Environmental Science: Atmospheres

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal | View Issue

Cite this: *Environ. Sci.: Atmos.*, 2023, **3**, 1820

Implementation and evaluation of the automated model reduction (AMORE) version 1.1 isoprene oxidation mechanism in GEOS-Chem†

Benjamin Yang, *[ab] Forwood C. Wiser,[c] V. Faye McNeill, [bc] Arlene M. Fiore,[d] Madankui Tao, [abd] Daven K. Henze,[e] Siddhartha Sen[f] and Daniel M. Westervelt*[a]

Detailed chemical mechanisms are computationally challenging to include in large-scale chemical transport models such as GEOS-Chem. Employing a graph theory-based automated model reduction (AMORE) algorithm, we developed a new reduced (12 species and 23 reactions) gas-phase isoprene oxidation mechanism. We performed GEOS-Chem simulations for a full year (June 2018–May 2019) with the default (BASE) and AMORE version 1.1 isoprene mechanisms at 2° × 2.5° horizontal resolution globally and 0.25° × 0.3125° resolution over the eastern United States (EUS). Additionally, we conducted BASE and AMORE sensitivity simulations in which biogenic isoprene and anthropogenic emissions were sequentially set to zero in the model. For the entire year simulated, GEOS-Chem was faster by 10% in total and 25% in the chemical reaction solver (KPP) with the AMORE mechanism. Evaluating GEOS-Chem against surface observations from the Air Quality System (AQS) and Interagency Monitoring of Protected Visual Environments (IMPROVE) networks as well as satellite columns from the Tropospheric Monitoring Instrument (TROPOMI) and Cross-track Infrared Sounder (CrIS), our results show comparable accuracy in BASE and AMORE nested-grid simulations of air pollutants, with annual mean model bias changes of 1% for both surface $PM_{2.5}$ and O_3 over the EUS. From the sensitivity simulations, we find that US biogenic isoprene contributes to 8–9% of surface $PM_{2.5}$ and 3–4% of surface O_3 on average in summer over the EUS. This study indicates that AMORE is an attractive option for future GEOS-Chem modeling studies, especially where detailed isoprene chemistry is not the focus.

Received 2nd August 2023
Accepted 2nd November 2023

DOI: 10.1039/d3ea00121k

rsc.li/esatmospheres

Environmental significance

Atmospheric oxidation of isoprene leads to the secondary formation of air pollutants, namely fine particulate matter and ozone, which affect health and climate. The full isoprene oxidation mechanism is highly complex, involving hundreds of species and over a thousand reactions. To accurately and efficiently model atmospheric chemistry, the optimal balance between small mechanism size and accuracy should be achieved. We find that using our automated model reduction (AMORE) isoprene mechanism over the default mechanism in the GEOS-Chem chemical transport model not only speeds up simulations but also roughly maintains the accuracy of simulated air pollutant concentrations compared with surface and satellite-based observations. Our AMORE approach has the potential to be applied to other chemical mechanisms and models.

1. Introduction

Ambient fine particulate matter ($PM_{2.5}$) and ozone (O_3) were responsible for 4.14 million and 0.37 million global premature deaths, respectively, in 2019.[1] $PM_{2.5}$ is particularly harmful as these aerosols ≤2.5 μm in aerodynamic diameter can be inhaled deep into the lungs, and exposure to elevated $PM_{2.5}$ concentrations increases susceptibility to respiratory diseases such as asthma, lung cancer, and chronic obstructive pulmonary disease.[2] In addition, air quality is intricately linked with climate change. While O_3 exerts a net positive radiative forcing and aerosols induce either a negative or positive forcing, depending on the balance of scattering, absorption, and cloud interactions, climate change is expected to degrade air quality

[a]*Lamont-Doherty Earth Observatory of Columbia University, Palisades, NY, USA. E-mail: benjamin.yang@columbia.edu; danielmw@ldeo.columbia.edu*

[b]*Department of Earth and Environmental Sciences, Columbia University, New York, NY, USA*

[c]*Department of Chemical Engineering, Columbia University, New York, NY, USA*

[d]*Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA*

[e]*Department of Mechanical Engineering, University of Colorado Boulder, Boulder, CO, USA*

[f]*Microsoft Research, New York, NY, USA*

† Electronic supplementary information (ESI) available. See DOI: https://doi.org/10.1039/d3ea00121k

in polluted regions overall, involving changes in meteorology, deposition, and chemistry.[3] To better understand regional health impacts and climate feedbacks, we need to improve high-resolution modeled estimates of pollution not only directly emitted (primary) from anthropogenic and biogenic sources but also that formed through anthropogenic–biogenic chemical interactions (secondary).[4]

Isoprene (C_5H_8) is the predominant non-methane biogenic volatile organic compound (BVOC), representing about 50% of 1 Pg of total global annual BVOC emissions.[5] Most isoprene-emitting plant species are located in the humid tropics, and, in temperate regions, oak and aspen trees emit significant amounts of isoprene.[6] Monoterpenes and sesquiterpenes are other important BVOCs, constituting 15% and 3% of total global BVOC emissions.[5] The southeastern United States (US) is an extensively-studied isoprene hotspot with significant anthropogenic emissions. There, isoprene is primarily emitted during the warm, growing season and dominates hydroxyl radical (OH) reactivity in summer, while monoterpenes remain more constant and dominate O_3 and nitrate radical (NO_3) reactivity throughout the year.[7] In GEOS-Chem, isoprene is mostly oxidized by OH (85%), followed by O_3 (11%) and NO_3 (4%) pathways, over the southeastern US.[8]

Atmospheric photooxidation of VOCs including isoprene by OH produces an organic peroxy radical that converts nitric oxide (NO) to nitrogen dioxide (NO_2).[9] An organic oxy radical reacts with oxygen to form hydroperoxyl radical (HO_2), and the HO_2 additionally converts NO to NO_2.[9] In the presence of sunlight, photolysis of NO_2 forms O_3, which means that both isoprene and total NO_x (NO + NO_2) levels partially control O_3 production. Formaldehyde (HCHO) and other oxygenated VOCs are formed in the oxidation process. Furthermore, biogenic isoprene and monoterpenes are responsible for 60% of organic aerosol (OA), a major component of $PM_{2.5}$, over the southeastern US.[10] Isoprene and its intermediates participate in an elaborate series of chemical reactions to form lower volatility products that partition into OA depending on the ambient conditions.[11] For example, isoprene species react with OH and HO_2 to form isoprene epoxydiols (IEPOX) under low-NO_x conditions.[12] It is estimated that IEPOX, organonitrates, and tetrafunctional compounds each contribute to one-third of global isoprene secondary organic aerosol (SOA).[8] Yet, past studies have indicated many unknowns in isoprene chemistry, particularly the fate of isoprene nitrates.[13–15]

In atmospheric chemistry modeling, we need innovative ways for representing complex, nonlinear, multiphase, and multiscale chemical processes. Currently, the master chemical mechanism (MCM version 3.3.1) for isoprene oxidation involves 602 species and 1926 reactions, derived from theoretical studies and chamber experiments.[16] Adding 86 species and 358 reactions to a detailed, comprehensive mechanism of isoprene and terpene chemistry (MOZART) was shown to decrease the daily maximum surface O_3 bias of the CESM/CAM-chem model generally by 3–4 ppb in the eastern US, relative to the previous, smaller MOZART mechanism, and increased computation time by 50%.[15] However, it is not always practical to capture the detailed chemistry in large-scale atmospheric chemical transport models (CTMs) including GEOS-Chem or chemistry-climate models (CCMs) due to excessive computational costs. Consequently, CTMs and CCMs often use reduced mechanisms while striving for accuracy within acceptable ranges as a trade-off. Even in reduced form, chemistry involving isoprene is one of the most computationally-intensive components in these models.

To date, chemical mechanism reduction has been mostly performed manually by expert air quality scientists and software developers through techniques such as chemical lumping and empirical parameterization. Condensing the mechanism may occur in several stages including isomer grouping, applying steady-state approximation, and lumping minor and latter pathways together.[17] Species are also lumped, resulting in different nomenclature across mechanisms employed in CTMs.[8] The reduced mechanisms typically lack flexibility to adapt to new environmental conditions and require significant labor, time, and resources to make further updates. Updates are infrequent and susceptible to errors. Therefore, automating the reduction of chemical mechanisms to speed up the integration of chemical kinetics is an attractive, viable alternative.[18,19] A few studies have applied graph theory in reducing mechanisms in atmospheric chemistry, although it is has not widely been used.[19–21] Recently, we used a directed-graph path-based automated model reduction (AMORE) approach to develop a novel reduced isoprene oxidation mechanism (version 1.0) which outperformed other common mechanisms in the Framework for 0-D Atmospheric Modeling (F0AM) and Community Multiscale Air Quality (CMAQ) models.[19]

Here, we incorporate an updated AMORE isoprene oxidation mechanism (version 1.1) into the chemical reaction solver of the GEOS-Chem chemical transport model. We describe the configuration of year-long global and nested-grid eastern US (EUS) simulations. Both ground-based (AQS/IMPROVE) and satellite observations (TROPOMI/CrIS) were used to evaluate the GEOS-Chem model with the AMORE mechanism. We primarily aim to understand the effects of AMORE on GEOS-Chem (1) computational efficiency and (2) accuracy in predictions of air pollutants and related chemical species ($PM_{2.5}$, O_3, NO_2, HCHO, and total organic carbon). Finally, we examine the sensitivities of summertime air pollutant levels to biogenic isoprene and anthropogenic emissions across the EUS.

2. Methods

2.1. AMORE isoprene oxidation mechanism

The new reduced gas-phase isoprene oxidation mechanism was developed using the AMORE algorithm.[19] This algorithm takes a large, high-fidelity input mechanism, such as the "Caltech full mechanism" based on Wennberg *et al.* (2018)[17] which was expanded in Wiser *et al.* (2023),[19] and reduces it to a much smaller size for 3D atmospheric models where time savings are critical. A set of priority species are given for the reduced mechanism to include, selected based on prior knowledge of important species which have high yields in the isoprene system and are usually included in atmospheric chemistry models. For the AMORE mechanism, these species were

isoprene, IEPOX, isoprene nitrates (lumped), formaldehyde, glyoxal, methylglyoxal, methacrolein, methyl vinyl ketone, peroxyacetyl nitrate, methyl peroxy radical, and peroxyacetyl radical.

The full isoprene mechanism produces various priority species which are highly sensitive to concentrations of atmospheric reactants. The AMORE algorithm was designed to create highly reduced mechanisms with similar sensitivity to atmospheric reactant concentrations as the full mechanism in regard to priority output products. AMORE reduces the original mechanism to a set of simple mechanistic pathways which are characterized by the types of reactions in the pathway. The type of reaction is defined by the reactive species involved (OH, HO_2, NO, NO_2, NO_3, O_3, or methyl peroxy radical) or as a photolytic reaction. AMORE selected the sequence of reaction types necessary in the reduced mechanism to account for sensitivity of the output of the full mechanism to the concentrations of these reactive species as well as photolysis. In order to measure the sensitivity of the output of the full mechanism to these conditions, yields of all priority species were measured under reduced and elevated levels of each of these species (ppb) and photolysis (unitless): OH (1×10^{-6}, 1×10^{-4}), HO_2 (0.04, 0.2), NO (1.17×10^{-6}, 0.53), NO_2 (1×10^{-4}, 0.01), NO_3 (2.3×10^{-4}, 0.02), O_3 (16.7, 100), methyl peroxy radical (0.1, 0.2), and photolysis scale factor (0, 1). Mechanistic pathways consisted of sequences of reaction types and were selected when the species or photolysis involved in the reactions were elevated. For example, the pathway consisting of a reaction with OH followed by a reaction with HO_2 was selected when OH and HO_2 concentrations were elevated, with all other species or photolysis reduced. The importance of a given pathway was quantified by measuring the change in yield of priority species between the pathway and those with one less step. For example, the yields of priority species in the OH and HO_2 two-step pathway were compared to those in the OH one-step pathway and the HO_2 one-step pathway, to assess the magnitude of change incurred by the two-step pathway over the one-step pathways.

A set of the most important mechanistic pathways was selected to form a reduced mechanism. In any instances in which two pathways shared common reaction types, the pathways were joined and branched from each other at the differentiating step. Any pathway with more than one step required an intermediate species for each additional step to join the reactions together. For example, the OH and HO_2 pathway required an intermediate product for the reaction of isoprene and OH, which then reacted with HO_2. All pathways had associated priority species yields, which were used to determine appropriate stoichiometric coefficients for priority species in the terminal reaction of the pathway. Reaction rate constants were selected based on reference values for each reaction type. The algorithmically generated mechanism was run through multiple rounds of box model testing, and manual edits were made to stoichiometric coefficients and rate constants. Some products were included that had been omitted from the algorithm's product selection process as atmospheric reactants. For a detailed description of the development of this AMORE version 1.0 isoprene oxidation mechanism, see Wiser *et al.* (2023).[19]

Minor updates were made to the version 1.0 mechanism based on initial GEOS-Chem testing. Most of the isoprene species were renamed to match existing, similar nomenclature accepted in GEOS-Chem. As excessive O_3 produced from NO_2 in the AMORE version 1.0 mechanism was observed, we opted to remove one artifact, NO_2 produced in the isoprene hydroxy hydroperoxide (RIPA) + OH reaction, which lowered O_3 to near the baseline levels. In implementing AMORE, carbon and nitrogen are not conserved, and therefore AMORE might not be an appropriate choice for carbon and nitrogen budget studies. Displayed in Tables 1 and 2, the AMORE isoprene version 1.1 mechanism comprises 12 species and 23 reactions. The default GEOS-Chem isoprene oxidation mechanism (mini Caltech) is described in Bates and Jacob (2019)[8] and will be referred to as BASE. Compared to BASE, AMORE is smaller by 43 species and 167 reactions. A copy of the BASE mechanism was replaced with our final AMORE version 1.1 mechanism in the Kinetic PreProcessor (KPP) version 2.3.3, the chemical reaction solver of GEOS-Chem.[22] Including non-isoprene chemistry, the entire BASE mechanism in KPP consists of 291 species and 903 reactions. KPP takes a list of species, reactions, and rate constants as inputs, solves the chemical kinetics for the given mechanism,

Table 1 A list of the 12 isoprene species participating in the AMORE mechanism. Dummy species to track oxidation of isoprene by OH (LISOPOH) and NO_3 ($LISOPNO_3$) are excluded. Deposition for each isoprene species is treated the same way it was before in GEOS-Chem

#	Species name	AMORE nomenclature	GEOS-Chem nomenclature
1	Isoprene	ISOP	ISOP
2	Isoprene nitrate peroxy radical	INO2	INO2B
3	Isoprene peroxy radical	IHOO	IHOO1
4	Isoprene hydroxy hydroperoxide	ISHP	RIPA
5	Isoprene hydroxy nitrate	IHN	IHN1
6	Isoprene peroxy nitrate	IPN	INPB
7	Lumped higher order isoprene nitrates	ISON	IDN
8	Isoprene lumped peroxy carbonyl + other functional groups	IPC	IDC
9	Isoprene epoxydiols	IEPOX	IEPOXA
10	Isoprene-4,1-hydroxyaldehyde	HC5A	HC5A
11	Aerosol-phase organic nitrate from isoprene precursors	IONITA	IONITA
12	Aerosol-phase IEPOX	SOAIE	SOAIE

Table 2 The AMORE isoprene oxidation mechanism (23 reactions) as integrated into KPP. Heterogeneous reactions are not counted. Photolysis of RIPA (s^{-1}) is calculated externally by the default FAST-JX version 7.0a photolysis mechanism using the cross section of methyl hydroperoxide (CH_3OOH). All other rate constants have units of $cm^3\ mol^{-1}\ s^{-1}$. T is temperature (K)

#	Reaction	Rate constant
1	ISOP + O_3 → 0.189MVK + 0.58CH_2O + 0.25OH + 0.25HO_2 + 0.08MO_2 + 0.1MCO_3 + 0.09H_2O_2 + 0.461MACR + 0.14CO	$1.58 \times 10^{-14} \exp(-2000/T)$
2	ISOP + NO_3 → INO2B + 0.3CH_2O + 0.3NO_2 + 0.3IDN + $LISOPNO_3$	$2.95 \times 10^{-12} \exp(-450/T)$
3	ISOP + OH → IHOO1 + 0.02MO_2 + LISOPOH	$2.69 \times 10^{-11} \exp(390/T)$
4	IHOO1 + HO_2 → RIPA + 0.6HO_2 + 0.15CH_2O	$4.5 \times 10^{-13} \exp(1300/T)$
5	IHOO1 + NO → 0.14IHN1 + 0.7CH_2O + 0.44MVK + 0.88HO_2 + 0.78NO_2 + 0.28MACR + 0.021GLYX	$2.7 \times 10^{-12} \exp(350/T)$
6	RIPA + OH → IHOO1	$4.6 \times 10^{-12} \exp(200/T)$
7	INO2B + HO_2 → INPB + OH	$3.14 \times 10^{-14} \exp(580/T)$
8	INO2B + NO → 0.9CH_2O + 0.5MGLY + 0.8MVK + 0.5NO_2 + HO_2 + 0.2IDN + 0.1MO_2	$9.42 \times 10^{-16} \exp(580/T)$
9	INPB + HO_2 → 0.8NO_2 + 0.4CH_2O + 0.05GLYX + 0.1MGLY + 0.4MACR + HO_2 + 0.94MVK + 0.2IDN + 0.1MO_2	$3.4 \times 10^{-11} \exp(390/T)$
10	IHN1 + OH → IDN + OH + 0.2IEPOXA	$2.4 \times 10^{-7} \exp(580/T)$
11	RIPA + OH → 0.15CH_2O + 0.05MGLY + 0.15MACR + 0.02GLYX + 0.2MVK + 0.05IDC + 0.58IEPOXA + 0.8OH	$2.97 \times 10^{-11} \exp(390/T)$
12	IDC + NO → 0.35NO_2 + 0.8NO	1×10^{-10}
13	IDN + OH → CO + 0.12NO_2	5×10^{-11}
14	IDN + NO_3 → CO	2×10^{-14}
15	IEPOXA + OH → OH	$5 \times 10^{-11} \exp(-400/T)$
16	IHN1 → HNO_3	2.3×10^{-5}
17	IHOO1 + MCO_3 → 0.5HO_2 + 0.5MO_2 + 1.048CH_2O + 0.219MACR + 0.305MVK	$8.4 \times 10^{-14} \exp(221/T)$
18	IHOO1 + IHOO1 → 2MVK + 2HO_2 + 2CH_2O	$6.92 \times 10^{-14} [1.1644 + (-7.0485 \times 10^{-4})T]$
19	IHOO1 + MO_2 → MVK + 2HO_2 + 2CH_2O	$2 \times 10^{-12} [1.1644 + (-7.0485 \times 10^{-4})T]$
20	IHOO1 + MO_2 → CH_2O + 0.5HC5A + 1.5HO_2 + 0.5MVKHP + 0.5CO + 0.5OH	$2 \times 10^{-12} [-0.1644 + (7.0485 \times 10^{-4})T]$
21	HC5A + OH → 1.065OH + 0.355CO_2 + 0.638CO + 0.355MGLY + 0.283HO_2 + 0.125MVKHP + 0.158MCRHP	$4.64 \times 10^{-12} \exp(650/T)$
22	Cl + ISOP → HCl + IHOO1	$7.6 \times 10^{-11} \exp(500/T)$
23	RIPA + $h\nu$ → 0.4CH_2O + 0.1MGLY + 0.06MCO_3	PHOTOL(RIPA)

and outputs Fortran 90 code that gets compiled in GEOS-Chem.[18,23,24] We built compilable code in KPP with the BASE and AMORE mechanisms separately.

2.2. GEOS-Chem model simulations

GEOS-Chem is a global 3D Eulerian CTM accessed freely and used by research groups worldwide to investigate atmospheric chemistry (**https://geoschem.github.io**, last access: 12 November 2023). We set up the GEOS-Chem "Classic" version 13.3.3 model[25] on Columbia University's "Ginsburg" High Performance Computing Cluster to conduct all simulations on a single node with two Intel Xeon Gold 6226 2.9 GHz central processing units, 32 cores, and 192 GB (for coarser-resolution simulations) or 768 GB (for higher-resolution simulations) of memory. All simulations were configured with full chemistry, complex SOA, and semi-volatile primary organic aerosol (POA) for 2018–2019. We designated the first five months as the model spin-up period and the period June 2018–May 2019 for our analysis. Complex SOA and semi-volatile POA were selected over simple SOA and non-volatile POA because the former combination includes heterogeneous reactions of isoprene oxidation products to form SOA and are thus more coupled to the gas-phase isoprene chemistry.[11,26] Output diagnostic files included monthly-averaged aerosol mass concentrations, species concentrations, and reaction rates as well as daily boundary conditions. Timers in GEOS-Chem were enabled for the BASE and AMORE isoprene mechanism simulations to compare total run (wall-clock) times and estimate computational speed gained by running AMORE over BASE. As presented in Table 3, we performed 10 out of the 12 simulations globally at 2° × 2.5° horizontal resolution. These simulations were configured with all 72 hybrid sigma vertical levels; 10 minutes timesteps for transport, cloud convection, planetary boundary layer (PBL) mixing, and wet deposition; and 20 minutes timesteps for chemistry, emissions, and dry deposition. To optimize computational efficiency and simulation accuracy, chemical operator duration is recommended to be twice the transport operator duration.[27] We used Modern-Era Retrospective analysis for

Table 3 All 12 GEOS-Chem sensitivity simulations conducted in this study

#	Simulation name	Isoprene chemistry	Emissions perturbation	Horizontal resolution
1	BASE	Default	None	2° × 2.5°
2	AMORE	AMORE	None	2° × 2.5°
3	BASE_HR	Default	None	0.25° × 0.3125°
4	AMORE_HR	AMORE	None	0.25° × 0.3125°
5	BASE_zUS_ISOP	Default	US biogenic isoprene set to zero	2° × 2.5°
6	AMORE_zUS_ISOP	AMORE	US biogenic isoprene set to zero	2° × 2.5°
7	BASE_zUS_ANTH	Default	US anthropogenic set to zero	2° × 2.5°
8	AMORE_zUS_ANTH	AMORE	US anthropogenic set to zero	2° × 2.5°
9	BASE_zUS_ISOP_ANTH	Default	US biogenic isoprene and anthropogenic set to zero	2° × 2.5°
10	AMORE_zUS_ISOP_ANTH	AMORE	US biogenic isoprene and anthropogenic set to zero	2° × 2.5°
11	BASE_zGLB_ANTH	Default	Global anthropogenic set to zero	2° × 2.5°
12	AMORE_zGLB_ANTH	AMORE	Global anthropogenic set to zero	2° × 2.5°

Research and Applications, version 2 (MERRA-2) meteorological products[28,29] to drive the coarser-resolution simulations.

Using boundary conditions from the BASE 2° × 2.5° run, we also performed two 0.25° × 0.3125° nested-grid simulations (BASE_HR and AMORE_HR), over the eastern half of the US: latitude = [24°, 49°] and longitude = [−100°, −66°]. Tracer concentrations are known to be unrealistic on the edges of the nested grid domain where there is no advection.[30] Therefore, we applied a buffer zone of 3° along each of the four boundaries. The goal was to focus on peak summertime isoprene emissions over the southeastern US, but we expanded this domain to include the more forested, humid, and densely-populated urban areas east of the 100th west meridian. Particularly in the northeastern US, anthropogenic emissions interact with biogenic emissions to form ground-level O_3 in the summertime.[31] The higher-resolution model had 47 vertical levels, extending from the surface to 0.01 hPa as in the 72-layer model, but with a coarser vertical resolution in the stratosphere to save computational resources. Transport, cloud convection, PBL mixing, and wet deposition cycled on 5 minutes time steps, while chemistry, emissions, and dry deposition cycled on 10 minutes time steps. Instead of MERRA-2, Goddard Earth Observing System-forward processing (GEOS-FP) meteorology[32] was used for the higher-resolution simulations. MERRA-2 and GEOS-FP are similar and both originate from the National Aeronautics and Space Administration (NASA) Global Modeling and Assimilation Office, though GEOS-FP offers the finer native horizontal resolution (0.25° × 0.3125°) that is not available with MERRA-2.

To quantify contributions from biogenic isoprene and anthropogenic emissions to air pollution levels in our EUS domain *via* both isoprene mechanisms, we perturbed emissions in eight sensitivity simulations (Table 3). Emissions are processed by the Harvard-NASA Emissions Component (HEMCO) module.[33] First, we defined a scale factor to zero out emissions over the contiguous US: latitude = [20°, 60°] and longitude [−140°, −50°]. Thereafter, we applied this scale factor to isoprene from the Model of Emissions of Gases and Aerosols from Nature (MEGAN) version 2.1 for the zUS_ISOP simulations.[5] Total biogenic isoprene emissions over the EUS domain in the coarser-resolution simulations (12.16 Tg C $year^{-1}$) and higher-resolution (12.55 Tg C $year^{-1}$) simulations are comparable. For the zUS_ANTH simulations, the same scale factor was applied to the following anthropogenic emissions sources: the Community Emissions Data System version 2,[34] ethane,[35] propane,[36] the Aviation Emissions Inventory Code,[37] ship emissions,[34] and anthropogenic fugitive, combustion, and industrial dust.[38] These anthropogenic emissions include sulfur dioxide (SO_2), carbon monoxide (CO), NO_x, PM, and VOCs. Soil NO_x emissions including fertilizer NO_x are left on and are similar in the coarser-resolution simulations (0.23 Tg N $year^{-1}$) and higher-resolution simulations (0.22 Tg N $year^{-1}$). Combining the aforementioned methods, both biogenic isoprene and anthropogenic emissions were zeroed out for the zUS_ISOP_ANTH simulations. Finally, we turned off the same anthropogenic emissions inventories globally for the zGLB_ANTH simulations.

2.3. Evaluating model against observations

Table 4 provides an overview of the observed data used to evaluate the model: (1) surface $PM_{2.5}$ and O_3 from the US Environmental Protection Agency's Air Quality System (AQS) network, (2) surface total organic carbon (OC) from the Interagency Monitoring of Protected Visual Environments (IMPROVE) network, (3) NO_2 and HCHO tropospheric vertical column densities (VCDs) from the Tropospheric Monitoring Instrument (TROPOMI), and (4) isoprene total VCD from the Cross-track Infrared Sounder (CrIS). We obtained AQS/IMPROVE data for June 2018–May 2019, removed any negative daily average concentrations, and selected days without any gaps in hourly measurements. For each AQS/IMPROVE variable, the concentrations were averaged by month. While the AQS/IMPROVE sites provide valuable *in situ* point measurements with high temporal coverage, they are unevenly distributed and have less spatial coverage compared to satellite remote sensing data which measure total atmospheric columns and not only the surface layer. In particular, AQS instruments are sensitive to local, urban pollution that tend to not be representative of entire model grid boxes for short-lived species including NO_2, HCHO, and isoprene.

The European Space Agency's Sentinel-5 Precursor (S5P) is a sun-synchronous, polar-orbiting satellite with a daily equator

Table 4 Summary of observed data processed from AQS/IMPROVE monitoring sites (FRM/FEM for $PM_{2.5}$) and TROPOMI/CrIS satellite retrievals

Variable	Units	Number of sites/spatial resolution	Data source
Surface $PM_{2.5}$	$\mu g\ m^{-3}$	606 sites	US Environmental Protection Agency AirData[45]
Surface O_3	ppb	765 sites	US Environmental Protection Agency AirData[45]
Surface total OC	$\mu g\ m^{-3}$	43 sites	Colorado State University Federal Land Manager Environmental Database[46]
NO_2 tropospheric VCD	molecules cm^{-2}	0.05° × 0.05°	NASA Goddard Earth Sciences Data and Information Services Center[47]
HCHO tropospheric VCD	molecules cm^{-2}	0.05° × 0.05°	NASA Goddard Earth Sciences Data and Information Services Center[48]
Isoprene total VCD	molecules cm^{-2}	0.5° × 0.625°	University of Minnesota Data Repository[44]

overpass time of 13:30 local time.[39] TROPOMI is a nadir-viewing spectrometer aboard the S5P, measuring NO_2 and HCHO columns in the ultraviolet-visible-near-infrared spectral bands at 3.5 km × 7 km resolution. In addition to higher spatial resolution, TROPOMI has a better signal-to-noise ratio compared with its predecessor, the Ozone Monitoring Instrument (OMI), but still exhibits regional biases including overestimating small HCHO columns and underestimating high HCHO columns.[40] We re-gridded TROPOMI Level 2 quality-controlled retrievals, transitioning from reprocessed (RPRO) and offline (OFFL) products on 28 November 2018, to 0.05° × 0.05° daily from June 2018–May 2019. As recommended,[41,42] we selected pixels with quality assurance (QA) values greater than 0.75 for NO_2 and QA > 0.5 for HCHO, removing errors and observations influenced by cloud, snow, or ice cover. TROPOMI NO_2 and HCHO tropospheric VCDs, averaging kernels, and vertical pressure levels were averaged monthly.

CrIS is a Fourier transform spectrometer aboard the US National Oceanic and Atmospheric Administration's Suomi-NPP and NOAA-20 sun-synchronous satellites which have daily overpass times of about 13:30 and 12:40 local time.[43] We use the Retrieval of Organics from CrIS Radiances (ROCR) isoprene retrievals[44] at 0.5° × 0.625° resolution from Wells *et al.* (2022)[43] who employ a feed-forward neural network to derive total isoprene columns from CrIS hyperspectral range indices. Averaging kernels or other satellite observation operators are not available through this machine learning framework, and the products are averaged each month of year over the period 2012–2020. As a result, GEOS-Chem and CrIS should be compared more qualitatively.

To evaluate GEOS-Chem against TROPOMI and CrIS, we first converted NO_2, HCHO, and isoprene mole fractions for each 3D grid box from all GEOS-Chem simulations to partial tropospheric (NO_2 and HCHO) or total (isoprene) columns. The calculation required air temperature, specific humidity, surface pressure, and tropopause pressure from both the MERRA-2 and GEOS-FP meteorological products for the coarser and higher-resolution simulations, respectively. For consistency between modeled and observed vertical profiles, we applied TROPOMI averaging kernels at TROPOMI vertical pressure levels nearest to GEOS-Chem vertical pressure levels. This was performed by multiplying partial tropospheric columns (re-gridded to 0.05° × 0.05°) by averaging kernels on the same grid and then integrating from the surface up to the nonuniform tropopause. Isoprene partial columns were simply integrated from the surface to the top of the atmosphere. For all variables listed in Table 4, we selected observations within latitude = [27°, 46°] and longitude = [−97°, −69°] to exclude the buffer zone of the nested-grid simulations. The nearest model grid box to each AQS/IMPROVE site or TROPOMI/CrIS grid box was found, differences between the monthly predicted and observed values were calculated, and these differences were averaged by season and over the year.

3. Results and discussion

3.1. Computational efficiency

Table 5 summarizes the speed of BASE *versus* AMORE GEOS-Chem simulations through total wall-clock times. AMORE was faster than BASE by about 19 hours for the entire year simulated at 2° × 2.5° resolution. This speedup was about 10% in total and 25% in the gas- and aqueous-phase chemical reaction solver (KPP), the latter within the expected range of 20–32% reported by Lin *et al.* (2023)[18] who used an adaptive solver option to reduce comprehensive oxidant-aerosol chemistry. At 0.25° × 0.3125° resolution, AMORE was again faster than BASE, by about 2% in total and 20% in KPP. In general, the higher-resolution simulations took approximately three times longer than the coarser-resolution simulations, within the range described in Philip *et al.* (2016).[27] Finer external time steps and boundary conditions partly contribute to longer wall-clock times at higher resolutions.[18,27] Only the rates of the first two reactions in both chemical mechanisms were saved as diagnostic output by default in the higher-resolution simulations, whereas additional rates in the BASE (38 reactions) and AMORE (25 reactions) isoprene oxidation mechanisms were saved in the coarser-resolution simulations. The difference in diagnostic output between mechanisms mostly explains the greater total GEOS-Chem speedup at 2° × 2.5° resolution. Our results demonstrate that reducing the isoprene oxidation mechanism alone can significantly expedite the model. Even with the same model configurations, wall-clock times may vary depending on other processes on the system and across different high-performance computing systems. Repeated model simulations

Table 5 GEOS-Chem timer statistics comparing BASE and AMORE. "GEOS-Chem" measures the total model run time, and "KPP" measures the chemical reaction solver run time. Total wall-clock times are sums of monthly simulation times from June 2018–May 2019, reported for both the coarser and higher-resolution simulations

Timer name	Horizontal resolution	BASE (hours)	AMORE (hours)	AMORE – BASE time difference (hours)	AMORE – BASE time change (%)
GEOS-Chem	2° × 2.5°	181.2	162.6	−18.5	−10.2
KPP	2° × 2.5°	44.0	33.1	−10.8	−24.6
GEOS-Chem	0.25° × 0.3125°	496.3	485.9	−10.4	−2.1
KPP	0.25° × 0.3125°	44.9	35.8	−9.1	−20.3

would increase the confidence in the wall-clock times presented here.

3.2. Air pollutants and related chemical species

In Fig. 1, model performance statistics compare BASE, AMORE, BASE_HR, and AMORE_HR against observational data averaged over June 2018–May 2019 across the EUS. Overall, there are minor differences between BASE and AMORE as well as between BASE_HR and AMORE_HR. Two-sample *t*-tests show that, over the entire EUS model domain, there is not a statistically significant difference (p-value > 0.05) between BASE and AMORE for $PM_{2.5}$, NO_2, and total OC as well as between BASE_HR and AMORE_HR for $PM_{2.5}$, O_3, NO_2, and total OC. All simulations have low to moderate correlation ($0.25 < r < 0.51$) with observed annual mean $PM_{2.5}$, O_3, and total OC and high correlation ($0.82 < r < 0.89$) for NO_2 and HCHO. Resolution is particularly important for $PM_{2.5}$, O_3, and total OC, considering the correlation coefficient (r) increases by over 0.1 from the coarser to higher-resolution simulations, greater than differences related to the isoprene mechanism. The $PM_{2.5}$ negative bias and total OC (or OA) positive bias have been documented before, revealing uncertainties in biogenic isoprene and other aerosol species.[10,11] In addition, previous studies have indicated that GEOS-Chem overestimates O_3, in part due to excessive NO_x emissions.[49] The O_3 and HCHO positive biases are lower in both AMORE and AMORE_HR, although improvements elsewhere in the model are still likely needed to significantly reduce these biases. Likewise, AMORE_HR is less biased than BASE_HR in simulating total OC.

Spatial distributions of the nested-grid EUS simulations compared with each other and against observations are presented in Fig. 2 and 3. On average over the entire time period and at the surface, $PM_{2.5}$ is lower by 0.08 μg m^{-3} (1.4%), O_3 is lower by 0.70 ppb (2.1%), total OC is lower by 0.02 μg m^{-3} (1.3%), NO_2 is higher by 0.02 ppb (1.7%), and HCHO is lower by 0.17 ppb (9.0%) in AMORE_HR relative to BASE_HR. Fig. 2 shows that $PM_{2.5}$ is highest in the Great Lakes region (including Pennsylvania) in the model, influenced by anthropogenic emissions (Fig. 2). Both simulations underpredict $PM_{2.5}$ on average, with the mean model bias just 0.05 μg m^{-3} (1%) lower in AMORE_HR than BASE_HR. O_3 is elevated along the coastline in the model and local areas such as within the Southern Appalachians in the observations. A shallow marine PBL and topography, among other factors, can trap O_3.[3] Both simulations overpredict O_3 on average, but AMORE_HR decreases the mean model bias by 0.55 ppb (1%). Total OC is highest in the southeastern US, and AMORE_HR reduces the mean model bias by 0.02 μg m^{-3} (1%).

Compared against TROPOMI, both BASE_HR and AMORE_HR slightly overpredict NO_2 tropospheric VCD on average to about the same degree, but there are several NO_2 hotspots in urban areas that the model underpredicts (Fig. 3). Across the southeastern US, where HCHO and isoprene are enhanced in the domain, both simulations overpredict HCHO tropospheric

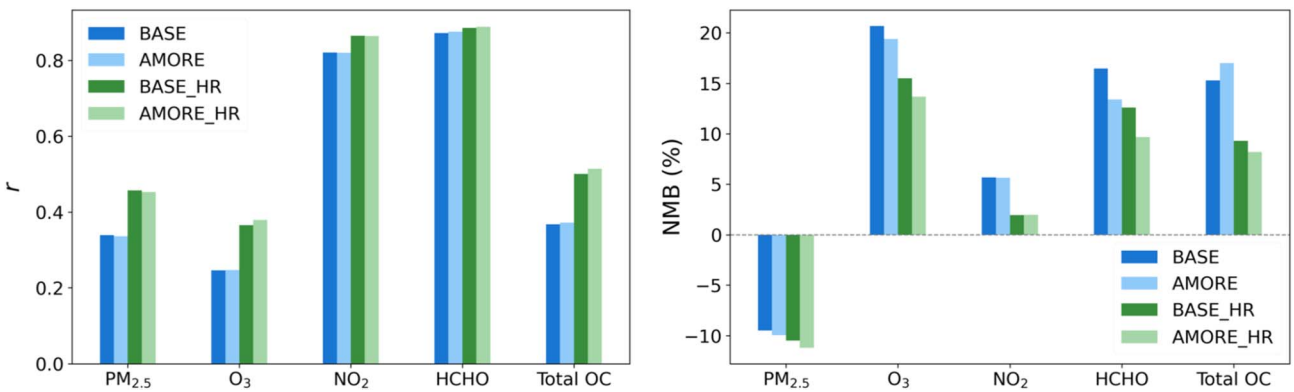

Fig. 1 GEOS-Chem model performance statistics for annual (June 2018–May 2019) mean surface $PM_{2.5}$, O_3, and total OC at AQS/IMPROVE sites and NO_2 and HCHO tropospheric VCDs at TROPOMI grid boxes across the EUS domain. Pearson's correlation coefficient (r) is on the left, and normalized mean bias (NMB) is on the right. For each of the five variables, the 2° × 2.5° (BASE and AMORE) and 0.25° × 0.3125° (BASE_HR and AMORE_HR) simulations were evaluated against observations.

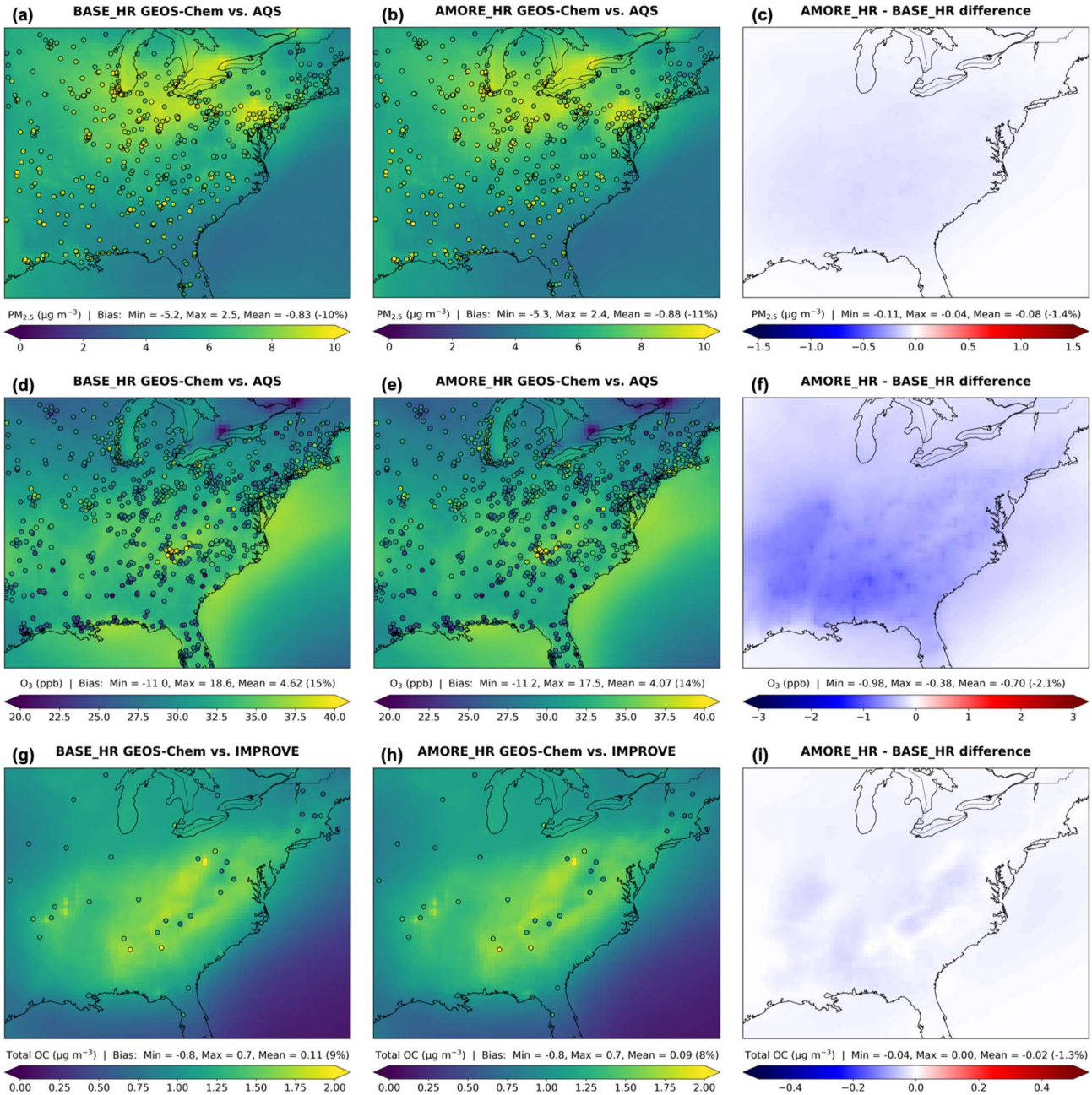

Fig. 2 EUS maps of 0.25° × 0.3125° GEOS-Chem simulated annual (June 2018–May 2019) mean surface (a–c) $PM_{2.5}$ (μg m^{-3}), (d–f) O_3 (ppb), and (g–i) total OC (μg m^{-3}). BASE_HR (a, d and g) and AMORE_HR (b, e and h) simulations are compared with AQS/IMPROVE observations (dots). Minimum, maximum, and mean model biases (modeled – observed) are calculated over EUS sites. Differences (c, f and i) between the simulations represent lower (blue) or higher (red) concentrations in AMORE_HR than BASE_HR.

VCD, but AMORE_HR decreases the EUS mean model bias by 0.17×10^{15} molecules cm^{-2} (3%). Relative to CrIS, both simulations largely underpredict isoprene total VCD across the EUS domain, possibly because of insufficient isoprene emissions and uncertainties in the satellite retrievals.[43] Isoprene is higher by 38.3% in AMORE_HR than BASE_HR at the surface, which reduces the mean model bias across the troposphere by 0.29×10^{15} molecules cm^{-2} (19%). Higher isoprene domain-wide is beneficial, including for the "isoprene volcano" or Missouri Ozarks region which features a high density of strongly isoprene-emitting oak trees.[50] However, AMORE_HR increases the existing high model bias for two prominent isoprene peaks to the south of the Missouri Ozarks.

Higher isoprene in AMORE_HR than BASE_HR suggests that less of it is oxidized in the AMORE mechanism. The main oxidant, OH, exhibits widespread lower annual mean concentrations, by 3.4% at the surface and by 1.2% across the troposphere on average globally, with the AMORE mechanism than BASE mechanism (see Fig. S1 in the ESI†). In low-NO_x conditions, reactions between isoprene peroxy radicals (IHOO1) and

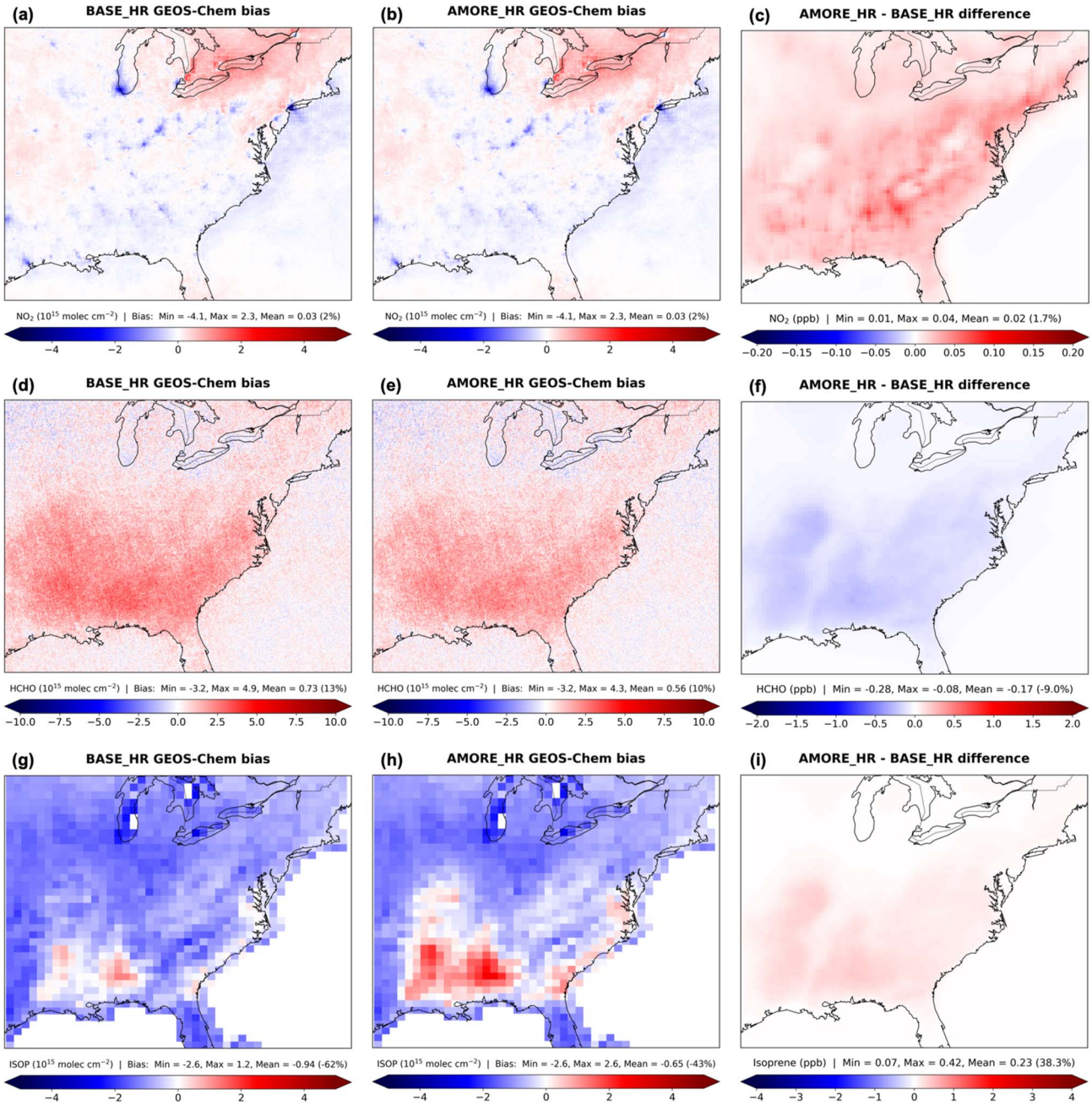

Fig. 3 A comparison between the EUS 0.25° × 0.3125° GEOS-Chem simulations and TROPOMI/CrIS observations for annual (June 2018–May 2019) mean (a and b) NO_2 tropospheric VCD, (d and e) HCHO tropospheric VCD, and (g and h) isoprene total VCD, all in units of 10^{15} molecules cm^{-2}. BASE_HR (a, d and g) and AMORE_HR (b, e and h) biases (modeled – observed) indicate either overprediction (red) or underprediction (blue). Differences in surface concentrations (c, f and i) between the simulations indicate that AMORE_HR is higher (red) or lower (blue) than BASE_HR.

HO_2 leads to OH consumption *via* hydroperoxide (RIPA) production which can be regenerated through more intramolecular hydrogen shifts.[8] Less OH recycling and thus lower OH concentrations are expected with fewer intermediate species in the AMORE mechanism, on par with other similarly sized reduced mechanisms.[19] Higher isoprene may also emphasize the low-NO_x pathway, leading to further OH depletion and hindering NO_2 production. Reaction rates for ISOP + OH and ISOP + O_3 are the same order of magnitude but lower and higher in the AMORE mechanism, respectively. Since the rate constants are about the same in both mechanisms, changes in species concentrations are likely responsible for the reaction rate differences. While the latter reaction depletes O_3 and is more competitive in the AMORE mechanism, ISOP + OH happens faster, and lower HCHO signifies the impact of less OH-initiated oxidation. A lower ISOP + NO_3 reaction rate in the AMORE mechanism than BASE mechanism, due to a rate constant discrepancy, could allow less NO_x to be lost to isoprene

nitrates. We find that the ISOP + NO_3 rate constant in the AMORE mechanism matches that in Wennberg *et al.* (2018),[17] lower by 1–2 orders of magnitude than in the BASE mechanism.

As displayed in Fig. S1 in the ESI,† global spatial differences in annual mean surface concentrations of $PM_{2.5}$, O_3, and other key chemical species between the BASE and AMORE simulations are mostly small and have the same signs as over the EUS. Isoprene chemistry most prominently differs between the two mechanisms over NO_x-limited tropical forest regions, namely the Amazon, Central Africa, and Indonesia, where isoprene is emitted in significant amounts and exhibits proportional differences. While $PM_{2.5}$, biogenic OA, and IEPOX OA are lower in these tropical forest regions with the AMORE mechanism, IEPOX OA is higher elsewhere, revealing the impact of missing non-IEPOX pathways on SOA formation. Comparing the two mechanisms for several other species of broad interest, we show that annual mean surface CO and peroxyacetyl nitrate (PAN) concentrations are lower by 2.5% and 3.1%, respectively, on average globally in AMORE. For NO_3, the sign of the difference varies from region to region, with NO_3 clearly lower in AMORE across India. Nitric acid (HNO_3) also varies regionally, but it appears to be lower in AMORE over parts of Asia and Africa where HNO_3 concentrations are relatively high.

Although higher isoprene with the AMORE mechanism reduces the global mean model bias (Fig. S2†), as background isoprene levels are underestimated, it enhances the existing overestimation for isoprene hotspots, particularly the Amazon region. Again, due to less oxidation by OH, there are widespread lower surface HCHO concentrations in AMORE, decreasing the existing positive tropospheric HCHO bias across these tropical forest regions. Globally, surface NO_2 levels are slightly higher in AMORE than BASE, especially over India, China, and other high-NO_x areas. We find that GEOS-Chem underestimates NO_2 tropospheric VCD over Europe, which may be due to other factors such as emissions, and that NO_2 in AMORE is in the right direction in such regions. Depending on the specific application of the model, the BASE mechanism might be a better choice over tropical forest regions to capture detailed chemistry, whereas the AMORE mechanism may suffice for estimating ambient air pollutants over much of the globe. However, more measurements of chemical species, including radical species such as OH, are needed to further evaluate the GEOS-Chem model.

3.3. Sensitivity of air pollutants to biogenic isoprene and anthropogenic emissions

Now, we focus our analysis on the summertime when isoprene emissions peak in the region of interest. For all seasons and sensitivity maps, see Fig. S3 and S4 in the ESI.† Fig. 4 displays average EUS summer 2018 percent changes in concentration responses of surface $PM_{2.5}$, O_3, NO_2, HCHO, and total OC to biogenic and anthropogenic emissions in BASE and AMORE compared with the eight sensitivity simulations in which emissions were zeroed out. On average, we estimate that US biogenic isoprene emissions (blue box plots) contribute to 0.5–0.6 μg m^{-3} (8–9%) of surface $PM_{2.5}$ and 1.1–1.7 ppb (3–4%) of surface O_3. US anthropogenic emissions (green box plots) contribute to about 2.2 μg m^{-3} (32–33%) of $PM_{2.5}$ and 17.6–18.1 ppb (43–45%) of O_3. Together, US biogenic isoprene and anthropogenic emissions (green box plots) contribute to about 2.6–2.7 μg m^{-3} (39–40%) of $PM_{2.5}$ and 15.6–16.1 ppb (39%) of O_3. These changes are not equal to those from separate isoprene and anthropogenic emissions summed together because of nonlinearities in the model. Adding in isoprene emissions subtly decreases O_3 levels over some of the southeastern US, perhaps scavenging OH that would otherwise produce more O_3 through reactions with other VOCs such as monoterpenes or sesquiterpenes.[15,51] Subtracting zGLB_ANTH from zUS_ANTH,

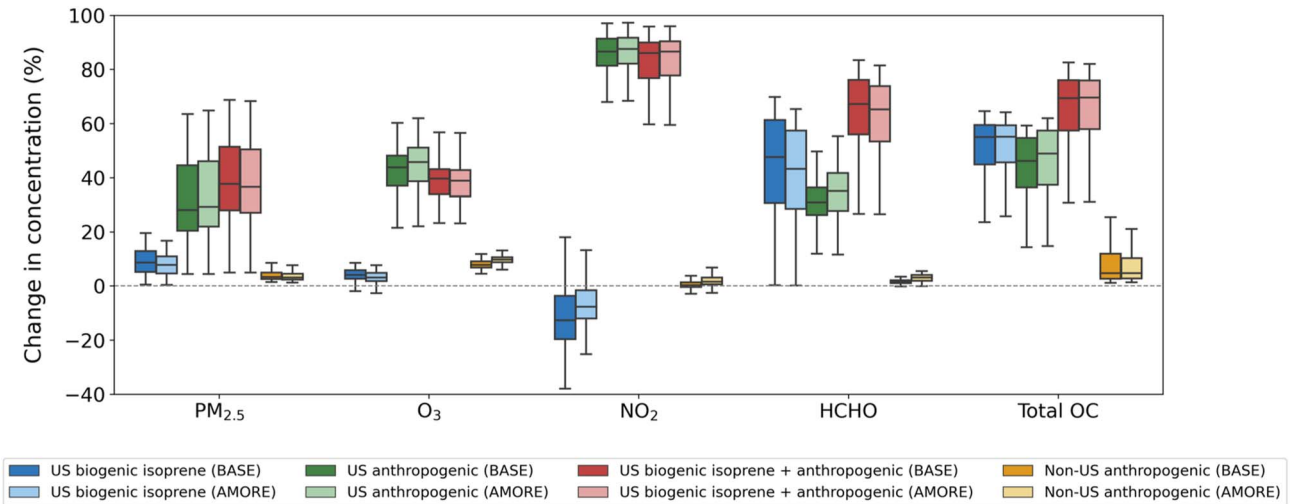

Fig. 4 Summer average (June–August 2018) sensitivities of ground-level atmospheric constituents to "adding in" different emissions across the EUS domain at 2° × 2.5° resolution. Percent changes of BASE minus BASE zero emissions (dark blue, green, and red) or AMORE minus AMORE zero emissions (light blue, green, and red) are relative to BASE or AMORE. Non-US anthropogenic (yellow) represents zUS_ANTH minus zGLB_ANTH. Each box plot shows the standard interquartile range (IQR) from the 25th (Q1) to 75th (Q3) percentiles, with whiskers extending to Q1 − (1.5 × IQR) and Q3 + (1.5 × IQR).

we estimate that non-US anthropogenic emissions (yellow box plots) contribute to 0.2 $\mu g\ m^{-3}$ (4%) of $PM_{2.5}$ and 1.8–2.1 ppb (8–9%) of O_3.

Unlike $PM_{2.5}$ and O_3, surface HCHO and total OC are much more sensitive to isoprene, with increases of 49–54% (HCHO) and 51% (total OC) from US biogenic isoprene emissions alone and increases of 68–71% (HCHO) and 64% (total OC) from the combination of US isoprene and anthropogenic emissions. US anthropogenic emissions are responsible for the majority of surface NO_2 (84–85%) which decreases slightly when adding in biogenic isoprene emissions. For all air pollutants, little differences exist between the BASE and AMORE simulations, suggesting that we can use the computationally-efficient AMORE approach without sacrificing process-level understanding gained from the perturbation simulations.

4. Conclusions

We have compared the new reduced (AMORE version 1.1) and default (BASE) isoprene oxidation mechanisms and their impacts on air pollutants in the GEOS-Chem model, both at 2° × 2.5° resolution globally and 0.25° × 0.3125° resolution over the EUS. Our study demonstrates that using the AMORE isoprene oxidation mechanism in GEOS-Chem not only saves computational resources but also achieves comparable model accuracy, including for model sensitivities to precursor emissions, over the EUS. While chemistry is important for predicting $PM_{2.5}$ and O_3, our results suggest that increasing the model resolution may improve the model performance more than changing the isoprene mechanism. Moreover, the relatively small changes in air pollutant concentrations due to isoprene support the need to continue to significantly reduce US anthropogenic emissions to improve air quality in the context of projected biogenic VOC increases in a warming climate.[4,13]

Future development of reduced chemical mechanisms can benefit from our automated graph theory-based approach. We tested the AMORE algorithm for gas-phase isoprene oxidation, but there are potential applications to aqueous-phase chemistry or other mechanisms. More work is needed to make AMORE fully automated and sufficiently flexible for use in a wide range of chemical mechanisms and atmospheric models. A limitation of this study was the amount and quality of observational data available to evaluate the GEOS-Chem model. This model evaluation would benefit from more *in situ* and remote sensing measurements of chemical species important to isoprene oxidation, such as OH, HCHO, and OA, in different photochemical regimes across the globe. For example, the Tropospheric Emissions Monitoring of Pollution (TEMPO) instrument, launched in April 2023, will provide daytime hourly air pollution information across North America from a geostationary orbit.[52] We are interfacing with the GEOS-Chem modeling team to make AMORE an option for all users in the standardized GEOS-Chem code. Improvements in computational speed and accuracy of CTMs and CCMs will allow for efficient air quality and climate forecasting, research, and management.

Data availability

The AMORE version 1.1 isoprene oxidation mechanism used in the Kinetic PreProcessor is available online: **https://github.com/benjaminyang93/amore_v1.1** (last access: 12 November 2023). Additional ESI files† including the AMORE algorithm code are available: **https://github.com/fcw2110/AMORE_supplementary_files** (last access: 12 November 2023).

Author contributions

VFM, AMF, and DMW designed the study and acquired the funding. FCW, SS, and VFM developed the automated model reduction algorithm and mechanism. BY and FCW implemented the algorithm in the GEOS-Chem model. BY and DMW set up the model and performed the simulations. BY led the data analysis. MT and AMF contributed to the model-satellite comparison. BY and FCW wrote the original draft of the paper. All authors participated in the review of the paper.

Conflicts of interest

There are no conflicts of interest to declare.

Acknowledgements

This work was funded by the US Environmental Protection Agency under grant number R840013. The authors would like to thank Kelvin H. Bates and Sidhant J. Pai, with whom we had helpful discussions.

References

1 R. Fuller, P. J. Landrigan, K. Balakrishnan, G. Bathan, S. Bose-O'Reilly, M. Brauer, J. Caravanos, T. Chiles, A. Cohen, L. Corra, M. Cropper, G. Ferraro, J. Hanna, D. Hanrahan, H. Hu, D. Hunter, G. Janata, R. Kupka, B. Lanphear, M. Lichtveld, K. Martin, A. Mustapha, E. Sanchez-Triana, K. Sandilya, L. Schaefli, J. Shaw, J. Seddon, W. Suk, M. M. Téllez-Rojo and C. Yan, Pollution and health: a progress update, *Lancet Planet. Health*, 2022, **6**, e535–e547.

2 T. Li, R. Hu, Z. Chen, Q. Li, S. Huang, Z. Zhu and L.-F. Zhou, Fine particulate matter ($PM_{2.5}$): the culprit for chronic lung diseases in China, *Chronic. Dis. Transl. Med.*, 2018, **4**, 176–186.

3 A. M. Fiore, V. Naik and E. M. Leibensperger, Air Quality and Climate Connections, *J. Air Waste Manage. Assoc.*, 2015, **65**, 645–685.

4 A. H. Goldstein, C. D. Koven, C. L. Heald and I. Y. Fung, Biogenic carbon and anthropogenic pollutants combine to form a cooling haze over the southeastern United States, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 8835–8840.

5 A. B. Guenther, X. Jiang, C. L. Heald, T. Sakulyanontvittaya, T. Duhl, L. K. Emmons and X. Wang, The Model of Emissions of Gases and Aerosols from Nature version 2.1 (MEGAN2.1): an extended and updated framework for

modeling biogenic emissions, *Geosci. Model Dev.*, 2012, **5**, 1471–1492.

6 T. D. Sharkey, A. E. Wiberley and A. R. Donohue, Isoprene Emission from Plants: Why and How, *Ann. Bot.*, 2008, **101**, 5–18.

7 D. F. McGlynn, L. E. R. Barry, M. T. Lerdau, S. E. Pusede and G. Isaacman-VanWertz, Measurement report: variability in the composition of biogenic volatile organic compounds in a Southeastern US forest and their role in atmospheric reactivity, *Atmos. Chem. Phys.*, 2021, **21**, 15755–15770.

8 K. H. Bates and D. J. Jacob, A new model mechanism for atmospheric oxidation of isoprene: global effects on oxidants, nitrogen oxides, organic products, and secondary organic aerosol, *Atmos. Chem. Phys.*, 2019, **19**, 9613–9640.

9 D. J. Jacob, *Introduction to Atmospheric Chemistry*, Princeton University Press, 1999.

10 P. S. Kim, D. J. Jacob, J. A. Fisher, K. Travis, K. Yu, L. Zhu, R. M. Yantosca, M. P. Sulprizio, J. L. Jimenez, P. Campuzano-Jost, K. D. Froyd, J. Liao, J. W. Hair, M. A. Fenn, C. F. Butler, N. L. Wagner, T. D. Gordon, A. Welti, P. O. Wennberg, J. D. Crounse, J. M. St. Clair, A. P. Teng, D. B. Millet, J. P. Schwarz, M. Z. Markovic and A. E. Perring, Sources, seasonality, and trends of southeast US aerosol: an integrated analysis of surface, aircraft, and satellite observations with the GEOS-Chem chemical transport model, *Atmos. Chem. Phys.*, 2015, **15**, 10411–10433.

11 S. J. Pai, C. L. Heald, J. R. Pierce, S. C. Farina, E. A. Marais, J. L. Jimenez, P. Campuzano-Jost, B. A. Nault, A. M. Middlebrook, H. Coe, J. E. Shilling, R. Bahreini, J. H. Dingle and K. Vu, An evaluation of global organic aerosol schemes using airborne observations, *Atmos. Chem. Phys.*, 2020, **20**, 2637–2665.

12 F. Paulot, J. D. Crounse, H. G. Kjaergaard, A. Kürten, J. M. St. Clair, J. H. Seinfeld and P. O. Wennberg, Unexpected Epoxide Formation in the Gas-Phase Photooxidation of Isoprene, *Science*, 2009, **325**, 730–733.

13 A. M. Fiore, L. W. Horowitz, D. W. Purves, H. Levy II, M. J. Evans, Y. Wang, Q. Li and R. M. Yantosca, Evaluating the contribution of changes in isoprene emissions to surface ozone trends over the eastern United States, *J. Geophys. Res.: Atmos.*, 2005, **110**, D1203.

14 D. J. Jacob and D. A. Winner, Effect of climate change on air quality, *Atmos. Environ.*, 2009, **43**, 51–63.

15 R. H. Schwantes, L. K. Emmons, J. J. Orlando, M. C. Barth, G. S. Tyndall, S. R. Hall, K. Ullmann, J. M. St. Clair, D. R. Blake, A. Wisthaler and T. P. V. Bui, Comprehensive isoprene and terpene gas-phase chemistry improves simulated surface ozone in the southeastern US, *Atmos. Chem. Phys.*, 2020, **20**, 3739–3776.

16 M. E. Jenkin, J. C. Young and A. R. Rickard, The MCM v3.3.1 degradation scheme for isoprene, *Atmos. Chem. Phys.*, 2015, **15**, 11433–11459.

17 P. O. Wennberg, K. H. Bates, J. D. Crounse, L. G. Dodson, R. C. McVay, L. A. Mertens, T. B. Nguyen, E. Praske, R. H. Schwantes, M. D. Smarte, J. M. St Clair, A. P. Teng, X. Zhang and J. H. Seinfeld, Gas-Phase Reactions of Isoprene and Its Major Oxidation Products, *Chem. Rev.*, 2018, **118**, 3337–3390.

18 H. Lin, M. S. Long, R. Sander, A. Sandu, R. M. Yantosca, L. A. Estrada, L. Shen and D. J. Jacob, An Adaptive Auto-Reduction Solver for Speeding Up Integration of Chemical Kinetics in Atmospheric Chemistry Models: Implementation and Evaluation in the Kinetic Pre-Processor (KPP) Version 3.0.0, *J. Adv. Model. Earth Syst.*, 2023, **15**, e2022MS003293.

19 F. Wiser, B. K. Place, S. Sen, H. O. T. Pye, B. Yang, D. M. Westervelt, D. K. Henze, A. M. Fiore and V. F. McNeill, AMORE-Isoprene v1.0: a new reduced mechanism for gas-phase isoprene oxidation, *Geosci. Model Dev.*, 2023, **16**, 1801–1821.

20 Z. M. Nikolaou, J.-Y. Chen, Y. Proestos, J. Lelieveld and R. Sander, Accelerating simulations using REDCHEM_v0.0 for atmospheric chemistry mechanism reduction, *Geosci. Model Dev.*, 2018, **11**, 3391–3407.

21 S. J. Silva, S. M. Burrows, M. J. Evans and M. Halappanavar, A Graph Theoretical Intercomparison of Atmospheric Chemical Mechanisms, *Geophys. Res. Lett.*, 2021, **48**, e2020GL090481.

22 The International GEOS-Chem User Community, *geoschem/KPP: KPP-for-GEOS-Chem 2.3.3_gc*, Zenodo, DOI: **10.5281/zenodo.5140218**.

23 V. Damian, A. Sandu, M. Damian, F. Potra and G. R. Carmichael, The kinetic preprocessor KPP-a software environment for solving chemical kinetics, *Comput. Chem. Eng.*, 2002, **26**, 1567–1579.

24 A. Sandu and R. Sander, Technical note: Simulating chemical systems in Fortran90 and Matlab with the Kinetic PreProcessor KPP-2.1, *Atmos. Chem. Phys.*, 2006, **6**, 187–195.

25 The International GEOS-Chem User Community, *geoschem/GCClassic: GEOS-Chem 13.3.3*, Zenodo, DOI: **10.5281/zenodo.5748260**.

26 E. A. Marais, D. J. Jacob, J. L. Jimenez, P. Campuzano-Jost, D. A. Day, W. Hu, J. Krechmer, L. Zhu, P. S. Kim, C. C. Miller, J. A. Fisher, K. Travis, K. Yu, T. F. Hanisco, G. M. Wolfe, H. L. Arkinson, H. O. T. Pye, K. D. Froyd, J. Liao and V. F. McNeill, Aqueous-phase mechanism for secondary organic aerosol formation from isoprene: application to the southeast United States and co-benefit of SO_2 emission controls, *Atmos. Chem. Phys.*, 2016, **16**, 1603–1618.

27 S. Philip, R. V. Martin and C. A. Keller, Sensitivity of chemistry-transport model simulations to the duration of chemical and transport operators: a case study with GEOS-Chem v10-01, *Geosci. Model Dev.*, 2016, **9**, 1683–1695.

28 Global Modeling and Assimilation Office (GMAO), *MERRA-2 inst3_3d_asm_Nv: 3d, 3-Hourly, Instantaneous, Model-Level, Assimilated Meteorological Fields V5.12.4, Goddard Space Flight Cent. Distrib. Act. Arch. Cent.*, GSFC DAAC, DOI: **10.5067/WWQSXQ8IVFW8**.

29 Global Modeling and Assimilation Office (GMAO), *MERRA-2 tavg1_2d_slv_Nx: 2d, 1-Hourly, Time-Averaged, Single-Level, Assimilated Diagnostics V5.12.4, Goddard Earth Sci. Data Inf. Serv. Cent*, GES DISC, DOI: **10.5067/VJAFPLI1CSIV**.

30 Y. X. Wang, M. B. McElroy, D. J. Jacob and R. M. Yantosca, A nested grid formulation for chemical transport over Asia: Applications to CO, *J. Geophys. Res.: Atmos.*, 2004, **109**, D22307.

31 M. Tao, A. M. Fiore, X. Jin, L. D. Schiferl, R. Commane, L. M. Judd, S. Janz, J. T. Sullivan, P. J. Miller, A. Karambelas, S. Davis, M. Tzortziou, L. Valin, A. Whitehill, K. Civerolo and Y. Tian, Investigating Changes in Ozone Formation Chemistry during Summertime Pollution Events over the Northeastern United States, *Environ. Sci. Technol.*, 2022, **56**, 15312–15327.

32 R. Lucchesi, *File Specification for GEOS FP, GMAO Off. Note No 4 Version 12*.

33 C. A. Keller, M. S. Long, R. M. Yantosca, A. M. Da Silva, S. Pawson and D. J. Jacob, HEMCO v1.0: a versatile, ESMF-compliant component for calculating emissions in atmospheric models, *Geosci. Model Dev.*, 2014, **7**, 1409–1417.

34 R. M. Hoesly, S. J. Smith, L. Feng, Z. Klimont, G. Janssens-Maenhout, T. Pitkanen, J. J. Seibert, L. Vu, R. J. Andres, R. M. Bolt, T. C. Bond, L. Dawidowski, N. Kholod, J. Kurokawa, M. Li, L. Liu, Z. Lu, M. C. P. Moura, P. R. O'Rourke and Q. Zhang, Historical (1750–2014) anthropogenic emissions of reactive gases and aerosols from the Community Emissions Data System (CEDS), *Geosci. Model Dev.*, 2018, **11**, 369–408.

35 Z. A. Tzompa-Sosa, E. Mahieu, B. Franco, C. A. Keller, A. J. Turner, D. Helmig, A. Fried, D. Richter, P. Weibring, J. Walega, T. I. Yacovitch, S. C. Herndon, D. R. Blake, F. Hase, J. W. Hannigan, S. Conway, K. Strong, M. Schneider and E. V. Fischer, Revisiting global fossil fuel and biofuel emissions of ethane, *J. Geophys. Res.: Atmos.*, 2017, **122**, 2493–2512.

36 Y. Xiao, J. A. Logan, D. J. Jacob, R. C. Hudman, R. Yantosca and D. R. Blake, Global budget of ethane and regional constraints on U.S. sources, *J. Geophys. Res.: Atmos.*, 2008, **113**, D21306.

37 M. E. J. Stettler, S. Eastham and S. R. H. Barrett, Air quality and public health impacts of UK airports. Part I: Emissions, *Atmos. Environ.*, 2011, **45**, 5415–5424.

38 S. Philip, R. V. Martin, G. Snider, C. L. Weagle, A. van Donkelaar, M. Brauer, D. K. Henze, Z. Klimont, C. Venkataraman, S. K. Guttikunda and Q. Zhang, Anthropogenic fugitive, combustion and industrial dust is a significant, underrepresented fine particulate matter source in global atmospheric models, *Environ. Res. Lett.*, 2017, **12**, 044018.

39 D. L. Goldberg, M. Harkey, B. de Foy, L. Judd, J. Johnson, G. Yarwood and T. Holloway, Evaluating NO_x emissions and their effect on O_3 production in Texas using TROPOMI NO_2 and HCHO, *Atmos. Chem. Phys.*, 2022, **22**, 10875–10900.

40 C. Vigouroux, B. Langerock, C. A. Bauer Aquino, T. Blumenstock, Z. Cheng, M. De Mazière, I. De Smedt, M. Grutter, J. W. Hannigan, N. Jones, R. Kivi, D. Loyola, E. Lutsch, E. Mahieu, M. Makarova, J.-M. Metzger, I. Morino, I. Murata, T. Nagahama, J. Notholt, I. Ortega, M. Palm, G. Pinardi, A. Röhling, D. Smale, W. Stremme, K. Strong, R. Sussmann, Y. Té, M. van Roozendael, P. Wang and H. Winkler, TROPOMI–Sentinel-5 Precursor formaldehyde validation using an extensive network of ground-based Fourier-transform infrared stations, *Atmos. Meas. Tech.*, 2020, **13**, 3751–3767.

41 H. Eskes and K.-U. Eichmann, *S5P Mission Performance Centre Nitrogen Dioxide [L2__NO2___] Readme*, Eur. Space Agency ESA, DOI: **10.5270/S5P-9bnp8q8**.

42 I. De Smedt, F. Romahn and K.-U. Eichmann, *S5P Mission Performance Centre Formaldehyde [L2__HCHO___] Readme*, Eur. Space Agency ESA, DOI: **10.5270/S5P-vg1i7t0**.

43 K. C. Wells, D. B. Millet, V. H. Payne, C. Vigouroux, C. a. B. Aquino, M. De Mazière, J. A. de Gouw, M. Graus, T. Kurosu, C. Warneke and A. Wisthaler, Next-Generation Isoprene Measurements From Space: Detecting Daily Variability at High Resolution, *J. Geophys. Res.: Atmos.*, 2022, **127**, e2021JD036181.

44 K. C. Wells and D. B. Millet, *ROCR Isoprene Retrievals from the CrIS Satellite Sensor, Univ. Minn. Data Repos.*, DOI: **10.13020/5n0j-wx73**.

45 US Environmental Protection Agency (EPA), *Air Data*, **https://www.epa.gov/outdoor-air-quality-data**, accessed 12 November 2023.

46 IMPROVE, *Federal Land Manager Environmental Database (FED)*, **http://views.cira.colostate.edu/fed/**, accessed 12 November 2023.

47 European Space Agency (ESA), *Sentinel-5P TROPOMI Tropospheric NO2 1-Orbit L2 7km x 3.5km V1, NASA Goddard Earth Sci. Data Inf. Serv. Cent*, GES DISC, DOI: **10.5270/S5P-s4ljg54**.

48 European Space Agency (ESA), *Sentinel-5P TROPOMI Tropospheric Formaldehyde HCHO 1-Orbit L2 7km x 3.5km V1, NASA Goddard Earth Sci. Data Inf. Serv. Cent*, GES DISC, DOI: **10.5270/S5P-tjlxfd2**.

49 K. R. Travis, D. J. Jacob, J. A. Fisher, P. S. Kim, E. A. Marais, L. Zhu, K. Yu, C. C. Miller, R. M. Yantosca, M. P. Sulprizio, A. M. Thompson, P. O. Wennberg, J. D. Crounse, J. M. St. Clair, R. C. Cohen, J. L. Laughner, J. E. Dibb, S. R. Hall, K. Ullmann, G. M. Wolfe, I. B. Pollack, J. Peischl, J. A. Neuman and X. Zhou, Why do models overestimate surface ozone in the Southeast United States?, *Atmos. Chem. Phys.*, 2016, **16**, 13561–13577.

50 C. Wiedinmyer, J. Greenberg, A. Guenther, B. Hopkins, K. Baker, C. Geron, P. I. Palmer, B. P. Long, J. R. Turner, G. Pétron, P. Harley, T. E. Pierce, B. Lamb, H. Westberg, W. Baugh, M. Koerber and M. Janssen, Ozarks Isoprene Experiment (OZIE): Measurements and modeling of the "isoprene volcano", *J. Geophys. Res.: Atmos.*, 2005, **110**, D18307.

51 G. McFiggans, T. F. Mentel, J. Wildt, I. Pullinen, S. Kang, E. Kleist, S. Schmitt, M. Springer, R. Tillmann, C. Wu, D. Zhao, M. Hallquist, C. Faxon, M. Le Breton, Å. M. Hallquist, D. Simpson, R. Bergström, M. E. Jenkin, M. Ehn, J. A. Thornton, M. R. Alfarra, T. J. Bannan,

C. J. Percival, M. Priestley, D. Topping and A. Kiendler-Scharr, Secondary organic aerosol reduced by mixture of atmospheric vapours, *Nature*, 2019, **565**, 587–593.

52 Smithsonian Astrophysical Observatory and National Aeronautics and Space Administration, *Tropospheric Emissions: Monitoring of Pollution (TEMPO)*, **https://tempo.si.edu**, accessed 12 November 2023.