JAMES | Journal of Advances in Modeling Earth Systems

RESEARCH ARTICLE
10.1029/2020MS002413

Description of the NASA GEOS Composition Forecast Modeling System GEOS-CF v1.0

Christoph A. Keller[1,2], K. Emma Knowland[1,2], Bryan N. Duncan[1], Junhua Liu[1,2], Daniel C. Anderson[1,2], Sampa Das[1,2], Robert A. Lucchesi[1,3], Elizabeth W. Lundgren[4], Julie M. Nicely[1,5], Eric Nielsen[1,3], Lesley E. Ott[1], Emily Saunders[1,3], Sarah A. Strode[1,2], Pamela A. Wales[1,2], Daniel J. Jacob[4], and Steven Pawson[1]

[1]NASA Goddard Space Flight Center, Greenbelt, MD, USA, [2]Universities Space Research Association, Columbia, MD, USA, [3]Science Systems and Applications, Inc., Lanham, MD, USA, [4]School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA, [5]Earth System Science Interdisciplinary Center, University of Maryland, College Park, Lanham, MD, USA

Key Points:

- GEOS-CF is a new modeling system that produces global forecasts of atmospheric composition at 25 km^2 horizontal resolution
- GEOS-CF model output is freely available and offers a new tool for academic researchers, air quality managers, and the public

Correspondence to:
C. A. Keller,
christoph.a.keller@nasa.gov

Citation:
Keller, C. A., Knowland, K. E., Duncan, B. N., Liu, J., Anderson, D. C., Das, S., et al. (2021). Description of the NASA GEOS composition forecast modeling system GEOS-CF v1.0. *Journal of Advances in Modeling Earth Systems*, *13*, e2020MS002413. https://doi.org/10.1029/2020MS002413

Received 16 NOV 2020
Accepted 16 MAR 2021

Abstract The Goddard Earth Observing System composition forecast (GEOS-CF) system is a high-resolution (0.25°) global constituent prediction system from NASA's Global Modeling and Assimilation Office (GMAO). GEOS-CF offers a new tool for atmospheric chemistry research, with the goal to supplement NASA's broad range of space-based and in-situ observations. GEOS-CF expands on the GEOS weather and aerosol modeling system by introducing the GEOS-Chem chemistry module to provide hindcasts and 5-days forecasts of atmospheric constituents including ozone (O_3), carbon monoxide (CO), nitrogen dioxide (NO_2), sulfur dioxide (SO_2), and fine particulate matter ($PM_{2.5}$). The chemistry module integrated in GEOS-CF is identical to the offline GEOS-Chem model and readily benefits from the innovations provided by the GEOS-Chem community. Evaluation of GEOS-CF against satellite, ozonesonde and surface observations for years 2018–2019 show realistic simulated concentrations of O_3, NO_2, and CO, with normalized mean biases of −0.1 to 0.3, normalized root mean square errors between 0.1–0.4, and correlations between 0.3–0.8. Comparisons against surface observations highlight the successful representation of air pollutants in many regions of the world and during all seasons, yet also highlight current limitations, such as a global high bias in SO_2 and an overprediction of summertime O_3 over the Southeast United States. GEOS-CF v1.0 generally overestimates aerosols by 20%–50% due to known issues in GEOS-Chem v12.0.1 that have been addressed in later versions. The 5-days forecasts have skill scores comparable to the 1-day hindcast. Model skills can be improved significantly by applying a bias-correction to the surface model output using a machine-learning approach.

Plain Language Summary Accurate forecasting of the compostion of the atmosphere is important for a variety of applications, including air pollution mitigation, support of satellite and other remote-sensing observations, and research applications. Producing such forecasts is computationally expensive due to the complexity of atmospheric chemistry, which interacts with weather on all scales. Here we present the NASA Goddard Earth Observing System composition forecast (GEOS-CF) system, which produces global forecasts of major atmospheric constituents such as ozone (O_3), nitrogen dioxide (NO_2), and fine particulate matter ($PM_{2.5}$). On a daily basis, the model tracks the atmospheric concentrations of more than 250 chemical species in more than 55 million model grid cells, computing the interactions between those species using the state-of-the-science GEOS-Chem chemistry model. We present an in-depth evaluation of the GEOS-CF model through comparison against independent observations. We show how the model captures many observed features of atmospheric composition, such as spatio-temporal variations in air pollution due to changes in pollutant emissions, weather, and chemistry. We also highlight some of the model deficiencies, for example, with respect to the simulation of aerosol particles. Finally, we demonstrate how surface observations and model data can be combined using machine learning to provide improved local air quality forecasts.

1. Introduction

Near real-time information of global atmospheric composition is invaluable for a wide range of applications, including academic research, airborne and satellite mission support, air quality forecasting, disaster

management, and ecosystem monitoring. However, the numerical simulation of atmospheric chemistry is computationally expensive because it involves hundreds of species that interact with each other on time scales from milliseconds to years, and the species are also influenced by dynamics across a wide range of spatiotemporal scales. This precludes the inclusion of detailed aerosol and reactive trace gases in standard operational numerical weather prediction (NWP) systems. Instead, real-time simulation of atmospheric composition is typically done within a simplified system in order to reduce the computational burden, for example, by running the model at reduced horizontal resolution or over a regional domain only, using a simplified representation of atmospheric composition, or by coupling a weather model with an offline chemical transport model (CTM) (e.g., Bhattacharjee et al., 2018; Emmons et al., 2020; Flemming et al., 2009; Marécal et al., 2015).

Here we present the NASA Goddard Earth Observing System (GEOS) composition forecast modeling system, GEOS-CF v1.0, which provides global near real-time estimates ("hindcasts") and forecasts of atmospheric composition such as ozone (O_3), carbon monoxide (CO), nitrogen dioxide (NO_2), sulfur dioxide (SO_2), and fine particulate matter ($PM_{2.5}$) in near real-time at a horizontal resolution of approximately 25 × 25 km^2. As summarized by the World Meteorological Organization (WMO) Global to Local Air Quality Forecast Inventory (https://community.wmo.int/activity-areas/gaw/science-for-services/gafis/air-quality-forecast-inventory), GEOS-CF is one of only a few global forecasting systems conducting gaseous and aerosol composition simulations in near real-time. The European Center for Medium-Range Weather Forecasts (ECMWF) offers 5-days global forecasts of aerosols and trace gases at approximately 40 × 40 km^2 horizontal resolution through the Copernicus Atmosphere Monitoring Service (CAMS, https://atmosphere.copernicus.eu/global-forecast-plots). The US National Center for Atmospheric Research (NCAR) conducts 10-days global forecasts at approximately 100 × 100 km^2 horizontal resolution based on offline simulations of the Model for Ozone and Related chemical Tracers (MOZART) Chemistry Mechanism in the Community Earth System Model Version 2 (CESM2) (Emmons et al., 2020) driven by GEOS meteorological forecasts (https://www2.acom.ucar.edu/acresp/forecasts-and-near-real-time-nrt-products). Finally, the Finnish Meteorological Institute provides daily 4-days global forecasts of atmospheric composition at approximately 35 × 35 km^2 resolution using the System for Integrated modeling of Atmospheric coMposition (SILAM v5.7, http://silam.fmi.fi/). The atmospheric composition forecasts from these models can vary considerably due to differences in the underlying meteorological fields, observational constraints, chemical mechanisms, or assumptions about pollutant emissions (see, e.g., Huijnen et al., 2019). The uncertainties associated with these processes can be difficult to quantify from a single model simulation alone, and the availability of multiple, independently developed models offers great potential to provide better air quality information through combination of a suite of models (Marécal et al., 2015). This is recognized by the Global Air Quality Forecasting and Information System (GAFIS) project, a new initiative by the WMO Global Atmospheric Watch (GAW) program with the aim to provide globally harmonized air quality forecasts tailored to the needs of end users. GEOS-CF can support such efforts by offering an independent but complementary composition forecast to the existing suite of global composition forecasting systems in the form of global 5-days forecasts of atmospheric composition based on the GEOS-Chem atmospheric chemistry module (http://www.geos-chem.org) within the GEOS Earth System Model (ESM).

One of the key aspects of the GEOS-CF system is the full integration of the offline GEOS-Chem CTM in the GEOS system, which allows for the simulation of reactive gases and aerosols at the same temporal and spatial resolution as the meteorology (Hu et al., 2018; Long et al., 2015). GEOS-Chem is a global CTM driven by archived meteorological inputs produced by GEOS. It is actively evaluated and developed by a large international research community. The chemical components of GEOS-Chem (including emission, deposition, and aerosol-oxidant chemistry) can also be used online in an ESM such as GEOS, bypassing the need for the offline transport scheme used by the CTM. Instead, meteorology and species transport is calculated directly within the ESM. The scientific validity of the GEOS-Chem chemistry module within GEOS has been demonstrated by Hu et al. (2018), who show that a one-year global simulation of atmospheric composition at 12.5 km^2 produces results consistent with the offline GEOS-Chem model. The online application has advantages over the CTM as it does not introduce transport errors related to temporal and spatial averaging of the archived meteorological input fields (Yu et al., 2018), but comes at the cost of increased complexity and computing requirements. GEOS-CF is an application of the online GEOS-Chem module embedded in GEOS, with a focus on daily operation and forecast capabilities. The GEOS-Chem chemistry module

incorporated in GEOS-CF uses the exact same codebase as the offline CTM. This enables seamless integration of scientific updates provided by the GEOS-Chem CTM community into GEOS-CF without the need to make any modifications to the source code.

The GEOS-CF modeling system is the latest in a series of research and applications products generated by the NASA Global Modeling and Assimilation Office (GMAO), including the GEOS forward processing weather and aerosol system (GEOS FP; Lucchesi, 2017), GEOS FP for instrument teams (GEOS FP-IT; Lucchesi, 2015), the second Modern-Era Retrospective analysis for Research and Applications (MERRA-2; Gelaro et al., 2017), and the Seasonal to Subseasonal Forecasting System (GEOS-S2S; Borovikov et al., 2019; Molod et al., 2020). It leverages GMAO's model infrastructure and directly builds on a number of development activities centered around the GEOS model, with the goal to extend these forecasting capabilities toward (short-lived) trace gases and aerosols.

The GEOS ESM is a General Circulation Model (GCM) and Data Assimilation System (DAS) consisting of a suite of model components that can be connected in a modular manner through the Earth System Modeling Framework (ESMF, Hill et al., 2004) and Modeling Analysis and Prediction Layer (MAPL, Suarez et al., 2007) software interface. The model can be configured to run with fully interactive chemistry so that the chemical constituents feed back to the dynamics ("online"), or as an offline model where external meteorological fields are used as input. A hybrid approach is the "replay" feature, where the model dynamics are nudged toward pre-computed analysis fields (from a previous DAS simulation) in a way that is consistent with the internal physics of the model (Orbe et al., 2017). This approach is particularly useful for chemistry simulations as it sidesteps the need to conduct a computationally costly meteorological data assimilation cycle. Several chemistry and aerosol modules of varying complexity are available in GEOS (Nielsen et al., 2017), enabling a wide range of applications including but not limited to near real-time simulation of aerosols using the GOCART module (Buchard et al., 2017; Colarco et al., 2010; Randles et al., 2017), computationally efficient analysis of stratospheric O_3 using parameterized chemistry in combination with 3D-Variational assimilation of satellite observations (Wargan et al., 2015); multi-decade simulation of tropospheric and stratospheric chemistry using the Global Modeling Initiative (GMI) chemistry module (Douglass et al., 2004; Duncan et al., 2007; Strahan et al., 2007); and fully coupled simulation of gas-phase and aerosol chemistry using the chemistry module of the GEOS-Chem CTM embedded in GEOS (Long et al., 2015). The latter is used as the principal driver of chemistry in GEOS-CF.

In this paper we describe the configuration of GEOS-CF version 1.0 (Section 2) and demonstrate the validity of the produced hindcasts through comparison against independent observations for years 2018 and 2019 (Sections 3 and 4). Model forecasts over the same time period are evaluated in Section 5 and we discuss how model skill scores can be improved by applying a bias-correction to the surface observations using a machine-learning approach. As described in detail below, the current version of GEOS-CF constitutes a hybrid model between an online weather and chemistry assimilation system and an offline CTM application, with a development pathway toward a fully coupled forecasting system with integrated trace gases and aerosols. Many of its design features were guided by practical considerations as well as computational limitations related to the simulation of atmospheric chemistry.

2. Model Description

2.1. General Description

A schematic of the GEOS-CF v1.0 modeling system is provided in Figure 1. Since 2018, GEOS-CF is operated in near real-time, producing a 5-days forecast once a day. The forecast initial conditions are provided by a 1-day replay simulation (referred to as "hindcast") constrained by pre-computed meteorological analysis fields. In the v1.0 configuration, the meteorological forcings that GEOS-CF uses during the hindcast step are obtained from the GEOS FP-IT system. GEOS FP-IT is a "frozen" model system that is comparable to MERRA-2 but—unlike MERRA-2—is available in near real-time to support retrievals by satellite instrument teams. The GEOS model in general, and the version of the GEOS GCM used by GEOS FP-IT (v5.12.4) in particular, has shown to be well suited for atmospheric chemistry applications as it realistically captures features critical to atmospheric composition, such as the seasonal climate of moisture and temperature and

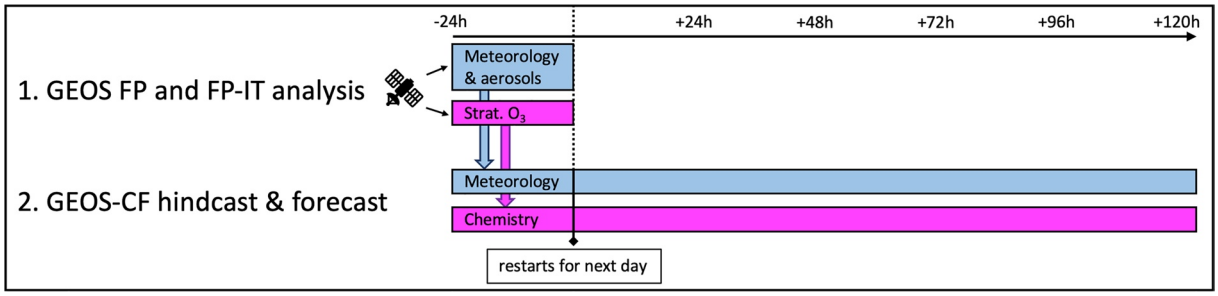

Figure 1. Schematic of the GEOS-CF modeling system approach, consisting of 1 day hindcast and 5-days forecast. This combination of simulations is conducted on a daily basis.

large-scale transport of constituents (e.g., Douglass et al., 2014; Gelaro et al., 2017; Molod et al., 2015; Oman and Douglass, 2014; Pawson et al., 2007).

In GEOS FP-IT, the GOCART module provides aerosol feedback on meteorology. To ensure consistent model physics between GEOS FP-IT and GEOS-CF, the physics component in GEOS-CF is coupled to the GOCART module, with GOCART aerosol optical depth (AOD) constrained by satellite observations as further described below. The GEOS-Chem chemistry module, which contains its own aerosol module coupled to the gas-phase chemistry, is run "passively" to provide coupled aerosol-oxidant chemistry in the troposphere and stratosphere (Hu et al., 2018). That is, the GEOS-Chem aerosols and trace gases are advected online (including turbulence), while emissions, dry and wet deposition, and chemistry are handled by the GEOS-Chem chemical module driven by instantaneous meteorological information from the GEOS physics modules. Deep convective transport of chemical species is performed by the GEOS-Chem convection scheme driven by instantaneous model diagnostics to capture scavenging in the updrafts. The GEOS-Chem aerosols and trace gases are currently integrated without observational constraints, with the exception of stratospheric ozone which is nudged to precomputed ozone fields from the GEOS FP system (described in more detail below).

A full snapshot of the model conditions at the end of the simulated hindcast day are saved and serve as input for the next day's hindcast step ("restart files"). This preserves the model chemical and meteorological state from one forecast cycle to the next, leading to a continuous atmospheric composition archive since January 2018.

GEOS-CF v1.0 uses a model physics package that is similar to MERRA-2 and GEOS FP-IT, consisting of the GEOS atmospheric model, version 5, described in Rienecker et al. (2008) with updates as described in Molod et al. (2012) and Molod et al. (2015). The model uses the finite-volume dynamical core of Lin (2004) with a cubed sphere grid discretization to avoid grid-cell singularities (Putman and Lin, 2007). It is run at cubed-sphere c360 horizontal resolution (roughly equivalent to $0.25° \times 0.25°$) and 72 hybrid-eta levels from the surface to 0.01 hPa. This results in 30–35 model levels in the troposphere and a surface model level depth of approximately 130 m. Model physics includes parameterizations for moist processes, radiation, turbulent mixing, land-surface processes, and gravity wave drag. The moist module contains parameterization of convection using the Relaxed Arakawa-Schubert scheme (Moorthi & Suarez, 1992), and the single-moment parameterization for large-scale precipitation and cloud cover described in Bacmeister et al. (2006). Note that in GEOS-CF v1.0, the moist processes in GEOS are not directly applied to the GEOS-Chem species but provide the diagnostic variables needed by GEOS-Chem to perform these tasks (Hu et al., 2018; Yu et al., 2018). The radiation module includes parameterization for long-wave (Chou 1990, 1992) and short-wave radiation processes (Chou and Suarez 1994). Turbulence is parameterized using the gradient Richardson number in the free atmosphere and the Lock scheme (Lock et al., 2000) interfaced with the scheme of Louis and Geleyn (1982) in the boundary layer. Exchange of heat, moisture and momentum between land, atmosphere, and ocean or sea ice surfaces are parameterized using Monin-Obukhov similarity theory (Helfand and Schubert, 1995; Molod et al., 2013), and the gravity wave drag parameterization contains orographic (McFarlane, 1987) and non-orographic (Garcia & Boville, 1994) waves.

In the GEOS-CF system, the GEOS physics components are coupled to the GOCART aerosol component to provide consistent physics with the GEOS FP-IT meteorology. Specifically, the GEOS-Chem chemistry module is run alongside the GOCART module, with no feedbacks to model physics.The GOCART aerosols, constrained during the hindcast step by AOD observations from the Moderate Resolution Imaging Spectroradiometer (MODIS) aboard the Terra and Aqua satellites (Randles et al., 2017), are used to compute the feedback between aerosols and dynamic while the GEOS-Chem chemistry module, run alongside the GOCART module, does not feedback to model physics. Currently, there is no chemical coupling between GOCART and GEOS-Chem, and no observations are directly assimilated into GEOS-CF. However, stratospheric O_3 in GEOS-Chem is nudged toward O_3 produced by GEOS FP, which is constrained by ozone measurements from the Microwave Limb Sounder (MLS), Ozone Monitoring Instrument (OMI), and NASA's Ozone Mapping and Profiler Suite (OMPS) and produces a realistic analysis of O_3 in the stratosphere (Wargan et al., 2015, 2020). In addition, near real-time MODIS observations of fire radiative power are used to constrain fire emissions, as produced by the Quick Fire Emissions Dataset (QFED) (Darmenov and Da Silva, 2015). For the forecasts, the latest available QFED estimates are assumed to persist over the entire forecasting period.

All computations are conducted on the Discover supercomputing cluster of the NASA Center for Climate Simulation (NCCS). Run on 3510 Intel Xeon Haswell processor cores, the 1-day hindcast and 5-days forecast takes approximately 8.5 wall-clock hours in total. GEOS-CF hindcast and forecast output includes chemistry and meteorology "surface" output every 15 min as well as hourly-average and instantaneous fields for surface, column-average, and 3-dimensional model output. The model output is publicly available at https://gmao.gsfc.nasa.gov/weather_prediction/GEOS-CF/data_access/in the form of on-demand figures or through access to the model output (in netCDF data format) via Hypertext Transfer Protocol (HTTP) file download or through the Open-source Project for a Network Data Access Protocol (OPeNDAP) remote access tool. The full 5-days model forecast output is publicly available for a duration of 14 days. Given the growing interest in air quality forecasting applications, the model forecasts for a selection of surface air pollutants are made available on the public portal indefinitely. Full details on available output and data access are available in the GEOS-CF File Specification document (Knowland et al., 2020).

2.2. Chemistry

GEOS-CF v1.0 uses the continually updated standard version of the GEOS-Chem chemistry module to simulate coupled aerosol-oxidant chemistry in the troposphere and stratosphere. The results presented here use GEOS-Chem version 12.0.1 in order to have a two-year record for comparison to observations, while the latest version 13.0 has been implemented into GEOS for a future version of GEOS-CF. GEOS-Chem is ESMF-compliant and its chemistry module is implemented here as an ESMF gridded component of GEOS, as described in Long et al. (2015) and Hu et al. (2018).

The gas-phase chemistry scheme includes detailed tropospheric chemistry of HO_x, NO_x, BrO_x, volatile organic compounds (VOC), and O_3, as originally described by Bey et al. (2001), with addition of halogen chemistry by Parrella et al. (2012) and Sherwen et al. (2016) plus updates to isoprene oxidation as described by Mao et al. (2013) and Miller et al. (2017). Stratospheric chemistry is fully coupled with tropospheric chemistry through the Unified tropospheric-stratospheric Chemistry eXtension (UCX; Eastham et al., 2014) and extends to the top of the atmosphere. Photolysis rates are computed by GEOS-Chem using the Fast-JX code (Bian and Prather, 2002). The gas-phase mechanism comprises 250 chemical species and 725 reactions and is solved using the Kinetic PreProcessor (KPP) Rosenbrock solver (Sandu and Sander, 2006). The aerosol simulation includes sulphate-nitrate-ammonia chemistry (Park et al., 2004), black carbon (Park et al., 2003; Wang et al., 2014), organic aerosols (Marais et al., 2016), mineral dust (Fairlie et al., 2007; Ridley et al., 2012), and sea salt aerosols (Jaeglé et al., 2011). Aerosol and gas-phase chemistry interact through gas-aerosol partitioning (Fountoukis and Nenes, 2007; Pye et al., 2009), heterogeneous chemistry on aerosol surface (Evans and Jacob, 2005; Mao et al., 2013), and aerosol impacts on photolysis (Martin et al., 2003). Methane concentrations are prescribed as monthly mean surface concentrations, spatially interpolated from NOAA GLOBALVIEW flask data (Dlugokencky et al., 1994).

Table 1
Emissions Used by GEOS-CF

Description	Reference	Comments
Offline inventories		
Anthropogenic NO, CO, black carbon (BC), organic carbon (OC), Ammonia (NH_3)	HTAP v2.2 (Janssens-Maenhout et al., 2015)	Global except Africa
Anthropogenic SO_2	OMI-HTAP (Liu et al., 2018)	Global except Africa
Anthropogenic VOCs	RETRO (Schultz et al., 2008)	Global except Africa
Anthropogenic NO, CO, SO_2, BC, OC, NH_3, VOCs	DICE-Africa (Marais and Wiedinmyer, 2016)	Africa
Arctic seabird NH_3	Croft et al. (2016)	
Volcanic SO_2	Carn (2019)	5% of the sulfur emitted as SO_4
Aircraft NO_x (=NO + NO_2), CO, SO_2, VOCs, BC, OC	AEIC (Stettler et al., 2011)	
Emissions calculated online based on real-time environment		
Biogenic VOCs	MEGAN v2.1 (Guenther et al., 2012)	
Biomass burning (wildfires) NO_x, CO, SO_2, VOCs, BC, OC	QFED v2.5 (Darmenov and da Silva, 2015)	35% emitted between 3.5 and 5.5 km altitude (Fischer et al., 2014).
Lightning NO_x	Murray et al., 2012	
Soil NO_x	Hudman et al., 2012	
Soil dust	Zender et al., 2003	
Sea salt aerosols	Gong, 2003; Jaeglé et al., 2011	
Oceanic DMS, CH_2O, C_3H_6O	Johnson, 2010; Nightingale et al., 2000	
Oceanic iodine	Carpenter et al., 2013	

2.3. Emissions and Deposition

The dry deposition scheme in GEOS-Chem is based on the resistance-in-series model of Wesely (1989), as implemented by Wang et al. (1998). Wet deposition of aerosols and soluble gases includes scavenging in convective updrafts, in-cloud rainout, and below-cloud washout (Amos et al., 2012; Liu et al., 2001). While these processes are not fully coupled to GEOS, they are all driven by instantaneous GEOS model diagnostics and thus directly respond to changes in the model environment.

All emission calculations are done using the Harmonic Emissions Component HEMCO v2.1.009 (Keller et al., 2014). Table 1 summarizes the emission configuration used by GEOS-CF v1.0. The emission inputs used by GEOS-CF are either calculated online based on current conditions (e.g., lightning NO_x, dust emissions, sea salt), or obtained from pre-computed emission inventories. For anthropogenic emissions, the Hemispheric Transport of Air Pollution emissions inventory (HTAP v2.2; Janssens-Maenhout et al., 2015) is used as the primary source due to its high spatial resolution (0.1°) and the availability of monthly emissions data (as opposed to annual emission averages). HTAP v2.2 provides harmonized global emission maps combining global emissions data from the Emission Database for Global Atmospheric Research (EDGAR) (Crippa et al., 2018) with regional emission inventories over Europe, Asia, and the US. For Africa, HTAP v2.2 anthropogenic emissions are overwritten with the DICE emissions inventory (Marais and Wiedinmyer, 2016) to take full advantage of local information. Anthropogenic emissions are further broken down into hourly values using sector-specific day-of-week and diurnal scale factors (Denier van der Gon et al., 2011). For CO, the HTAP v2.2 base emissions, which are representative for year 2010, are adjusted to more recent conditions by applying an annual gridded scale factor based on the latest available Open-source Data Inventory for Anthropogenic CO_2 (ODIAC; Oda & Maksyutov, 2017), assuming the same relative change in anthropogenic CO emissions as for CO_2. This results in an increase in anthropogenic CO emissions by a factor of 1.2–1.4 over China, India, and South America and decreases by 5%–10% over Europe and North America for year 2018 (relative to 2010). For other anthropogenic emissions, no attempts are currently made in GEOS-CF to adjust the HTAP v2.2 base emissions. This is an obvious simplification, especially for species and regions undergoing rapid emission changes, such

Table 2
Overview of Observation Data Sets Used for GEOS-CF Model Validation

Description	Species	# Sites	Reference
Ozonesonde	O_3	24	Tilmes et al., 2012; Thompson et al., 2017
NASA OMI NO_2 v4.0	Tropospheric column NO_2	global	Lamsal et al., 2021
MOPITT v8	Total column CO	global	Deeter et al., 2019
MODIS	AOD at 550 nm	Global	Remer et al., 2005; Levy et al., 2010, 2015
AERONET	AOD at 550 nm	195	Giles et al., 2019
GAW WDCGG	CO	54	https://gaw.kishou.go.jp/
GAW WDCRG	O_3, NO_2, SO_2	48 (O_3), 6 (NO_2), 9 (SO_2)	https://www.gaw-wdcrg.org/
OpenAQ	O_3, NO_2, SO_2, $PM_{2.5}$	3151 (O_3), 2789 (NO_2), 1221 (SO_2), 2667 ($PM_{2.5}$)	https://openaq.org

as SO_2 and NO_2 over China or India over the last decade. Incorporation of real-time information into the emissions calculation is currently under development and slated to be included in future versions of GEOS-CF.

2.4. Observations Used for Model Evaluation

GEOS-CF is intended to supplement NASA's broad range of space-based and in-situ observations, providing a new tool for researchers, government scientists, and air quality managers. We therefore focus our evaluation on the species most pertinent to these applications, including O_3, NO_2, CO, SO_2, and aerosols. These species are also constantly evaluated by the broader GEOS-Chem community using the standard CTM simulations (e.g., Hu et al., 2017, 2018; Travis & Jacob, 2019). However, it should be noted that GEOS-CF simulations can differ from standard GEOS-Chem simulations due to the higher horizontal resolution (Hu et al., 2018), the use of online meteorological information instead of time-averaged meteorological fields (Yu et al., 2018), and differences in the emission inputs, in particular wildfire emissions from QFED versus the Global Fire Emissions Database (GFED) typically used in the GEOS-Chem CTM (Carter et al., 2020).

We first evaluate the ability of the GEOS-CF hindcast to provide a realistic representation of atmospheric composition based on the hourly-average hindcast fields (Section 4). The validation covers the two-year time period from the start of the GEOS-CF record on 1 Jan 2018 to 31 Dec 2019. Differences between the 5-days model forecasts and the model hindcast are discussed in Section 5. All aerosol comparisons are for the aerosol fields calculated by the GEOS-Chem module, thus providing aerosol information that is consistent with the gas-phase chemistry. A detailed comparison of GOCART AOD and $PM_{2.5}$ is given in Buchard et al. (2017).

The data sets used for model validation are summarized in Table 2. Briefly, we evaluate the global tropospheric distribution of O_3 against ozonesonde observations obtained from the World Ozone and Ultraviolet Data Center (WOUDC, http://www.woudc.org) and the NOAA Earth System Research Laboratory—Global Monitoring Division (ftp://ftp.cmdl.noaa.gov/ozww/Ozonesonde/), NO_2 against the OMI NASA standard tropospheric column NO_2 product v4.0 (Lamsal et al., 2021) and CO total columns against satellite retrievals from the Measurements Of Pollution In The Troposphere (MOPITT) V8 (Deeter et al., 2019). All satellite comparisons use model output sampled at satellite overpass time and with averaging kernels applied to it. For AOD, we compare the model simulation against the Aqua MODIS combined dark target and deep blue aerosol level 2.0 product (https://modaps.modaps.eosdis.nasa.gov/services/about/products/c6/MYD04_L2.html; Remer et al., 2005; Levy et al., 2010, 2015) as well as Aerosol Robotic Network (AERONET) level 2.0 daily average data from the version three data product (https://aeronet.gsfc.nasa.gov/; Giles et al., 2019). Simulated surface concentrations are compared against in-situ observations from the Global Atmospheric Watch (GAW) World Data Center for Greenhouse Gases (WDCGG, https://gaw.kishou.go.jp/) and World Data Center for Reactive Gases (WDCRG, https://www.gaw-wdcrg.org/), as well as observations from the OpenAQ database (https://openaq.org). We only include OpenAQ sites with at least 80% of data coverage during the analyzed 2-year period. The analysis of surface concentrations is

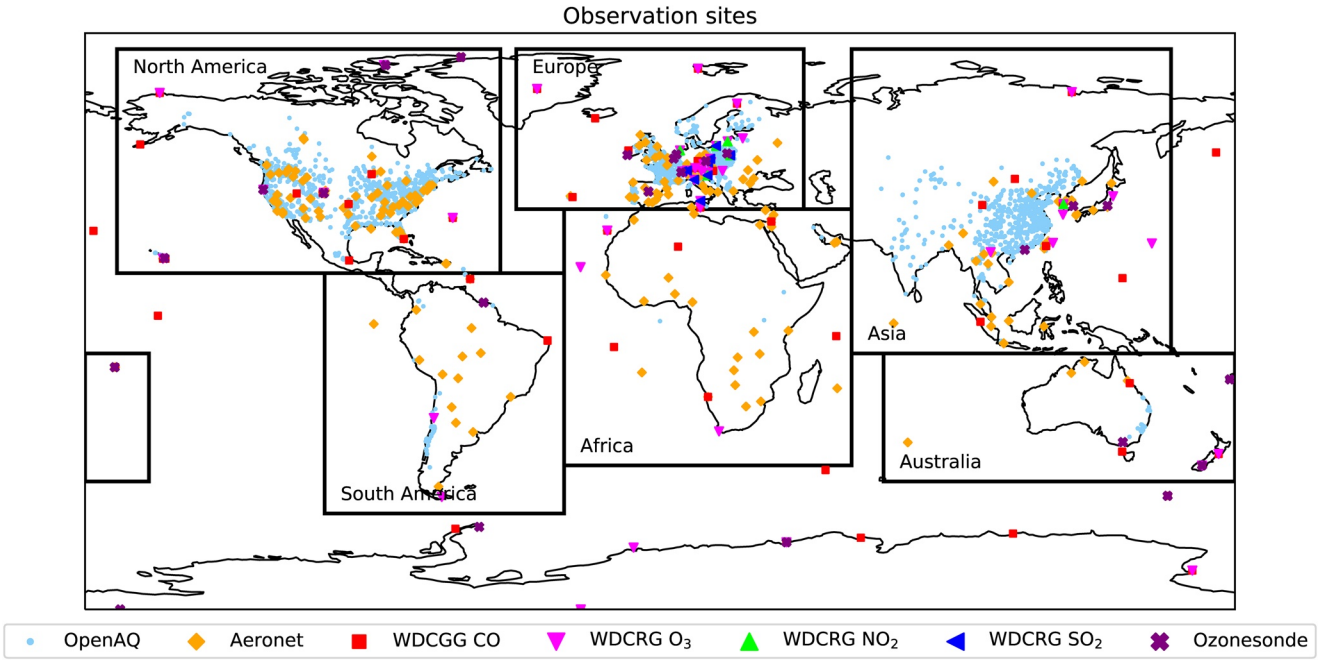

Figure 2. Overview of observation sites used for model validation. Black boxes show the six regions used for aggregation of OpenAQ and Aeronet observations.

based on daily mean values. For NO_2, we omit mountainous GAW sites given that the model resolution of 25 × 25 km^2 is not high enough to resolve the fine-scale vertical gradients around mountain slopes. On average, the depth of the GEOS-CF model surface layer is 130 meters and we use this value without attempting to adjust for sub-grid vertical gradients. Accounting for the subgrid vertical gradient can improve model-observation comparisons for ozone (Travis & Jacob, 2019) but is more challenging to implement for species with both sources and sinks, such as NO_2 or $PM_{2.5}$. All aerosol evaluation is based on the GEOS-Chem aerosols, and model $PM_{2.5}$ is calculated for a relative humidity (RH) of 35%, following standard practice in GEOS-Chem.

Figure 2 shows the global distribution of all in-situ observations used in the evaluation. The GAW observation sites (WDCGG and WDCRG) are located away from the major pollution sources, representing the global background concentrations. In contrast, most observations available on the OpenAQ platform are located in densely populated areas, providing information about local air pollution. With a model grid size of 25 × 25 km^2, GEOS-CF is not expected to resolve many of the local features observed at monitoring sites available on OpenAQ. However, the use of model forecasts for local air quality applications is a popular request and the OpenAQ comparisons shown here are intended to serve as a benchmark for such types of applications. We discuss in Section 5 how bias correction methods can be used to overcome some of the model representativeness errors. It should also be emphasized that the observations provided on OpenAQ are reported in near real-time and thus do not undergo the same level of quality control as the GAW observations. While this makes OpenAQ an excellent platform to monitor a system in near real-time, it can introduce analysis errors related to observation errors that can be difficult to quantify. In an ad-hoc attempt to exclude obviously errorenous observations (such as ozone concentrations of several 1000 ppbv), we remove all OpenAQ observations that are 3 standard deviations beyond the mean monthly value of that site. This results in a removal of less than 0.1% of all OpenAQ observations.

The OpenAQ and AERONET observations are further grouped into six geographical regions, as shown in Figure 2. This subset of observations provides good coverage for North America, Europe, and Eastern Asia (especially China) but a limited view of other regions such as South America or Africa. For those under-represented regions, our analysis relies heavily on the global daily satellite observations.

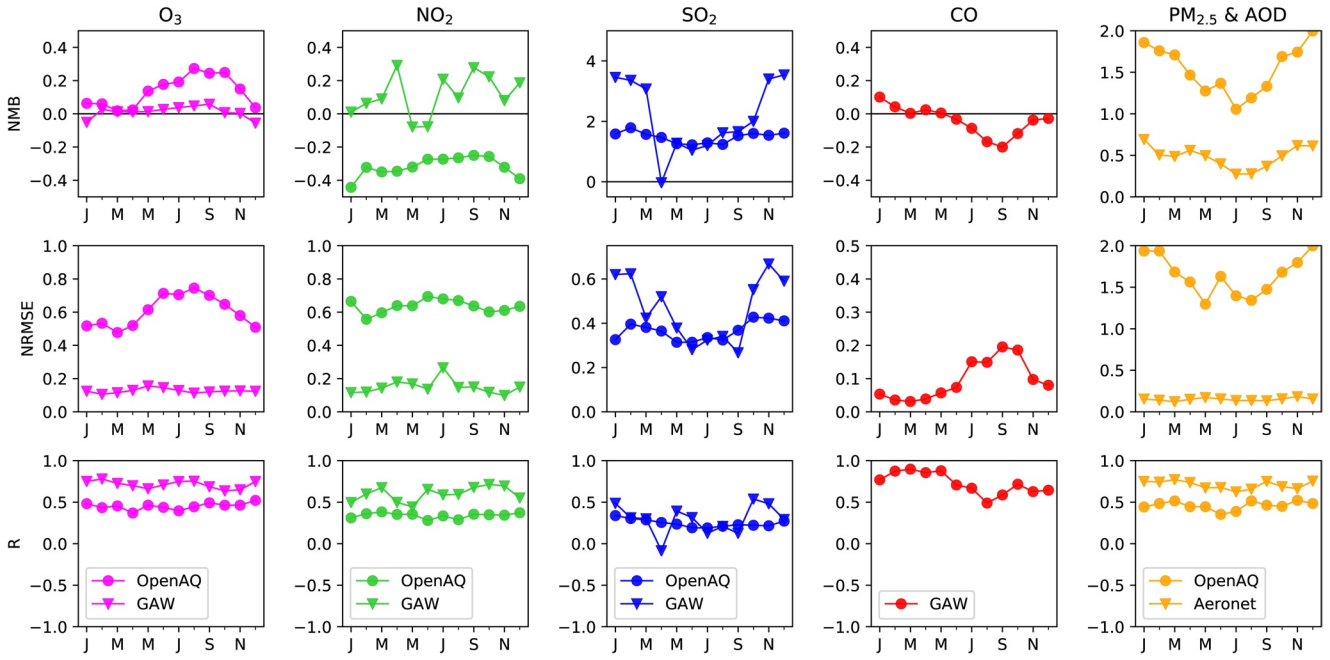

Figure 3. Overview of NMB, NRMSE and *R* for daily mean O_3, NO_2, SO_2, CO, and $PM_{2.5}$ (OpenAQ only) and AOD (Aeronet only). Shown are the monthly skill score means, as calculated from daily mean concentrations, aggregated by observation type. *Note*. the different scale for SO_2, $PM_{2.5}$ and AOD.

3. Evaluation of Model Hindcast

3.1. Model Skill Scores

We first highlight the model skill scores against all surface observations before discussing the individual species in more detail below. Figure 3 shows monthly model skill scores for O_3, NO_2, SO_2, CO, and $PM_{2.5}$ and/or AOD, aggregated by observation type across all measurement sites. The skill scores used are the normalized mean bias (NMB), normalized root mean square error (NRMSE), and Pearson correlation coefficient (*R*):

$$\mathrm{NMB} = \frac{\Sigma_{n=1}^{N}\left(y_n - o_n\right)}{\overline{o}} \tag{1}$$

$$\mathrm{NRMSE} = \frac{\sqrt{\frac{1}{N}\Sigma_{n=1}^{N}\left(y_n - o_n\right)^2}}{o_{0.95} - o_{0.05}} \tag{2}$$

$$R = \frac{\Sigma_{n=1}^{N}\left(y_n - \overline{y}\right)\left(o_n - \overline{o}\right)}{\sqrt{\Sigma_{n=1}^{N}\left(y_n - \overline{y}\right)^2}\sqrt{\Sigma_{n=1}^{N}\left(o_n - \overline{o}\right)^2}} \tag{3}$$

where y_n denotes an indidivual model estimate, $\overline{y}$ is the model average, o_n is the observation associated with y_n, and $\overline{o}$ is the observation average; $o_{0.05}$ and $o_{0.95}$ are the 5th and 95th percentile, respectively, of the observations sample, and N is the total number of observation-model pairs. All skill scores are calculated using daily average concentrations. This removes the impact of the diurnal cycle, which otherwise can dominate the correlation signal. Differences between night-time and day-time skill scores are small on average (not shown).

For daily average surface O_3, the model shows a normalized mean annual bias of 0.011 compared to the GAW sites, with little seasonal variability (Figure 3, first column). The GAW sites are typically located in remote locations and the small model bias indicates that the model adequately captures atmospheric background O_3 throughout the year. This is confirmed by the low NRMSE of 0.13 and the correlation coefficient of 0.71,

again with small seasonal variability. Compared to the GAW sites, the simulated O_3 shows weaker skill scores relative to the OpenAQ observations with an NMB of 0.14, an NRMSE of 0.61 and *R* of 0.45. A lower model skill against the observations from the OpenAQ data portal is not surprising given that OpenAQ sites are typically located close to polluted sites, and the model representativeness error is therefore expected to be higher. As further discussed below, the OpenAQ comparisons also highlight a pronounced model overestimation of surface O_3 during the summer months.

For NO_2 (second column of Figure 3), the normalized model bias is 0.11 compared against the GAW sites, and −0.32 against the OpenAQ sites. As for O_3, the larger differences between GEOS-CF and OpenAQ observations (compared to GEOS-CF vs. GAW) can likely be attributed to the spatial representation mismatch between the 25×25 km^2 model output and the OpenAQ observations. The NMB at the GAW sites varies strongly from month-to-month, reflecting the small sample size (6 stations). The OpenAQ bias shows the strongest negative bias between November through January, with a reduction of the negative bias from −0.4 to −0.25 during June–August. Given that the OpenAQ sites are predominantly located in the NH (Figure 2), this reflects a stronger model underestimation during winter. The NRMSE is 0.15 at the GAW sites and 0.64 at the OpenAQ sites, similar to O_3. The *R* values for NO_2 are 0.6 and 0.34 at the GAW and OpenAQ sites, respectively. Analysis of NO_2 is complicated by the potential NO_2 observation interference with other nitrogen compounds (Grosjean and Harrison, 1985; Steinbacher et al., 2007; Winer et al., 1974), as well as the short atmospheric lifetime of NO_2, which makes it challenging for the model to capture the spatiotemporal variability in surface NO_2 in full detail. This explains some of the low model bias for NO_2 relative to OpenAQ observations, which are commonly located in urban locations close to combustion sources.

For SO_2, the model—observation comparison shows a large model bias of up to a factor of 4, with an annual mean bias of 2.1 at GAW sites and 1.5 at OpenAQ locations. The NRMSE is 0.47 for the GAW observations and 0.37 for OpenAQ data, and the correlations are similar with 0.29 and 0.25 for GAW and OpenAQ, respectively. An important factor for the model overestimation is the use of outdated anthropogenic SO_2 emissions data, as further discussed below.

The simulated CO compares well against the 54 GAW sites, with an NMB of −0.042, an NRMSE of 0.096 and *R* of 0.73. The model error is largest during the NH summer months, with a decrease in NMB to −0.14, an increase in NRMSE to 0.2, and a reduction of *R* to 0.5.

The model comparisons against AERONET AOD and OpenAQ $PM_{2.5}$ observations reveal a systematic model overestimation of aerosol concentrations in GEOS-CF, with an average NMB of 0.48 for AOD and 1.5 for surface $PM_{2.5}$. Excessive sulphate aerosol as well as known model issues in the treatment of model emissions and wet scavenging contribute to these biases, as will be discussed in more detail below. In addition, some $PM_{2.5}$ observations on the OpenAQ platform represent dry particulate matter while model $PM_{2.5}$ assumes 35% RH, which also contributes to the mismatch. At the AERONET sites, the model shows a good correlation score of 0.71 and the NMRSE of 0.15 is comparable to that of other species at background sites, indicating that the model captures the relative changes in AOD well despite the high model bias. For surface $PM_{2.5}$, the average model NRMSE is 1.65 and *R* skill score is 0.46.

3.2. Ozone

Comparisons of GEOS-CF against ozonesondes and GAW surface observations demonstrate that GEOS-CF reproduces well-known features of background surface O_3, such as the local maxima in the NH during spring.

Figure 4 shows annual mean (2018–2019) vertical O_3 profiles as observed by ozonesondes at 24 locations across the world (purple crosses in Figure 2), compared against corresponding GEOS-CF model concentrations. While GEOS-CF generally captures the observed vertical structure of O_3, the model tends to underestimate free tropospheric O_3 (approx. 800–300 hPa) over the NH midlatitudes. For these profiles (e.g., De Bilt, Payerne, Madrid, Tateno), the simulated vertical gradient is much less pronounced than observed and the model consistently underestimates O_3 concentrations. This pattern is consistent with previous comparisons of the online GEOS-Chem module against observations (Hu et al., 2018) and implies two potential model errors: (1) excessive vertical mixing resulting in an overestimation in the lower altitudes and underestimation

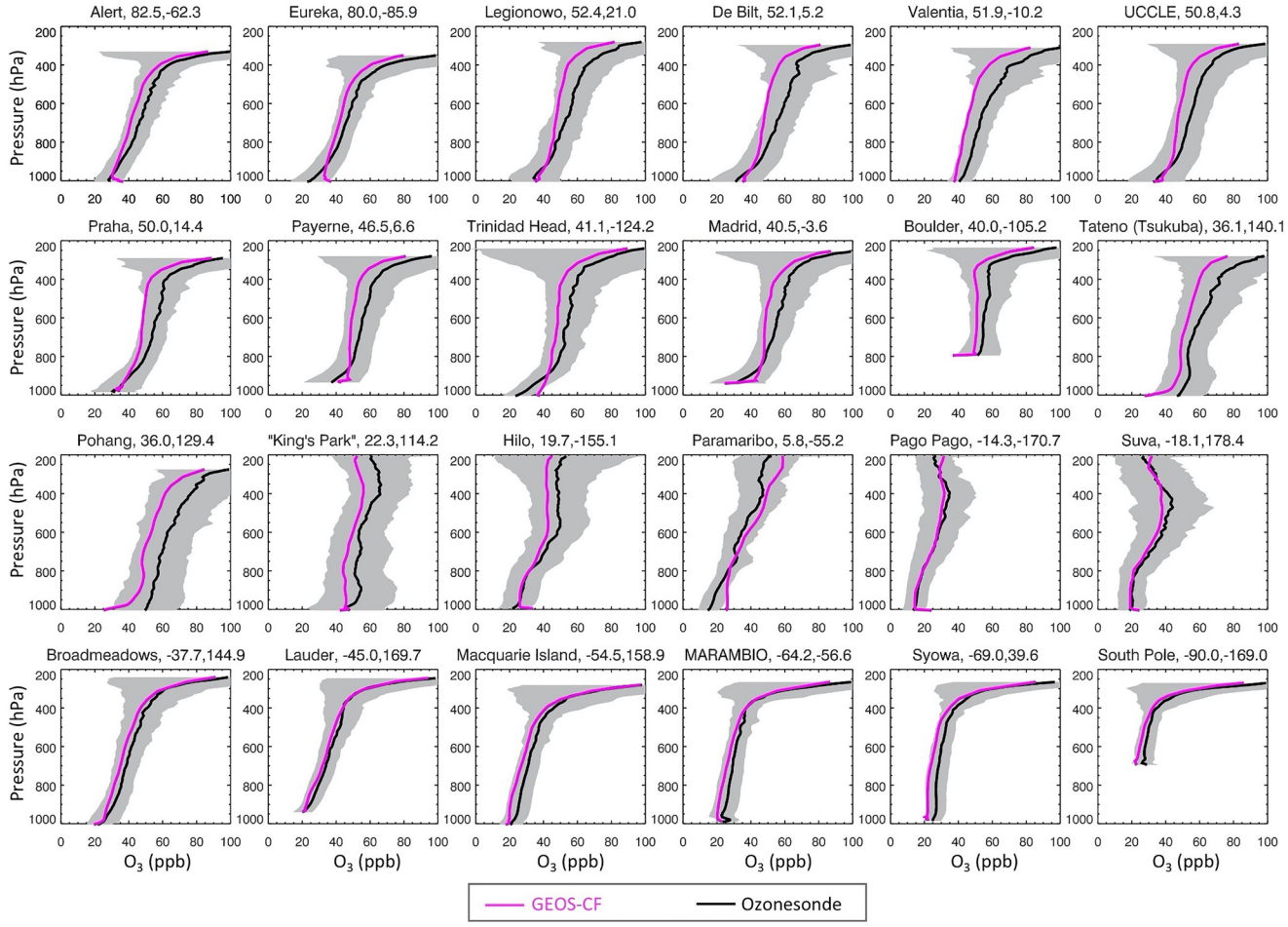

Figure 4. Vertical profile of annual average O_3 (2018–2019) at 24 sites as observed by ozonesondes (black) and simulated by GEOS-CF (magenta). The gray shaded area indicates sonde observation standard deviation.

at higher altitudes, as evident over Payerne, Trinidad Head, and Madrid (Figure 4); and (2) an underestimation of lightning NO_x over the NH midlatitudes, resulting in an underestimation of O_3 production in the upper troposphere. The global lightning NO_x source in GEOS-CF is 5.9 Tg N yr^{-1}, in good agreement with other studies (Schumann and Huntrieser, 2007) and the 6.0 Tg N yr^{-1} reported for the GEOS-Chem CTM (Murray et al., 2012). However, due to the real-time nature of the system, the GEOS-CF lightning parameterization uses the unconstrained cloud top height parameterization (Price and Rind, 1992, 1993, 1994) and does not apply time-dependent, regional redistribution factors based on Lightning Imaging Sensor (LIS) and the Optical Transient Detector (OTD) satellite observations as is standard in GEOS-Chem. As described in Murray et al. (2012), this results in an underestimation of simulated lightning flash rates over the NH extratropics, which is likely one of the main reasons for the model underestimation of O_3 in the NH upper troposphere. We also note that the model vertical resolution is approximately 500 m (20 hPa) in the mid-to upper troposphere, which might be insufficient to resolve the strong vertical gradients across the tropopause boundary and contribute to the model-observation mismatches in the upper troposphere. Over the Southern Hemisphere (SH), the simulated O_3 profiles are in good agreement with the ozonesonde observations and show an improved O_3 distribution compared to the offline GEOS-Chem model (Hu et al., 2018).

The GAW observations show a pronounced seasonal cycle for surface O_3 with a maximum of 30–50 ppbv during spring and summer and a low of 10–30 ppbv in winter, a feature that is well reproduced by GEOS-CF (Figure 5). At remote locations such as the high-latitude NH GAW sites (Figure 5 top row), the seasonal cycle is more distinct with a peak in early spring. The O_3 maximum develops more slowly at locations that are more heavily influenced by anthropogenic emissions. As already observed in the ozonesondes, the model

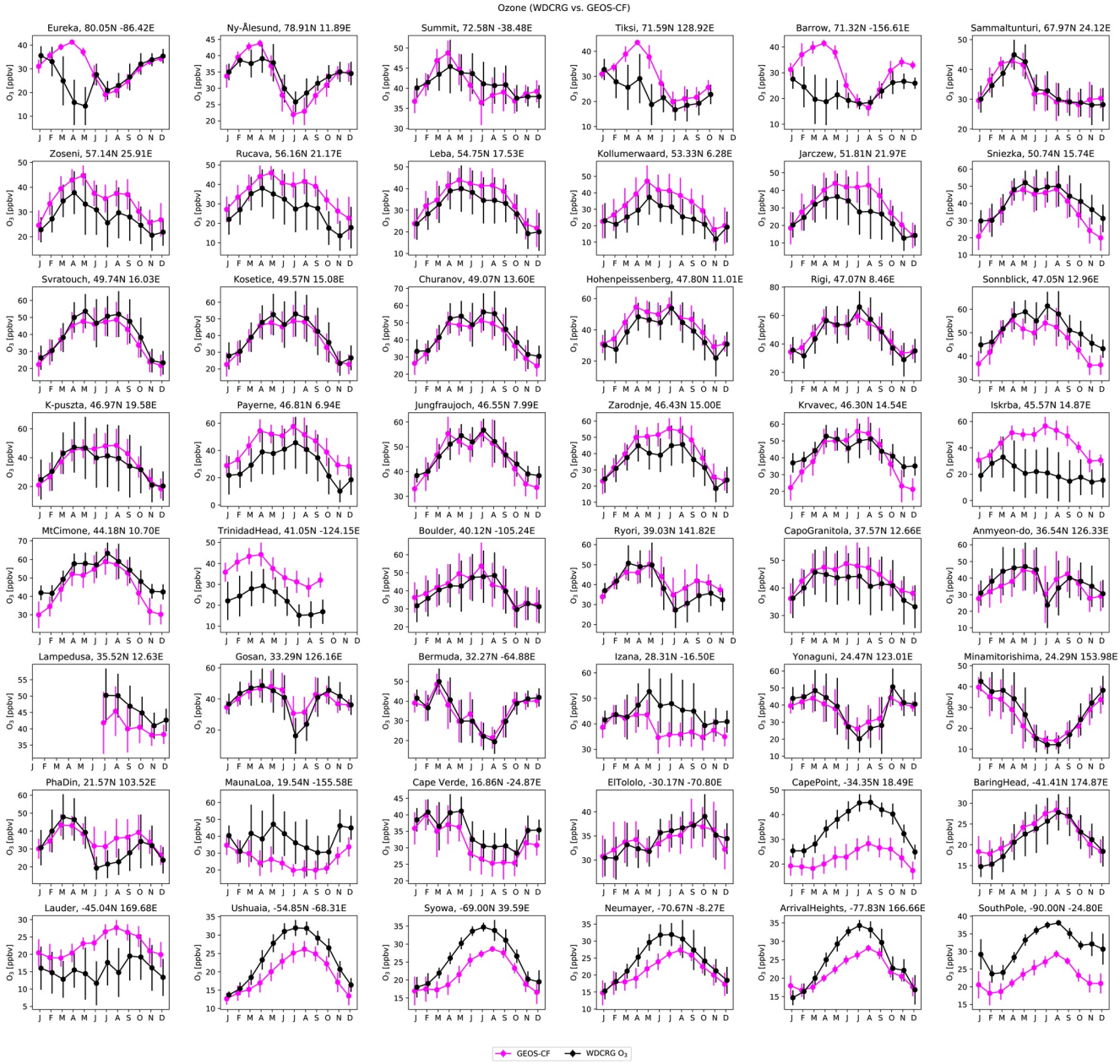

Figure 5. Monthly average surface O_3 as observed at 48 GAW sites (black) and simulated by GEOS-CF (magenta). Vertical bars represent the standard deviation of hourly variability. *Y*-axis ranges vary by station.

underestimates O_3 over the Southern Ocean by up to 10 ppbv. This is likely due to excessive O_3 deposition over seawater (Pound et al., 2020), a problem since corrected in newer versions of GEOS-Chem.

Figure 6 shows monthly average surface O_3 by region, as captured by observations in the OpenAQ database and simulated by GEOS-CF. The model generally overestimates surface O_3 over Europe, North and South America and Australia by 5–10 ppbv and underestimates it over Africa and Asia by up to 20 ppbv. The model-observation mismatch is larger than for the remote GAW sites, suggesting that sub-grid factors such as vertical and horizontal model representation errors and nighttime loss of O_3 by reaction with NO contribute to this difference (Dacic et al., 2020; Travis & Jacob, 2019). Global models tend to underestimate O_3 nighttime depletion due to difficulties to resolve small-scale horizontal NO_2 gradients, and evaluation of surface O_3 is thus often restricted to daytime or afternoon O_3 (Millet et al., 2015; Hu et al., 2018; Travis

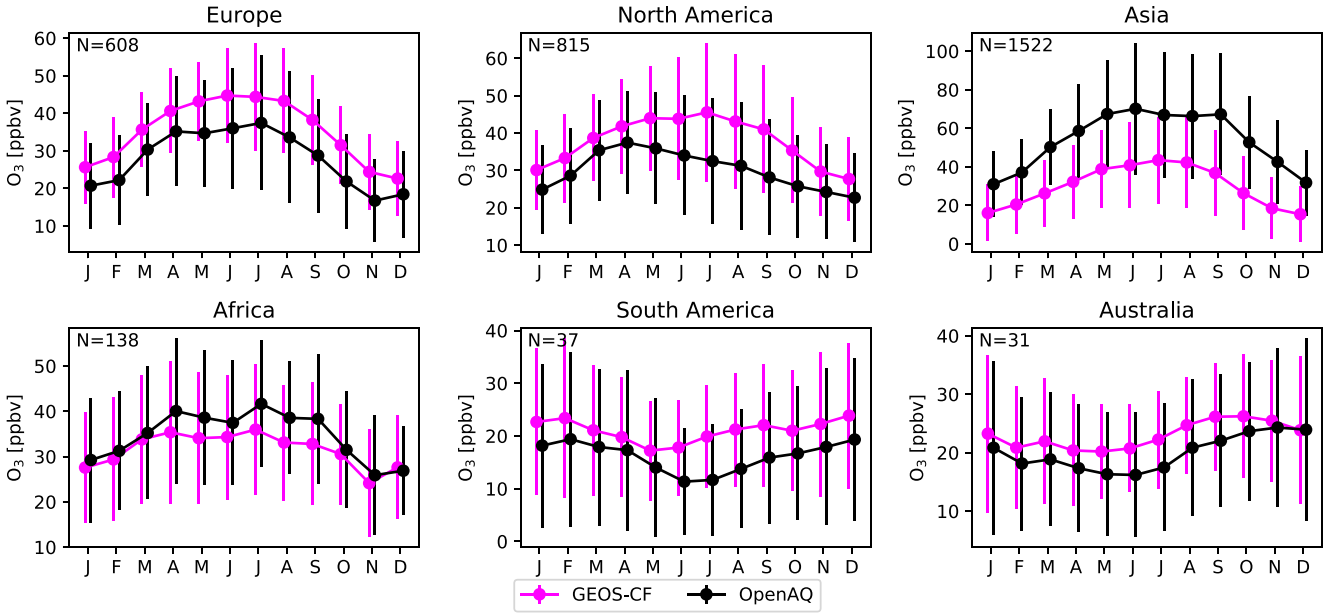

Figure 6. Monthly average surface O_3 grouped into six regions (see Figure 2) as obtained from the OpenAQ database (black) and simulated by GEOS-CF (magenta). Vertical bars represent the standard deviation of hourly variability across all sites. Number of sites is given in the inset. *Y*-axis ranges vary by region.

& Jacob, 2019). As shown in Figure 7, restricting the analysis to daytime O_3 (12:00–16:00 local time) does indeed improve the comparison and reduce the bias by up to 5 ppbv, in particular over Europe, South America and Australia. The model still shows a systematic positive bias over the US during summer and fall, a known issue in GEOS-Chem (Hu et al., 2018; Travis et al., 2016; Travis & Jacob, 2019). Results for the mean values of the daily maximum 8 h average (MDA8) instead of afternoon O_3 yield the same qualitative result (not shown).

Factors that likely contribute to the high NH surface O_3 in GEOS-CF are uncertainties in the production of O_3 from the oxidation of isoprene (Bates & Jacob, 2019; Travis et al., 2016) and errors in O_3 deposition to wet

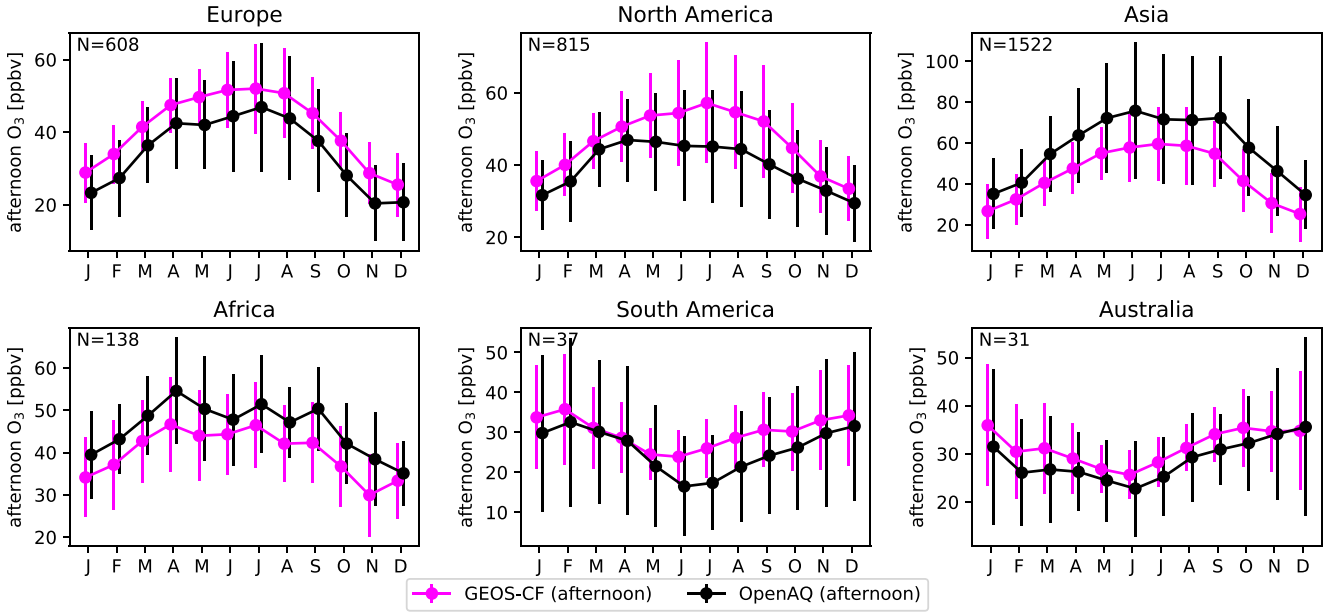

Figure 7. As Figure 6 but using afternoon O_3 only (12:00-16:00 local time).

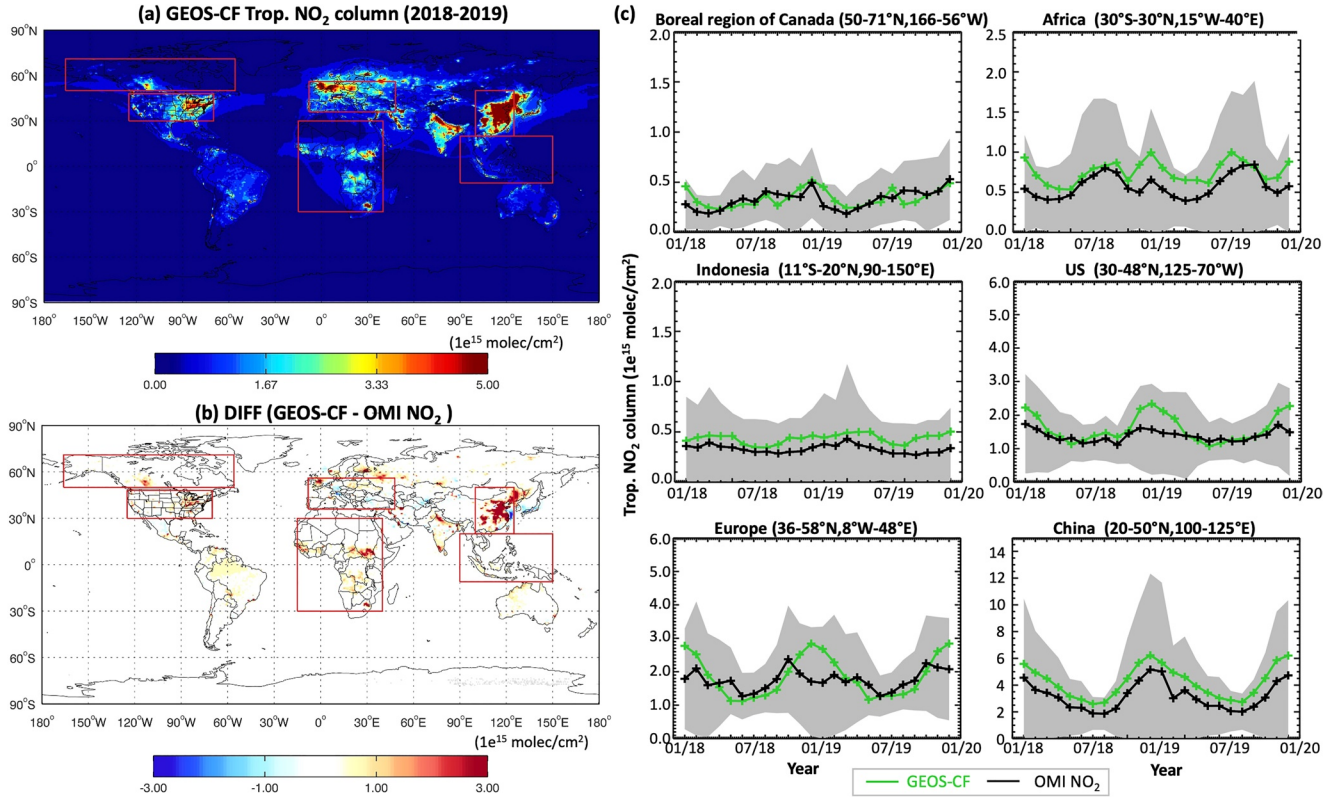

Figure 8. Comparison of GEOS-CF against OMI tropospheric column NO_2. (a) the 2018–2019 average tropospheric NO_2 column (1–2 pm local time) as simulated by GEOS-CF. (b) The difference between model simulated tropospheric column and NASA OMI NO_2 (v4.0) observations. (c) Right panel shows the tropospheric NO_2 time series averaged for six regions, as shown in (a) and (b) (n.b., *y*-axis intervals are not the same for each time-series).

surfaces (Travis & Jacob, 2019). In addition, most OpenAQ observation sites are located near densely populated areas, and the 25 × 25 km^2 model simulation cannot fully capture the fine-scale features characteristic for these environments (Keller et al., 2020).

3.3. Nitrogen Dioxide

As shown in Figure 8, GEOS-CF captures major features of the global distribution of tropospheric NO_2. The model simulated NO_2 columns (Figure 8a) show distinct hot spots over urban areas (e.g., Eastern China, Europe, and Eastern US), reflecting the dominant contribution of fossil fuel combustion coupled with the short atmospheric lifetime of NO_2 (Duncan et al., 2016; Streets et al., 2013;). Additional hot spots resulting from biomass burning are found over Africa and the boreal region of Canada (Figure 8a). The spatial pattern simulated by GEOS-CF agrees well with the NO_2 columns observed by OMI. Over Eastern China, the model simulated NO_2 columns are up to 3.0 × 10^{15} molecules cm^{-2} (or approx. 40%) higher than the OMI observations (Figure 8b), suggesting a potential overestimation of NO_x emissions or a longer NO_x lifetime in the model (Shah et al., 2020). However, the OMI retrieval algorithm v4.0 tends to underestimate tropospheric NO_2 over polluted areas (Lamsal et al., 2021), which complicates the analysis. As further discussed in the next section, the comparison against surface observations does not support the view of a broad-based overestimation of surface NO_2 over Asia.

As shown in Figure 8c, the simulated seasonality of tropospheric NO_2 columns is in good agreement with OMI observations. Over areas dominated by anthropogenic activities, such as the US, Europe, and China, the simulated NO_2 columns show a distinct seasonal cycle with the minimum during summer and peak during winter, driven by the seasonal variation in NO_x lifetime against oxidation (Shah et al., 2020). The seasonal cycle observed over China is well captured by GEOS-CF, while the simulated wintertime peak over the US and Europe is higher than observed by OMI. Chemical loss of NO_x during winter is dominated

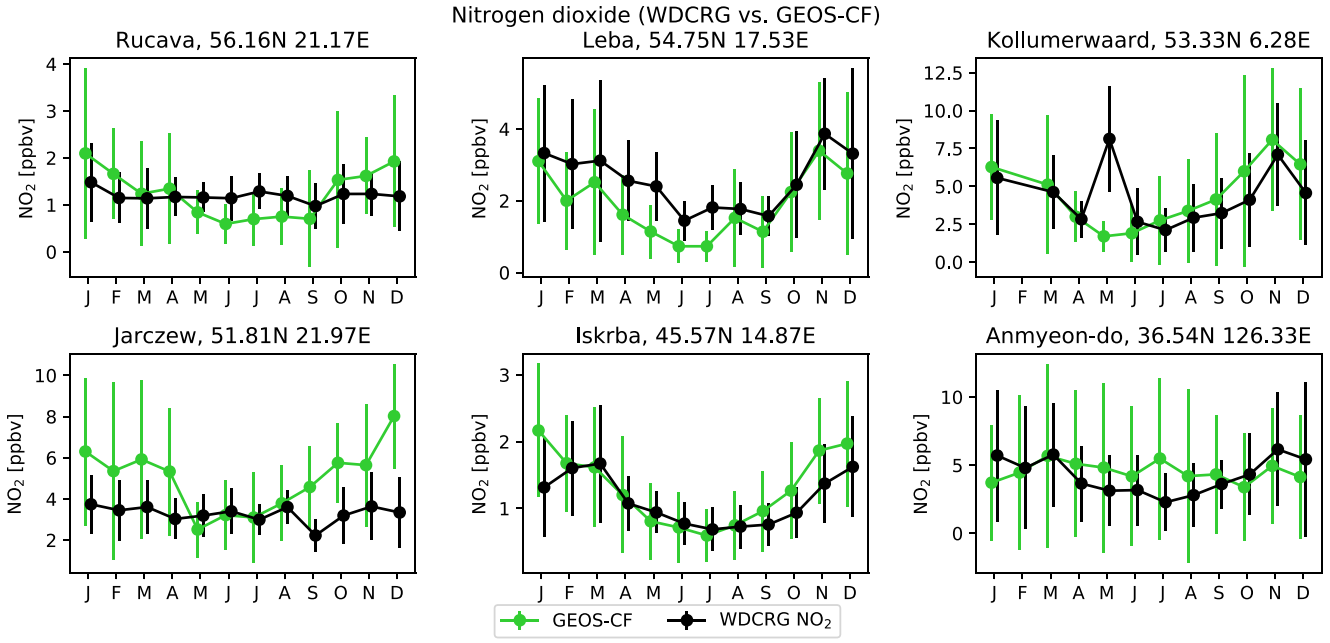

Figure 9. Monthly average surface NO_2 as observed at six GAW sites (black) and simulated by GEOS-CF (green). Vertical bars represent the standard deviation of hourly variability. *Y*-axis ranges vary by station.

by N_2O_5 hydrolysis, which is sensitive to NO_x emissions and O_3 concentrations (Jaeglé et al., 2018; Shah et al., 2020). Over Africa, the seasonal cycle is dominated by summer biomass burning, which is well captured by the model.

Figure 9 shows comparisons of monthly mean surface NO_2 at six GAW sites (n.b., all located in Europe) against model simulated concentrations. These comparisons show that GEOS-CF is in good agreement with the (non-mountainous) GAW sites, suggesting that it provides a realistic simulation of background surface NO_2 over Europe. Figure 10 further evaluates the simulated distribution of global surface NO_2 in

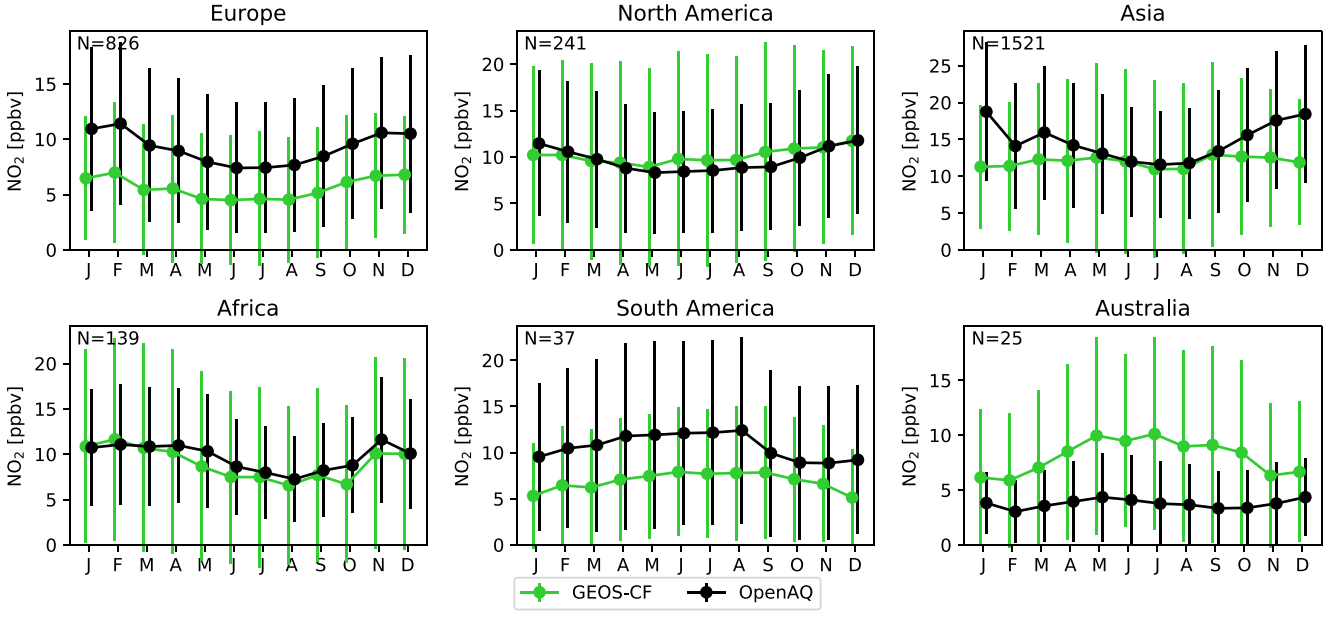

Figure 10. Monthly average surface NO_2 grouped into six regions (see Figure 2) as obtained from the OpenAQ database (black) and simulated by GEOS-CF (green). Vertical bars represent the standard deviation of hourly variability across all sites. Number of sites is given in the inset. *Y*-axis ranges vary by region.

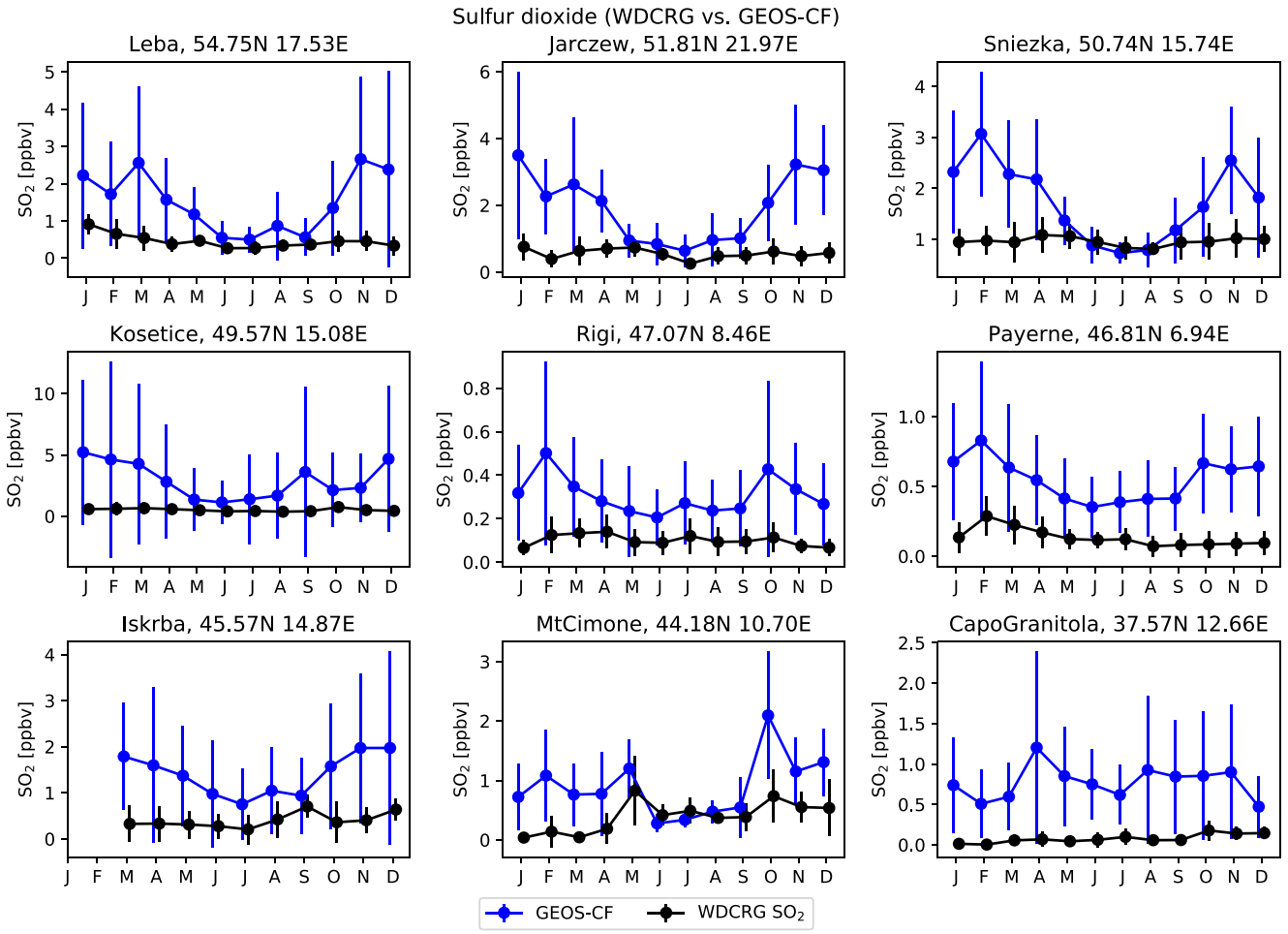

Figure 11. Monthly average surface SO_2 as observed at nine GAW sites (black) and simulated by GEOS-CF (blue). Vertical bars represent the standard deviation of hourly variability. *Y*-axis ranges vary by station.

comparison to observations in OpenAQ. The model is in excellent agreement with observations over North America and Africa, and Asia during the summertime, but underestimates concentrations over Europe and South America, as well as Asia during the wintertime. The apparent low bias over Europe and Asia is inconsistent with the tropospheric column comparisons shown in Figure 8 and also the comparison against the GAW observations (Figure 9), which do not show such a systematic underestimation of NO_2 by GEOS-CF.

The comparison of surface concentrations of NO_2 is complicated by the fact that most surface observations are based on chemiluminescence using a molybdenum converter, which can result in an overestimation of reported NO_2 concentrations due to interference with other oxidized nitrogen compounds (Grosjean and Harrison, 1985; Steinbacher et al., 2007; Winer et al., 1974). This might explain some of the model underestimation of NO_2 relative to the OpenAQ observations. In addition, since the OpenAQ observations tend to be located in relative proximity to urban areas they often do not represent the regional average NO_2 concentrations captured by GEOS-CF, which can introduce a systematic bias. While this is an issue for all analyzed species, the problem is particularly pronounced for NO_2 given its large spatial and temporal variability.

3.4. Sulfur Dioxide

GEOS-CF overestimates surface SO_2 concentrations compared against observations from the GAW network (Figure 11) as well as the OpenAQ database (Figure 12). The GAW sites are all located in Europe, and the model has a mean bias of 0.9 ppbv across the nine stations. Against the low average background concentration of 0.5 ppbv observed at the GAW sites, this indicates an overestimation by the model by a factor of more

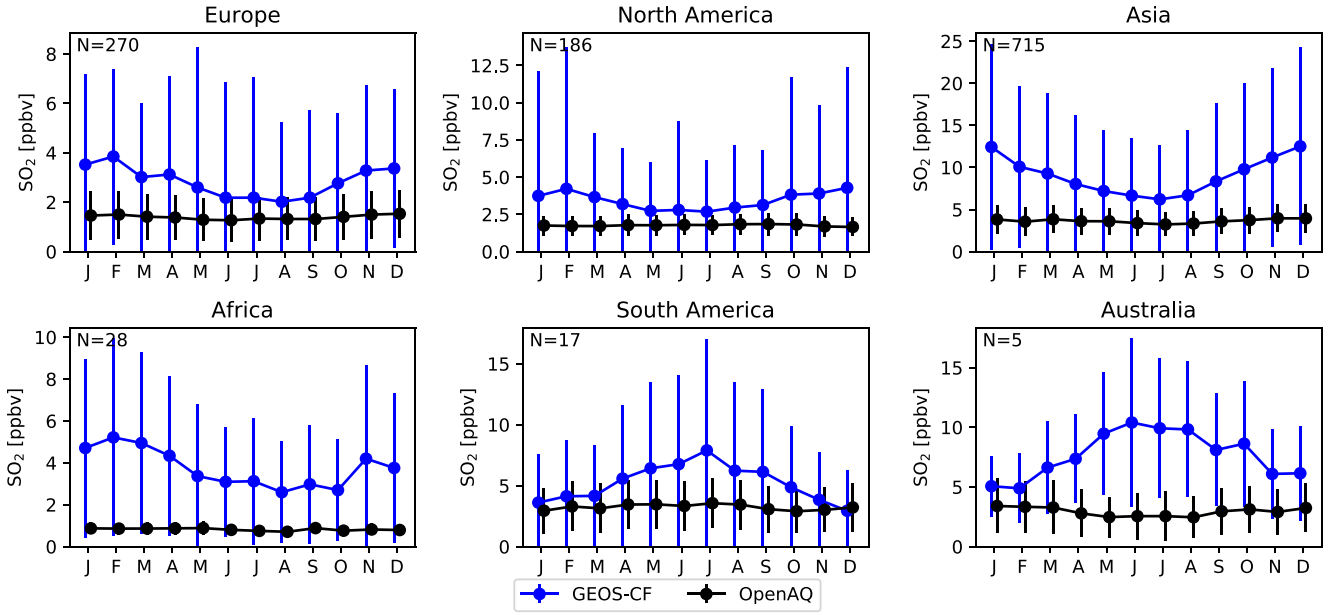

Figure 12. Monthly average surface SO_2 grouped into six regions (see Figure 2) as obtained from the OpenAQ database (black) and simulated by GEOS-CF (blue). Vertical bars represent the standard deviation of hourly variability across all sites. Number of sites is given in the inset. *Y*-axis ranges vary by region.

than two. A similar bias is found at the OpenAQ sites. We attribute the high model bias in SO_2 to a large extent the use of outdated emissions from the OMI-HTAP emission inventory. Anthropogenic emissions in OMI-HTAP are based on HTAP v2.2, which is representative for year 2010 (Liu et al., 2018). Since then, anthropogenic emissions of SO_2 have declined significantly in most parts of the world (Feng et al., 2020; Fioletov et al., 2016; Li et al., 2020; Qu et al., 2019). The model overestimation is particularly severe during wintertime, when the atmospheric lifetime of SO_2 is longest (Lee et al., 2011) and the majority of anthropogenic SO_2 is emitted (according to OMI-HTAP). Another likely explanation is that the model may be under-accounting for non-photochemical SO_2 oxidants in winter (Shao et al., 2019).

3.5. Carbon Monoxide

Our analysis of simulated CO shows that the model has no significant global CO bias compared against satellite and surface observations. Figure 13 shows the global distribution and seasonal cycle of total column CO retrieved from MOPITT and derived from GEOS-CF. Local MOPITT averaging kernels were applied to the GEOS-CF CO profiles to obtain the column values. The model simulated CO pattern is in good agreement with MOPITT, with local maxima over major polluted areas (e.g., East China and India) and over biomass burning regions (South America and Central Africa). Similar to NO_2, the simulated CO columns over China are larger than the observations, possibly due to an overestimation of CO emissions over that region. The baseline anthropogenic CO emissions in GEOS-CF are scaled based on ODIAC emission trends, with strongest increases over China and India (Oda et al., 2018). Our scaling methodology assumes a constant CO/CO_2 emissions ratio, and any decrease in the CO/CO_2 emission ratio, for example, due to a technology shift, would result in an overestimation of CO emissions.

The seasonal cycle of CO is determined by its shorter chemical lifetime during summer due to increased photochemical activity, which is reflected in lower NH concentrations during July and August where the model underestimates the MOPITT concentrations in the NH middle and high latitudes by 10%–20% (Figure 13c). This underestimation is driven by a stronger than observed decrease of simulated total column CO during summertime, a pattern that is confirmed by comparisons against the GAW surface observations (Figure 14). While the model underestimates summertime surface CO in the NH, the opposite is true for the SH where model simulated concentrations during summer are higher than the observations. A low bias in CO is a long-standing issue in atmospheric chemistry models, commonly attributed to an underestimation of direct emissions, differences in the CO yield from VOC oxidation, and inconsistencies in the simulated

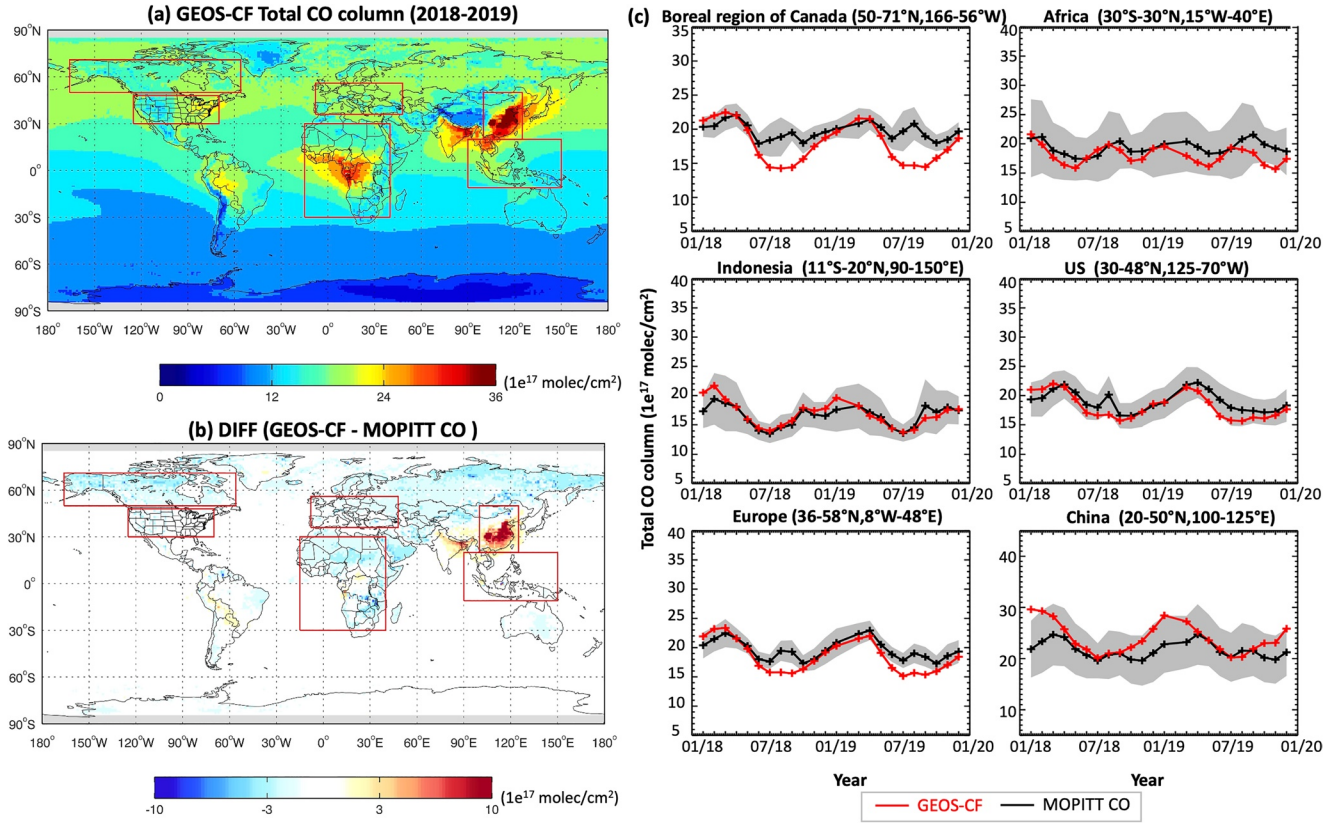

Figure 13. Comparison of GEOS-CF against MOPITT V8 total column CO. Top left panel shows the 2018–2019 average total CO column as simulated by GEOS-CF. Bottom left panel shows the difference between model simulated total column and MOPITT observations. Right panel shows the total CO time series averaged for six regions, as shown on the left.

distribution of the hydroxyl radical OH, the main atmospheric oxidant (Flemming et al., 2015; Gaubert et al., 2016; Monks et al., 2015; Shindell et al., 2006; Strode et al., 2015). The air mass-weighted mean tropospheric OH in GEOS-CF is 11.9×10^5 molecules cm^{-3}, in good agreement with other models as well as estimates derived from methyl chloroform observations (Montzka et al., 2000; Naik et al., 2013; Spivakovsky et al., 2000). The inter-hemispheric (NH/SH) ratio is 1.33, again in good agreement with other model estimates (Naik et al., 2013) but higher than observation-derived values of 0.8–1.0 (Krol and Lelieveld 2003; Montzka et al., 2000; Patra et al., 2014; Prinn et al., 2001). This is consistent with an underestimation of summertime CO in the northern high-latitudes—as shown by the comparisons against MOPITT and the GAW surface observations (Figures 13 and 14)—and an overestimation of SH CO (Figure 14). An overestimation of NH OH in GEOS-CF is also supported by inversion studies using the GEOS-Chem CTM, which find that a similar OH correction is needed to match methane observations (Zhang et al., 2020).

3.6. Aerosols

The evaluation of model simulated O_3, NO_2 and CO has shown that GEOS-CF reproduces many of the features of the tropospheric distribution of these trace gases. With respect to aerosols simulated by GEOS-Chem, our validation shows that GEOS-CF has a high bias but still captures many of the observed spatial and temporal patterns. Some of the high aerosol bias can be attributed to the high SO_2 in GEOS-CF (Section 4.4), resulting in excessive formation of sulphate. In addition, a high bias in aerosols is a known issue in GEOS-Chem v12.0.1 used in GEOS-CF v1.0, and recent versions of GEOS-Chem show improved simulation of aerosols including surface $PM_{2.5}$ (Carter et al., 2020; Luo et al, 2019, 2020).

Figure 15 shows the global distribution and seasonality of simulated AOD at 550 nm against Aqua MODIS satellite observations. GEOS-CF captures the seasonal cycle of AOD well but tends to overestimate the

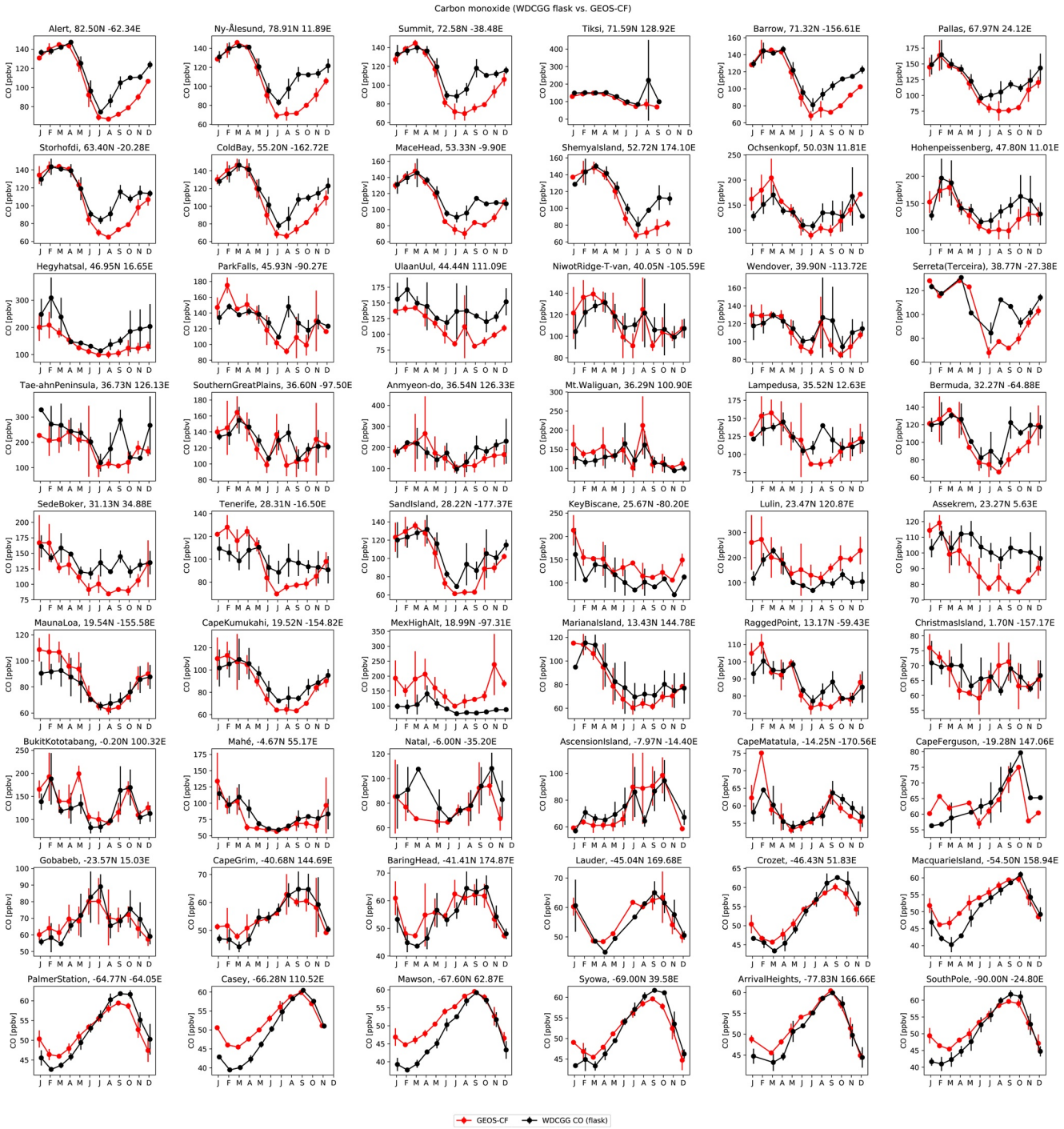

Figure 14. Monthly average surface CO as observed at 54 GAW sites (black) and simulated by GEOS-CF (red). Vertical bars represent the standard deviation of hourly variability. *Y*-axis ranges vary by station.

aerosol burden, especially over the US, Europe, and China. These regions have seen large declines in SO_2 emissions since 2010 that are not captured by GEOS-CF and we attribute the high model AOD at least partly to the excessive formation of sulphate aerosols. Further, inadequate treatment of wet scavenging processes is a known issue in GEOS-Chem v12.0.1, resulting in an overprediction of aerosol nitrate and ammonium, especially over Asia (Luo et al., 2019, 2020).

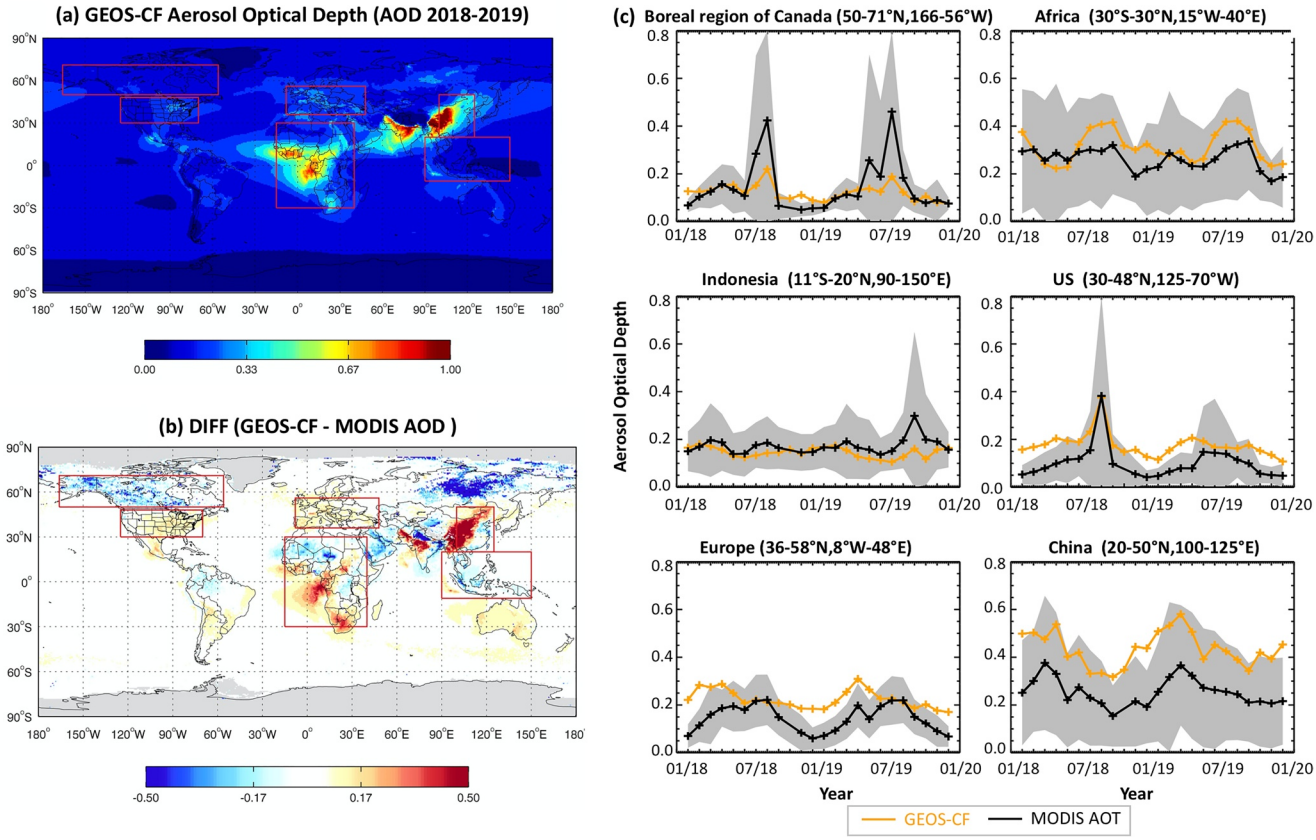

Figure 15. Comparison of GEOS-CF against MODIS Aerosol Optical Depth (AOD). Top left panel shows the 2018–2019 average total AOD as simulated by GEOS-CF. Bottom left panel shows the difference between model simulated total column and MODIS observations. Right panel shows the total AOD time series averaged for six regions, as shown on the left.

GEOS-CF tends to underestimate AOD related to strong wildfires events, such as over Canada during the summers of 2018 and 2019, the US in 2018, or Indonesia in 2019 (Figure 15). This points to a potential underestimation of severe wildfire emissions by QFED, possibly due to the thick smoke obscuring fires from satellite detection (Eck et al., 2019; Liu et al., 2020). However, the QFED emissions inventory has been shown to lead to an overall high aerosol bias when used in GEOS-Chem as the enhancement factors used in QFED (to match MODIS-observed AOD) are based on the GOCART model (Darmenov and Da Silva, 2015), which can result in an overestimation of biomass burning emissions if the differences between the GEOS-Chem and GOCART aerosol parameterizations are not accounted for (Carter et al., 2020). This factor likely contributes to the overestimation of AOD over Africa and South America.

Figure 16 compares model simulated AOD at 550 nm against AERONET observations, confirming the systematic model overprediction of AOD. The model bias is highest over Europe and North America, consistent with an overprediction of sulphate aerosols due to excessive SO_2 emissions in GEOS-CF. To provide an upper bound on the contribution of the SO_2 error to total AOD, we analyzed GEOS-CF AOD excluding sulphate. As shown by the dashed orange line in Figure 16, this indicates that sulphate indeed constitutes up to 50% of total AOD over Europe, North America, and Asia, while its contribution to total aerosol is smaller in other regions.

Figure 17 shows model-observation comparisons for surface $PM_{2.5}$. As for AOD, GEOS-CF generally overestimates surface $PM_{2.5}$. On a relative basis, the overprediction is most pronounced for Europe and North America, and the largest absolute bias of surface $PM_{2.5}$ is found over Asia.

While we conclude that the main reasons for the high aerosol bias in GEOS-CF is related to three main factors—excessive sulphate aerosols, an underestimation of aerosol wet scavenging, and overestimated biomass burning emissions outside of high-intensity wildfire episodes—there are other factors that could

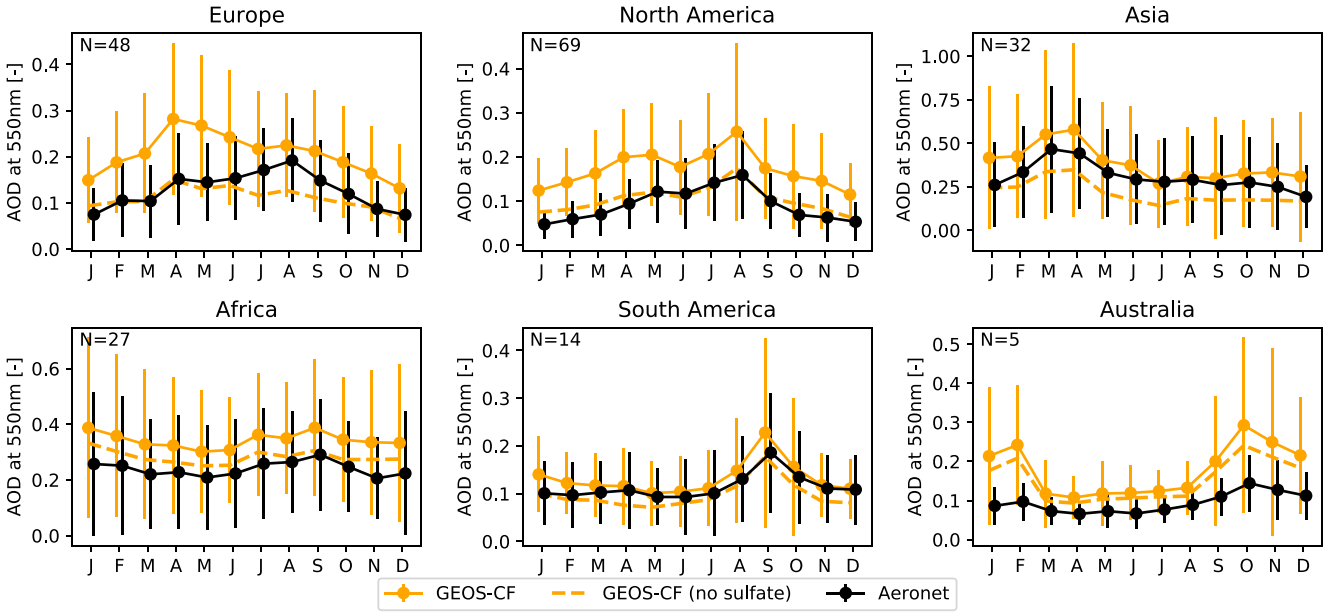

Figure 16. Monthly average Aerosol optical depth (AOD) at 550 nm grouped into six regions (see Figure 2) as observed at AERONET sites (black) and simulated by GEOS-CF (orange). Vertical bars represent the standard deviation of daily variability. Dashed line indicates GEOS-CF AOD when excluding sulphate aerosols. Number of sites is given in the inset. *Y*-axis ranges vary by region.

contribute to the observed model-observation mismatches. For instance, a general overestimation of wildfire emissions could directly affect modeled organic aerosol as well as the formation of secondary organic aerosol (SOA). Organic aerosol constitutes an important fraction of total aerosol mass—also in GEOS-Chem (Kim et al., 2015)—but its sources and atmospheric processes are complex and poorly constrained (e.g., Bates et al., 2019; De Gouw and Jimenez, 2009; Marais et al., 2016). Biases in dust emissions could be another contributor to model biases, especially over Africa. The dust emissions calculated by GEOS-CF total

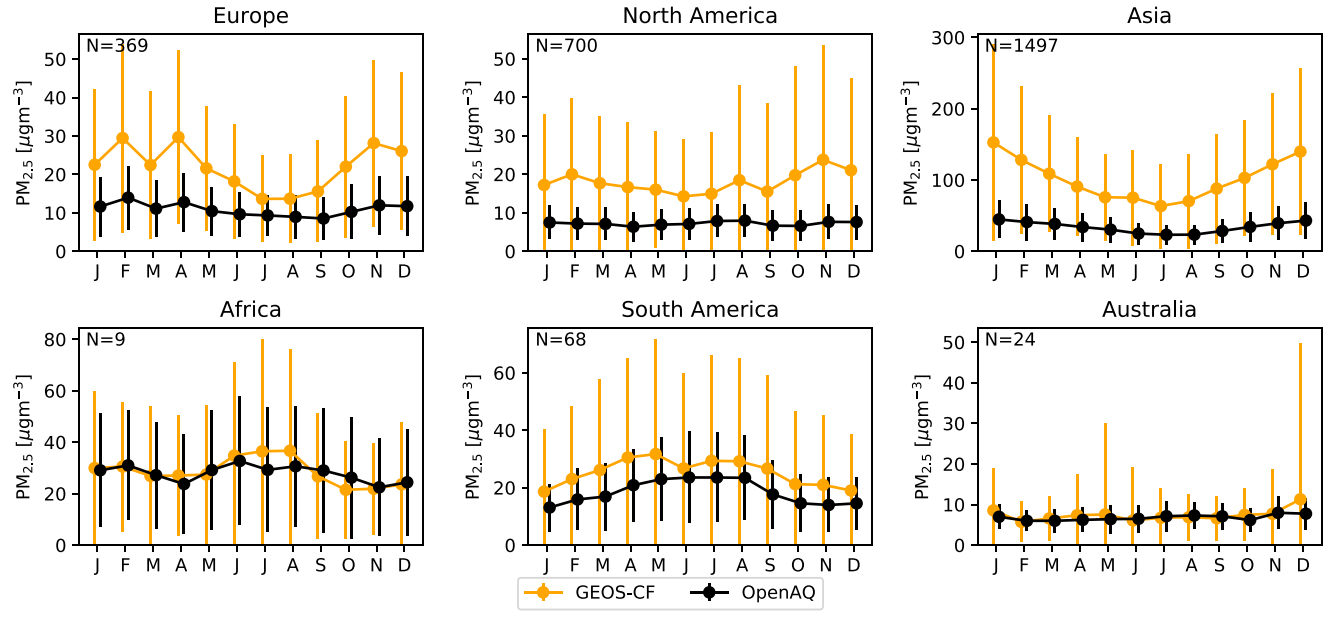

Figure 17. Monthly average surface $PM_{2.5}$ grouped into six regions (see Figure 2) as obtained from the OpenAQ database (black) and simulated by GEOS-CF (orange). Vertical bars represent the standard deviation of hourly variability. Number of sites is given in the inset. *Y*-axis ranges vary by region.

770 Tg/yr and 630 Tg/yr for 2018 and 2019, respectively, which is at the lower range of current estimates of 426–2726 Tg/yr (Huneeus et al., 2011) and thus unlikely a major contributor to the high aerosol bias.

4. Evaluation of Model Forecasts

4.1. Comparison of Model Forecasts Against Observations

GEOS-CF v1.0 does not directly assimilate trace gas observations and differences between the 1-day hindcast and the model forecasts are thus mainly driven by variations between the forecasted and analyzed meteorological state. The meteorology not only impacts the flow of the constituents but also affects deposition and dynamically calculated emissions, including lightning NO_x, biogenic VOCs, sea salt aerosols, and dust (see Table 1). Further, the model forecasts assume persistence in the biomass burning emissions, meaning the fires observed during the hindcast are assumed to continue burning and emitting the same amount for the next five days. The model thus does not capture changes in occurrence or intensity of wildfires, both of which can lead to significant changes in surface air pollution close to and downwind from the fires.

Figure 18 shows model-observation skill scores for the model hindcast (forecast day −1) and the 5-days forecasts (forecast days +1 to +5) for O_3, NO_2, and $PM_{2.5}$ relative to OpenAQ and GAW surface observations. Skill scores were calculated at each observation site individually using daily mean concentrations. Similar to the results discussed in Section 4.1 (Figure 3), the skill scores were computed for each month separately before aggregating them in the form of boxplots (as shown in Figure 18) in order to remove the impact of seasonality. For all three evaluated species, the model hindcast shows the best agreement with the observations. The median NMB of the model hindcast is 0.1 for O_3, -0.31 for NO_2, and 1.56 for $PM_{2.5}$, with almost no difference between the hindcast and the model forecasts. Similarly, the forecasted NRMSE are little changed relative to the hindcast; the median NRMSE for the hindcast is 0.61 for O_3, 0.63 for NO_2 and 1.61 for $PM_{2.5}$, and the NRMSE only slightly deteriorates with increasing forecast lead time, resulting in NRMSE's for the 5-days forecast of 0.62 for O_3, 0.65 for NO_2 and 1.65 for $PM_{2.5}$.

The largest change in skill score between hindcast and forecasts is found for R, which drops from 0.49 for the hindcast to 0.38 for the 5-days forecast for O_3, 0.39 to 0.26 for NO_2, and 0.52 to 0.36 for $PM_{2.5}$. This indicates that errors in the meteorological forecasts (and biomass burning emissions) indeed impact the quality of the surface air quality forecasts. Of the three analyzed skill scores, R is most sensitive to errors in the temporal pattern and we attribute a large fraction in the deterioration in R to transport errors in the forecasts, such as the evolution of frontal systems or the dispersion of smoke plumes. $PM_{2.5}$ is particularly sensitive to these factors given its large spatiotemporal gradients, comparatively long atmospheric lifetime, and strong sensitivity to changes in biomass burning emissions.

4.2. Bias-Corrected Local Forecasts

As discussed above, model-predicted concentrations of O_3, NO_2 and $PM_{2.5}$ can differ from the observations for a number of reasons, including model representation errors, uncertainties in the meteorology, or model biases arising from errors in the model treatment of emissions, deposition, or atmospheric chemistry. One approach to deal with these issues is to quantify and correct these systematic model errors in a post-processing step. Such bias correction methods can be applied to near real-time model forecasts and have been found to be an effective tool to significantly improve local model predictions, for example, by using mean subtraction (McKeen et al., 2005; Wilczak et al., 2006), historical analogs (Hamill and Whitaker, 2006), Kalman-filtering (Delle Monache et al., 2006; Djalalova et al., 2015), or kriging (Honoré et al., 2008). More recently, machine learning (ML) approaches have become popular to relate model output to air quality observations (e.g., Grange et al., 2018; Grange and Carslaw, 2019; Ivatt and Evans, 2020; Petetin et al., 2020). As discussed in Keller et al. (2020), bias-correction using ML can significantly reduce GEOS-CF model biases compared to surface observations. This is illustrated in Figure 19, which shows the GEOS-CF model skill scores for the same surface observations analyzed above (GAW and OpenAQ) but using bias-corrected model concentrations instead of the original model output. The bias correction methodology is described in detail in Keller et al. (2020). It uses the XGBoost algorithm (Chen and Guestrin, 2016) to correct the original hourly-average model predictions of O_3, NO_2, and $PM_{2.5}$ based on local meteorology and atmospheric composition, as predicted by the GEOS-CF model. The input features fed into the ML algorithm include 9 meteorological pa-

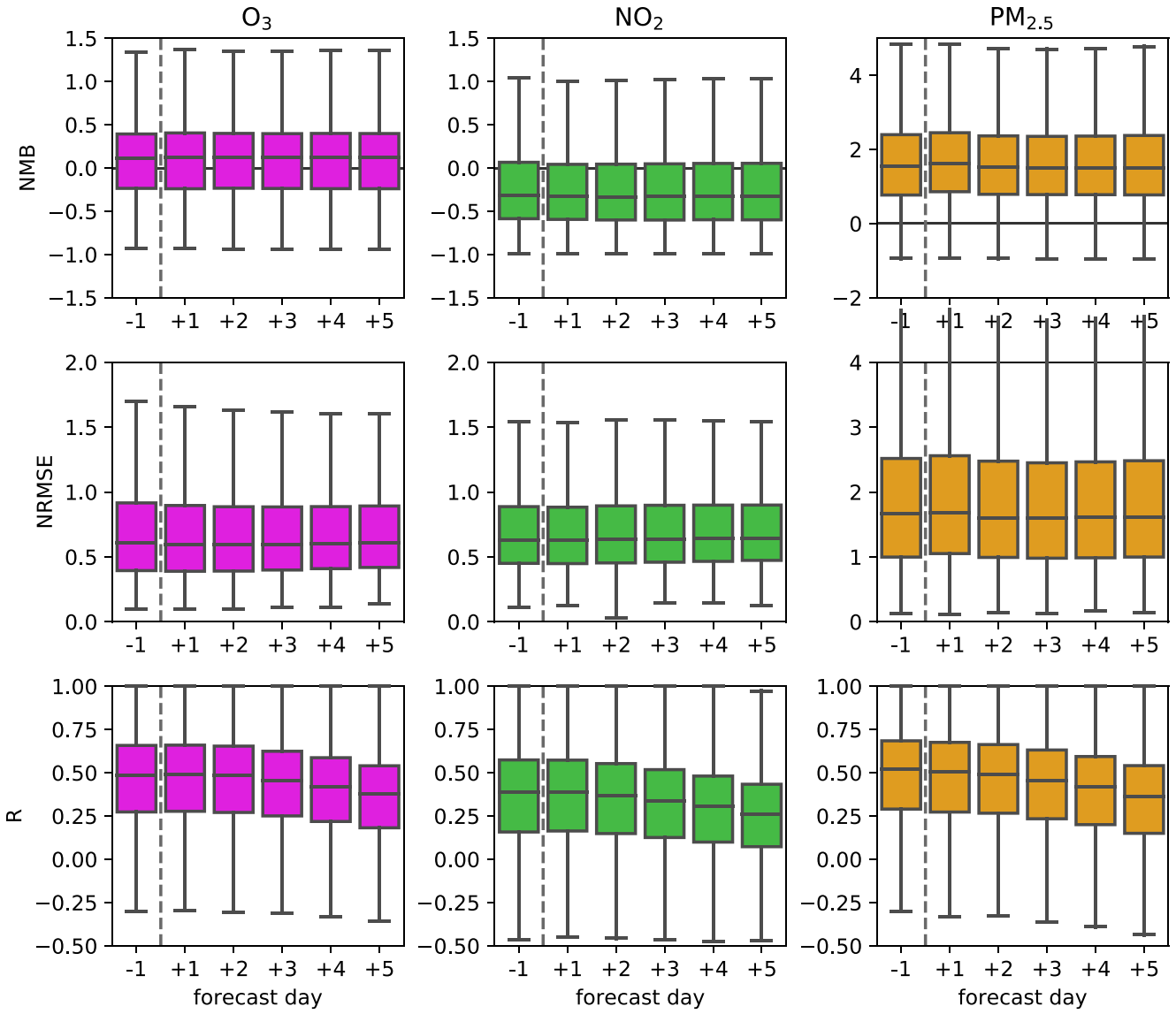

Figure 18. GEOS-CF model skill scores for the hindcast (forecast day −1) and the 5-days forecasts (forecast day +1 to +5) relative to surface observations (GAW and OpenAQ). Boxplots show the variation in the NMB, NRMSE, and *R* across all surface sites for daily mean O_3, NO_2, and $PM_{2.5}$.

rameters, 51 chemical species, 21 modeled emissions, and calendar information including hour of the day, weekday, and month of the year. The latter allow the ML model to identify systematic observation-model mismatches related to the diurnal, weekly, and seasonal cycle of the pollutants. Since model biases are often site-specific—especially the representation error—a separate ML model is trained for each site. Using the hourly model hindcast—observation pairs of year 2018, the ML models are trained to predict the systematic bias between model predicted pollutant concentrations and the actual observation. We then apply the ML algorithm to the 2019 model output and compare the resulting bias-corrected model concentrations against observations (Figure 19). It is important to emphasize that the ML correction is site-specific and varies as a function of local meteorological and chemical conditions, as well as time of the year, weekday, and hour of day. As such, it it designed to provide a "localized" model prediction only representative for the given observation site, not the 25×25 km^2 grid cell as a whole.

Compared to the uncorrected model output (indicated by the diamonds in Figure 19), the bias-corrected model values agree better with the observations for all species, skill scores, and lead times. For all three species, the NMB of the bias-corrected values is close to zero for both the hindcast and the 5-days forecasts.

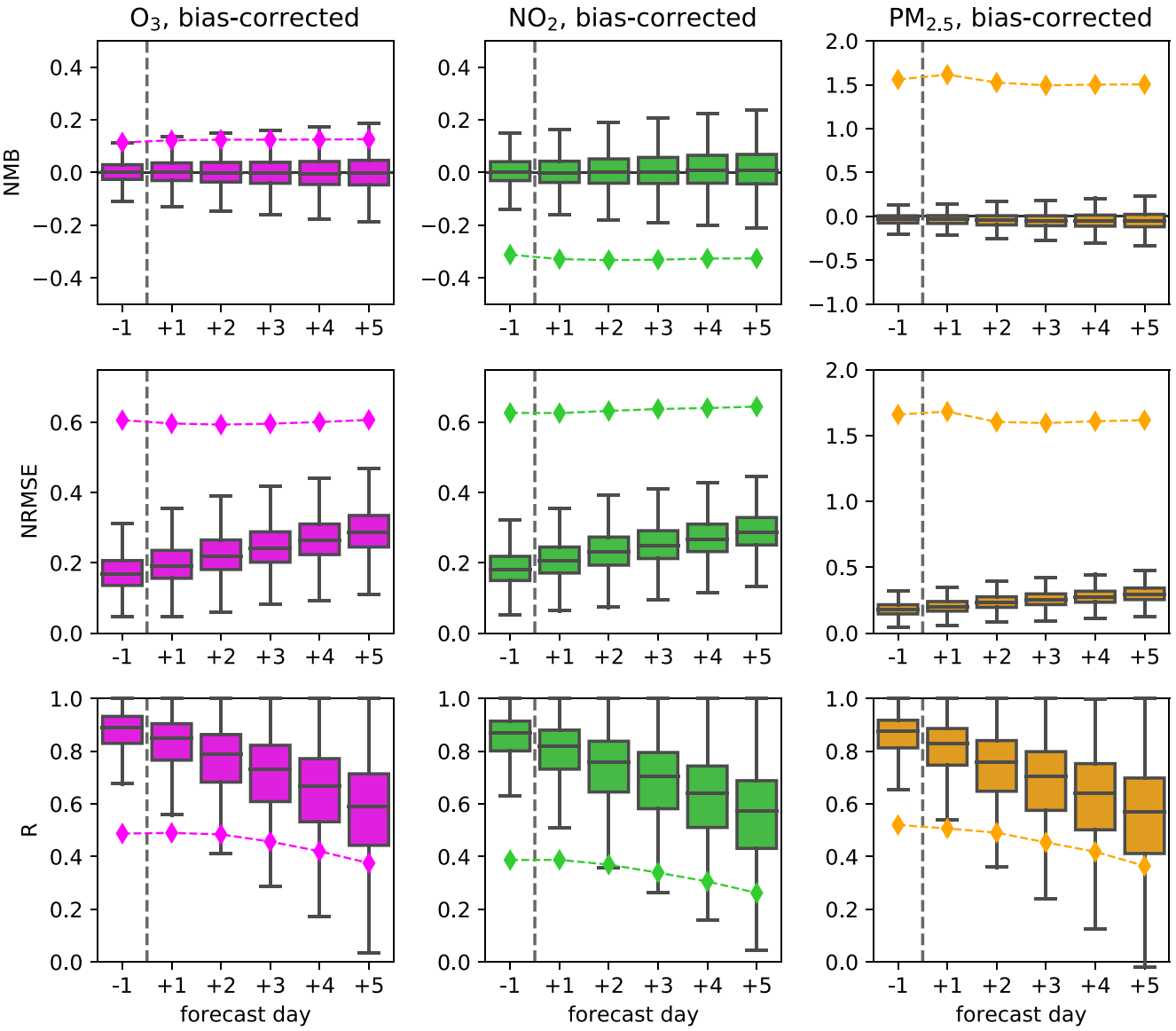

Figure 19. Similar to Figure 18, however the boxplots indicate the GEOS-CF model skill scores for the bias-corrected hindcast (forecast day −1) and the bias-corrected 5-days forecasts (forecast day +1 to +5) relative to surface observations. The median statistics of the non-corrected model values from Figure 18 shown as diamonds.

This result is not unexpected given that the ML algorithm is designed to minimize the mean model bias. The NRMSE and *R* scores are also significantly improved compared to the original data, with median RMSE values of approximately 0.2 and *R* values of 0.9 for the hindcast. Relative to the hindcast, the NRMSE and *R* become incrementally worse for the 1-day to 5-days lead forecasts relative to the actually observed values. Also, the skill scores of the bias-corrected forecasts deteriorate more rapidly than is the case for the uncorrected output but still outperform the original forecasts for all five lead days. Since the ML algorithm was only trained on the hindcast data, model errors in the meteorology forecasts will also negatively impact the quality of the bias correction applied to the (forecast) baseline, which explains the steady decline in the skill scores for the bias-corrected forecasts.

5. Conclusions

The GEOS-CF system provides global, near real-time estimates of current conditions ("hindcast") and 5-days forecast simulations of atmospheric composition and meteorology at the global resolution of 0.25° based on the GEOS-Chem chemistry module run within the GEOS GCM. Comparisons against a suite of satellite, ozonesonde and surface observations demonstrate that the model realistically captures the global distribution of O_3, NO_2, and CO. For these three species, the model shows little biases at background locations, with NMB values ranging between −0.1 to +0.1. Like other atmospheric chemistry models, the model overestimates surface O_3 in the NH during summer, especially over the south eastern US (Emmons et al., 2020; Flemming et al., 2015; Travis et al., 2016). While the horizontal resolution of 0.25° is one of the highest for a global atmospheric chemistry forecast model, it is still not fine enough to capture local-scale features of urban air pollution, which can lead to substantial model-observation mismatches when compared against urban air quality observations. This is reflected in the poorer model skill against the OpenAQ observations, which are typically located in the proximity of urban areas. The model representation error is an important limitation of global composition models such as GEOS-CF and needs to be taken into account for air quality applications. One possible solution to this problem is to apply a bias-correction algorithm to the gridded model output, based on ML and historical observation-model comparisons. This can lead to significant improvements of the model predictions. This procedure, which can be automated as a post-processing step, offers the option to provide highly localized forecasts anywhere in the world where surface observations are available. However, it should be emphasized that such localized forecasts are only valid for the given observation site and do not represent a model grid cell average (25×25 km^2) anymore.

The 5-days forecasts produced by the GEOS-CF system have comparable NMB and NRMSE as the 1-day hindcast (compared to hourly observations), but show a deterioration in temporal correlation with increasing forecast lead time. The current version of GEOS-CF (v1.0) does not directly assimilate trace gas observations, and the reduction in correlation skill score is thus directly related to errors in meteorological forecasts—which feed back on chemistry, deposition, and dynamically updated emissions—as well as errors in the wildfire emission forecasts which are assumed to persist over the entire 5-days forecast window.

GEOS-CF v1.0 systematically overpredicts SO_2 due to the use of outdated emissions data, highlighting the critical need of timely emission information. In it's current form, GEOS-CF does not account for real-time changes in anthropogenic emissions, for example, due to reduced human activities in the wake of the COVID-19 pandemic. GEOS-CF v1.0 thus represents a business-as-usual estimate of the atmosphere (Keller et al., 2020). The inclusion of near real-time information for dynamic emission adjustment, for example, based on traffic data or satellite observations, is a focus of future model development for GEOS-CF.

The GEOS-CF modeling system leverages components developed by the GEOS and the GEOS-Chem modeling communities and directly benefits from the continuous advancements provided by these groups. The current version of GEOS-CF (v1.0) incorporates GEOS-Chem v12.0.1. Several important updates have been added to GEOS-Chem since, including updates to the chemistry of isoprene (Bates and Jacob, 2019) and halogens (Wang et al., 2019), improved wet scavenging of aerosols (Luo et al., 2019, 2020), and updated O_3 deposition over seawater (Pound et al., 2020). When implemented in future versions of GEOS-CF, these updates are expected to reduce the high bias observed in $PM_{2.5}$ and AOD as well as the high bias in surface O_3 over the SH.

Another model development focus will center around the assimilation of satellite observations of atmospheric constituents, which has been shown to lead to improved representation of atmospheric composition, in particular for longer-lived species such as O_3 and CO (Benedetti et al., 2009; Inness et al., 2015). In its current form, GEOS-CF does not directly assimilate tropospheric trace gas observations, and the variability in constituent distribution is thus driven by the anthropogenic emission inventories, real-time biomass burning information, and the current meteorological state and its impact on dynamic emission sources such as biogenic emissions or sea salt aerosols.

GEOS-CF offers a new tool for academic researchers, air quality managers, and the public. Applications include flight campaign planning, support of satellite and other remote-sensing observations, interpretation of field campaign data (Dacic et al., 2020; Johnson et al., 2021), and air quality research (Keller et al., 2020).

Data Availability Statement

All model output is centrally stored at the NASA Center for Climate Simulation (NCCS). Public access to these archives is provided by the GMAO at https://gmao.gsfc.nasa.gov/weather_prediction/GEOS-CF/data_access/in the form of weather maps and through model output access tools, including OPeNDAP and Hypertext Transfer Protocol (HTTP).

Acknowledgments
Resources supporting the model simulations were provided by the NASA Center for Climate Simulation at the Goddard Space Flight Center (https://www.nccs.nasa.gov/services/discover). CAK, KEK, DJJ, EWL, and SP acknowledge support by the NASA Modeling, Analysis and Prediction (MAP) Program.

References

Amos, H. M., Jacob, D. J., Holmes, C. D., Fisher, J. A., Wang, Q., Yantosca, R. M., et al. (2012). Gas-particle partitioning of atmospheric Hg(II) and its effect on global mercury deposition. *Atmospheric Chemistry and Physics*, *12*, 591–603. https://doi.org/10.5194/acp-12-591-2012

Bacmeister, J. T., Suarez, M. J., & Robertson, F. R. (2006). Rain reevaporation, boundary layer-convection interactions, and Pacific rainfall patterns in an AGCM. *Journal of the Atmospheric Sciences*, *63*, 3383–3403. https://doi.org/10.1175/jas3791.1

Bates, K. H., & Jacob, D. J. (2019). A new model mechanism for atmospheric oxidation of isoprene: Global effects on oxidants, nitrogen oxides, organic products, and secondary organic aerosol. *Atmospheric Chemistry and Physics*, *19*, 9613–9640. https://doi.org/10.5194/acp-19-9613-2019

Benedetti, A., Morcrette, J.-J., Boucher, O., Dethof, A., Engelen, R. J., Fisher, M., et al. (2009). Aerosol analysis and forecast in the European Center for Medium-Range Weather Forecasts Integrated Forecast System: Data assimilation. *Journal of Geophysical Research*, *D13205*, 114. https://doi.org/10.1020/2008JD011115

Bey, I., Jacob, D. J., Yantosca, R. M., Logan, J. A., Field, B. D., Fiore, A. M., et al. (2001). Global modeling of tropospheric chemistry with assimilated meteorology: Model description and evaluation. *Journal of Geophysical Research*, *106*, 23073–23095. https://doi.org/10.1029/2001jd000807

Bhattacharjee, P. S., Wang, J., Lu, C.-H., & Tallapragada, V. (2018). The implementation of NEMS GFS Aerosol Component (NGAC) Version 2.0 for global multispecies forecasting at NOAA/NCEP—Part 2: Evaluation of aerosol optical thickness. *Geoscientific Model Development*, *11*, 2333–2351. https://doi.org/10.5194/gmd-11-2333-2018

Bian, H., & Prather, M. J. (2002). Fast-J2: Accurate simulation of stratospheric photolysis in global chemical models. *Journal of Atmospheric Chemistry*, *41*, 281–296. https://doi.org/10.1023/a:1014980619462

Borovikov, A., Cullather, R., Kovach, R., Marshak, J., Vernieres, G., Vikhliaev, Y., et al. (2019). GEOS-5 seasonal forecast system. *Climate Dynamics*, *53*, 7335–7361. https://doi.org/10.1007/s00382-017-3835-2

Buchard, V., Randles, C. A., da Silva, A. M., Darmenov, A., Colarco, P. R., Govindaraju, R., et al. (2017). The MERRA-2 aerosol reanalysis, 1980 onward. Part II: Evaluation and case studies. *Journal of Climate*, *30*, 6851–6872. https://doi.org/10.1175/jcli-d-16-0613.1

Carn, S. (2019). *Multi-satellite volcanic sulfur dioxide L4 long-term global database V3*. https://doi.org/10.5067/measures/so2/data404

Carpenter, L. J., MacDonald, S. M., Shaw, M. D., Kumar, R., Saunders, R. W., Parthipan, R., et al. (2013). Atmospheric iodine levels influenced by sea surface emissions of inorganic iodine. *Nature Geoscience*, *6*(2), 108–111. https://doi.org/10.1038/ngeo1687

Carter, T. S., Heald, C. L., Jimenez, J. L., Campuzano-Jost, P., Kondo, Y., Moteki, N., et al. (2020). How emissions uncertainty influences the distribution and radiative impacts of smoke from fires in North America. *Atmospheric Chemistry and Physics*, *20*, 2073–2097. https://doi.org/10.5194/acp-20-2073-2020

Chen, T., & Guestrin, C. (2016). *XGBoost: A scalable tree boosting system*. *CoRR*, abs/1603 (pp. 785–794). 02754. https://doi.org/10.1145/2939672.2939785

Chou, M.-D. (1990). Parameterizations for the absorption of solar radiation by O2 and CO_2 with application to climate studies. *Journal of Climate*, *3*, 209–217. https://doi.org/10.1175/1520-0442(1990)003<0209:pftaos>2.0.co;2

Chou, M.-D. (1992). A solar radiation model for use in climate studies. *Journal of the Atmospheric Sciences*, *49*, 762–772. https://doi.org/10.1175/1520-0469(1992)049<0762:asrmfu>2.0.co;2

Chou, M.-D., & Suarez, M. J. (1994). An efficient thermal infrared radiation parameterization for use in general circulation models. NASA Tech.Memorandum, NASA/TM-1994-104606 (Vol. 3, p. 85). Greenbelt, MD: NASA Goddard Space Flight Center.

Colarco, P., da Silva, A., Chin, M., & Diehl, T. (2010). Online simulations of global aerosol distributions in the NASA GEOS-4 model and comparisons to satellite and ground-based aerosol optical depth. *Journal of Geophysical Research: Atmospheres*, *115*, D14. https://doi.org/10.1029/2009jd012820

Crippa, M., Guizzardi, D., Muntean, M., Schaaf, E., Dentener, F., van Aardenne, J. A., et al. (2018). Gridded emissions of air pollutants for the period 1970-2012 within EDGAR v4.3.2. *Earth System Science Data*, *10*, 1987–2013. https://doi.org/10.5194/essd-10-1987-2018

Croft, B., Wentworth, G., Martin, R., Leaitch, W. R., Murphy, J. G., Murphy, B. N., et al. (2016). Contribution of Arctic seabird-colony ammonia to atmospheric particles and cloud-albedo radiative effect. *Nature Communications*, *7*, 13444. https://doi.org/10.1038/ncomms13444

Dacic, N., Sullivan, J. T., Knowland, K. E., Wolfe, G. M., Oman, L. D., Berkoff, T. A., & Gronoff, G. P. (2020). Evaluation of NASA's high-resolution global composition simulations: Understanding a pollution event in the Chesapeake Bay during the summer 2017 OWLETS campaign. *Atmospheric Environment*, *222*, 117133. ISSN 1352-2310. https://doi.org/10.1016/j.atmosenv.2019.117133

Darmenov, A. S., & da Silva, A. (2015). The Quick Fire Emissions Dataset (QFED)—Documentation of versions 2.1, 2.2 and 2.4. Technical report series on Global Modeling and Data Assimilation, NASA//TM-2015-104606 (Vol. 38, p. 212).

De Gouw, J., & Jimenez, J. L. (2009). Organic aerosols in the Earth's atmosphere. *Environmental Science & Technology*, *43*(20), 7614–7618. https://doi.org/10.1021/es9006004

Deeter, M. N., Edwards, D. P., Francis, G. L., Gille, J. C., Mao, D., Martínez-Alonso, S., et al. (2019). Radiance-based retrieval bias mitigation for the MOPITT instrument: The version 8 product. *Atmospheric Measurement Techniques*, *12*, 4561–4580. https://doi.org/10.5194/amt-12-4561-2019

Delle Monache, L., Nipen, T., Deng, X., Zhou, Y., & Stull, R. (2006). Ozone ensemble forecasts: 2. A Kalman filter predictor bias correction. *Journal of Geophysical Research*, *111*, D05308. https://doi.org/10.1029/2005JD006311

Denier van der Gon, H., Hendriks, C., Kuenen, J., Segers, A., & Visschedijk, A. (2011). Description of current temporal emission patterns and sensitivity of predicted AQ for temporal emission patterns, *TNO EU FP7 MACC deliverable report D_D-EMIS_1.3*, Utrecht, The Netherlands. Retrieved from https://atmosphere.copernicus.eu/sites/default/files/2019-07/MACC_TNO_del_1_3_v2.pdf

Djalalova, I., Delle Monache, L., & Wilczak, J. (2015). $PM_{2.5}$ analog forecast and Kalman filter post-processing for the Community Multiscale Air Quality (CMAQ) model. *Atmospheric Environment*, *108*, 76–87. ISSN 1352-2310. https://doi.org/10.1016/j.atmosenv.2015.02.021

Dlugokencky, E. J., Steele, L. P., Lang, P. M., & Masarie, K. A. (1994). The growth rate and distribution of atmospheric methane. *Journal of Geophysical Research*, *99*, 17021–17043. https://doi.org/10.1029/94JD01245

Douglass, A. R., Stolarski, R. S., Strahan, S. E., & Connell, P. S. (2004). Radicals and reservoirs in the GMI chemistry and transport model: Comparison to measurements. *Journal of Geophysical Research*, *109*, D16302. https://doi.org/10.1029/2004jd004632

Douglass, A. R., Strahan, S. E., Oman, L. D., & Stolarski, R. S. (2014). Understanding differences in chemistry climate model projections of stratospheric ozone. *Journal of Geophysical Research D: Atmospheres*, *119*, 4922–4939. https://doi.org/10.1002/2013jd021159

Duncan, B. N., Lamsal, L. N., Thompson, A. M., Yoshida, Y., Lu, Z., Streets, D. G., et al. (2016). A space-based, high-resolution view of notable changes in urban NO_x pollution around the world (2005–2014). *Journal of Geophysical Research D: Atmospheres*, *121*, 976–996. https://doi.org/10.1002/2015jd024121

Duncan, B. N., Logan, J. A., Bey, I., Megretskaia, I. A., Yantosca, R. M., Novelli, P. C., & Rinsland, C. P. (2007). The global budget of CO, 1988–1997: Source estimates and validation with a global model. *Journal of Geophysical Research*, *112*, D22301. https://doi.org/10.1029/2007jd008459

Eastham, S. D., Weisenstein, D. K., & Barrett, S. R. H. (2014). Development and evaluation of the unified tropospheric-stratospheric chemistry extension (UCX) for the global chemistry-transport model GEOS-Chem. *Atmospheric Environment*, *89*, 52–63. https://doi.org/10.1016/j.atmosenv.2014.02.001

Eck, T. F., Holben, B. N., Giles, D. M., Slutsker, I., Sinyuk, A., Schafer, J. S., et al. (2019). AERONET remotely sensed measurements and retrievals of biomass burning aerosol optical properties during the 2015 Indonesian burning season. *Journal of Geophysical Research D: Atmospheres*, *124*, 4722–4740. https://doi.org/10.1029/2018jd030182

Emmons, L. K., Schwantes, R. H., Orlando, J. J., Tyndall, G., Kinnison, D., Lamarque, J.-F., et al. (2020). The Chemistry Mechanism in the Community Earth System Model version 2 (CESM2). *Journal of Advances in Modeling Earth Systems*, *12*, e2019MS001882. https://doi.org/10.1029/2019ms001882

Evans, M. J., & Jacob, D. (2005). Impact of new laboratory studies of N_2O_5 hydrolysis on global model budgets of tropospheric nitrogen oxides, ozone, and OH. *Geophysical Research Letters*, *32*. https://doi.org/10.1029/2005gl022469

Fairlie, T. D., Jacob, D. J., & Park, R. J. (2007). The impact of transpacific transport of mineral dust in the United States. *Atmospheric Environment*, *41*, 1251–1266. https://doi.org/10.1016/j.atmosenv.2006.09.048

Feng, L., Smith, S. J., Braun, C., Crippa, M., Gidden, M. J., Hoesly, R., et al. (2020). The generation of gridded emissions data for CMIP6. *Geoscientific Model Development*, *13*, 461–482. https://doi.org/10.5194/gmd-13-461-2020

Fioletov, V. E., McLinden, C. A., Krotkov, N., Li, C., Joiner, J., Theys, N., et al. (2016). A global catalog of large SO_2 sources and emissions derived from the Ozone Monitoring Instrument. *Atmospheric Chemistry and Physics*, *16*, 11497–11519. https://doi.org/10.5194/acp-16-11497-2016

Fischer, E. V., Jacob, D. J., Yantosca, R. M., Sulprizio, M. P., Millet, D. B., Mao, J., et al. (2014). Atmospheric peroxyacetyl nitrate (PAN): A global budget and source attribution. *Atmospheric Chemistry and Physics*, *14*, 2679–2698. https://doi.org/10.5194/acp-14-2679-2014

Flemming, J., Huijnen, V., Arteta, J., Bechtold, P., Beljaars, A., Blechschmidt, A.-M., et al. (2015). Tropospheric chemistry in the Integrated Forecasting System of ECMWF. *Geoscientific Model Development*, *8*, 975–1003. https://doi.org/10.5194/gmd-8-975-2015

Flemming, J., Inness, A., Flentje, H., Huijnen, V., Moinat, P., Schultz, M. G., & Stein, O. (2009). Coupling global chemistry transport models to ECMWF's integrated forecast system. *Geoscientific Model Development*, *2*, 253–265. https://doi.org/10.5194/gmd-2-253-2009

Fountoukis, C., & Nenes, A. (2007). ISORROPIA II: A computationally efficient thermodynamic equilibrium model for K+-Ca2+-Mg2+-NH4+-Na+-SO42−-NO3−-Cl−-H2O aerosols. *Atmospheric Chemistry and Physics*, *7*, 4639–4659. https://doi.org/10.5194/acp-7-4639-2007

Garcia, R. R., & Boville, B. A. (1994). "Downward control" of the mean meridional circulation and temperature distribution of the polar winter stratosphere. *Journal of the Atmospheric Sciences*, *51*, 2238–2245. https://doi.org/10.1175/1520-0469(1994)051<2238:cotmmc>2.0.co;2

Gaubert, B., Arellano, A. F., Barré, J., Worden, H. M., Emmons, L. K., Tilmes, S., et al. (2016). Toward a chemical reanalysis in a coupled chemistry-climate model: An evaluation of MOPITT CO assimilation and its impact on tropospheric composition. *Journal of Geophysical Research D: Atmospheres*, *121*, 7310–7343. https://doi.org/10.1002/2016JD024863

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

Giles, D. M., Sinyuk, A., Sorokin, M. G., Schafer, J. S., Smirnov, A., Slutsker, I., et al. (2019). Advancements in the Aerosol Robotic Network (AERONET) Version 3 database—Automated near-real-time quality control algorithm with improved cloud screening for Sun photometer aerosol optical depth (AOD) measurements. *Atmospheric Measurement Techniques*, *12*, 169–209. https://doi.org/10.5194/amt-12-169-2019

Gong, S. L. (2003). A parameterization of sea-salt aerosol source function for sub- and super-micron particles. *Global Biogeochemical Cycles*, *17*, 1097. https://doi.org/10.1029/2003GB002079

Grange, S. K., & Carslaw, D. C. (2019). Using meteorological normalization to detect interventions in air quality time series. *The Science of the Total Environment*, *653*, 578–588. https://doi.org/10.1016/j.scitotenv.2018.10.344

Grange, S. K., Carslaw, D. C., Lewis, A. C., Boleti, E., & Hueglin, C. (2018). Random forest meteorological normalization models for Swiss PM10 trend analysis. *Atmospheric Chemistry and Physics*, *18*, 6223–6239. https://doi.org/10.5194/acp-18-6223-2018

Grosjean, D., & Harrison, J. (1985). Response of chemiluminescence NO_x analyzers and ultraviolet ozone analyzers to organic air pollutants. *Environmental Science & Technology*, *19*, 862–865. https://doi.org/10.1021/es00139a016

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

Hamill, T. M., & Whitaker, J. S. (2006). Probabilistic quantitative precipitation forecasts based on reforecast analogs: Theory and application. *Monthly Weather Review*, *134*, 3209–3229. https://doi.org/10.1175/MWR3237.1

Helfand, H. M., & Schubert, S. D. (1995). Climatology of the simulated Great Plains low-level jet and its contribution to the continental moisture budget of the United States. *Journal of Climate*, *8*, 784–806. https://doi.org/10.1175/1520-0442(1995)008<0784:cotsgp>2.0.co;2

Hill, C., DeLuca, C., Balaji, V., Suarez, M., & da Silva, A. (2004). The architecture of the earth system modeling framework. *Computing in Science & Engineering*, *6*(1), 18–28. https://doi.org/10.1109/mcise.2004.1255817

Honoré, C., Rouïl, L., Vautard, R., Beekmann, M., Bessagnet, B., Dufour, A., et al. (2008). Predictability of European air quality: Assessment of 3 years of operational forecasts and analyses by the PREV'AIR system. *Journal of Geophysical Research*, *113*, D04301. https://doi.org/10.1029/2007JD008761

Hu, L., Jacob, D. J., Liu, X., Zhang, Y., Zhang, L., Kim, P. S., et al. (2017). Global budget of tropospheric ozone: Evaluating recent model advances with satellite (OMI), aircraft (IAGOS), and ozonesonde observations. *Atmospheric Environment*, *167*, 323–334. https://doi.org/10.1016/j.atmosenv.2017.08.036

Hu, L., Keller, C. A., Long, M. S., Sherwen, T., Auer, B., Da Silva, A., et al. (2018). Global simulation of tropospheric chemistry at 12.5 km resolution: Performance and evaluation of the GEOS-Chem chemical module (v10-1) within the NASA GEOS Earth system model (GEOS-5 ESM). *Geoscientific Model Development*, *11*, 4603–4620. https://doi.org/10.5194/gmd-11-4603-2018

Hudman, R. C., Moore, N. E., Mebust, A. K., Martin, R. V., Russell, A. R., Valin, L. C., & Cohen, R. C. (2012). Steps toward a mechanistic model of global soil nitric oxide emissions: Implementation and space based-constraints. *Atmospheric Chemistry and Physics*, *12*, 7779–7795. https://doi.org/10.5194/acp-12-7779-2012

Huijnen, V., Pozzer, A., Arteta, J., Brasseur, G., Bouarar, I., Chabrillat, S., et al. (2019). Quantifying uncertainties due to chemistry modeling—Evaluation of tropospheric composition simulations in the CAMS model (cycle 43R1). *Geoscientific Model Development*, *12*, 1725–1752. https://doi.org/10.5194/gmd-12-1725-2019

Huneeus, N., Schulz, M., Balkanski, Y., Griesfeller, J., Prospero, J., Kinne, S., et al. (2011). Global dust model intercomparison in AeroCom phase I. *Atmospheric Chemistry and Physics*, *11*(15), 7781–7816. https://doi.org/10.5194/acp-11-7781-2011

Inness, A., Blechschmidt, A.-M., Bouarar, I., Chabrillat, S., Crepulja, M., Engelen, R. J., et al. (2015). Data assimilation of satellite-retrieved ozone, carbon monoxide and nitrogen dioxide with ECMWF's Composition-IFS. *Atmospheric Chemistry and Physics*, *15*, 5275–5303. https://doi.org/10.5194/acp-15-5275-2015

Ivatt, P. D., & Evans, M. J. (2020). Improving the prediction of an atmospheric chemistry transport model using gradient-boosted regression trees. *Atmospheric Chemistry and Physics*, *20*, 8063–8082. https://doi.org/10.5194/acp-20-8063-2020

Jaeglé, L., Quinn, P. K., Bates, T. S., Alexander, B., & Lin, J.-T. (2011). Global distribution of sea salt aerosols: New constraints from in situ and remote sensing observations. *Atmospheric Chemistry and Physics*, *11*, 3137–3157. https://doi.org/10.5194/acp-11-3137-2011

Jaeglé, L., Shah, V., Thornton, J. A., Lopez-Hilfiker, F. D., Lee, B. H., McDuffie, E. E., et al. (2018). Nitrogen oxides emissions, chemistry, deposition, and export over the Northeast United States during the WINTER Aircraft Campaign. *Journal of Geophysical Research D: Atmospheres*, *123*, 12368–12393. https://doi.org/10.1029/2018jd029133

Janssens-Maenhout, G., Crippa, M., Guizzardi, D., Dentener, F., Muntean, M., Pouliot, G., et al. (2015). HTAP_v2.2: A mosaic of regional and global emission grid maps for 2008 and 2010 to study hemispheric transport of air pollution. *Atmospheric Chemistry and Physics*, *15*, 11411–11432. https://doi.org/10.5194/acp-15-11411-2015

Johnson, M. S., Strawbridge, K., Knowland, K. E., Keller, C., & Travis, M. (2021). Long-range transport of Siberian biomass burning emissions to North America during FIREX-AQ. *Atmospheric Environment*, *252*, 118241.

Johnson, M. T. (2010). A numerical scheme to calculate temperature and salinity dependent air–water transfer velocities for any gas. *Ocean Science*, *6*, 913–932. https://doi.org/10.5194/os-6-913-2010

Keller, C. A., Evans, M. J., Knowland, K. E., Hasenkopf, C. A., Modekurty, S., Lucchesi, R. A., et al. (2020). Global impact of COVID-19 restrictions on the surface concentrations of nitrogen dioxide and ozone. *Atmospheric Chemistry and Physics Discussions*, *21*, 3555–3592. https://doi.org/10.5194/acp-2020-685

Keller, C. A., Long, M. S., Yantosca, R. M., Da Silva, A. M., Pawson, S., & Jacob, D. J. (2014). HEMCO v1.0: A versatile, ESMF-compliant component for calculating emissions in atmospheric models. *Geoscientific Model Development*, *7*, 1409–1417. https://doi.org/10.5194/gmd-7-1409-2014

Kim, P. S., Jacob, D. J., Fisher, J. A., Travis, K., Yu, K., Zhu, L., et al. (2015). Sources, seasonality, and trends of southeast US aerosol: An integrated analysis of surface, aircraft, and satellite observations with the GEOS-Chem chemical transport model. *Atmospheric Chemistry and Physics*, *15*, 10411–10433. https://doi.org/10.5194/acp-15-10411-2015

Krol, M., & Lelieveld, J. (2003). Can the variability in tropospheric OH be deduced from measurements of 1,1,1-trichloroethane (methyl chloroform)? *Journal of Geophysical Research*, *108*, 4125. https://doi.org/10.1029/2002JD002423

Lamsal, L. N., Krotkov, N. A., Vasilkov, A., Marchenko, S., Qin, W., Yang, E.-S., et al. (2021). Ozone Monitoring Instrument (OMI) Aura nitrogen dioxide standard product version 4.0 with improved surface and cloud treatments. *Atmospheric Measurement Techniques*, *14*, 455–479. https://doi.org/10.5194/amt-14-455-2021

Lee, C., Martin, R. V., van Donkelaar, A., Lee, H., Dickerson, R. R., Hains, J. C., et al. (2011). SO2emissions and lifetimes: Estimates from inverse modeling using in situ and global, space-based (SCIAMACHY and OMI) observations. *Journal of Geophysical Research*, *116*, D06304. https://doi.org/10.1029/2010JD014758

Levy, R. C., Hsu, C., et al. (2015). *MODIS atmosphere L2 aerosol product. NASA MODIS Adaptive processing system*. USA: Goddard Space Flight Center. http://dx.doi.org/10.5067/MODIS/MYD04_L2.061

Levy, R. C., Remer, L. A., Kleidman, R. G., Mattoo, S., Ichoku, C., Kahn, R., & Eck, T. F. (2010). Global evaluation of the Collection 5 MODIS dark-target aerosol products over land. *Atmospheric Chemistry and Physics*, *10*, 10399–10420. https://doi.org/10.5194/acp-10-10399-2010

Li, C., Krotkov, N. A., Leonard, P. J. T., Carn, S., Joiner, J., Spurr, R. J. D., & Vasilkov, A. (2020). Version 2 Ozone Monitoring Instrument SO2 product (OMSO2 V2): New anthropogenic SO2 vertical column density dataset. *Atmospheric Measurement Techniques*, *13*, 6175–6191. https://doi.org/10.5194/amt-13-6175-2020

Lin, S.-J. (2004). A "vertically Lagrangian" finite-volume dynamical core for global models. *Monthly Weather Review*, *132*, 2293–2307. https://doi.org/10.1175/1520-0493(2004)132<2293:avlfdc>2.0.co;2

Liu, F., Choi, S., Li, C., Fioletov, V. E., McLinden, C. A., Joiner, J., et al. (2018). A new global anthropogenic SO2 emission inventory for the last decade: A mosaic of satellite-derived and bottom-up emissions. *Atmospheric Chemistry and Physics*, *18*, 16571–16586. https://doi.org/10.5194/acp-18-16571-2018

Liu, H., Jacob, D. J., Bey, I., & Yantosca, R. M. (2001). Constraints from 210Pb and 7Be on wet deposition and transport in a global three-dimensional chemical tracer model driven by assimilated meteorological fields. *Journal of Geophysical Research*, *106*, 12109–12128. https://doi.org/10.1029/2000jd900839

Liu, T., Mickley, L. J., Marlier, M. E., DeFries, R. S., Firoz Khan, M., Latif, M. T., and Karambelas, A. (2020). Diagnosing spatial biases and uncertainties in global fire emissions inventories: Indonesia as regional case study, *Remote Sensing of Environment, 237* (111557), https://doi.org/10.1016/j.rse.2019.111557

Lock, A. P., Brown, A. R., Bush, M. R., Martin, G. M., & Smith, R. N. B. (2000). A new boundary layer mixing scheme. Part I: Scheme description and single-column model tests. *Monthly Weather Review*, *128*, 3187–3199. https://doi.org/10.1175/1520-0493(2000)128<3187:anblms>2.0.co;2

Long, M. S., Yantosca, R., Nielsen, J. E., Keller, C. A., da Silva, A., Sulprizio, M. P., et al. (2015). Development of a grid-independent GEOS-Chem chemical transport model (v9-02) as an atmospheric chemistry module for Earth system models. *Geoscientific Model Development*, *8*, 595–602. https://doi.org/10.5194/gmd-8-595-2015

Louis, J., & Geleyn, J. (1982). A short history of the PBL parameterization at ECMWF. In *Proceedings of the ECMWF workshop on planetary boundary layer parameterization* (pp. 59–80). Reading, UK: ECMWF.

Lucchesi, R. (2015). *File specification for GEOS-5 FP-it*. GMAO Office Note No. 2 (version 1.4) (58 pp.). Retrieved from http://gmao.gsfc.nasa.gov/pubs/office_notes.php

Lucchesi, R. (2017). *File specification for GEOS-5 FP*. GMAO Office Note No. 4 (version1.1) (61 pp.). Retrieved from http://gmao.gsfc.nasa.gov/pubs/office_notes

Luo, G., Yu, F., & Moch, J. M. (2020). Further improvement of wet process treatments in GEOS-Chem v12.6.0: Impact on global distributions of aerosols and aerosol precursors. *Geoscientific Model Development*, *13*, 2879–2903. https://doi.org/10.5194/gmd-13-2879-2020

Luo, G., Yu, F., & Schwab, J. (2019). Revised treatment of wet scavenging processes dramatically improves GEOS-Chem 12.0.0 simulations of surface nitric acid, nitrate, and ammonium over the United States. *Geoscientific Model Development*, *12*, 3439–3447. https://doi.org/10.5194/gmd-12-3439-2019

Mao, J., Paulot, F., Jacob, D. J., Cohen, R. C., Crounse, J. D., Wennberg, P. O., et al. (2013). Ozone and organic nitrates over the eastern United States: Sensitivity to isoprene chemistry. *Journal of Geophysical Research: Atmospheres*, *118*, 11256–11268. https://doi.org/10.1002/jgrd.50817

Marais, E. A., Jacob, D. J., Jimenez, J. L., Campuzano-Jost, P., Day, D. A., Hu, W., et al. (2016). Aqueous-phase mechanism for secondary organic aerosol formation from isoprene: Application to the southeast United States and co-benefit of SO2 emission controls. *Atmospheric Chemistry and Physics*, *16*, 1603–1618. https://doi.org/10.5194/acp-16-1603-2016

Marais, E. A., & Wiedinmyer, C. (2016). Air quality impact of Diffuse and Inefficient Combustion Emissions in Africa (DICE-Africa). *Environmental Science & Technology*, *50*(19), 10739–10745. https://doi.org/10.1021/acs.est.6b02602

Marécal, V., Peuch, V.-H., Andersson, C., Andersson, S., Arteta, J., Beekmann, M., et al. (2015). A regional air quality forecasting system over Europe: The MACC-II daily ensemble production. *Geoscientific Model Development*, *8*, 2777–2813. https://doi.org/10.5194/gmd-8-2777-2015

Martin, R. V., Jacob, D. J., Yantosca, R. M., Chin, M., & Ginoux, P. (2003). Global and regional decreases in tropospheric oxidants from photochemical effects of aerosols. *Journal of Geophysical Research*, *108*, 4097. https://doi.org/10.1029/2002jd002622

McFarlane, N. A. (1987). The effect of orographically excited gravity wave drag on the general circulation of the lower stratosphere and troposphere. *Journal of the Atmospheric Sciences*, *44*, 1775–1800. https://doi.org/10.1175/1520-0469(1987)044<1775:teooeg>2.0.co;2

McKeen, S., Wilczak, J., Grell, G., Djalalova, I., Peckham, S., Hsie, E.-Y., et al. (2005). Assessment of an ensemble of seven real-time ozone forecasts over eastern North America during the summer of 2004. *Journal of Geophysical Research*, *110*, D21307. https://doi.org/10.1029/2005JD005858

Miller, C., Jacob, D. J., Marais, E. A., Yu, K., Travis, K. R., Kim, P. S., et al. (2017). Glyoxal yield from isoprene oxidation and relation to formaldehyde: Chemical mechanism, constraints from SENEX aircraft observations, and interpretation of OMI satellite data. *Atmospheric Chemistry and Physics*, *17*, 8725–8738. https://doi.org/10.5194/acp-17-8725-2017

Millet, D. B., Baasandorj, M., Farmer, D. K., Thornton, J. A., Baumann, K., Brophy, P., et al. (2015). A large and ubiquitous source of atmospheric formic acid. *Atmospheric Chemistry and Physics*, *15*, 6283–6304. https://doi.org/10.5194/acp-15-6283-2015

Molod, A., Hackert, E., Vikhliaev, Y., Zhao, B., Barahona, D., Vernieres, G., et al. (2020). GEOS-S2S version 2: The GMAO high-resolution coupled model and assimilation system for seasonal prediction. *Journal of Geophysical Research D: Atmospheres*, *125*, e2019JD031767. https://doi.org/10.1029/2019JD031767

Molod, A., Suarez, M., & Partyka, G. (2013). The impact of limiting ocean roughness on GEOS-5 AGCM tropical cyclone forecasts. *Geophysical Research Letters*, *40*, 411–416. https://doi.org/10.1029/2012gl053979

Molod, A., Takacs, L. L., Suarez, M. J., Bacmeister, J. T., Song, I.-S., & Eichmann, A. (2012). The GEOS-5 atmospheric general circulation model: mean climate and development from MERRA to fortuna. NASA Tech. Memo. 104606. In M. J. Suarez (Ed.), Tech. Rep. Series on global modeling and data assimilation (Vol. 28, 117 pp.).

Molod, A., Takacs, L., Suarez, M., & Bacmeister, J. (2015). Development of the GEOS-5 atmospheric general circulation model: Evolution from MERRA to MERRA2. *Geoscientific Model Development*, *8*, 1339–1356. https://doi.org/10.5194/gmd-8-1339-2015

Monks, S. A., Arnold, S. R., Emmons, L. K., Law, K. S., Turquety, S., Duncan, B. N., et al. (2015). Multi-model study of chemical and physical controls on transport of anthropogenic and biomass burning pollution to the Arctic. *Atmospheric Chemistry and Physics*, *15*, 3575–3603. https://doi.org/10.5194/acp-15-3575-2015

Montzka, S. A., Spivakovsky, C. M., Butler, J. H., Elkins, J. W., Lock, L. T., & Mondeel, D. J. (2000). New observational constraints for atmospheric hydroxyl on global and hemispheric scales. *Science*, *288*, 500–503. https://doi.org/10.1126/science.288.5465.500

Moorthi, S., & Suarez, M. J. (1992). Relaxed Arakawa-Schubert. A parameterization of moist convection for general circulation models. *Monthly Weather Review*, *120*, 978–1002. https://doi.org/10.1175/1520-0493(1992)120<0978:rasapo>2.0.co;2

Murray, L. T., Jacob, D. J., Logan, J. A., Hudman, R. C., & Koshak, W. J. (2012). Optimized regional and interannual variability of lightning in a global chemical transport model constrained by LIS/OTD satellite data. *Journal of Geophysical Research*, *117*, D20307. https://doi.org/10.1029/2012JD017934

Naik, V., Voulgarakis, A., Fiore, A. M., Horowitz, L. W., Lamarque, J.-F., Lin, M., et al. (2013). Preindustrial to present-day changes in tropospheric hydroxyl radical and methane lifetime from the Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP). *Atmospheric Chemistry and Physics*, *13*, 5277–5298. https://doi.org/10.5194/acp-13-5277-2013

Nielsen, J. E., Pawson, S., Molod, A., Auer, B., da Silva, A. M., Douglass, A. R., et al. (2017). Chemical mechanisms and their applications in the Goddard Earth Observing System (GEOS) Earth System Model. *Journal of Advances in Modeling Earth Systems*, *9*, 3019–3044. https://doi.org/10.1002/2017MS001011

Nightingale, P. D., Malin, G., Law, C. S., Watson, A. J., Liss, P. S., Liddicoat, M. I., et al. (2000). In situ evaluation of air-sea gas exchange parameterizations using novel conservative and volatile tracers. *Global Biogeochemical Cycles*, *14*, 373–387. https://doi.org/10.1029/1999GB900091

Oda, T., & Maksyutov, S. (2017). *Open-source Data Inventory for Anthropogenic CO_2 (ODIAC) emission dataset*. National Institute for Environmental Studies. https://doi.org/10.17595/20170411.001

Oda, T., Maksyutov, S., & Andres, R. J. (2018). The Open-source Data Inventory for Anthropogenic CO_2, version 2016 (ODIAC2016): A global monthly fossil fuel CO_2 gridded emissions data product for tracer transport simulations and surface flux inversions. *Earth System Science Data*, *10*, 87–107. https://doi.org/10.5194/essd-10-87-2018

Oman, L. D., & Douglass, A. R. (2014). Improvements in total column ozone in GEOSCCM and comparisons with a new ozone-depleting substances scenario. *Journal of Geophysical Research D: Atmospheres*, *119*, 5613–5624. https://doi.org/10.1002/2014jd021590

Orbe, C., Oman, L. D., Strahan, S. E., Waugh, D. W., Pawson, S., Takacs, L. L., & Molod, A. M. (2017). Large-scale atmospheric transport in GEOS replay simulations. *Journal of Advances in Modeling Earth Systems*, *9*(7), 2545–2560. https://doi.org/10.1002/2017MS001053

Park, R. J., Jacob, D. J., Chin, M., & Martin, R. V. (2003). Source of carbonaceous aerosols over the United States and implications for natural visibility. *Journal of Geophysical Research*, *108*, 4355. https://doi.org/10.1029/2002jd003190

Park, R. J., Jacob, D. J., Field, B. D., Yantosca, R. M., & Chin, M. (2004). Natural and transboundary pollution influences on sulphate-nitrate-ammonium aerosols in the United States: Implications for policy. *Journal of Geophysical Research D: Atmospheres*, *109*. https://doi.org/10.1029/2003jd004473

Parrella, J. P., Jacob, D. J., Liang, Q., Zhang, Y., Mickley, L. J., Miller, B., et al. (2012). Tropospheric bromine chemistry: Implications for present and pre-industrial ozone and mercury. *Atmospheric Chemistry and Physics*, *12*, 6723–6740. https://doi.org/10.5194/acp-12-6723-2012

Patra, P. K., Krol, M. C., Montzka, S. A., Arnold, T., Atlas, E. L., Lintner, B. R., et al. (2014). Observational evidence for interhemispheric hydroxyl-radical parity. *Nature*, *513*, 219. https://doi.org/10.1038/nature13721

Pawson, S., Stajner, I., Kawa, S. R., Hayashi, H., Tan, W., Nielsen, J. E., et al. (2007). Stratospheric transport using six-hour averaged winds from a data assimilation system. *Journal of Geophysical Research*, *112*, D23103. https://doi.org/10.1029/2006jd007673

Petetin, H., Bowdalo, D., Soret, A., Guevara, M., Jorba, O., Serradell, K., & Pérez García-Pando, C. (2020). Meteorology-normalized impact of the COVID-19 lockdown upon NO2 pollution in Spain. *Atmospheric Chemistry and Physics*, *20*, 11119–11141. https://doi.org/10.5194/acp-20-11119-2020

Pound, R. J., Sherwen, T., Helmig, D., Carpenter, L. J., & Evans, M. J. (2020). Influences of oceanic ozone deposition on tropospheric photochemistry. *Atmospheric Chemistry and Physics*, *20*, 4227–4239. https://doi.org/10.5194/acp-20-4227-2020

Price, C., & Rind, D. (1992). A simple lightning parameterization for calculating global lightning distributions. *Journal of Geophysical Research*, *97*, 9919–9933. https://doi.org/10.1029/92JD00719

Price, C., & Rind, D. (1993). What determines the cloud-to-ground lightning fraction in thunderstorms?. *Geophysical Research Letters*, *20*(6), 463–466. https://doi.org/10.1029/93GL00226

Price, C., & Rind, D. (1994). Modeling global lightning distributions in a general circulation model. *Monthly Weather Review*, *122*(8), 1930–1939. https://doi.org/10.1175/1520-0493(1994)122<1930:MGLDIA>2.0.CO;2

Prinn, R. G., Huang, J., Weiss, R. F., Cunnold, D. M., Fraser, P. J., Simmonds, P. G., et al. (2001). Evidence for substantial variations of atmospheric hydroxyl radicals in the past two decades. *Science*, *292*, 1882–1888. https://doi.org/10.1126/science.1058673

Putman, W. M., & Lin, S.-J. (2007). Finite-volume transport on various cubed-sphere grids. *Journal of Computational Physics*, *227*, 55–78. https://doi.org/10.1016/j.jcp.2007.07.022

Pye, H. O. T., Liao, H., Wu, S., Mickley, L. J., Jacob, D. J., Henze, D. K., & Seinfeld, J. H. (2009). 36 Effect of changes in climate and emissions on future sulphate-nitrate-ammonium aerosol levels in the United States. *Journal of Geophysical Research: Atmospheres*, *114*, D01205. https://doi.org/10.1029/2008jd010701

Qu, Z., Henze, D. K., Li, C., Theys, N., Wang, Y., Wang, J., et al. (2019). SO2 emission estimates using OMI SO2 retrievals for 2005–2017. *Journal of Geophysical Research D: Atmospheres*, *124*, 8336–8359. https://doi.org/10.1029/2019jd030243

Randles, C. A., Buchard, A. M. V., Colarco, P. R., Darmenov, A., Govindaraju, R., Smirnov, A., et al. (2017). The MERRA-2 aerosol reanalysis, 1980 onward. Part I: System description and data assimilation evaluation. *Journal of Climate*, *30*, 6823–6850. https://doi.org/10.1175/jcli-d-16-0609.1

Remer, L. A., Kaufman, Y. J., Tanré, D., Mattoo, S., Chu, D. A., Martins, J. V., et al. (2005). The MODIS aerosol algorithm, products, and validation. *Journal of the Atmospheric Sciences*, *62*(4), 947–973. https://doi.org/10.1175/jas3385.1

Ridley, D. A., Heald, C. L., & Ford, B. J. (2012). North African dust export and impacts: An integrated satellite and model perspective. *Journal of Geophysical Research*, *117*, D02202. https://doi.org/10.1029/2011jd016794

Rienecker, M. M., Suarez, M. J., Todling, R., Bacmeister, J., Takacs, L., Liu, H.-C., et al. (2008). The GEOS-5 data assimilation system—documentation of versions 5.0.1, 5.1.0, and 5.2.0. *Technical Report Series on Global Modeling and Data Assimilation 104606* (Vol. *27)*.

Sandu, A., & Sander, R. (2006). Technical note: Simulating chemical systems in Fortran90 and Matlab with the Kinetic PreProcessor KPP-2.1. *Atmospheric Chemistry and Physics*, *6*, 187–195. https://doi.org/10.5194/acp-6-187-2006

Schultz, M. G., Heil, A., Hoelzemann, J. J., Spessa, A., Thonicke, K., Goldammer, J. G., et al. (2008). Global wildland fire emissions from 1960 to 2000. *Global Biogeochemical Cycles*, *22*, a. GB2002. https://doi.org/10.1029/2007GB003031

Schumann, U., & Huntrieser, H. (2007). The global lightning-induced nitrogen oxides source. *Atmospheric Chemistry and Physics*, *7*(14), 3823–3907. https://doi.org/10.5194/acp-7-3823-2007

Shah, V., Jacob, D. J., Li, K., Silvern, R. F., Zhai, S., Liu, M., et al. (2020). Effect of changing NOx lifetime on the seasonality and long-term trends of satellite-observed tropospheric NO2 columns over China. *Atmospheric Chemistry and Physics*, *20*, 1483–1495. https://doi.org/10.5194/acp-20-1483-2020

Shao, J., Chen, Q., Wang, Y., Lu, X., He, P., Sun, Y., et al. (2019). Heterogeneous sulphate aerosol formation mechanisms during wintertime Chinese haze events: Air quality model assessment using observations of sulphate oxygen isotopes in Beijing. *Atmospheric Chemistry and Physics*, *19*, 6107–6123. https://doi.org/10.5194/acp-19-6107-2019

Sherwen, T., Schmidt, J. A., Evans, M. J., Carpenter, L. J., Großmann, K., Eastham, S. D., et al. (2016). Global impacts of tropospheric halogens (Cl, Br, I) on oxidants and composition in GEOS-Chem. *Atmospheric Chemistry and Physics*, *16*, 12239–12271. https://doi.org/10.5194/acp-16-12239-2016

Shindell, D. T., Faluvegi, G., Stevenson, D. S., Krol, M. C., Emmons, L. K., Lamarque, J.-F. et al. (2006). Multimodel simulations of carbon monoxide: Comparison with observations and projected near-future changes. *Journal of Geophysical Research*, *111*, D19306. https://doi.org/10.1029/2006JD00710010.1029/2005jd006348

Spivakovsky, C. M., Logan, J. A., Montzka, S. A., Balkanski, Y. J., Foreman-Fowler, M., Jones, D. B. A., et al. (2000). Three-dimensional climatological distribution of tropospheric OH: Update and evaluation. *Journal of Geophysical Research*, *105*, 8931–8980. https://doi.org/10.1029/1999jd901006

Steinbacher, M., Zellweger, C., Schwarzenbach, B., Bugmann, S., Buchmann, B., Ordóñez, C., et al. (2007). Nitrogen oxide measurements at rural sites in Switzerland: Bias of conventional measurement techniques. *Journal of Geophysical Research*, *112*, D11307. https://doi.org/10.1029/2006JD007971

Stettler, M. E. J., Eastham, S., & Barrett, S. R. H. (2011). Air quality and public health impacts of UK airports. Part I: Emissions. *Atmospheric Environment*, *45*, 5415–5424. https://doi.org/10.1016/j.atmosenv.2011.07.012

Strahan, S. E., Duncan, B. N., & Hoor, P. (2007). Observationally derived transport diagnostics for the lowermost stratosphere and their application to the GMI chemistry and transport model. *Atmospheric Chemistry and Physics*, *7*, 2435–2445. https://doi.org/10.5194/acp-7-2435-2007

Streets, D. G., Canty, T., Carmichael, G. R., de Foy, B., Dickerson, R. R., Duncan, B. N., et al. (2013). Emissions estimation from satellite retrievals: A review of current capability. *Atmospheric Environment*, *77*, 1011–1042. https://doi.org/10.1016/j.atmosenv.2013.05.051

Strode, S. A., Duncan, B. N., Yegorova, E. A., Kouatchou, J., Ziemke, J. R., & Douglass, A. R. (2015). Implications of carbon monoxide bias for methane lifetime and atmospheric composition in chemistry climate models. *Atmospheric Chemistry and Physics*, *15*, 11789–11805. https://doi.org/10.5194/acp-15-11789-2015

Suarez, M., Trayanov, A., Hill, C., Schopf, P., & Vikhliaev, Y. (2007). MAPL: A high-level programming paradigm to support more rapid and robust encoding of hierarchical trees of interacting high-performance components. In *Proceedings of the 2007 symposium on component and framework technology in high-performance and scientific computing*. (pp. 11–20). https://doi.org/10.1145/1297385.1297388

Thompson, A. M., Witte, J. C., Sterling, C., Jordan, A., Johnson, B. J., Oltmans, S. J., et al. (2017). First reprocessing of southern hemisphere additional ozonesondes (SHADOZ) ozone profiles (1998–2016): 2. Comparisons with satellites and ground-based instruments. *Journal of Geophysical Research D: Atmospheres*, *122*, 13000–13025. https://doi.org/10.1002/2017jd027406

Tilmes, S., Lamarque, J.-F., Emmons, L. K., Conley, A., Schultz, M. G., Saunois, M., et al. (2012). Technical note: Ozonesonde climatology between 1995 and 2011: Description, evaluation and applications. *Atmospheric Chemistry and Physics*, *12*, 7475–7497. https://doi.org/10.5194/acp-12-7475-2012

Travis, K. R., & Jacob, D. J. (2019). Systematic bias in evaluating chemical transport models with maximum daily 8 h average (MDA8) surface ozone for air quality applications: A case study with GEOS-Chem v9.02. *Geoscientific Model Development*, *12*, 3641–3648. https://doi.org/10.5194/gmd-12-3641-2019

Travis, K. R., Jacob, D. J., Fisher, J. A., Kim, P. S., Marais, E. A., Zhu, L., et al. (2016). Why do models overestimate surface ozone in the Southeast United States? *Atmospheric Chemistry and Physics*, *16*, 13561–13577. https://doi.org/10.5194/acp-16-13561-2016

Wang, Q., Jacob, D. J., Spackman, J. R., Perring, A. E., Schwarz, J. P., Moteki, N., et al. (2014). Global budget and radiative forcing of black carbon aerosol: Constraints from pole-to-pole (HIPPO) observations across the Pacific. *Journal of Geophysical Research D: Atmospheres*, *119*, 195–206. https://doi.org/10.1002/2013jd020824

Wang, X., Jacob, D. J., Eastham, S. D., Sulprizio, M. P., Zhu, L., Chen, Q., et al. (2019). The role of chlorine in global tropospheric chemistry. *Atmospheric Chemistry and Physics*, *19*, 3981–4003. https://doi.org/10.5194/acp-19-3981-2019

Wang, Y., Jacob, D. J., & Logan, J. A. (1998). Global simulation of tropospheric O3-NOx-hydrocarbon chemistry: 3. Origin of tropospheric ozone and effects of nonmethane hydrocarbons. *Journal of Geophysical Research*, *103*, 10757–10767. https://doi.org/10.1029/98jd00156

Wargan, K., Kramarova, N., Weir, B., Pawson, S., & Davis, S. M. (2020). Toward a Reanalysis of Stratospheric Ozone for Trend Studies: Assimilation of the Aura Microwave Limb Sounder and Ozone Mapping and Profiler Suite Limb Profiler Data. *Journal of Geophysical Research D: Atmospheres*, *125*, e2019JD031892. https://doi.org/10.1029/2019JD031892

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

Wesely, M. L. (1989). Parameterization of surface resistances to gaseous dry deposition in regional-scale numerical models. *Atmospheric Environment*, *23*, 1293–1304. https://doi.org/10.1016/0004-6981(89)90153-4

Wilczak, J., McKeen, S., Djalalova, I., Grell, G., Peckham, S., Gong, W., et al. (2006). Bias-corrected ensemble and probabilistic forecasts of surface ozone over eastern North America during the summer of 2004. *Journal of Geophysical Research*, *111*(D23S28). https://doi.org/10.1029/2006JD007598

Winer, A. M., Peters, J. W., Smith, J. P., & Pitts, J. N., Jr (1974). Response of commercial chemiluminescent nitric oxide-nitrogen dioxide analyzers to other nitrogen-containing compounds. *Environmental Science & Technology*, *8*, 1118–1121. https://doi.org/10.1021/es60098a004

Yu, K., Keller, C. A., Jacob, D. J., Molod, A. M., Eastham, S. D., & Long, M. S. (2018). Errors and improvements in the use of archived meteorological data for chemical transport modeling: An analysis using GEOS-Chem v11-01 driven by GEOS-5 meteorology. *Geoscientific Model Development*, *11*, 305–319. https://doi.org/10.5194/gmd-11-305-2018

Zender, C. S., Bian, H., & Newman, D. (2003). Mineral Dust Entrainment and Deposition (DEAD) model: Description and 1990s dust climatology. *Journal of Geophysical Research*, *108*, 4416. https://doi.org/10.1029/2002JD002775

Zhang, Y., Jacob, D. J., Lu, X., Maasakkers, J. D., Scarpelli, T. R., Sheng, J.-X., et al. (2020). Attribution of the accelerating increase in atmospheric methane during 2010–2018 by inverse analysis of GOSAT observations. *Atmospheric Chemistry and Physics Discussions*. https://doi.org/10.5194/acp-2020-964 in review.