Role of chemical production and depositional losses on formaldehyde in the Community Regional Atmospheric Chemistry Multiphase Mechanism (CRACMM)

T. Nash Skipper[1,2], Emma L. D'Ambro[2], Forwood C. Wiser[3], V. Faye McNeill[3,4], Rebecca H. Schwantes[5], Barron H. Henderson[6], Ivan R. Piletic[2], Colleen B. Baublitz[6], Jesse O. Bash[2], Andrew R. Whitehill[2], Lukas C. Valin[2], Asher P. Mouat[7], Jennifer Kaiser[7,8], Glenn M. Wolfe[9], Jason M. St. Clair[9,10], Thomas F. Hanisco[9], Alan Fried[11], Bryan K. Place[9,12], and Havala O.T. Pye[2]

[1] Oak Ridge Institute for Science and Education, Office of Research and Development, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina, USA
[2] Office of Research and Development, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina, USA
[3] Department of Chemical Engineering, Columbia University, New York, New York, USA
[4] Department of Earth and Environmental Sciences, Columbia University, New York, New York, USA
[5] Chemical Sciences Laboratory, National Oceanic and Atmospheric Administration, Boulder, Colorado, USA
[6] Office of Air and Radiation, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina, USA
[7] School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA, USA
[8] School of Earth and Atmospheric Sciences, Georgia Institute of Technology, Atlanta, GA, USA
[9] Atmospheric Chemistry and Dynamics Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD, USA
[10] Joint Center for Earth Systems Technology, University of Maryland Baltimore County, Baltimore, MD, USA
[11] Institute of Arctic and Alpine Research (INSTAAR), University of Colorado, Boulder, CO, USA
[12] SciGlob Instruments and Services, LLC, Columbia, MD, USA

Correspondence to: Havala O.T. Pye (pye.havala@epa.gov)

Abstract. Formaldehyde (HCHO) is an important air pollutant due to its direct health effects as an air toxic that contributes to elevated cancer risk, its role in ozone formation, and its role as a product from oxidation of most gas phase reactive organic carbon (ROC). We make several updates affecting secondary production of HCHO from ROC in the Community Regional Atmospheric Chemistry Multiphase Mechanism (CRACMM) in the Community Multiscale Air Quality (CMAQ) model. ~~Secondary~~ Production of HCHO from isoprene and monoterpenes is increased, correcting an underestimate in the current version. Simulated 2019 June–August surface HCHO during peak photochemical production (11am–3pm) increased by 0.6 ppb (32%) over the southeastern US and by 0.2 ppb (13%) over the entire contiguous US. The increased HCHO compares more favorably with satellite-based observations from TROPOMI and observations from an aircraft campaign. Evaluation against hourly surface observations indicates a missing nighttime sink for HCHO which can be ~~ameliorated~~ improved by adding bidirectional exchange of HCHO and a leaf wetness dependent deposition process which increases nighttime deposition, decreasing 2019 June–August nocturnal (8pm–4am) surface HCHO by 1.1 ppb (36%) over the southeastern US and 0.5 ppb (29%) over the entire contiguous US. The ability of CRACMM to capture peak levels of HCHO at midday is improved, particularly at sites in the northeastern US, while peak levels at southeastern US sites are improved though still lower than observed. Using established risk assessment methods, lifetime exposure of the contiguous U.S. population (~320

million) to ambient HCHO levels predicted here may result in 6200 lifetime cancer cases, 40% of which are from controllable anthropogenic emissions of nitrogen oxides and reactive organic compounds. Chemistry updates will be available in CRACMM version 2 (CRACMM2) in CMAQv5.5.

1 Introduction

Formaldehyde (HCHO) is a gas-phase reactive organic compound designated as a hazardous air pollutant (HAP) by the U.S. Environmental Protection Agency (EPA). It is among the top three species contributing to noncancer health risk and the leading driver of cancer risk from ambient exposure to inhaled air toxics in the United States (Scheffe et al., 2016; Strum and Scheffe, 2016). EPA's 2019 AirToxScreen assessment estimates a nationwide average cancer risk of ~15 in a million for HCHO, about half of the total national average cancer risk from ambient exposure to air toxics (https://www.epa.gov/AirToxScreen). HCHO is also an important oxidation product and indicator of gas-phase chemistry. Once formed, HCHO can be a source of radicals that modulate cycling of nitrogen oxides (NO_X=NO+NO_2) and thus formation of the criteria pollutant ozone (O_3). HCHO is quantified through remote sensing and has been used to provide top-down constraints on emissions of isoprene and other precursor species through inverse modeling (Fortems-Cheiney et al., 2012; Kaiser et al., 2018; Oomen et al., 2024) and, along with satellite-based observations of NO_2, to characterize O_3 chemical regimes (Martin et al., 2004; Duncan et al., 2010; Tao et al., 2022).

The abundance of ambient HCHO is influenced by both primary emissions of HCHO and its precursors as well as atmospheric chemistry. Primary HCHO is emitted by many sources as a combustion byproduct as well as from natural sources. Biogenic sources such as vegetation are the largest source of primary HCHO in the US (~1200 Gg yr^{-1}) with other major sources including fires (~300 Gg yr^{-1}), mobile sources (~40 Gg yr^{-1}), oil and gas (~20 Gg yr^{-1}), and wood burning for residential heating (~20 Gg yr^{-1}) (Foley et al. (2023) based on 2017 National Emissions Inventory (NEI), Fig. S1). HCHO's short lifetime of only a few hours against photolysis and reaction with the OH radical means impacts of primary HCHO are typically localized near source (characteristic transport of ~30 km for 3 h lifetime with 3 m s^{-1} wind speed). Secondary production tends to dominate over primary sources in driving total abundance, particularly in warmer months when HCHO levels are the highest (Dix et al., 2023). HCHO is produced from oxidation of nearly every gas-phase reactive organic carbon (ROC) species with isoprene being the biggest source of secondary HCHO. Other important precursors include methane and alkenes (Luecken et al., 2012).

EPA's AirToxScreen as well as inverse modeling for emission estimation rely on chemical transport models (CTMs) to simulate HCHO. Specifically, as part of AirToxScreen, ambient exposure levels of air toxics are obtained from concentrations predicted by the Community Multiscale Air Quality (CMAQ) CTM combined with local scale information from a dispersion model (U.S. EPA, 2022a), and CMAQ alone provides the estimates of secondary HCHO. CMAQ has been previously reported to underestimate HCHO (Luecken et al., 2012; Luecken et al., 2018) which could propagate to errors in predictions of health risk. Combined with the national population for 2019, the AirToxScreen nationwide cancer risk from HCHO (2019 value) implies ~4800 cancer cases result from lifetime exposure. However, Zhu et al. (2017) estimated between 6600 and 12500 cancer cases based on exposures derived from satellite-based HCHO observations. A more accurate representation of secondary HCHO could improve inverse modeling estimates of emissions as well as our understanding of the role of ambient HCHO in inhalation health risks.

Here, we focus on the representation of secondary HCHO production in the Community Regional Atmospheric Chemistry Multiphase Mechanism (CRACMM). CRACMM is designed to integrate modeling of O_3, $PM_{2.5}$, and HAPs and has been primarily applied in CMAQ (Pye et al., 2023). We make several updates to CRACMM version 1 (CRACMM1), leading to CRACMM version 2 (CRACMM2). Most of the updates in CRACMM2 target HCHO, and additional updates for completeness are documented here for users of CMAQ and CRACMM2. Chemistry updates were screened with a box model, the Framework for 0-D Atmospheric Modeling (F0AM) (Wolfe et al., 2016a), and then tested in a series of regional CMAQ simulations covering the contiguous US (CONUS). The performance of CRACMM (1 and 2) in CMAQ ~~are~~ is evaluated with a suite of observations including satellite based HCHO from TROPOspheric Monitoring Instrument (TROPOMI), observations from an aircraft campaign, and hourly surface observations from several field deployments. Based on the evaluation, sensitivity simulations are conducted to explore areas for future improvement of HCHO in CMAQ CRACMM. These include simulations with updates to HCHO dry deposition to reduce high nocturnal biases. Estimates of cancer risk from ambient exposure to HCHO derived from CMAQ CRACMM are provided along with an estimate of the portion of cancer risk that is controllable through reductions in anthropogenic NO_X and ROC emissions.

2 Chemistry updates

CRACMM2 includes updates to several chemical systems which are discussed below. CRACMM1 is described in detail by Pye et al. (2023) and has been applied within CMAQ for the northeastern US to investigate O_3 (Place et al., 2023), CONUS during summer to investigate $PM_{2.5}$ (Vannucci et al., 2024), and CONUS to investigate SOA from asphalt paving (Seltzer et al., 2023). Some relevant details on CRACMM1 chemistry are given here for comparison with CRACMM2. A list of all reactions that have been updated or added in CRACMM2 ~~are~~ is provided in Table S~~3~~1. In total, the number of reactions is increased from 508 to 531 and the number of species (gases and particles) is increased from 229 to 239 in CRACMM2 compared to CRACMM1.

~~3~~2.1 AMORE isoprene

CRACMM1 included two options for isoprene chemistry. The main mechanism used isoprene chemistry based on RACM2 (Goliff et al., 2013; Sarwar et al., 2013) with additional IEPOX chemistry which is not included in the native RACM2 representation. A separate version of CRACMM included the Automated Model Reduction (AMORE) version 1.0 (Wiser et al., 2023) condensation of a detailed isoprene mechanism (Wennberg et al., 2018) and was referred to as CRACMM1AMORE in CMAQ. The development of the AMORE mechanism condensation technique is described in detail by Wiser et al. (2023). Briefly, AMORE takes the full mechanism along with a list of priority species, estimates the sensitivity of the full mechanism to variations in important species, and builds a reduced mechanism which emulates the sensitivity of the full mechanism. The AMOREv1.1 representation of isoprene chemistry was implemented in GEOS-Chem which yielded similar model performance with improved computational speed compared to the default GEOS-Chem mechanism (Yang et al., 2023). AMOREv1.2 is the default isoprene representation in CRACMM2, merging the base CRACMM and CRACMM-AMORE mechanisms, to better represent isoprene degradation productions and SOA precursors. AMOREv1.2, developed here (see SI for additional details), is intended to produce better NO_X cycling and O_3 compared to CRACMM1AMORE and adds new SOA precursors. New gas phase species are INO2 (isoprene nitrooxy peroxy radical), IPX (lumped isoprene tetrafunctional compounds), and INALD (lumped isoprene nitrates). Two new SOA species were added as products of heterogeneous uptake of IPX and INALD (see Sect. ~~3~~2.3 for details). In addition, HCHO yields were updated to more closely follow the detailed mechanism by Wennberg et al. (2018) based on box model testing (Fig. S~~5~~3).

~~3~~2.2 Methane

CMAQ specifies a fixed global background methane value of 1850 ppb by default, although the level can be modified by the user. CRACMM1 includes one methane reaction with OH, but the Carbon Bond family of mechanisms in CMAQ also include a reactive tracer species for emitted methane (ECH4) to capture the effects of local methane emissions on top of the global background. CRACMM2 adds the ECH4 species and includes a single ECH4 reaction with OH that is identical to the global methane reaction with OH from CRACMM1. Adding ECH4 can have small effects on secondary HCHO production as the methyl peroxy radical (MO2) produced from ECH4 + OH is a source of HCHO through reaction with NO and peroxy radical (RO_2) cross reactions. ECH4 is typically small compared to the global background methane value of 1850 ppb and only has notable impacts on other model species near sources with high ECH4 emissions.

~~3~~2.3 Heterogeneous uptake

Four new heterogeneous uptake pathways have been added to CRACMM2. Two of these are heterogeneous uptake of isoprene-derived species from AMORE to form SOA. Lumped isoprene tetrafunctional compounds (IPX) form AISO4, and lumped isoprene nitrates (INALD) form AISO5 and nitric acid. Uptake of these species is expected to behave similarly to IEPOX uptake, so we base their uptake rates on the existing IEPOX uptake rate in CMAQ (Pye et al., 2013; Pye et al., 2017). Uptake of IPX is scaled to two times the IEPOX uptake rate, and uptake of INALD is scaled to half of the IEPOX uptake rate. We also add heterogeneous uptake of HO_2 and nitrate radicals in CRACMM2. Heterogeneous uptake of HO_2 radicals has been included in other CTMs but not yet in any released version of CMAQ. Uptake of HO_2 tends to reduce O_3 and may be particularly important when aerosol concentrations are very high (Ivatt et al., 2022). CRACMM2 uses an uptake coefficient of $\gamma=0.2$ and assumes that HO_2 produces only water (Ivatt et al., 2022). More complex parameterizations involving HO_2 uptake catalyzed by copper and iron have sometimes been employed (Mao et al., 2013), but the simpler version that we have opted for in CRACMM2 is commonly employed in other CTMs such as GEOS-Chem. Heterogeneous uptake is a potential sink for nitrate radicals which may influence nighttime chemistry when oxidation by nitrate radicals takes place. CRACMM2 uses an uptake coefficient of $\gamma=10^{-3}$ for nitrate and assumes that nitrate radical uptake produces nitric acid (Jacob, 2000; Zhu et al., 2024). These heterogeneous uptake additions are not targeted towards improving HCHO but are implemented primarily for their effects on SOA (uptake of isoprene-derived compounds) and O_3 (radical uptake).

~~3~~2.4 Monoterpenes

Monoterpenes in CRACMM are categorized based on their number of double bonds as either API (represented with the chemistry of α-pinene) or LIM (represented with the chemistry of limonene). Monoterpene chemistry in CRACMM1 was largely based on the Master Chemical Mechanism (MCM) (Jenkin et al., 1997; Saunders et al., 2003) with additional updates including autoxidation pathways. After α-pinene, β-pinene is one of the most abundant monoterpenes from biogenic emissions (Guenther et al., 2012). In CRACMM, both α-pinene and β-pinene are represented by the lumped species API; however, the yield of HCHO from these monoterpenes differs significantly as the presence of the exocyclic terminal double bond in β-pinene leads to greater HCHO production. Experimental yields of HCHO from α-pinene have been reported as 0.16–0.23 (Nozière et al., 1999; Orlando et al., 2000; Lee et al., 2006) while yields from β-pinene have been reported as 0.45–0.53 (Hatakeyama et al., 1991; Orlando et al., 2000; Lee et al., 2006). HCHO from API in CRACMM1 is underestimated in part because the larger yield from β-pinene is not accounted for. Limonene HCHO yields are also likely underestimated as the yield in CRACMM1 (0.28) is less than what has been reported in experimental results (0.43) (Lee et al., 2006).

In CRACMM2, monoterpene chemistry has been updated based on work by Schwantes et al. (2020) which primarily used experimental results to determine monoterpene oxidation products and yields as part of the development of an update to the Model of Ozone And Related chemical Tracers (MOZART) chemical mechanism (Emmons et al., 2020). In the updated MOZART mechanism (MOZART-TS2), monoterpenes are grouped into four species represented by α-pinene, β-pinene, limonene, and myrcene which each have unique detailed chemical evolution. Some deviations and simplifications are made in porting the MOZART-TS2 chemistry to CRACMM2. To manage the computational burden of CRACMM2, we retain the two monoterpene species from CRACMM1 (API and LIM) and map species from MOZART-TS2 to existing species from CRACMM1. We retain the behavior from CRACMM1 such that a fraction of the peroxy radicals formed from oxidation of a monoterpene by OH or nitrate (2.5% for API; 5.5% for LIM) undergo rapid autoxidation based on Piletic and Kleindienst (2022). The products from the remaining monoterpene peroxy radicals (i.e., those that do not undergo rapid autoxidation) and from ozonolysis of monoterpenes are updated based on Schwantes et al. (2020).

API products are updated to include products from both α-pinene and β-pinene. We assume 65% of products are from α-pinene and 35% are from β-pinene based on the relative emissions of these species estimated by the Biogenic Emission Inventory

System (BEIS) over the 12 km CONUS modeling domain (Fig. S~~8~~6). We avoid adding a new β-pinene species to CRACMM2 because it requires adding around 30 new reactions to represent β-pinene oxidation and RO_2 fate which was deemed too computationally expensive and because the reactivity of α-pinene and β-pinene are similar enough to be represented with the same reaction for both species. The updates to monoterpene reactions and products are detailed in Table S~~3~~1. Most notably for HCHO, the effective yield from API+OH RO_2 (APIP1) has increased from 0 to 0.31 for RO_2+NO; from 0 to 0.40 for RO_2+NO_3, and from 0 to 0.06 for RO_2+HO_2. HCHO yields for RO_2+RO_2 cross reactions involving APIP1 have also increased. HCHO yields increased from 0 to 0.46 for API ozonolysis. In CRACMM1, HCHO from API was exclusively due to later generation chemistry involving pinonaldehyde (species PINAL). CRACMM2 forms HCHO in earlier generations and brings HCHO yields more in line with experimental yields. LIM products are updated based on the limonene representation from MOZART-TS2 where the most significant updates for HCHO are an increase in the yield of HCHO from LIM+OH RO_2 (LIMP1) from 0.28 to 0.43 for RO_2+NO and an increase from 0 to 0.33 for LIM ozonolysis.

CRACMM1 includes two monoterpene aldehydes based on pinonaldehyde (species PINAL) and limonaldehyde (species LIMAL) which react with OH to produce peroxy radicals (PINALP and LIMALP) and acyl peroxy radicals (species RCO3). PINALP and LIMALP react with NO and HO_2 but can also form highly oxygenated organic molecules (HOM) with an autoxidation rate of 1 s^{-1} in CRACMM1. Box model testing indicated that this autoxidation rate made the bimolecular NO and HO_2 channels uncompetitive at typical atmospheric levels of NO and HO_2 (i.e., essentially all PINALP and LIMALP would autoxidize and make HOM), so the autoxidation rates are updated for CRACMM2. The autoxidation of PINALP and LIMALP will proceed via multiple steps involving slightly different mechanistic pathways because of differences in chemical structure. All H-shift rates are approximated using the structure activity relationships developed by Vereecken and Nozière (2020). Specifically, an OH initiation reaction with PINAL will produce an acyl peroxide radical as the dominant product (represented by CRACMM species PINALP) while the same reaction will produce a tertiary peroxy radical via OH addition to the double bond in LIMAL. For LIMAL, a subsequent 1,6-H shift that abstracts the aldehyde H at a rate of 0.29 s^{-1} gives rise to an analogous albeit more oxidized acyl peroxide radical (represented by CRACMM species LIMALP). At this point, both acyl peroxy radicals will likely abstract from a tertiary carbon via a 1,5-H shift that is fairly rapid (0.7 s^{-1}) and a subsequent 1,5-H shift from the β-oxo site produces HOM radicals at a rate of 0.02 – 0.03 s^{-1}. Given that the latest generation autoxidation reaction is the slowest, it was used to approximate the overall autoxidation rate. This approximation simplifies the modeled autoxidation process because alternative pathways may exist including cyclobutyl ring opening following H abstraction for PINAL (Iyer et al., 2021) or peroxy radical ring closure reactions for LIMAL if the initiation step extracts the aldehyde H (Piletic and Kleindienst, 2022). Within this approximation, the autoxidation rates of monoterpene aldehydes have been updated in CRACMM2 to 0.029 s^{-1} for PINALP and 0.024 s^{-1} for LIMALP. At these autoxidation rates, reaction with NO or HO_2 becomes competitive with HOM formation. The rates and products of PINALP and LIMALP reactions with NO and HO_2 have also been updated based on parameterizations from Wennberg et al. (2018). For the monoterpene systems, autoxidation occurs in both the first and second (through aldehydes) generation chemistry. Since autoxidation is an efficient source of SOA in monoterpene systems, balancing the role of autoxidation across generations is needed to ensure reasonable SOA production. API ozonolysis in CRACMM2 retains a prompt (first generation) autoxidation channel with a yield of 0.21 for an RO_2 with an autoxidation rate set to the PINALP rate rather than a fixed yield (no competition with biomolecular RO_2 reactions) of 5% for HOM-RO_2 as in CRACMM1. Aldehyde yields are significantly higher in the LIM ozonolysis system compared to API, and all autoxidation from LIM ozonolysis was tied to further aldehyde reaction. Future work should aim to improve the representation of autoxidation across monoterpene ozonolysis and aldehyde systems.

CRACMM1 contains one monoterpene nitrate species (TRPN) which forms primarily from reactions of API and LIM derived peroxy radicals with NO. Further oxidation of TRPN in CRACMM1 results in a 100% yield of HOM, though deposition of TRPN is a competing fate which reduces the effective SOA yield from TRPN in CTMs. In CRACMM2, several additional chemical fates are added for TRPN. Photolysis of TRPN is added, resulting in recycling of NO_X and smaller organic products (species KET and UALD). Oxidation of TRPN no longer produces HOM; instead, we assume that oxidation of TRPN results in a 33% yield of a second-generation monoterpene nitrate species (new species HONIT). The remaining 67% of products result in the release of the nitrate group to NO_2 plus fragmentation products. The 1/3 to 2/3 split to HONIT assumes that 1/3 of TRPN is unsaturated (i.e., contains a double bond) while the remaining two thirds are saturated following a monoterpene nitrate mechanism previously implemented in GEOS-Chem by Fisher et al. (2016) based on a mechanism by Browne et al. (2014). Unsaturated monoterpene nitrates are expected to retain the nitrate group and form a more oxygenated monoterpene nitrate upon reaction while saturated monoterpene nitrates are expected to release the nitrate group to form NO_2 plus other fragmentation products. Limonene oxidation by OH is expected to produce only unsaturated products based on Fisher et al. (2016), so no fragmentation products from limonene derived nitrates are expected. Thus, fragmentation products are based on oxidation products of α-pinene and β-pinene derived nitrates in MCM. The α-pinene nitrate products from MCM indicate a 62% yield of pinonaldehyde (species PINAL) and 38% yield of a ketone (species KET). The β-pinene nitrate products from MCM indicate a 92% yield of a ketone (species KET) and HCHO and 8% yield of an aldehyde (species ALD). Since α-pinene and β-pinene are lumped in species API, we apply a 65/35 split of α-pinene and β-pinene based on the biogenic emissions of these species to calculate the total yields of these additional products.

The new second generation monoterpene species HONIT can be lost through photolysis, reaction with OH, deposition, or hydrolysis. Both TRPN and HONIT are treated as semivolatile species in CRACMM2 with C* of ~1400 $\mu g\ m^{-3}$ and ~0.04 $\mu g\ m^{-3}$ respectively based on their structures (Pankow and Asher, 2008). The resulting SOA from TRPN and HONIT are tracked as two new aerosol species (ATRPN and AHONIT). These monoterpene nitrate aerosol species also undergo hydrolysis with 3-h lifetime to form aerosol HOM (species AHOM) and nitric acid (Pye et al., 2015). The updates to monoterpene nitrates do not have significant effects on HCHO, but the updates to other parts of the monoterpene system offered an opportunity to address these additional areas that were known to be missing from CRACMM1.

~~3~~2.5 Aromatics

The chemistry of aromatics in CRACMM1 is generally based on MCM and the work of Xu et al. (2020) as described in Pye et al. (2023). In CRACMM2 most aromatic species are unchanged from CRACMM1, but we make some updates to how emissions of aromatic compounds are mapped to lumped mechanism species. CRACMM1 includes two lumped xylene-based species defined by a range in OH reactivity: XYE includes ethylbenzene, o- and p-xylene, and other aromatic species with chemistry based on ethylbenzene and XYM includes m-xylene and other aromatic species with chemistry based on m-xylene. In CRACMM1, single ring aromatic species benzene, toluene, and those in the intermediate-volatility range are separately represented from XYE and XYM. In CRACMM2, XYE is renamed to EBZ to represent ethylbenzene explicitly and no longer includes any xylene isomers or other species. XYM is renamed to XYL and is now used to represent all isomers of xylene plus other single ring aromatic species that are not otherwise represented. Isomers of xylene are commonly reported in measurements as a mixture of o-, m-, and/or p-xylene. Lumping all xylenes into XYL prevents artificial distinctions in reactivity that are not actually available from measurements or the emission inventories informed by them. There are no changes in the chemistry of XYE (now EBZ) or XYM (now XYL) between CRACMM1 and CRACMM2 – only changes in how emissions are mapped onto these species.

Styrene is lumped into XYM in CRACMM1 but has been separated as an explicit species STY in CRACMM2. Styrene was added because it is a HAP and because it also has a much higher yield of secondary HCHO than m-xylene which led to underestimates in secondary HCHO estimated by box modeling (Sect. ~~2~~3). Styrene chemistry is based exclusively on MCM and proceeds through one route in CRACMM2: OH addition to the exocyclic double bond (Jenkin et al., 2003; Bloss et al., 2005). Molteni et al. (2018) quantified HOM yields from aromatics but did not include styrene in their tests. Since autoxidation in aromatic systems likely occurs for bicyclic RO_2 (Molteni et al., 2018; Xu et al., 2020) which does not occur in the styrene system, we assume first generation styrene products do not undergo autoxidation. No organic nitrates are predicted. Major products include HCHO, benzaldehyde, and peroxides. The peroxide is predicted to have a C* of 1.5×10^3 μg m^{-3} according to EPISuite vapor pressure so it is mapped to an oxygenated IVOC with O:C=0.2 (species VROCP3OXY2). VROCP3OXY2 undergoes multigenerational oxidation leading to fragmentation products as well as SOA. Previous work (Tajuelo et al., 2019; Yu et al., 2022) suggests styrene produces SOA in small amounts which are not considered here, although VROCP3OXY2 can go on to make SOA in further generation chemistry.

~~2~~3 Box model simulations

Box model simulations were conducted using F0AM (Wolfe et al., 2016a) v4.3 to explore the representation of secondary production of HCHO in CRACMM1 compared to ~~the Master Chemical Mechanism (~~MCM~~)~~ v3.3.1 (Jenkin et al., 1997; Saunders et al., 2003; Jenkin et al., 2003; Bloss et al., 2005; Jenkin et al., 2012; Jenkin et al., 2015). Results from MCM are used as a benchmark to compare with CRACMM1 since it provides a much more detailed representation of chemistry (17224 reactions and 5832 species in MCM compared to 508 reactions and 229 species in CRACMM1). While MCM is much more detailed than the chemical mechanisms typically used in CTMs, it has limitations and uncertainties. Deviations between MCM and CRACMM in the box model simulations serve as a screening process to identify precursors systems for further investigation and updates in CRACMM. ~~The box model simulations serve as a screening level identification of precursor systems that may not produce sufficient secondary HCHO in CRACMM1.~~ Discrepancies between MCM and CRACMM1 indicate differences in mechanism assumptions but not necessarily an error in CRACMM1. Emission sectors and/or precursor~~s~~ systems that showed meaningful differences from MCM were used to prioritize chemical systems for further analysis and development in CRACMM2. F0AM was run as a batch simulation with pressure, relative humidity (RH), and temperature held at 1013 mbar, 10%, and 298 K, respectively. Photolysis rates from CRACMM1 were matched to existing MCM photolysis rates in F0AM. Simulations were run for 8 hours of photochemical processing with NO_X initialized at an atmospherically relevant value of 1 ppb of NO_2 and allowed to evolve freely during the simulation. Effects of OH-initiated oxidation and ozonolysis were tested separately (oxidant concentrations held constant at 10^6 molecules cm^{-3} OH and zero O_3 in Fig. 1 or 30 ppb O_3 and zero OH in Fig. S~~3~~7). Simulations including both OH and O_3 were also conducted (Fig. S~~4~~8).

Initial concentrations of ROC precursors were set based on grouping emissions in two different ways: by emissions sector and by precursor system. Emissions from each anthropogenic emissions sector for individual species available in the EPA SPECIATE database (Simon et al., 2010) as previously compiled by Pye et al. (2023) were mapped to species available in CRACMM and MCM. The concentrations of ROC precursors were initialized based on the emissions of each species with 100 Gg of annual emissions represented by 1 ppb (except for primary HCHO which was excluded). For the emission sector simulations (Fig. 1a), all emitted ROC species from each of 20 emissions sectors (Table S~~1~~2) were initialized at their emission-weighted value. For the precursor system simulations (Fig. 1b), the total emissions across all sectors were divided into 19 distinct precursor groups (Table S~~2~~3), and a simulation was conducted with initial concentrations for only the species belonging to a particular precursor group. Secondary HCHO from biogenic emissions was similarly assessed except that initial precursor concentrations were set with 1000 Gg of annual emissions represented as 1 ppb. Fig. S2 shows the annual emissions for each emission sector and precursor group which were used to set the initial ROC precursor concentrations. At the end of 8 hours of photochemical processing, the ending HCHO concentrations simulated by MCM and CRACMM1 were used to compare the representation of secondary HCHO from CRACMM1 and MCM. We also provide here for comparison the secondary HCHO simulated in CRACMM2 after all chemistry updates (Sect. ~~3~~2) were added. These ending concentrations of HCHO are not intended to represent the expected contribution to ambient HCHO from a particular emissions sector or precursor group; they are only intended to serve as a convenient metric to compare secondary production of HCHO across mechanisms and to identify systems requiring further investigation. The emissions-weighted approach used here for setting the initial ROC precursor concentrations means that the magnitude of the ending concentration of HCHO depends on two major factors: the total ROC emissions from the individual emission sector or precursor system (Fig. S2) and the yield of HCHO from the ROC species included in the simulations. The box model setup employed here is limited in its ability to assess some atmospheric processes, such as transport or interactions between emissions from different sectors and does not capture the range of NO_X and ROC precursor concentrations in the atmosphere. However, it offers an efficient way to conduct idealized tests of HCHO production with different chemical mechanisms.

Results from the F0AM box model simulations with OH oxidation are summarized in Fig. 1. Secondary HCHO from biogenic sources is much higher in MCM compared to CRACMM1. This discrepancy is mostly from isoprene which has much lower HCHO production in CRACMM1, though monoterpenes also contribute. Secondary formation of HCHO from isoprene in CRACMM1 is inherited from RACM2, which has been found to produce less HCHO from isoprene compared to other mechanisms (Wolfe et al., 2016a; Wiser et al., 2023). Production from sesquiterpenes is also underestimated, but this is less influential since sesquiterpene emissions are small relative to isoprene and monoterpenes. Isoprene updates in CRACMM2, specifically an increased HCHO yield from isoprene oxidation, drive substantially higher (~a factor of 6) secondary HCHO from total biogenic emissions compared to CRACMM1 which is more consistent with MCM. MCM has been previously found to underestimate the rate of isomerization of isoprene hydroxy peroxy radicals based on comparisons to experimental results (Novelli et al., 2020) which may also affect isoprene products, including HCHO. ~~HCHO from monoterpenes is also increased in CRACMM2 and is more in line with what is predicted by MCM.~~ In addition to the comparisons to MCM, the production of HCHO from isoprene in CRACMM2 has also been compared to the more detailed representation from ~~The production of HCHO from isoprene with the AMOREv1.2 condensed mechanism has also been compared with the detailed isoprene mechanism from~~ Wennberg et al. (2018) and compares favorably (see SI and Fig. S3~~ for more details~~). This is expected since CRACMM2 uses the AMOREv1.2 condensation of the Wennberg et al. (2018) isoprene mechanism (see Sect. 2.1). HCHO from monoterpenes is also increased in CRACMM2 and is more in line with what is predicted by MCM.

Other than biogenic emissions, fires have the highest secondary HCHO production by sector in these tests because they have large total ROC emissions. Secondary HCHO simulated by MCM for fire sectors is higher than CRACMM1, primarily due to differences in secondary HCHO from alkenes mostly in the form of terminal olefins. HCHO from volatile chemical products (VCPs) was identified as an important source of difference between mechanisms where HCHO from CRACMM1 was low compared to MCM. The largest source of secondary HCHO for VCPs was from limonene. Updates to the limonene system (Sect. ~~3~~2.4) resulted in better agreement between CRACMM2 and MCM-estimated secondary HCHO. Gasoline mobile sources (onroad gas and nonroad gas) and the nonpt sector (a miscellaneous sector for area sources that do not have their own sector) were also low in CRACMM compared to MCM, mostly due to alkenes. HCHO from non-EGU point sources (also sometimes called the ptnonipm sector) was also underestimated in the box model testing. Part of the underestimate for the non-EGU point sector was from the representation of styrene which was lumped with CRACMM1 species XYM (represented with the chemistry of m-xylene) but added as a new explicit species in CRACMM2 (Sect. ~~3~~2.5). Styrene made up 65% of emissions mapped to XYM for the ptnonipm sector, which was a much larger fraction than for other sectors (e.g., 12% for fires, 6% for VCPs, and 1% for gasoline-powered mobile sources). The addition of explicit styrene improved the comparison between MCM and CRACMM2 because the HCHO yield from styrene is much greater than that of m-xylene; however, secondary HCHO from this sector is still low compared to MCM. The other sectors mostly had good agreement between MCM and CRACMM with most of the secondary HCHO production driven by alkenes. One exception is the agricultural sector (ag) where dimethyl sulfide (DMS) contributed to HCHO for MCM but is not currently represented in CRACMM.

When total emissions across all sectors (excluding biogenic emissions) are separated into compound precursor groups, alkenes, such as ethene and propene, make up the largest contribution to secondary HCHO. Ethene is represented explicitly in CRACMM and has very similar HCHO production in MCM and CRACMM. However, the lumped terminal alkene species (OLT) in CRACMM has a lower HCHO yield (0.78) than the effective HCHO yield of propene in MCM (0.98), leading to lower secondary HCHO from alkenes in CRACMM1 compared to MCM which contributes to the low secondary HCHO seen in several source sectors. Alkene chemistry for terminal and internal olefins has not been modified in CRACMM1 or 2 since the original RACM2 implementation but is an area where future development may be needed.

Secondary HCHO from monoterpenes (which are represented in MCM by α-pinene, β-pinene, and limonene) is low in CRACMM1 compared to MCM and has been improved with CRACMM2. Some other groups with lower HCHO in CRACMM include furans, alcohols, and ketones. These are responsible for a smaller fraction of total ROC emissions and were not prioritized for updates in CRACMM2. Aldehydes stand out as a group where secondary HCHO in CRACMM was higher than in MCM which was a result of higher production in CRACMM of methylperoxy radicals, which produce~~s~~ HCHO through reaction with NO. The “other ROC” group is dominated by semi-volatile and intermediate volatility compounds (generally C12 and larger species) which are important for SOA formation in CRACMM but do not exist in MCM and thus do not produce HCHO in MCM.

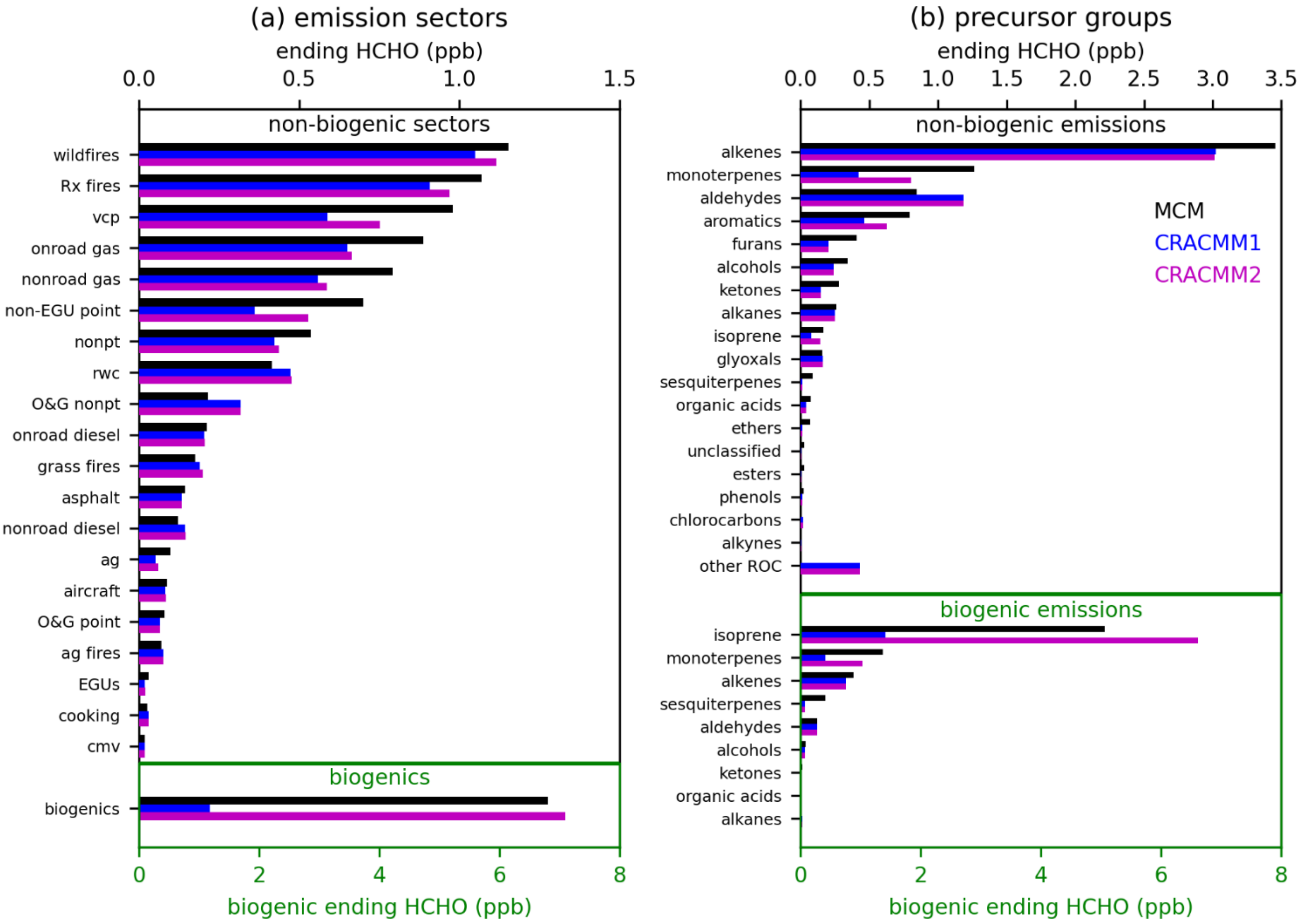


Figure 1. Ending HCHO concentration after 8-h box model simulations for MCM, CRACMM1, and CRACMM2 separated by emissions sector (a) and ROC precursor group (b). Results shown here are for a simulation where OH was held constant at 10^6 molecules cm^{-3} while O$_3$ was held at zero. Each bar represents a separate box model simulation with initial ROC precursor concentrations dependent on the emissions sector or precursor group. Descriptions of the emission sectors and of the species included in each precursor group are given in Tables S~~1~~2-S~~2~~3.

4 CMAQ simulations

CTM simulations were conducted using CMAQv5.4 (U.S. EPA, 2022b) and model inputs from the EQUATES (EPA's Air Quality TimE Series) modeling framework (Foley et al., 2023). The CMAQ model setup is the same as described in Vannucci et al. (2024). The modeling domain covers the CONUS with a horizontal resolution of 12 km. Meteorological inputs are from the Weather Research Forecasting (WRF) model version 4.1.1 (Skamarock et al., 2019) processed through the Meteorology-Chemistry Interface Processor (MCIP) (Otte and Pleim, 2010) for use in CMAQ. Boundary and initial conditions were from a

2019 northern hemispheric simulation from EQUATES with species from the Carbon Bond 6 mechanism mapped to corresponding CRACMM species. Emissions from EQUATES were processed through SMOKE to generate model-ready emission inputs with CRACMM emission speciation. Mapping of emissions species to model species uses the Detailed Emissions Scaling, Isolation, and Diagnostic (DESID) module in CMAQ (Murphy et al., 2021). The emissions mapping step is particularly important in CRACMM for applying appropriate volatility profiles to emissions of primary organic carbon and non-carbon organic matter as operational inventories currently lack that information. Biogenic emissions are computed inline in CMAQ using the BEIS module (Bash et al., 2016). The Surface Tiled Aerosol and Gaseous Exchange (STAGE) dry deposition model is used (Appel et al., 2021; Clifton et al., 2023). Annual simulations for 2019 were conducted using the base CRACMM1 mechanism and the updated CRACMM2 mechanism with one month spin up in December 2018 to reduce the influence of initial conditions. The incremental impacts of chemistry updates (Sect. ~~3~~2) were documented with simulations covering summer when secondary HCHO is highest.

Simulated HCHO is highest in the southeastern US (Fig. 2) in the summer (Fig. S9) due to secondary HCHO from biogenic emissions and photochemical activity. High levels of HCHO are also simulated in California in forested areas surrounding the Central Valley. HCHO in CRACMM2 is higher compared to CRACMM1 in most areas, with the largest increases in summer, though there are some places with seasonal reductions in HCHO of up to -0.1 ppb. Besides the southeastern US and in parts of California where biogenic emissions of isoprene are highest, summer HCHO is also increased across the eastern US broadly. Changes in HCHO in the western US (outside of California) are small (<0.2 ppb). CRACMM2 simulates increased HCHO in the summer across the boreal forests of Canada and forested areas of Mexico within the modeling domain. Predicted spring and fall HCHO also increases in CRACMM2 for the eastern US, California, and Mexico, but to a lesser degree than in the summer (Fig. S9). Overall, 2019 June–August surface HCHO during peak photochemical production (11am–3pm) is increased by 0.6 ppb (32%) over the southeastern US and by 0.2 ppb (13%) over the entire CONUS.

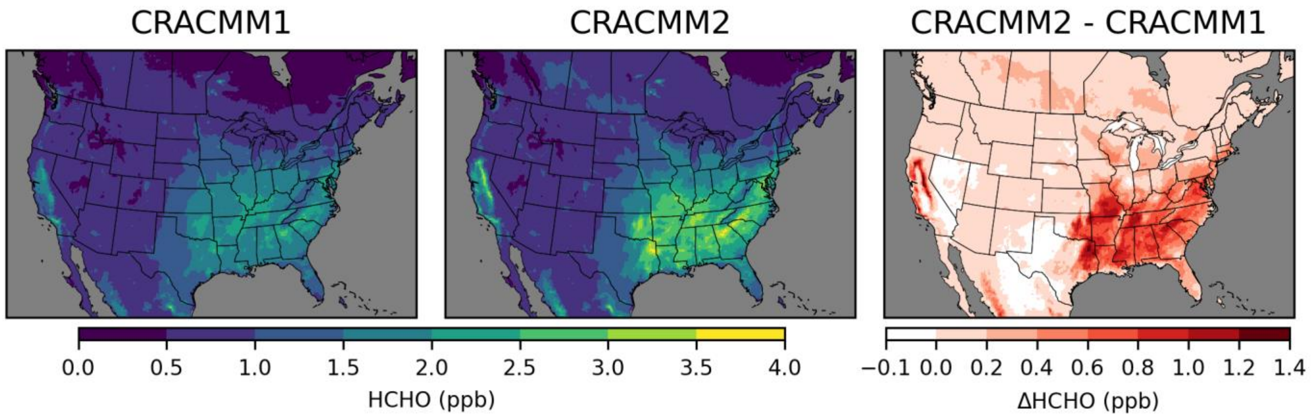
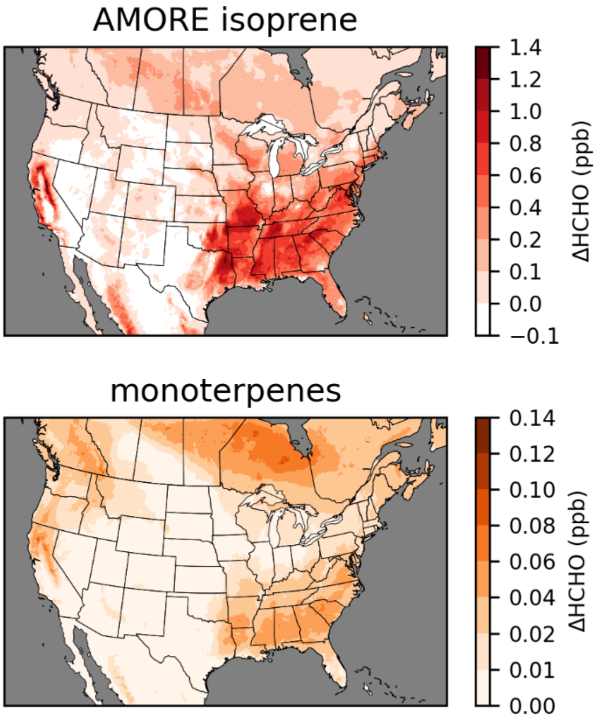


Figure 2. Surface layer 11 am–3 pm local time June–August 2019 average HCHO concentrations simulated with CRACMM1 (left) and CRACMM2 (middle) and the change in CRACMM2 compared to CRACMM1 (right). Analogous results for other seasons are provided in Fig. S9.

Chemistry updates were implemented in stages to track the incremental effects of updates to different chemical systems (Fig. 3; Fig. S10). The update of isoprene chemistry to the AMOREv1.2 isoprene condensation from the RACM2-based isoprene chemistry of CRACMM1 had by far the largest impact on HCHO, and the impacts of the isoprene updates dominate the difference in HCHO between CRACMM1 and CRACMM2. HCHO concentrations most dramatically increase in the southeastern U.S. where biogenic emissions, dominated by isoprene, are highest. Widespread increases in HCHO of ~0.5 ppb occur throughout much of the rest of the eastern US and the boreal forests of Canada as a result of the increased isoprene HCHO yields. Isoprene itself is decreased in CRACMM2 compared to CRACMM1 because of increased reactivity. The summer average reductions in isoprene in the southeastern US are around 0.5 to 1 ppb (20-30%). After isoprene, the monoterpene chemistry updates had the largest impact on HCHO, accounting for ~10% of the total increase in HCHO in CRACMM2 compared to CRACMM1. The impacts on HCHO are spatially representative of biogenic monoterpene emissions with the largest increases in the southeastern US and smaller increases extending to much of the rest of the eastern US. On the west coast, monoterpene impacts have a different spatial pattern than was seen for the isoprene updates as the forests in the Pacific Northwest have larger fraction of total biogenic emissions from monoterpenes compared to the southeastern US. The two CRACMM monoterpene species are decreased due to slightly increased reactivity. In the southeastern US, reductions in CRACMM species API (which represents α-pinene, β-pinene, and other monoterpenes with one double bond) is reduced by around 50-100 ppt (5-15%) while reductions in CRACMM species LIM (which represents limonene and other monoterpenes with more than one double bond) is reduced by around 5-20 ppt (5-15%).


AMORE isoprene

monoterpenes

Figure 3. Incremental impacts on surface layer 11 am–3 pm local time 2019 June–August average HCHO resulting from AMORE isoprene chemistry updates and monoterpene chemistry updates. Incremental impacts of other chemistry updates are provided in Fig. S10.

Effects on HCHO from other CRACMM2 chemistry updates are small in comparison to the isoprene and monoterpene updates (Fig. S10). The inclusion of ECH4 results in some localized increases in HCHO near extremely large methane sources. ECH4 emissions included in CMAQ here do not ~~include~~ use ~~all~~ the anthropogenic methane emissions that are available from the gridded EPA U.S. methane greenhouse gas inventory (Maasakkers et al., 2023) and do not include any natural methane emissions such as from wetlands but only include emissions for traditional NEI sources. The updated heterogeneous chemistry results in small (<40 ppt) increases in HCHO in the southeastern US. There are two likely contributing factors. One is ~~due to~~ decreased HO_X from uptake of HO_2 marginally increasing the lifetime of HCHO. The other is a decrease in the favorability of the RO_2+HO_2 channel with reduced HO_2 and resulting increase in the favorability of the RO_2+NO channel which has higher HCHO yields compared to the RO_2+HO_2 route. The aromatic chemistry updates result in small (<10 ppt) increases in HCHO which are localized to areas with high styrene emissions. More detail on the effects on HCHO from these updates is given in the SI and Fig. S10.

Many of the updates in CRACMM2 have been targeted at secondary HCHO, but the updates also affect $PM_{2.5}$ and O_3. Since $PM_{2.5}$ and O_3 are not the focus of this work, we provide only a brief overview here. Many more details on $PM_{2.5}$ and O_3 impacts are documented in the SI for interested readers. $PM_{2.5}$ decreased across the CONUS in CRACMM2 compared to CRACMM1. For $PM_{2.5}$, the annual mean bias across sites in the Air Quality System (AQS) database went from -0.5 µg m^{-3} in CRACMM1 to -0.8 µg m^{-3} in CRACMM2 driven by reductions in organic aerosol in CRACMM2 from reduced HOM formation from monoterpene nitrates in CRACMM2 (Sect. ~~3~~2.4). These decreases are partially offset by new SOA pathways through heterogeneous uptake of isoprene-derived compounds (Sect. ~~3~~2.1 and ~~3~~2.3). The changes improve the performance of organic carbon which is biased high for the annual average (both in CRACMM1 and in CRACMM2). Low biases in $PM_{2.5}$ mass come from low biases in other $PM_{2.5}$ species including sulfate (Vannucci et al., 2024), nitrate, ammonium, and elemental carbon. Annual average max daily 8-h average (MDA8) O_3 increased in CRACMM2 in the eastern US (particularly in the southeastern US) and in California. MDA8 O_3 decreased slightly (<0.5 ppb) in western Texas and throughout the central US. The changes in O_3 come primarily from changes in HO_X resulting from the implementation of the AMORE isoprene chemistry condensation and from increased NO_X recycling from monoterpene nitrates. Annual mean bias in MDA8 O_3 across AQS sites improved from -1.1 ppb in CRACMM1 to -0.7 ppb in CRACMM2, though there are spatial and seasonal differences in biases that offset each other. On average across all sites, underestimates in MDA8 O_3 in the spring improve in CRACMM2 and a high bias in summer to early fall MDA8 O_3 becomes slightly worse in CRACMM2.

5 Comparisons to observations

CMAQ HCHO results are compared against several different sources of observations to evaluate the impacts of the CRACMM2 updates. Observational data includes satellite based-observations from TROPOMI, aircraft-based observations from the Fire Influence on Regional to Global Environments and Air Quality (FIREX-AQ) campaign, and surface-level hourly observations.

5.1 TROPOMI

TROPOMI onboard the Sentinel-5 Precursor satellite provides once daily coverage at around 13:30 local solar time. We use the TROPOMI HCHO tropospheric vertical column density (VCD) and compare with the HCHO VCD simulated by CMAQ. TROPOMI and CMAQ data are processed for comparison using the cmaqsatproc python tool (https://github.com/barronh/cmaqsatproc). We use a reprocessed TROPOMI HCHO dataset with a resolution of 5.5 km × 3.5 km which uses version 2 of the level 2 processor for all of 2019. TROPOMI data are filtered to include only data with a quality assurance (QA) value > 0.75 (stricter than the QA value > 0.5 recommended minimum). A QA value > 0.5 indicates no error flag, cloud radiance fraction at 340 nm < 0.5, solar zenith angle <= 70°, surface albedo <= 0.2, no snow/ice warning, and air mass factor (AMF) > 0.1 (KNMI, 2023). TROPOMI data are gridded onto the 12 km × 12 km CMAQ model grid and are updated with an AMF based on the CMAQ HCHO vertical profile. For each comparison of a CMAQ simulation to TROPOMI, the AMF derived from that specific CMAQ simulation is used. The CMAQ data are sampled so that CMAQ VCDs are only retained for model grid cells and days when there is valid TROPOMI data. The daily TROPOMI HCHO VCDs are scaled up by 25% when the HCHO VCD exceeds 8×10^{15} molecules cm^{-2} to account for a low bias in TROPOMI HCHO at high HCHO VCD levels (De Smedt et al., 2021) and then averaged seasonally. The uniform scaling of 25% is a simplification. Previous comparisons of TROPOMI HCHO against ground-based Fourier-transform infrared (FTIR) observations of HCHO have found that the negative bias of TROPOMI at higher levels of HCHO increases with increasing FTIR HCHO (Vigouroux et al., 2020; Oomen et al., 2024). So, for areas with the highest HCHO, the correction of 25% scaling may still be too low. ~~The CMAQ data are sampled so that CMAQ VCDs are only retained for model grid cells and days when there is valid TROPOMI data.~~

The summer average HCHO VCD from CMAQ (with CRACMM1 and CRACMM2) broadly reproduces the spatial distribution of TROPOMI, with the highest HCHO occuring in the southeastern US along with another area of high HCHO surrounding the Central Valley of California (Fig. 4, Fig. S11). The updates introduced in CRACMM2 increase column HCHO in the eastern US, particularly in the southeastern US, and in California. These increases are mostly from increased HCHO from isoprene from biogenic emissions with some additional increases from monoterpene HCHO yields which are also mostly from biogenic sources. CMAQ becomes closer to TROPOMI with these increases, though HCHO is still consistently lower than TROPOMI throughout the CONUS. In most areas, however, the HCHO VCD simulated by CMAQ is within the range of TROPOMI uncertainty (Fig. S12). The largest underestimates in HCHO occur in the western US. HCHO is significantly underestimated in the Permian Basin, a major oil and gas producing area in western Texas and New Mexico. HCHO is also underestimated over other oil producing areas in Texas and Oklahoma, specifically over the Ft. Worth and Anadarko Basins which could be due to underestimates in primary HCHO, other ROC precursor emissions, and/or secondary production. CTM simulations with WRF-Chem using the fuel-based inventory of oil and gas (FOG) (Gorchov Negron et al., 2018; Francoeur et al., 2021) showed higher HCHO VCDs over the Permian Basin than our simulations here (Dix et al., 2023). Comparisons of FOG to the 2014 NEI have shown that FOG had lower NO_X emissions and higher non-methane VOC emissions (Francoeur et al., 2021). The emissions inventory used in our simulations is based on the 2017 NEI with some updates (see EQUATES, Foley et al. (2023)), and more recent versions of the NEI may show different results. Emissions of ~~both~~ NO_X and ROC precursors will both affect HCHO production in this area (Dix et al., 2023). A sensitivity simulation in which NO_X and ROC emissions from oil and gas sources were doubled resulted in increases in summer average HCHO VCD at the TROPOMI overpass time of up to 1.4×10^{15} molecules cm^{-2} and increases of surface level 11am–3pm summer average HCHO of up to 0.5 ppb (Fig. S13).

Comparison to the TROPOMI column HCHO indicates some regional biases in CMAQ. TROPOMI column HCHO is consistently higher than CMAQ values in the Mountain West and the southwestern US. A large underestimate is seen in Arizona over the Tonto National Forest to the northeast of Phoenix. Large underestimates in California occur over the Los Angeles metropolitan area and over national forest land east of the Central Valley. Underestimates over the national forest land in Arizona and California could result from underestimated biogenic emissions. The underestimated HCHO in Los Angeles is more likely related to anthropogenic precursors and could result from either underestimated precursor emissions or secondary production. More detailed data and analysis of ~~these~~ individual areas, such as might be possible with a field campaign, are likely needed to explore the specific reasons for the underestimates of HCHO. In the part of the modeling domain covering Canada, CMAQ HCHO is consistently higher than TROPOMI. HCHO is extremely overestimated (by $>10^{16}$ molecules cm^{-2}) by CMAQ in parts of Manitoba and Ontario due to excessive primary HCHO from wildfires which likely resulted from inaccurate representation of the emissions and/or plume trajectories from these fires in the model. Updates to HCHO production in CRACMM2 increase the HCHO VCD in the eastern US by $\sim 1\times10^{15}$ molecules cm^{-2} on average with increases of up to $\sim 4\times10^{15}$ molecules cm^{-2} in the southeastern US, leading to a better comparison with TROPOMI HCHO. However, several additional areas with underestimated HCHO (e.g., the Permian Basin and parts of Arizona and California) still need more exploration in future work. More detailed analysis is needed to understand the roles of precursor emissions, secondary HCHO production, and the diurnal variability of HCHO as compared to observations.

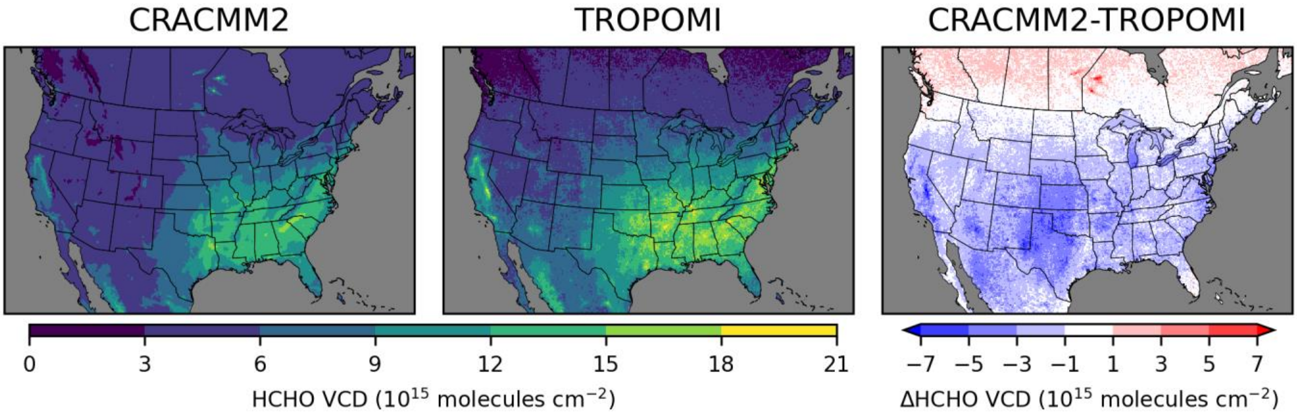


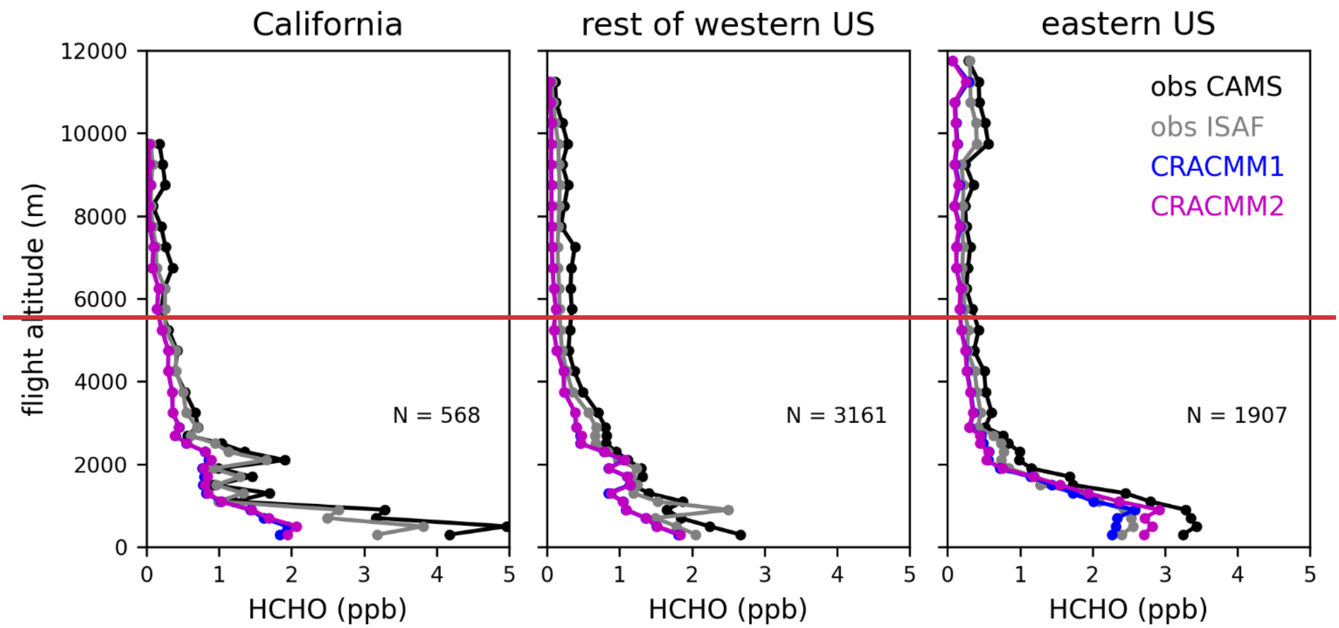
Figure 4. June–August 2019 average tropospheric vertical column densities from CMAQ with CRACMM2 (left) and from TROPOMI (middle) and the difference between CRACMM2 and TROPOMI (right). Similar comparisons for other seasons are provided in Fig. S11.

5.2 FIREX

As part of the FIREX-AQ experiment, in-situ measurements of HCHO (among many other trace gas and aerosol measurements) were taken to assess the chemical evolution of fire plumes by sampling from the NASA DC-8 aircraft during the summer of 2019 (Liao et al., 2021; Warneke et al., 2023). While FIREX-AQ was targeted towards fires, measurements also include conditions outside of wildfire plumes. A significant amount of data was collected outside of fire plumes and is more representative of background conditions than fire conditions. We use HCHO data from two instruments onboard the DC-8 during FIREX-AQ. One is the In Situ Airborne Formaldehyde (ISAF) instrument (Cazorla et al., 2015) which uses laser-induced fluorescence to measure HCHO. The second is the Compact Atmospheric Multispecies Spectrometer (CAMS)

(Richter et al., 2015) which is a mid-IR laser-based spectrometer. During FIREX-AQ, HCHO measured by the ISAF and CAMS instruments were highly correlated with an r^2 of 0.99 and an intercept near zero but with a slope of 1.27 based on an orthogonal regression between the two. Follow up studies indicated that this discrepancy was due to differences in the calibration standards employed (Liao et al., 2021). We include both the ISAF and CAMS observations in our analysis and interpret their difference as an indicator of measurement uncertainty. FIREX-AQ observations at 1 Hz frequency were averaged up to the minute and were paired with the CMAQ model outputs coincident in space and time with the flight track by matching the observation time to the nearest hourly model output time step, the radar altitude to the model vertical layer height, and the aircraft coordinates to the corresponding model horizontal grid cell. Paired observation-model data are then separated into “smoke” or “background” categories based on a smoke indicator flag which is based on CO and black carbon enhancements above background concentrations. Starting from a total of 9084 paired data points available, 7568 (83%) had measurements available for both ISAF and CAMS HCHO. Of these, 1932 (26%) were flagged as smoke with the remaining 5636 (74%) taken as background.

We focus on the background (i.e., not in fire plumes) data since HCHO from fires and within fire plumes was not a focus of the CRACMM2 updates. (See Pye et al. in prep. for an evaluation of CMAQ-CRACMM1AMORE predictions of HCHO during FIREX-AQ.) Although these data are sampled outside of fire plumes, there still may be some influences from fire emissions even in the background observations since the data are collected in fire-affected regions during periods with active fires. Data are further separated geographically to highlight differences in CMAQ performance in California, the rest of the western US, and the eastern US with a longitude of -97 °W defining the east-west boundary. The data in California primarily sample the Central Valley and the Los Angeles area. The data in the rest of the western US sample within the states of Arizona, Idaho, Utah, Washington, and Montana. The data in the eastern US are exclusively in the southeastern US. The vertical profile of HCHO in CMAQ is evaluated with the FIREX-AQ HCHO measurements (Fig. 5). Data are aggregated by altitude in bins of 200 m below 3000 m and 500 m above 3000 m to generate a campaign average HCHO vertical profile in each geographic region. Across all regions that were sampled during FIREX-AQ, the simulated vertical profile of HCHO follows the basic shape of the observations with the highest values in the boundary layer and decreases with altitude. Above ~2 km, the CMAQ vertical profile is biased low across all regions, and the CRACMM2 updates have negligible effects. ~~The modeled near-surface concentrations are very low in California (1-2 ppb below observations depending on the instrument).~~ The low bias in HCHO aloft may be from underestimated precursor abundance aloft and/or from underestimated secondary production from the dominant aloft precursors. The low bias in HCHO aloft may also explain some of the low biases in HCHO VCDs from CMAQ compared to TROPOMI (Sect. 5.1) since TROPOMI has a greater sensitivity at higher altitudes. The modeled near-surface concentrations are very low in California (1-2 ppb below observations depending on the instrument). Near-surface HCHO is also biased low in the rest of the western US, though with a smaller magnitude. The updates in CRACMM2 have only small effects even near the surface in the western US. In the southeastern US, however, CRACMM2 updates lead to an increase in HCHO below 2 km which improves the low bias in CRACMM1. The CRACMM2 southeastern US predictions at lower altitudes are consistent with measurements as they fall between the ISAF and CAMS measurements. The CRACMM2 updates primarily affect secondary HCHO from biogenic emissions, so increases in HCHO in the southeastern US are expected and are consistent with the impacts shown in previous sections.



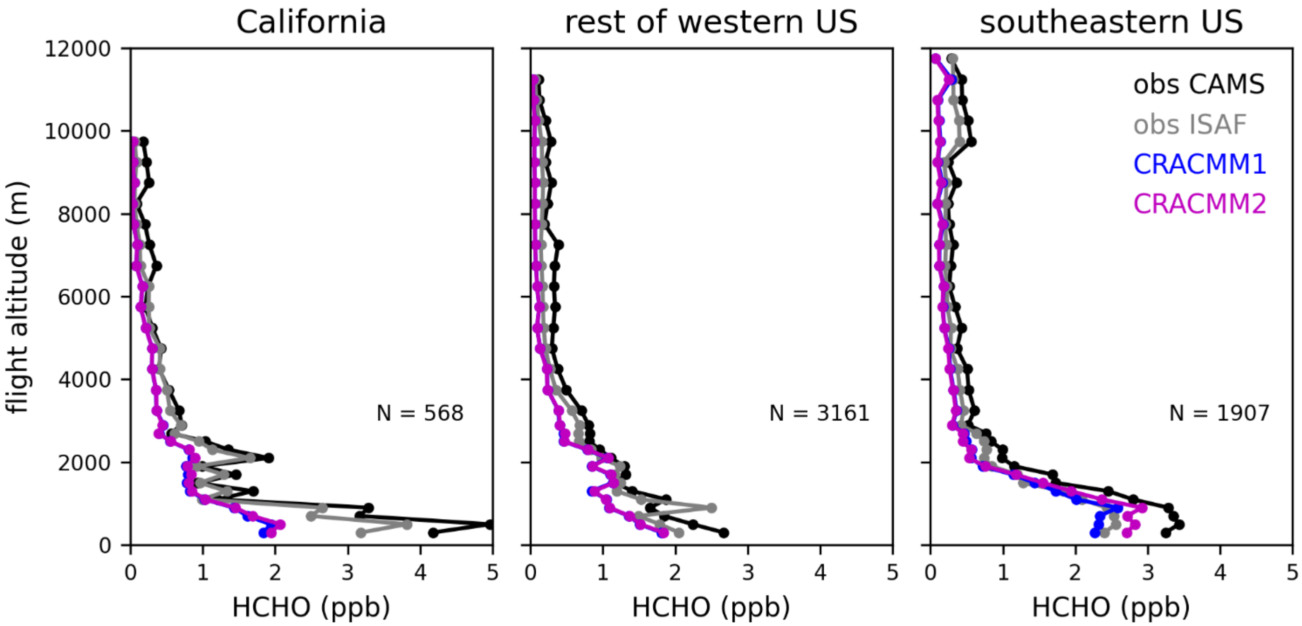


Figure 5. FIREX-AQ campaign average vertical profiles of observed (CAMS and ISAF) and simulated (CRACMM1 and CRACMM2) HCHO. Data flagged as within smoke plumes is excluded here. Profiles are separated into western and eastern US using a longitude of -97 °W. Data over California is further separated from the rest of the western US. Flights east of -97 °W were exclusively in the southeastern US and so are labeled as southeastern US in the rightmost panel title. The vertical profiles show the average HCHO over altitude bins of 200 m below 3000 m and 500 m above 3000 m. The number of observations (N) in each geographical area is also provided.

5.3 Hourly surface observations

HCHO observations from federal, state, local, and tribal air quality monitoring networks are available from the AQS database. Many HCHO observations from AQS are based on a 24-h sample collection (i.e., daily average) with offline characterization (method TO-11A), though some sites collect three 8-h samples over the course of a day on a once per three days schedule during the summer. The lack of hourly data for evaluation of the HCHO diurnal variability in CMAQ is a limitation of the AQS HCHO observations. In addition, previous work indicates offline network measurements of HCHO can be biased high or low (Zhu et al., 2017; Mouat et al., 2024), and we find AQS measurements show a summer, regional maximum in HCHO in the Carolinas (Fig. S17-S18) rather than in the northern Georgia region, in contrast to CMAQ and TROPOMI (Fig. 2 and 4). Here, we focus on surface HCHO observations with hourly resolution from episodic field intensives to better understand drivers of concentrations. In several cases, due to data limitations, we leverage observations from a year other than our 2019 modeling year. As temperature is a strong driver of isoprene emissions and can modulate chemistry, some deviation between the model predictions and observational data is expected (more analysis of HCHO variation with temperature is provided in Fig. S20-S21). Rather than evaluating the performance of the hourly HCHO in CMAQ quantitatively, we use the hourly measurements available in other years as a qualitative indication of how well CRACMM2 in CMAQ represents the typical diurnal variability of HCHO. Data are paired by hour and date across observed and modeled years, and hourly data points with missing observations are dropped before averaging to the diurnal cycle. Comparisons with routine AQS data are available in the supplement, and details on sampling locations, dates, and instrumentation used here are provided in Table S4.

The Salt Lake City, UT, data from winter 2017 covers periods with persistent cold-air pool (PCAP) events which are characterized by extremely shallow mixed layers that prevent vertical mixing. These events are often not well-captured by meteorological models that drive CTMs, so we exclude data collected during three PCAP events (13-20 January, 27 January - 4 February, and 13-18 February). The Salt Lake City observations show a relatively flat diurnal profile with slight peaks in the late morning and in the evening (Fig. 6). The CRACMM diurnal profile is also flat with small peaks in the morning and in the early afternoon. The magnitude of the simulated HCHO diurnal profile is lower by about a factor of 2.5 on average compared to the observations. Although the comparison uses different observation and simulation years, it suggests a missing anthropogenic source in the model emission inventory since biogenic emissions would not be a major factor during the winter sampling period. Previous work suggested primary HCHO emissions are underestimated in the Salt Lake City area based on data collected during the Salt Lake Regional Smoke, Ozone and Aerosol Study (SAMOZA) campaign in summer 2022

(Ninneman et al., 2023; Jaffe et al., 2024). Primary HCHO is expected to contribute relatively more to overall HCHO in the winter as compared to warmer seasons due to the longer lifetime of HCHO in the winter and reduced biogenic precursor emissions. Model simulations have estimated primary HCHO fractions in the winter of 25-50% (Luecken et al., 2012). Secondary production is still important in winter, and photochemistry can be enhanced through increased albedo in snow-covered areas (Edwards et al., 2014). While this data is suggestive of underestimated anthropogenic emissions in the area, the missing driver cannot be identified beyond a combination of primary HCHO and/or ROC precursors.

For the several locations in the northeastern US (Westport, New Brunswick, and Flax Pond) in summer (see Fig. S22-S23 for other seasons), the comparisons of continuously sampling online techniques (in 2023) to simulation predictions (in 2019) are generally consistent and indicate the model captures the correct broad features of HCHO. The simulated HCHO reaches about the same midday peak level as the observations when the CRACMM2 updates are added. While the model does reflect a daytime increase in HCHO at these sites, the simulated diurnal profile shows less diurnal variation than the observations. The observations show a sharp rise from the early morning to a midday peak, followed by a sharp decline over the late afternoon and into the night. The comparisons of observed diurnal variability of HCHO with CRACMM1 and CRACMM2 in CMAQ indicate HCHO in CMAQ tends to be too high at night.

The error in the HCHO diurnal profile during summer in CMAQ is pronounced for two sites in Atlanta, GA, where data has been collected as part of a longer-term HCHO sampling effort (Mouat et al., 2024). One site is co-located with a Photochemical Assessment Monitoring Stations (PAMS) network site, known as the South DeKalb (SDK) monitoring site, located in a suburban part of the Atlanta metro area. The other site is located on the campus of Georgia Tech (GT) which is within the urban core of the city of Atlanta. The two Atlanta sites are located ~15 km away from one another and are in adjacent grid cells of the 12 km CMAQ modeling domain. At both Atlanta sites, the observed diurnal profile begins increasing at 6 am until it reaches peak levels around 11 am to 3 pm before dropping again into the late afternoon and overnight (Fig. 6). The overnight lows at the SDK site are lower than at the GT site, though the diurnal variation (i.e., the difference between the high and low values) at each site is similar. The modeled diurnal profile does not reproduce the observed shape at either site. The model correctly reflects the start of the rise in HCHO at 6 am; however, predicted HCHO in CRACMM1 and CRACMM2 declines in the late morning, remains flat as the afternoon progresses, then has a slight rise at night. Similar discrepancies occur for other seasons (Fig. S23).

Across seasons, CMAQ does not capture the peak HCHO during midday for several possible reasons. Biogenic isoprene emissions could be low in CMAQ. The observed HCHO diurnal profile largely follows the typical daily cycle of isoprene emissions, and secondary HCHO from isoprene is expected to be the dominant contributor to HCHO in the southeastern US. A comparison of the modeled diurnal profile of isoprene in 2019 to observations in 2022 and 2023 (hourly isoprene measurements are not available at the SDK site for 2019) shows that simulated isoprene is within the range of interannual variability (Fig. S24). However, the 2019 simulated isoprene diurnal profile decreases between noon and 5 pm whereas the observed isoprene in 2022 and 2023 continues to increase or remains near its peak during this period. The June–August 2019 total of isoprene emissions over the southeastern US (75-100 °W, 26-42 °N following Müller et al. (2024)) from the CMAQ inline implementations of BEIS (5.7 Tg C) and MEGAN (5.6 Tg C) are in good agreement. However, compared to an inversion optimizing isoprene emissions based on HCHO column totals from Ozone Monitoring Instrument (OMI) by Müller et al. (2024), the June–August total of isoprene emissions from BEIS over the southeastern US are slightly high (+8%) compared to an inversion where OMI HCHO was not bias corrected (5.3 Tg C) but low (-41%) compared to an inversion where OMI HCHO was corrected for a low bias in the OMI HCHO retrievals (9.7 Tg C). The emission totals reported here from the inverse analysis by Müller et al. (2024) are in a different year (2013) than our simulations (2019). While there is interannual variability in the emissions of isoprene, this would not account for such a large difference (-4.0 Tg C). Besides isoprene, another potential contributing factor to the low midday HCHO could be that the loss rate of HCHO is too high so that HCHO is lost faster that it can be produced, contributing to the lack of peak during midday. For instance, if cloud coverage is underestimated in the model, the photolysis losses could be too high. In all seasons except winter (when HCHO is very low at all times), the modeled Atlanta nighttime values are typically higher than the observations, especially after the CRACMM2 updates. The high nocturnal HCHO does not seem to result from a shallow modeled boundary layer. Modeled CO (used here as an indicator for boundary layer depth) decreases at night while observed CO increases, indicating that the modeled boundary layer is too deep rather than too shallow (Fig. S24).

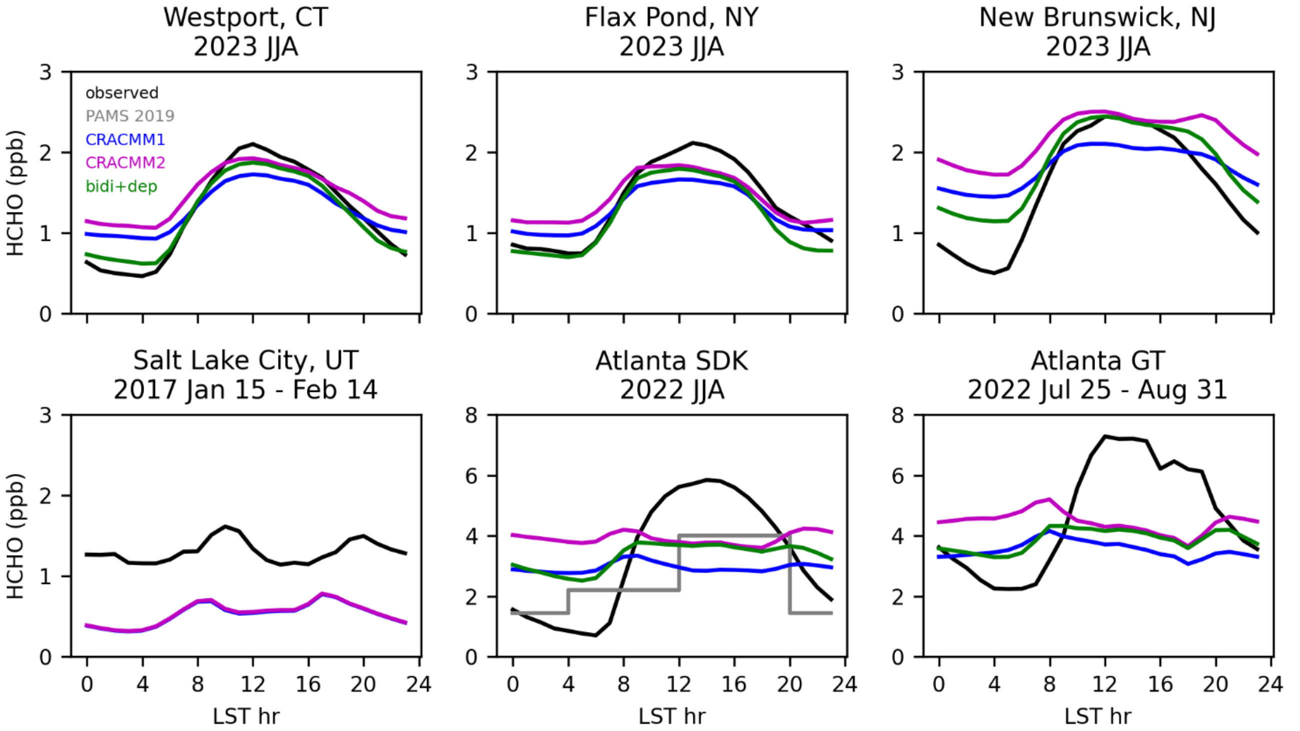


Figure 6. Diurnal profiles of observations in several years at several sites compared to CMAQ simulations in 2019 using CRACMM1, CRACMM2, and CRACMM2 with updated HCHO bidirectional flux and deposition (bidi+dep). Sampling locations and dates are provided above each panel. PAMS 2019 (grey line) shows the average of 8-h HCHO samples collected using method TO-11A during summer 2019 at the SDK monitoring site.

5.4 Deposition updates

HCHO is expected to decline at night, as is seen in the hourly observations, since HCHO production is primarily driven by photochemistry. The consistently high predicted nighttime HCHO levels compared with observations from multiple locations suggest a missing nighttime loss process for HCHO in the model. Bidirectional exchange of HCHO on plant surfaces has been proposed and measured in a laboratory setting (Shutter et al., 2024). Bidirectional exchange of formic acid has also been previously implemented in CMAQ, resulting in improvement of the diurnal variability from a previously flat modeled diurnal profile to one more consistent with surface observations (Gao et al., 2022). We performed a sensitivity simulation for summer 2019 where the STAGE dry deposition model in CMAQ was updated to add a bidirectional flux for HCHO based on the HCHO stomatal compensation point parameterization of Shutter et al. (2024) and a relative humidity (RH) dependence to leaf wetness for dry deposition (Altimir et al., 2006; Burkhardt et al., 2009). The stomatal compensation point is taken as the internal concentration of HCHO in the leaf and represents the ambient HCHO concentration at which there is no net flux via the stomata. When the ambient concentration exceeds the compensation point, there is deposition. When the ambient concentration is below the compensation point, there is emission. The stomatal compensation point is then incorporated into bidirectional flux calculations within the STAGE deposition module (see Clifton et al. (2023) for more details on the implementation of STAGE in CMAQ). The addition of the stomatal bidirectional flux parameterization of Shutter et al. (2024) tends to slightly increase HCHO (typically 0-50 ppt daily average). In addition to the stomatal bidirectional flux implementation, an additional sink was added by accounting for the role of plant surface wetness in deposition based on Altimir et al. (2006). Plant surface wetness was parameterized using RH based on experimental results by Burkhardt et al. (2009). This new surface wetness dependent deposition process decreased HCHO at night when RH is higher.

These updates to deposition lead to better agreement of the modeled and observed diurnal profiles (Fig. 6). The addition of the bidirectional flux of HCHO tends to slightly increase HCHO throughout all hours of the day. The leaf wetness deposition tends to reduce HCHO throughout all hours of the day with smaller decreases during the day and larger decreases at night, consistent with the typical diurnal variability of RH which is higher at night. The increased HCHO from the bidirectional flux mostly offsets the increased deposition losses during the day. At night, the increase in deposition reduces HCHO, leading to better agreement with nighttime observations. For two of the northeastern US sites (Westport and Flax Pond), the HCHO at night becomes very close to the observations after the deposition updates are added (Fig. 6). At the New Brunswick site, HCHO is reduced at night which better matches observations but is still higher than observed. For the two Atlanta sites, the addition of the bidirectional flux of HCHO and the increased deposition leads to better agreement with the observed diurnal profile. The shape of the diurnal profile becomes more like the observations, falling at night and peaking during the day. However, the model still does not quite capture the lows at night, particularly at the SDK site, or the height of the peak during midday. The bidirectional flux and deposition updates slightly reduce surface and column HCHO by up to ~~0.15~~0.14 ppb (June–August 11 am–3 pm average) and 0.3×10^{15} molecules cm^{-2} (June–August average at TROPOMI overpass) (Fig. S25). The June–August nocturnal (8pm–4am) surface HCHO is reduced on average by 1.1 ppb (36%) over the southeastern US and 0.5 ppb (29%) over the entire contiguous US.

6 Implications

The increased HCHO in CRACMM2 has implications for the estimation of cancer risk as HCHO is a leading driver of cancer risk from ambient exposure to HAPs (Strum and Scheffe, 2016). A significant amount of HCHO originates from oxidation of biogenic ROC, primarily isoprene. However, anthropogenic emissions of ROC precursors also contribute to HCHO, and anthropogenic NO$_X$ affects the secondary production of HCHO (Valin et al., 2016; Wolfe et al., 2016b). Here, we estimate a controllable fraction of HCHO and its resulting cancer risk by performing an annual simulation where anthropogenic emissions of NO$_X$ and ROC (excluding anthropogenic fire emissions) within our 12 km CMAQ modeling domain are set to zero. Using the anthropogenic zero out simulation, we estimate the controllable fraction of HCHO simulated in CMAQ with CRACMM2 as the HCHO concentration in the CRACMM2 base simulation minus the HCHO concentration in the zero anthropogenic simulation divided by the HCHO concentration in the base simulation (Fig. 7). The controllable portion of the annual average over land is 24% on average and ranges from 2% to 97% over the CONUS. Seasonally, the highest controllable fraction occurs in winter (average over land of 46%), particularly in the northern portions of the domain (Fig. S26) consistent with increased primary HCHO from residential heating along with longer HCHO lifetimes and reduction in biogenic emissions in winter. The controllable fraction is lowest in the summer (average over land of 17%) when photochemistry is most active, biogenic precursors are highest, and HCHO concentrations are at their highest. The lifetime of HCHO against photolysis is also shortest during this time which limits the impact of primary HCHO. Here we define controllable to include anthropogenic emissions of short-lived precursors NO_x and ROC, but this definition neglects the effects of global background methane oxidation on HCHO. Methane has more than doubled in concentration since the preindustrial era and has a lifetime of ~12 years (Prather et al., 2012), such that reductions in methane could impact HCHO concentrations over large spatial scales in the near term. Future work may consider the role of methane in the fraction of controllable HCHO.

The increased cancer risk from a lifetime of exposure to ambient HCHO is estimated as the annual average concentration times the unit risk estimate (URE). The URE of HCHO of 1.3×10^{-5} $(\mu g\ m^{-3})^{-1}$ indicates 13 more people might be expected to develop cancer per one million people exposed daily for a lifetime to 1 $\mu g\ m^{-3}$ of HCHO. For purposes of estimating risk, we apply an assumed lifetime of exposure of 70 years to our predicted annual average concentrations. The gridded cancer risk estimate is used along with 2019 American Community Survey (ACS) block group level population estimates which are gridded onto the 12 km model domain to calculate the CONUS population-weighted cancer risk and to make an estimate of the total number of CONUS cancer cases estimated from HCHO. Cancer cases are calculated as the sum over CONUS grid cells of the gridded cancer risk times the gridded population (equivalently: the population-weighted cancer risk over CONUS grid cells times the CONUS population). The population-weighted cancer risk (not mortality) for exposure to HCHO in ambient air predicted by CMAQ increases from 17 in a million with CRACMM1 to 19 in a million with CRACMM2, of which 8 in a million (~40%) is estimated to be controllable. The estimate of CONUS cancer cases increases from 5400 with CRACMM1 to 6200 in CRACMM2, of which 2500 are estimated to be controllable (Fig. 7). Although the estimated number of cancer cases has increased with CRACMM2, the number may be underestimated since comparison to observations indicates that CRACMM2 is biased low. CRACMM2 was particularly biased low compared to TROPOMI HCHO in some western US oil and gas producing areas, including the Permian, Ft. Worth, and Anadarko Basins. CRACMM2 was also biased low in parts of California, including in the Los Angeles area (based on comparisons to TROPOMI) and in the Central Valley (based on comparisons to FIREX-AQ aircraft observations) which are two highly populated parts of the state. For reference, the national average risk from exposure to ambient HCHO from the 2019 AirToxScreen assessment implies a lifetime risk of ~4800 cancer cases (using the ACS 2019 population estimate). While the results from AirToxScreen are typically rounded to one significant digit, we retain two significant digits here to better compare results from different simulations. Some differences between AirToxScreen and this work are expected given differences in the CMAQ model version (5.3.2 in AirToxScreen vs. 5.4 here), the chemical mechanism (cb6r3 in AirToxScreen vs. CRACMM1 and CRACMM2 here), the WRF version (3.8 in

AirToxScreen vs. 4.1.1 here), the baseline anthropogenic emissions inventory (2017 NEI in AirToxScreen vs. EQUATES emissions here), and the use of a hybrid CTM and dispersion model approach in AirToxScreen vs. CTM results alone here.

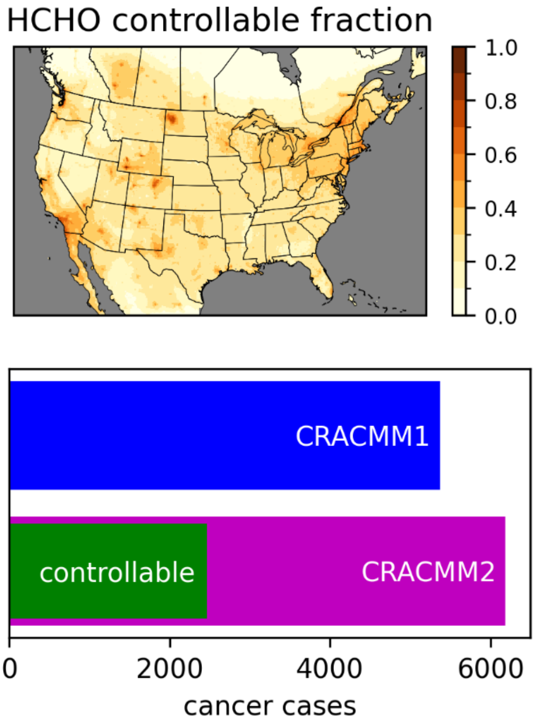


Figure 7. Annual average controllable fraction of HCHO calculated as the difference between the CRACMM2 base simulation and a simulation in which US anthropogenic NO_X and ROC emissions were set to zero divided by the base simulation (top). Estimated lifetime cancer cases from exposure to ambient HCHO estimated using CRACMM1 and CRACMM2 as well as the controllable portion in CRACMM2 calculated as the difference in risk between the base simulation and the zero US anthropogenic NO_X and ROC simulation (bottom).

In this work, an updated representation of deposition and secondary production of HCHO improves our ability to simulate ambient HCHO and its consistency with observations from satellite remote sensing, FIREX-AQ field data, and hourly surface measurements. The investigation here and upcoming data indicate avenues for future work to further improve our understanding of drivers of ambient concentrations. For example, comparison of the diurnal variability of HCHO against hourly surface observations showed that CRACMM2 was typically too high at night, pointing to the potential for a missing nighttime loss pathway for HCHO in CMAQ. The ability of nocturnal leaf wetness to modulate dry deposition and therefore abundance of HCHO suggests concentrations of other soluble species could also be improved by updates to dry deposition or bidirectional exchange. In CRACMM2, peak HCHO levels were near observed levels for surface sites in the northeastern US; however, for daytime in the southeastern US and across the free troposphere, values in CRACMM2 were lower than observed.

This suggests improvements to precursor abundance and/or secondary production is still needed. More in-depth explorations of HCHO and its precursors may be possible with data from the 2023 AGES+ field campaigns (https://csl.noaa.gov/projects/ages) and with the new geostationary satellite-based HCHO data from Tropospheric Emissions: Monitoring of Pollution (TEMPO) mission (https://tempo.si.edu) which will provide daytime variation in HCHO and could enable further improvements in HCHO and its precursors in CRACMM. In addition, we focused exclusively on ambient air in this work, but indoor air concentrations of HCHO can be substantial (Salthammer et al., 2010). A more complete representation of inhaled HCHO health risk will require further improvements to predictions for ambient air as well as characterizing exposure for the indoor environment and extending this analysis to health endpoints beyond cancer.

Code and data availability

- The CMAQ source code is available from GitHub (github.com/USEPA/CMAQ) and Zenodo (https://doi.org/10.5281/zenodo.7218076).
- The CRACMM GitHub site (github.com/USEPA/CRACMM) provides files needed to run CRACMM2 in F0AM, the complete CRACMM2 mechanism, and CRACMM2 species descriptions and properties.
- The F0AM code is available from GitHub (github.com/AirChem/F0AM).
- FIREX-AQ observational data are available from the FIREX-AQ data archive (https://www-air.larc.nasa.gov/cgi-bin/ArcView/firexaq). CAMS HCHO data is revision R3. ISAF HCHO data is revision R0. Navigational data is revision R1.
- HCHO observational data for Atlanta are available from GitHub (github.com/KaiserLab-GeorgiaTech/long-term-HCHO-monitoring_efforts_datasets) and Zenodo (https://doi.org/10.5281/zenodo.10855090).
- HCHO observational data for summer 2023 at Westport, Flax Pond, and New Brunswick sites are available from the following data archive: https://www-air.larc.nasa.gov/cgi-bin/ArcView/listos.2023. Data from all three sites are revision R0.
- HCHO observational data for winter 2017 in Salt Lake City are available from the following data archive: https://csl.noaa.gov/groups/csl7/measurements/2017uwfps/Ground/DataDownload/. Data is revision R0.
- Additional supporting data will be available at data.gov upon publication of the final manuscript.
- Code to recreate figures in the main text will also be available at https://github.com/tnskipper/hcho_cracmm2_paper in addition to the data.gov archive upon publication of the final manuscript.

Author contributions

TNS performed all simulations and analyses and wrote the initial draft. TNS and HOTP designed the research. TNS, HOTP, ELD, RHS, and IRP developed monoterpene chemistry. FCW and VFM developed the AMORE isoprene chemistry in consultation with HOTP and TNS. HOTP, BHH, JOB, and BKP provided additional code and analysis. ARW, LCV, APM, JK, GMW, JMS, TFH, and AF provided ambient data. All coauthors contributed to reviewing and editing the manuscript.

Acknowledgements

This work was supported by the U.S. Environmental Protection Agency Office of Research and Development. This research was supported in part by an appointment to the U.S. Environmental Protection Agency ~~(EPA)~~ Research Participation Program administered by the Oak Ridge Institute for Science and Education (ORISE) through an interagency agreement between the U.S. DOE and the U.S. Environmental Protection Agency. ORISE is managed by ORAU under DOE contract number DE-SC0014664. The views expressed in this paper are those of the authors and do not necessarily represent the views or policies of the U.S. Environmental Protection Agency, the U.S. DOE, or ORISE. We thank TROPOMI and AQS teams for providing data and the CMAQ team for additional discussion. We thank Golam Sarwar and Doris Chen for comments on a draft version of the paper. GMW, JMS, and TFH acknowledge support from the NASA Tropospheric Composition Program and NOAA Climate Program Office's Atmospheric Chemistry, Carbon Cycle and Climate (AC4) program (NA17OAR4310004).

Disclaimer

The authors have no competing interests to declare.

References

Altimir, N., Kolari, P., Tuovinen, J. P., Vesala, T., Bäck, J., Suni, T., Kulmala, M., and Hari, P.: Foliage surface ozone deposition: a role for surface moisture?, Biogeosciences, 3, 209-228, 10.5194/bg-3-209-2006, 2006.

Appel, K. W., Bash, J. O., Fahey, K. M., Foley, K. M., Gilliam, R. C., Hogrefe, C., Hutzell, W. T., Kang, D., Mathur, R., Murphy, B. N., Napelenok, S. L., Nolte, C. G., Pleim, J. E., Pouliot, G. A., Pye, H. O. T., Ran, L., Roselle, S. J., Sarwar, G., Schwede, D. B., Sidi, F. I., Spero, T. L., and Wong, D. C.: The Community Multiscale Air Quality (CMAQ) model versions 5.3 and 5.3.1: system updates and evaluation, Geosci. Model Dev., 14, 2867-2897, 10.5194/gmd-14-2867-2021, 2021.

Bash, J. O., Baker, K. R., and Beaver, M. R.: Evaluation of improved land use and canopy representation in BEIS v3.61 with biogenic VOC measurements in California, Geosci. Model Dev., 9, 2191-2207, 10.5194/gmd-9-2191-2016, 2016.

Bloss, C., Wagner, V., Jenkin, M. E., Volkamer, R., Bloss, W. J., Lee, J. D., Heard, D. E., Wirtz, K., Martin-Reviejo, M., Rea, G., Wenger, J. C., and Pilling, M. J.: Development of a detailed chemical mechanism (MCMv3.1) for the atmospheric oxidation of aromatic hydrocarbons, Atmos. Chem. Phys., 5, 641-664, 10.5194/acp-5-641-2005, 2005.

Browne, E. C., Wooldridge, P. J., Min, K. E., and Cohen, R. C.: On the role of monoterpene chemistry in the remote continental boundary layer, Atmos. Chem. Phys., 14, 1225-1238, 10.5194/acp-14-1225-2014, 2014.

Burkhardt, J., Flechard, C. R., Gresens, F., Mattsson, M., Jongejan, P. A. C., Erisman, J. W., Weidinger, T., Meszaros, R., Nemitz, E., and Sutton, M. A.: Modelling the dynamic chemical interactions of atmospheric ammonia with leaf surface wetness in a managed grassland canopy, Biogeosciences, 6, 67-84, 10.5194/bg-6-67-2009, 2009.

Cazorla, M., Wolfe, G. M., Bailey, S. A., Swanson, A. K., Arkinson, H. L., and Hanisco, T. F.: A new airborne laser-induced fluorescence instrument for in situ detection of formaldehyde throughout the troposphere and lower stratosphere, Atmos. Meas. Tech., 8, 541-552, 10.5194/amt-8-541-2015, 2015.

Clifton, O. E., Schwede, D., Hogrefe, C., Bash, J. O., Bland, S., Cheung, P., Coyle, M., Emberson, L., Flemming, J., Fredj, E., Galmarini, S., Ganzeveld, L., Gazetas, O., Goded, I., Holmes, C. D., Horváth, L., Huijnen, V., Li, Q., Makar, P. A., Mammarella, I., Manca, G., Munger, J. W., Pérez-Camanyo, J. L., Pleim, J., Ran, L., San Jose, R., Silva, S. J., Staebler, R., Sun, S., Tai, A. P. K., Tas, E., Vesala, T., Weidinger, T., Wu, Z., and Zhang, L.: A single-point modeling approach for the intercomparison and evaluation of ozone dry deposition across chemical transport models (Activity 2 of AQMEII4), Atmos. Chem. Phys., 23, 9911-9961, 10.5194/acp-23-9911-2023, 2023.

De Smedt, I., Pinardi, G., Vigouroux, C., Compernolle, S., Bais, A., Benavent, N., Boersma, F., Chan, K. L., Donner, S., Eichmann, K. U., Hedelt, P., Hendrick, F., Irie, H., Kumar, V., Lambert, J. C., Langerock, B., Lerot, C., Liu, C., Loyola, D., Piters, A., Richter, A., Rivera Cárdenas, C., Romahn, F., Ryan, R. G., Sinha, V., Theys, N., Vlietinck, J., Wagner, T., Wang, T., Yu, H., and Van Roozendael, M.: Comparative assessment of TROPOMI and OMI formaldehyde observations and validation against MAX-DOAS network column measurements, Atmos. Chem. Phys., 21, 12561-12593, 10.5194/acp-21-12561-2021, 2021.

Dix, B., Li, M., Roosenbrand, E., Francoeur, C., Brown, S. S., Gilman, J. B., Hanisco, T. F., Keutsch, F., Koss, A., Lerner, B. M., Peischl, J., Roberts, J. M., Ryerson, T. B., St. Clair, J. M., Veres, P. R., Warneke, C., Wild, R. J., Wolfe, G. M., Yuan, B., Veefkind, J. P., Levelt, P. F., McDonald, B. C., and de Gouw, J.: Sources of Formaldehyde in U.S. Oil and Gas Production Regions, ACS Earth and Space Chemistry, 7, 2444-2457, 10.1021/acsearthspacechem.3c00203, 2023.

Duncan, B. N., Yoshida, Y., Olson, J. R., Sillman, S., Martin, R. V., Lamsal, L., Hu, Y., Pickering, K. E., Retscher, C., Allen, D. J., and Crawford, J. H.: Application of OMI observations to a space-based indicator of NOx and VOC controls on surface ozone formation, Atmospheric Environment, 44, 2213-2223, https://doi.org/10.1016/j.atmosenv.2010.03.010, 2010.

Edwards, P. M., Brown, S. S., Roberts, J. M., Ahmadov, R., Banta, R. M., deGouw, J. A., Dubé, W. P., Field, R. A., Flynn, J. H., Gilman, J. B., Graus, M., Helmig, D., Koss, A., Langford, A. O., Lefer, B. L., Lerner, B. M., Li, R., Li, S.-M., McKeen, S. A., Murphy, S. M., Parrish, D. D., Senff, C. J., Soltis, J., Stutz, J., Sweeney, C., Thompson, C. R., Trainer, M. K., Tsai, C., Veres, P. R., Washenfelder, R. A., Warneke, C., Wild, R. J., Young, C. J., Yuan, B., and Zamora, R.: High winter ozone pollution from carbonyl photolysis in an oil and gas basin, Nature, 514, 351-354, 10.1038/nature13767, 2014.

Emmons, L. K., Schwantes, R. H., Orlando, J. J., Tyndall, G., Kinnison, D., Lamarque, J.-F., Marsh, D., Mills, M. J., Tilmes, S., Bardeen, C., Buchholz, R. R., Conley, A., Gettelman, A., Garcia, R., Simpson, I., Blake, D. R., Meinardi, S., and Pétron, G.: 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, 2020.

Fisher, J. A., Jacob, D. J., Travis, K. R., Kim, P. S., Marais, E. A., Chan Miller, C., Yu, K., Zhu, L., Yantosca, R. M., Sulprizio, M. P., Mao, J., Wennberg, P. O., Crounse, J. D., Teng, A. P., Nguyen, T. B., St. Clair, J. M., Cohen, R. C., Romer, P., Nault, B. A., Wooldridge, P. J., Jimenez, J. L., Campuzano-Jost, P., Day, D. A., Hu, W., Shepson, P. B., Xiong, F., Blake, D. R., Goldstein, A. H., Misztal, P. K., Hanisco, T. F., Wolfe, G. M., Ryerson, T. B., Wisthaler, A., and Mikoviny, T.: Organic nitrate chemistry and its implications for nitrogen budgets in an isoprene- and monoterpene-rich atmosphere: constraints from aircraft (SEAC4RS) and ground-based (SOAS) observations in the Southeast US, Atmos. Chem. Phys., 16, 5969-5991, 10.5194/acp-16-5969-2016, 2016.

Foley, K. M., Pouliot, G. A., Eyth, A., Aldridge, M. F., Allen, C., Appel, K. W., Bash, J. O., Beardsley, M., Beidler, J., Choi, D., Farkas, C., Gilliam, R. C., Godfrey, J., Henderson, B. H., Hogrefe, C., Koplitz, S. N., Mason, R., Mathur, R., Misenis, C., Possiel, N., Pye, H. O. T., Reynolds, L., Roark, M., Roberts, S., Schwede, D. B., Seltzer, K. M., Sonntag, D., Talgo, K., Toro, C., Vukovich, J., Xing, J., and Adams, E.: 2002–2017 anthropogenic emissions data for air quality modeling over the United States, Data in Brief, 47, 109022, https://doi.org/10.1016/j.dib.2023.109022, 2023.

Fortems-Cheiney, A., Chevallier, F., Pison, I., Bousquet, P., Saunois, M., Szopa, S., Cressot, C., Kurosu, T. P., Chance, K., and Fried, A.: The formaldehyde budget as seen by a global-scale multi-constraint and multi-species inversion system, Atmos. Chem. Phys., 12, 6699-6721, 10.5194/acp-12-6699-2012, 2012.

Francoeur, C. B., McDonald, B. C., Gilman, J. B., Zarzana, K. J., Dix, B., Brown, S. S., de Gouw, J. A., Frost, G. J., Li, M., McKeen, S. A., Peischl, J., Pollack, I. B., Ryerson, T. B., Thompson, C., Warneke, C., and Trainer, M.: Quantifying Methane and Ozone Precursor Emissions from Oil and Gas Production Regions across the Contiguous US, Environmental Science & Technology, 55, 9129-9139, 10.1021/acs.est.0c07352, 2021.

Gao, Z., Vasilakos, P., Nah, T., Takeuchi, M., Chen, H., Tanner, D. J., Ng, N. L., Kaiser, J., Huey, L. G., Weber, R. J., and Russell, A. G.: Emissions, chemistry or bidirectional surface transfer? Gas phase formic acid dynamics in the atmosphere, Atmospheric Environment, 274, 118995, https://doi.org/10.1016/j.atmosenv.2022.118995, 2022.

Goliff, W. S., Stockwell, W. R., and Lawson, C. V.: The regional atmospheric chemistry mechanism, version 2, Atmospheric Environment, 68, 174-185, https://doi.org/10.1016/j.atmosenv.2012.11.038, 2013.

Gorchov Negron, A. M., McDonald, B. C., McKeen, S. A., Peischl, J., Ahmadov, R., de Gouw, J. A., Frost, G. J., Hastings, M. G., Pollack, I. B., Ryerson, T. B., Thompson, C., Warneke, C., and Trainer, M.: Development of a Fuel-Based Oil and Gas Inventory of Nitrogen Oxides Emissions, Environmental Science & Technology, 52, 10175-10185, 10.1021/acs.est.8b02245, 2018.

Guenther, A. B., Jiang, X., Heald, C. L., Sakulyanontvittaya, T., Duhl, T., Emmons, L. K., and Wang, X.: 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., 5, 1471-1492, 10.5194/gmd-5-1471-2012, 2012.

Hatakeyama, S., Izumi, K., Fukuyama, T., Akimoto, H., and Washida, N.: Reactions of OH with α-pinene and β-pinene in air: Estimate of global CO production from the atmospheric oxidation of terpenes, Journal of Geophysical Research: Atmospheres, 96, 947-958, https://doi.org/10.1029/90JD02341, 1991.

Ivatt, P. D., Evans, M. J., and Lewis, A. C.: Suppression of surface ozone by an aerosol-inhibited photochemical ozone regime, Nature Geoscience, 15, 536-540, 10.1038/s41561-022-00972-9, 2022.

Iyer, S., Rissanen, M. P., Valiev, R., Barua, S., Krechmer, J. E., Thornton, J., Ehn, M., and Kurtén, T.: Molecular mechanism for rapid autoxidation in α-pinene ozonolysis, Nature Communications, 12, 878, 10.1038/s41467-021-21172-w, 2021.

Jacob, D. J.: Heterogeneous chemistry and tropospheric ozone, Atmospheric Environment, 34, 2131-2159, https://doi.org/10.1016/S1352-2310(99)00462-8, 2000.

Jaffe, D. A., Ninneman, M., Nguyen, L., Lee, H., Hu, L., Ketcherside, D., Jin, L., Cope, E., Lyman, S., Jones, C., O'Neil, T., and Mansfield, M. L.: Key results from the salt lake regional smoke, ozone, and aerosol study (SAMOZA), Journal of the Air & Waste Management Association, 74, 163-180, 10.1080/10962247.2024.2301956, 2024.

Jenkin, M. E., Saunders, S. M., and Pilling, M. J.: The tropospheric degradation of volatile organic compounds: a protocol for mechanism development, Atmospheric Environment, 31, 81-104, https://doi.org/10.1016/S1352-2310(96)00105-7, 1997.

Jenkin, M. E., Young, J. C., and Rickard, A. R.: The MCM v3.3.1 degradation scheme for isoprene, Atmos. Chem. Phys., 15, 11433-11459, 10.5194/acp-15-11433-2015, 2015.

Jenkin, M. E., Saunders, S. M., Wagner, V., and Pilling, M. J.: Protocol for the development of the Master Chemical Mechanism, MCM v3 (Part B): tropospheric degradation of aromatic volatile organic compounds, Atmos. Chem. Phys., 3, 181-193, 10.5194/acp-3-181-2003, 2003.

Jenkin, M. E., Wyche, K. P., Evans, C. J., Carr, T., Monks, P. S., Alfarra, M. R., Barley, M. H., McFiggans, G. B., Young, J. C., and Rickard, A. R.: Development and chamber evaluation of the MCM v3.2 degradation scheme for β-caryophyllene, Atmos. Chem. Phys., 12, 5275-5308, 10.5194/acp-12-5275-2012, 2012.

Kaiser, J., Jacob, D. J., Zhu, L., Travis, K. R., Fisher, J. A., González Abad, G., Zhang, L., Zhang, X., Fried, A., Crounse, J. D., St. Clair, J. M., and Wisthaler, A.: High-resolution inversion of OMI formaldehyde columns to quantify isoprene emission on ecosystem-relevant scales: application to the southeast US, Atmos. Chem. Phys., 18, 5483-5497, 10.5194/acp-18-5483-2018, 2018.

KNMI: S5P MPC Product Readme Formaldehyde V02.05.00, 2023.

Lee, A., Goldstein, A. H., Kroll, J. H., Ng, N. L., Varutbangkul, V., Flagan, R. C., and Seinfeld, J. H.: Gas-phase products and secondary aerosol yields from the photooxidation of 16 different terpenes, Journal of Geophysical Research: Atmospheres, 111, https://doi.org/10.1029/2006JD007050, 2006.

Liao, J., Wolfe, G. M., Hannun, R. A., St. Clair, J. M., Hanisco, T. F., Gilman, J. B., Lamplugh, A., Selimovic, V., Diskin, G. S., Nowak, J. B., Halliday, H. S., DiGangi, J. P., Hall, S. R., Ullmann, K., Holmes, C. D., Fite, C. H., Agastra, A., Ryerson, T. B., Peischl, J., Bourgeois, I., Warneke, C., Coggon, M. M., Gkatzelis, G. I., Sekimoto, K., Fried, A., Richter, D., Weibring, P., Apel, E. C., Hornbrook, R. S., Brown, S. S., Womack, C. C., Robinson, M. A., Washenfelder, R. A., Veres, P. R., and Neuman, J. A.: Formaldehyde evolution in US wildfire plumes during the Fire Influence on Regional to Global Environments and Air Quality experiment (FIREX-AQ), Atmos. Chem. Phys., 21, 18319-18331, 10.5194/acp-21-18319-2021, 2021.

Luecken, D. J., Hutzell, W. T., Strum, M. L., and Pouliot, G. A.: Regional sources of atmospheric formaldehyde and acetaldehyde, and implications for atmospheric modeling, Atmospheric Environment, 47, 477-490, https://doi.org/10.1016/j.atmosenv.2011.10.005, 2012.

Luecken, D. J., Napelenok, S. L., Strum, M., Scheffe, R., and Phillips, S.: Sensitivity of Ambient Atmospheric Formaldehyde and Ozone to Precursor Species and Source Types Across the United States, Environmental Science & Technology, 52, 4668-4675, 10.1021/acs.est.7b05509, 2018.

Maasakkers, J. D., McDuffie, E. E., Sulprizio, M. P., Chen, C., Schultz, M., Brunelle, L., Thrush, R., Steller, J., Sherry, C., Jacob, D. J., Jeong, S., Irving, B., and Weitz, M.: A Gridded Inventory of Annual 2012–2018 U.S. Anthropogenic Methane Emissions, Environmental Science & Technology, 57, 16276-16288, 10.1021/acs.est.3c05138, 2023.

Mao, J., Fan, S., Jacob, D. J., and Travis, K. R.: Radical loss in the atmosphere from Cu-Fe redox coupling in aerosols, Atmos. Chem. Phys., 13, 509-519, 10.5194/acp-13-509-2013, 2013.

Martin, R. V., Fiore, A. M., and Van Donkelaar, A.: Space-based diagnosis of surface ozone sensitivity to anthropogenic emissions, Geophysical Research Letters, 31, https://doi.org/10.1029/2004GL019416, 2004.

Molteni, U., Bianchi, F., Klein, F., El Haddad, I., Frege, C., Rossi, M. J., Dommen, J., and Baltensperger, U.: Formation of highly oxygenated organic molecules from aromatic compounds, Atmos. Chem. Phys., 18, 1909-1921, 10.5194/acp-18-1909-2018, 2018.

Mouat, A. P., Siegel, Z. A., and Kaiser, J.: Evaluation of Aeris mid-infrared absorption (MIRA), Picarro CRDS (cavity ring-down spectroscopy) G2307, and dinitrophenylhydrazine (DNPH)-based sampling for long-term formaldehyde monitoring efforts, Atmos. Meas. Tech., 17, 1979-1994, 10.5194/amt-17-1979-2024, 2024.

Müller, J. F., Stavrakou, T., Oomen, G. M., Opacka, B., De Smedt, I., Guenther, A., Vigouroux, C., Langerock, B., Aquino, C. A. B., Grutter, M., Hannigan, J., Hase, F., Kivi, R., Lutsch, E., Mahieu, E., Makarova, M., Metzger, J. M., Morino, I., Murata, I., Nagahama, T., Notholt, J., Ortega, I., Palm, M., Röhling, A., Stremme, W., Strong, K., Sussmann, R., Té, Y., and Fried, A.: Bias correction of OMI HCHO columns based on FTIR and aircraft measurements and impact on top-down emission estimates, Atmos. Chem. Phys., 24, 2207-2237, 10.5194/acp-24-2207-2024, 2024.

Murphy, B. N., Nolte, C. G., Sidi, F., Bash, J. O., Appel, K. W., Jang, C., Kang, D., Kelly, J., Mathur, R., Napelenok, S., Pouliot, G., and Pye, H. O. T.: The Detailed Emissions Scaling, Isolation, and Diagnostic (DESID) module in the Community Multiscale Air Quality (CMAQ) modeling system version 5.3.2, Geosci. Model Dev., 14, 3407-3420, 10.5194/gmd-14-3407-2021, 2021.

Ninneman, M., Lyman, S., Hu, L., Cope, E., Ketcherside, D., and Jaffe, D.: Investigation of Ozone Formation Chemistry during the Salt Lake Regional Smoke, Ozone, and Aerosol Study (SAMOZA), ACS Earth and Space Chemistry, 7, 2521-2534, 10.1021/acsearthspacechem.3c00235, 2023.

Novelli, A., Vereecken, L., Bohn, B., Dorn, H. P., Gkatzelis, G. I., Hofzumahaus, A., Holland, F., Reimer, D., Rohrer, F., Rosanka, S., Taraborrelli, D., Tillmann, R., Wegener, R., Yu, Z., Kiendler-Scharr, A., Wahner, A., and Fuchs, H.: Importance of isomerization reactions for OH radical regeneration from the photo-oxidation of isoprene investigated in the atmospheric simulation chamber SAPHIR, Atmos. Chem. Phys., 20, 3333-3355, 10.5194/acp-20-3333-2020, 2020.

Nozière, B., Barnes, I., and Becker, K.-H.: Product study and mechanisms of the reactions of α-pinene and of pinonaldehyde with OH radicals, Journal of Geophysical Research: Atmospheres, 104, 23645-23656, https://doi.org/10.1029/1999JD900778, 1999.

Oomen, G. M., Müller, J. F., Stavrakou, T., De Smedt, I., Blumenstock, T., Kivi, R., Makarova, M., Palm, M., Röhling, A., Té, Y., Vigouroux, C., Friedrich, M. M., Frieß, U., Hendrick, F., Merlaud, A., Piters, A., Richter, A., Van Roozendael, M., and Wagner, T.: Weekly derived top-down volatile-organic-compound fluxes over Europe from TROPOMI HCHO data from 2018 to 2021, Atmos. Chem. Phys., 24, 449-474, 10.5194/acp-24-449-2024, 2024.

Orlando, J. J., Nozière, B., Tyndall, G. S., Orzechowska, G. E., Paulson, S. E., and Rudich, Y.: Product studies of the OH- and ozone-initiated oxidation of some monoterpenes, Journal of Geophysical Research: Atmospheres, 105, 11561-11572, https://doi.org/10.1029/2000JD900005, 2000.

Otte, T. L. and Pleim, J. E.: The Meteorology-Chemistry Interface Processor (MCIP) for the CMAQ modeling system: updates through MCIPv3.4.1, Geosci. Model Dev., 3, 243-256, 10.5194/gmd-3-243-2010, 2010.

Pankow, J. F. and Asher, W. E.: SIMPOL.1: a simple group contribution method for predicting vapor pressures and enthalpies of vaporization of multifunctional organic compounds, Atmos. Chem. Phys., 8, 2773-2796, 10.5194/acp-8-2773-2008, 2008.

Piletic, I. R. and Kleindienst, T. E.: Rates and Yields of Unimolecular Reactions Producing Highly Oxidized Peroxy Radicals in the OH-Induced Autoxidation of α-Pinene, β-Pinene, and Limonene, The Journal of Physical Chemistry A, 126, 88-100, 10.1021/acs.jpca.1c07961, 2022.

Place, B. K., Hutzell, W. T., Appel, K. W., Farrell, S., Valin, L., Murphy, B. N., Seltzer, K. M., Sarwar, G., Allen, C., Piletic, I. R., D'Ambro, E. L., Saunders, E., Simon, H., Torres-Vasquez, A., Pleim, J., Schwantes, R. H., Coggon, M. M., Xu, L., Stockwell, W. R., and Pye, H. O. T.: Sensitivity of northeastern US surface ozone predictions to the representation of atmospheric chemistry in the Community Regional Atmospheric Chemistry Multiphase Mechanism (CRACMMv1.0), Atmos. Chem. Phys., 23, 9173-9190, 10.5194/acp-23-9173-2023, 2023.

Prather, M. J., Holmes, C. D., and Hsu, J.: Reactive greenhouse gas scenarios: Systematic exploration of uncertainties and the role of atmospheric chemistry, Geophysical Research Letters, 39, https://doi.org/10.1029/2012GL051440, 2012.

Pye, H. O. T., Luecken, D. J., Xu, L., Boyd, C. M., Ng, N. L., Baker, K. R., Ayres, B. R., Bash, J. O., Baumann, K., Carter, W. P. L., Edgerton, E., Fry, J. L., Hutzell, W. T., Schwede, D. B., and Shepson, P. B.: Modeling the Current and Future Roles of Particulate Organic Nitrates in the Southeastern United States, Environmental Science & Technology, 49, 14195-14203, 10.1021/acs.est.5b03738, 2015.

Pye, H. O. T., Pinder, R. W., Piletic, I. R., Xie, Y., Capps, S. L., Lin, Y.-H., Surratt, J. D., Zhang, Z., Gold, A., Luecken, D. J., Hutzell, W. T., Jaoui, M., Offenberg, J. H., Kleindienst, T. E., Lewandowski, M., and Edney, E. O.: Epoxide Pathways Improve Model Predictions of Isoprene Markers and Reveal Key Role of Acidity in Aerosol Formation, Environmental Science & Technology, 47, 11056-11064, 10.1021/es402106h, 2013.

Pye, H. O. T., Murphy, B. N., Xu, L., Ng, N. L., Carlton, A. G., Guo, H., Weber, R., Vasilakos, P., Appel, K. W., Budisulistiorini, S. H., Surratt, J. D., Nenes, A., Hu, W., Jimenez, J. L., Isaacman-VanWertz, G., Misztal, P. K., and Goldstein, A. H.: On the implications of aerosol liquid water and phase separation for organic aerosol mass, Atmos. Chem. Phys., 17, 343-369, 10.5194/acp-17-343-2017, 2017.

Pye, H. O. T., Place, B. K., Murphy, B. N., Seltzer, K. M., D'Ambro, E. L., Allen, C., Piletic, I. R., Farrell, S., Schwantes, R. H., Coggon, M. M., Saunders, E., Xu, L., Sarwar, G., Hutzell, W. T., Foley, K. M., Pouliot, G., Bash, J., and Stockwell, W. R.: Linking gas, particulate, and toxic endpoints to air emissions in the Community Regional Atmospheric Chemistry Multiphase Mechanism (CRACMM), Atmos. Chem. Phys., 23, 5043-5099, 10.5194/acp-23-5043-2023, 2023.

Richter, D., Weibring, P., Walega, J. G., Fried, A., Spuler, S. M., and Taubman, M. S.: Compact highly sensitive multi-species airborne mid-IR spectrometer, Applied Physics B, 119, 119-131, 10.1007/s00340-015-6038-8, 2015.

Salthammer, T., Mentese, S., and Marutzky, R.: Formaldehyde in the Indoor Environment, Chemical Reviews, 110, 2536-2572, 10.1021/cr800399g, 2010.

Sarwar, G., Godowitch, J., Henderson, B. H., Fahey, K., Pouliot, G., Hutzell, W. T., Mathur, R., Kang, D., Goliff, W. S., and Stockwell, W. R.: A comparison of atmospheric composition using the Carbon Bond and Regional Atmospheric Chemistry Mechanisms, Atmos. Chem. Phys., 13, 9695-9712, 10.5194/acp-13-9695-2013, 2013.

Saunders, S. M., Jenkin, M. E., Derwent, R. G., and Pilling, M. J.: Protocol for the development of the Master Chemical Mechanism, MCM v3 (Part A): tropospheric degradation of non-aromatic volatile organic compounds, Atmos. Chem. Phys., 3, 161-180, 10.5194/acp-3-161-2003, 2003.

Scheffe, R. D., Strum, M., Phillips, S. B., Thurman, J., Eyth, A., Fudge, S., Morris, M., Palma, T., and Cook, R.: Hybrid Modeling Approach to Estimate Exposures of Hazardous Air Pollutants (HAPs) for the National Air Toxics Assessment (NATA), Environmental Science & Technology, 50, 12356-12364, 10.1021/acs.est.6b04752, 2016.

Schwantes, R. H., Emmons, L. K., Orlando, J. J., Barth, M. C., Tyndall, G. S., Hall, S. R., Ullmann, K., St. Clair, J. M., Blake, D. R., Wisthaler, A., and Bui, T. P. V.: Comprehensive isoprene and terpene gas-phase chemistry improves simulated surface ozone in the southeastern US, Atmos. Chem. Phys., 20, 3739-3776, 10.5194/acp-20-3739-2020, 2020.

Seltzer, K. M., Rao, V., Pye, H. O. T., Murphy, B. N., Place, B. K., Khare, P., Gentner, D. R., Allen, C., Cooley, D., Mason, R., and Houyoux, M.: Anthropogenic secondary organic aerosol and ozone production from asphalt-related emissions, Environmental Science: Atmospheres, 3, 1221-1230, 10.1039/D3EA00066D, 2023.

Shutter, J. D., Cox, J. L., and Keutsch, F. N.: Leaf-Level Bidirectional Exchange of Formaldehyde on Deciduous and Evergreen Tree Saplings, ACS Earth and Space Chemistry, 10.1021/acsearthspacechem.3c00325, 2024.

Simon, H., Beck, L., Bhave, P. V., Divita, F., Hsu, Y., Luecken, D., Mobley, J. D., Pouliot, G. A., Reff, A., Sarwar, G., and Strum, M.: The development and uses of EPA’s SPECIATE database, Atmospheric Pollution Research, 1, 196-206, https://doi.org/10.5094/APR.2010.026, 2010.

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Liu, Z., Berner, J., Wang, W., Powers, J. G., Duda, M. G., Barker, D., and Huang, X.-y.: A Description of the Advanced Research WRF Model Version 4.1 (No. NCAR/TN-556+STR), 10.5065/1dfh-6p97, 2019.

Strum, M. and Scheffe, R.: National review of ambient air toxics observations, Journal of the Air & Waste Management Association, 66, 120-133, 10.1080/10962247.2015.1076538, 2016.

Tajuelo, M., Rodríguez, D., Baeza-Romero, M. T., Díaz-de-Mera, Y., Aranda, A., and Rodríguez, A.: Secondary organic aerosol formation from styrene photolysis and photooxidation with hydroxyl radicals, Chemosphere, 231, 276-286, https://doi.org/10.1016/j.chemosphere.2019.05.136, 2019.

Tao, M., Fiore, A. M., Jin, X., Schiferl, L. D., Commane, R., Judd, L. M., Janz, S., Sullivan, J. T., Miller, P. J., Karambelas, A., Davis, S., Tzortziou, M., Valin, L., Whitehill, A., Civerolo, K., and Tian, Y.: Investigating Changes in Ozone Formation Chemistry during Summertime Pollution Events over the Northeastern United States, Environmental Science & Technology, 56, 15312-15327, 10.1021/acs.est.2c02972, 2022.

U.S. EPA: Technical Support Document EPA's Air Toxics Screening Assessment 2018 AirToxScreen TSD, 2022a.

U.S. EPA: CMAQ (Version 5.4) [Software]. Available from https://doi.org/10.5281/zenodo.7218076, 2022b.

Valin, L. C., Fiore, A. M., Chance, K., and González Abad, G.: The role of OH production in interpreting the variability of CH2O columns in the southeast U.S, Journal of Geophysical Research: Atmospheres, 121, 478-493, https://doi.org/10.1002/2015JD024012, 2016.

Vannucci, P. F., Foley, K., Murphy, B. N., Hogrefe, C., Cohen, R. C., and Pye, H. O. T.: Temperature-Dependent Composition of Summertime PM2.5 in Observations and Model Predictions across the Eastern U.S, ACS Earth and Space Chemistry, 8, 381-392, 10.1021/acsearthspacechem.3c00333, 2024.

Vereecken, L. and Nozière, B.: H migration in peroxy radicals under atmospheric conditions, Atmos. Chem. Phys., 20, 7429-7458, 10.5194/acp-20-7429-2020, 2020.

Vigouroux, C., Langerock, B., Bauer Aquino, C. A., Blumenstock, T., Cheng, Z., De Mazière, M., De Smedt, I., Grutter, M., Hannigan, J. W., Jones, N., Kivi, R., Loyola, D., Lutsch, E., Mahieu, E., Makarova, M., Metzger, J. M., Morino, I., Murata, I., Nagahama, T., Notholt, J., Ortega, I., Palm, M., Pinardi, G., Röhling, A., Smale, D., Stremme, W., Strong, K., Sussmann, R., Té, Y., van Roozendael, M., Wang, P., and Winkler, H.: TROPOMI–Sentinel-5 Precursor formaldehyde validation using an extensive network of ground-based Fourier-transform infrared stations, Atmos. Meas. Tech., 13, 3751-3767, 10.5194/amt-13-3751-2020, 2020.

Warneke, C., Schwarz, J. P., Dibb, J., Kalashnikova, O., Frost, G., Al-Saad, J., Brown, S. S., Brewer, W. A., Soja, A., Seidel, F. C., Washenfelder, R. A., Wiggins, E. B., Moore, R. H., Anderson, B. E., Jordan, C., Yacovitch, T. I., Herndon, S. C., Liu, S., Kuwayama, T., Jaffe, D., Johnston, N., Selimovic, V., Yokelson, R., Giles, D. M., Holben, B. N., Goloub, P., Popovici, I., Trainer, M., Kumar, A., Pierce, R. B., Fahey, D., Roberts, J., Gargulinski, E. M., Peterson, D. A., Ye, X., Thapa, L. H., Saide, P. E., Fite, C. H., Holmes, C. D., Wang, S., Coggon, M. M., Decker, Z. C. J., Stockwell, C. E., Xu, L., Gkatzelis, G., Aikin, K., Lefer, B., Kaspari, J., Griffin, D., Zeng, L., Weber, R., Hastings, M., Chai, J., Wolfe, G. M., Hanisco, T. F., Liao, J., Campuzano Jost, P., Guo, H., Jimenez, J. L., Crawford, J., and Team, T. F.-A. S.: Fire Influence on Regional to Global Environments and Air Quality (FIREX-AQ), Journal of Geophysical Research: Atmospheres, 128, e2022JD037758, https://doi.org/10.1029/2022JD037758, 2023.

Wennberg, P. O., Bates, K. H., Crounse, J. D., Dodson, L. G., McVay, R. C., Mertens, L. A., Nguyen, T. B., Praske, E., Schwantes, R. H., Smarte, M. D., St Clair, J. M., Teng, A. P., Zhang, X., and Seinfeld, J. H.: Gas-Phase Reactions of Isoprene and Its Major Oxidation Products, Chemical Reviews, 118, 3337-3390, 10.1021/acs.chemrev.7b00439, 2018.

Wiser, F., Place, B. K., Sen, S., Pye, H. O. T., Yang, B., Westervelt, D. M., Henze, D. K., Fiore, A. M., and McNeill, V. F.: AMORE-Isoprene v1.0: a new reduced mechanism for gas-phase isoprene oxidation, Geosci. Model Dev., 16, 1801-1821, 10.5194/gmd-16-1801-2023, 2023.

Wolfe, G. M., Marvin, M. R., Roberts, S. J., Travis, K. R., and Liao, J.: The Framework for 0-D Atmospheric Modeling (F0AM) v3.1, Geosci. Model Dev., 9, 3309-3319, 10.5194/gmd-9-3309-2016, 2016a.

Wolfe, G. M., Kaiser, J., Hanisco, T. F., Keutsch, F. N., de Gouw, J. A., Gilman, J. B., Graus, M., Hatch, C. D., Holloway, J., Horowitz, L. W., Lee, B. H., Lerner, B. M., Lopez-Hilifiker, F., Mao, J., Marvin, M. R., Peischl, J., Pollack, I. B., Roberts, J. M., Ryerson, T. B., Thornton, J. A., Veres, P. R., and Warneke, C.: Formaldehyde production from isoprene oxidation across NOx regimes, Atmos. Chem. Phys., 16, 2597-2610, 10.5194/acp-16-2597-2016, 2016b.

Xu, L., Møller, K. H., Crounse, J. D., Kjaergaard, H. G., and Wennberg, P. O.: New Insights into the Radical Chemistry and Product Distribution in the OH-Initiated Oxidation of Benzene, Environmental Science & Technology, 54, 13467-13477, 10.1021/acs.est.0c04780, 2020.

Yang, B., Wiser, F. C., McNeill, V. F., Fiore, A. M., Tao, M., Henze, D. K., Sen, S., and Westervelt, D. M.: Implementation and evaluation of the automated model reduction (AMORE) version 1.1 isoprene oxidation mechanism in GEOS-Chem, Environmental Science: Atmospheres, 3, 1820-1833, 10.1039/D3EA00121K, 2023.

Yu, S., Jia, L., Xu, Y., and Pan, Y.: Molecular composition of secondary organic aerosol from styrene under different NOx and humidity conditions, Atmospheric Research, 266, 105950, https://doi.org/10.1016/j.atmosres.2021.105950, 2022.

Zhu, L., Jacob, D. J., Keutsch, F. N., Mickley, L. J., Scheffe, R., Strum, M., González Abad, G., Chance, K., Yang, K., Rappenglück, B., Millet, D. B., Baasandorj, M., Jaeglé, L., and Shah, V.: Formaldehyde (HCHO) As a Hazardous Air Pollutant: Mapping Surface Air Concentrations from Satellite and Inferring Cancer Risks in the United States, Environmental Science & Technology, 51, 5650-5657, 10.1021/acs.est.7b01356, 2017.

Zhu, Q., Schwantes, R. H., Coggon, M., Harkins, C., Schnell, J., He, J., Pye, H. O. T., Li, M., Baker, B., Moon, Z., Ahmadov, R., Pfannerstill, E. Y., Place, B., Wooldridge, P., Schulze, B. C., Arata, C., Bucholtz, A., Seinfeld, J. H., Warneke, C., Stockwell, C. E., Xu, L., Zuraski, K., Robinson, M. A., Neuman, J. A., Veres, P. R., Peischl, J., Brown, S. S., Goldstein, A. H., Cohen, R. C., and McDonald, B. C.: A better representation of volatile organic compound chemistry in WRF-Chem and its impact on ozone over Los Angeles, Atmos. Chem. Phys., 24, 5265-5286, 10.5194/acp-24-5265-2024, 2024.

Supplemental information for:

Role of chemical production and depositional losses on formaldehyde in the Community Regional Atmospheric Chemistry Multiphase Mechanism (CRACMM)

T. Nash Skipper, et al.

Correspondence to: Havala O.T. Pye (pye.havala@epa.gov)

Primary HCHO and total ROC emissions

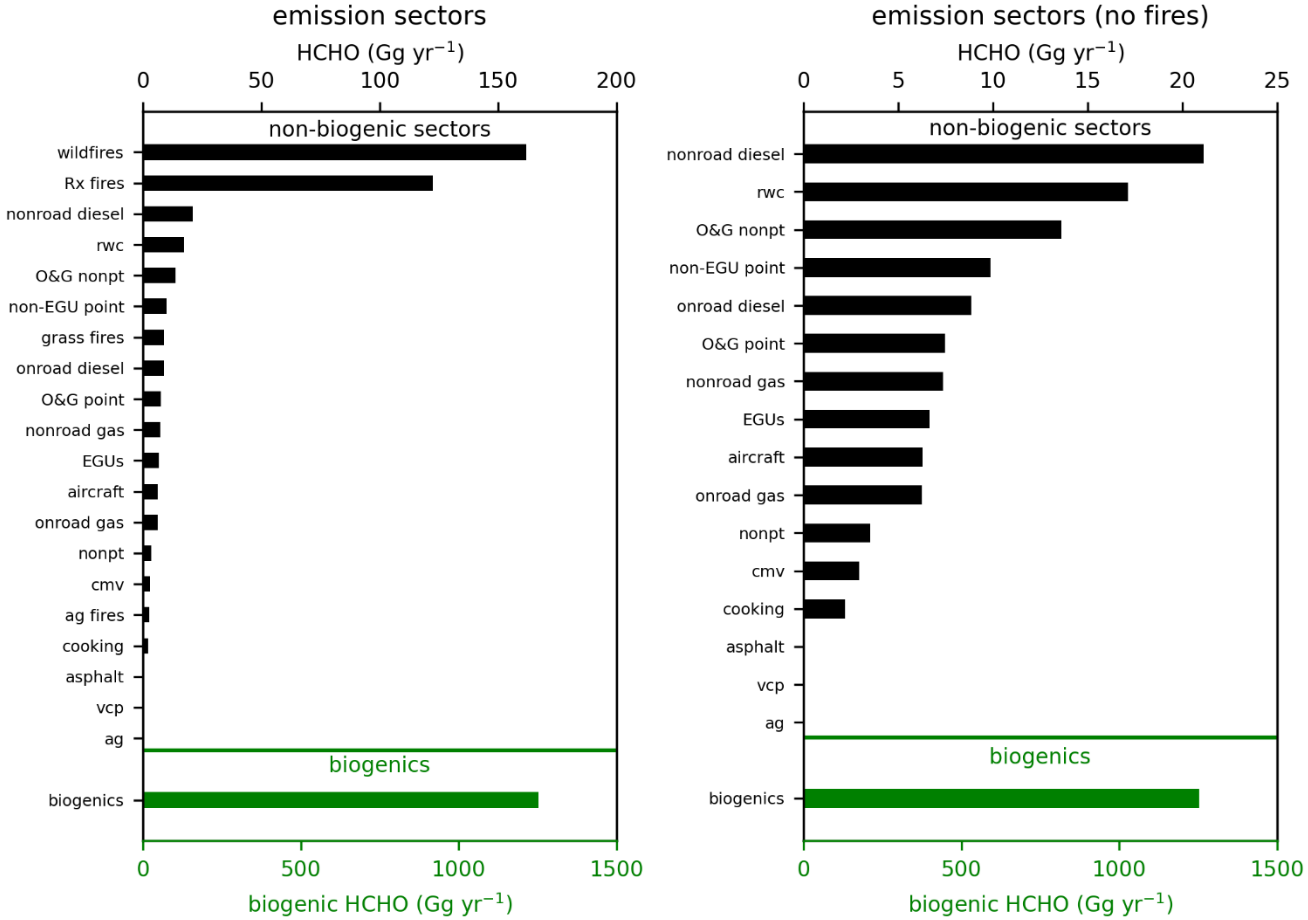


Figure S1. Primary HCHO by emission sector. Fires are included in the left panel and excluded on the right panel so that non-fire emission sector contributions can be seen in more detail.

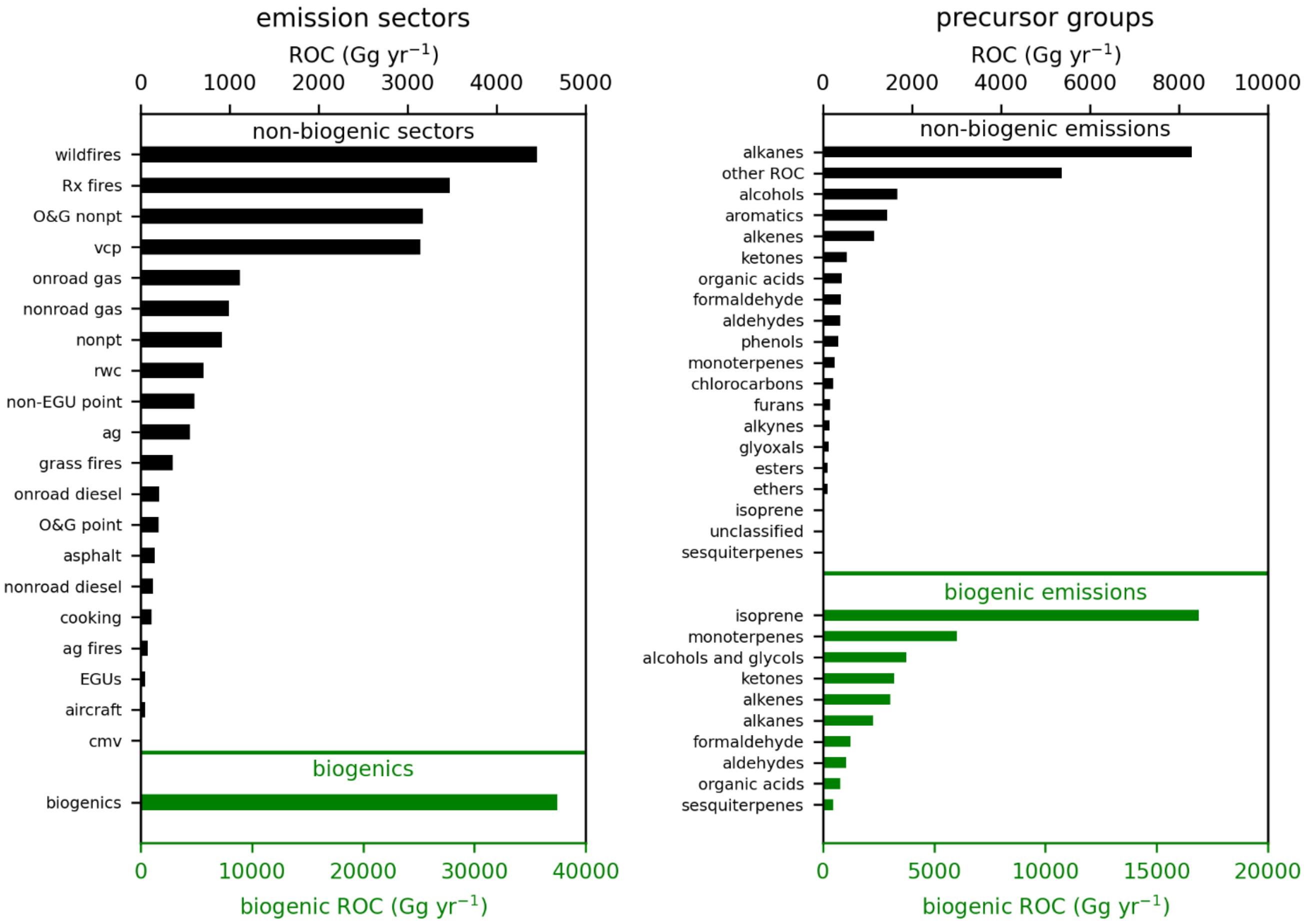


Figure S2. Total gas phase reactive organic carbon (ROC) emissions by emissions sector (left) and precursor group (right). Totals by emission sector (left) exclude primary HCHO which is shown separately in Fig.~~ure~~ S1. Totals by precursor group (right) show primary HCHO separately from other aldehydes.

CRACMM2 mechanism updates

Table S1. Reactions updated in CRACMM2.

	CRACMM1				CRACMM2				
Reaction label	**reactants**		**products**	**rate constant**	**reactants**		**products**	**rate constant**	**Note**
API system									
R189	APIP1 + NO	-->	0.82*HO2 + 0.82*NO2 + 0.82*PINAL + 0.18*TRPN	4.00E-12	APIP1 + NO	-->	0.65*(0.23*TRPN + 0.77*HO2 + 0.77*NO2 + 0.3*PINAL + 0.27*ALD + 0.09*ACT + 0.09*LIMAL + 0.21*HCHO + 0.11*OPB) + 0.35*(0.08*ACT + 0.49*HCHO + 0.2*LIMAL + 0.24*KET + 0.25*TRPN + 0.31*ALD + 0.75*HO2 + 0.75*NO2)	2.70E-12*exp(360/T)	1, 3
R339	APIP1 + NO3	-->	HO2 + NO2 + ALD + KET	1.20E-12	APIP1 + NO3	-->	0.65*(NO2 + HO2 + 0.39*PINAL + 0.35*ALD + 0.12*ACT + 0.12*LIMAL + 0.27*HCHO + 0.14*OPB) + 0.35*(0.11*ACT + 0.65*HCHO + 0.27*LIMAL + 0.32*KET + 0.41*ALD + HO2 + NO2)	2.30E-12	1, 3
R229	APIP1 + HO2	-->	OPB	1.50E-11	APIP1 + HO2	-->	0.65*(0.06*ACT + 0.06*LIMAL + 0.08*HCHO + 0.65*OPB + 0.48*HO2 + 0.29*PINAL + 0.35*HO) + 0.35*(0.97*OPB + 0.03*HO + 0.03*KET + 0.03*HCHO + 0.03*HO2)	2.60E-13*exp(1300/T)	1, 3
R265	APIP1 + MO2	-->	HO2 + 0.68*HCHO + 0.60*PINAL + 0.07*KET + 0.32*MOH + 0.25*ROH	3.56E-14*exp(708/T)	APIP1 + MO2	-->	0.65*(0.83*HCHO + 0.14*LIMAL + 0.42*PINAL + 0.2*ALD + 0.13*OPB + 0.17*MOH + 0.11*KET + 0.06*ACT + 1.16*HO2) + 0.35*(1.4*HCHO + 0.37*LIMAL + 0.32*KET + 1.5*HO2 + 0.08*ACT + 0.31*ALD)	2.00E-12	1, 3
R301	APIP1 + ACO3	-->	0.63*HO2 + 0.70*MO2 + 0.60*PINAL + 0.30*ORA2 + 0.07*KET + 0.25*ROH	7.40E-13*exp(765/T)	APIP1 + ACO3	-->	0.65*(0.39*PINAL + 0.35*ALD + 0.14*OPB + 0.12*ACT + 0.12*LIMAL + 0.27*HCHO + HO2 + MO2) + 0.35*(0.32*KET + 0.27*LIMAL + 0.41*ALD + 0.11*ACT + 0.65*HCHO + HO2 + MO2)	2.00E-12*exp(500/T)	1, 3

	CRACMM1				CRACMM2				
Reaction label	reactants		products	rate constant	reactants		products	rate constant	Note
R131	API + O3	-->	0.90*HO + 0.90*APIP1 + 0.05*APIP2 + 0.05*PINAL + 0.05*H2O2 + 0.14*CO	5.00E-16*exp(-530/T)	API + O3	-->	0.65*(0.77*HO + 0.33*PINALP + 0.22*H2O2 + 0.39*PINAL + 0.01*ORA2 + 0.17*HO2 + 0.17*CO + 0.27*HCHO + 0.27*RCO3) + 0.35*(0.51*KET + 0.3*HO + 0.3*RCO3 + 0.32*H2O2 + 0.19*HC3 + 0.81*HCHO + 0.11*OP1 + 0.08*ORA1)	8.05E-16*exp(-640/T)	1, 3
TRP14	APINP1 + NO	-->	2.00*NO2 + PINAL	4.00E-12	APINP1 + NO	-->	0.65*(1.86*NO2 + 0.07*TRPN + 0.07*ONIT + 0.93*PINAL) + 0.35*(0.54*TRPN + 0.07*ONIT + 1.39*NO2 + 0.44*ALD + 0.02*KET + 0.02*HCHO + 0.47*HO2)	2.7E-12*exp(360/T)	1, 3
TRP53	---	-->	---	---	APINP1 + NO3	-->	0.65*(2*NO2 + PINAL) + 0.35*(1.5*NO2 + 0.5*TRPN + 0.48*ALD + 0.02*KET + 0.02*HCHO + 0.5*HO2)	2.30E-12	1, 3
TRP22	APINP1 + HO2	-->	TRPN	1.50E-11	APINP1 + HO2	-->	0.65*(0.3*TRPN + 0.7*PINAL + 0.7*NO2 + 0.7*HO) + 0.35*(0.47*HO + 0.76*TRPN + 0.22*ALD + 0.02*KET + 0.24*NO2 + 0.02*HCHO)	2.71E-13*exp(1300/T)	1, 3
TRP30	APINP1 + MO2	-->	0.37*HO2 + 0.86*NO2 + 0.68*HCHO + 0.86*PINAL + 0.32*MOH + 0.14*TRPN	3.56E-14*exp(708/T)	APINP1 + MO2	-->	0.65*(0.18*TRPN + 0.95*HCHO + 0.05*MOH + 0.82*HO2 + 0.82*NO2 + 0.82*PINAL) + 0.35*(0.64*TRPN + 0.02*KET + 0.34*ALD + 0.36*NO2 + 1.1*HO2 + 0.99*HCHO + 0.03*MOH)	2.00E-12	1, 3
TRP36	APINP1 + ACO3	-->	0.86*NO2 + 0.14*TRPN + 0.86*PINAL + 0.70*MO2 + 0.30*ORA2	7.40E-13*exp(765/T)	APINP1 + ACO3	-->	0.65*(NO2 + PINAL + MO2) + 0.35*(MO2 + 0.5*NO2 + 0.5*TRPN + 0.48*ALD + 0.02*KET + 0.02*HCHO + 0.5*HO2)	2.0E-12*exp(500/T)	1, 3
TRP19	PINALP + NO	-->	0.95*HO2 + 0.95*NO2 + 0.05*TRPN + 0.95*HCHO + 0.95*KET	2.70E-12*exp(360/T)	PINALP + NO	-->	0.36*TRPN + 0.64*HOM + 0.64*NO2	2.7E-12*exp(360/T)	1, 4, 5
TRP27	PINALP + HO2	-->	OPB	2.91E-13*exp(1300/T)	PINALP + HO2	-->	0.75*OPB + 0.25*HO + 0.25*HOM	2.71E-13*exp(1300/T)	1, 4, 5
TRP11	PINALP	-->	HOM	1	PINALP	-->	HOM	0.029	6
LIM system									
R190	LIMP1 + NO	-->	0.77*HO2 + 0.77*NO2 + 0.49*LIMAL + 0.28*HCHO + 0.28*UALD + 0.23*TRPN	4.00E-12	LIMP1 + NO	-->	0.23*TRPN + 0.77*NO2 + 0.77*LIMAL + 0.77*HO2 + 0.43*HCHO	2.70E-12*exp(360/T)	2, 3

	CRACMM1				CRACMM2				
Reaction label	**reactants**		**products**	**rate constant**	**reactants**		**products**	**rate constant**	**Note**
R340	LIMP1 + NO3	-->	HO2 + NO2 + 0.38500*OLI + 0.38500*HCHO + 0.61500*MACR	1.20E-12	LIMP1 + NO3	-->	NO2 + LIMAL + HO2 + 0.56*HCHO	2.30E-12	2, 3
R230	LIMP1 + HO2	-->	OPB	1.50E-11	LIMP1 + HO2	-->	0.9*OPB + 0.1*LIMAL + 0.1*HO + 0.1*HO2 + 0.06*HCHO	2.60E-13*exp(1300/T)	2, 3
R266	LIMP1 + MO2	-->	HO2 + HCHO + 0.42*LIMAL + 0.30*KET + 0.32*MOH + 0.27*ROH	3.560E-14*exp(708/T)	LIMP1 + MO2	-->	0.25*MOH + LIMAL + 1.03*HCHO + HO2	2.00E-12	2, 3
R302	LIMP1 + ACO3	-->	0.63*HO2 + 0.70*MO2 + 0.42*LIMAL + 0.30*KET + 0.30*ORA2 + 0.32*HCHO + 0.27*ROH	7.40E-13*exp(765/T)	LIMP1 + ACO3	-->	LIMAL + 0.56*HCHO + HO2 + MO2	2.00E-12*exp(500/T)	2, 3
R132	LIM + O3	-->	0.84*HO + 0.84*LIMP1 + 0.11*LIMP2 + 0.05*LIMAL + 0.05*H2O2 + 0.14*CO	2.95E-15*exp(-783/T)	LIM + O3	-->	0.66*HO + 0.66*LIMAL + 0.33*ACO3 + 0.33*HCHO + 0.33*RCO3 + 0.33*H2O2 + 0.01*ORA2	2.8E-15*exp(-770/T)	2, 3
TRP17	LIMNP1 + NO	-->	2.00*NO2 + LIMAL	4.00E-12	LIMNP1 + NO	-->	0.57*TRPN + 0.07*ONIT + 1.36*NO2 + 0.43*LIMAL + 0.5*HO2	2.7E-12*exp(360/T)	2, 3
TRP54	---	-->	---	---	LIMNP1 + NO3	-->	1.46*NO2 + 0.46*LIMAL + 0.54*TRPN + 0.54*HO2	2.30E-12	2, 3
TRP25	LIMNP1 + HO2	-->	TRPN	1.50E-11	LIMNP1 + HO2	-->	0.77*TRPN + 0.5*HO + 0.23*LIMAL + 0.23*NO2 + 0.27*HO2	2.71E-13*exp(1300/T)	2, 3
TRP33	LIMNP1 + MO2	-->	0.37*HO2 + 0.68*HCHO + 0.70*LIMAL + 0.70*NO2 + 0.32*MOH + 0.30*TRPN	3.56E-14*exp(708/T)	LIMNP1 + MO2	-->	0.69*TRPN + 0.91*HCHO + 0.09*MOH + 1.01*HO2 + 0.31*LIMAL + 0.31*NO2	2.00E-12	2, 3
TRP39	LIMNP1 + ACO3	-->	0.70*NO2 + 0.70*LIMAL + 0.30*TRPN + 0.70*MO2 + 0.30*ORA2	7.40E-13*exp(765/T)	LIMNP1 + ACO3	-->	MO2 + 0.46*NO2 + 0.46*LIMAL + 0.54*TRPN + 0.54*HO2	2.0E-12*exp(500/T)	2, 3
TRP05	LIMAL + HO	-->	0.70*LIMALP + 0.30*RCO3	1.00E-10	LIMAL + HO	-->	0.83*LIMALP + 0.17*RCO3	1.10E-10	2, 3
TRP20	LIMALP + NO	-->	0.94*HO2 + 0.94*NO2 + 0.06*TRPN + 0.94*HCHO + 0.94*KET	2.70E-12*exp(360/T)	LIMALP + NO	-->	0.64*TRPN + 0.36*NO2 + 0.36*HO2 + 0.36*HCHO + 0.36*PAA	2.7E-12*exp(360/T)	2, 4, 7
TRP28	LIMALP + HO2	-->	OPB	2.91E-13*exp(1300/T)	LIMALP + HO2	-->	0.9*OPB + 0.1*HO + 0.1*HO2 + 0.1*HCHO + 0.1*PAA	2.73E-13*exp(1300/T)	2, 4, 7
TRP12	LIMALP	-->	HOM	1	LIMALP	-->	HOM	0.024	6
TRP08	LIMAL + O3	-->	0.04*HO + 0.67*HC10P + 0.79*HCHO + 0.33*KET + 0.04*HO2 + 0.20*CO	8.30E-18	LIMAL + O3	-->	0.09*HO + ALD + 0.62*HCHO + 0.23*OP1 + 0.02*H2O2 + 0.15*ORA1	8.30E-18	2, 3
Monoterpene nitrate system									
TRP07	TRPN + HO	-->	HOM	4.80E-12	TRPN + HO	-->	0.33*HONIT + 0.67*NO2 + 0.27*PINAL + 0.38*KET + 0.21*HCHO + 0.02*ALD	4.80E-12	8
TRP09	TRPN + O3	-->	HOM	1.67E-16	TRPN + O3	-->	0.33*HONIT + 0.67*NO2 + 0.27*PINAL + 0.38*KET + 0.21*HCHO + 0.02*ALD	1.67E-16	8

	CRACMM1				CRACMM2				
Reaction label	**reactants**		**products**	**rate constant**	**reactants**		**products**	**rate constant**	**Note**
TRP10	TRPN + NO3	-->	HOM	3.15E-14*exp(-448/T)	TRPN + NO3	-->	0.33*HONIT + 0.67*NO2 + 0.27*PINAL + 0.38*KET + 0.21*HCHO + 0.02*ALD	3.15E-13*exp(-448/T)	8
TRP55	---	-->	---	---	TRPN + hv	-->	NO2 + 0.67*KET + 0.33*UALD	ONIT_RACM2	9
TRP56	---	-->	---	---	HONIT + hv	-->	HKET + NO2	ONIT_RACM2	9
TRP57	---	-->	---	---	HONIT + HO	-->	HKET + NO3	same as HNO3 + OH: xk0=2.40E-14*exp(460/T) xk2=2.70E-17*exp(2199/T) xk3=6.50E-34*exp(1335/T) k = xk0+xk3*M/(1.0+xk3*M/xk2)	10
TRP58	---	-->	---	---	ATRPNJ	-->	AHOMJ + HNO3	9.26E-05	11
TRP59	---	-->	---	---	AHONITJ	-->	AHOMJ + HNO3	9.26E-05	11
STY system									
ROCARO71	---	-->	---	---	STY + HO	-->	STYP	5.80E-11	12
ROCARO72	---	-->	---	---	STYP + HO2	-->	VROCP3OXY2	2.91E-13*exp(1300/T)	12, 13
ROCARO73	---	-->	---	---	STYP + NO	-->	NO2 + HO2 + HCHO + BALD	2.7E-12*exp(360/T)	12
ROCARO74	---	-->	---	---	STYP + NO3	-->	NO2 + HO2 + HCHO + BALD	2.30E-12	12
ROCARO75	---	-->	---	---	STYP + MO2	-->	HO2 + HCHO + BALD + 0.68*HCHO + 0.37*HO2 + 0.32*MOH	2.50E-13	12
ROCARO76	---	-->	---	---	STYP + ACO3	-->	HO2 + HCHO + BALD + 0.7*MO2 + 0.3*ORA2	2.50E-13	12
ISO system									
R130/RAM01	ISO + O3	-->	0.25*HO + 0.25*HO2 + 0.08*MO2 + 0.10*ACO3 + 0.10*MACP + 0.09*H2O2 + 0.14*CO + 0.58*HCHO + 0.46100*MACR + 0.18900*MVK + 0.28*ORA1 + 0.15300*OLT	7.86E-15*exp(-1913/T)	ISO + O3	-->	0.25*HO + 0.25*HO2 + 0.40*MO2 + 0.01800*ACO3 + 0.10*MACP + 0.09*H2O2 + 0.22*CO + HCHO + 0.30*MACR + 0.14*MVK + 0.28*ORA1 + 0.15300*OLT	1.58E-14*exp(-2000/T)	14
R145/RAM02	ISO + NO3	-->	ISON	3.03E-12*exp(-446/T)	ISO + NO3	-->	0.40*NO2 + 0.04500*ISON + 0.35*HCHO + 0.55500*INO2 + 0.26*MVK + 0.02800*MACR	2.95E-12*exp(-450/T)	14
R086/RAM03	ISO + HO	-->	ISOP	2.70E-11*exp(390/T)	ISO + HO	-->	ISOP + 0.25*HCHO + 0.03*MACR + 0.05*MGLY	2.69E-11*exp(390/T)	14
R228/RAM04	ISOP + HO2	-->	ISHP	2.05E-13*exp(1300/T)	ISOP + HO2	-->	ISHP + 0.07*HO2 + 0.50*HO	4.50E-13*exp(1300/T)	14
R188/RAM05	ISOP + NO	-->	0.88*HO2 + 0.88*NO2 + 0.20*HCHO + 0.28*MACR + 0.44*MVK + 0.12*ISON + 0.02100*GLY + 0.02900*HKET + 0.02700*ALD	2.43E-12*exp(360/T)	ISOP + NO	-->	0.13*ISON + 0.40*HCHO + 0.88*HO2 + 0.87*NO2 + 0.18*MACR + 0.51*MVK	6.00E-12*exp(350/T)	14
R115/RAM06	ISHP + HO	-->	HO + MACR + 0.90400*IEPOX	1.00E-10	ISHP + HO	-->	ISOP	4.60E-12*exp(200/T)	14

	CRACMM1				CRACMM2				
Reaction label	**reactants**		**products**	**rate constant**	**reactants**		**products**	**rate constant**	**Note**
RAM07	---	-->	---	---	ISHP + HO	-->	0.04*MGLY + 0.02*GLY + 0.13*MVK + 0.44*IEPOX + 0.11*ACO3 + 0.03*MACR + 2.00*HO + 0.34*HO2 + 0.14*IPX + CO	2.97E-11*exp(390/T)	14
RAM08	---	-->	---	---	INO2 + HO2	-->	0.45*HO + 0.95*INALD + 0.02*IPX	3.14E-11*exp(580/T)	14
RAM09	---	-->	---	---	INO2 + NO	-->	0.15*MVK + 0.65*INALD + 0.05*ISON + 0.20*HCHO + 1.30*NO2	9.42E-12*exp(580/T)	14
R125/RAM10	ISON + HO	-->	INALD + 0.07*HKET + 0.07*HCHO	1.30E-11	ISON + HO	-->	HO + 0.35*INALD + 0.17*IEPOX + 0.65*NO2	2.40E-11*exp(390/T)	14
R124/RAM11	NALD + HO	-->	NO2 + XO2 + HKET	5.60E-12*exp(270/T)	INALD + HO	-->	CO + NO2 + 0.30*HO2 + HCHO	1.50E-11	14
RAM12	---	-->	---	---	ISON	-->	HNO3 + ROH	4.00E-05	14
RAM13	---	-->	---	---	IPX + HO	-->	0.57*MACR + 0.43*MVK	3.00E-12	14
Hetereogeneous reactions									
HET_IPX	---	-->	---	---	IPX	-->	AISO4J	2*$k_{\text{het IEPOX}}$	15
HET_INALD	---	-->	---	---	INALD	-->	AISO5J + HNO3	0.5*$k_{\text{het IEPOX}}$	15
HET_HO2	---	-->	---	---	HO2	-->	H2O	$k_{\text{het HO2}}$ (γ = 0.2)	16
HET_NO3	---	-->	---	---	NO3	-->	HNO3	$k_{\text{het NO3}}$ (γ = 1E-3)	17
Methane									
R364	---	-->	---	---	ECH4 + HO	-->	MO2	2.45E-12*exp(-1775/T)	18

Notes for Table S~~3~~1.

1. Products in the API system have been updated to include products from α-pinene and β-pinene, assuming a 65/35 split of α-/β-pinene (based on biogenic emission totals over the contiguous US). Updated reactions are based on α-pinene and β-pinene chemistry in the MOZART-TS2 mechanism developed by Schwantes et al. (2020).

2. Reactions and products in the LIM system have been updated based on limonene chemistry in the MOZART-TS2 mechanism developed by Schwantes et al. (2020).

3. Species from MOZART-TS2 have been mapped to existing CRACMM species. Mapping from MOZART-TS2 to CRACMM2 includes the following:
    a. TS2 monoterpene nitrate species were mapped to CRACMM species TRPN.
        i. Except TS2 species TERPFDN (monoterpene derived di-nitrate) is mapped to CRACMM species TRPN+ONIT for conservation of nitrogen.
    b. TS2 species TERPA is mapped to CRACMM species PINAL.
    c. TS2 species TERPA3 is mapped to CRACMM species ALD.
    d. TS2 species TERPF1 is mapped to CRACMM species LIMAL.
    e. TS2 species TERP1OOH and TERPOOH are mapped to CRACMM species OPB.
    f. TS2 species TERPK is mapped to CRACMM species KET.
    g. TS2 species TERPA2O2 is mapped to CRACMM species PINALP.
    h. TS2 species TERPACID is mapped to CRACMM species ORA2.
    i. TS2 species TERPA2 is mapped to CRACMM species PINAL.
    j. TS2 species TERPA2CO3 is mapped to CRACMM species RCO3.
    k. TS2 species BIGALK is mapped to CRACMM species HC3.
    l. TS2 species HMHP is mapped to CRACMM species OP1.

4. The following characteristics of monoterpene aldehyde derived peroxy radical (PINALP and LIMALP) reactions are based on Wennberg et al. (2018):
    a. Nitrate (CRACMM species TRPN) branching ratio from reaction with NO
    b. Rate of reaction with HO2
    c. Yield of peroxide (CRACMM species OPB) from reaction with HO2

5. Products of PINALP reaction with NO and HO2 that do not go to TRPN or OPB, respectively, are expected to form a ring-opening peroxy radical that undergoes autoxidation which is mapped to CRACMM species HOM.

6. Unimolecular autoxidation rates of PINALP and LIMALP are calculated based on Vereecken and Nozière (2020) and assuming that the first competitive step rather than the first step best reflects the rate of HOM formation in a reduced mechanism. Details:
    a. For PINALP rate:
        i. From Table 1 k(298 K) = 6.2E-4 s^{-1} (1,5-H migration; substitution pattern –CH< >C(OO•)-)
        ii. From Table 5 k(298 K) correction factor = 47 (exo-β-oxo 1,5)
        iii. Overall k = 6.2E-4 × 47 = 0.029 s^{-1}
    b. For LIMALP rate both 1,5-H shift and 1,6-H shift are possible. We consider both options and add the rate constant from the two options to calculate the overall rate constant:
        i. From Table1:
            1. k_1(298 K) = 4.49E-4 s^{-1} (1,5-H migration; substitution pattern –CH_2– >C(OO•)-)
            2. k_2(298 K) = 3.12E-4 s^{-1} (1,6-H migration; substitution pattern –CH_2– >C(OO•)-)
        ii. From Table 5:
            1. k_1(298 K) correction factor = 47 (exo-β-oxo 1,5)
            2. k_2(298 K) correction factor = 10 (exo-β-oxo 1,6)
        iii. Overall k = 4.49E-4 × 47 + 3.12E-4 × 10 = 0.024 s^{-1}

7. Products of LIMALP reaction with NO and HO2 that do not go to TRPN or OPB, respectively, are expected to form a peroxy acid (mapped to CRACMM species PAA) and formaldehyde (CRACMM species HCHO).

8. Oxidation of TRPN assumes 1/3 is unsaturated and retains the nitrate group upon oxidation to from HONIT while the remaining 2/3 is saturated and releases the nitrate group to form NO2 and other products. The 1/3 to 2/3 split is based on Fisher et al. (2016). The coefficients of other products are based on products from α-pinene and β-pinene derived nitrates in MCM with a 65/35 weighting as described in Note 1 above. From MCM products, we find a 62% yield of PINAL and a 38% yield of a ketone (KET) for α-pinene and a 92% yield of a ketone (KET) + HCHO and an 8% yield of an aldehyde (ALD) for β-pinene.

9. Photolysis rates of monoterpene nitrates are set to the same rate as the other organic nitrate species in CRACMM (ONIT). Photolysis products are NO2 and fragmentation products.

10. Oxidation of HONIT is based on Fisher et al. (2016) and Browne et al. (2014).

11. ATRPN and AHONIT are expected to undergo hydrolysis with a 3-h lifetime based on Pye et al. (2015).

12. Styrene chemistry is adopted from MCM for CRACMM2.

13. MCM predicts a peroxide formed from the reaction of the styrene peroxy radical with HO2. In CRACMM2, this peroxide is mapped to an oxygenated IVOC with $C^*=10^3$ μg/m^3 and O:C=0.2 (CRACMM species VROCP3OXY2).

14. From AMOREv1.2 isoprene condensation of a detailed isoprene mechanism by Wennberg et al. (2018). For more details on the initial development of the AMORE technique and condensed isoprene mechanism (i.e., AMOREv1.0), Wiser et al. (2023). For more details specific to AMOREv1.2, see Section ~~3~~2.1 of the main text.

15. Heterogeneous uptake of isoprene derived nitrates (INALD) and isoprene derived tetrafunctional compounds (IPX) are set to be proportional to IEPOX uptake.

16. HO2 radical uptake is based on Ivatt et al. (2022).

17. NO3 radical uptake is based on Jacob (2000) and Zhu et al. (2024).

18. Emitted methane (ECH4) is tracked separately to account for effects of methane emissions on top of the global average background methane (set to 1850 ppb in CMAQ by default).

AMORE v1.2 updates overview

The AMORE v1.0 Isoprene Mechanism was developed from the Caltech Isoprene Mechanism using a graph theory-based algorithm for the reduction of large chemical mechanisms (Wiser et al., 2023). This mechanism was added as an option in CRACMM1. Subsequent updates were made from new GEOS-Chem data (Yang et al., 2023), CMAQ data and box model results (this work). The updated AMORE v1.2 isoprene mechanism is included with this mechanism update, and includes improved NOx chemistry, improved VOC yields, including formaldehyde, and additional SOA pathways from isoprene derived tetrafunctional compounds.

GEOS-Chem simulation results showed that the AMORE v1.0 mechanism had elevated NOx production leading to increased ozone bias. This was confirmed with CMAQ simulations and replicated in box model simulations, and the NOx chemistry was adjusted by reducing the yield of NO2 to better represent the full reference mechanism, leading to improved NOx chemistry and reduced ozone production, which resulted in stronger model agreement with atmospheric data. For further information, see Yang et al. (2023).

CMAQ simulations showed that the AMORE v1.0 formaldehyde yield was notably higher than base CRACMM and higher than TROPOMI measurements, despite the AMORE v1.0 mechanism having strong agreement with box model runs. We found that the cause of this discrepancy was in the tail end production of formaldehyde, which occurs at timescales on the order of multiple days. This long tail formaldehyde production was captured in select box model runs but not in measured data, where transport and condensation would likely limit tail end formaldehyde production. Running the box model with diurnal cycle data reduced the discrepancy between the box model and the measured data. The primary source of formaldehyde in the AMORE mechanism is from the reaction of isoprene with OH. The molar yield of formaldehyde from this reaction was adjusted to determine which value best represented the full mechanism and the intended use in transport models. As shown in Fig.~~ure~~ S~~5~~3, in a diurnal cycle run, the 0.25 molar yield mechanism has much stronger agreement with the reference mechanism after one day. From two to three days, the mechanisms begin to diverge, with the reference mechanism having a higher formaldehyde yield. This figure shows the overall net production of formaldehyde, as dilution and formaldehyde decomposition reactions were turned off.

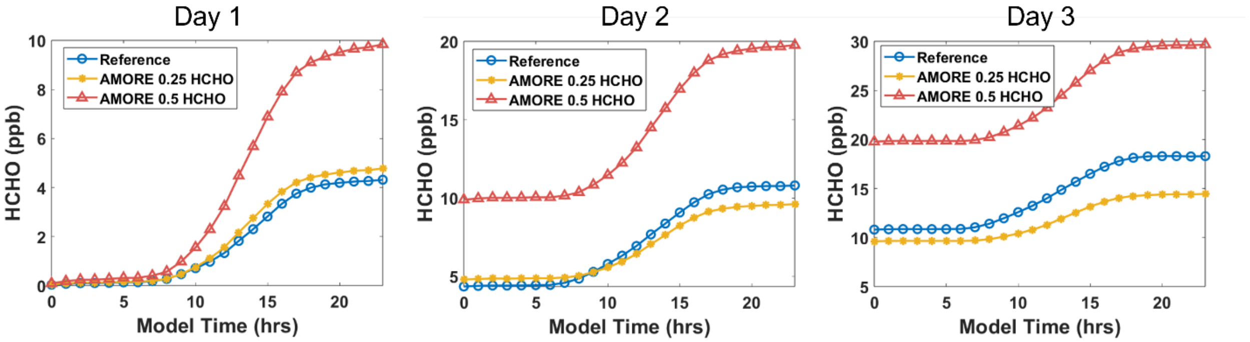


Figure S~~5~~3. Concentration of formaldehyde for the 0.25 molar yield AMORE mechanism vs. the 0.5 molar yield AMORE mechanism in comparison to the Caltech Full Reference Isoprene mechanism. These plots show three consecutive days using the F0AM diurnal cycle run with hourly concentration data for isoprene and reactive atmospheric gases along with solar intensity values from 6/30/2013 in Centerville, Alabama. Data is from the SOAS field campaign. Ozone and NOx concentrations are set to zero to focus on the primary formaldehyde pathway from isoprene reacting with OH. Dilution is set to zero and formaldehyde decomposition reactions are removed to demonstrate accumulated yield of formaldehyde over the entire run.

We conducted a run with several different formaldehyde molar yields in CMAQ, ranging from 0.25 to 0.5. The results are shown in Fig.~~ure~~ S~~6~~4, with comparisons to the CRACMM1 baseline mechanism and TROPOMI data. All AMORE mechanisms showed increased formaldehyde production from the CRACMM1 baseline mechanism. The CRACMM1 baseline mechanism had negative formaldehyde bias compared to the TROPOMI data. The overall bias was reduced for each AMORE mechanism except for the 0.5 molar yield mechanism, which had significant positive formaldehyde bias.

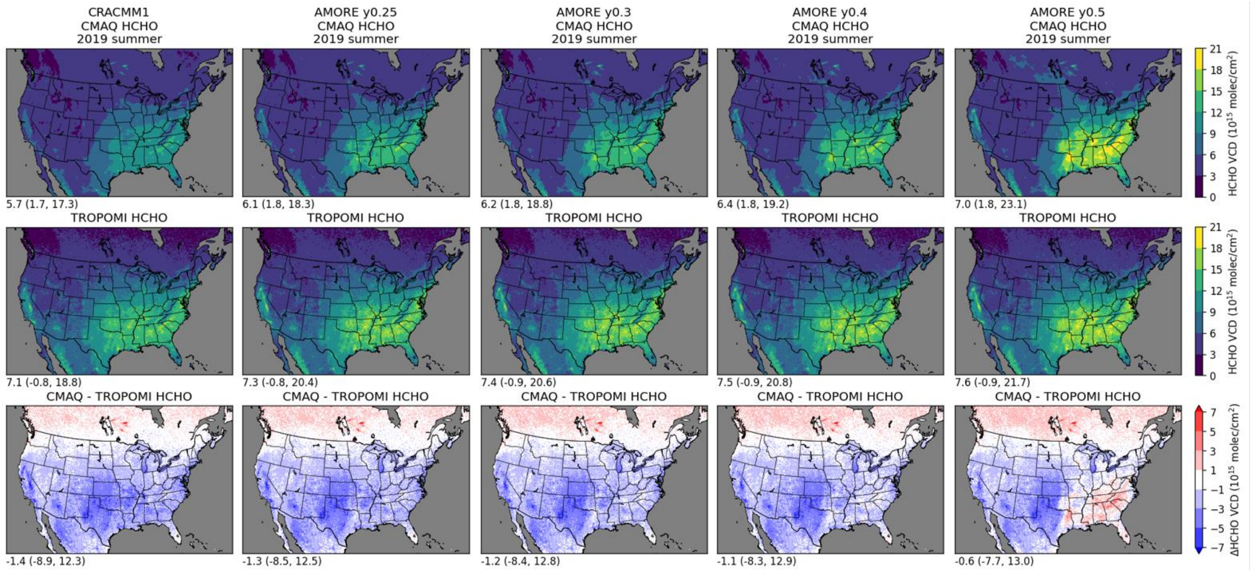


Figure S~~6~~4. CMAQ simulations of 4 different formaldehyde yields for the AMORE mechanism, in comparison to the CRACMM1 baseline mechanism and TROPOMI data. The simulation and data are taken from the summer of 2019. The first row shows the simulated concentration of formaldehyde for each mechanism. The second row shows the measured TROPOMI values, and the final row shows the bias between the simulation and the TROPOMI values.

Ultimately, the TROPOMI data is a useful comparison point but cannot be used to calibrate the formaldehyde yield, as there is no way to isolate isoprene derived formaldehyde from this data. According to the box model results, any discrepancies between the AMORE 0.25 yield mechanism and the reference mechanism would only show up after 2 days, and only reach significant levels after 3 days. Most of this tail end formaldehyde production is the result of aged, highly oxidized isoprene derived species, which would suggest that other processes such as deposition, condensation, and transport might dominate at this timescale, leading to suppressed production of formaldehyde in the gas phase. Thus, the 0.25 formaldehyde molar yield was chosen for the AMORE v1.2 mechanism used in this work. We made similar adjustments to other VOC species such as methylglyoxal and glyoxal.

Isoprene derived tetrafunctional compounds, including multifunctional epoxydiols which are similar but distinct from isoprene epoxy-diol, are potentially significant sources of SOA from isoprene (Bates et al). They estimate that IEPOX, organonitrates, and isoprene tetrafunctional compounds all contribute approximately 30% each of isoprene derived SOA. Of these, isoprene tetrafunctional compounds have the least studied chemistry, resulting in some uncertainty in their SOA yields. A lumped species was added to the AMORE v1.2 isoprene mechanism to represent this group, and an SOA formation pathway was added with the same mechanism as IEPOX, but with a lower SOA yield due to uncertainties in the behavior of these species. The molar yield of this new species was determined by comparison to the set of tetrafunctional compounds represented in the Caltech full isoprene mechanism. The concentration plot comparison is shown in Fig.~~ure~~ S7.

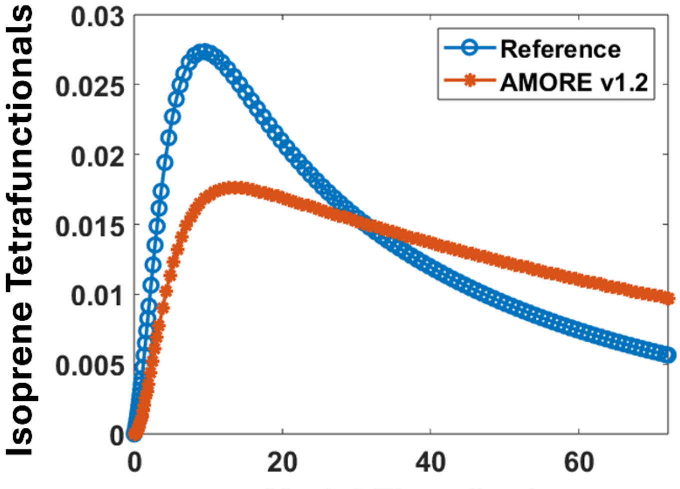


Figure S~~7~~5. This figure shows the concentration of isoprene tetrafunctional compounds for the AMORE v1.2 mechanism and the Caltech reference full isoprene mechanism over a 72-hour F0AM box model run with moderate light conditions, 1 ppb of isoprene, 4e-5 ppb OH, 10 ppb O_3, 0.005 ppb HO2, and 0.05 ppb NO.

Pinene biogenic emissions

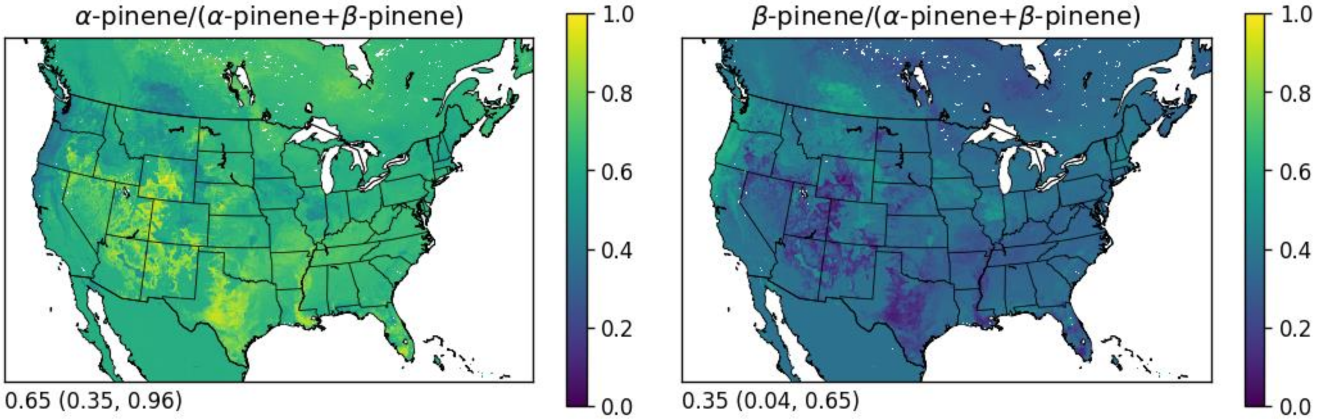


Figure S~~8~~6. Ratio of α-pinene and β-pinene to total pinenes based on the annual sum of biogenic emissions from BEIS. The values for α-pinene and β-pinene include additional monoterpene emissions from BEIS that would map to these species. Monoterpenes mapped to β-pinene are those that have one double bond which is a terminal double bond and include β -pinene, camphene, and sabinene. Monoterpenes mapped to α-pinene are all other monoterpenes with one double bond that is not terminal and include α-pinene, δ-3-carene, and α-thujene. The annotations below each panel show the mean (min, max) over land grid cells.

Additional information on F0AM simulations

Table S~~1~~2. Description of each emission sector used in the F0AM box modeling simulations. Most descriptions are based on information provided with the 2019 Emissions Modeling Platform (U.S. EPA, 2022).

Sector	Description
biogenics	Emissions from BEIS; typically simulated in-line in CMAQ
wildfires	Wildfire emissions
Rx fires	Prescribed (Rx) burning emissions
grass fires	Grassland burning emissions
ag fires	Agricultural burning emissions
rwc	Residential wood combustion emissions
O&G point	Point source oil and gas emissions
O&G nonpt	Area source oil and gas emissions
vcp	Volatile chemical product emissions
onroad gas	Highway mobile source emissions from gasoline vehicles
nonroad gas	Off highway mobile source emissions from gasoline vehicles
onroad diesel	Highway mobile source emissions from diesel vehicles
nonroad diesel	Off highway mobile source emissions from diesel vehicles
nonpt	Area source emissions not included in other sectors
non-EGU point	Point source emissions from industrial activities (also known as ptnonipm)
ag	Agricultural emissions
asphalt	Asphalt paving emissions
cooking	Cooking emissions
EGUs	Electric generating unit emissions
aircraft	Aircraft emissions
cmv	Class 1, 2, and 3 commercial marine vessel emissions

Table S~~2~~3. Precursor groups used in F0AM simulations. The MCM column indicates the MCM species included in each group. The CRACMM1 and CRACMM2 columns indicate the species from those mechanisms that best corresponds with each MCM species. Species mappings that differ in CRACMM1 and CRACMM2 are colored blue.

group	MCM	CRACMM1	CRACMM2
alcohols	CH3OH	MOH	MOH
	C2H5OH	EOH	EOH
	NPROPOL	ROH	ROH
	IPROPOL	ROH	ROH
	NBUTOL	ROH	ROH
	BUT2OL	ROH	ROH
	IBUTOL	ROH	ROH
	TBUTOL	ROH	ROH
	PECOH	ROH	ROH
	IPEAOH	ROH	ROH
	ME3BUOL	ROH	ROH
	IPECOH	ROH	ROH
	IPEBOH	ROH	ROH

group	MCM	CRACMM1	CRACMM2
	CYHEXOL	ROH	ROH
	MIBKAOH	HKET	HKET
	ETHGLY	ETEG	ETEG
	PROPGLY	PROG	PROG
	MBO	OLT	OLT
formaldehyde	HCHO	HCHO	HCHO
aldehydes	CH3CHO	ACD	ACD
	C2H5CHO	ALD	ALD
	C3H7CHO	ALD	ALD
	IPRCHO	ALD	ALD
	C4H9CHO	ALD	ALD
	ACR	ACRO	ACRO
	MACR	MACR	MACR
	C4ALDB	MACR	MACR
alkanes	CH4	n/a	ECH4
	C2H6	ETH	ETH
	C3H8	HC3	HC3
	NC4H10	HC3	HC3
	IC4H10	HC3	HC3
	NC5H12	HC5	HC5
	IC5H12	HC5	HC5
	NEOP	HC3	HC3
	NC6H14	HC5	HC5
	M2PE	HC5	HC5
	M3PE	HC5	HC5
	M22C4	HC3	HC3
	M23C4	HC5	HC5
	NC7H16	HC10	HC10
	M2HEX	HC5	HC5
	M3HEX	HC5	HC5
	NC8H18	HC10	HC10
	NC9H20	HC10	HC10
	NC10H22	HC10	HC10
	NC11H24	HC10	HC10
	NC12H26	ROCP6ALK	ROCP6ALK
	CHEX	HC10	HC10
alkenes	C2H4	ETE	ETE
	C3H6	OLT	OLT
	BUT1ENE	OLT	OLT
	CBUT2ENE	OLI	OLI
	TBUT2ENE	OLI	OLI
	MEPROPENE	OLT	OLT

group	MCM	CRACMM1	CRACMM2
	PENT1ENE	OLT	OLT
	CPENT2ENE	OLI	OLI
	TPENT2ENE	OLI	OLI
	ME2BUT1ENE	OLT	OLT
	ME3BUT1ENE	OLT	OLT
	ME2BUT2ENE	OLI	OLI
	HEX1ENE	OLT	OLT
	CHEX2ENE	OLI	OLI
	THEX2ENE	OLI	OLI
	DM23BU2ENE	OLI	OLI
	C4H6	BDE13	BDE13
alkynes	C2H2	ACE	ACE
aromatics	BENZENE	BEN	BEN
	TOLUENE	TOL	TOL
	OXYL	XYE	XYL
	MXYL	XYM	XYL
	PXYL	XYE	XYL
	EBENZ	XYE	EBZ
	PBENZ	XYE	XYL
	IPBENZ	XYE	XYL
	TM123B	XYM	XYL
	TM124B	XYM	XYL
	TM135B	XYM	XYL
	OETHTOL	XYE	XYL
	METHTOL	XYM	XYL
	PETHTOL	XYE	XYL
	DIME35EB	XYE	XYL
	DIET35TOL	XYM	XYL
	STYRENE	XYM	STY
	BENZAL	BALD	BALD
	MCATECHOL	MCT	MCT
	n/a (use MXYL)	NAPH	NAPH
chlorocarbons	CH3CL	SLOWROC	SLOWROC
	CH2CL2	SLOWROC	SLOWROC
	CHCL3	SLOWROC	SLOWROC
	CH3CCL3	SLOWROC	SLOWROC
	TCE	SLOWROC	SLOWROC
	TRICLETH	OLI	OLI
	CDICLETH	OLI	OLI
	TDICLETH	OLI	OLI
	CH2CLCH2CL	SLOWROC	SLOWROC

group	MCM	CRACMM1	CRACMM2
	CCL2CH2	OLT	OLT
	CL12PROP	HC3	HC3
	CHCL2CH3	SLOWROC	SLOWROC
	CH3CH2CL	HC3	HC3
	CHCL2CHCL2	SLOWROC	SLOWROC
	CH2CLCHCL2	SLOWROC	SLOWROC
	VINCL	OLT	OLT
isoprene	C5H8	ISO	ISO
esters	CH3OCHO	SLOWROC	SLOWROC
	METHACET	SLOWROC	SLOWROC
	ETHACET	HC3	HC3
	NPROACET	HC3	HC3
	IPROACET	HC3	HC3
	NBUTACET	HC5	HC5
	SBUTACET	HC5	HC5
	TBUACET	HC3	HC3
ethers	CH3OCH3	HC3	HC3
	DIETETHER	HC10	HC10
	MTBE	HC3	HC3
	DIIPRETHER	HC10	HC10
	ETBE	HC10	HC10
	MO2EOL	ROH	ROH
	EOX2EOL	ROH	ROH
	PR2OHMOX	ROH	ROH
	BUOX2ETOH	ROH	ROH
	BOX2PROL	ROH	ROH
ketones	CH3COCH3	ACT	ACT
	MEK	MEK	MEK
	MPRK	KET	KET
	DIEK	KET	KET
	MIPK	KET	KET
	HEX2ONE	KET	KET
	HEX3ONE	KET	KET
	MIBK	KET	KET
	MTBK	KET	KET
	CYHEXONE	KET	KET
	MVK	MVK	MVK
monoterpenes	APINENE	API	API
	BPINENE	API	API
	LIMONENE	LIM	LIM
sesquiterpenes	BCARY	SESQ	SESQ
organic acids	HCOOH	ORA1	ORA1

group	MCM	CRACMM1	CRACMM2
	CH3CO2H	ORA2	ORA2
	PROPACID	ORA2	ORA2
unclassified	DMM	HC5	HC5
	DMC	SLOWROC	SLOWROC
	DMS	HC5	HC5
	ETHOX	SLOWROC	SLOWROC
phenols	CRESOL	CSL	CSL
	PHENOL	PHEN	PHEN
furans	PXYFUONE	FURAN	FURAN
glyoxals	GLYOX	GLY	GLY
	MGLYOX	MGLY	MGLY

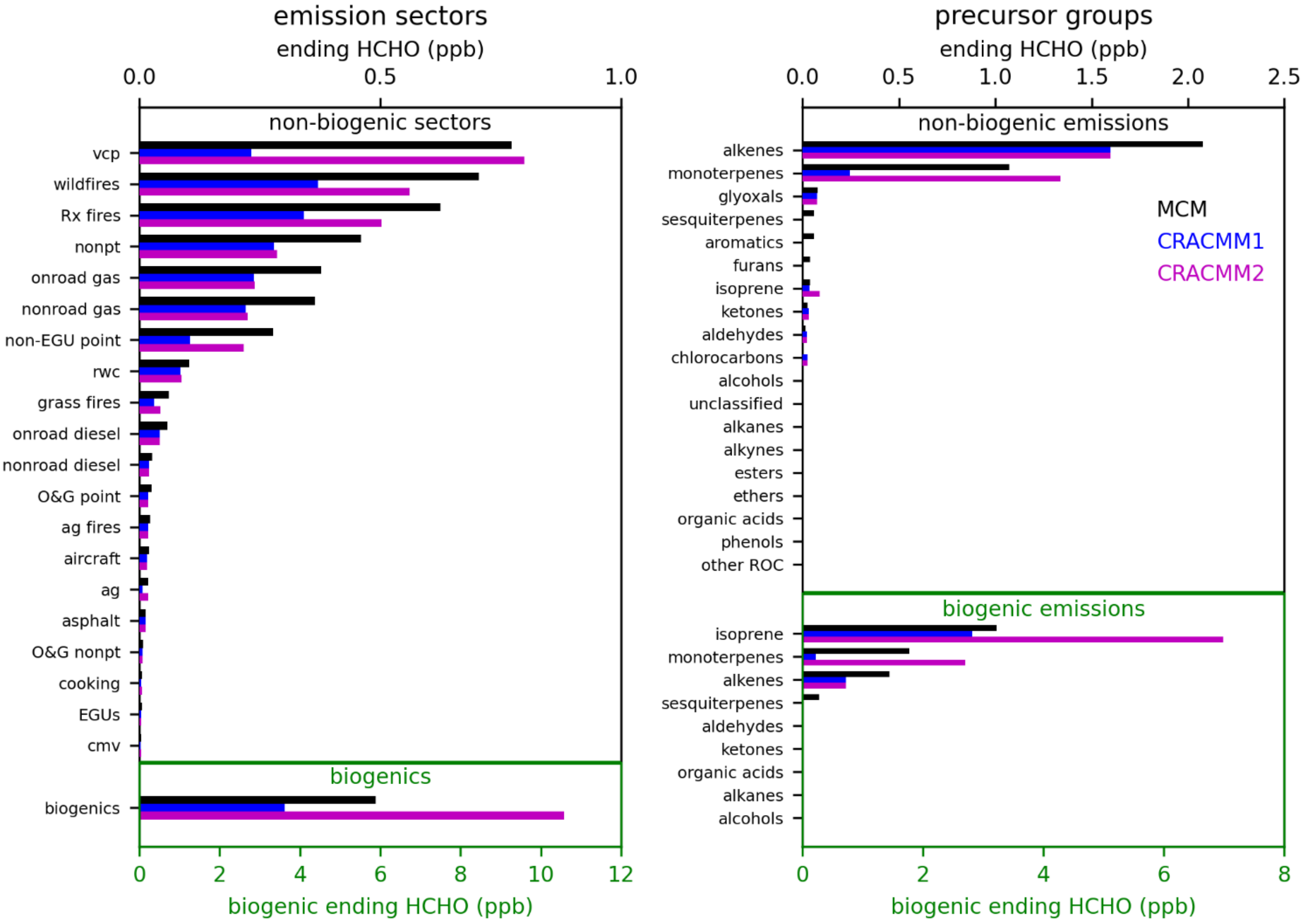


Figure S~~3~~7. Like Figure 1, except that the F0AM simulation holds OH constant at zero and holds ozone constant at 30 ppb. Several categories of precursors have no ozonolysis channel so do not produce any HCHO in these simulations.

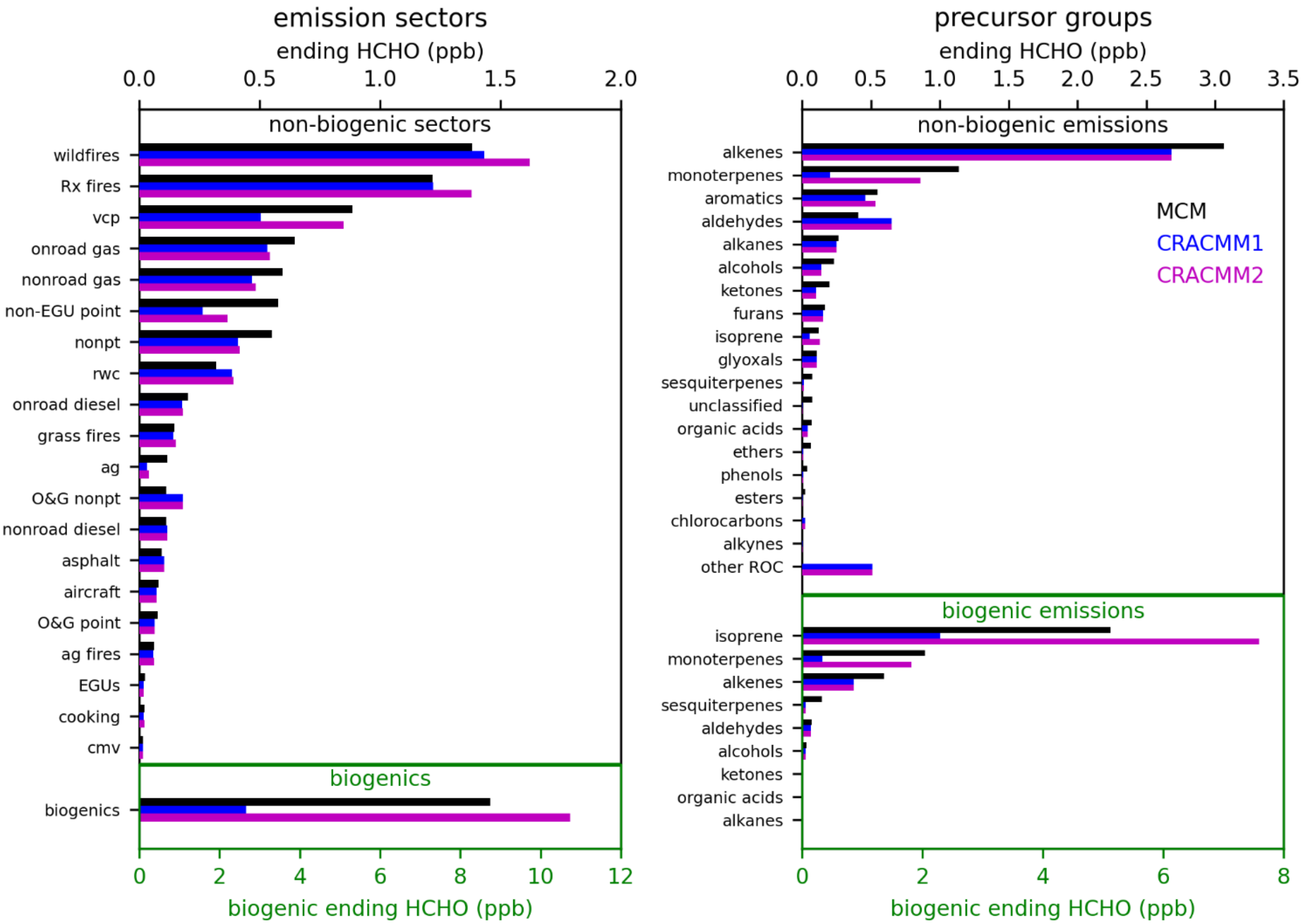


Figure S~~4~~8. Like Figure 1, except that the F0AM simulation holds OH constant at 10^6 molecules cm^{-3} and holds ozone constant at 30 ppb.

Additional CMAQ surface concentration plots

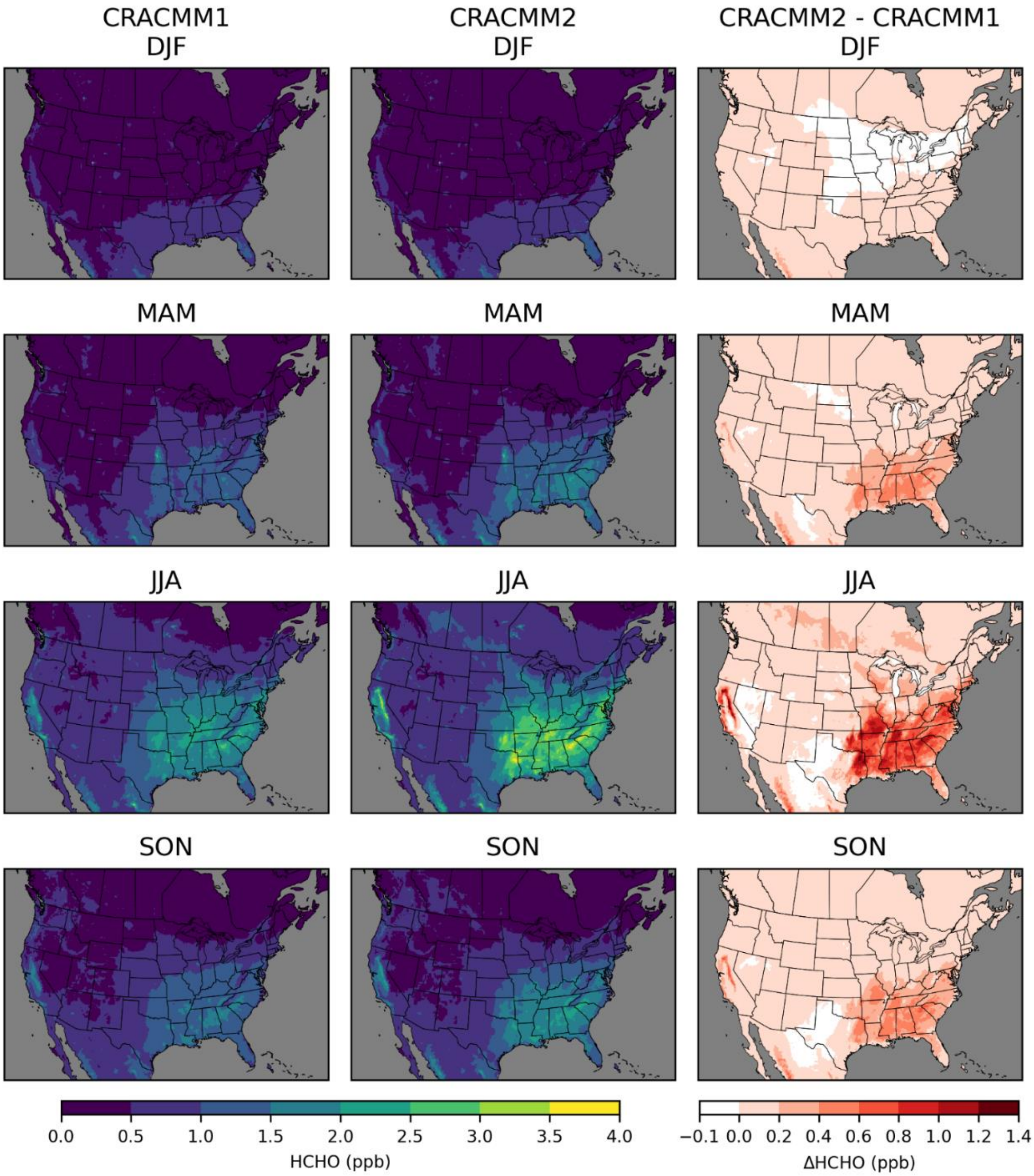


Figure S9. Surface layer 2019 seasonal average HCHO concentrations averaged over 11 am–3 pm local time simulated with CRACMM1 (left column) and CRACMM2 (middle column) and the change in CRACMM2 compared to CRACMM1 (right column).

The effects of adding ECH4 to CRACMM2 are mostly localized in areas with extremely high ECH4 in the emissions inventory (Fig.~~ure~~ S10). The largest impacts are seen in northeastern Wyoming and around the border between New Mexico and Colorado. Impacts are also seen in Alberta, Canada, where there is substantial oil and gas activity. Small impacts from ECH4 can also be seen in the Central Valley of California and in southern California. ECH4 is small compared to the fixed CMAQ global background methane concentration of 1850 ppb, so we do not expect to see significant effects on HCHO from ECH4 other than from extremely large sources. ECH4 as it is represented in the emissions inventory also does not provide a full accounting of methane emissions. Not all sources of methane are inventoried in the NEI (e.g., waterbodies). Effects of emitted methane will therefore not be captured for the sources that are not traditionally included in the emissions inventory. Impacts on HCHO from adding heterogeneous uptake of HO_2 and nitrate are small and affect mostly the southeastern US. ~~The effect is primarily through the reduction of HOx due to HO_2 uptake which decreases the oxidation of isoprene thus reducing secondary HCHO.~~ There are two likely contributing factors. One is decreased HO_X from uptake of HO_2 marginally increasing the lifetime of HCHO. The other is a decrease in the favorability of the RO_2+HO_2 channel with reduced HO_2 and resulting increase in the favorability of the RO_2+NO channel which has higher HCHO yields compared to the RO_2+HO_2 route. The addition of new heterogeneous uptake pathways was not aimed at improving HCHO but were opportunistic updates targeted for implementation in CRACMM2.

The impacts on HCHO from the updates to aromatic chemistry are small and extremely localized compared to some of the other updates. These effects come from the newly added explicit styrene species (STY) in CRACMM2. Most styrene in the emissions inventory (~55%) comes from non-EGU point sources which are mostly made up of industrial sources. During the typical peak of photochemistry from ~~noon~~11am-3pm, the largest impacts are seen in the Los Angeles, California, area and along the border with Mexico near Tijuana. Impacts can also be seen around the Great Lakes region, particularly near Chicago, Illinois, and near the Puget Sound in Washington. More impacted areas can be seen if we look instead at the change in HCHO over all hours rather than focusing only on the peak of photochemistry. The other updates to aromatics do not change any chemistry. They only involve changes in how emissions are mapped, and these changes have little impact on HCHO. Although the effects on HCHO from the aromatic chemistry updates are relatively small, there is additional value in the addition of two new HAPs, ethylbenzene (EBZ) and styrene (STY), which are now represented explicitly in CRACMM2.

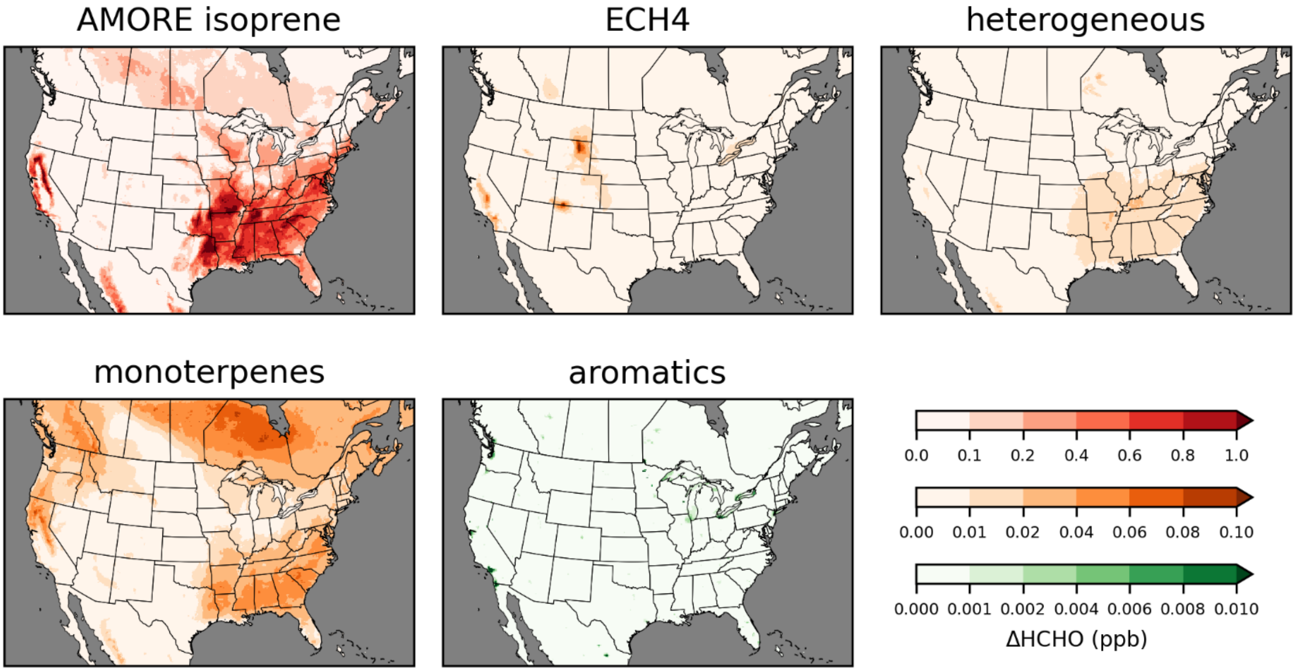


Figure S10. Incremental impacts on surface layer 11 am–3 pm local time 2019 June–August average HCHO concentration resulting from chemistry updates. Color bar scales for incremental impacts differ depending on the magnitude of the impacts. The red color scale is used for the AMORE isoprene impacts; the orange color scale is used for the ECH4, heterogeneous uptake, and monoterpene impacts; the green color scale is used for aromatic impacts.

Additional TROPOMI results

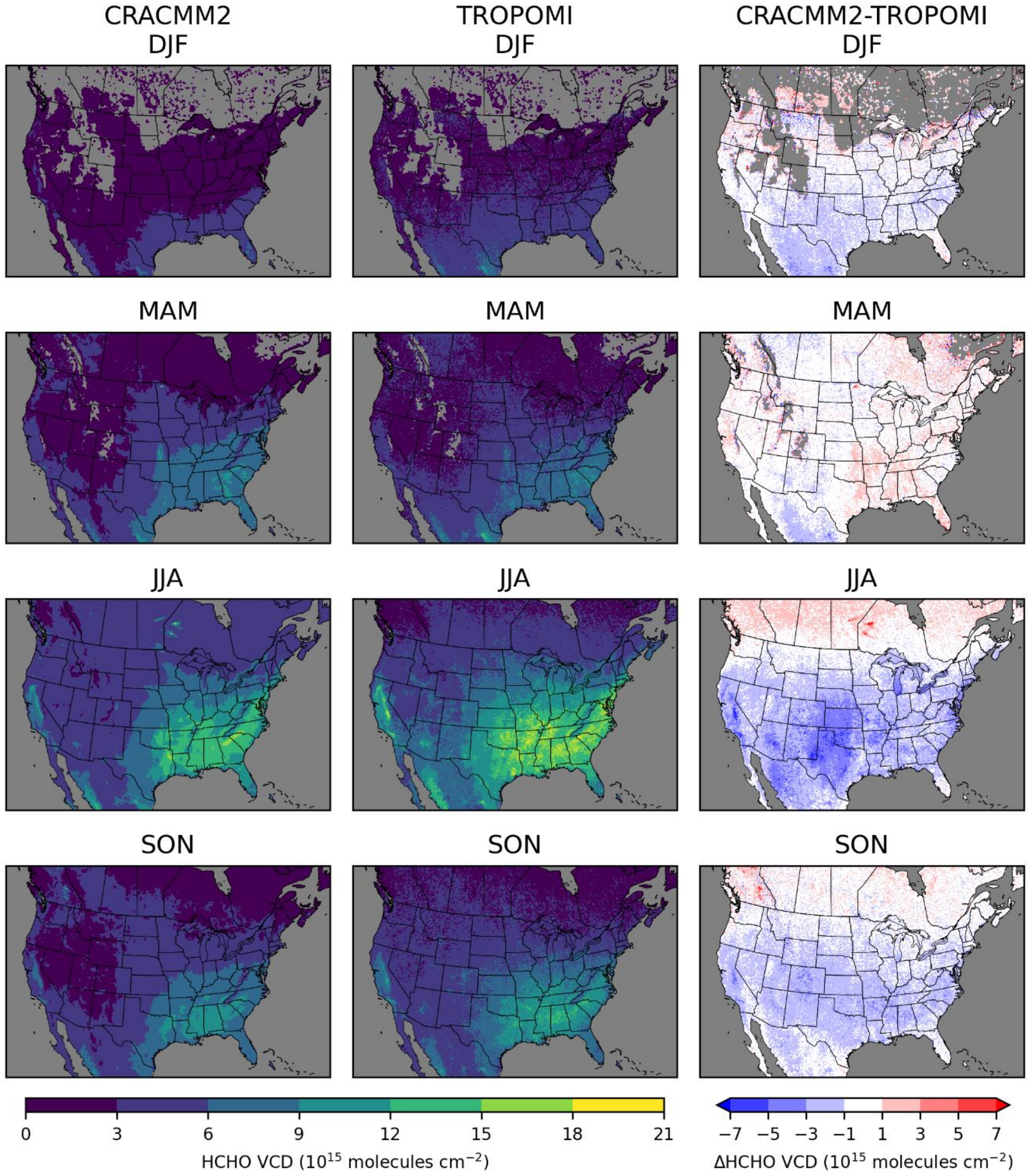


Figure S11. 2019 seasonal average tropospheric vertical column densities from CMAQ with CRACMM2 (left) and from TROPOMI (middle), and the difference between CRACMM2 and TROPOMI (right).

We incorporate an uncertainty estimate into the comparison of HCHO VCD from CMAQ with CRACMM2 to TROPOMI using uncertainty estimates provided in the TROPOMI HCHO algorithm theoretical basis document (ATBD) (KNMI, 2022). Table 13 of the ATBD reports uncertainties in monthly averaged columns for low ($<=1\times10^{16}$ molecules cm^{-2}) and elevated ($>1\times10^{16}$ molecules cm^{-2}) columns as 50% and 25%, respectively, as the total uncertainty neglecting uncertainties in the AMF. We neglect the uncertainties in the AMF since we have incorporated information from the TROPOMI averaging kernel into the CMAQ VCD which is the suggested approach in the ATBD. The uncertainties for low and elevated columns are applied to the June-August 2019 seasonal averages to calculate lower and upper bounds for TROPOMI HCHO VCD which are compared to CRACMM2 along with the original (i.e., without uncertainty estimate) TROPOMI HCHO VCD (Fig.~~ure~~ S12). The HCHO VCD in CMAQ with CRACMM2 is mostly within the limits of uncertainty of the TROPOMI HCHO VCD. Notable exceptions to this are in the Permian Basin; near Phoenix, Arizona; Los Angeles, California; and other parts of California east of the Central Valley.

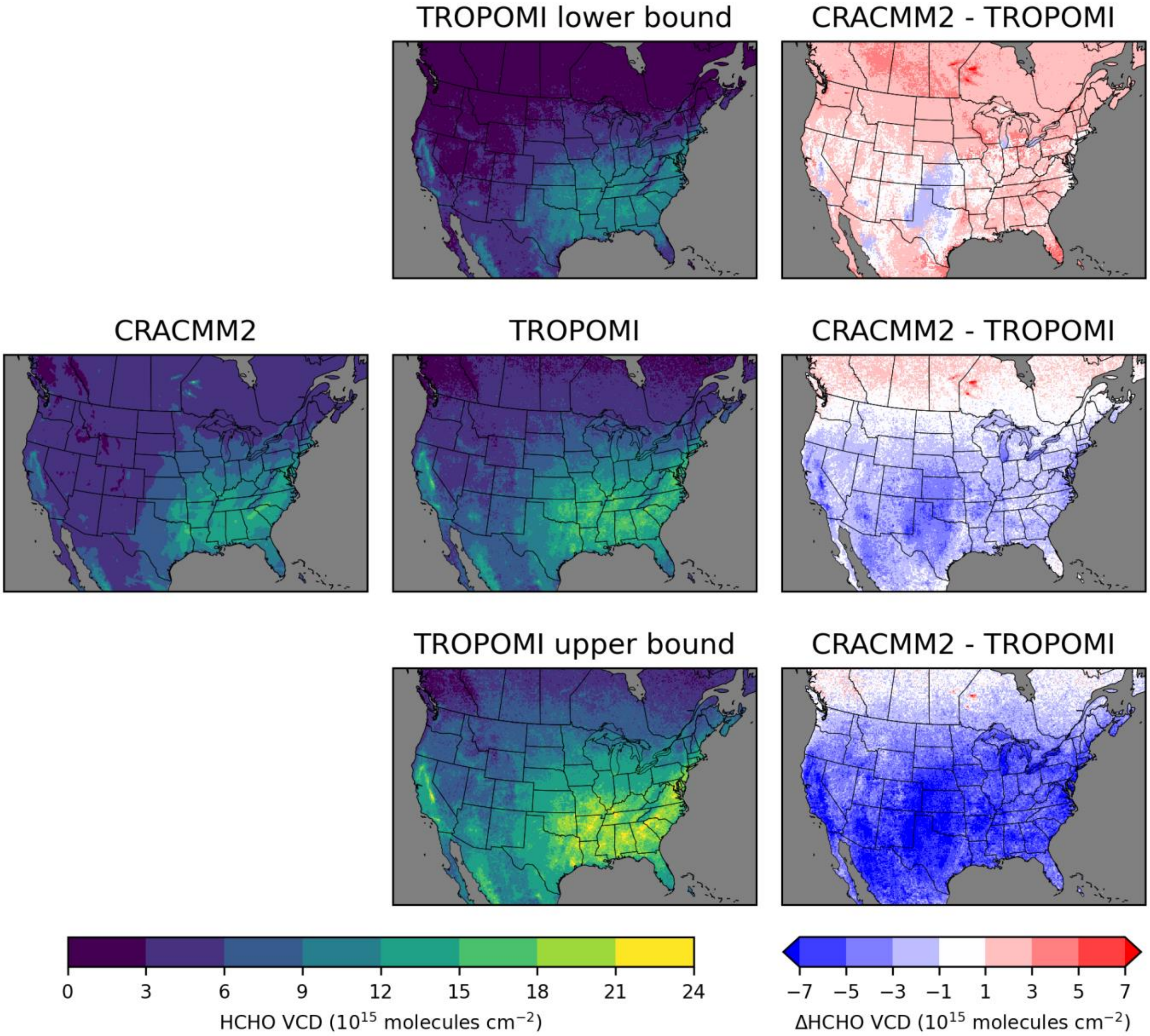


Figure S12. 2019 June-August average tropospheric vertical column densities from CMAQ with CRACMM2 (left column), TROPOMI with CRACMM2-based AMF (middle column), and the difference between CRACMM2 and TROPOMI (right column). The top row uses the lower bound for TROPOMI based on the uncertainty analysis. The middle row does not incorporate uncertainty. The bottom row uses the upper bound for TROPOMI based on the uncertainty analysis.

275 **Oil and gas emission sensitivity simulation results**

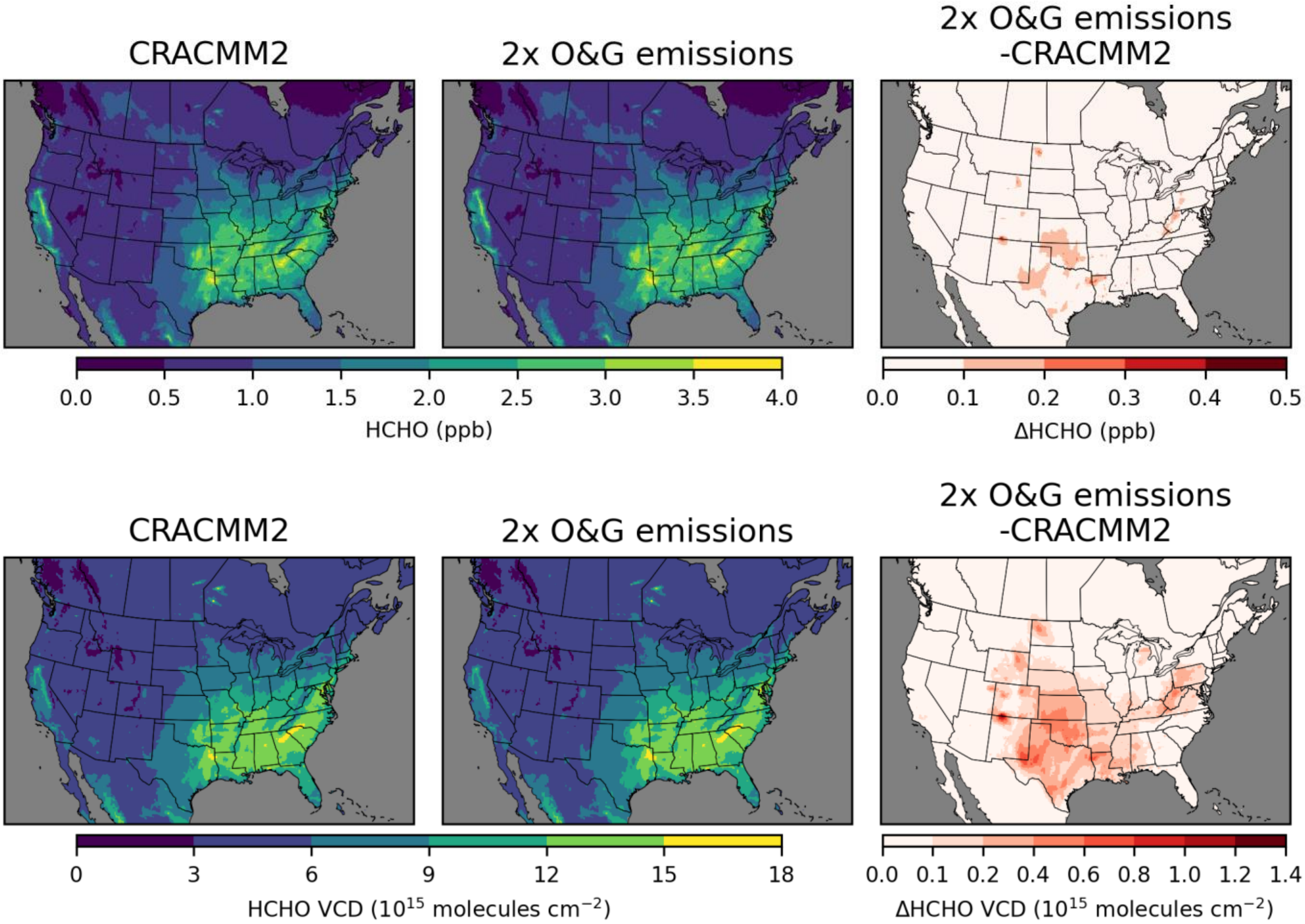


Figure S13. Average June-August 2019 change in HCHO in CMAQ at the surface from 11am-3pm (top row) and in the vertical column at the TROPOMI overpass time (~1:30pm local solar time) from doubling oil and gas NO_X and ROC emissions.

280

HCHO AQS evaluation

HCHO observations are available from the Air Quality System (AQS) database which contains data from federal, state, local, and tribal air quality monitoring stations. Routine HCHO measurements are primarily taken at sites in the Photochemical Assessment Monitoring Stations (PAMS) and National Air Toxics Trends Sites (NATTS) networks. HCHO is measured using method TO-11A in which HCHO is collected on 2,4-dinitrophenylhydrazine (DNPH) coated cartridges from which HCHO derivative products are measured offline using high performance liquid chromatography (HPLC). HCHO data shown here are obtained from the AQS website (https://aqs.epa.gov/aqsweb/airdata/download_files.html) and have not been corrected for field blanks; however, the values of field blanks are expected to be small enough that any qualitative conclusions are unaffected. PAMS and NATTS sites typically have sample collection schedules of three 8-h samples or one 24-h sample which limits the usefulness for evaluation of the diurnal variability of HCHO. The DNPH measurement technique has also recently been found to measure lower HCHO values compared to other instruments in Atlanta, GA, (Mouat et al., 2024) and in Salt Lake City, UT, (Jaffe et al., 2024). While there is some uncertainty in these measurements, we report a comparison between the daily average HCHO observations available in AQS and daily averages from CMAQ using CRACMM1 and CRACMM2 (Figure S14). A total of 5424 daily average observations are available from a total of 109 sites in 2019. HCHO is underestimated, but the bias is improved by about 15%. There are, however, a limited number of sites in the southeastern US (Figure S15) where HCHO in CRACMM2 increased the most which limits the usefulness of the comparisons to AQS data in evaluating the CRACMM2 updates.

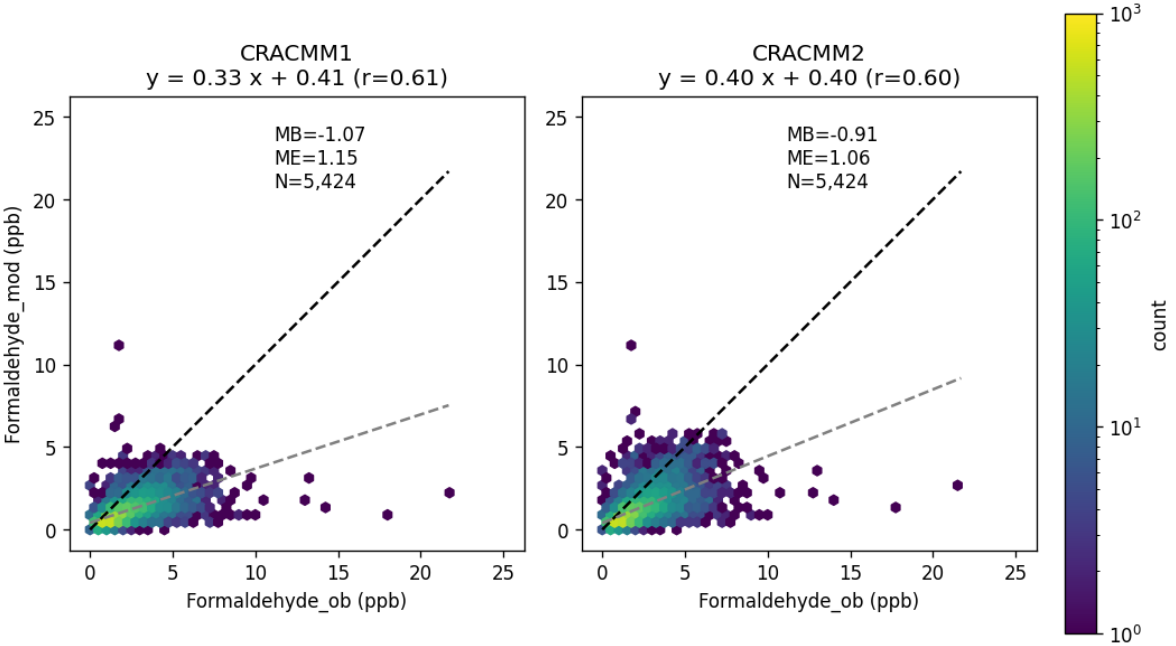


Figure S14. Density scatter plots of daily average HCHO observations from AQS compared to simulated daily averages from CMAQ using CRACMM1 (left) and CRACMM2 (right).

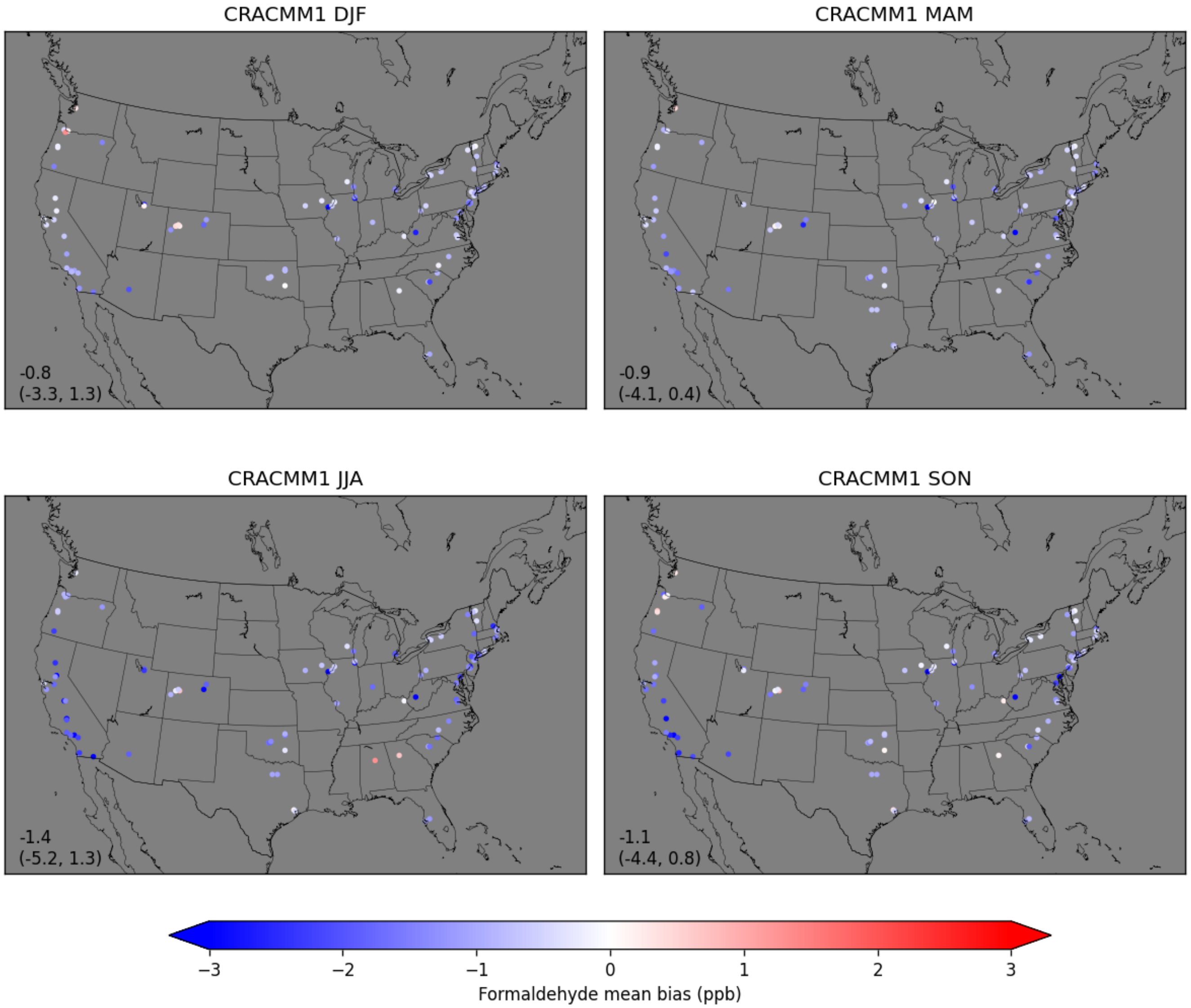


Figure S15. Seasonal average bias in daily average HCHO with CRACMM1 compared to observed HCHO from AQS sites. Annotations in the lower left of each panel show the mean (min, max) of seasonal average biases over all sites.

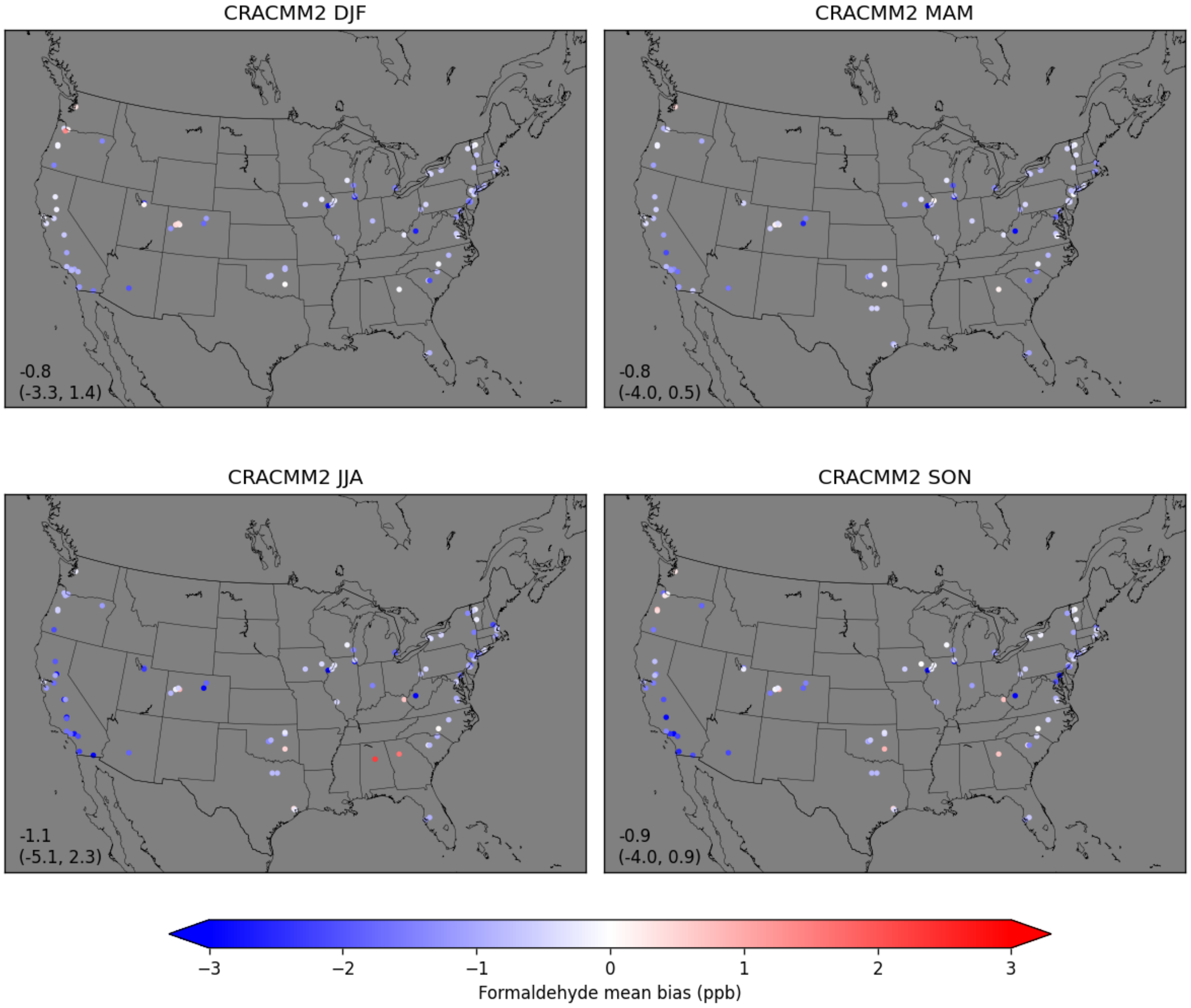


Figure S16. Seasonal average bias in daily average HCHO with CRACMM2 compared to observed HCHO from AQS sites. Annotations in the lower left of each panel show the mean (min, max) of seasonal average biases over all sites.

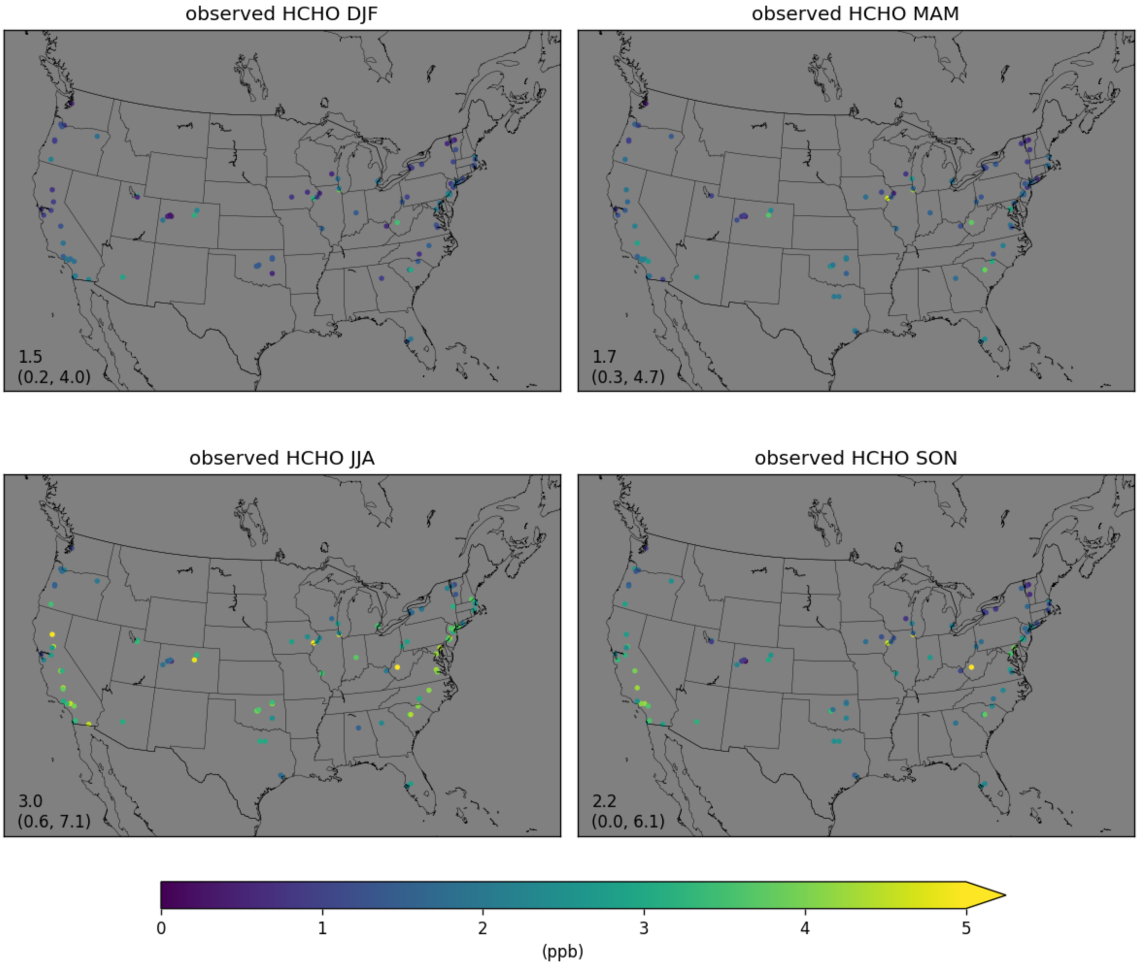


Figure S17. Seasonal average observed daily average HCHO from AQS sites. Annotations in the lower left of each panel show the mean (min, max) of seasonal average observations over all sites.

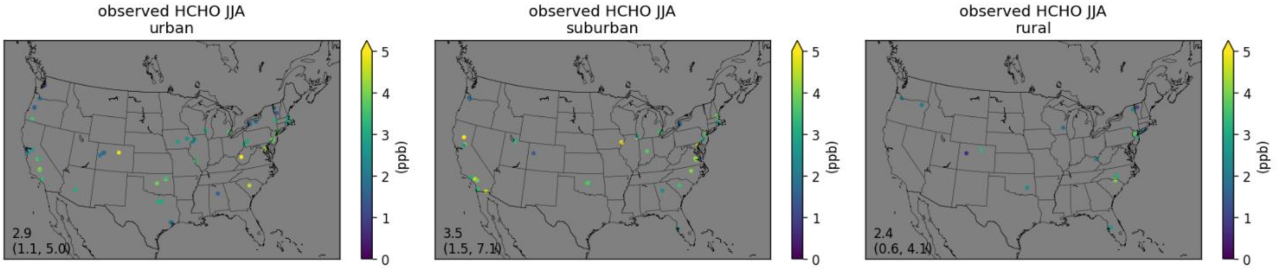


Figure S18. June-August 2019 average observed daily average HCHO from AQS sites, separated by their location in urban, suburban, and rural areas. Annotations in the lower left of each panel show the mean (min, max) of seasonal average observations over all sites.

Hourly observations

Table S4. Hourly HCHO measurement locations, sampling dates, and instrumentation.

Location	Sampling dates	Instrument	Uncertainty	Additional notes
Westport, CT	2023 May 22 – Nov 14	Aerodyne TILDAS	Maximum of 0.6 ppb or 15%	Collected in support of 2023 Synergistic TEMPO Air Quality Science (STAQS). Available from https://www-air.larc.nasa.gov/cgi-bin/ArcView/listos.2023?GROUND-WESTPORT=1
	2018 Jun 21 – Sep 4	Aerodyne TILDAS	10%	Collected during the 2018 Long Island Sound Tropospheric Ozone Study (LISTOS). Available from https://www-air.larc.nasa.gov/cgi-bin/ArcView/listos?GROUND-WESTPORT=1
Flax Pond, NY	2023 May 22 – Nov 16	Aerodyne TILDAS	Maximum of 0.6 ppb or 15%	Collected in support of 2023 STAQS. Available from https://www-air.larc.nasa.gov/cgi-bin/ArcView/listos.2023?GROUND-FLAX-POND=1
New Brunswick, NJ	2023 May 22 – Nov 8	Picarro G2307	Maximum of 0.6 ppb or 15%	Collected in support of 2023 STAQS. Available from https://www-air.larc.nasa.gov/cgi-bin/ArcView/listos.2023?GROUND-RUTGERS=1
Sheboygan, WI	2017 Jun 26 – Jul 22	Aerodyne TILDAS	10%	Collected during the 2017 Lake Michigan Ozone Study (LMOS). Available from https://www-air.larc.nasa.gov/cgi-bin/ArcView/lmos?GROUND-SHEBOYGAN=1
Salt Lake City, UT	2017 Jan 15 – Feb 14	Aerodyne TILDAS	10%	Collected during the 2017 Utah Winter Fine Particle Study (UWFPS). Available from https://csl.noaa.gov/groups/csl7/measurements/2017uwfps/Ground/DataDownload/index.php?page=/groups/csl7/measurements/2017uwfps/Ground/DataDownload/
Research Triangle Park, NC	2016 Aug 16 – Dec 11	Aerodyne TILDAS	10%	Collected near EPA campus
Atlanta, GA (South DeKalb)	2022 Apr 26 – 2023 Oct 31	Picarro G2307	10%	See Mouat et al., 2024
Atlanta, GA (Georgia Tech)	2022 Jul 25 – 2023 Jan 31	MIRA Ultra	14% + 0.3 ppb	See Mouat et al., 2024

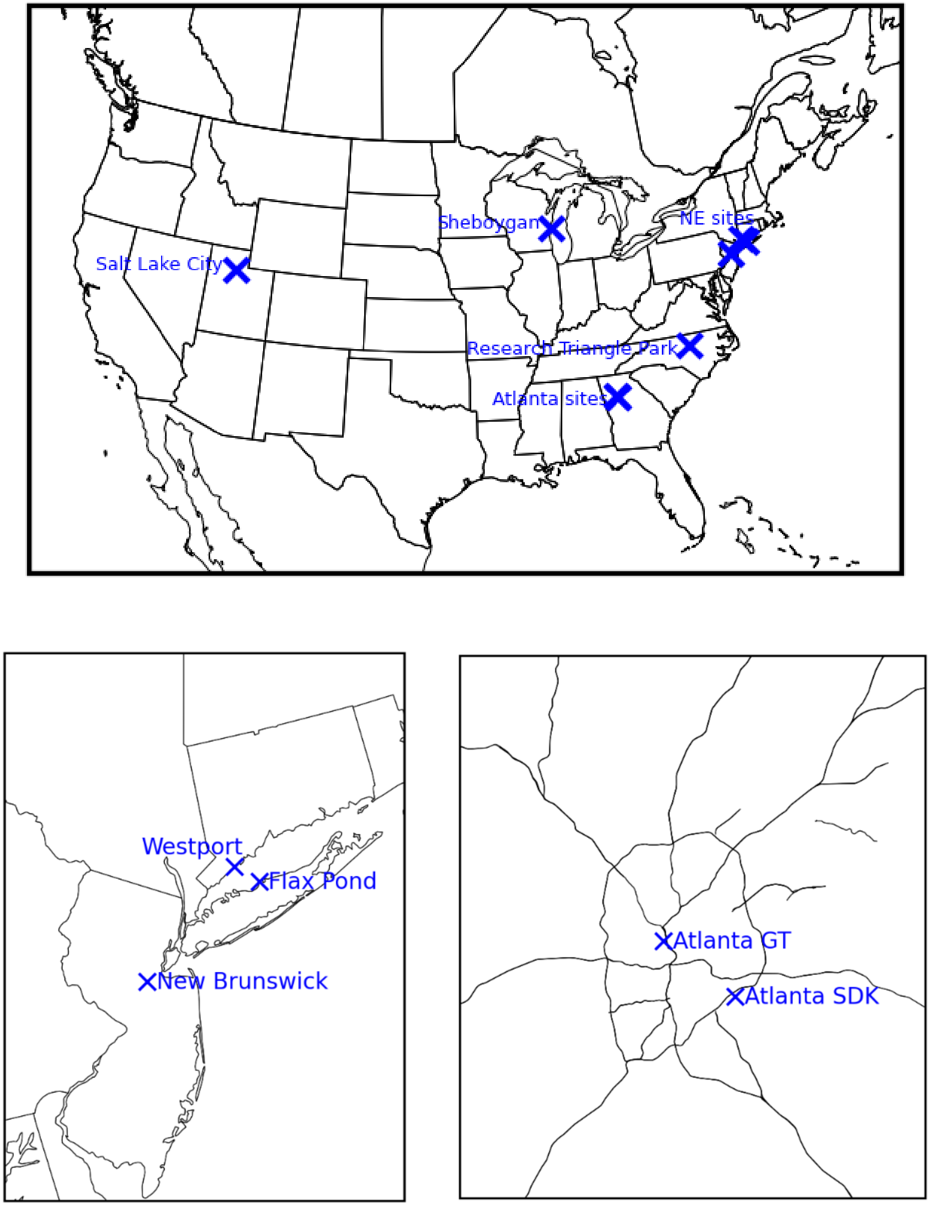


Figure S19. Map of sampling locations listed in Table S4. The two maps in the bottom row show a zoomed in view of the northeastern (NE) US sites and the Atlanta sites. On the map for the Atlanta sites, major roadways are displayed to provide additional geographic context.

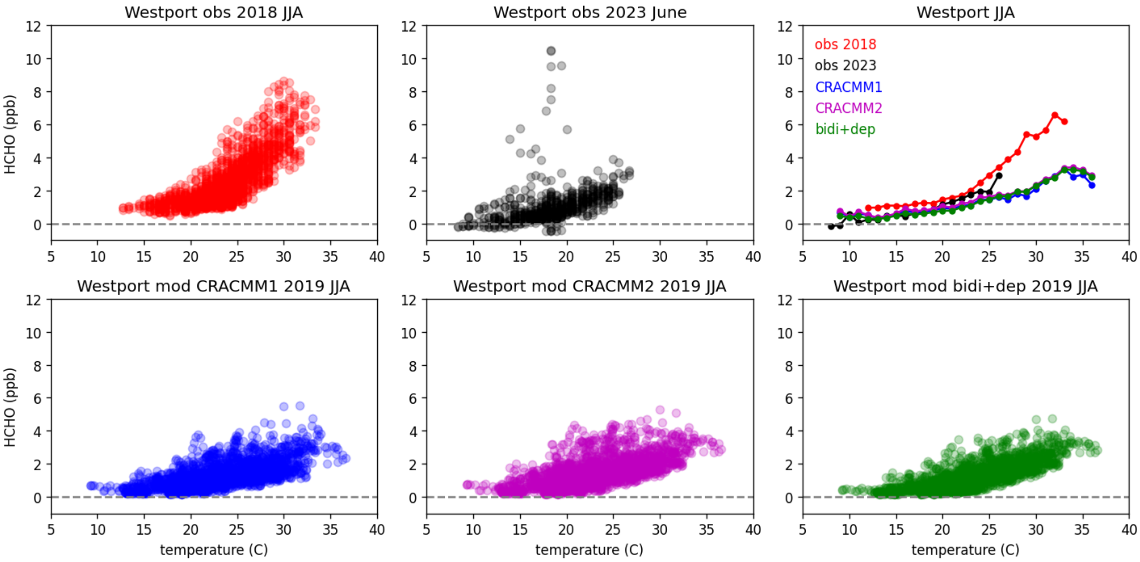


Figure S20. Hourly HCHO concentration by temperature at the Westport, CT, sampling location during summer. All available hourly data points are provided for the 2018 and 2023 observations and for the CRACMM1, CRACMM2, and bidi+dep CMAQ simulations. The upper right panel shows the median temperature in 1 degree C temperature bins for each set of observations and modeled concentrations.

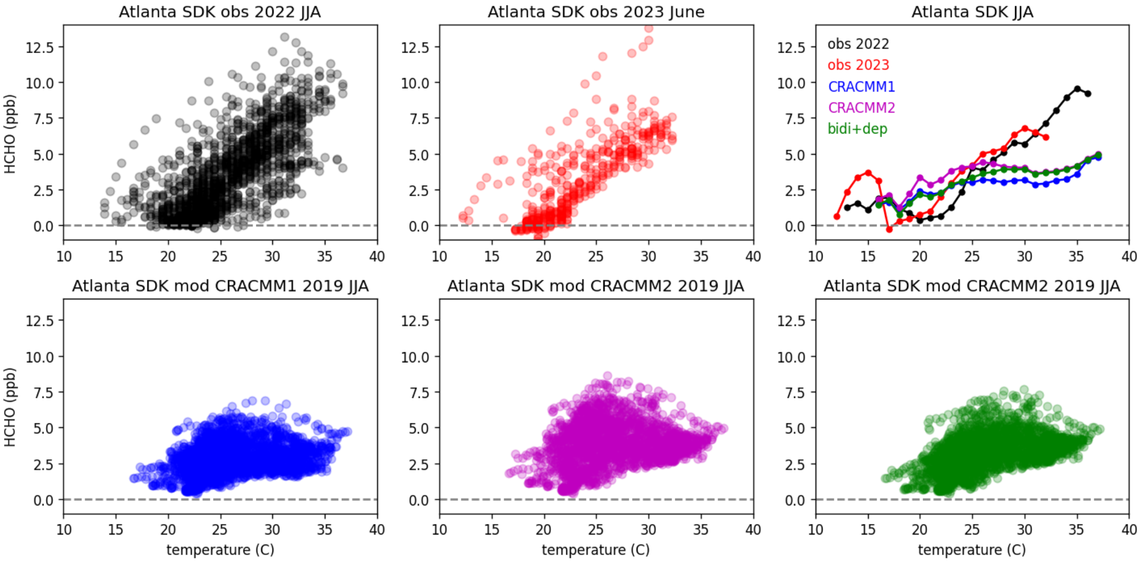


Figure S21. Hourly HCHO concentration by temperature at the Atlanta, GA South DeKalb (SDK) sampling location during summer. All available hourly data points are provided for the 2022 and 2023 observations and for the CRACMM1, CRACMM2, and bidi+dep CMAQ simulations. The upper right panel shows the median temperature in 1 degree C temperature bins for each set of observations and modeled concentrations.

Observations from fall 2016 in Research Triangle Park, NC, show a rapid rise in the early morning leading to a midday peak which is not captured in the CMAQ simulations. The CRACMM2 updates do increase the HCHO at midday (as well as throughout the entire day) which makes the simulated values closer to the observed at midday but results in higher HCHO at night. The diurnal profile of the CMAQ simulations is much flatter than is seen in the observations. The other observations included in Figure S22 are during summer. For the Sheboygan, WI, data the diurnal profile is mostly flat with small late morning and evening peaks. The peak values simulated with CRACMM2 are close to those seen in the observations, despite the shape of the diurnal profile not aligning well with the observations. We note also that this sampling location is on the shore of Lake Michigan which has complex land-sea breeze effects which are not always well simulated, so meteorology could play an important role here. The observations for Westport, CT, during 2018 are similar to the 2023 observations at the same site with HCHO beginning to rise in the early morning, peaking around noon, and then falling as the afternoon progresses and into night. The major difference between the 2018 and 2023 data is the higher midday peak in the 2018 observations. The CMAQ data shown for Westport, CT, below are identical to what is shown in Figure 6 for the Westport 2023 data.

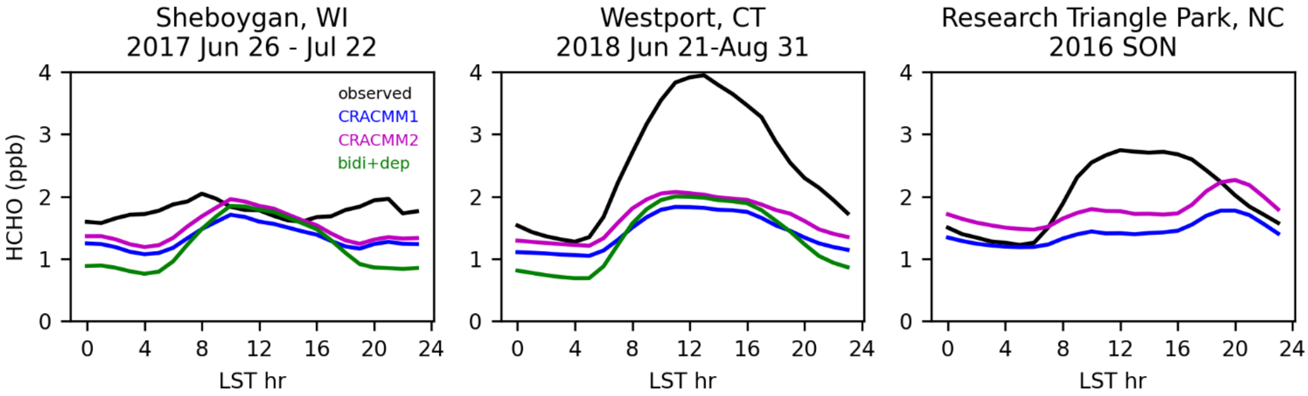


Figure S22. Diurnal profiles of observations in several years at several sites compared to CMAQ simulations in 2019 using CRACMM1, CRACMM2, and CRACMM2 with updated HCHO bidirectional flux and deposition (bidi+dep). The bidi+dep results are not shown for the Research Triangle Park, NC, site since this was a sensitivity simulation conducted for summer only. Sampling locations and dates are provided above each panel.

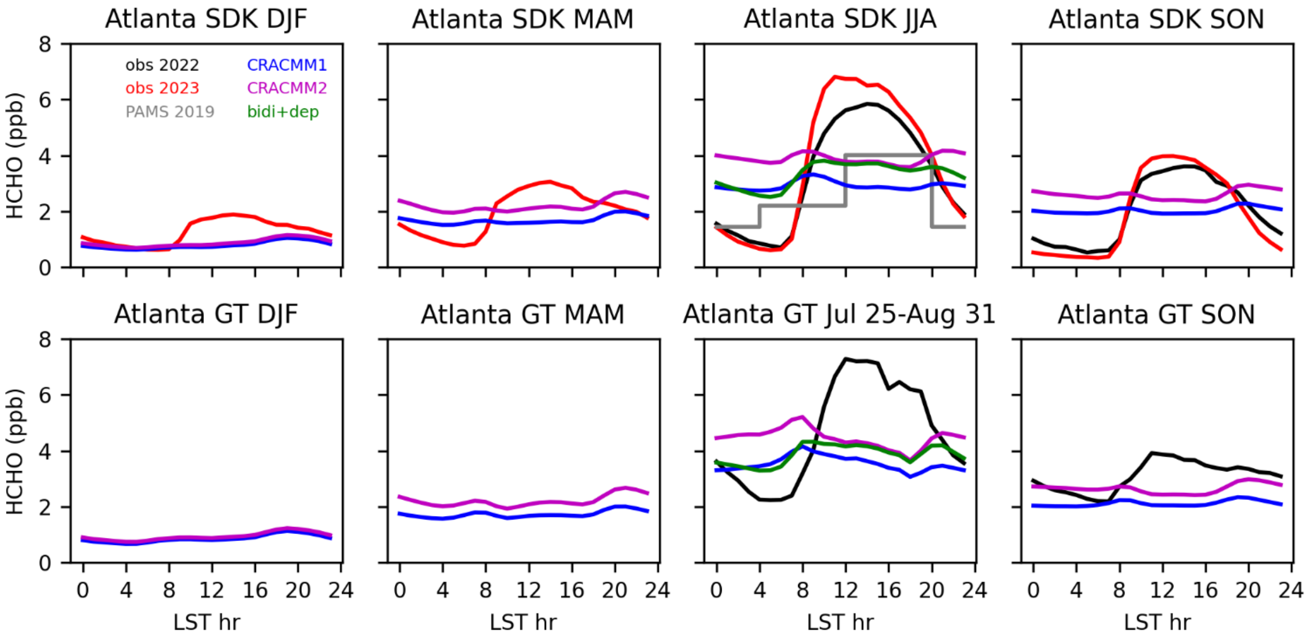


Figure S23. Seasonal average diurnal variability at South DeKalb (SDK) and Georgia Tech (GT) sites.

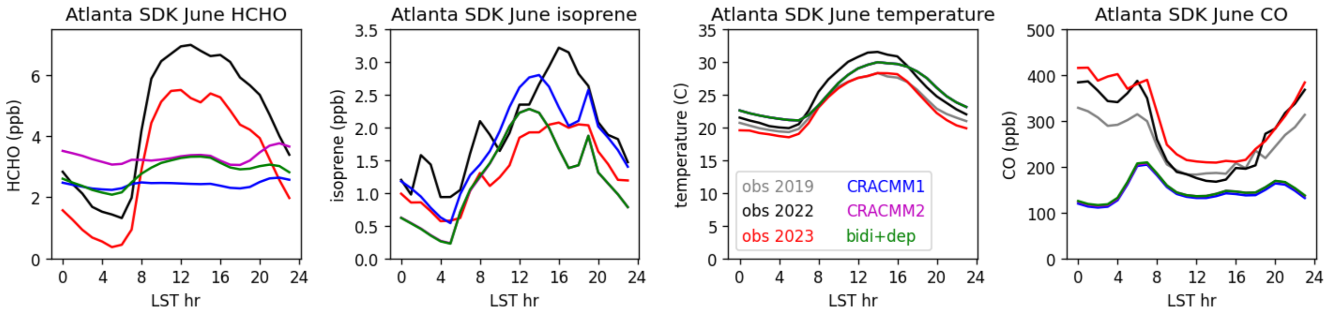
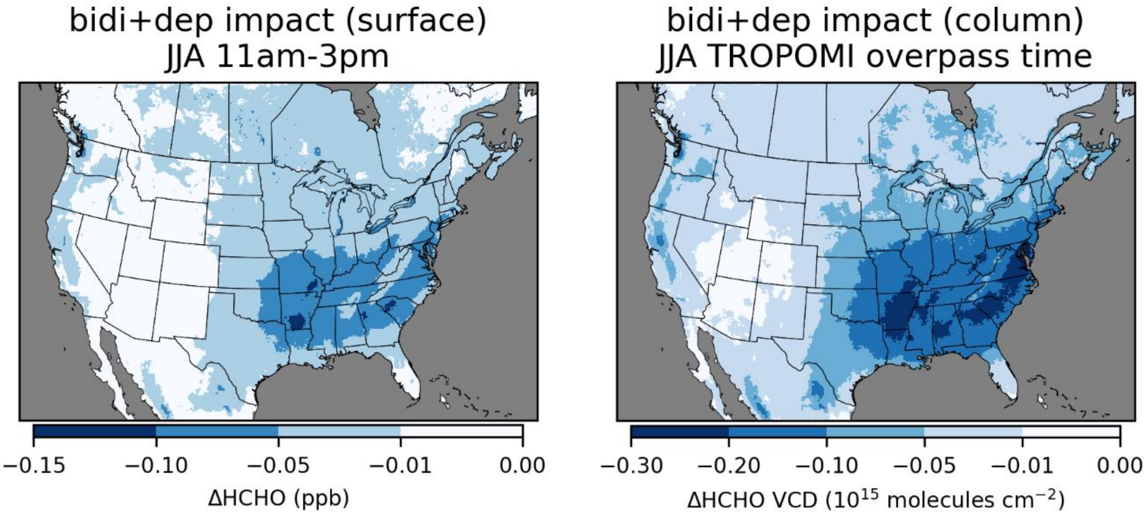


Figure S24. Hourly HCHO, isoprene, temperature, and CO during June for various years at Atlanta SDK PAMS site.

Effects of HCHO bidirectional flux and deposition updates



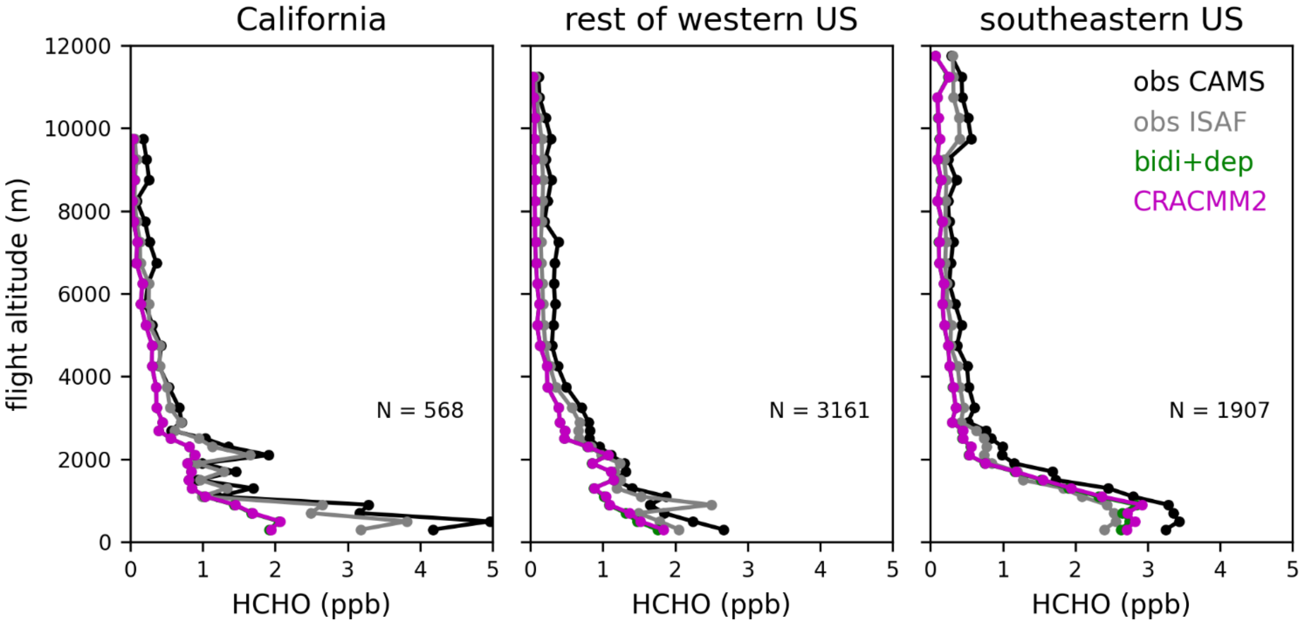


Figure S25. Effects of bidirectional flux and deposition updates on HCHO for 2019 June-August 11am-3pm surface HCHO (top left), 2019 June-August average HCHO VCD at the TROPOMI overpass time (top right), and on the HCHO vertical profile comparison to FIREX (bottom).

Seasonal average HCHO controllable fractions

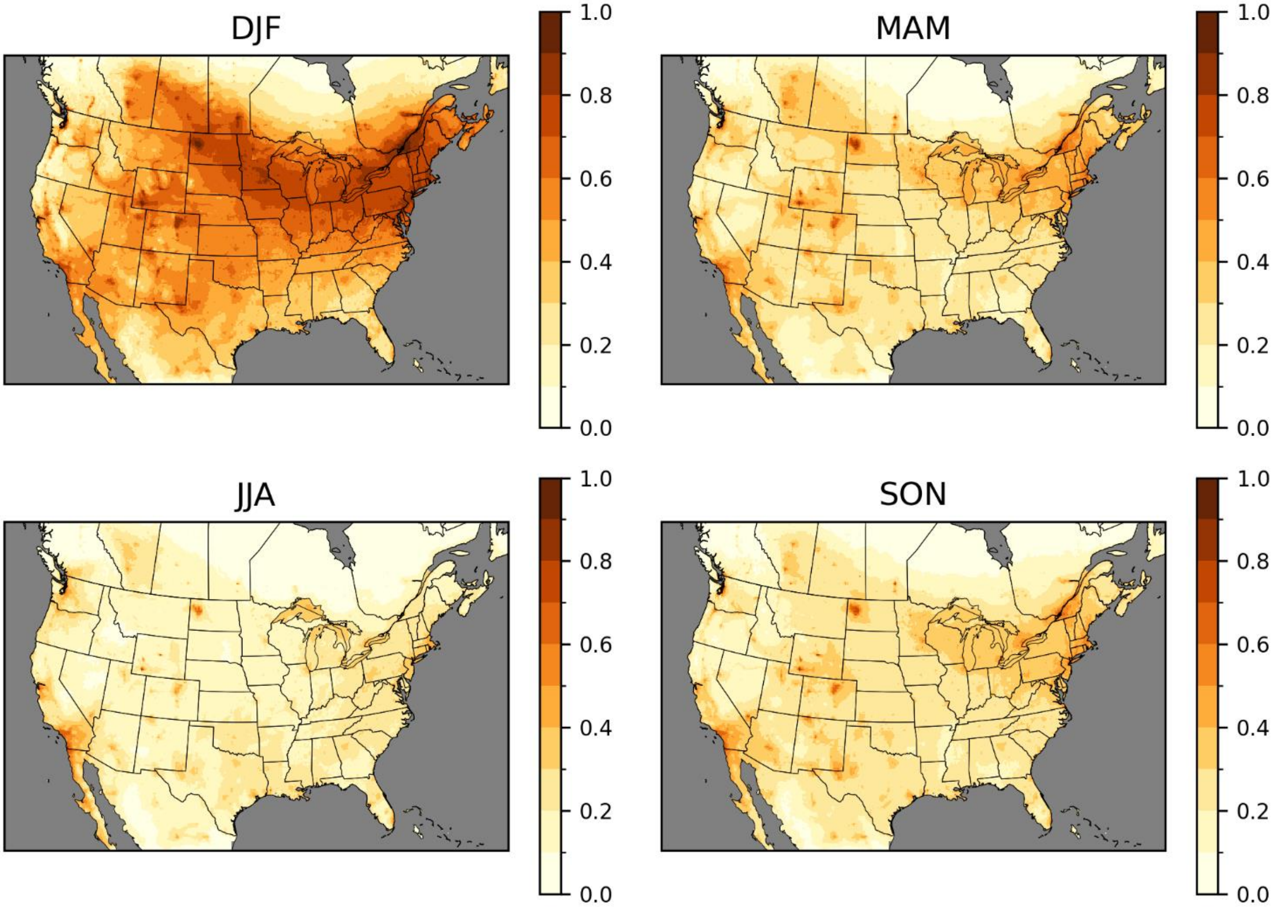


Figure S26. Seasonal average controllable fraction of HCHO calculated as the difference between the base simulation and the zero US anthropogenic NOx and ROC simulation divided by the base simulation.

CRACMM2 ozone evaluation

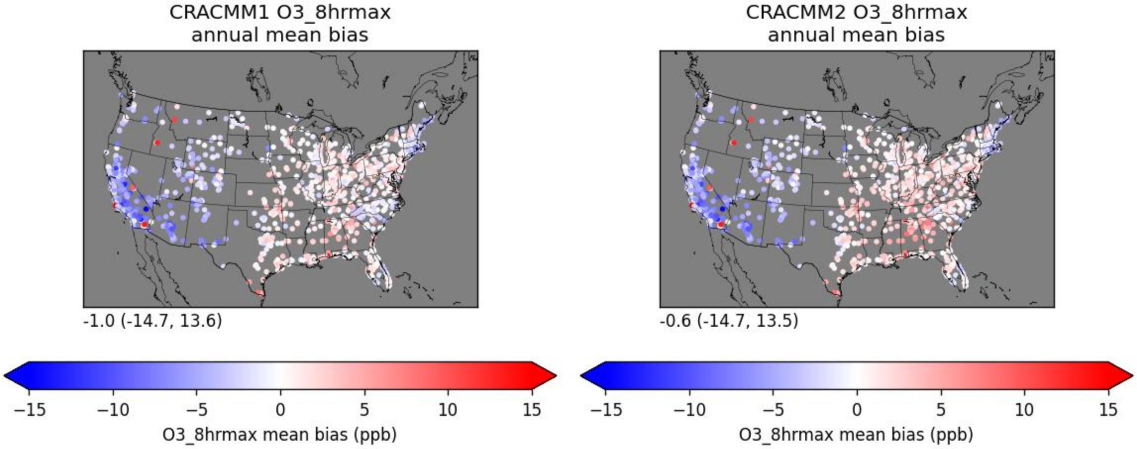


Figure S27. Annual mean bias in MDA8 O_3 for 2019 at AQS monitoring sites for CRACMM1 (left) and CRACMM2 (right). The annotations in the lower left show the mean (min, max) site average biases.

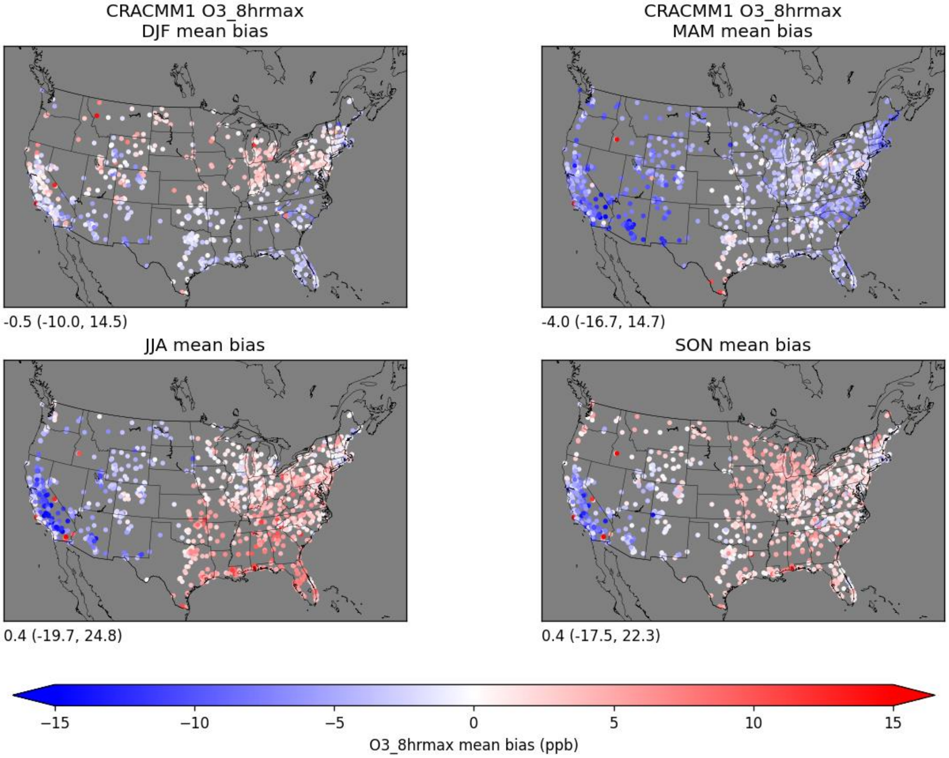


Figure S28. Seasonal mean bias in MDA8 O_3 for 2019 at AQS monitoring sites for CRACMM1. The annotations in the lower left show the mean (min, max) site average biases.

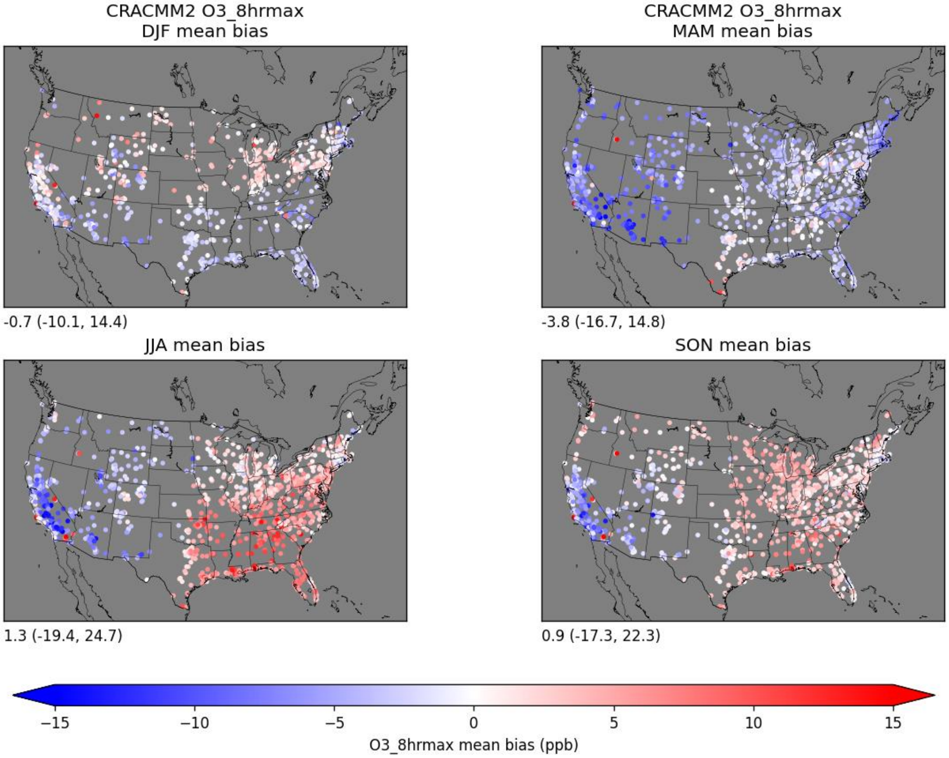


Figure S29. Seasonal mean bias in MDA8 O_3 for 2019 at AQS monitoring sites for CRACMM2. The annotations in the lower left show the mean (min, max) site average biases.

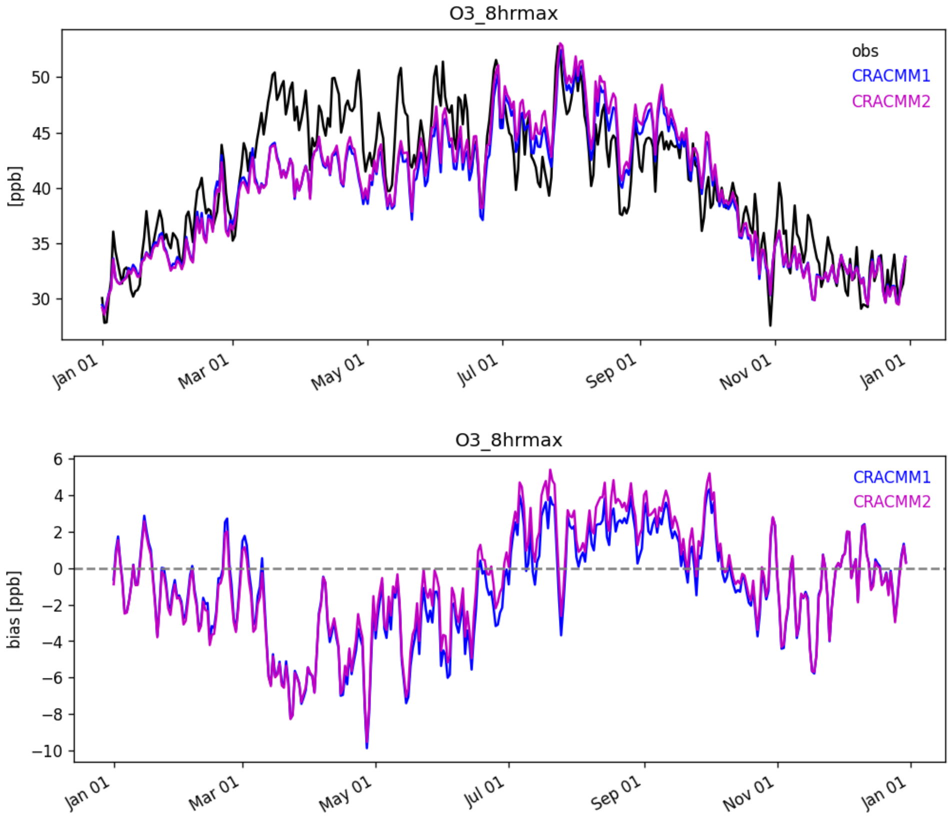


Figure S30. Daily average MDA8 O_3 (top) and bias (bottom) for 2019 averaged over all AQS monitoring sites for CRACMM1 and CRACMM2.

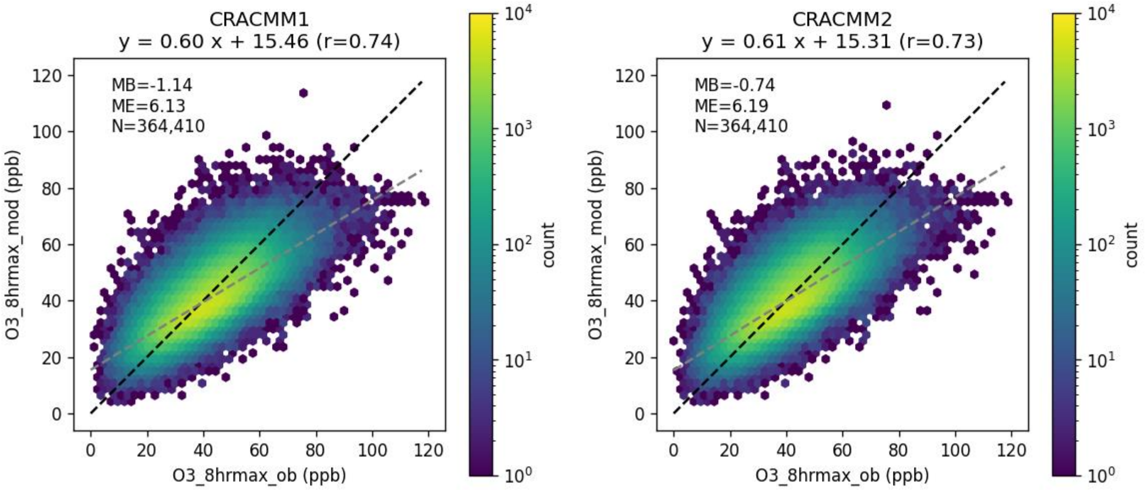


Figure S31. Density scatter plot of observed MDA8 O_3 vs. CRACMM1 (left) and CRACMM2 (right). The mean bias (MB), mean error (ME), and number of observations (N) are indicated on each panel.

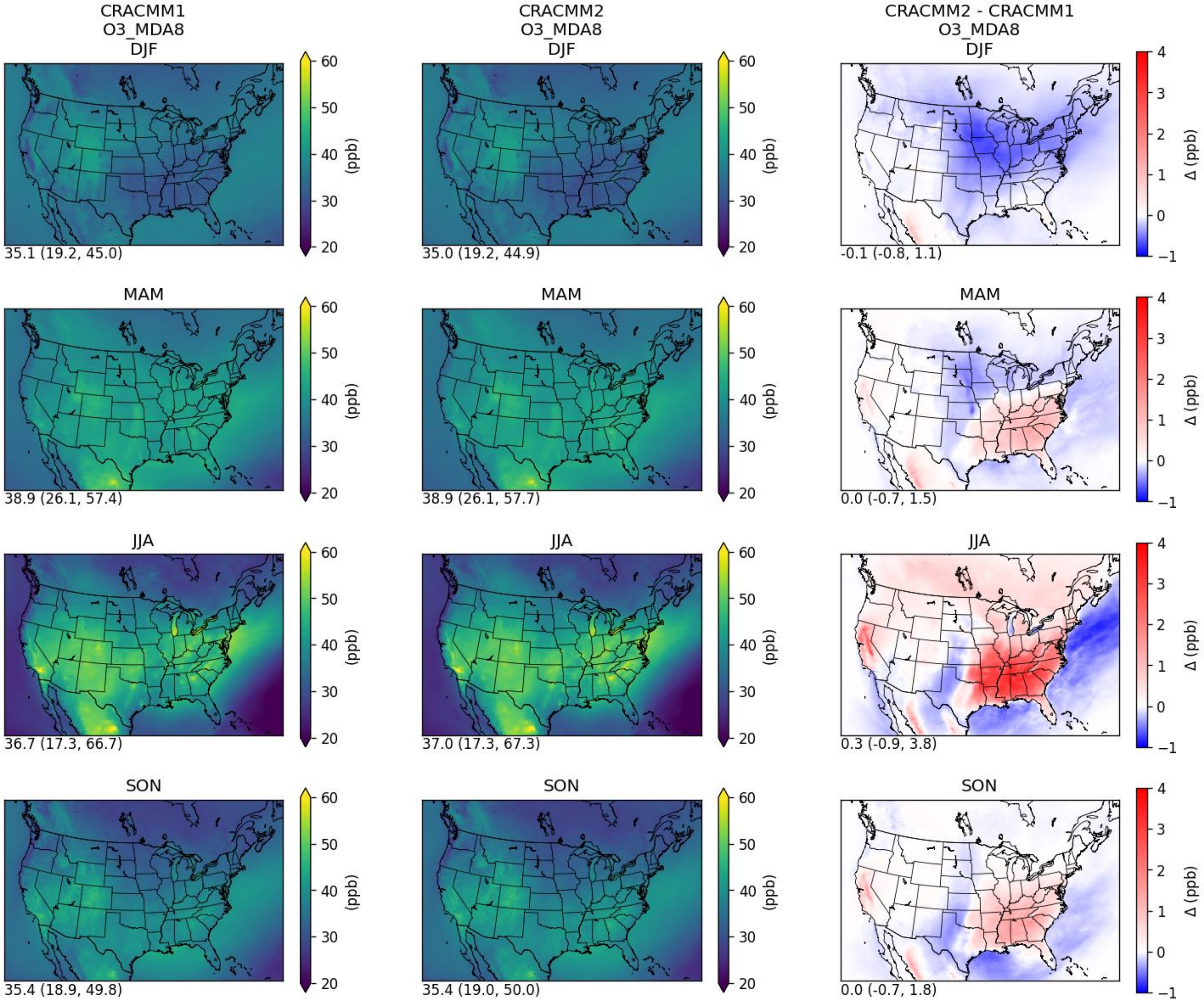


Figure S32. Seasonal average MDA8 O_3 in CRACMM1 (left column) and CRACMM2 (middle column) and the change in CRACMM2 compared to CRACMM1 (right column). The annotations in the lower left show the mean (min, max) over the modeling domain.

CRACMM2 PM2.5 evaluation

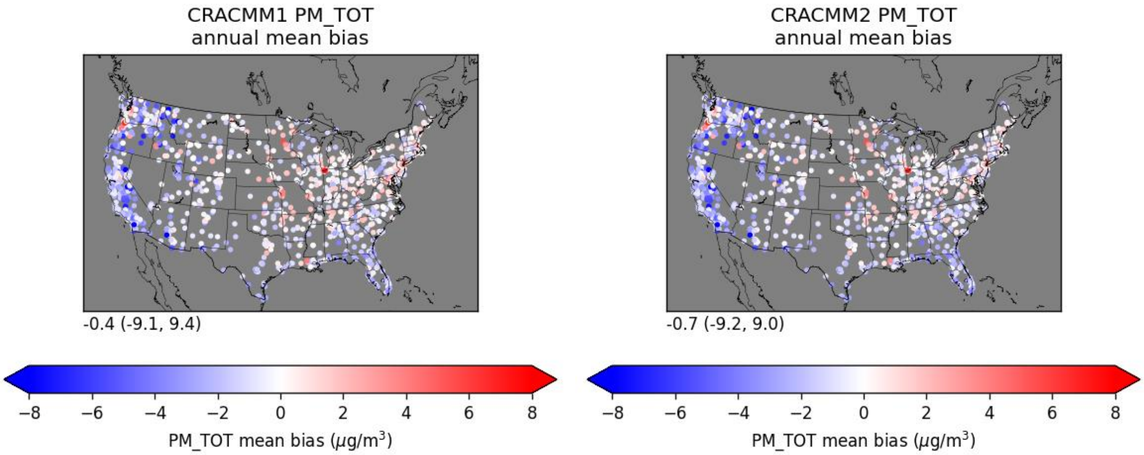


Figure S33. Annual mean bias in daily average PM2.5 for 2019 at AQS monitoring sites for CRACMM1 (left) and CRACMM2 (right). The annotations in the lower left show the mean (min, max) site average biases.

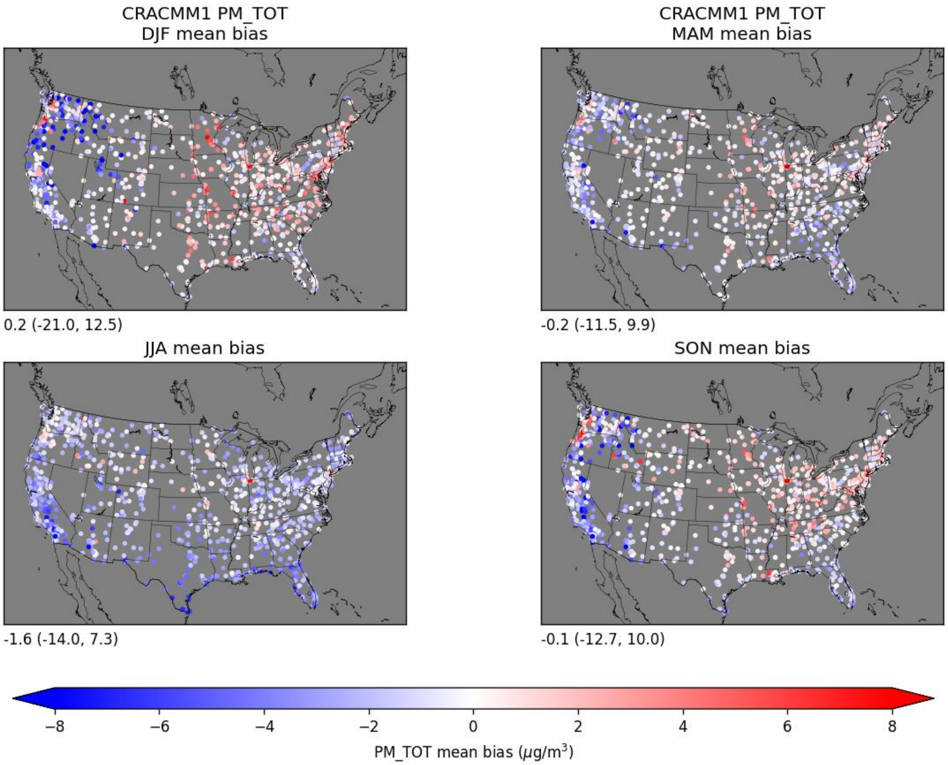


Figure S34. Seasonal mean bias in daily average PM2.5 for 2019 at AQS monitoring sites for CRACMM1. The annotations in the lower left show the mean (min, max) site average biases.

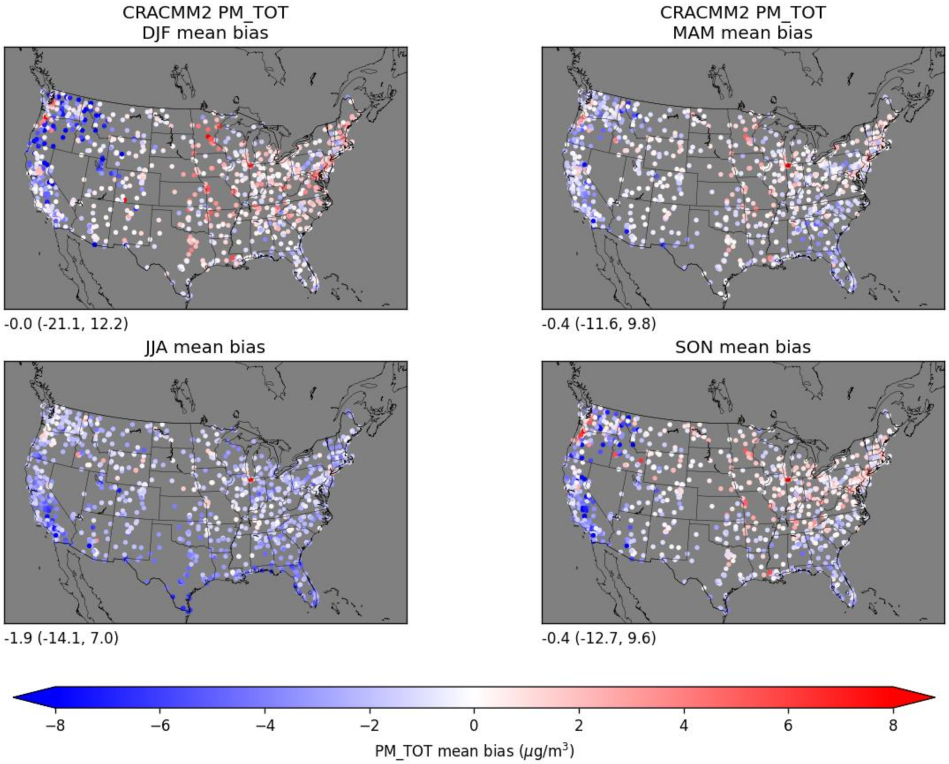


Figure S35. Seasonal mean bias in daily average $PM_{2.5}$ for 2019 at AQS monitoring sites for CRACMM2. The annotations in the lower left show the mean (min, max) site average biases.

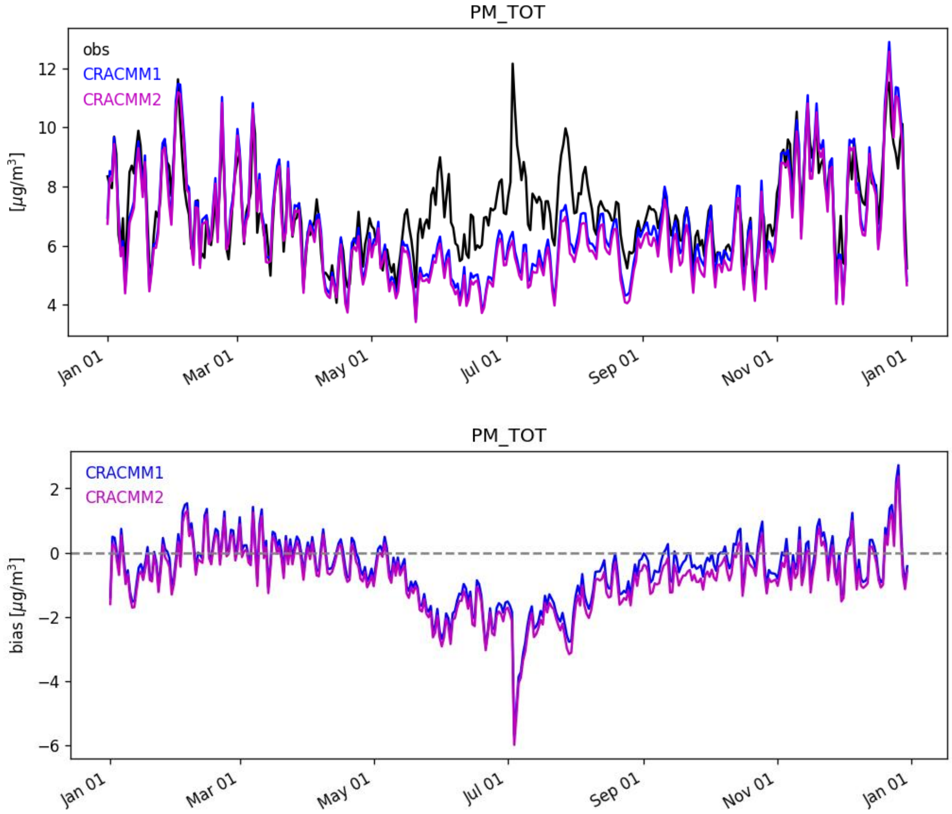


Figure S36. Daily average PM2.5 (top) and bias (bottom) for 2019 averaged over all AQS monitoring sites for CRACMM1 and CRACMM2.

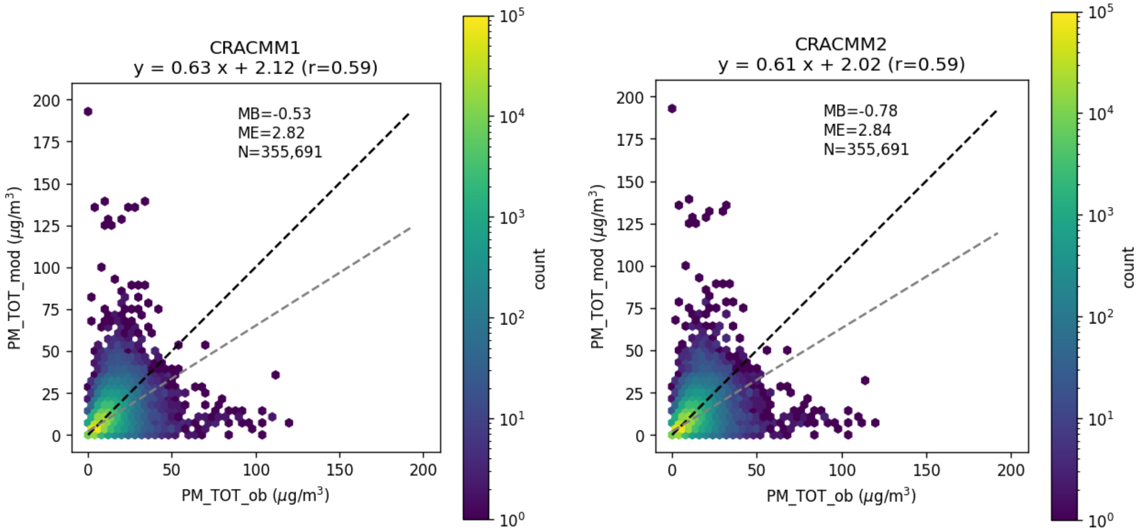
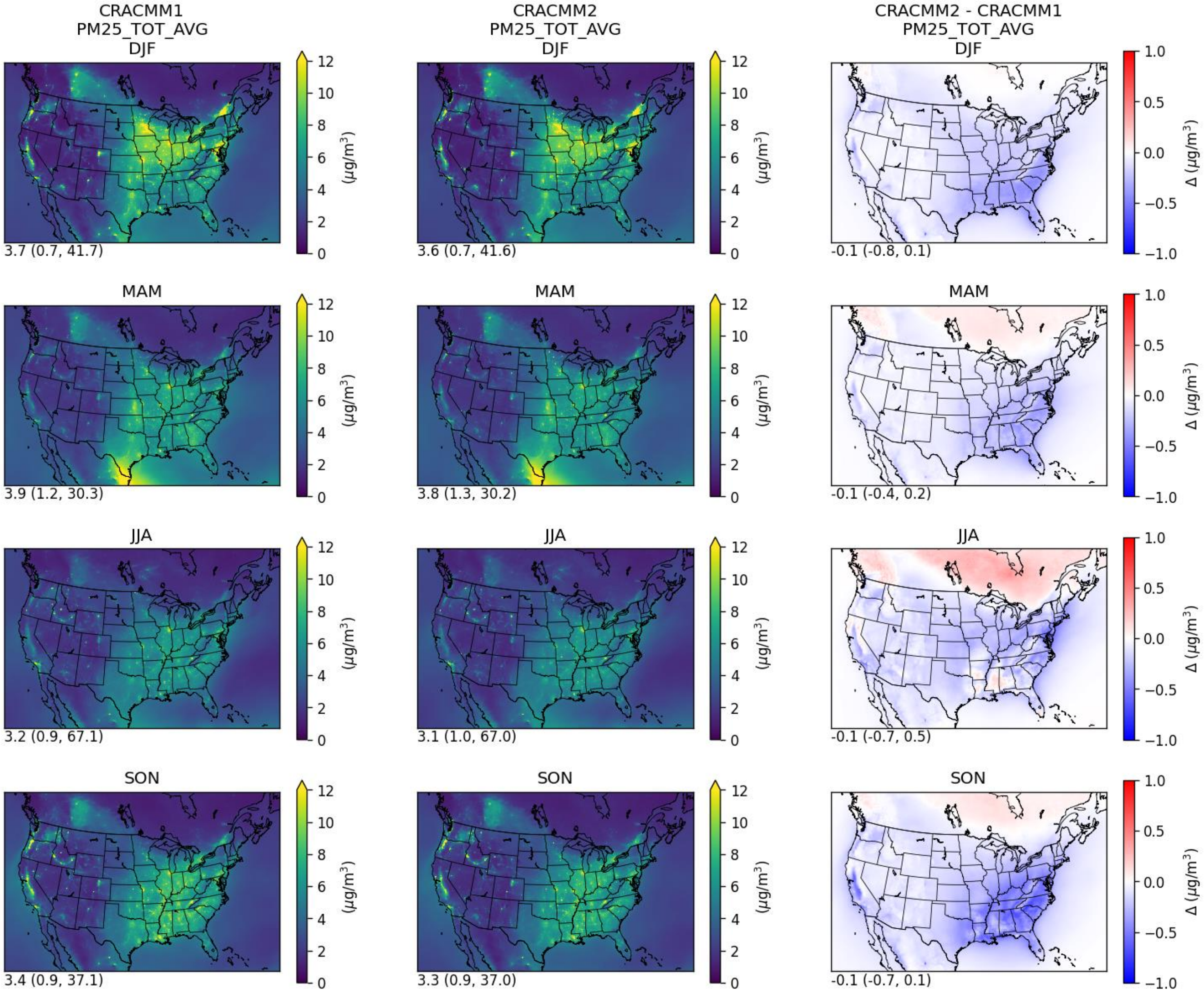


Figure S37. Density scatter plot of observed daily average PM2.5 vs. CRACMM1 (left) and CRACMM2 (right). The mean bias (MB), mean error (ME), and number of observations (N) are indicated on each panel.



440 **Figure S38. Seasonal average $PM_{2.5}$ in CRACMM1 (left column) and CRACMM2 (middle column) and the change in CRACMM2 compared to CRACMM1 (right column). The annotations in the lower left show the mean (min, max) over the modeling domain.**

References

Browne, E. C., Wooldridge, P. J., Min, K. E., and Cohen, R. C.: On the role of monoterpene chemistry in the remote continental boundary layer, Atmos. Chem. Phys., 14, 1225-1238, 10.5194/acp-14-1225-2014, 2014.

Fisher, J. A., Jacob, D. J., Travis, K. R., Kim, P. S., Marais, E. A., Chan Miller, C., Yu, K., Zhu, L., Yantosca, R. M., Sulprizio, M. P., Mao, J., Wennberg, P. O., Crounse, J. D., Teng, A. P., Nguyen, T. B., St. Clair, J. M., Cohen, R. C., Romer, P., Nault, B. A., Wooldridge, P. J., Jimenez, J. L., Campuzano-Jost, P., Day, D. A., Hu, W., Shepson, P. B., Xiong, F., Blake, D. R., Goldstein, A. H., Misztal, P. K., Hanisco, T. F., Wolfe, G. M., Ryerson, T. B., Wisthaler, A., and Mikoviny, T.: Organic nitrate chemistry and its implications for nitrogen budgets in an isoprene- and monoterpene-rich atmosphere: constraints from aircraft (SEAC4RS) and ground-based (SOAS) observations in the Southeast US, Atmos. Chem. Phys., 16, 5969-5991, 10.5194/acp-16-5969-2016, 2016.

Ivatt, P. D., Evans, M. J., and Lewis, A. C.: Suppression of surface ozone by an aerosol-inhibited photochemical ozone regime, Nature Geoscience, 15, 536-540, 10.1038/s41561-022-00972-9, 2022.

Jacob, D. J.: Heterogeneous chemistry and tropospheric ozone, Atmospheric Environment, 34, 2131-2159, https://doi.org/10.1016/S1352-2310(99)00462-8, 2000.

Jaffe, D. A., Ninneman, M., Nguyen, L., Lee, H., Hu, L., Ketcherside, D., Jin, L., Cope, E., Lyman, S., Jones, C., O'Neil, T., and Mansfield, M. L.: Key results from the salt lake regional smoke, ozone, and aerosol study (SAMOZA), Journal of the Air & Waste Management Association, 74, 163-180, 10.1080/10962247.2024.2301956, 2024.

KNMI: S5P/TROPOMI HCHO ATBD, 2022.

Mouat, A. P., Siegel, Z. A., and Kaiser, J.: Evaluation of Aeris mid-infrared absorption (MIRA), Picarro CRDS (cavity ring-down spectroscopy) G2307, and dinitrophenylhydrazine (DNPH)-based sampling for long-term formaldehyde monitoring efforts, Atmos. Meas. Tech., 17, 1979-1994, 10.5194/amt-17-1979-2024, 2024.

Pye, H. O. T., Luecken, D. J., Xu, L., Boyd, C. M., Ng, N. L., Baker, K. R., Ayres, B. R., Bash, J. O., Baumann, K., Carter, W. P. L., Edgerton, E., Fry, J. L., Hutzell, W. T., Schwede, D. B., and Shepson, P. B.: Modeling the Current and Future Roles of Particulate Organic Nitrates in the Southeastern United States, Environmental Science & Technology, 49, 14195-14203, 10.1021/acs.est.5b03738, 2015.

Schwantes, R. H., Emmons, L. K., Orlando, J. J., Barth, M. C., Tyndall, G. S., Hall, S. R., Ullmann, K., St. Clair, J. M., Blake, D. R., Wisthaler, A., and Bui, T. P. V.: Comprehensive isoprene and terpene gas-phase chemistry improves simulated surface ozone in the southeastern US, Atmos. Chem. Phys., 20, 3739-3776, 10.5194/acp-20-3739-2020, 2020.

U.S. EPA: Technical Support Document (TSD): Preparation of Emissions Inventories for the 2019 North American Emissions Modeling Platform, 2022.

Vereecken, L. and Nozière, B.: H migration in peroxy radicals under atmospheric conditions, Atmos. Chem. Phys., 20, 7429-7458, 10.5194/acp-20-7429-2020, 2020.

Wennberg, P. O., Bates, K. H., Crounse, J. D., Dodson, L. G., McVay, R. C., Mertens, L. A., Nguyen, T. B., Praske, E., Schwantes, R. H., Smarte, M. D., St Clair, J. M., Teng, A. P., Zhang, X., and Seinfeld, J. H.: Gas-Phase Reactions of Isoprene and Its Major Oxidation Products, Chemical Reviews, 118, 3337-3390, 10.1021/acs.chemrev.7b00439, 2018.

Wiser, F., Place, B. K., Sen, S., Pye, H. O. T., Yang, B., Westervelt, D. M., Henze, D. K., Fiore, A. M., and McNeill, V. F.: AMORE-Isoprene v1.0: a new reduced mechanism for gas-phase isoprene oxidation, Geosci. Model Dev., 16, 1801-1821, 10.5194/gmd-16-1801-2023, 2023.

Yang, B., Wiser, F. C., McNeill, V. F., Fiore, A. M., Tao, M., Henze, D. K., Sen, S., and Westervelt, D. M.: Implementation and evaluation of the automated model reduction (AMORE) version 1.1 isoprene oxidation mechanism in GEOS-Chem, Environmental Science: Atmospheres, 3, 1820-1833, 10.1039/D3EA00121K, 2023.

Zhu, Q., Schwantes, R. H., Coggon, M., Harkins, C., Schnell, J., He, J., Pye, H. O. T., Li, M., Baker, B., Moon, Z., Ahmadov, R., Pfannerstill, E. Y., Place, B., Wooldridge, P., Schulze, B. C., Arata, C., Bucholtz, A., Seinfeld, J. H., Warneke, C., Stockwell, C. E., Xu, L., Zuraski, K., Robinson, M. A., Neuman, J. A., Veres, P. R., Peischl, J., Brown, S. S., Goldstein, A. H., Cohen, R. C., and McDonald, B. C.: A better representation of volatile organic compound chemistry in WRF-Chem and its impact on ozone over Los Angeles, Atmos. Chem. Phys., 24, 5265-5286, 10.5194/acp-24-5265-2024, 2024.