x emission source in cities, as well as decreases in industrial activities and power generation (Myllyvirta, 2020; Schuman, 2020; Zara, 2020). Simulations of chemistry transport models are needed if to isolate the benefit of emission reduction from variations of transport (Valin et al., 2013) or NO_x lifetime (Laughner and Cohen, 2019). Note that since we are comparing 2020 NO_2 columns to 2021, part of the lockdown related NO_2 reduction might be canceled out by the lower emission rate in 2021 due to the emission declines benefited from environmental regulations with each advancing year (Wu et al., 2019). Therefore, the actual NO_2 decreases could be larger if we were to compare with 2019 NO_2, as shown in the TROPOMI study of Bauwens et al. (2020) Table 1.

Table 6.2: NO_2 tropospheric vertical column densities (TVCDs) reduction observed during the COVID-19 lockdown period, starting on the Reference date and lasting for 21 days, relative to the same period in 2021, with the exception of China and Iran, where it lasts for 11 and 16 days respectively, in order to avoid the interference with the New Year Holidays. The percentage change is defined as (TVCDs2020 - TVCDs$_{2021}$)/TVCDs$_{2021}$ $\times$ 100%. The numbers in the brackets are the standard error of the mean.



City	Lat	Lon	Reference date	NOAA-20 OMPS
Beijing	39.9	116.4	31-Jan-20	-27(±4)%
Tianjin	39.3	117.4	31-Jan-20	-33(±3)%
Shenyang	41.8	123.4	31-Jan-20	-21(±4)%
Zhengzhou	34.7	113.6	31-Jan-20	-29(±3)%
Jinan	36.7	117.1	31-Jan-20	-46(±3)%
Shanghai	31.2	121.5	31-Jan-20	3(±7)%
Chengdu	30.6	104.1	31-Jan-20	-50(±6)%
Guangzhou	23.1	113.3	31-Jan-20	-68(±3)%
Shenzhen	22.5	114.1	31-Jan-20	-56(±4)%
Hong Kong	22.3	114.2	31-Jan-20	-54(±4)%
New Delhi	28.6	77.2	25-Mar-20	-16(±2)%
Mumbai	19.1	72.9	25-Mar-20	-12(±4)%
Milan	45.5	9.2	23-Feb-20	-23(±4)%
Venice	45.4	12.3	23-Feb-20	-16(±4)%
Madrid	40.4	3.7	15-Mar-20	-32(±3)%
Barcelona	41.4	2.2	15-Mar-20	-15(±4)%
Moscow	55.8	37.6	30-Mar-20	-37(±3)%
Tehran	35.7	51.4	04-Mar-20	12(±7)%
New York	40.7	-74.0	24-Mar-20	-22(±4)%
Washington DC	38.9	-77.0	24-Mar-20	-18(±4)%
Chicago	41.9	-87.6	24-Mar-20	-17(±4)%

Note: We used OMPS global daily gridded Level-3 data at 0.25° × 0.25° and the reductions are calculated based on pixels within a 100-km radius around the city center with cloud fractions of 40% or less.

Chapter 7: Validation of tropospheric NO_2 from SNPP OMPS: data, model, and method

7.1 SNPP OMPS NO_2 observations

The Ozone Mapping and Profiler Suite - Nadir Mapper (OMPS-NM) aboard Suomi-NPP (SNPP) spacecraft was launched in October 2011, under NASA/NOAA Joint Polar Satellite System (JPSS) mission. Like other legacy LEO instruments including the Ozone Monitoring Instrument (OMI, 2004 to present) on Aura (Levelt et al., 2006), GOME-2 on METOP-A (2006 to present) & METOP-B (2012 to present) (Callies et al., 2000; Munro et al., 2016), and the TROPOspheric Monitoring Instrument (TROPOMI) on S5P (Veefkind et al., 2012), OMPS-NM is a nadir-viewing hyperspectral UV spectrometer flying in a sun-synchronized orbit with an ascending equator overpass local time of 1:30 P.M., same as NOAA-20 OMPS and close in time to OMI & TROPOMI. OMPS measures backscattered UV radiance in the 300 – 380 nm range with a high signal-to-noise ratio of approximately 2500. The OMPS signal-to-noise ratio is about twice as large as that for OMI, compensating for the lower NO_2 measurement sensitivity in OMPS UV spectral range (Yang et al., 2014). OMPS NM scans Earth's atmosphere in about 14 orbits a day, with a 2800 km cross-track swath (110° field of view) divided into 35 instantaneous Field Of Views (iFOVs) or pixels, leading to a spatial resolution of 50×50 km^2 at nadir. The ground pixel of SNPP OMPS is smaller than GOME-2 (80 × 40 km^2) but larger than OMI (13 × 24 km^2) and TROPOMI (5.5 × 3.5 km^2).

The NO_2 column retrievals from SNPP OMPS UV spectra are achieved through the DVCF algorithm and a novel orbit-based Stratosphere-Troposphere-Separation (STS) approach developed by Yang et al. (2014) and applied in Huang et al. (2022). For UV spectra, since the Rayleigh scattering is stronger than in the visible bands, fewer photons reach the lower atmosphere and consequently, there is reduced NO_2 measurement sensitivity. Due to the combined effect of the smaller differential structures of NO_2 absorption cross sections and the stronger Rayleigh scattering in UV, it is more challenging to retrieve NO_2 from the UV spectra than the violet-blue (410-460 nm) wavelength range that is typically used in NO_2 retrievals from other polar-orbiting satellite platforms. The DVCF algorithm is uniquely designed to overcome the challenge, enable a better spectral fit to the measured spectra, and achieve a more accurate retrieval of NO_2 vertical columns than the traditional DOAS retrieval algorithm. Details about the DVCF algorithm have been explained in Section 1.3.2 and Section 5.2.

7.2 Ground-based MAX-DOAS NO_2 observations

SNPP OMPS NO_2 retrievals are validated against NO_2 tropospheric columns measured by the ground-based MAX-DOAS (Multi-AXis Differential Optical Absorption Spectroscopy) instrument. MAX-DOAS instrument is used to derive vertical profiles of trace gases and aerosols in the troposphere from the measurements of scattered sunlight at multiple elevation angles (Vlemmix et al., 2010, 2011; Wagner et al., 2004). It provides high-temporal resolution long-term observations that are often used for evaluating satellite retrievals (Kanaya et al., 2014; Wang et al., 2017a). We compare SNPP OMPS tropospheric NO_2 vertical columns to those retrieved from the ground-based MAX-DOAS instruments deployed in two cities of China: Xingtai and Wuxi (Figure 7.1, Table 7.1). The Xingtai station is located near a large industrial area with several cement and steel industries and coal coking facilities on North China Plain. With the Taihang Mountain located about 30 km to the west of Xingtai station, local NO_x pollution does not get transported far from the site and get accumulated along the mountain range (Wang et al., 2019). The other MAX-DOAS station at Wuxi is located in the Yangtze River delta economic zone, close to the urban center of Wuxi city with a population of 6 million (Wang et al., 2017b). Both ground stations are located in one of the most populous and industrialized regions in China, so they provide a broad range of NO_2 data for OMPS evaluation. In this study, we utilized both the MAX-DOAS measured NO_2 columns and aerosol extinction profiles to validate and improve OMPS NO_2 products. More details about these two MAX-DOAS stations are summarized in Table 7.1.

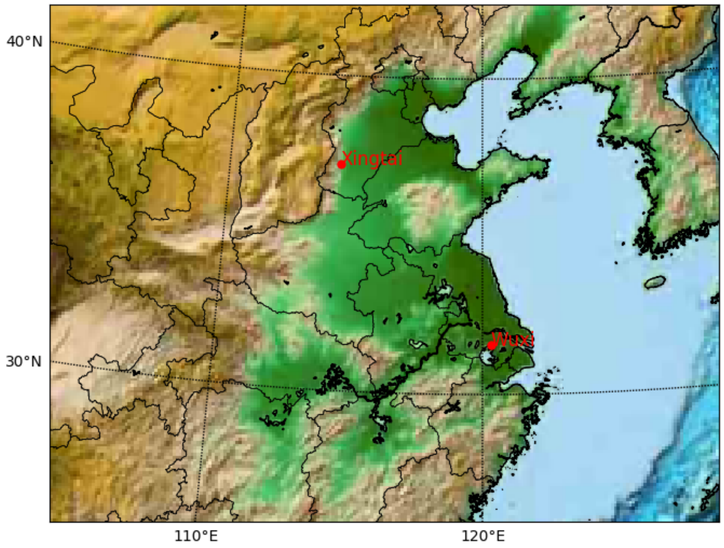


Figure 7.1: Locations of two ground-based MAX-DOAS stations in China for SNPP OMPS NO_2 validation (marked in red): Xingtai (37.18° N, 114.37° E) and Wuxi (31.57° N, 120.31° E), colors in the background indicate terrain height.

To retrieve NO_2 and aerosol extinction (AE) profiles from the MAX-DOAS instruments,

Table 7.1: An overall summary of the data and model used to evaluate SNPP OMPS NO_2 vertical columns.

	MAX-DOAS Site I	**MAX-DOAS Site II**
Site name	Xingtai (China)	Wuxi (China)
Location	(37.18° N, 114.37° E)	(31.57° N, 120.31° E)
Time operated	May 8 – Jun 10, 2016	May 2011 – Nov 2014
If a field campaign is deployed during MAX-DOAS operation	Yes ARIAs (Benish et al., 2020)	No
MAX-DOAS elevation angles	1, 2, 3, 4, 6, 8, 10, 15, 20, 30, 90° (11 elevation angles)	5, 10, 20, 30, 90° (5 elevation angles)
MAX-DOAS NO_2 and aerosol retrieval spectra range	351 – 390 nm	
MAX-DOAS retrieved profile extent (vertical grid resolution)	4 (0.2) km	
References	Wang et al. (2019)	Wang et al. (2017b)
OMPS observation time span	Starting Feb 01, 2012, to present	
OMPS & MAX-DOAS overlap time period	May 8 – Jun 10, 2016	Feb 1, 2012 – Nov 21, 2014
Number of coincidences between MAX-DOAS and OMPS (standard product)	36 (all clouds) 23 (< 30% cloud fraction)	1012 (all clouds) 647 (< 30% cloud fraction)
Chemistry transport model	CMAQ v5.2	CMAQ v5.0.2
Model resolution	12 km	
Model simulation time span	May and Jun of 2016	Jan, Mar, Jul, Oct of 2014
OMPS, MAX-DOAS & Model overlap time period	May 8 – Jun 10, 2016 (34 days)	Jan, Mar, Jul, Oct of 2014 (4 months)
Number of coincidences between MAX-DOAS, CMAQ and OMPS (model improved)	36 (all clouds) 23 (< 30% cloud fraction)	107 (all clouds) 79 (< 30% cloud fraction)

first, the differential slant column densities (dSCDs) of the oxygen dimer (O_4) and NO_2 are retrieved from the UV spectra of scattered sunlight measured by MAX-DOAS using the DOAS technique. Then, the tropospheric vertical profiles of AE and NO_2 volume mixing ratios are retrieved from the elevation-dependent dSCDs using the PriAM profile inversion algorithm (Wang et al., 2017b). The PriAM inversion application has a two-step process: 1) vertical profiles of AE are retrieved from the measured O_4 dSCDs, and 2) the retrieved AE profiles are used for the inversion calculations of all trace gas profiles including NO_2. The profile inversion iterations require radiative transfer simulations and a priori profile assumptions. Since clouds can strongly impact MAX-DOAS retrievals, different cloud conditions are identified from the MAX-DOAS observations of the color index using a cloud classification scheme introduced by Wagner et al. (2014, 2016) and verified by Wang et al. (2015). One important message from the studies is that a robust trace gas profile/column can be retrieved not only in clear sky conditions (including both high and low aerosol loadings), but also in most cloudy situations as long as they are neither fog nor optically thick clouds. On the other hand, for aerosol retrievals, only near-surface AE is in good quality under cloudy conditions. The complete recommendations of whether to use the MAX-DOAS retrieved trace gases and aerosol quantities under different cloud conditions can be found in Table 3 of Wang et al. (2017b). In this study, we use assured high-quality MAX-DOAS NO_2 data and exclude the cloud-contaminated retrievals of aerosol extinction profiles & AODs in the following analysis.

In order to achieve a robust retrieval, quality control is programmed into MAX-DOAS retrieval procedure to filter the unreliable dSCDs and screen suspicious profile inversion results of large differences between measured and modeled dSCDs, based on the criteria explained in Wang et al. (2017b, 2019). In addition, the PriAM retrieval algorithm has explicitly corrected the seasonal variations of ambient temperature and pressure in the forward model, while the general practices use fixed temperature and pressure profiles (US summer standard). This adjustment improves MAX-DOAS retrieval accuracy and is especially important for retrieving long-term observations like Wuxi. In previous studies (Wang et al., 2017b, 2019), the MAX-DOAS measurements of trace gases and aerosols have been cross verified with multiple collocated independent instruments, including an AERONET sun photometer, a visibility meter, and a long path DOAS for the Wuxi station, with the additions of a Ramen lidar and ARIAs aircraft spirals for the Xingtai station. The comparisons show good agreement between MAX-DOAS and coincident measurements from other independent instruments, which justifies the data quality that our evaluation analysis relies on. Note that the MAX-DOAS retrieved profiles go up to 4 km with a grid resolution of 200m. Also, all the reported vertical column densities are derived by integrating the retrieved profiles which yields more accurate results than the geometric approximation.

7.3 CMAQ model

We used the Community Multiscale Air Quality (CMAQ) model to simulate high-resolution NO_2 spatial distributions over China. The model outputs were employed to calculate NO_2 vertical profile shapes and derive NO_2 horizontal averaging kernels to further advance OMPS retrieval accuracy. Since the reference data were measured during two distinct time periods, two CMAQ simulations were conducted separately to cover each MAX-DOAS observation record spatially and temporally. For Xingtai (2016) period, two nested model domains with grid resolutions of 36- and 12-km were simulated between May and June 2016, and Xingtai station was covered under the 12-km domain (Figure 7.2a). For Wuxi (2014) period, CMAQ simulations were

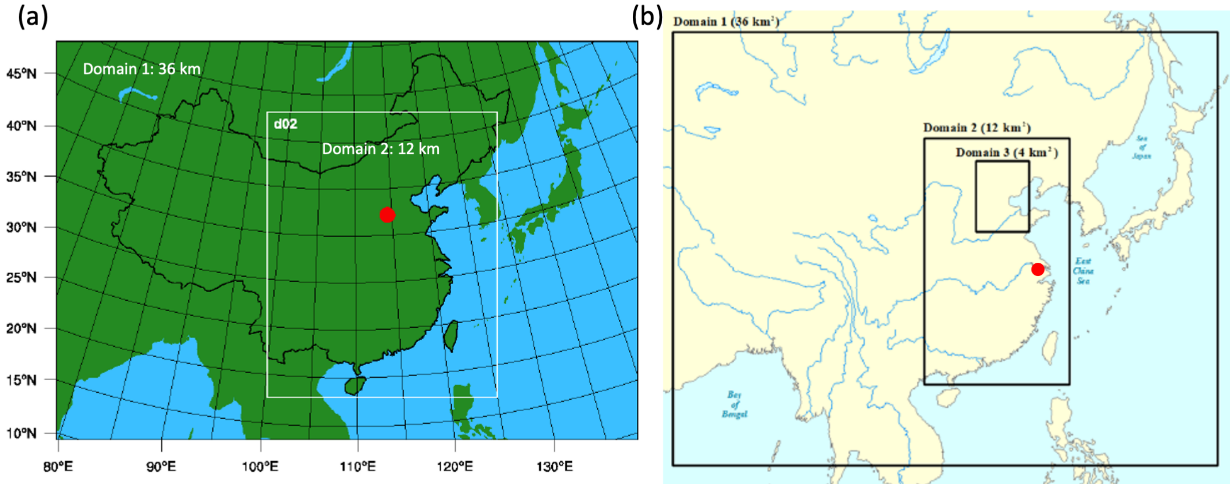


Figure 7.2: CMAQ model domains for **(a)** Xingtai 2016-episode and **(b)** Wuxi 2014-episode, with indications of domain grid spacings on the plots. The Red dots indicate the locations of Xingtai station in **(a)** and Wuxi station in **(b)**.

conducted in 3 nested domains with spatial resolutions of 36-, 12- and 4-km, note that the inner 4-km domain focused on a small region near Beijing, and our validation site Wuxi was covered under the 12-km domain (Figure 7.2b). Therefore, the 12-km CMAQ model outputs were used to validate both sites. Note that although our ground site at Wuxi has an extended 2.5-years observation, the CMAQ simulations were only conducted in January, March, July, and October of 2014 (4 months), representing the average conditions in winter, spring, summer and fall respectively (Zhao et al., 2017). Since OMPS algorithm improvements (detailed later) rely on model simulations, the improved NO_2 columns were only applied to these 4 months overlap between satellite-, ground- and model-based data at Wuxi (Table 7.1), indicative of a year-round situation.

Generally, simulations of regional air quality are accomplished in a comprehensive three-step process via meteorology-emission-chemistry modeling framework. For both CMAQ simulations, the Weather Research and Forecasting (WRF) model was used to generate hourly meteorological fields to drive emission and air quality modeling. A summary of main model configurations

Table 7.2: CMAQ model configurations used to evaluate OMPS NO_2 in this analysis. Two model simulations were conducted over East China during the MAX-DOAS operated time periods: Xingtai 2016 period and Wuxi 2014 period.

METEOROLOGY-EMISSION-CHEMISTRY MODEL FRAMEWORK		
Simulation period	May – Jun 2016 (Xingtai)	Jan, Mar, Jul, Oct 2014 (Wuxi)
WRF version	WRF v3.8.1	WRF v3.7.1
Meteorology initial/ boundary condition	ECMWF (ds627.0)	NCEP 1°x 1° (6h)
WRF main configurations	KF cumulus	KF cumulus
	WSM6 microphysics	Morrison microphysics
	RRTM SW & LW radiation	RRTMG SW & LW radiation
	Pleim-Xiu land surface	Pleim-Xiu land surface
	ACM2 PBL	ACM2 PBL
Emission	EDGAR v4.2 0.1°x 0.1° of year 2010	High-resolution emission inventory with unit-based industrial sources
		MIX emission inventory 2010 (outside China)
CMAQ version	CMAQ v5.2	CMAQ v5.0.2
CMAQ spatial resolution	36-12 km (12 km was used)	36-12-4 km (12 km was used)
CMAQ vertical layers	35 total layers ~20 layers in the lowest 2 km	23 total layers ~12 layers in the lowest 2 km
CMAQ spin-up period	14 days	5 days
CMAQ main configurations	CB06 AERO6	CB05 AERO6
	Fixed stack height and plume rise ~200 m	Unit-based stack height and Inline plume rise
	BEIS biogenic emissions	MEGEN biogenic emissions
References	He et al. (2019)	Zhao et al. (2017, 2018) Zheng et al. (2019) Chang et al. (2019)

with references for the two simulation episodes is provided in Table 7.2. The anthropogenic emission inputs for CMAQ Xingtai (2016) simulations were from the Emissions Database for Global Atmospheric Research (EDGAR) Version 4.2 of year 2010 (He et al., 2019). The emission fields were then processed by the Sparse Matrix Operator Kernel Emissions (SMOKE) model and speciated onto CMAQ grids. The CMAQ Wuxi (2014) simulations employed a high-resolution anthropogenic emission inventory with unit-based point sources developed by Tsinghua University using an emission factor method (Cai et al., 2017; Fu et al., 2013; S. X. Wang et al., 2014; Zhao et al., 2013a,b, 2018). The unit-based inventory results in improvements in air quality modeling compared with the proxy-based emission inventory, where all sectors are allocated as area sources using spatial proxies such as population, GDP, road map, and land use data (Chang et al., 2019; Zheng et al., 2019). The emission inventory outside China for Wuxi (2014) episode was obtained from the most recent MIX emission inventory of the year 2010 (Li et al., 2017b). With the meteorology and emission inputs, the spatial and temporal variations of NO_2 were subsequently simulated by CMAQ with the major configurations listed in Table 7.2. The CMAQ simulated NO_2 profiles have 20 layers (Xingtai) and 12 layers (Wuxi) below 2 km, allowing us to obtain fine NO_2 vertical profiles to improve satellite a priori assumptions.

7.4 ARIAs field campaign

During the time of MAX-DOAS operation at Xingtai (May to June 2016), the Air Chemistry Research in Asia (ARIAs) field campaign was deployed over the same region, providing vertically resolved observations of atmospheric trace gases and aerosols (Dickerson et al., 2007). The airborne measurements of NO_2 vertical profiles were used to validate CMAQ simulated profiles

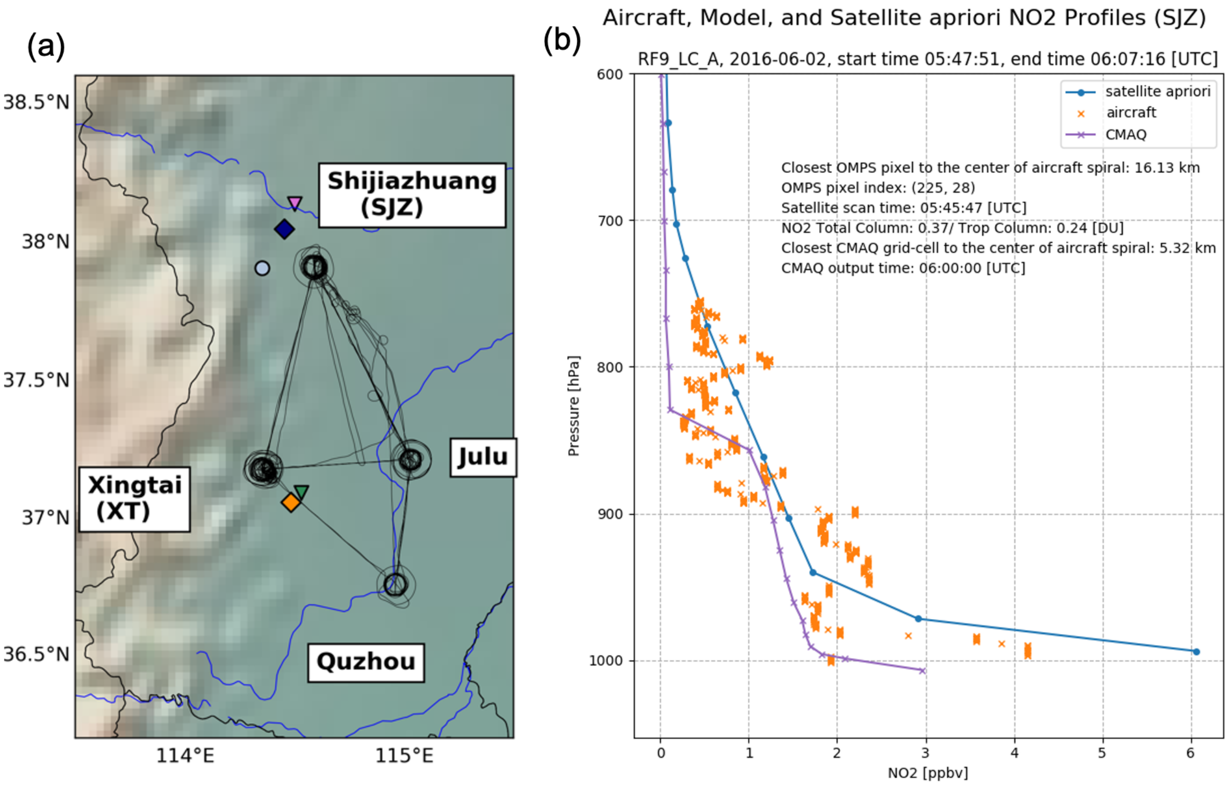


Figure 7.3: **(a)** Map of 11 ARIAs flight tracks, spiraling over four cities (Shijiazhuang, Xingtai, Quzhou, Julu) on North China Plain. Image credit: Benish et al. (2020). **(b)** NO_2 vertical profile measured by the ARIAs aircraft spiral over Shijiazhuang (SJZ) from Research Flight 9 Ascending mode, on June 2, 2016 around 6 UTC (2 PM Beijing Time) (denoted in orange cross), compared with CMAQ simulated profile (in purple) and OMPS a priori NO_2 profile (in blue) at nearest coincidence.

that were later used to improve OMPS a priori NO_2 assumptions. The campaign included 11 research flights measuring the vertical distributions of multiple trace gases, including O_3, SO_2, NO_2, NO/NO_y, CO, VOC etc., as well as aerosol optical properties (scattering and absorption) in the lower troposphere (Benish et al., 2020; F. Wang et al., 2018). The aircraft flew vertical spirals from 300 to 3500 m over four locations: Shijiazhuang - Luancheng Airport (37.91°N, 114.59°E; 58m a.s.l.), Julu (37.22°N, 115.02°E; 30m a.s.l.), Quzhou (36.76°N, 114.96°E; 40m a.s.l.), and Xingtai (37.18°N, 114.36°E; 182m a.s.l.) on North China Plain. See Figure 7.3a for the flight tracks map.

Measurements taken from the aircraft spirals over Xingtai site have previously been used to validate the MAX-DOAS NO_2 column retrievals, and good agreement between them has been demonstrated in Wang et al. (2019). In this study, we used the airborne samplings of NO_2 vertical profiles from 11 research flights to evaluate CMAQ model simulated NO_2 profiles over the region (Figure 7.3b). In general, CMAQ simulated profile shapes (normalized profile) agree better with the observed NO_2 vertical distribution than OMPS a priori profiles from the global GEOS-Chem model. The comparison confirms that CMAQ profiles are more accurately representative of time-varying atmosphere than the climatological a priori profiles used in OMPS NO_2 standard product.

7.5 Comparison methodology

We evaluate the agreement between SNPP OMPS and MAX-DOAS NO_2 vertical column densities by calculating the mean difference (MD), the mean relative difference (MRD), the dispersion (i.e., the standard deviation) of the differences (σ), the correlation coefficient (r) the slope of the ordinary least square (OLS) linear regression fit for the measurements as well as those of reduced major axis (RMA) linear regression fit. The MD is defined as the average difference between the OMPS and MAX-DOAS VCDs in Eq. 7.1, whereas the MRD is the average of these differences when normalized to MAX-DOAS's VCDs (reference data), as shown in Eq. 7.2.

$$\text{MD} \quad = \frac{1}{n}\sum_{i=1}^{n}(VCD_{OMPS_i} - VCD_{MAX-DOAS_i}) \tag{7.1}$$

$$\text{MRD (\%)} = \frac{1}{n}\sum_{i=1}^{n}\frac{(VCD_{OMPS_i} - VCD_{MAX-DOAS_i})}{VCD_{MAX-DOAS_i}} \times 100\% \tag{7.2}$$

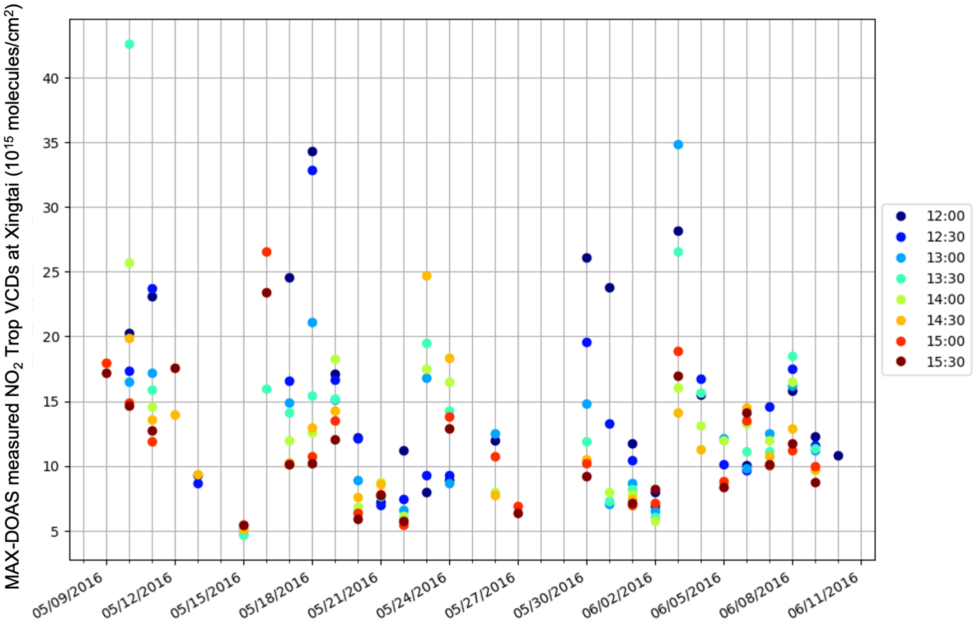


Figure 7.4: Time series of MAX-DOAS measured tropospheric NO_2 vertical columns (lower 4 km vertical columns) at Xingtai station during ARIAs campaign. Each vertical line corresponds to a day and the colors represent local times from 12:00 (noon) to 15:30 in half an hour measurement time intervals. OMPS can possibly overpass the ground station at any of these given early afternoon times once daily or twice daily. If twice daily, the first overpass of the day would correspond to the left side of the swath (small cross-track index) and the second overpass to the right side (large cross-track index).

A positive (negative) MD or MRD is thus indicative of OMPS overestimation (underestimation). The RMA linear regression is used since it is proven to solve the linear fitting problem when both x and y variables have measurement errors (Judd et al., 2020).

The MAX-DOAS measurements have a high temporal resolution, whereas OMPS provides global once-daily or at most twice-daily coverage (edge pixels) at any given location. Figure 7.4 illustrates the temporal variability of MAX-DOAS measured column NO_2 amounts at Xingtai station during the early afternoon hours when a satellite overpass could possibly happen. It shows MAX-DOAS measured NO_2 columns can vary significantly within a short amount of time, due to the effects of advection or local emission changes. To reconcile the spatiotemporal resolutions of the two datasets, MAX-DOAS data within ± 60 min of OMPS overpass were averaged to evaluate

OMPS observed columns. Other coincidence criteria including the nearest coincidence alongside different average time windows were also investigated and results are presented in Section 8.5. Spatially, OMPS pixels that encompass the Wuxi/Xingtai MAX-DOAS station with the smallest pixel-center-to-station distance were considered for the comparison. Both OMPS and MAX-DOAS data were filtered separately based on a set of quality assurance criteria, after which the remaining temporally collocated measurements were compared with each other. MAX-DOAS data are filtered according to cloud classification schemes and only assured high-quality NO_2 data were included in the comparison. For OMPS, we filtered pixels with radiative cloud fraction $> 30\%$ for general-purpose evaluations, and also investigated different radiative cloud fraction scenes and results are summarized in Section 8.5. Since the MAX-DOAS retrievals measured NO_2 columns extend up to 4 km, to ensure a fair comparison, OMPS NO_2 tropospheric VCDs (TVCDs) are also integrated in the same range based on a priori profile shapes when OMPS is compared with MAX-DOAS. Apart from that, TVCDs in this thesis generally refer to the standard tropospheric vertical columns.

7.6 Evaluation of tropospheric NO_2 standard product

Figure 7.5 compares SNPP NO_2 standard product with nearest coincidence measurements from MAX-DOAS at Wuxi station from Feb 1, 2012, to Nov 21, 2014. MAX-DOAS observed tropospheric NO_2 vertical columns in the range from surface to 4 km altitude are used to evaluate SNPP NO_2 standard product measured in the same range. SNPP OMPS biased low against MAX-DOAS, with mean difference of -9.18×10^{15} molecules/cm^2 and mean relative difference of -26.53%. It is moderately correlated with MAX-DOAS columns (r = 0.59, N = 346) at the eastern

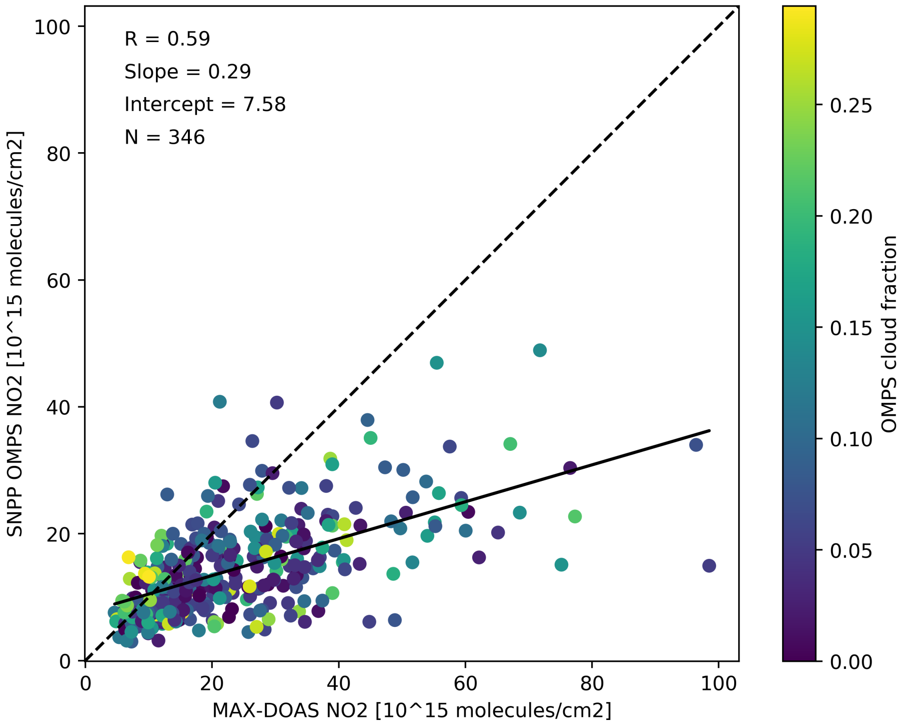


Figure 7.5: Scatter plot of SNPP OMPS NO_2 standard product vs. MAX-DOAS observed NO_2 tropospheric vertical columns (surface to 4 km altitude) over Wuxi station from Feb 2012 to Nov 2014, colored by OMPS cloud fraction. MAX-DOAS reference columns on this plot are clear-sky measurements only and are matched as temporally closest to the OMPS pixels with cloud fraction less than 30%.

China city station (Wuxi), as indicated by the OLS linear regression statistics. Although the campaign length of MAX-DOAS observations at Xingtai station is too short for any comprehensive evaluation (38 pairs of coincidence), it shows similar statistical significance between SNPP OMPS SP and MAX-DOAS measured NO_2 TVCDs (r = 0.53, $\text{slope}_{OLS} = 0.34$, N = 38, nearest coincidence, all clouds). This suggests the SNPP NO_2 SP is moderately correlated with independent reference data in China with a correlation between 0.5 and 0.6.

Like what we have found in Section 6.3, satellite NO_2 columns tend to underestimate the ground-based reference observations for relatively large columns, corresponding to polluted urban regions and or episodes of relatively elevated pollution. While NOAA-20 OMPS NO_2

total columns overestimate at some remote Pandora locations in the US, the SNPP NO_2 column retrievals mostly show underestimation compared to the MAX-DOAS coincidence measurements in China. This is probably because these two MAX-DOAS sites are located in polluted urban cities and measured much higher column densities than the Pandora US stations. We also color the coincidence comparison by OMPS cloud fraction and we found no clear relationship between smaller cloud fraction and better-quality column retrieval in the point-to-point coincidental basis.

Part of the OMPS underestimation is explained by the nature of spatial representativity mismatch between satellite and ground-based measurements when an area-averaged column density over a relatively large satellite pixel is compared with a ground-based instrument that has limited FOV. Other possible reasons as well as solutions for the satellite underestimation of high NO_2 columns are exploited in Chapter 8.

Chapter 8: Algorithm improvements for tropospheric NO_2 retrieval from SNPP OMPS

8.1 Algorithm improvement I: NO_2 a priori profile adjustment

The measurement sensitivity to tropospehreic NO_2 is proportional to the altitude-dependent photon path length, which usually decreases rapidly towards the surface. As a result, the observed tropospheric NO_2 column depends significantly on the prescribed NO_2 profile $\mathbf{X}_a$. Algebraically, the dependence of retrieved NO_2 column (V) on the $(\mathbf{X}_a)$ can be expressed as

$$V = V_a + \mathbf{AK} \cdot (\mathbf{X}_t - \mathbf{X}_a), \tag{8.1}$$

where V and V_a are respectively the retrieved vertical column and the *a priori* vertical NO_2 column, $\mathbf{AK}$ the averaging kernel of the DVCF algorithm, and $\mathbf{X}_t$ the actual NO_2 profile. For convenience, a profile shape factor $\mathbf{S}$ is defined as $\mathbf{X} = V\mathbf{S}$, which specifies a shape factor as a normalized profile with its vertical integrated sum equal to 1 (i.e., $\sum \mathbf{S} = 1$). For instance, *a priori* profile $\mathbf{X}_a = V_a\mathbf{S}_a$, is simply the product of *a priori* column V_a and the *a priori* shape factor $\mathbf{S}_a$.

Eq. 8.1 shows that the DVCF column (V) retrieval depends on the *a priori* assumption $(\mathbf{X}_a)$. The forward modeling of the DVCF algorithm includes explicitly the sensitivity of the measurement to NO_2 concentrations from the top of atmosphere down to the underlying surface. These sensitivities are weighted with the assumed NO_2 profile to produce the radiance changes measured by the instrument (see Eq. 5.1). The $\mathbf{AK}$ describes this measurement sensitivity profile (e.g., Figure 8.1), and relates retrieved quantity with the actual NO_2 profile.

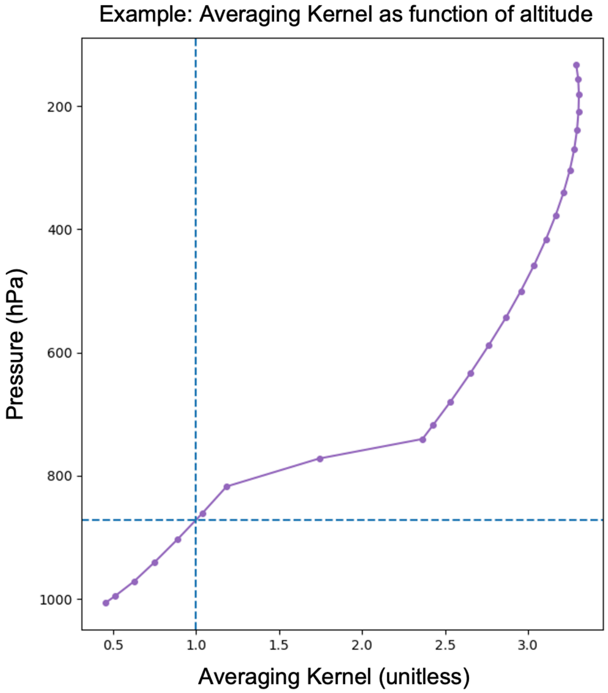


Figure 8.1: An example of how Averaging Kernel (AK) typically varies with regard to altitude. The AK in this example is output at one satellite iFOV over Shijiazhuang, China, on 2016-05-19 05:07 UTC.

Mathematically, $\mathbf{AK}$ is a matrix consisting of partial derivatives of spectral measurements at different wavelengths with respect to (wrt) the state vectors and wrt NO_2 a priori profiles at different altitudes. In DVCF NO_2 retrieval, the elements of state vectors include surface reflectivity and NO_2 vertical column. Conceptually, $\mathbf{AK}$ can be seen as a total sensitivity term. It not only includes the scattering weight which is part of the DOAS method air mass factor (AMF) calculation, but also accounts for the impact of a priori profile change on surface reflectivity. The central difference between the DVCF $\mathbf{AK}$ and DOAS AK is that DVCF $\mathbf{AK}$ correctly represents the measurement kernel of the spectral window used in the retrieval, while DOAS $\mathbf{AK}$ is a simplification estimated at a single wavelength in the middle of this spectral window.

Trace gas shape profiles provided by model simulations are a critical input to satellite retrievals. The operational SNPP OMPS NO_2 products currently use a year-specific (2012) and monthly-averaged GEOS-Chem global model (2° latitude × 2.5° longitude) to provide the a priori NO_2 profiles. The profiles provided by GEOS-Chem model are at a much coarser spatial resolution than the OMPS observations and lack of short-term temporal variabilities. To illustrate the deficiency of the *a priori* based on a monthly climatology, Figure 8.2 shows typical CMAQ simulated NO_2 temporal and spatial variability during a month. Compared to the monthly-averaged GEOS-Chem profiles used in OMPS standard product, the updated finer resolution a priori profiles from 12-km CMAQ model run can resolve the spatial heterogeneities in urban areas and capture the day-to-day variability, providing more accurate representations of NO_2 vertical distributions.

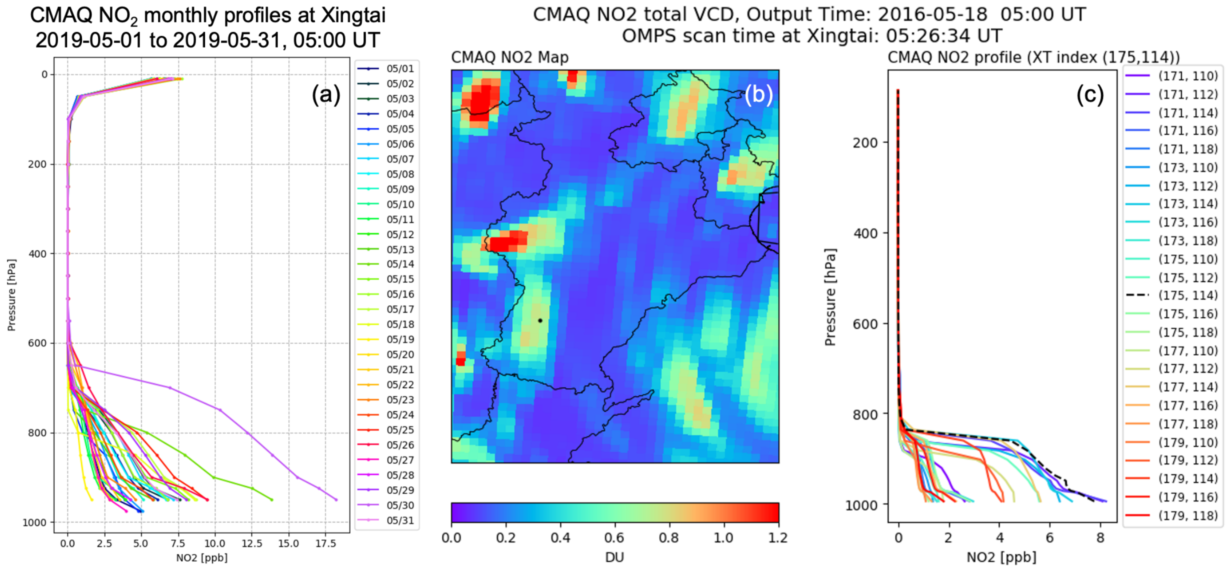


Figure 8.2: CMAQ model simulated NO$_2$ spatial and temporal variability at 12-km resolution over Xingtai (China) station. **(a)** CMAQ daily NO$_2$ vertical profiles over Xingtai at 5 UTC (1 PM local time) during the month of May 2016, which closely resembles OMPS early-afternoon overpassing time. The colors indicate different days during the month. Taking one day (5 UTC 18 May 2016) as an example, panel **(b)** shows the spatial variations of CMAQ simulated NO$_2$ total vertical columns over the Hebei area. The black dot on the map indicates the location of Xingtai Station. **(c)** The corresponding CMAQ NO$_2$ profiles at Xingtai (dash line) as well as the adjacent grid cells (indicated by colors) as shown in (b). We show a 9- by 9-gridboxes region centered at Xingtai, with the first color (purple) in the legend depicting the lower-left corner and the last color (red) depicting the upper-right corner of the region. The first index in the parenthesis is the row index, increasing from south to north, followed by the column index increasing from west to east.

Replacing the *a priori* shape $\mathbf{S}_a$ with a more accurate NO$_2$ shape factor $\mathbf{S}_i$ (from the CMAQ model), the corrected NO$_2$ column (V_i) is obtained by applying the AKs to the difference between the two shape factors

$$V_i = V\left(1 + (\mathbf{AK} - \mathbf{1}) \cdot (\mathbf{S_a} - \mathbf{S_i})\right). \tag{8.2}$$

where V is retrieved NO$_2$ column based on the initial shape factor $\mathbf{X}_a$. Equation 8.2 is derived from the difference between two equations (based on Eq. 8.1): $V = V_a + \mathbf{AK} \cdot (\mathbf{X_t} - \mathbf{X_a})$ and $V_i = V_a + \mathbf{AK} \cdot (\mathbf{X_t} - \mathbf{X_i})$.

Figure 8.3 shows an example of OMPS NO_2 tropospheric columns before and after a priori profile adjustment using high-resolution (12 km) model simulated profiles. The daily CMAQ NO_2 shape profiles from the hourly output are matched temporally and spatially to the OMPS pixel and then applied in Eq. 8.2. The adjustment reveals elevated NO_2 pollution in Tianjin and south Hebei areas, as well as industrial centers in Shandong and Henan provinces. Since the $\mathbf{AK}$ generally increases with altitude (Figure 8.1), a positive adjustment means that the new a priori profile shape factors put more weight in the lower levels. Note that $\mathbf{1}$ is a row vector. $(\mathbf{AK} - \mathbf{1})$ is negative below the pressure altitude of 920 hPa in the example of Figure 8.1, if $(\mathbf{S}_a - \mathbf{S}_i)$ is negative in the same layer, it results in a positive adjustment. $\mathbf{AK}$ typically increases with altitude and reaches 1 in the lower-mid Planetary Boundary Layer (PBL). A positive adjustment means the new profile shape factor $\mathbf{S}_i$ is larger than the initial *a priori* profile $\mathbf{S}_a$ at the AK-below-1 altitudes located in the PBL. Although the satellite retrievals require information from the model as a priori assumptions, the satellite retrieved and adjusted NO_2 fields do not necessarily resemble modeled NO_2 in either spatial patterns or quantities. Actually, for most cases, they do not. As long as the a priori vertical distributions from other auxiliary sources (either high-resolution models or observations) are more accurately representing the true atmospheric heterogeneous state than the climatological values used in the standard product, the a priori profile adjustment helps improve OMPS NO_2 retrievals, especially in heavily polluted areas.

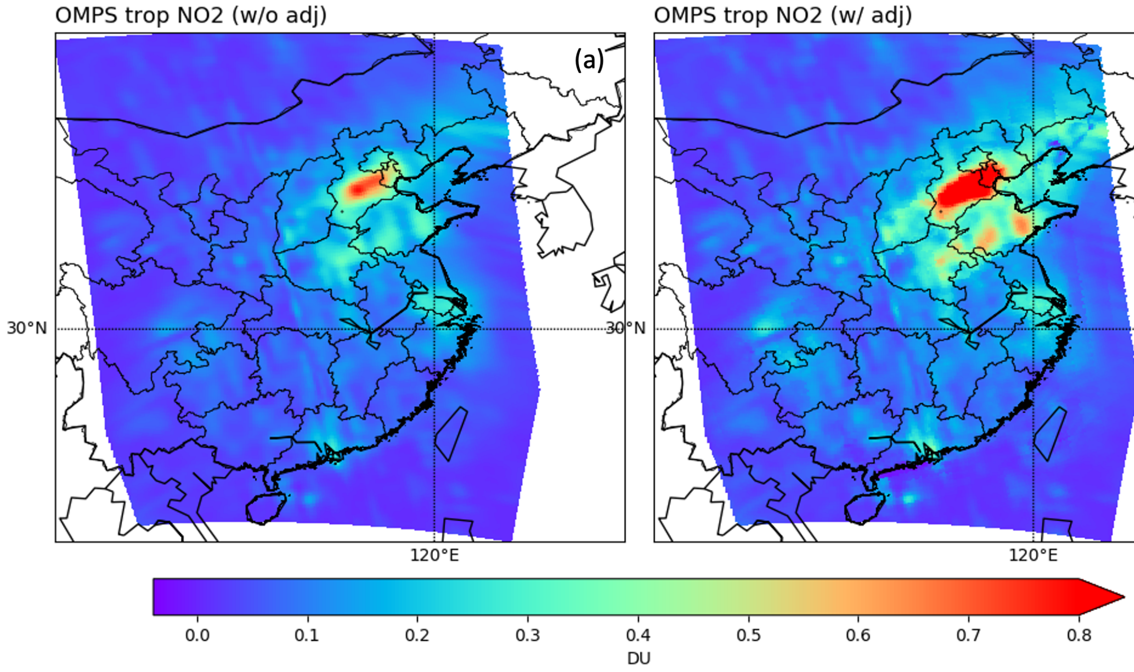


Figure 8.3: OMPS tropospheric NO_2 vertical columns **(a)** before and **(b)** after a priori profile adjustment. OMPS scan time was 2016-05-12 05:33 to 05:42 UTC for the regions shown in the maps. We replaced OMPS a priori profiles with CMAQ (12 km × 12 km) simulated NO_2 profiles at 06:00 UTC.

8.2 Algorithm improvement II: spatial downscaling

It is common to see a satellite pixel of a relatively large footprint not able to resolve the subpixel variability, especially when it is within heterogeneous urban areas. Figure 8.4 illustrates such situations by comparing SNPP measured NO_2 column density with CMAQ simulated column NO_2 over the same satellite pixel area (black polygon) that covers the Wuxi station (circle). Obviously, OMPS measurement is homogenous within its own iFOV (i.e., pixel), however, with a higher resolution, CMAQ sees a strong gradient within one satellite iFOV. There is much more NO_2 pollution in Wuxi city than in the nearby suburban areas. For the next algorithm improvement, we applied a spatial weighting kernel to downscale OMPS NO_2 TVCDs and enhance OMPS spatial resolution.

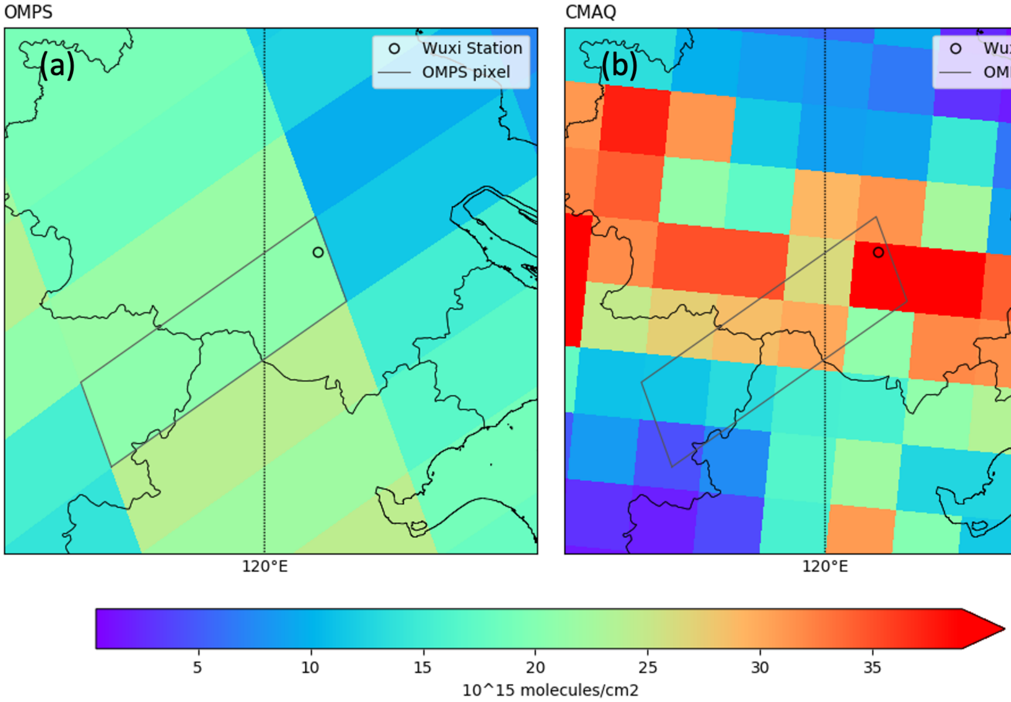


Figure 8.4: Illustration of OMPS subpixel inhomogeneity and downscaling method, **(a)** OMPS tropospheric NO_2 vertical columns, **(b)** CMAQ simulated NO_2 columns. The OMPS pixel that encompasses Wuxi station is highlighted in black on both panels. The OMPS scan time at Wuxi station is 2014-01-14 04:18 UTC and CMAQ output time is 04:00 UTC.

Downscaling is a common concept in meteorological simulations, which is often used in global circulation models to provide initial and boundary conditions for regional models. Using a similar approach as described by Kim et al. (2016), we created a spatial weighting kernel based on the values of CMAQ grid cells that intersect with the satellite pixel (shaded grey cells in Figure 8.5). We collected the CMAQ simulated column densities (integrated sum over altitude) and then normalized them so that the average of the model grid cells used for downscaling equals 1. This whole set of CMAQ grid cells that are made of fractional values is called a spatial weighing kernel. We then applied the spatial kernel to the target OMPS pixel. As a result, we generated a reconstructed OMPS pixel with a finer structure, and at the same time, it conserves the original quantity, which means averaging the reconstructed column densities yields the originally measured OMPS NO_2 column density. Please note our satellite pixel data are in density units (e.g., molecules/cm^2), thus we make sure the average of the spatial kernel equals 1 and the average of the reconstructed satellite pixel equals the original satellite pixel measurement. If the satellite pixel data are in the mass units, the spatial kernel calculation is slightly different.

In the example of Figure 8.4, Wuxi station is located in a polygon where CMAQ indicates the highest amount of NO_2 simulated within the target satellite pixel. The spatial weighting kernel will distribute more weight to the polygon that contains the Wuxi station, so downscaling will adjust the column measurement higher at that location. Therefore, using the downscaling method could potentially alleviate the satellite's underestimation at the urban cores and overestimation at the suburban regions due to the resolution-caused geometry effect (Judd et al., 2019; Kim et al., 2016). However, since this method relies on the accuracy of the model simulated NO_2 relative distribution, it clearly has its own limitations. The wind field simulation and the inputs of emission source locations are important factors for simulating NO_2 plume transport and distribution,

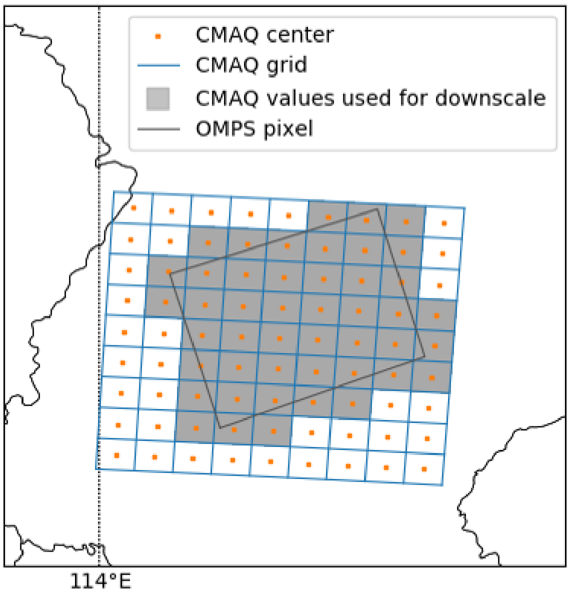


Figure 8.5: Illustration of satellite pixel, CMAQ grid, and how the downscaling method applies: any CMAQ grids that intersect with the OMPS pixel are used for creating the spatial-weighting kernel (shaded in grey).

and their uncertainties could degrade the model performance and thus negatively impact the efficacy of the downscaling technique.

8.3 Algorithm improvement III: explicit aerosol correction scheme

To summarize what previous algorithm improvements have done so far to the retrieval quality, Figure 8.6 shows scatter plots of SNPP OMPS NO_2 SP columns (panel a) and SP columns after the two algorithm improvements (panel b) against MAX-DOAS measured NO_2 TVCDs. Although both SNPP OMPS NO_2 columns underestimate compared with MAX-DOAS, OMPS TVCDs after the a priori profile and spatial downscaling adjustments certainly improves over the OMPS SP columns, in terms of the correlation coefficient as well as the slope of the OLS linear regression statistics. To further explore the satellite underestimation bias, we overlay MAX-DOAS measured aerosol optical depth (AOD) on top of Figure 8.6b. We find that OMPS

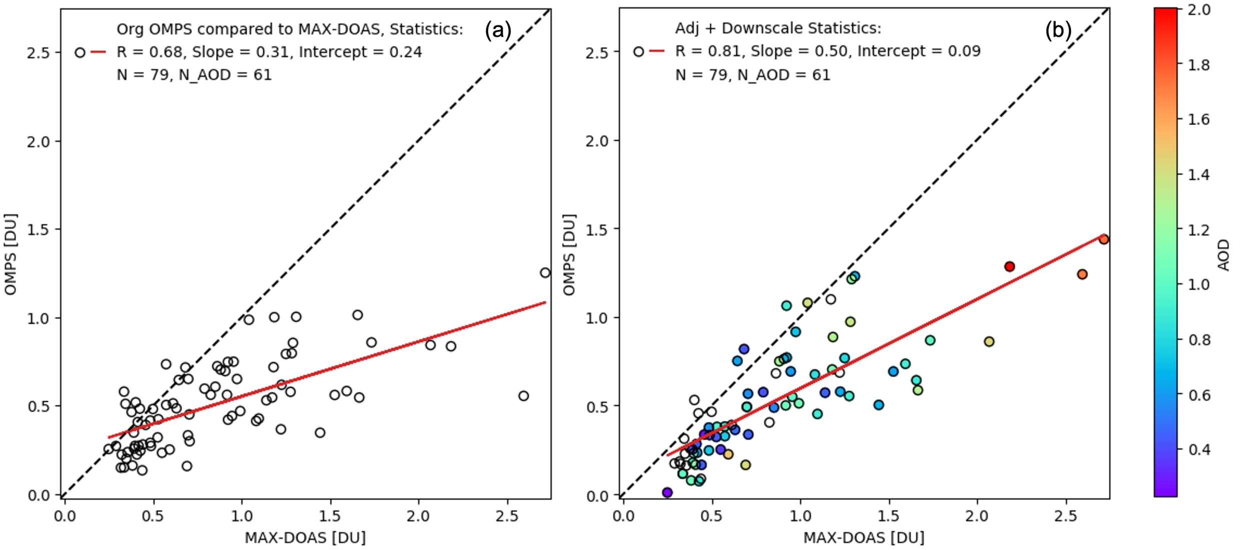


Figure 8.6: **(a)** SNPP OMPS standard product (SP) and **(b)** Improved OMPS SP NO_2 columns (surface to 4 km altitude) with a priori profile & downscale adjustments are compared to MAX-DOAS observed NO_2 tropospheric vertical columns over Wuxi station during 4 months of 2014: January, March, July, and October. MAX-DOAS columns are averaged within ±60 min of OMPS overpass and OMPS data with cloud fraction $>30\%$ are filtered. Points in panel **(b)** are colored by MAX-DOAS AOD at the nearest coincidence, with the open circles indicative of no quality-assured AOD observations available within ±60 min of OMPS overpass. The statistics of OLS linear regression fit are shown on the plots (N represents the number of coincidences).

underestimation trend was predominated by a few high-AOD-high-NO_2 points on the far right of the scatter plot. This systematic worsening of satellite underestimation in measured NO_2 columns with increasing AOD is evident in other satellite instruments as well (Chimot et al., 2016; Lin et al., 2014; Wang et al., 2017a).

Tropospheric NO_2 is usually co-emitted with aerosols (e.g., soot), its transport is frequently mixed with dust and sulfate, and its atmospheric transformation turns it into fine particles (nitrate aerosol). In short, tropospheric NO_2 in polluted areas is usually accompanied by aerosols. The presence of particles (i.e., aerosols, clouds, or both) changes the photon contributions to the satellite-measured UV-Vis radiation compared to a molecular (i.e., particle-free) atmosphere, leading to significant changes in NO_2 measurement sensitivity depending on its relative position with respect to the particles. For instance, particles may reduce the measurement sensitivity of NO_2 when NO_2 is below but enhance the NO_2 sensitivity when NO_2 is above the particles.

Currently, most satellite NO_2 products use implicit aerosol corrections (i.e., the MLER treatment) by assuming aerosols to be effective clouds. For most scattering aerosols at high altitudes, the implicit aerosol correction can largely account for the aerosol effect on the trace gas products (Boersma et al., 2011). However, in other cases, for example, low altitude aerosols with high AOD and small Single Scattering Albedo (SSA), the implicit correction might even increase the errors of the AMF (Castellanos et al., 2015). Frequently, aerosols' maximum optical signatures (total extinction) are located above the NO_2 peak profiles under polluted conditions. For instance, at Xingtai MAX-DOAS station during the ARIAs campaign, 75% of integrated aerosol profiles are below 1.4 km versus 75% of NO_2 columns are below 0.5 km (see Figure 8.7). This implies that the implicit aerosol corrections often fail to account for the particle effects on NO_2 sensitivity, leading to large errors (up to 60%) in the retrieved NO_2 vertical columns.

For most satellite observations, complete aerosol information (loading, vertical distribution, and optical properties) is in general not available. In deriving NO_2 profile from MAX-DOAS observations, aerosol profiles are retrieved simultaneously assuming that the aerosol is represented by a highly simplified aerosol model, whose optical properties are characterized by a fixed single-scattering albedo (SSA) of 0.95 and an aerosol phase function parameterized according to Henyey and Greenstein (1941) with an asymmetry parameter of 0.72 (based on the measurements of sun photometer operated at the Xingtai ground station) (Wang et al., 2019). In our explicit aerosol scheme, we use the carbonaceous aerosol model (BIO, Torres et al. 2007) to provide a realistic representation of air pollution in the urban area. Given the aerosol type and its vertical distribution taken from the MAX-DOAS measurement, the aerosol optical depth and single

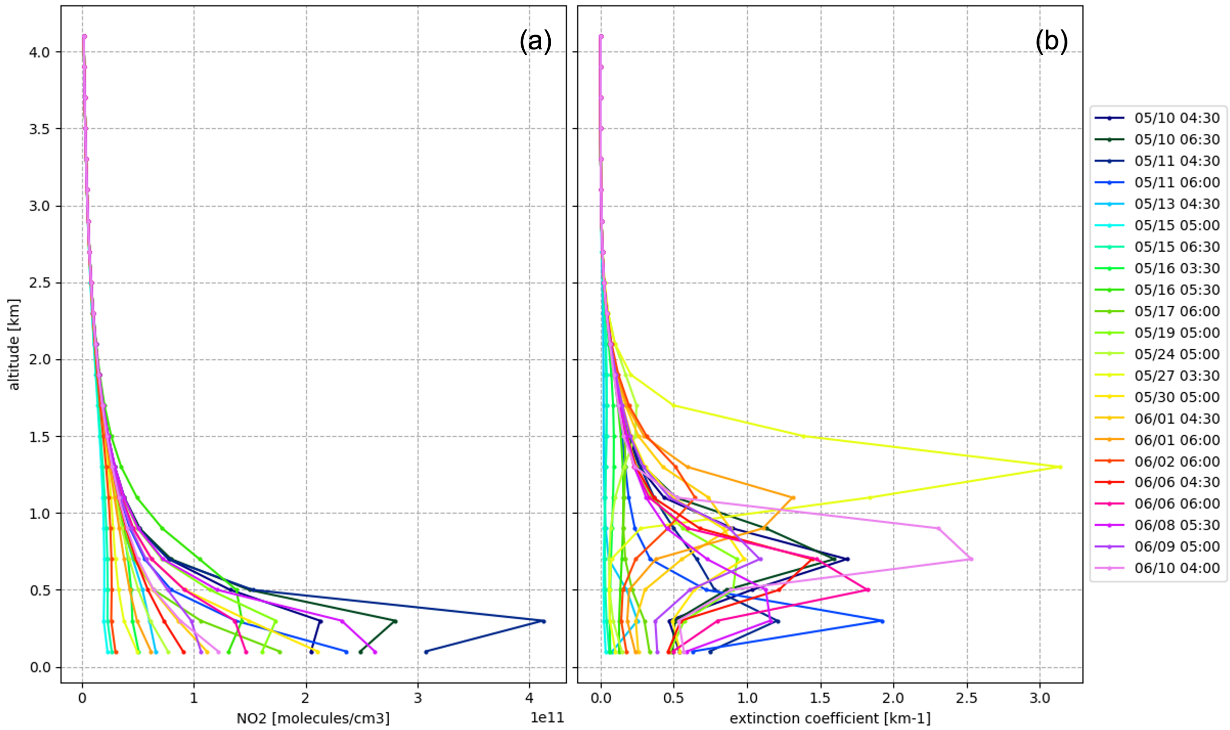


Figure 8.7: MAX-DOAS measured vertical profile of **(a)** NO_2 number density and **(b)** aerosol total extinction at Xingtai station during ARIAs campaign (May to June 2016).

scattering albedo are estimated from the radiance spectra Torres et al. 2007. Radiative transfer modeling with explicit inclusion of the retrieved aerosol loading and its scattering and absorption optical properties is performed next to obtain a more accurate WF. The WF ratio between explicit and MLER treatment yields the multiplication factor to correct standard NO_2 retrieval. Here, we provide an example to illustrate how explicit aerosol treatment improves the accuracy of NO_2 retrieval over the commonly used MLER treatment. In this case, there is a significant amount of NO_2 and UV-absorbing aerosol present in the IFOV. Their vertical distributions are displayed in Figure 8.8(ab), showing that the bulk of NO_2 is below the peak of aerosol extinction (optical thickness). The measured radiance is slightly higher than that of a clean (i.e., aerosol-free) atmosphere due to the backscattering by aerosol. Thus, the MLER treatment detects clouds in the IFOV and retrieves a small cloud fraction due to aerosol absorption reducing the cloud albedo effect (i.e., suppress the backscattering radiance enhancement). Because of the small cloud fraction, the NO_2 measurement sensitivity from the MLER treatment below the aerosol decreases only slightly. On the other hand, including the aerosol explicitly in the radiative transfer calculation reveals a large drop in the NO_2 measurement sensitivity below the aerosol peak, due to both aerosol absorption and scattering. Figure 8.8(c) shows that the WF of MLER treatment is about 1.7 times higher than that of the explicit aerosol treatment, implying that retrieved NO_2 based on the MLER treatment is severely underestimated. This example represents the typical case of aerosol and NO_2 vertical distributions encountered in remote sensing of polluted areas.

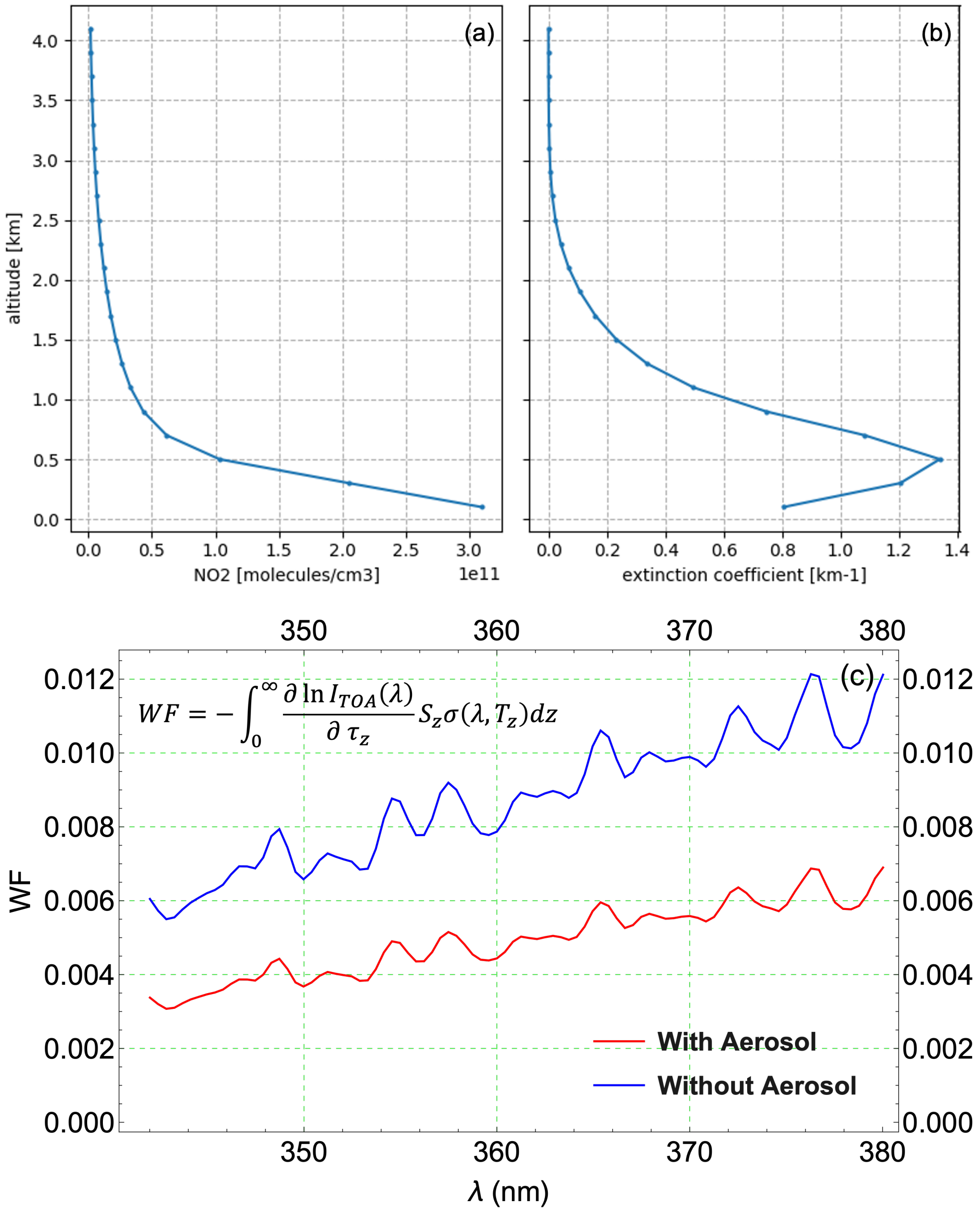


Figure 8.8: MAX-DOAS measured vertical profiles of NO_2 number density **(a)** and aerosol total extinction **(b)** at Wuxi station on 2014-07-07 04:51 UTC. **(c)** The weighting functions (WF) as a function of wavelength with the explicit aerosol treatment (red curve) and MLER treatment (blue curve) of the corresponding aerosol-laden atmosphere as shown in (a) and (b).

8.4 Summary of the aforementioned 3 algorithm improvements: how do they apply?

Section 8.1 to 8.3 have discussed the theoretical basis behind these three algorithm improvements and demonstrated the impact of each individual technique could have on the NO_2 retrievals. In this section, we discuss how these algorithm improvements apply collaboratively in the DVCF retrieval scheme.

The order to apply the aforementioned algorithmic improvements follows a general rule: applying the most intrinsic technique first and applying the posterior process last. While the downscaling technique is useful in terms of resolving satellite sub-pixel variability, it uses model data to adjust existing satellite data but does not address issues inherent with satellite retrievals. Therefore, the downscaling technique should be implemented as the last step. In the standard product, $I_{TOA}(\lambda)$ in Eq. 5.1 is the modeled reflectance spectrum without including absorptions of trace gases and aerosols. Applying the aerosol correction scheme means inherently adding aerosol in the I_{TOA} modeling and thus modifying the averaging kernel, which is an important satellite parameter and will then be used in the a priori profile adjustment. Therefore, the explicit aerosol correction scheme must be applied in the first step of the three algorithm improvements, followed by a priori profile adjustment, and spatial downscaling implemented as the last step.

8.5 Validation of the improved NO_2 tropospheric columns

In the previous sections, we have explained the theoretical principle and established the framework for the three algorithm improvements. In this section, we applied the techniques to one-month data (January 2014) to assess the real-world impact of the algorithm improvements on OMPS column NO_2 retrievals. January is a typical month that we see high particular matter pollution due to increased fossil fuel combustion from higher energy demand, as well as high NO_2 loading in the atmosphere due to longer NO_2 lifetime during the winter. This makes it a perfect test case to evaluate the efficacy of the explicit aerosol correction scheme.

Figure 8.9 compares OMPS baseline and improved NO_2 TVCDs to those measured from MAX-DOAS. The algorithm improvements were applied incrementally to OMPS NO_2 SP: the aerosol correction was applied first in panel b (green dots); then the a priori profile adjustment was applied on top of aerosol correction, showing in panel c (blue dots); lastly, the spatial downscaling was applied on top of the previous two improvements, showing in panel d (red dots). The comparison suggests that the algorithm improvements acting collaboratively make OMPS columns have a better agreement with the MAX-DOAS measured NO_2 columns, in terms of closer relation to the 1:1 line and improved correlation coefficient.

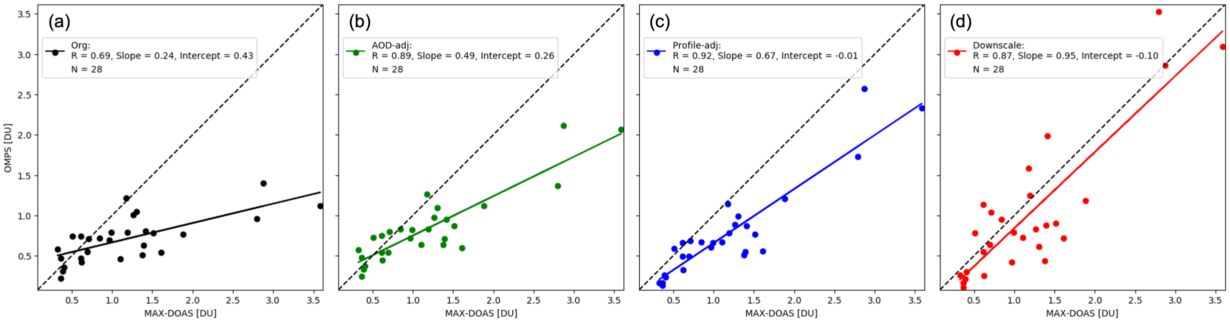


Figure 8.9: Scatter plots of SNPP OMPS vs. MAX-DOAS observed NO_2 tropospheric vertical columns (surface to 4 km altitude) at Wuxi station during January 2014, for **(a)** OMPS SP, **(b)** OMPS SP with explicit aerosol correction, **(c)** OMPS SP with explicit aerosol correction plus a priori profile adjustments, **(d)** OMPS SP with explicit aerosol correction, a priori profile and downscaling adjustments. Temporally closest MAX-DOAS coincidence and OMPS pixels with cloud fraction of 30% or less were included in the comparison.

To visualize the impact of the algorithm improvements on OMPS retrievals progressively, we plot the linear regression statistics of this case validation in the slope-correlation coordinate in Figure 8.10. Each color in Figure 8.10 corresponds to the improvements made in the scatter plots: black for Figure 8.9(a), green for Figure 8.9(b), blue for Figure 8.9(c), and red for Figure 8.9(d). The best case scenario is indicated by 'Target' on the plot at the coordinate of slope = 1 and r = 1. From 'SP' to 'Aerosol', 'Profile', and 'Downscale', we can clearly see the algorithm improvements are moving progressively towards the 'Target' (i.e., best performance where slope = 1 and r = 1). The first algorithm improvement (aerosol correction) has advanced both linear regression slope and correlation significantly compared to SP. The second algorithm improvement (a priori profile adjustment) shows decent improvement in the slope yet less improvement as seen in terms of correlation. For the third algorithm improvement (spatial downscaling), although it has improved the linear regression slope significantly, the correlation coefficient is slightly lowed down compared with the previous two algorithm improvements. This is possibly because downscaling uses model information to redistribute the observed column density within the satellite pixel, the reconstructed field may deviate from the measurement truth, causing the reduced correlation. In summary, aerosol correction and a priori profile adjustment inherently improve correlation to the reference data, as well as the linear regression slope. Spatial downscaling mainly improves slope and might have a detrimental effect on correlation depending on model performance. Among these three algorithm improvements, the explicit aerosol correction scheme has the biggest impact on the satellite NO_2 column retrievals, as it helps improve the measurement correlation the most significantly, especially in the PM-polluted regions. These results highlight the path of algorithm improvements for operational NO_2 retrievals from other satellite observation platforms.

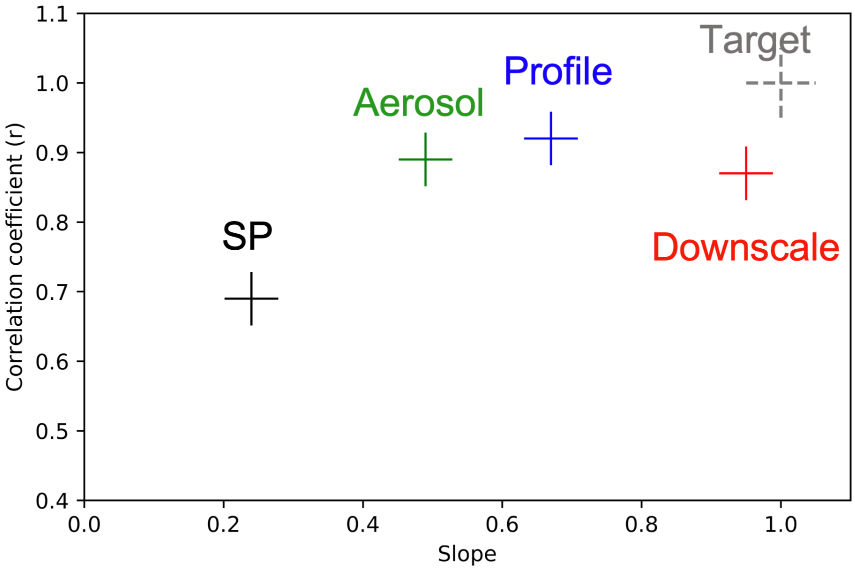


Figure 8.10: Linear regression statistics visualized for the example case validation in Figure 8.9. Each color corresponds to the improvements made in scatter plots: black for Figure 8.9(a), green for Figure 8.9(b), blue for Figure 8.9(c), and red for Figure 8.9(d). Performance is presented in terms of the OLS regression slope (x-axis) and correlation coefficient r (y-axis). The best case scenario is indicated by Target on the plot at the coordinate of slope = 1 and r = 1.

For the explicit aerosol correction, we currently only performed one month (January 2014) of data for testing purposes. For the other two algorithm improvements including a priori profile adjustment and spatial downscaling, we have expanded our validation to Xingtai and Wuxi stations during the time periods where CMAQ simulations are available. We also explored the column comparisons under different coincidence criteria, including different temporal windows to average the MAX-DOAS observations, as well as different cloud filters for the OMPS data. The statistics of the comparisons between SP, a priori profile adjusted, and spatially downscaled columns are summarized in Table 8.1 for Xingtai and Table 8.2 for Wuxi. From the comparisons, we can see both algorithm improvements have ubiquitously improved the column statistics no matter what temporal windows or cloud filters are used, although the degree of improvement may vary between different coincidence criteria. In general, the best linear regression statistics

Table 8.1: SNPP OMPS vs. MAX-DOAS column statistics based on coincidence criteria at Xingtai station during 2016 ARIAs campaign. The highest correlation score after two algorithm improvements is highlighted in bold.

Coincidence Criteria		**Statistics**						
Temporal window for MAX-DOAS Mean	Filter by OMPS Radiative Cloud Fraction	Original		Profile adjusted		Downscaled		Number of samples
		R	Slope	R	Slope	R	Slope	
± 90 min	30%	0.49	0.30	0.56	0.38	0.77	1.02	23
± 60 min		0.46	0.27	0.54	0.34	0.77	0.97	23
± 15 min		0.61	0.46	0.68	0.57	**0.88**	1.34	19
Nearest Coincidence		0.55	0.39	0.64	0.50	0.86	1.31	23
Nearest Coincidence (exclude if greater than ± 90 min)	10%	0.42	0.20	0.59	0.31	0.75	0.98	12
	20%	0.54	0.39	0.63	0.50	0.80	1.17	20
	35%	0.53	0.37	0.63	0.49	0.83	1.27	24
	All Clouds	0.56	0.35	0.65	0.45	0.61	0.99	36

are provided by MAX-DOAS temporal window within ± 60 min and OMPS cloud fraction filter $\sim 30\%$. These results consistently demonstrate the effectiveness of the three proposed algorithm improvements, indicating a clear direction toward the retrieval algorithmic development for future satellite product releases to achieve better accuracy.

Table 8.2: Similar to Table 8.1, but for Wuxi station during January, March, July, and October of 2014.

Coincidence Criteria		**Statistics**						
Temporal window for MAX-DOAS Mean	Filter by OMPS Radiative Cloud Fraction	Original		Profile adjusted		Downscaled		Number of samples
		R	Slope	R	Slope	R	Slope	
± 90 min	30%	0.66	0.31	0.76	0.39	0.79	0.51	79
± 60 min		0.68	0.31	0.78	0.39	**0.81**	0.50	79
± 15 min		0.65	0.27	0.74	0.34	0.74	0.40	68
Nearest Coincidence		0.58	0.20	0.68	0.26	0.71	0.34	79
Nearest Coincidence (exclude if greater than ± 90 min)	10%	0.56	0.19	0.63	0.22	0.63	0.23	40
	20%	0.57	0.21	0.65	0.25	0.67	0.31	67
	35%	0.58	0.21	0.68	0.26	0.72	0.34	82
	All Clouds	0.48	0.20	0.59	0.26	0.66	0.34	107

Chapter 9: Summary and future work

9.1 Summary

In this dissertation, I have presented the retrievals of atmospheric trace gases from the UV measurements made from operational satellite platforms, including retrieving O_3 and SO_2 from DISCOVR EPIC, and retrieving NO_2 from NOAA-20 OMPS and SNPP OMPS instruments. All the trace gas retrievals are built upon the DVCF algorithm, which is designed to maximize the absorption signature from the Earth's atmosphere in the UV spectral range. We performed algorithm improvements, detailed error analysis, validation against ground-based measurements, inter-comparisons with OMI instrument and MERRA-2 reanalysis data, a case study of the volcanic plumes using EPIC observations, and a case study of the drastic NO_2 changes during COVID-19 using NOAA-20 OMPS observations. The algorithm improvements have demonstrated enhanced performance for retrieving O_3, SO_2, and NO_2 from the spaceborne UV measurements. Overall, trace gas retrievals provide long-term high-quality observations to monitor changes in atmospheric composition/ air quality and play a crucial role in chemical data assimilation, air quality modeling and forecast, and regulatory decision-making.

The first part of my Ph.D. research discusses the algorithmic theoretical basis and algorithm advances for retrieving the O_3 and SO_2 from EPIC UV measurements. The algorithm advances presented in the thesis, including using the improved O_3 profile representation, the regulated direct fitting inversion technique, and the SOE (Spatial Optimal Estimation) scheme, significantly improve the accuracy of O_3 and SO_2 retrievals from the multi-channel measurements of DSCOVR EPIC.

A thorough error analysis is provided to quantify O_3 and SO_2 retrieval uncertainties due to various error sources and simplified algorithm physics treatments. Error analysis findings indicate that the MLER treatment of UV-absorbing aerosols leads to significant uncertainties in retrieved O_3 and SO_2 columns. Future improvements may include explicit aerosol treatment or other schemes for radiance or product corrections. The GLER treatment of anisotropic surface reflections introduces small errors in the retrieved total O_3 and SO_2 columns, primarily because surface reflection is a minor component of measured radiance in the UV. However, this GLER treatment does not generally provide a more accurate tropospheric AMF. Hence, explicit BRDF treatment of surface reflection is needed for accurate retrievals of tropospheric gases.

The EPIC total O_3 columns are validated against coincident ground-based Brewer measurements and compared with coincident O_3 data from MERRA-2 assimilation. The findings show that EPIC total O_3 is highly accurate, capturing the short-term O_3 variability realistically while maintaining long-term consistency over the entire record. The EPIC SO_2 loading of volcanic plumes is evaluated against those from hyperspectral measurements of the same eruptions, showing that EPIC provides accurate SO_2 quantifications from large volcanic eruptions. EPIC's high-cadence observations allow better identification of the peak loading of volcanic SO_2 plume compared to polar-orbiting instruments.

In the second part of this dissertation, I have presented a suite of product development behind the new NOAA-20 OMPS tropospheric NO_2 columns, covering retrieval algorithm, instrument measurement sensitivity assessment, inter-comparison with OMI, validation against Pandora direct-sun measurements, and application during COVID-19. We applied the DVCF algorithm and an effective stratosphere-troposphere-separation (STS) scheme to the UV measurements from NOAA-20 OMPS-NM, which methods were successfully used to retrieve NO_2 from its predecessors: SNPP OMPS and Aura OMI.

To evaluate NOAA-20 OMPS NO_2 column retrievals, we first compared the stratospheric NO_2 vertical columns derived from OMPS to those from OMI. The comparison shows excellent agreement in detecting the stratospheric background columns between the two instruments, which facilitates the accuracy of the remained OMPS tropospheric NO_2 retrievals. The result also validates the sliding-median STS scheme that is adopted in NOAA-20 OMPS, especially given the agreement relies on independent spectral measurements at different wavelengths using very different retrieval methods. We compared NOAA-20 OMPS with OMI monthly mean TVCDs observations for July and December 2019. It shows similar spatial distributions and good quantitative agreement. We then preliminarily validated OMPS NO_2 columns against the independent NO_2 measurements from 4 ground-based Pandora spectrometers over the NYC metro area. NOAA-20 NO_2 observations are biased low against (-28%) and are moderately correlated (r = 0.45) with Pandora total columns. The evaluation was then extended to other U.S. Pandora stations, with a total of 13 stations compared with NOAA-20 OMPS. The results suggest that OMPS NO_2 total columns underestimate for relatively large Pandora NO_2 total columns, corresponding to polluted urban regions and episodes of elevated pollution, while overestimating for relatively small NO_2 total columns. Part of the low biases is expected and can be explained by spatial representativity mismatch between satellite and ground-based measurements, when an area-averaged quantity over a relatively large satellite pixel is compared with Pandora observations that have small iFOV. Such kind of spatial representativity mismatch is often associated with localized large pollution enhancements observed by Pandora and OMPS is spatially averaged with nearby less-polluted locations within the larger satellite pixel area. Apart from that, the biases (both underestimation and overestimation) are possibly caused by the coarse a priori profiles currently used in the NOAA-20 NO_2 retrievals. Replacing the a priori NO_2 profiles from high-resolution chemical transport models could potentially improve the agreement. Finally, with the new NOAA-20 OMPS NO_2 retrievals, we investigated the impact of the COVID-19 lockdown on urban NO_2 air pollution. It shows a 20-40% drastic decline in tropospheric NO_2 around the world in January-April 2020 during COVID-19 precautions, supporting the analyses from other satellite-based studies (Bauwens et al., 2020; Goldberg et al., 2020; Liu et al., 2020). These results demonstrate the high sensitivity of NOAA-20 OMPS to tropospheric NO_2 and validate its potential use for extending the long-term global NO_2 record on the series of OMPS-NMs aboard JPSS satellites.

The third part of my dissertation has focused on validation and algorithm improvements for the tropospheric NO_2 column retrievals from SNPP OMPS UV measurements. We first validated SNPP NO_2 TVCDs standard product (SP) against the MAX-DOAS NO_2 measurements acquired from the Wuxi (China) station during 2011 - 2014 and from the Xingtai (China) station between May and June 2016. Both ground stations are located in one of the most urbanized and industrialized regions in China and provide a broad range of NO_2 data for the uncertainty assessment of OMPS products. Like what we have found in Huang et al. (2022) that NOAA-20 NO_2 total columns tend to be lower than ground-based MAX-DOAS in polluted urban regions and higher in clean areas/episodes associated with relatively small NO_2 columns. The SNPP NO_2 column retrievals generally show underestimation compared to the coincidence measurements from MAX-DOAS, because these two MAX-DOAS sites are located in the urban regions of Eastern China and measured much higher column densities than the Pandora US stations. The

SNPP standard columns are modestly correlated with NO_2 columns measured by MAX-DOAS (r = 0.59, N = 346). In order to achieve higher retrieval accuracy, we developed and implemented a series of algorithm improvements, including the replacement of climatological a priori NO_2 profile with the more accurate NO_2 vertical distribution obtained from high-resolution (12×12 km^2) CMAQ simulations, the explicit aerosol corrections to account for changes in measurement sensitivity caused by aerosol scattering and absorption, and the application of model-derived spatial weighting kernel to account for the effect of heterogeneous subpixel distribution. These improvements yield more accurate OMPS NO_2 retrievals in better agreement with MAX-DOAS NO_2 measurements. From our testing case validation during January 2014, we found that aerosol correction and a priori profile adjustment consistently improve correlation to the reference data as well as the linear regression slope (r = 0.92, N = 28). Spatial downscaling mainly improves slope (bias) and helps resolve NO_2 column subpixel variations. It might have a detrimental effect on correlation depending on input model accuracy. Among the three algorithm improvements, the explicit aerosol correction scheme is crucial for improving satellite NO_2 observations in highly polluted regions, as it helps improve both OMPS correlation and bias to the reference data the most significantly. Overall, our results demonstrate the high sensitivity of SNPP OMPS to NO_2 in the troposphere and validate its use for monitoring the spatiotemporal variability of NO_2 in urban environments and extending the long-term global NO_2 record.

9.2 Future work

This dissertation addresses the current status & progress regarding trace gas retrievals from spaceborne UV measurements via the DVCF algorithm. The algorithm improvements we have proposed in the thesis are not limited to one instrument or one type of trace gas retrieval, they can be well explored for other applications. Some of the future directions are listed as follows:

1) The theoretical principle of the SOE approach, introduced to reduce retrieval artifacts due to EPIC's band-to-band misregistration, can be exploited for other applications, such as the separation of a spatially smooth data field (e.g., stratospheric O_3) from that (e.g., tropospheric O_3) with higher spatial variations.

2) A high-resolution regional model with the most up-to-date emission inventory is certainly better in capturing the spatial heterogeneities in urban areas, using it in lieu of the GEOS-chem global model to provide NO_2 shape profiles can yield more accurate retrieval results. However, the modeled climatological profiles invariably differ from the actual measured profiles. In the future, we will develop a retrieval scheme to improve the profile representation by selecting from a group of likely NO_2 profiles, instead of a single shape factor, based on the column amount and profile information contained in spectral measurements.

3) The MAX-DOAS measured aerosol total extinction profiles at the validation sites are currently employed in our explicit aerosol correction scheme to test the efficacy of the method. In the future, we can apply model simulated total extinction profiles into the established explicit aerosol scheme to account for the impact of aerosol absorption and scattering on trace gas retrievals over a region or entire globe. The total extinction profiles can be calculated based on model simulations of aerosol type (refractive index), number density and size distribution .etc.

4) The development of the explicit BRDF treatment of surface reflection is needed for the next-generation accurate retrieval of tropospheric trace gases, as the GLER treatment does not generally provide a more accurate tropospheric AMF.

Satellite remote sensing has been a crucial and reliable technology to provide extensive global earth observation data that fuels the research efforts of atmospheric science communities. The continuous and ongoing advancements in retrieval algorithms will enhance the accuracy of these measurements, offering valuable insights for researchers, stakeholders, and governments to inform policy decisions.

Appendix A: Miscellaneous supplementary information for Chapter 6

A.1 NOAA-20 OMPS Stratospheric NO_2 comparison with OMI

When we compare the stratospheric NO_2 vertical columns between NOAA-20 OMPS and OMI in Section 6.1, we grid OMI Level-2 data in the same way as we grid OMPS Level-2 data. It is identical sampling in terms of filter criterion, gridding method, and date duration. The only sampling difference that could lead to discrepancy is that OMI data have row anomaly and thus some cross-track positions are unavailable, while OMPS has full cross-track data available.

To investigate the possible OMI-row-anomaly caused sampling mismatch, we compare OMI cross-track positions that are not affected by row anomaly and bad VCD quality flags against the equivalent OMPS cross-track positions. Specifically, we compare OMI orbits with row index from 0 to 20 against OMPS orbits with row index from 0 to 47. The selection of row indexes is based on similar Viewing Zenith Angle (VZA). We also make sure the descending part of the orbits are excluded from both instruments. We then do the same seasonal average with zonal bins of 2°, and plot Figure A.1. Compared to Figure 6.1, it shows slight improvements over the high latitudes. We can conclude that the row anomaly caused sampling mismatch are not the main reason for the large discrepancy at high latitudes, as the stratospheric columns are relatively homogeneous across longitudes. The larger discrepancy in the high latitudes is likely due to the coupling of large solar zenith angle (SZA) and the strong sunlight sensitivity of the Nitrogen

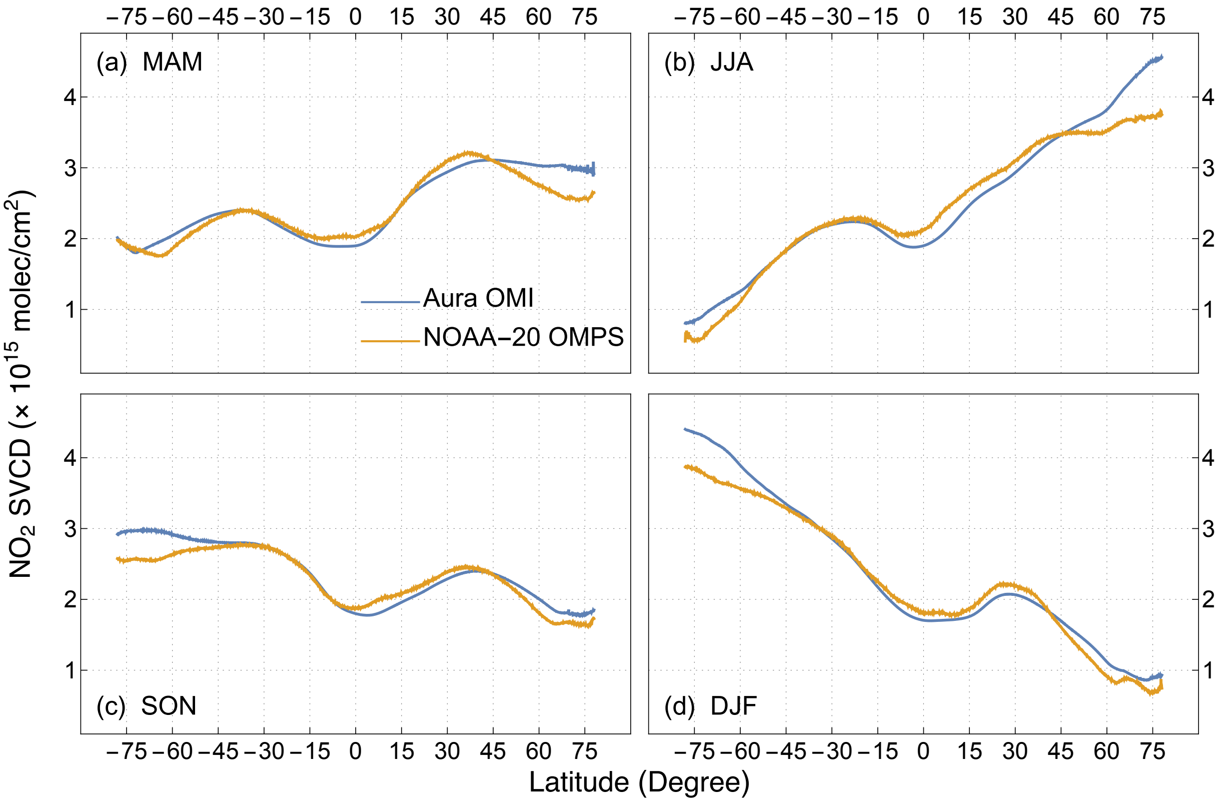


Figure A.1: Seasonal averaged stratospheric NO_2 vertical columns observed from NOAA-20 OMPS (orange curve) and OMI (blue curve) as a function of latitude for (a) MAM, (b) JJA, (c) SON, (d) DJF, over the period from 2019-03-01 to 2020-04-30. OMI cross-track positions that are not affected by row anomaly and bad VCD quality flags (row index from 0 to 20) are compared against the equivalent OMPS cross-track positions (row index from 0 to 47) based on VZA.

chemistry.

A.2 Brief summary of OMI and Pandora column NO_2 comparisons in the literature

Pandora have been widely used to evaluate OMI retrievals in the past. The results of OMI Pandora column NO_2 agreement from the literature can well serve as a reference for the OMPS NO_2 product. Below is a brief summary of some of the previous OMI studies.

1. OMI standard product (SP) total NO_2 columns are fairly correlated (r = 0.25, N = 52) with Pandora observations at 6 DISCOVER-AQ Maryland sites during July 2011 and in

agreement to within 30%. OMI SP total NO_2 columns are moderately correlated (r = 0.5, N = 163) with Pandora at the CAPABLE site at NASA Langley in 2010-2012 (Lamsal et al., 2014).

2. OMI SP and Pandora total NO_2 columns have moderate correlation r = 0.51 in Helsinki in 2012 (Ialongo et al., 2016).

3. Comparisons of OMI SP and Pandora total NO_2 columns are investigated at different locations around the world (Herman et al., 2019), the results of correlations are listed as follows:

    - Boulder, CO, r = 0.20, 2014-2018
    - NASA HQ, Washington DC, r = 0.41, 2014-208
    - Mauna Loa, Hawaii, r = 0.10, 2015-2018
    - Waterflow, New Mexico, r = 0.36, 2012-2018
    - Seoul, South Korea, r = 0.33, 2012-2018
    - Busan, South Korea, r = 0.30, 2012-2018
    - Buenos Aires, Argentina, r = 0.40, 2017-2018

These results suggest that OMPS and OMI have comparable agreement with regard to Pandora measurements. Similar to OMPS, OMI column NO_2 shows underestimation over polluted city stations and overestimation over cleaner areas/episode as well.

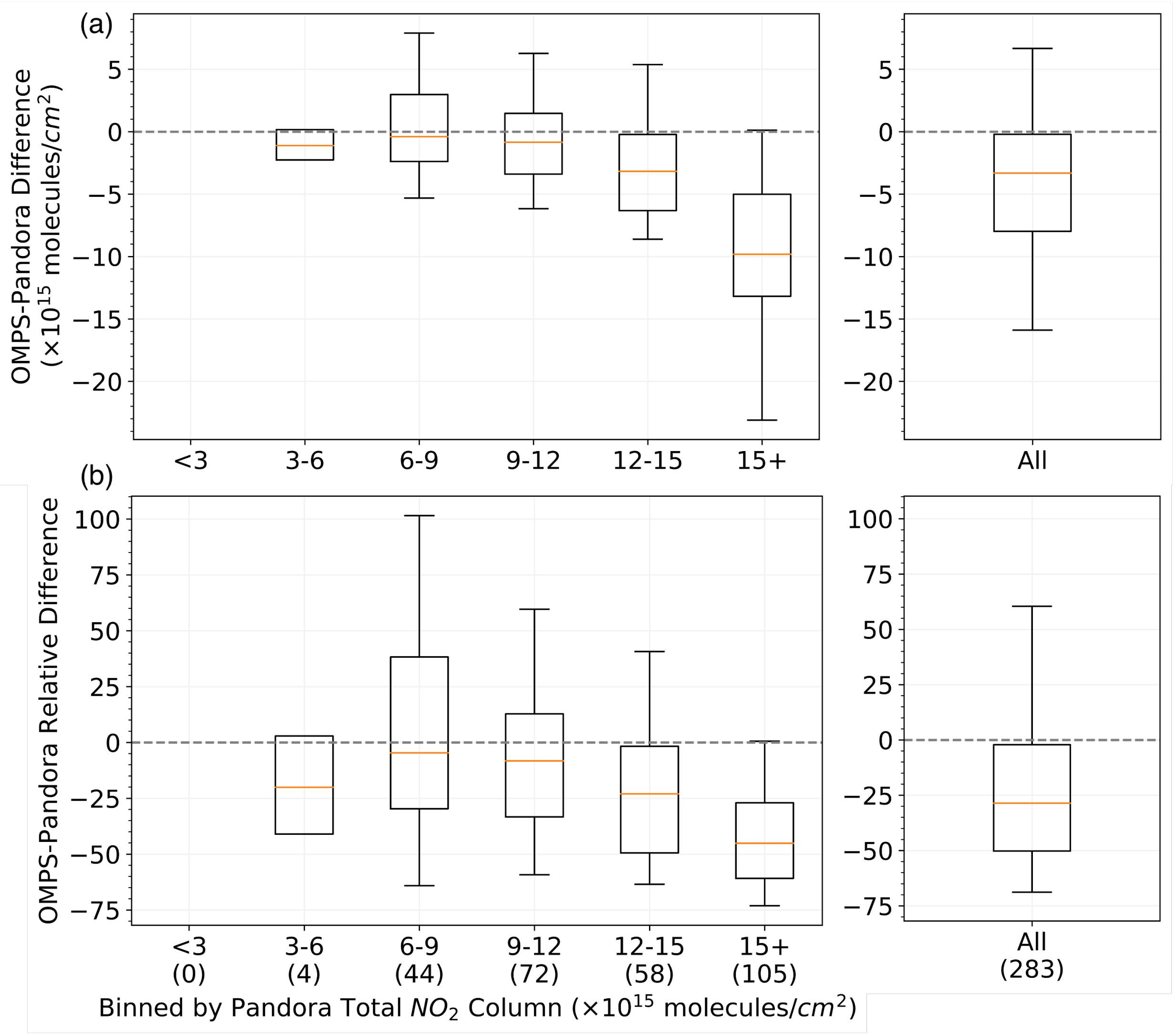


Figure A.2: Box-whisker plots (95-75-50-25-5 percentiles) showing the **(a)** absolute difference and **(b)** relative difference between NOAA-20 OMPS and Pandora measured total NO_2 columns, binned by Pandora columns at the labeled thresholds (left), as well as all data points (right). The number of points in each bin and all data are indicated by the numbers in parentheses. The data used in this plot are collected from the 4 Pandora stations in the NYC metro area as shown in Figure 6.4(b), over a period from 2019-02-14 to 2020-04-30.

Bibliography

Ahmad, Z., Bhartia, P. K., and Krotkov, N. A.: Spectral properties of backscattered UV radiation in cloudy atmospheres, Journal of Geophysical Research, 109, D01 201, https://doi.org/10.1029/2003JD003395, 2004.

Bak, J., Liu, X., Wei, J. C., Pan, L. L., Chance, K. V., and Kim, J. H.: Improvement of omi ozone profile retrievals in the upper troposphere and lower stratosphere by the use of a tropopause-based ozone profile climatology, Atmospheric Measurement Techniques, 6, 2239–2254, https://doi.org/10.5194/amt-6-2239-2013, 2013.

Bak, J., Liu, X., Birk, M., Wagner, G., Gordon, I. E., and Chance, K.: Impact of using a new ultraviolet ozone absorption cross-section dataset on OMI ozone profile retrievals, Atmospheric Measurement Techniques, 13, 5845–5854, https://doi.org/10.5194/amt-13-5845-2020, 2020.

Bauwens, M., Compernolle, S., Stavrakou, T., Müller, J. F., van Gent, J., Eskes, H., Levelt, P. F., van der A, R., Veefkind, J. P., Vlietinck, J., Yu, H., and Zehner, C.: Impact of Coronavirus Outbreak on NO_2 Pollution Assessed Using TROPOMI and OMI Observations, Geophysical Research Letters, 47, 1–9, https://doi.org/10.1029/2020GL087978, 2020.

Beirle, S., Boersma, K. F., Platt, U., Lawrence, M. G., and Wagner, T.: Megacity Emissions and Lifetimes of Nitrogen Oxides Probed from Space, Science, 333, 1737–1739, https://doi.org/10.1126/science.1207824, 2011.

Beirle, S., Borger, C., Dörner, S., Li, A., Hu, Z., Liu, F., Wang, Y., and Wagner, T.: Pinpointing nitrogen oxide emissions from space, Science Advances, 5, 1–7, https://doi.org/10.1126/sciadv.aax9800, 2019.

Bell, M. L., McDermott, A., Zeger, S. L., Samet, J. M., and Dominici, F.: Ozone and Short-term Mortality in 95 US Urban Communities, 1987-2000, JAMA, 292, 2372–2378, https://doi.org/10.1001/jama.292.19.2372, 2004.

Benish, S., He, H., Ren, X., J. Roberts, S., J. Salawitch, R., Li, Z., Wang, F., Wang, Y., Zhang, F., Shao, M., Lu, S., and R. Dickerson, R.: Measurement report: Aircraft observations of ozone, nitrogen oxides, and volatile organic compounds over Hebei Province, China, Atmospheric Chemistry and Physics, 20, 14 523–14 545, https://doi.org/10.5194/acp-20-14523-2020, 2020.

Bhartia, P. K. and Wellemeyer, C. G.: TOMS-V8 Total O_3 Algorithm, in: OMI Algorithm Theoretical Basis Document, edited by Bhartia, P. K., vol. II, chap. 2, pp. 15–32, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA, 2 edn., https://eospso.gsfc.nasa.gov/sites/default/files/atbd/ATBD-OMI-02.pdf, 2002.

Birk, M. and Wagner, G.: ESA SEOM-IAS – Measurement and ACS database SO2 UV region, https://doi.org/10.5281/zenodo.1492582, 2018.

Birk, M. and Wagner, G.: ESA SEOM-IAS – Measurement and ACS database O3 UV region, https://doi.org/10.5281/zenodo.4423918, 2021.

Blank, K.: EPIC Geolocation and Color Imagery Algorithm Revision 6, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA, https://doi.org/10.5067/EPIC/DSCOVR/L1B.003, 2019.

Blank, K., Huang, L.-K., Herman, J., and Marshak, A.: Earth Polychromatic Imaging Camera Geolocation; Strategies to Reduce Uncertainty, Frontiers in Remote Sensing, 2, 1–9, https://doi.org/10.3389/frsen.2021.715296, 2021.

Boersma, K. F., Eskes, H. J., Veefkind, J. P., Brinksma, E. J., Van Der A, R. J., Sneep, M., Van Den Oord, G. H., Levelt, P. F., Stammes, P., Gleason, J. F., and Bucsela, E. J.: Near-real time retrieval of tropospheric NO2 from OMI, Atmospheric Chemistry and Physics, 7, 2103–2118, https://doi.org/10.5194/acp-7-2103-2007, 2007.

Boersma, K. F., Eskes, H. J., Dirksen, R. J., van der A, R. J., Veefkind, J. P., Stammes, P., Huijnen, V., Kleipool, Q. L., Sneep, M., Claas, J., Leitão, J., Richter, A., Zhou, Y., and Brunner, D.: An improved tropospheric NO_2 column retrieval algorithm for the Ozone Monitoring Instrument, Atmospheric Measurement Techniques, 4, 1905–1928, https://doi.org/10.5194/amt-4-1905-2011, 2011.

Bogumil, K., Orphal, J., Homann, T., Voigt, S., Spietz, P., Fleischmann, O., Vogel, A., Hartmann, M., Kromminga, H., Bovensmann, H., Frerick, J., and Burrows, J.: Measurements of molecular absorption spectra with the SCIAMACHY pre-flight model: instrument characterization and reference data for atmospheric remote-sensing in the 230–2380 nm region, Journal of Photochemistry and Photobiology A: Chemistry, 157, 167–184, https://doi.org/https://doi.org/10.1016/S1010-6030(03)00062-5, atmospheric Photochemistry, 2003.

Bosilovich, M., Akella, S., Coy, L., Cullather, R., Draper, C., Gelaro, R., Kovach, R., Liu, Q., Molod, A., Norris, P., Wargan, K., Chao, W., Reichle, R., Takacs, L., Vikhliaev, Y., Bloom, S., Collow, A., Firth, S., Labow, G., Partyka, G., Pawson, S., Reale, O., Schubert, S. D., and Suarez, M.: MERRA-2 : Initial Evaluation of the Climate, in: NASA Technical Report Series on Global Modeling and Data Assimilation NASA/TM–2015-104606, edited by Randal D. Koster, vol. 43, p. 139, Goddard Space Flight Center, Greenbelt, Maryland, USA, https://gmao.gsfc.nasa.gov/pubs/docs/Bosilovich803.pdf, 2015.

Bovensmann, H., Burrows, J. P., Buchwitz, M., Frerick, J., Noël, S., Rozanov, V. V., Chance, K. V., and Goede, A. P. H.: SCIAMACHY: Mission Objectives and Measurement Modes, Journal of the Atmospheric Sciences, 56, 127–150, https://doi.org/10.1175/1520-0469(1999)056⟨0127:SMOAMM⟩2.0.CO;2, 1999.

Brennan, B. and Bandeen, W. R.: Anisotropic Reflectance Characteristics of Natural Earth Surfaces, Applied Optics, 9, 405, https://doi.org/10.1364/AO.9.000405, 1970.

Brion, J., Chakir, A., Daumont, D., Malicet, J., and Parisse, C.: High-resolution laboratory absorption cross section of O3. Temperature effect, Chemical Physics Letters, 213, 610–612, https://doi.org/https://doi.org/10.1016/0009-2614(93)89169-I, 1993.

Brodzik, M. J. and Stewart, J. S.: Near-Real-Time SSM/I-SSMIS EASE-Grid Daily Global Ice Concentration and Snow Extent, Version 3, NASA National Snow and Ice Data Center Distributed Active Archive Center, Boulder, Colorado, USA, https://doi.org/10.5067/JAQDJKPX0S60, 2021.

Bréon, F.-M., Tanré, D., and Generoso, S.: Aerosol Effect on Cloud Droplet Size Monitored from Satellite, Science, 295, 834–838, https://doi.org/10.1126/science.1066434, 2002.

Cai, S., Wang, Y., Zhao, B., Wang, S., Chang, X., and Hao, J.: The impact of the “Air Pollution Prevention and Control Action Plan” on PM2.5 concentrations in Jing-Jin-Ji region during 2012–2020, Science of The Total Environment, 580, 197–209, https://doi.org/https://doi.org/10.1016/j.scitotenv.2016.11.188, 2017.

Callies, J., Corpaccioli, E., Eisinger, M., Hahne, A., and Lefebvre, A.: GOME-2 – Metop's Second-Generation Sensor for Operational Ozone Monitoring, ESA Bulletin, http://www.esa.int/esapub/bulletin/bullet102/Callies102.pdf, 2000.

Carn, S. A., Krueger, A. J., Krotkov, N. A., Yang, K., and Evans, K.: Tracking volcanic sulfur dioxide clouds for aviation hazard mitigation, Natural Hazards, 51, 325–343, https://doi.org/10.1007/s11069-008-9228-4, 2009.

Castellanos, P., Boersma, K. F., Torres, O., and de Haan, J. F.: OMI tropospheric NO_2 air mass factors over South America: effects of biomass burning aerosols, Atmospheric Measurement Techniques, 8, 3831–3849, https://doi.org/10.5194/amt-8-3831-2015, 2015.

Caudill, T. R., Flittner, D. E., Herman, B. M., Torres, O., and McPeters, R. D.: Evaluation of the pseudo-spherical approximation for backscattered ultraviolet radiances and ozone retrieval, Journal of Geophysical Research: Atmospheres, 102, 3881–3890, https://doi.org/https://doi.org/10.1029/96JD03266, 1997.

Cede, A., Kang Huang, L., McCauley, G., Herman, J., Blank, K., Kowalewski, M., and Marshak, A.: Raw EPIC Data Calibration, Frontiers in Remote Sensing, 2, 1–18, https://doi.org/10.3389/frsen.2021.702275, 2021.

Chance, K. V.: Spectroscopic Measurements of Tropospheric Composition from Satellite Measurements in the Ultraviolet and Visible: Steps Toward Continuous Pollution Monitoring from Space, in: Remote Sensing of the Atmosphere for Environmental Security, edited by Perrin, A., Ben Sari-Zizi, N., and Demaison, J., pp. 1–25, Springer Netherlands, Dordrecht, 2006.

Chance, K. V. and Spurr, R. J. D.: Ring effect studies: Rayleigh scattering, including molecular parameters for rotational Raman scattering, and the Fraunhofer spectrum, Applied Optics, 36, 5224, https://doi.org/10.1364/AO.36.005224, 1997.

Chang, X., Wang, S., Zhao, B., Xing, J., Liu, X., Wei, L., Song, Y., Wu, W., Cai, S., Zheng, H., Ding, D., and Zheng, M.: Contributions of inter-city and regional transport to PM 2.5 concentrations in the Beijing-Tianjin-Hebei region and its implications on regional joint air pollution control, Science of the Total Environment, 660, 1191–1200, https://doi.org/10.1016/j.scitotenv.2018.12.474, 2019.

Chauhan, A., Inskip, H. M., Linaker, C. H., Smith, S., Schreiber, J., Johnston, S. L., and Holgate, S. T.: Personal exposure to nitrogen dioxide (NO2) and the severity of virus-induced asthma in children, The Lancet, 361, 1939–1944, https://doi.org/https://doi.org/10.1016/S0140-6736(03)13582-9, 2003.

Chimot, J., Vlemmix, T., Veefkind, J. P., de Haan, J. F., and Levelt, P. F.: 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, 359–382, https://doi.org/10.5194/amt-9-359-2016, 2016.

Conrad, R.: Soil microorganisms as controllers of atmospheric trace gases (H_2, CO, CH_4, OCS, N_2O, and NO), Microbiological Reviews, 60, 609–640, https://doi.org/10.1128/mr.60.4.609-640.1996, 1996.

Coulson, K. L. and Reynolds, D. W.: The Spectral Reflectance of Natural Surfaces, Journal of Applied Meteorology (1962-1982), 10, 1285–1295, http://www.jstor.org/stable/26175651, 1971.

Cox, C. and Munk, W.: Measurement of the Roughness of the Sea Surface from Photographs of the Sun's Glitter, J. Opt. Soc. Am., 44, 838–850, https://doi.org/10.1364/JOSA.44.000838, 1954a.

Cox, C. and Munk, W.: Statistics of the sea surface derived from Sun glitter, Journal of Marine Research, 13, 198–227, n/a, 1954b.

Daumont, D., Brion, J., Charbonnier, J., and Malicet, J.: Ozone UV spectroscopy I: Absorption cross-sections at room temperature, Journal of Atmospheric Chemistry, 15, 145–155, https://doi.org/10.1007/BF00053756, 1992.

Dave, J. V.: Meaning of Successive Iteration of the Auxiliary Equation in the Theory of Radiative Transfer., The Astrophysical Journal, 140, 1292, https://doi.org/10.1086/148024, 1964.

Davis, S. M., Hegglin, M. I., Fujiwara, M., Dragani, R., Harada, Y., Kobayashi, C., Long, C., Manney, G. L., Nash, E. R., Potter, G. L., Tegtmeier, S., Wang, T., Wargan, K., and Wright, J. S.: Assessment of upper tropospheric and stratospheric water vapor and ozone in reanalyses as part of S-RIP, Atmospheric Chemistry and Physics, 17, 12 743–12 778, https://doi.org/10.5194/acp-17-12743-2017, 2017.

Deirmendjian, D.: Electromagnetic scattering on spherical polydispersions, New York : American Elsevier Pub. Co., https://www.rand.org/pubs/reports/R0456.html, 1969.

Dickerson, R. R., Li, C., Li, Z., Marufu, L. T., Stehr, J. W., McClure, B., Krotkov, N., Chen, H., Wang, P., Xia, X., Ban, X., Gong, F., Yuan, J., and Yang, J.: Aircraft observations of dust and pollutants over northeast China: Insight into the meteorological mechanisms of transport, Journal of Geophysical Research: Atmospheres, 112, https://doi.org/https://doi.org/10.1029/2007JD008999, 2007.

Doda, D. D. and Green, A. E. S.: Surface reflectance measurements in the UV from an airborne platform Part 1, Applied Optics, 19, 2140, https://doi.org/10.1364/AO.19.002140, 1980.

Doda, D. D. and Green, A. E. S.: Surface reflectance measurements in the ultraviolet from an airborne platform Part 2, Applied Optics, 20, 636, https://doi.org/10.1364/AO.20.000636, 1981.

Eck, T. F., Bhartia, P. K., Hwang, P. H., and Stowe, L. L.: Reflectivity of Earth's surface and clouds in ultraviolet from satellite observations, Journal of Geophysical Research, 92, 4287, https://doi.org/10.1029/JD092iD04p04287, 1987.

Edlén, B.: The Refractive Index of Air, Metrologia, 2, 71–80, https://doi.org/10.1088/0026-1394/2/2/002, 1966.

Eskes, H. J., van der A, R. J., Brinksma, E. J., Veefkind, J. P., de Haan, J. F., and Valks, P. J. M.: Retrieval and validation of ozone columns derived from measurements of SCIAMACHY on Envisat, Atmospheric Chemistry and Physics Discussions, 5, 4429–4475, https://doi.org/10.5194/acpd-5-4429-2005, 2005.

F. Wang, F., Li, Z., Ren, X., Jiang, Q., He, H., Dickerson, R. R., Dong, X., and Lv, F.: Vertical distributions of aerosol optical properties during the spring 2016 ARIAs airborne campaign in the North China Plain, Atmospheric Chemistry and Physics, 18, 8995–9010, https://doi.org/10.5194/acp-18-8995-2018, 2018.

Farman, J. C., Gardiner, B. G., and Shanklin, J. D.: Large losses of total ozone in Antarctica reveal seasonal ClO_x/NO_x interaction, Nature, 315, 207–210, https://doi.org/10.1038/315207a0, 1985.

Feister, U. and Grewe, R.: SPECTRAL ALBEDO MEASUREMENTS IN THE UV and VISIBLE REGION OVER DIFFERENT TYPES OF SURFACES, Photochemistry and Photobiology, 62, 736–744, https://doi.org/10.1111/j.1751-1097.1995.tb08723.x, 1995.

Flynn, L. E., Long, C., Wu, X., Evans, R., Beck, C. T., Petropavlovskikh, I., McConville, G., Yu, W., Zhang, Z., Niu, J., Beach, E., Hao, Y., Pan, C., Sen, B., Novicki, M., Zhou, S., and Seftor, C. J.: Performance of the Ozone Mapping and Profiler Suite (OMPS) products, Journal of Geophysical Research Atmospheres, 119, 6181–6195, https://doi.org/10.1002/2013JD020467, 2014.

Flynn, L. E., Zhang, Z., Mikles, V., Das, B., Niu, J., Beck, T. C., and Beach, E.: Algorithm Theoretical Basis Document for NOAA NDE OMPS Version 8 Total Column Ozone (V8TOz) Environmental Data Record (EDR) Version 1.0, ATBD, https://www.star.nesdis.noaa.gov/jpss/documents/ATBD/ATBD_OMPS_TC_V8TOz_v1.1.pdf, 2016.

Fortuin, J. P. F. and Kelder, H.: An ozone climatology based on ozonesonde and satellite measurements, Journal of Geophysical Research: Atmospheres, 103, 31 709–31 734, https://doi.org/10.1029/1998JD200008, 1998.

Fu, X., Wang, S., Zhao, B., Xing, J., Cheng, Z., Liu, H., and Hao, J.: Emission inventory of primary pollutants and chemical speciation in 2010 for the Yangtze River Delta region, China, Atmospheric Environment, 70, 39–50, https://doi.org/10.1016/j.atmosenv.2012.12.034, 2013.

Gassó, S.: Satellite observations of the impact of weak volcanic activity on marine clouds, Journal of Geophysical Research: Atmospheres, 113, https://doi.org/https://doi.org/10.1029/2007JD009106, 2008.

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C. A., Darmenov, A., Bosilovich, M. G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., da Silva, A. M., Gu, W., Kim, G. K., Koster, R., Lucchesi, R., Merkova, D., Nielsen, J. E., Partyka, G., Pawson, S., Putman, W., Rienecker, M., Schubert, S. D., Sienkiewicz, M., and Zhao, B.: The modern-era retrospective analysis for research and applications, version 2 (MERRA-2), Journal of Climate, 30, 5419–5454, https://doi.org/10.1175/JCLI-D-16-0758.1, 2017.

Goldberg, D. L., Anenberg, S. C., Griffin, D., McLinden, C. A., Lu, Z., and Streets, D. G.: Disentangling the Impact of the COVID-19 Lockdowns on Urban NO_2 From Natural Variability, Geophysical Research Letters, 47, https://doi.org/10.1029/2020GL089269, 2020.

Grainger, J. F. and Ring, J.: Anomalous Fraunhofer Line Profiles, Nature, 193, 762–762, https://doi.org/10.1038/193762a0, 1962.

He, H., Ren, X., Benish, S., Li, Z., Wang, F., Wang, Y., Canty, T., Dong, X., Lv, F., Hu, Y., Zhu, T., and Dickerson, R.: Evaluation of Anthropogenic Emissions and Ozone Pollution in the North China Plain: Insights from the Air Chemistry Research in Asia (ARIAs) Campaign, Evaluation of Anthropogenic Emissions and Ozone Pollution in the North China Plain: Insights from the Air Chemistry Research in Asia (ARIAs) Campaign, pp. 1–45, https://doi.org/10.5194/acp-2019-248, 2019.

Henyey, L. G. and Greenstein, J. L.: Diffuse radiation in the Galaxy., Astrophysical Journal, 93, 70–83, https://doi.org/10.1086/144246, 1941.

Herman, J., Huang, L., McPeters, R., Ziemke, J., Cede, A., and Blank, K.: Synoptic Ozone, Cloud Reflectivity, and Erythemal Irradiance from Sunrise to Sunset for the Whole Earth as viewed by the DSCOVR spacecraft from Lagrange-1, Atmospheric Measurement Techniques Discussions, pp. 1–48, https://doi.org/10.5194/amt-2017-155, 2017.

Herman, J., Huang, L., McPeters, R., Ziemke, J., Cede, A., and Blank, K.: Synoptic ozone, cloud reflectivity, and erythemal irradiance from sunrise to sunset for the whole earth as viewed by the DSCOVR spacecraft from the earth–sun Lagrange 1 orbit, Atmospheric Measurement Techniques, 11, 177–194, https://doi.org/10.5194/amt-11-177-2018, 2018.

Herman, J., Abuhassan, N., Kim, J., Kim, J., Dubey, M., Raponi, M., and Tzortziou, M.: Underestimation of column NO2 amounts from the OMI satellite compared to diurnally varying ground-based retrievals from multiple PANDORA spectrometer instruments, Atmospheric Measurement Techniques, 12, 5593–5612, https://doi.org/10.5194/amt-12-5593-2019, 2019.

Herman, J. R. and Celarier, E. a.: Earth surface reflectivity climatology at 340-380 nm from TOMS data, Journal of Geophysical Research: Atmospheres, 102, 28 003–28 011, https://doi.org/10.1029/97JD02074, 1997.

Herman, J. R., Bhartia, P. K., Torres, O., Hsu, C., Seftor, C., and Celarier, E.: Global distribution of UV-absorbing aerosols from Nimbus 7/TOMS data, Journal of Geophysical Research: Atmospheres, 102, 16 911–16 922, https://doi.org/10.1029/96JD03680, 1997.

Huang, X., Yang, K., Kondragunta, S., Wei, Z., Valin, L., Szykman, J., and Goldberg, M.: NO_2 retrievals from NOAA-20 OMPS: Algorithm, evaluation, and observations of drastic changes during COVID-19, Atmospheric Environment, 290, 119 367, https://doi.org/https://doi.org/10.1016/j.atmosenv.2022.119367, 2022.

Ialongo, I., Herman, J., Krotkov, N., Lamsal, L., Folkert Boersma, K., Hovila, J., and Tamminen, J.: Comparison of OMI NO_2 observations and their seasonal and weekly cycles with ground-based measurements in Helsinki, Atmospheric Measurement Techniques, 9, 5203–5212, https://doi.org/10.5194/amt-9-5203-2016, 2016.

Ialongo, I., Virta, H., Eskes, H., Hovila, J., and Douros, J.: Comparison of TROPOMI/Sentinel-5 Precursor NO_2 observations with ground-based measurements in Helsinki, Atmospheric Measurement Techniques, 13, 205–218, https://doi.org/10.5194/amt-13-205-2020, 2020.

Jacob, D. J., Logan, J. a., Gardner, G. M., Yevich, R. M., Spivakovsky, C. M., Wofsy, S. C., Sillman, S., and Prather, M. J.: Factors regulating ozone over the United States and its export to the global atmosphere, 98, 817–826, https://doi.org/10.1029/98JD01224, 1993.

Joiner, J. and Bhartia, P. K.: Accurate determination of total ozone using SBUV continuous spectral scan measurements, Journal of Geophysical Research: Atmospheres, 102, 12 957–12 969, https://doi.org/10.1029/97JD00902, 1997.

Joiner, J. and Vasilkov, A. P.: First results from the OMI rotational Raman scattering cloud pressure algorithm, IEEE Transactions on Geoscience and Remote Sensing, 44, 1272–1282, https://doi.org/10.1109/TGRS.2005.861385, 2006.

Joiner, J., Bhartia, P. K., Cebula, R. P., Hilsenrath, E., McPeters, R. D., and Park, H.: Rotational Raman scattering (Ring effect) in satellite backscatter ultraviolet measurements, Applied Optics, 34, 4513, https://doi.org/10.1364/AO.34.004513, 1995.

Judd, L. M., Al-Saadi, J. A., Janz, S. J., Kowalewski, M. G., Pierce, R. B., Szykman, J. J., Valin, L. C., Swap, R., Cede, A., Mueller, M., Tiefengraber, M., Abuhassan, N., and Williams, D.: Evaluating the impact of spatial resolution on tropospheric NO_2 column comparisons within urban areas using high-resolution airborne data, Atmospheric Measurement Techniques, 12, 6091–6111, https://doi.org/10.5194/amt-12-6091-2019, 2019.

Judd, L. M., Al-Saadi, J. A., Szykman, J. J., Valin, L. C., Janz, S. J., Kowalewski, M. G., Eskes, H. J., Veefkind, J. P., Cede, A., Mueller, M., Gebetsberger, M., Swap, R., Pierce, R. B., Nowlan, C. R., Abad, G. G., Nehrir, A., and Williams, D.: Evaluating Sentinel-5P TROPOMI tropospheric NO_2 column densities with airborne and Pandora spectrometers near New York City and Long Island Sound, Atmospheric Measurement Techniques, 13, 6113–6140, https://doi.org/10.5194/amt-13-6113-2020, 2020.

Kanaya, Y., Irie, H., Takashima, H., Iwabuchi, H., Akimoto, H., Sudo, K., Gu, M., Chong, J., Kim, Y. J., Lee, H., Li, A., Si, F., Xu, J., Xie, P.-H., Liu, W.-Q., Dzhola, A., Postylyakov, O., Ivanov, V., Grechko, E., Terpugova, S., and Panchenko, M.: Long-term MAX-DOAS network observations of NO_2 in Russia and Asia (MADRAS) during the period 2007ndash;2012: instrumentation, elucidation of climatology, and comparisons with OMI satellite observations and global model simulations, Atmospheric Chemistry and Physics, 14, 7909–7927, https://doi.org/10.5194/acp-14-7909-2014, 2014.

Kim, H. C., Lee, P., Judd, L., Pan, L., and Lefer, B.: OMI NO_2 column densities over North American urban cities:
the effect of satellite footprint resolution, Geoscientific Model Development, 9, 1111–1123, https://doi.org/10.5194/gmd-9-1111-2016, 2016.

King, M. D., Platnick, S., Menzel, W. P., Ackerman, S. A., and Hubanks, P. A.: Spatial and Temporal Distribution of Clouds Observed by MODIS Onboard the Terra and Aqua Satellites, IEEE Transactions on Geoscience and Remote Sensing, 51, 3826–3852, https://doi.org/10.1109/TGRS.2012.2227333, 2013.

Kleipool, Q. L., Dobber, M. R., de Haan, J. F., and Levelt, P. F.: Earth surface reflectance climatology from 3 years of OMI data, Journal of Geophysical Research, 113, D18 308, https://doi.org/10.1029/2008JD010290, 2008.

Koelemeijer, R. B. A.: A database of spectral surface reflectivity in the range 335–772 nm derived from 5.5 years of GOME observations, Journal of Geophysical Research, 108, 4070, https://doi.org/10.1029/2002JD002429, 2003.

Koelemeijer, R. B. A. and Stammes, P.: Effects of clouds on ozone column retrieval from GOME UV measurements, Journal of Geophysical Research, 104, 8281 – 8294, https://doi.org/10.1029/1999JD900012, 1999.

Labow, G. J., Ziemke, J. R., McPeters, R. D., Haffner, D. P., and Bhartia, P. K.: A total ozone-dependent ozone profile climatology based on ozonesondes and Aura MLS data, Journal of Geophysical Research: Atmospheres, 120, 2537–2545, https://doi.org/10.1002/2014JD022634, 2015.

Lamsal, L. N., Weber, M., Tellmann, S., and Burrows, J. P.: Ozone column classified climatology of ozone and temperature profiles based on ozonesonde and satellite data, Journal of Geophysical Research D: Atmospheres, 109, 1–15, https://doi.org/10.1029/2004JD004680, 2004.

Lamsal, L. N., Krotkov, N. A., Celarier, E. A., Swartz, W. H., Pickering, K. E., Bucsela, E. J., Gleason, J. F., Martin, R. V., Philip, S., Irie, H., Cede, A., Herman, J., Weinheimer, A., Szykman, J. J., and Knepp, T. N.: Evaluation of OMI operational standard NO_2 column retrievals using in situ and surface-based NO_2 observations, Atmospheric Chemistry and Physics, 14, 11 587–11 609, https://doi.org/10.5194/acp-14-11587-2014, 2014.

Landgraf, J., Hasekamp, O., van Deelen, R., and Aben, I.: Rotational Raman scattering of polarized light in the Earth atmosphere: a vector radiative transfer model using the radiative transfer perturbation theory approach, Journal of Quantitative Spectroscopy and Radiative Transfer, 87, 399–433, https://doi.org/10.1016/j.jqsrt.2004.03.013, 2004.

Laughner, J. L. and Cohen, R. C.: Direct observation of changing NO_x lifetime in North American cities, Science, 366, 723–727, https://doi.org/10.1126/science.aax6832, 2019.

Lelieveld, J., Evans, J. S., Fnais, M., Giannadaki, D., and Pozzer, A.: The contribution of outdoor air pollution sources to premature mortality on a global scale, Nature, 525, 367–371, https://doi.org/10.1038/nature15371, 2015.

Lerot, C., Van Roozendael, M., Lambert, J.-C., Granville, J., van Gent, J., Loyola, D., and Spurr, R. J. D.: The GODFIT algorithm: a direct fitting approach to improve the accuracy of total ozone measurements from GOME, International Journal of Remote Sensing, 31, 543–550, https://doi.org/10.1080/01431160902893576, 2010.

Lerot, C., Van Roozendael, M., Spurr, R. J. D., Loyola, D., Coldewey-Egbers, M. R., Kochenova, S., Van Gent, J., Koukouli, M., Balis, D., Lambert, J. C., Granville, J., and Zehner, C.: Homogenized total ozone data records from the European sensors GOME/ERS-2, SCIAMACHY/Envisat, and GOME-2/MetOp-A, Journal of Geophysical Research: Atmospheres, 119, 1639–1662, https://doi.org/10.1002/2013JD020831, 2014.

Levelt, P. F., Hilsenrath, E., Leppelmeier, G. W., van den Oord, G. H. J., Bhartia, P. K., Tamminen, J., de Haan, J. F., and Veefkind, J. P.: Science objectives of the ozone monitoring instrument, IEEE Transactions on Geoscience and Remote Sensing, 44, 1199–1208, https://doi.org/10.1109/TGRS.2006.872336, 2006.

Li, C., McLinden, C., Fioletov, V., Krotkov, N., Carn, S., Joiner, J., Streets, D., He, H., Ren, X., Li, Z., et al.: India is overtaking China as the world's largest emitter of anthropogenic sulfur dioxide, Scientific reports, 7, 14 304, 2017a.

Li, M., Zhang, Q., Kurokawa, J.-I., Woo, J.-H., He, K., Lu, Z., Ohara, T., Song, Y., Streets, D. G., Carmichael, G. R., Cheng, Y., Hong, C., Huo, H., Jiang, X., Kang, S., Liu, F., Su, H., and Zheng, B.: MIX: a mosaic Asian anthropogenic emission inventory under the international collaboration framework of the MICS-Asia and HTAP, Atmos. Chem. Phys., 17, 935–963, https://doi.org/10.5194/acp-17-935-2017, 2017b.

Lin, J.-T., Martin, R. V., Boersma, K. F., Sneep, M., Stammes, P., Spurr, R., Wang, P., Van Roozendael, M., Clémer, K., and Irie, H.: 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, 1441–1461, https://doi.org/10.5194/acp-14-1441-2014, 2014.

Liu, F., Page, A., Strode, S. A., Yoshida, Y., Choi, S., Zheng, B., Lamsal, L. N., Li, C., Krotkov, N. A., Eskes, H., Ronald van der, A., Veefkind, P., Levelt, P. F., Hauser, O. P., and Joiner, J.: Abrupt decline in tropospheric nitrogen dioxide over China after the outbreak of COVID-19, Science Advances, 6, 1–10, https://doi.org/10.1126/sciadv.abc2992, 2020.

Liu, G., Tarasick, D. W., Fioletov, V. E., Sioris, C. E., and Rochon, Y. J.: Ozone correlation lengths and measurement uncertainties from analysis of historical ozonesonde data in North America and Europe, Journal of Geophysical Research, 114, D04 112–D04 112, https://doi.org/10.1029/2008JD010576, 2009.

Lobell, D. B., Tommaso, S. D., and Burney, J. A.: Globally ubiquitous negative effects of nitrogen dioxide on crop growth, Science Advances, 8, eabm9909, https://doi.org/10.1126/sciadv.abm9909, 2022.

Loyola, D. G., Koukouli, M. E., Valks, P., Balis, D. S., Hao, N., Van Roozendael, M., Spurr, R. J. D., Zimmer, W., Kiemle, S., Lerot, C., and Lambert, J.-C.: The GOME-2 total column ozone product: Retrieval algorithm and ground-based validation, Journal of Geophysical Research, 116, D07 302, https://doi.org/10.1029/2010JD014675, 2011.

Lucht, W., Schaaf, C., and Strahler, A.: An algorithm for the retrieval of albedo from space using semiempirical BRDF models, IEEE Transactions on Geoscience and Remote Sensing, 38, 977–998, https://doi.org/10.1109/36.841980, 2000.

Malicet, J., Daumont, D., Charbonnier, J., Parisse, C., Chakir, A., and Brion, J.: Ozone UV spectroscopy. II. Absorption cross-sections and temperature dependence, Journal of Atmospheric Chemistry, 21, 263–273, https://doi.org/10.1007/BF00696758, 1995.

McPeters, R. D. and Labow, G. J.: Climatology 2011: An MLS and sonde derived ozone climatology for satellite retrieval algorithms, Journal of Geophysical Research Atmospheres, 117, 1–8, https://doi.org/10.1029/2011JD017006, 2012.

McPeters, R. D., Labow, G. J., and Logan, J. A.: Ozone climatological profiles for satellite retrieval algorithms, Journal of Geophysical Research Atmospheres, 112, 1–9, https://doi.org/10.1029/2005JD006823, 2007.

Munro, R., Lang, R., Klaes, D., Poli, G., Retscher, C., Lindstrot, R., Huckle, R., Lacan, A., Grzegorski, M., Holdak, A., Kokhanovsky, A., Livschitz, J., and Eisinger, M.: The GOME-2 instrument on the Metop series of satellites: Instrument design, calibration, and level 1 data processing - An overview, Atmospheric Measurement Techniques, 9, 1279–1301, https://doi.org/10.5194/amt-9-1279-2016, 2016.

Myllyvirta, L.: Analysis: Coronavirus temporarily reduced China's CO_2 emissions by a quarter, https://www.carbonbrief.org/analysis-coronavirus-has-temporarily-reduced-chinas-co2-emissions-by-a-quarter, 2020.

NASA/LARC/SD/ASDC: DSCOVR EPIC L2 Ozone (O_3), Sulfur Dioxide (SO_2) Aerosol Index (AI) with Epic L1B V03 Input, Version 2, https://doi.org/10.5067/EPIC/DSCOVR/L2_O3SO2AI.002, 2018.

Platt, U.: Air Monitoring by Differential Optical Absorption Spectroscopy, in: Encyclopedia of Analytical Chemistry, iii, pp. 1–28, John Wiley & Sons, Ltd, Chichester, UK, https://doi.org/10.1002/9780470027318.a0706.pub2, 2017.

Qu, Z., Jacob, D. J., Silvern, R. F., Shah, V., Campbell, P. C., Valin, L. C., and Murray, L. T.: US COVID-19 Shutdown Demonstrates Importance of Background NO_2 in Inferring NO_x Emissions From Satellite NO_2 Observations, Geophysical Research Letters, 48, 1–8, https://doi.org/10.1029/2021GL092783, 2021.

Ridley, B. A., Dye, J. E., Walega, J. G., Zheng, J., Grahek, F. E., and Rison, W.: On the production of active nitrogen by thunderstorms over New Mexico, Journal of Geophysical Research: Atmospheres, 101, 20 985–21 005, https://doi.org/https://doi.org/10.1029/96JD01706, 1996.

Riemer, N., Vogel, H., Vogel, B., Schell, B., Ackermann, I., Kessler, C., and Hass, H.: Impact of the heterogeneous hydrolysis of N2O5 on chemistry and nitrate aerosol formation in the lower troposphere under photosmog conditions, Journal of Geophysical Research: Atmospheres, 108, https://doi.org/https://doi.org/10.1029/2002JD002436, 2003.

Rivas, M. B., Veefkind, P., Boersma, F., Levelt, P., Eskes, H., and Gille, J.: Intercomparison of daytime stratospheric NO_2 satellite retrievals and model simulations, Atmospheric Measurement Techniques, 7, 2203–2225, https://doi.org/10.5194/amt-7-2203-2014, 2014.

Robock, A.: Volcanic eruptions and climate, Reviews of Geophysics, 38, 191–219, https://doi.org/https://doi.org/10.1029/1998RG000054, 2000.

Rodgers, C. D.: Inverse Methods for Atmospheric Sounding, WORLD SCIENTIFIC, https://doi.org/10.1142/3171, 2000.

S. X. Wang, S. X., Zhao, B., Cai, S. Y., Klimont, Z., Nielsen, C. P., Morikawa, T., Woo, J. H., Kim, Y., Fu, X., Xu, J. Y., Hao, J. M., and He, K. B.: Emission trends and mitigation options for air pollutants in East Asia, Atmospheric Chemistry and Physics, 14, 6571–6603, https://doi.org/10.5194/acp-14-6571-2014, 2014.

Sandermann Jr, H.: OZONE AND PLANT HEALTH, Annual Review of Phytopathology, 34, 347–366, https://doi.org/10.1146/annurev.phyto.34.1.347, pMID: 15012547, 1996.

Schaaf, C. B., Liu, J., Gao, F., and Strahler, A. H.: Aqua and Terra MODIS Albedo and Reflectance Anisotropy Products, pp. 549–561, Springer New York, New York, NY, https://doi.org/10.1007/978-1-4419-6749-7_24, 2011.

Schuman, R.: INRIX U.S. National Traffic Volume Synopsis, https://inrix.com/blog/covid19-us-traffic-volume-synopsis/https://inrix.com/blog/covid19-us-traffic-volume-synopsis-2/https://inrix.com/blog/covid19-us-traffic-volume-synopsis-3/https://inrix.com/blog/covid19-us-traffic-volume-synopsis-4/, 2020.

Seinfeld, J. H. and Pandis, S. N.: Atmospheric Chemistry and Physics : From Air Pollution to Climate Change Third ed., Hoboken New Jersey: John Wiley & Sons, 2016.

Shah, V., J. Jacob, D., Li, K., Silvern, R., Zhai, S., Liu, M., Lin, J., and Zhang, Q.: Effect of changing NO_x lifetime on the seasonality and long-term trends of satellite-observed tropospheric NO_2 columns over China, Atmospheric Chemistry and Physics, 20, 1483–1495, https://doi.org/10.5194/acp-20-1483-2020, 2020.

Sillman, S.: The relation between ozone, NOx and hydrocarbons in urban and polluted rural environments, Atmospheric Environment, 33, 1821–1845, https://doi.org/https://doi.org/10.1016/S1352-2310(98)00345-8, 1999.

Silvern, R. F., Jacob, D. J., Mickley, L. J., Sulprizio, M. P., Travis, K. R., Marais, E. A., Cohen, R. C., Laughner, J. L., Choi, S., and Joiner, J.: Using satellite observations of tropospheric NO_2 columns to infer long-term trends in US NO_x emissions: the importance of accounting for the free tropospheric NO_2 background, 19, 1–16, https://doi.org/10.5194/acp-19-8863-2019, 2019.

Sofieva, V. F., Tamminen, J., Kyrölä, E., Mielonen, T., Veefkind, P., Hassler, B., and Bodeker, G. E.: A novel tropopause-related climatology of ozone profiles, Atmospheric Chemistry and Physics, 14, 283–299, https://doi.org/10.5194/acp-14-283-2014, 2014.

Solomon, S. and Garcia, R. R.: On the distribution of nitrogen dioxide in the high-latitude stratosphere, Journal of Geophysical Research: Oceans, 88, 5229–5239, https://doi.org/10.1029/JC088IC09P05229, 1983.

Spurr, R. J.: 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, 316–342, https://doi.org/10.1016/j.jqsrt.2006.05.005, 2006.

Spurr, R. J. D., de Haan, J., van Oss, R., and Vasilkov, A. P.: Discrete-ordinate radiative transfer in a stratified medium with first-order rotational Raman scattering, Journal of Quantitative Spectroscopy and Radiative Transfer, 109, 404–425, https://doi.org/10.1016/j.jqsrt.2007.08.011, 2008.

Stajner, I., Wargan, K., Pawson, S., Hayashi, H., Chang, L.-P., Hudman, R. C., Froidevaux, L., Livesey, N., Levelt, P. F., Thompson, A. M., Tarasick, D. W., Stübi, R., Andersen, S. B., Yela, M., König-Langlo, G., Schmidlin, F. J., and Witte, J. C.: Assimilated ozone from EOS-Aura: Evaluation of the tropopause region and tropospheric columns, Journal of Geophysical Research, 113, D16S32, https://doi.org/10.1029/2007JD008863, 2008.

Stenchikov, G., Ukhov, A., Osipov, S., Ahmadov, R., Grell, G., Cady-Pereira, K., Mlawer, E., and Iacono, M.: How Does a Pinatubo-Size Volcanic Cloud Reach the Middle Stratosphere?, Journal of Geophysical Research: Atmospheres, 126, e2020JD033 829, https://doi.org/https://doi.org/10.1029/2020JD033829, e2020JD033829 2020JD033829, 2021.

Tilstra, L. G., Tuinder, O. N. E., Wang, P., and Stammes, P.: Surface reflectivity climatologies from UV to NIR determined from Earth observations by GOME-2 and SCIAMACHY, Journal of Geophysical Research: Atmospheres, 122, 4084–4111, https://doi.org/10.1002/2016JD025940, 2017.

Torres, O., Bhartia, P. K., Herman, J. R., Ahmad, Z., and Gleason, J.: Derivation of aerosol properties from satellite measurements of backscattered ultraviolet radiation: Theoretical basis, Journal of Geophysical Research Atmospheres, 103, 17 099–17 110, https://doi.org/10.1029/98JD00900, 1998.

Torres, O., Tanskanen, A., Veihelmann, B., Ahn, C., Braak, R., Bhartia, P. K., Veefkind, P., and Levelt, P.: Aerosols and surface UV products from Ozone Monitoring Instrument observations: An overview, Journal of Geophysical Research: Atmospheres, 112, https://doi.org/https://doi.org/10.1029/2007JD008809, 2007.

Val Martín, M., Honrath, R. E., Owen, R. C., Pfister, G., Fialho, P., and Barata, F.: Significant enhancements of nitrogen oxides, black carbon, and ozone in the North Atlantic lower free troposphere resulting from North American boreal wildfires, Journal of Geophysical Research: Atmospheres, 111, https://doi.org/https://doi.org/10.1029/2006JD007530, 2006.

Valin, L. C., Russell, A. R., Bucsela, E. J., Veefkind, J. P., and Cohen, R. C.: Observation of slant column NO_2 using the super-zoom mode of AURA-OMI, Atmospheric Measurement Techniques, 4, 1929–1935, https://doi.org/10.5194/amt-4-1929-2011, 2011.

Valin, L. C., Russell, A. R., and Cohen, R. C.: Variations of OH radical in an urban plume inferred from NO_2 column measurements, Geophysical Research Letters, 40, 1856–1860, https://doi.org/10.1002/GRL.50267, 2013.

Van Roozendael, M., Loyola, D., Spurr, R., Balis, D., Lambert, J.-C., Livschitz, Y., Valks, P., Ruppert, T., Kenter, P., Fayt, C., and Zehner, C.: Ten years of GOME/ERS-2 total ozone data—The new GOME data processor (GDP) version 4: 1. Algorithm description, Journal of Geophysical Research, 111, D14 311, https://doi.org/10.1029/2005JD006375, 2006.

Van Roozendael, M., Spurr, R. J. D., Loyola, D., Lerot, C., Balis, D., Lambert, J. C., Zimmer, W., Van Gent, J., Van Geffen, J., Koukouli, M., Granville, J., Doicu, A., Fayt, C., and Zehner, C.: Sixteen years of GOME/ERS-2 total ozone data: The new direct-fitting GOME Data Processor (GDP) version 5Algorithm description, Journal of Geophysical Research Atmospheres, 117, 1–18, https://doi.org/10.1029/2011JD016471, 2012.

van Vuuren, D. P., Bouwman, L. F., Smith, S. J., and Dentener, F.: Global projections for anthropogenic reactive nitrogen emissions to the atmosphere: an assessment of scenarios in the scientific literature, Current Opinion in Environmental Sustainability, 3, 359–369, https://doi.org/https://doi.org/10.1016/j.cosust.2011.08.014, carbon and nitrogen cycles, 2011.

Vasilkov, A. P., Joiner, J., Spurr, R. J. D., Bhartia, P. K., Levelt, P., and Stephens, G.: Evaluation of the OMI cloud pressures derived from rotational Raman scattering by comparisons with other satellite data and radiative transfer simulations, Journal of Geophysical Research, 113, 1–12, https://doi.org/10.1029/2007JD008689, 2008.

Veefkind, J., de Haan, J., Brinksma, E., Kroon, M., and Levelt, P.: Total ozone from the ozone monitoring instrument (OMI) using the DOAS technique, IEEE Transactions on Geoscience and Remote Sensing, 44, 1239–1244, https://doi.org/10.1109/TGRS.2006.871204, 2006.

Veefkind, J. P., Aben, I., McMullan, K., Förster, H., de Vries, J., Otter, G., Claas, J., Eskes, H. J., de Haan, J. F., Kleipool, Q., van Weele, M., Hasekamp, O., Hoogeveen, R., Landgraf, J., Snel, R., Tol, P., Ingmann, P., Voors, R., Kruizinga, B., Vink, R., Visser, H., and Levelt, P. F.: TROPOMI on the ESA Sentinel-5 Precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications, Remote Sensing of Environment, 120, 70–83, https://doi.org/10.1016/j.rse.2011.09.027, 2012.

Vernier, J. P., Pommereau, J. P., Garnier, A., Pelon, J., Larsen, N., Nielsen, J., Christensen, T., Cairo, F., Thomason, L. W., Leblanc, T., and McDermid, I. S.: Tropical stratospheric aerosol layer from CALIPSO lidar observations, Journal of Geophysical Research: Atmospheres, 114, https://doi.org/https://doi.org/10.1029/2009JD011946, 2009.

Vlemmix, T., Piters, A. J. M., Stammes, P., Wang, P., and Levelt, P. F.: Retrieval of tropospheric NO_2 using the MAX-DOAS method combined with relative intensity measurements for aerosol correction, Atmospheric Measurement Techniques, 3, 1287–1305, https://doi.org/10.5194/amt-3-1287-2010, 2010.

Vlemmix, T., Piters, A. J. M., Berkhout, A. J. C., Gast, L. F. L., Wang, P., and Levelt, P. F.: Ability of the MAX-DOAS method to derive profile information for NO_2: can the boundary layer and free troposphere be separated?, Atmospheric Measurement Techniques, 4, 2659–2684, https://doi.org/10.5194/amt-4-2659-2011, 2011.

Vountas, M., Rozanov, V., and Burrows, J.: RING EFFECT: IMPACT OF ROTATIONAL RAMAN SCATTERING ON RADIATIVE TRANSFER IN EARTH'S ATMOSPHERE, Journal of Quantitative Spectroscopy and Radiative Transfer, 60, 943–961, https://doi.org/10.1016/S0022-4073(97)00186-6, 1998.

Wagner, T., Dix, B., Friedeburg, C. v., Frieß, U., Sanghavi, S., Sinreich, R., and Platt, U.: MAX-DOAS O_4 measurements: A new technique to derive information on atmospheric aerosols—Principles and information content, Journal of Geophysical Research: Atmospheres, 109, 2004.

Wagner, T., Beirle, S., Deutschmann, T., and Penning de Vries, M.: A sensitivity analysis of Ring effect to aerosol properties and comparison to satellite observations, Atmospheric Measurement Techniques, 3, 1723–1751, https://doi.org/10.5194/amt-3-1723-2010, 2010.

Wagner, T., Apituley, A., Beirle, S., Dörner, S., Friess, U., Remmers, J., and Shaiganfar, R.: Cloud detection and classification based on MAX-DOAS observations, Atmospheric Measurement Techniques, 7, 1289–1320, https://doi.org/10.5194/amt-7-1289-2014, 2014.

Wagner, T., Beirle, S., Remmers, J., Shaiganfar, R., and Wang, Y.: Absolute calibration of the colour index and O_4 absorption derived from Multi AXis (MAX-)DOAS measurements and their application to a standardised cloud classification algorithm, Atmospheric Measurement Techniques, 9, 4803–4823, https://doi.org/10.5194/amt-9-4803-2016, 2016.

Wang, Y., Penning de Vries, M., Xie, P. H., Beirle, S., Dörner, S., Remmers, J., Li, A., and Wagner, T.: Cloud and aerosol classification for 2.5 years of MAX-DOAS observations in Wuxi (China) and comparison to independent data sets, Atmospheric Measurement Techniques, 8, 5133–5156, https://doi.org/10.5194/amt-8-5133-2015, 2015.

Wang, Y., Beirle, S., Lampel, J., Koukouli, M., De Smedt, I., Theys, N., Li, A., Wu, D., Xie, P., Liu, C., Van Roozendael, M., Stavrakou, T., Müller, J.-F., and Wagner, T.: Validation of OMI, GOME-2A and GOME-2B tropospheric NO_2, SO_2 and HCHO products using MAX-DOAS observations from 2011 to 2014 in Wuxi, China: investigation of the effects of priori profiles and aerosols on the satellite products, Atmospheric Chemistry and Physics, 17, 5007–5033, https://doi.org/10.5194/acp-17-5007-2017, 2017a.

Wang, Y., Lampel, J., Xie, P., Beirle, S., Li, A., Wu, D., and Wagner, T.: Ground-based MAX-DOAS observations of tropospheric aerosols, NO_2, SO_2 and HCHO in Wuxi, China, from 2011 to 2014, Atmospheric Chemistry and Physics, 17, 2189–2215, https://doi.org/10.5194/acp-17-2189-2017, 2017b.

Wang, Y., Dörner, S., Donner, S., Böhnke, S., De Smedt, I., Dickerson, R. R., Dong, Z., He, H., Li, Z., Li, Z., Li, D., Liu, D., Ren, X., Theys, N., Wang, Y., Wang, Y., Wang, Z., Xu, H., Xu, J., and Wagner, T.: Vertical profiles of NO_2, SO_2, HONO, HCHO, CHOCHO and aerosols derived from MAX-DOAS measurements at a rural site in the central western North China Plain and their relation to emission sources and effects of regional transport, Atmospheric Chemistry and Physics, 19, 5417–5449, https://doi.org/10.5194/acp-19-5417-2019, 2019.

Wargan, K., Pawson, S., Olsen, M. A., Witte, J. C., Douglass, A. R., Ziemke, J. R., Strahan, S. E., and Nielsen, J. E.: The global structure of upper troposphere-lower stratosphere ozone in GEOS-5: A multiyear assimilation of EOS Aura data, Journal of Geophysical Research: Atmospheres, 120, 2013–2036, https://doi.org/10.1002/2014JD022493, 2015.

Wargan, K., Labow, G., Frith, S., Pawson, S., Livesey, N., and Partyka, G.: Evaluation of the Ozone Fields in NASA's MERRA-2 Reanalysis, Journal of Climate, 30, 2961–2988, https://doi.org/10.1175/JCLI-D-16-0699.1, 2017.

Wassmann, A., Borsdorff, T., Aan De Brugh, J. M. J., Hasekamp, O. P., Aben, I., and Landgraf, J.: The direct fitting approach for total ozone column retrievals: a sensitivity study on GOME-2/MetOp-A measurements, Atmospheric Measurement Techniques, 8, 4429–4451, https://doi.org/10.5194/amt-8-4429-2015, 2015.

Wei, J. C., Pan, L. L., Maddy, E., Pittman, J. V., Divarkarla, M., Xiong, X., and Barnet, C.: Ozone profile retrieval from an advanced infrared sounder: experiments with tropopause-based climatology and optimal estimation approach, Journal of Atmospheric and Oceanic Technology, 27, 1123–1139, https://doi.org/10.1175/2010JTECHA1384.1, 2010.

Wellemeyer, C. G., Taylor, S. L., Seftor, C. J., Mcpeters, R. D., and Bhartia, P. K.: A correction for total ozone mapping spectrometer profile shape errors at high latitude, Journal of Geophysical Research, 102, 9029, https://doi.org/10.1029/96JD03965, 1997.

Wennberg, P. O., Cohen, R. C., Stimpfle, R. M., Koplow, J. P., Anderson, J. G., Salawitch, R. J., Fahey, D. W., Woodbridge, E. L., Keim, E. R., Gao, R. S., Webster, C. R., May, R. D., Toohey, D. W., Avallone, L. M., Proffitt, M. H., Loewenstein, M., Podolske, J. R., Chan, K. R., and Wofsy, S. C.: Removal of Stratospheric O_3 by Radicals: In Situ Measurements of OH, HO_2, NO, NO_2, ClO, and BrO, Science, 266, 398–404, https://doi.org/10.1126/science.266.5184.398, 1994.

Wu, R., Liu, F., Tong, D., Zheng, Y., Lei, Y., Hong, C., Li, M., Liu, J., Zheng, B., Bo, Y., Chen, X., Li, X., and Zhang, Q.: Air quality and health benefits of China's emission control policies on coal-fired power plants during 2005–2020, Environmental Research Letters, 14, 094 016, https://doi.org/10.1088/1748-9326/ab3bae, 2019.

Yang, K. and Liu, X.: Ozone Profile Climatology for Remote Sensing Retrieval Algorithms, Atmospheric Measurement Techniques Discussions, pp. 1–39, https://doi.org/10.5194/amt-2019-116, 2019.

Yang, K., Fleig, A., Wolfe, R., and Nishihama, M.: MODIS band-to-band registration, in: IGARSS 2000. IEEE 2000 International Geoscience and Remote Sensing Symposium. Taking the Pulse of the Planet: The Role of Remote Sensing in Managing the Environment. Proceedings (Cat. No.00CH37120), vol. 2, pp. 887–889, IEEE, https://doi.org/10.1109/IGARSS.2000.861735, 2000.

Yang, K., Bhartia, P. K., Wellemeyer, C. G., Qin, W., Spurr, R. J. D., Veefkind, J. P., and de Haan, J. F.: Application of spectral fitting method to GOME and comparison with OMI-DOAS and TOMS-V8 total ozone, in: Proceedings of the XX Quadrennial Ozone Symposium, 1-8 June 2004, Kos, Greece, edited by Zerefos, C. S., pp. 510–511, International Ozone Commission, https://www.researchgate.net/publication/354387548_Application_of_Spectral_Fitting_Method_to_GOME_and_Comparison_With_OMI_DOAS_and_TOMS-V8_Total_Ozone, 2004.

Yang, K., Krotkov, N. A., Krueger, A. J., Carn, S. A., Bhartia, P. K., and Levelt, P. F.: Retrieval of large volcanic SO_2 columns from the Aura Ozone Monitoring Instrument: Comparison and limitations, Journal of Geophysical Research Atmospheres, 112, 1–14, https://doi.org/10.1029/2007JD008825, 2007.

Yang, K., Krotkov, N. A., Krueger, A. J., Carn, S. A., Bhartia, P. K., and Levelt, P. F.: Improving retrieval of volcanic sulfur dioxide from backscattered UV satellite observations, Geophysical Research Letters, 36, L03 102–L03 102, https://doi.org/10.1029/2008GL036036, 2009a.

Yang, K., Liu, X., Krotkov, N. a., Krueger, A. J., and Carn, S. a.: Estimating the altitude of volcanic sulfur dioxide plumes from space borne hyper-spectral UV measurements, Geophysical Research Letters, 36, L10 803–L10 803, https://doi.org/10.1029/2009GL038025, 2009b.

Yang, K., Liu, X., Bhartia, P. K., Krotkov, N. A., Carn, S. A., Hughes, E. J., Krueger, A. J., Spurr, R. J. D., and Trahan, S. G.: Direct retrieval of sulfur dioxide amount and altitude from spaceborne hyperspectral UV measurements: Theory and application, Journal of Geophysical Research, 115, D00L09–D00L09, https://doi.org/10.1029/2010JD013982, 2010.

Yang, K., Dickerson, R. R., Carn, S. A., Ge, C., and Wang, J.: First observations of SO2 from the satellite Suomi NPP OMPS: Widespread air pollution events over China, Geophysical Research Letters, 40, 4957–4962, https://doi.org/10.1002/grl.50952, 2013.

Yang, K., Carn, S. a., Ge, C., Wang, J., and Dickerson, R. R.: Advancing Measurements of Tropospheric NO_2 from Space: New Algorithm and First Global Results from OMPS, Geophysical Research Letters, 41, 4777–4786, https://doi.org/10.1002/2014GL060136, 2014.

Y.Yang, Y., Meyer, K., Wind, G., Zhou, Y., Marshak, A., Platnick, S., Min, Q., Davis, A. B., Joiner, J., Vasilkov, A., Duda, D., and Su, W.: Cloud products from the Earth Polychromatic Imaging Camera (EPIC): algorithms and initial evaluation, Atmospheric Measurement Techniques, 12, 2019–2031, https://doi.org/10.5194/amt-12-2019-2019, 2019.

Zara, C.: Watch flight traffic literally disappear from the skies as the coronavirus hits travel demand, https://www.fastcompany.com/90473146/watch-flight-traffic-literally-disappear-from-the-skies-as-the-coronavirus-hits-travel-demand, 2020.

Zhao, B., Wang, S., Dong, X., Wang, J., Duan, L., Fu, X., Hao, J., and Fu, J.: Environmental effects of the recent emission changes in China: Implications for particulate matter pollution and soil acidification, Environmental Research Letters, 8, https://doi.org/10.1088/1748-9326/8/2/024031, 2013a.

Zhao, B., Wang, S., Wang, J., Fu, J. S., Liu, T., Xu, J., Fu, X., and Hao, J.: Impact of national NO_x and SO_2 control policies on particulate matter pollution in China, Atmospheric Environment, 77, 453–463, https://doi.org/https://doi.org/10.1016/j.atmosenv.2013.05.012, 2013b.

Zhao, B., Wu, W., Wang, S., Xing, J., Chang, X., Liou, K. N., Jiang, J. H., Gu, Y., Jang, C., Fu, J. S., Zhu, Y., Wang, J., Lin, Y., and Hao, J.: A modeling study of the nonlinear response of fine particles to air pollutant emissions in the Beijing-Tianjin-Hebei region, Atmospheric Chemistry and Physics, 17, 12 031–12 050, https://doi.org/10.5194/acp-17-12031-2017, 2017.

Zhao, B., Zheng, H., Wang, S., Smith, K. R., Lu, X., Aunan, K., Gu, Y., Wang, Y., Ding, D., Xing, J., Fu, X., Yang, X., Liou, K. N., and Hao, J.: Change in household fuels dominates the decrease in PM2.5 exposure and premature mortality in China in 2005–2015, Proceedings of the National Academy of Sciences of the United States of America, 115, 12 401–12 406, https://doi.org/10.1073/pnas.1812955115, 2018.

Zheng, B., Tong, D., Li, M., Liu, F., Hong, C., Geng, G., Li, H., Li, X., Peng, L., Qi, J., Yan, L., Zhang, Y., Zhao, H., Zheng, Y., He, K., and Zhang, Q.: Trends in China's anthropogenic emissions since 2010 as the consequence of clean air actions, Atmospheric Chemistry and Physics, 18, 14 095–14 111, https://doi.org/10.5194/acp-18-14095-2018, 2018.

Zheng, H., Cai, S., Wang, S., Zhao, B., Chang, X., and Hao, J.: Development of a unit-based industrial emission inventory in the Beijing-Tianjin-Hebei region and resulting improvement in air quality modeling, Atmospheric Chemistry and Physics, 19, 3447–3462, https://doi.org/10.5194/acp-19-3447-2019, 2019.