Geosci. Model Dev., 16, 1801–1821, 2023
https://doi.org/10.5194/gmd-16-1801-2023


Geoscientific Model Development
Open Access
EGU



# AMORE-Isoprene v1.0: a new reduced mechanism for gas-phase isoprene oxidation

**Forwood Wiser[1], Bryan K. Place[2], Siddhartha Sen[3], Havala O. T. Pye[2], Benjamin Yang[4,8], Daniel M. Westervelt[4,5], Daven K. Henze[6], Arlene M. Fiore[7,8], and V. Faye McNeill[1,8]**

[1]Department of Chemical Engineering, Columbia University, New York, NY 10027, USA
[2]Office of Research and Development, Environmental Protection Agency, Research Triangle Park, NC 27711, USA
[3]Microsoft Research, New York, NY 10012, USA
[4]Lamont-Doherty Earth Observatory of Columbia University, Palisades, NY 10964, USA
[5]NASA Goddard Institute for Space Studies, New York, NY 10025, USA
[6]Department of Mechanical Engineering, University of Colorado, Boulder, Boulder, CO 80309, USA
[7]Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[8]Department of Earth and Environmental Sciences, Columbia University, New York, NY 10027, USA

**Correspondence:** V. Faye McNeill (vfm2103@columbia.edu)

Received: 4 October 2022 – Discussion started: 6 October 2022
Revised: 1 February 2023 – Accepted: 23 February 2023 – Published: 29 March 2023

**Abstract.** Gas-phase oxidation of isoprene by ozone ($O_3$) and the hydroxyl (OH) and nitrate ($NO_3$) radicals significantly impacts tropospheric oxidant levels and secondary organic aerosol formation. The most comprehensive and up-to-date chemical mechanism for isoprene oxidation consists of several hundred species and over 800 reactions. Therefore, the computational expense of including the entire mechanism in large-scale atmospheric chemical transport models is usually prohibitive, and most models employ reduced isoprene mechanisms ranging in size from ∼10 to ∼200 species. We have developed a new reduced isoprene oxidation mechanism using a directed-graph path-based automated model reduction approach, with minimal manual adjustment of the output mechanism. The approach takes as inputs a full isoprene oxidation mechanism, the environmental parameter space, and a list of priority species which are protected from elimination during the reduction process. Our reduced mechanism, AMORE-Isoprene (where AMORE stands for Automated Model Reduction), consists of 12 species which are unique to the isoprene mechanism as well as 22 reactions. We demonstrate its performance in a box model in comparison with experimental data from the literature and other current isoprene oxidation mechanisms. AMORE-Isoprene's performance with respect to predicting the time evolution of isoprene oxidation products, including isoprene epoxydiols (IEPOX) and formaldehyde, is favorable compared with other similarly sized mechanisms. When AMORE-Isoprene is included in the Community Regional Atmospheric Chemistry Multiphase Mechanism 1.0 (CRACMM1AMORE) in the Community Multiscale Air Quality Model (CMAQ, v5.3.3), $O_3$ and formaldehyde agreement with Environmental Protection Agency (EPA) Air Quality System observations is improved. $O_3$ bias is reduced by 3.4 ppb under daytime conditions for $O_3$ concentrations over 50 ppb. Formaldehyde bias is reduced by 0.26 ppb on average for all formaldehyde measurements compared with the base CRACMM1. There was no significant change in computation time between CRACMM1AMORE and the base CRACMM. AMORE-Isoprene shows a 35 % improvement in agreement between simulated IEPOX concentrations and chamber data over the base CRACMM1 mechanism when compared in the Framework for 0-D Atmospheric Modeling (F0AM) box model framework. This work demonstrates a new highly reduced isoprene mechanism and shows the potential value of automated model reduction for complex reaction systems.

**Published by Copernicus Publications on behalf of the European Geosciences Union.**

## 1 Introduction

Isoprene is the most abundant non-methane hydrocarbon in the atmosphere. It has a major impact on tropospheric oxidant levels (Butler et al., 2008) and contributes to secondary organic aerosol (SOA) formation (and therefore fine particulate matter, $PM_{2.5}$) in many parts of the US and the world (Kroll et al., 2006; Henze and Seinfeld, 2006; Farmer et al., 2010; Liu et al., 2016; Fu et al., 2009). During the warm season, isoprene emissions enhance both regional and hemispheric ozone abundances at northern midlatitudes (Fiore et al., 2011; Guo et al., 2018). Isoprene oxidation chemistry contributes to natural background ozone and particulate matter over much of the US during the warm season (Fiore et al., 2014). Different representations of isoprene chemistry lead to uncertainty in air pollutant responses to anthropogenic emission reductions (Carlton et al., 2009) and to differences in model estimates of the background versus anthropogenic fractions (Fiore et al., 2005, 2014).

Knowledge of the isoprene oxidation reaction mechanism, including key pathways for both ozone and aerosol formation, has advanced rapidly over the last 2 decades (Wennberg et al., 2018). The full chemical mechanism for isoprene oxidation, as it is currently understood, consists of several hundred species (up to 602 in the Master Chemical Mechanism, MCM, v3.3.1; Jenkin et al., 2015) and $\sim$ 1000 reactions (see Table 1). Due to its size and complexity, including every known intermediate species and reaction in the isoprene oxidation network in 3-D air quality and atmospheric chemistry models is not feasible. Therefore, most models employ reduced isoprene mechanisms. For reduced mechanisms, there is a trade-off between mechanism size (i.e., the number of species and reactions represented) and accuracy. The goal is to find the smallest possible reduced mechanism that still provides the accuracy required for the modeling application. Commonly used reduced isoprene mechanisms range in size from $\sim$ 10 to $\sim$ 200 species. Table 1 shows the size of a select set of isoprene mechanisms currently being used in atmospheric chemistry models. The reduced models, including the Common Representative Intermediates (CRI) mechanism, the Caltech Reduced Plus mechanism, the Regional Atmospheric Chemistry Mechanism (RACM), and the Carbon Bond (CB) mechanism, have been developed manually by expert air quality scientists using techniques such as surrogate mechanisms (lumped structure – Yarwood et al., 2005, or lumped species – Aumont et al., 2005; Goliff et al., 2013; Jenkin et al., 2019) and empirical parameterization, along with expert knowledge of the reaction system. While these approaches have been successful with respect to representing atmospheric chemistry for the specific chemical and environmental scenarios for which they were developed, the resulting models tend to lack flexibility to be adapted to new scenarios or to be rapidly updated. Their implementation is also labor-intensive.

Automated chemical mechanism reduction techniques provide the opportunity to flexibly and rapidly generate accurate reduced chemical mechanisms, and they lower the barrier to updating the mechanism as new knowledge becomes available. While automated mechanism reduction has been applied in limited studies in atmospheric chemistry (Whitehouse et al., 2004a, b; Watson et al., 2008; Xia et al., 2009; Nikolaou et al., 2018; Sturm, 2021; Kelp et al., 2022; Lin et al., 2023), it has been further developed in the field of combustion (Wei and Kuo, 1969; Tomlin et al., 1992; Tomlin et al., 1997; Massias et al., 1999; Lu et al., 2001; Lu and Law, 2005; Pepiot-Desjardins and Pitsch, 2008; Sun et al., 2010). Combustion mechanisms have a number of features in common with the isoprene oxidation mechanism, including their complexity and the large number of intermediates involved. Thus, techniques developed for the application to combustion mechanisms may be applicable to the isoprene oxidation mechanism as well.

The methods of model reduction, whether automated or manual, fall into two main categories. The first is reduction by removing less-important species or reactions. The other method is to group species and reactions together which may participate in similar reaction pathways (chemical lumping). Each method aims to reduce the computational cost for simulating the mechanism by reducing the complexity and size of the reaction network while also retaining accuracy within a given tolerance.

Graph theory has been used as a framework for many model reduction algorithms, including the ones used in this work. An influential method for reduction is the directed relation graph (DRG) method, developed by Lu and Law (2005). In this method, a graph representing the reaction mechanism is created, consisting of nodes (carbon-containing chemical species) connected to each other by directed edges (reactions). Each edge is given a weight based on the strength of the relationship between the two nodes, which is a function of the kinetic rate laws and parameters. In the model reduction process, edges are removed in order of the weightings. Other methods include variations on the edge-weighting calculation to include more indirect relationships between species. One such method, the directed relation graph with error propagation (DRGEP), was used to reduce the RACM (Nikolaou et al., 2018). This method was successful at significantly reducing the number of species while maintaining accuracy for simulating $O_3$. We tested the DRG method and found it to be unsuitable for application to the isoprene mechanism (see Sect. S3). Briefly, the DRG method is successful when a mechanism features a significant number of species that can be removed without a major consequence with respect to the desired accuracy. The scale of reduction required for the isoprene mechanism and the breadth of important priority species (Sect. 2.2) make it incompatible with the DRG method.

Here, we present a new reduced isoprene oxidation mechanism that we have developed using a novel graph-theory-

**Table 1.** Sizes of select isoprene mechanisms, including the mechanism in this work. Larger mechanism sizes are self-reported. For smaller mechanisms, species and reactions were recounted for this work using the following criteria: (i) only species unique to isoprene chemistry are included, which excludes species that do not contain the isoprene carbon backbone; (ii) all reactions involving species unique to isoprene are counted; and (iii) heterogeneous reactions involving isoprene species are not counted.

| Mechanism | No. of species | No. of reactions | Reference |
|---|---|---|---|
| MCM v3.3.1 | 602 | 1926 | Jenkin et al. (2015) |
| CRI 2.2 | 56 | 186 | Jenkin et al. (2019) |
| Caltech | 404 | 897 | Wennberg et al. (2018) |
| Caltech Reduced Plus | 131 | 220 | Wennberg et al. (2018) |
| RACM2 | 9 | 12 | Sarwar et al. (2013); Goliff et al. (2013) |
| CB6r3 | 10 | 17 | Yarwood et al. (2010); Emery et al. (2015) |
| AMORE-Isoprene | 12 | 22 | This work |

based Automated Model Reduction approach (AMORE), with minimal manual adjustment of the output mechanism (AMORE-Isoprene). We describe the model reduction algorithm and then demonstrate the performance of AMORE-Isoprene compared to experimental data in the literature and other isoprene oxidation mechanisms using a box model as well as when incorporated into the Community Multiscale Air Quality (CMAQ) modeling system (US EPA Office of Research and Development, 2021) as part of the Community Regional Atmospheric Chemistry Multiphase Mechanism (CRACMM1AMORE).

## 2 Methods

In this section, we describe our approach for model reduction and inputs to the process as well as the procedure used for testing the reduced mechanism.

In brief, an algorithm was developed to reduce the full isoprene mechanism to a smaller more manageable mechanism that can be used in 3-D chemical transport models. The output mechanism from the AMORE algorithm was subsequently adjusted manually to optimize its performance for use in atmospheric modeling. In order to test the AMORE-Isoprene mechanism, a mechanism error metric was devised.

The AMORE-Isoprene mechanism was the product of this methodology. Our novel algorithm was essential in the creation of this mechanism, but it requires further work before it can be used for other mechanisms and without manual adjustment.

### 2.1 Full mechanism input

A "full" chemical mechanism is required for the input to the reduction algorithm. The full mechanism also serves as a benchmark for the accuracy of the reduced mechanism. In this study, the reference isoprene oxidation mechanism was based on Wennberg et al. (2018). The Wennberg mechanism is a comprehensive compilation of isoprene oxidation chemistry from laboratory and computational studies published up to 2018, including the formation of isoprene epoxydiols (IEPOX) (Paulot et al., 2009), intramolecular $RO_2$ chemistry (autoxidation) (Teng et al., 2017), and recent advances in isoprene nitrate chemistry (Schwantes et al., 2015). Despite its size and complexity, some branches of the oxidation cascade are truncated in the Wennberg et al. (2018) mechanism due to a lack of published experimental constraints, specifically degradation pathways for some later-generation intermediates with two, three, or four functional groups (Bates and Jacob, 2019). Therefore, modeled on the approach used by Wennberg et al. (2018) in preparing the Caltech Reduced Plus mechanism, we expanded the Wennberg mechanism to include degradation of these species. Further details are available in Sect. S1 in the Supplement, including box model comparisons of original and extended mechanisms to EUROCHAMP data (Muñoz and Gómez-Alvarez, 2008; Muñoz, 2021a, b). In addition, the extended mechanism is listed in its entirety in Sect. S18. Briefly, the intermediates were mapped to lumped species in the Caltech Reduced Plus mechanism or to species in MCM v.3.3.1 and were assigned the corresponding degradation pathway, products, and rate constants from that mechanism. For the rest of this paper, we refer to this updated mechanism as the *Caltech full mechanism*. This mechanism was chosen instead of the MCM isoprene mechanism (Jenkin et al., 2015), which is of a similar size, because it includes the results of more-recent isoprene chamber studies which were not yet published at the time that the current MCM mechanism was developed (e.g., Teng et al., 2017).

### 2.2 Priority species

Given that model reduction necessarily involves removing or lumping chemical species from the mechanism, we identified a set of nine important organic species and eight important oxidant and nitrogen oxide species to be protected from elimination during the model reduction process. This priority species list was an input to the model reduction algorithm. A full table of these species is available in Sect. S2.

Besides isoprene, these species were chosen for their importance for SOA or brown carbon formation and/or expected impact on gas-phase photochemistry (isoprene epoxydiols (lumped), isoprene nitrates (lumped), glyoxal, methylglyoxal, methacrolein, methyl vinyl ketone, peroxyacetyl nitrate, methyl radical, and peroxyacetyl radical). Formaldehyde was also included in the protected species list due to its status as an air toxic (EPA, 2018; Zhu et al., 2017; Scheffe et al., 2016) and for its potential to indicate oxidant levels (Travis et al., 2022). Other species such as $NO_x$, $HO_x$, $O_3$, and other oxidants are included in the mechanism as well. The accuracy of the reduced isoprene mechanism is measured by its ability to simulate the time evolution of the concentrations of the priority species and of oxidants and nitrogen oxides under different conditions.

### 2.3 Reduction algorithm development

In general, a new reduced isoprene oxidation mechanism will be a good candidate for application in large-scale models if it provides gains in accuracy or computational efficiency. As a trade-off exists between mechanism size (and therefore computational efficiency) and accuracy, improvements in one aspect are sought which avoid sacrifices in the other. Thus, the mechanism should be of similar size and complexity to existing mechanisms (or smaller) as well as of equal or better accuracy. The most compact isoprene mechanisms, including those currently used in the CMAQ modeling suite (version 2 of the Regional Atmospheric Chemistry Mechanism, RACM2, and revision 3 of Carbon Bond 6, CB6r3), include roughly 10 species unique to the isoprene mechanism and up to 20 reactions (Table 1). Note that this list of species does not include all priority species; some, such as IEPOX and isoprene nitrates, are included, whereas others, such as formaldehyde and glyoxal, which lack the isoprene carbon backbone and are also formed through non-isoprene pathways, are not. Thus, an isoprene mechanism of comparable size to existing reduced mechanisms will have around 10 isoprene-specific species, around 4 of which (isoprene, isoprene nitrates, IEPOX, and methyl vinyl ketone), are already priority species. The remaining six species are isoprene intermediates which are not considered priority species themselves but play an important role in the dynamics of the isoprene mechanism and the production of priority species.

The AMORE algorithm represents the full mechanism as a graph. Many prior works have utilized graph theory to analyze chemical mechanisms (Ratkiewicz and Truong, 2003; Lu and Law, 2005; Pepiot-Desjardins and Pitsch, 2008; Sun et al., 2010; Nikolaou et al., 2018; Silva et al., 2021). In this work, nodes represent species, and edges represent a directed relationship between two species, in which one is a reactant and the other a product of the same set of reactions. Prior graph-based reduction methods have focused solely on removing nonessential components of the mechanism ("pruning" the graph). This work focuses instead on determining the optimal graphical structure of the final reduced mechanism, as constrained by target mechanism size. This is done by determining the essential mechanistic pathways needed to accurately represent the full mechanism in a reduced structure, as discussed below.

A mechanistic pathway consists of a set of reactions joined by intermediate species. For a path of $N$ reactions, there are $N - 1$ intermediates. With the constraint of six intermediate species, this allows for roughly six paths with two reactions, each with one intermediate, or for three paths with three reactions, each with two intermediates, with both options having six intermediates. If some pathways are able to share intermediates, then more pathways can be included. It is also our goal that the reduced mechanism structure maps as closely as possible to known reactions with measured rates.

A new algorithm was designed specifically to develop optimal mechanisms of roughly 10 total species, 6 intermediate species, and 20 reactions. At a high level, the algorithm identifies a small set of the most important mechanistic pathways in the full mechanism and concatenates them in order to reduce the number of intermediate species. The algorithm estimates the importance of a given mechanistic pathway by determining the impact each possible pathway has on the yields of priority species. The mechanism reduction algorithm has four main components: (1) a sub-algorithm to rapidly estimate the yields from isoprene of priority species under constant oxidant and nitrogen oxide concentrations and atmospheric conditions (yield estimation algorithm; Sect. 2.3.1), (2) a sub-algorithm to assess the importance of different pathways given the yields of priority species (pathway importance algorithm; Sect. 2.3.2), (3) a sub-algorithm for optimally combining pathways to reduce intermediate species (pathway combination algorithm; Sect. 2.3.3), and (4) a sub-algorithm to estimate yields of priority species for each pathway in the mechanism (priority species yield determination; Sect. 2.3.4). The overall AMORE algorithm process is shown in Fig. 1. All sub-algorithms are described in detail in the following sections.

#### 2.3.1 Yield estimation algorithm

The yield estimation algorithm utilizes graph theory and takes advantage of the relatively small number of cycles (a path in the graph that starts and ends at the same species) and the small number of reactions with two carbon-containing reactants in the isoprene oxidation scheme. It rapidly estimates the yields of all species from isoprene in the full mechanism, assuming the complete oxidation of isoprene and its products. The algorithm emulates the full mechanism so that the numerical simulation need not be run repeatedly during sensitivity testing. The algorithm begins by representing the full mechanism as a directed graph. The directed nature of the representative graph delineates the direction of the flow of carbon over time. Cycles are unique instances in this context, in which carbon flows in two different directions and

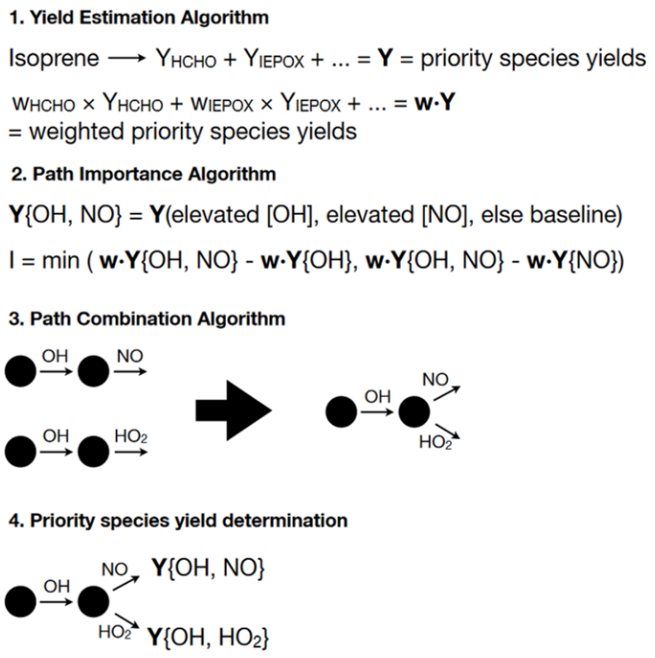


**Figure 1.** Schematic of the AMORE algorithm. The sub-algorithms are shown in order of implementation. Brackets are used specify a pathway within the mechanism, with each oxidant or nitrogen oxide within the brackets representing a reaction in a sequence involving that oxidant. For example, the pathway {OH, NO} represents a sequence of two reactions joined by an intermediate, in which OH and NO are reactants in the two reactions. The pathway {OH, NO} is shown as an example for sub-algorithms 2–4.

it is not necessarily evident which direction dominates. The algorithm takes oxidant and nitrogen oxide concentrations (OH, $HO_2$, $O_3$, $MO_2$, NO, $NO_2$, and $NO_3$), which are treated as constant, solar intensity, temperature, and pressure as inputs, and calculates the flux of carbon through the mechanism pathways using the rate law information provided. As this algorithm is dependent on oxidant and nitrogen oxide concentrations as well as other atmospheric parameters, it can be used to determine how yields are impacted by relevant atmospheric conditions.

The full mechanism is approximated using a directed acyclic graph (DAG). In order for a mechanism to be represented as a DAG, it must contain no cycles and reactions with two reactants must be broken into two sets of edges for each reactant, because edges can only represent the relationship between two species. For example, a reaction with two reactants and one product would become two edges, one for each reactant connecting to the product. Oxidant and nitrogen oxide concentrations are approximated to be constant, so reactions involving them are treated as pseudo-first-order reactions. Cycling in the isoprene oxidation system mainly takes place among oxidant and nitrogen oxide species, which are only represented implicitly in the graph. For cycles involving isoprene oxidation products, all species in the cycle are combined into one "super node". The incoming and outgoing edges of the super node include all edges of all species that it represents. The method used to reduce cycles to super nodes is described in Sect. S4.1.

The DAG structure is then utilized to calculate the partitioning of carbon between branches within the graph, ultimately giving an estimated yield for each oxidation product species. This novel approach takes advantage of the graphical representation of the mechanism to rapidly approximate yields which would otherwise require a box model to calculate. The resulting time savings allow a much larger set of input conditions to be tested than would be feasible with a box model. The yield is defined as the moles of each species produced per mole of isoprene reacted. Starting from a species of known yield, the yield of a direct product can be calculated as the rate constant involving said product over the sum of all rate constants reacting with the starting species. The yield for a species A from isoprene, $Y_A$, is calculated as follows:

$$Y_A = \sum_{i=1, i\neq A}^{N} Y_{i,\mathrm{isop}} Y_{A,i}, \tag{1}$$

$$Y_{A,i} = \frac{\sum_{r=1}^{R} k_r^I \max(\nu_{A,r}, 0)(-\min(\nu_{i,r}, 0))}{\sum_{r=1}^{R} k_r^I (-\min(\nu_{i,r}, 0))}. \tag{2}$$

Here, $Y_i$ is the yield of species $i$ from isoprene, $Y_{A,i}$ is the yield of species A from species $i$, $N$ is the number of species, $R$ is the number of reactions, $\nu_{i,r}$ is the stoichiometric coefficient of species $i$ in reaction r, and $k^I$ is the (pseudo-) first-order rate constant, that is, $k^I = k^{II}$[oxidant] for oxidation reactions, or else the first-order rate constant for photolysis and isomerization reactions. The yield of any species can be estimated once the yields of all its parent species in the graph are determined. Thus, with the assumptions and inputs outlined above, an estimate of the yield of all species from isoprene can be obtained for a given set of inputs. Running in a Jupyter notebook environment on a personal computer with a 1.8 GHz dual-core Intel Core i5 processor, it takes roughly 0.06 s for the algorithm to estimate all yields for a given set of conditions (50× shorter than a box model runtime of the full mechanism). This is a valuable tool for rapidly probing large mechanisms to study their outputs under a variety of inputs.

The yield estimation algorithm was tested for accuracy by comparing estimated yields to box model simulated yields for the range of conditions used for model development. A detailed analysis of the yield estimation algorithm accuracy is available in Sect. S4.2. A visualization of the yield estimation algorithm is shown in Fig. S7 in the Supplement.

### 2.3.2 Pathway importance algorithm

With the yield estimation algorithm in place, we developed a method to identify and evaluate the importance of paths within the mechanism. Given the constraints on the size of the final reduced mechanism discussed, the total number of

paths will vary depending on the number of intermediates that can be shared between paths. The full mechanism contains long, highly branched paths with multiple end products. Thus, no existing pathways within the full mechanism satisfy the design constraints. Instead, model paths were created in which each path was represented by a sequence of reactions with one of the possible oxidants or nitrogen oxides: OH, NO, $NO_2$, $NO_3$, $HO_2$, $O_3$, methyl peroxy radical ($MO_2$), or else photolysis. There was no requirement for a given path to be in the full mechanism; rather, paths recreate the oxidant- and nitrogen-oxide-dependent outcomes for the priority oxidation products. Each path was constrained to contain only irreversible reactions, with each oxidant or nitrogen oxide appearing no more than once (this constraint was lifted during the manual adjustment process). The justification for these simple paths is that isoprene oxidation product concentrations can be thought of as functions of isoprene, oxidant, and nitrogen oxide concentrations, and each path represents a scenario in which a set of oxidants and nitrogen oxides are favored. Thus, by containing multiple different paths, the priority species yields can be varied based on the oxidant and nitrogen oxide concentrations. For example, a path of {OH, NO} represents the reaction of OH or NO with isoprene to create a hypothetical intermediate and the reaction of the other oxidant or nitrogen oxide (either NO or OH) with that intermediate to form isoprene oxidation products. This path would be favored when OH and NO concentrations are high, and it allows for a unique distribution of priority species yields under these conditions. There were 256 possible paths, represented by non-duplicate combinations of the possible oxidants or nitrogen oxides. Temperature and pressure are other parameters that significantly influence isoprene chemistry. However, these parameters are implicit to the graph as inputs to calculate rate constants. Thus, temperature and pressure were not represented explicitly in the algorithm, leaving rate constants to be determined either through calibration or through direct reaction analogues in the full mechanism. The default temperature and pressure for yield estimates were 292 K and 1000 hPa, respectively.

Using the yield estimation algorithm, a measure of the importance of each path was determined by evaluating the product yields for a sequence of inputs designed to probe the sensitivity to each oxidant, nitrogen oxide, or photolysis. Each oxidant and nitrogen oxide was assigned a baseline concentration or intensity, determined from atmospherically relevant ranges in which rates of reactions involving each species were similar. For example, the baseline concentrations of $O_3$, OH, and $NO_3$ were set such that the rate of reaction of isoprene with each oxidant or nitrogen oxide would be the same. Input sequences were created in which the concentration of each oxidant or nitrogen oxide within the path, or photolysis intensity, is elevated in turn, roughly 1 order of magnitude above the baseline. Table 2 shows the input values used for the path importance algorithm. All possible combinations of each high and low value were used as an input space, resulting in 256 different input conditions. The goal in selecting input conditions was to find values that were relatively low and relatively high without biasing the algorithm with extreme values. They do not represent the full range of values that each input takes. The AMORE-Isoprene mechanism performs satisfactorily under more extreme conditions than those that were used as input conditions to the algorithm, but it would be possible to create a mechanism optimized for a more extreme scenario using the AMORE algorithm. We conducted a sensitivity test of the pathway importance algorithm to a select set of changes to the inputs shown in Table 2. Specifically, elevated concentrations of OH, NO, and $NO_2$ were adjusted to reflect realistic upper values for these species. The results of this test are given in Sect. S12 and Table S5.

**Table 2.** Baseline and elevated values of input parameters used in the pathway importance algorithm, a component of the AMORE algorithm.

| Parameter | Baseline value | Elevated value |
|---|---|---|
| Temperature (K) | 292 | – |
| Pressure (hPa) | 1000 | – |
| Solar intensity (unitless) | 0 | 1 |
| OH (ppb) | $1 \times 10^{-6}$ | $1 \times 10^{-4}$ |
| NO (ppb) | $1.17 \times 10^{-6}$ | 0.53 |
| $NO_2$ (ppb) | $1 \times 10^{-4}$ | 0.01 |
| $NO_3$ (ppb) | $2.3 \times 10^{-4}$ | 0.02 |
| $HO_2$ (ppb) | 0.04 | 0.2 |
| $O_3$ (ppb) | 16.7 | 100 |
| $MO_2$ (ppb) | 0.1 | 0.2 |

For each hypothetical path, the yield of priority species from that path was determined by elevating the input values of the oxidants or nitrogen oxides in the path. A path was considered important if this process resulted in yield estimates that differed significantly from the baseline. Multistep-path yields were evaluated in comparison to paths with one less elevated oxidant or nitrogen oxide step. For example, the path {OH, $HO_2$, NO} was compared to the paths {OH, $HO_2$}, {OH, NO}, and {$HO_2$, NO}. If the yield of priority species differed significantly from all of the compared paths, then the path was deemed important. The importance of each path was ranked in terms of the magnitude of difference in yield of the path to the least different shorter path. This method ensured that every component of the path was necessary to produce unique yields compared with the baseline. Equation (3) shows the importance metric used to choose the most important paths:

$$I = \min\left(\sum_{i=1}^{N} \frac{\mathrm{abs}(Y^0_{i,\mathrm{isop}} - Y^x_{i,\mathrm{isop}})}{\max(Y_{i,\mathrm{isop}})}, \forall x \in \text{path list}\right), \quad (3)$$

where $I$ is the path importance, $N$ is the number of species in the important species list, $i$ is an individual important species

in the list, $Y^0_{i,\mathrm{isop}}$ is the yield of species $i$ from isoprene for the path being measured, $Y^x_{i,\mathrm{isop}}$ is the yield of species $i$ from isoprene in path $x$, $\max(Y_{i,\mathrm{isop}})$ is the maximum yield obtained by species $i$ from the yields of all paths, and the path list is the set of paths with one less elevated oxidant or nitrogen oxide than the path being measured.

#### 2.3.3 Pathway combination algorithm

Using this path analysis, the following eight paths were identified and incorporated in the mechanism: {$O_3$}, {$NO_3$}, {$NO_3$, $HO_2$}, {$NO_3$, $HO_2$, photolysis intensity ($h\nu$)}, {$NO_3$, NO}, {OH}, {OH, $HO_2$}, and {OH, NO}. The number of paths was chosen based on the desired mechanism size, but the paths were determined by the pathway importance algorithm above. In order to reduce the number of intermediates, paths were joined together such that any shared oxidant or nitrogen oxide within paths had a shared intermediate. For example, all paths involving OH were structured so that the first reaction was with isoprene and OH which then formed a shared intermediate. The reaction paths were algorithmically structured to share as many intermediates as possible. The pathway combination algorithm started by grouping paths by a shared intermediate. For example, the paths {$NO_3$}, {$NO_3$, $HO_2$}, {$NO_3$, $HO_2$, $h\nu$}, and {$NO_3$, NO} all share a common $NO_3$ reaction step. There are instances in which there are multiple ways to group pathways. For example, {OH, NO} can either be grouped with other OH-containing pathways or with NO pathways. There was no algorithmic way to prioritize these two options. This is an instance in which manual intervention is required to assign preference between pathway groupings. This can be done by simply choosing the order in which the pathway reactions should occur. For example, choosing the order {OH, NO} would group this pathway with other OH pathways, whereas choosing the order {NO, OH} would group this pathway with other NO pathways.

Once the groupings are formed, an initial reaction step is created in which isoprene reacts with the commonly shared oxidant or nitrogen oxide to form an intermediate that is shared by all of the pathways. For example, in the $NO_3$ pathway grouping, the reaction of isoprene with $NO_3$ is shared with all pathways, which subsequently branch from each other. Pathways that share two oxidants or nitrogen oxides, such as {$NO_3$, $HO_2$} and {$NO_3$, $HO_2$, $h\nu$}, share two intermediates. By grouping pathways by shared oxidants and nitrogen oxides, and creating sub-groupings for multiple shared oxidants and nitrogen oxides, the pathway combination algorithm creates a reduced mechanism structure. This algorithm does not allow for the recombination of branched pathways, meaning that the resulting reduced graphs are, necessarily, trees. Figure S5 in Sect. S5 demonstrates the combination of all of the identified paths to form the reduced mechanism structure.

#### 2.3.4 Priority species yield determination

The yield of each priority species is measured for each path using the yield estimation algorithm (see Sect. 2.3.1). These yields are used as stoichiometric coefficients for the product terms of the terminal reaction of each path. All priority species are considered eligible as product terms of the terminal reaction of a given path. For each path, the terminal reaction is defined as the reaction in which no additional intermediates were produced. For example, the path {OH} contains the reaction of isoprene with OH to form isoprene hydroxy peroxy radical as an intermediate (Reaction SR1 in Table S2). This pathway was then given a terminal reaction, involving the first-order decomposition of the isoprene peroxy radical in order to produce the final priority oxidation products (Reaction SR2 in Table S2). The stoichiometric coefficients of each oxidation product were the yields as estimated by the yield estimation algorithm.

This algorithm completed the automated portion of the mechanism development process. The fully automated mechanism is described in Table S2. The assignment of reaction rate constants and species naming are discussed in the following section. The subsequent manual optimization process for direct implementation into 3-D atmospheric models is described in Sect. 2.4.

#### 2.3.5 Rate parameter identification and species naming

Once the skeletal reduced mechanism was established, rate parameters and species names were identified manually. The first step was to identify any direct analogues between mechanisms in the reduced mechanism and known reactions (i.e., those in the Caltech full mechanism). There were many reactions with direct analogues, including all reactions involving isoprene. In these cases, the rate law and parameters assigned were identical to the original.

For reactions without direct analogues, the reaction was typified by the oxidant or nitrogen oxide involved. In the Caltech full mechanism, reaction rate laws with the same oxidant or nitrogen oxide tend to have a similar form and fall under a limited range of parameter values. Where there were multiple possible reaction forms, the most common form was chosen. After choosing the form of the rate laws, parameters were tuned by running box model simulations under conditions that favored the reaction being tested. The parameters were calibrated to match the concentration profiles of dominant products in comparison to the Caltech full mechanism. A list of all rate laws and parameters, their analogues, and the method of selection is given in Table S3.

All species names listed in the AMORE-Isoprene mechanism were manually identified after the completion of the automated mechanism reduction process. As with the rate law selection process, the first step was to identify direct analogues in the full mechanism. As the AMORE-Isoprene mechanism is highly reduced, all species with the excep-

tion of isoprene represent groups of species in the Caltech full mechanism. Thus, direct analogues were generally analogous groups of species. For species without a clear analogue, naming was based on the oxidants and nitrogen oxides that reacted to form the species. From this information, a name was assigned based on the predicted functional groups present in the species. For CMAQ modeling, the naming convention is different for some species due to their prior existence in the model. Table S4 gives each species name for this paper and for CMAQ, the analogues that the species represents, and the functional groups involved.

### 2.4 Manual mechanism optimization and evaluation

The algorithmically generated isoprene mechanism was manually optimized for use in the CMAQ modeling environment and evaluated for its performance compared to other reduced mechanisms. The optimization process was done using the Framework for 0-D Atmospheric Modeling (F0AM) box model (Sect. 2.4.1) and the CMAQ testing environment (Sect. 2.4.2), and the manual optimization process is described in Sect. 2.4.3. The graph theoretical framework helped inform our decisions in this process. For example, the conceptualization of the mechanism as a set of unique pathways connected by sequences of reactions, which is rooted in graph theory, helped us to categorize reactions and how adjustments to their parameters would impact end results under different testing conditions.

In the process of evaluating the mechanism, an error metric was developed and used for quantitative comparisons between mechanisms (Sect. 2.4.4). In the optimization and evaluation phase, the Caltech full mechanism was used as a baseline for comparison, along with experimental chamber data for further corroboration (Paulot et al., 2009).

Higher priority was put on mechanism accuracy rather than retention of the original algorithmically generated mechanism structure. Thus, changes were made that deviated from the algorithmically generated mechanism; however, the core components of the algorithmically generated mechanism, including a majority of the identified important paths, were retained, and the algorithmically generated mechanism provided an essential functional starting point from which to improve the final mechanism performance.

#### 2.4.1 Box model testing

F0AM (Wolfe et al., 2016) was used to simulate isoprene mechanisms for the purpose of evaluating the AMORE-Isoprene mechanism. The 0-D box model testing was done in two primary phases. The first phase was aimed at optimizing the AMORE-Isoprene mechanism. The Caltech full mechanism was taken as the most accurate mechanism for ground truth, and RACM2 was used as a benchmark for comparison. Simulated concentration profiles of key species such as $NO_x$, $HO_x$, IEPOX, $O_3$, and formaldehyde were analyzed in order to assess the AMORE-Isoprene mechanism. The goal was to match both the magnitude and form of each species concentration in the Caltech full mechanism. A detailed description of the matching process is provided in Sect. 2.4.3.

The second phase of box model testing involved quantitative comparisons between mechanisms for demonstration of the performance of the AMORE-Isoprene mechanism. The mechanism was tested in the F0AM environment alongside the Caltech full mechanism, the RACM2 isoprene mechanism used in base CRACMM1, CB6r3, and the Caltech reduced isoprene mechanism. An error metric was created to determine the degree of matching between two concentration curves. This error metric was averaged over many species and conditions to create an overall mechanism error metric. Section 2.4.3 gives a detailed description of the error metric developed for this study.

A set of six input conditions was devised to simulate the mechanisms. Given that isoprene oxidation is split into three main pathways of reaction with OH, $NO_3$, and $O_3$, these three pathways must be represented in the chosen testing conditions. Subsequent oxidation with NO is particularly important in the OH pathway, and low-light conditions are important in the $NO_3$ pathway. Given this, the first five conditions were low $NO_x$, high $NO_x$, high $O_3$, high $NO_3$, and high $NO_3$ with low *hv*. The final input condition was set to simulate the chamber study of Paulot et al. (2009), allowing for the pairing of box model results with experimental results. In that study, $H_2O_2$ photolysis was used as the source of OH, and small amounts of $NO_x$ were measured as well. For all F0AM simulations, $H_2O_2$ was used as the source of OH (which allowed for OH to be a dynamic quantity), and NO was used as a source of $NO_x$. For ozone and $NO_3$ the concentrations were set directly. Due to the lack of $NO_3$ cycling and the resulting rapid decay of $NO_3$, $NO_3$ concentrations were held constant for high-$NO_3$ conditions in order to favor this pathway for the duration of the simulation. Temperature and pressure were held at 292 K and 1000 hPa for all conditions. This corresponds to low-elevation, warm conditions that are most relevant for isoprene chemistry. The rate of photolysis reactions are scaled by a unitless parameter labeled as *hv*. The value of this parameter was calibrated to match the results of Paulot et al. (2009) chamber data for high-photolysis conditions. Table 3 shows the inputs for each of the six conditions.

#### 2.4.2 CMAQ modeling

CMAQ v5.3.3 (Appel et al., 2021) with additional updates, as outlined in Place et al. (2023), was used to conduct simulations over the northeastern US for June through August 2018 (2–31 May used as spin-up) at a 4 km × 4 km horizontal resolution. Baseline gas- and aerosol-phase chemistry was specified by version 1 of CRACMM (Pye et al., 2022) which uses the RACM2 representation of isoprene chemistry (Sarwar et al., 2013). Additional simulations were con-

**Table 3.** F0AM box model testing input conditions used for calculating the error metric and evaluating the AMORE-Isoprene mechanism. Bold values represent species concentrations that were held constant. All other concentrations varied with time after initiation of the simulation.

| Species | Chamber comparison | Low $NO_x$ | High $NO_x$ | High $NO_3$, low *hv* | $NO_3$ | High $O_3$ |
|---|---|---|---|---|---|---|
| Isoprene (ppb) | 92.5 | 10 | 10 | 10 | 10 | 10 |
| $H_2O_2$ (ppb) | 1660 | 200 | 200 | 0 | 100 | 200 |
| NO (ppb) | 1 | 0.5 | 5 | 2 | 1 | 0.5 |
| $O_3$ (ppb) | 0 | 0 | 0 | 0 | 0 | 100 |
| $NO_3$ (ppb) | 0 | 0 | 0 | **0.02** | **0.02** | 0 |
| *hv* (unitless) | 3.5 | 3.5 | 3.5 | 0.5 | 3.5 | 3.5 |

ducted in which CRACMM's isoprene chemistry was replaced with AMORE-Isoprene. Meteorology was obtained from v4.1.2 of the Weather Research and Forecasting (WRF) model (Torres-Vazquez et al., 2022) and processed through version 5 of the Meteorology–Chemistry Interface Processor (Otte and Pleim, 2010). Boundary and initial conditions were mapped from previous work using CB6r3 (Torres-Vazquez et al., 2022), and emissions were respeciated for CRACMM with additional updates for volatile chemical products (Seltzer et al., 2021). Biogenic emissions were estimated with the Biogenic Emission Inventory System (BEIS) (Bash et al., 2016) with M3dry (Pleim et al., 2019) used for deposition. CMAQ output was compared to EPA Air Quality System (AQS) and other monitoring network data using the Atmospheric Model Evaluation Tool (AMET) (Appel et al., 2011). CRACMM was selected as a baseline mechanism due to concurrent development of AMORE-Isoprene and the CRACMM mechanism for use in EPA research. CRACMM indicated relatively consistent predictions of gas-phase ozone chemistry compared to other current mechanisms (Place et al., 2023), signifying that the choice of CRACMM as the baseline mechanism for 3-D modeling was unlikely to confound the AMORE-Isoprene results.

IEPOX has heterogeneous chemistry in CMAQ (reactive uptake leading to SOA) following Pye et al. (2013) with updates in Pye et al. (2017, 2022). The first-generation isoprene organic nitrate heterogeneous chemistry (leading to $HNO_3$ and gas-phase alcohols) was implemented in this work and is specific to AMORE (not in base CRACMM1).

In CMAQ, the species in AMORE undergo deposition. All species that were already present in the base CRACMM1 mechanism were treated the same as in CRACMM1. IPN and IPC (briefly described in Sect. 3) were both wet deposited with Henry's law coefficients predicted by OPERA (Mansouri et al., 2018). In addition, the species were dry deposited using species-specific diffusivities, mesophyll resistances, and LeBas molar volumes (Pye et al., 2017).

### 2.4.3 Manual mechanism adjustment

In this section, we discuss manual adjustments to the algorithmically generated mechanism. To make adjustments,

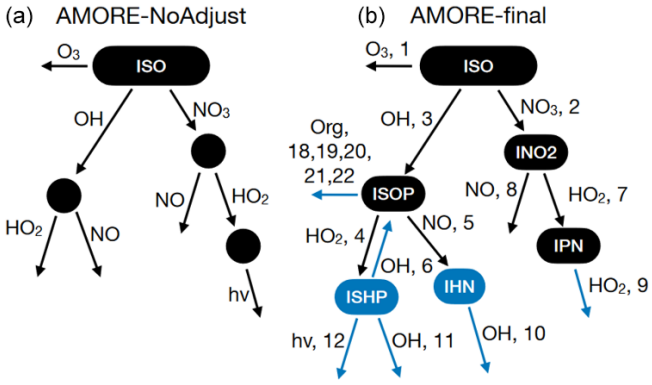


**Figure 2.** The original AMORE algorithmically generated mechanism prior to manual adjustment **(a)** and the final AMORE-Isoprene mechanism **(b)**. Changes are highlighted in blue.

we tested the AMORE mechanism in box model simulations (Sect. 2.4.1) and 3-D chemical transport simulations (Sect. 2.4.2). The testing process highlighted issues with the mechanism initially produced by the reduction algorithm that could be corrected via manual adjustments. This process has informed future algorithm development, as the ultimate goal is to automatically generate mechanisms which require no manual adjustment. The structural differences between the automated (labeled as AMORE-NoAdjust) and manually adjusted (labeled as AMORE-final) mechanisms are shown in Fig. 2. The corresponding reaction numbers from Table 4 are shown in the AMORE-final structure. Reactions (R13)–(R17) are not shown in the structure because they represent degradation schemes for end product species (IEPOX; isoprene hydroxy nitrate, IHN; and lumped multifunctional isoprene nitrates, ISON) or are used for oxidant and nitrogen oxide cycling and do not directly contribute to the production of priority species (Reaction R13).

The first issue to be addressed was that, because of the DAG assumptions, mechanistic pathways were constrained to have only forward reactions, and, because of the pathway identification algorithm, each oxidant or nitrogen oxide was only able to appear once. Adjustments were made to the mechanism to allow reversible reactions and repeat ap-

**Table 4.** The AMORE-Isoprene mechanism reactions and rate constants. Mechanism-specific species are listed in the text.

| No. | Reaction | Rate constant |
|---|---|---|
| R1 | ISO + O3 = 0.07 MACR + 0.189 MVK + 0.25 HO + 0.25 HO2 + 0.58 HCHO + 0.08 MO2 + 0.1 ACO3 + 0.09 H2O2 + 0.1 MACP + 0.461 MACR + 0.14 CO + 0.28 ORA1 + 0.15 OLT | $1.58 \times 10^{-14} \exp(-2000/T)\ \mathrm{cm^3\ mol^{-1}\ s^{-1}}$ |
| R2 | ISO + NO3 = INO2 + 0.3 HCHO + 0.3 NO2 + 0.3 ISON | $2.95 \times 10^{-12} \exp(-450/T)\ \mathrm{cm^3\ mol^{-1}\ s^{-1}}$ |
| R3 | ISO + HO = ISOP + 0.02 MO2 | $2.69 \times 10^{-11} \exp(390/T)\ \mathrm{cm^3\ mol^{-1}\ s^{-1}}$ |
| R4 | ISOP + HO2 = ISHP + 0.6 HO2 + 0.15 HCHO | $4.5 \times 10^{-13} \exp(1300/T)\ \mathrm{cm^3\ mol^{-1}\ s^{-1}}$ |
| R5 | ISOP + NO = 0.14 IHN + 0.7 HCHO + 0.44 MVK + 0.88 HO2 + 0.78 NO2 + 0.28 MACR + 0.021 GLY | $2.7 \times 10^{-12} \exp(350/T)\ \mathrm{cm^3\ mol^{-1}\ s^{-1}}$ |
| R6 | ISHP + HO = ISOP | $4.6 \times 10^{-12} \exp(200/T)\ \mathrm{cm^3\ mol^{-1}\ s^{-1}}$ |
| R7 | INO2 + HO2 = IPN + HO | $3.14 \times 10^{-14} \exp(580/T)\ \mathrm{cm^3\ mol^{-1}\ s^{-1}}$ |
| R8 | INO2 + NO = 0.2 ISON + 0.9 HCHO + 0.5 MGLY + 0.8 MVK + 0.5 NO2 + HO2 + 0.1 MO2 | $9.42 \times 10^{-16} \exp(580/T)\ \mathrm{cm^3\ mol^{-1}\ s^{-1}}$ |
| R9 | IPN + HO2 = 0.2 ISON + 0.8 NO2 + 0.4 HCHO + 0.05 GLY + 0.1 MGLY + 0.4 MACR + HO2 + 0.94 MVK + 0.1 MO2 | $3.4 \times 10^{-11} \exp(390/T)\ \mathrm{cm^3\ mol^{-1}\ s^{-1}}$ |
| R10 | IHN + HO = ISON + HO + 0.2 IEPOX | $2.4 \times 10^{-7} \exp(580/T)\ \mathrm{cm^3\ mol^{-1}\ s^{-1}}$ |
| R11 | ISHP + HO = 0.05 IPC + 0.15 HCHO + 0.05 MGLY + 0.15 MACR + 0.02 GLY + 0.2 MVK + 0.4 NO2 + 0.58 IEPOX + 0.8 HO | $2.97 \times 10^{-11} \exp(390/T)\ \mathrm{cm^3\ mol^{-1}\ s^{-1}}$ |
| R12 | ISHP = 0.4 HCHO + 0.1 MGLY + 0.06 ACO3 | Photol(HCHO_RAD_RACM2) $\mathrm{s^{-1}}$ |
| R13 | IPC + NO = 0.35 NO2 + 0.8 NO | $1 \times 10^{-10}\ \mathrm{cm^3\ mol^{-1}\ s^{-1}}$ |
| R14 | ISON + HO = CO + 0.12 NO2 | $5 \times 10^{-11}\ \mathrm{cm^3\ mol^{-1}\ s^{-1}}$ |
| R15 | ISON + NO3 = CO | $2 \times 10^{-14}\ \mathrm{cm^3\ mol^{-1}\ s^{-1}}$ |
| R16 | IHN = HNO3 | $2.3 \times 10^{-5}\ \mathrm{s^{-1}}$ |
| R17 | IEPOX + HO = HO | $5 \times 10^{-11} \exp(-400/T)\ \mathrm{cm^3\ mol^{-1}\ s^{-1}}$ |
| R18 | ISOP + MO2 = HO2 + 1.31 HCHO + 0.159 MACR + 0.250 MVK + 0.250 MOH + 0.250 ROH + 0.023 ALD + 0.018 GLY + 0.016 HKET | $3.4 \times 10^{-14} \exp(221/T)\ \mathrm{cm^3\ mol^{-1}\ s^{-1}}$ |
| R19 | ISOP + ACO3 = 0.5 HO2 + 0.5 MO2 + 1.048 HCHO + 0.219 MACR + 0.305 MVK + 0.5 ORA2 | $8.4 \times 10^{-14} \exp(221/T)\ \mathrm{cm^3\ mol^{-1}\ s^{-1}}$ |
| R20 | ISOP + APIP2 = 0.96 HOM + 0.48 ROH + 0.48 HCHO + 0.48 MVK + 0.48 HO + 0.48 HO2 + 0.04 ELHOM | $1 \times 10^{-10}\ \mathrm{cm^3\ mol^{-1}\ s^{-1}}$ |
| R21 | ISOP + APINP2 = 0.96 HOM + 0.48 ROH + 0.48 HCHO + 0.48 MVK + 0.48 NO2 + 0.48 HO2 + 0.04 ELHOM | $1 \times 10^{-10}\ \mathrm{cm^3\ mol^{-1}\ s^{-1}}$ |
| R22 | ISOP + LIMNP2 = 0.96 HOM + 0.48 ROH + 0.48 HCHO + 0.48 MVK + 0.48 NO2 + 0.48 HO2 + 0.04 ELHOM | $1 \times 10^{-10}\ \mathrm{cm^3\ mol^{-1}\ s^{-1}}$ |

A complete species list can be found in Table S1 in the Supplement.

pearances of a single oxidant or nitrogen oxide where there was a strong case for the adjustment based on the Caltech full mechanism. One of the most important instances of this is the isoprene OH oxidation pathway. In this pathway, OH and $HO_2$ are the most important oxidants as predicted by the AMORE algorithm; however, there is a reversible OH reaction (Reaction R6 in Table 4) which plays a significant role in the cycling of $HO_2$ and OH. This reversible reaction was added and was instrumental in improving the accuracy of the AMORE-Isoprene mechanism. The change is shown

in the {OH, $HO_2$} pathway of Fig. 2. This reversible reaction is absent from RACM2 but present in CB6r3. In order for the reversible reaction to terminate into final products, a reaction of the second intermediate with OH was added. The addition of these two reactions did not change the overall nature of the path {OH, $HO_2$} in terms of the oxidants present but added necessary complexity to the dynamics of the path, resulting in more accurate product differentiation in OH-dominant conditions.

It was also observed that $NO_x$ concentrations were relatively low compared with the Caltech full mechanism in low-$NO_x$ regimes in which the {OH, $HO_2$} pathway was dominant. To ameliorate the lower $NO_x$ concentrations, an additional intermediate called IPC was created for the purpose of reacting with NO to create additional $NO_2$ and NO. This addition is shown in Reaction (R11) (Table 4) where IPC is a product, and Reaction (R13) (Table 4), where NO and $NO_2$ are cycled. The effect of this addition is to increase $NO_x$ under low-$NO_x$ conditions, and thus increase ozone, leading to reduced ozone underestimation compared with the Caltech full mechanism.

In addition to box model testing, 3-D chemical transport modeling using CMAQ (Sect. 2.4.2) informed structural adjustments to the AMORE mechanism. These adjustments centered on the treatment of IHN (isoprene hydroxy nitrate, the intermediate of the reaction of ISOP (isoprene hydroperoxy radical) and NO (Reaction R5 in Table 4, as part of the {OH, NO} pathway) in the mechanism. IHN was not initially identified as a priority species during algorithm development and, thus, was not included as an intermediate. Instead, the reaction of ISOP and NO contained no intermediates and led directly to the production of priority end products. However, it was determined that IHN should be given priority based on recent research highlighting its importance in $NO_x$ cycling (Vasquez et al., 2020). Thus, IHN was added as an intermediate, and an additional decomposition reaction with OH was added (Reaction R10 in Table 4). This decomposition reaction led to the production of IEPOX and isoprene nitrates, which were originally produced directly from the reaction of ISOP with NO. Thus, the {OH, NO} pathway was expanded on by adding an additional OH reaction step for the decomposition of IHN. This change is shown in the {OH, NO} pathway with the addition of an OH reaction step in Fig. 2. In addition to decomposition into other organics, IHN acts as a sink for $NO_x$. This was represented by the addition of Reaction (R16) in Table 4, which did not involve any oxidants or nitrogen oxides as reactants. It was observed that the reaction of IPC with NO (Reaction R13 in Table 4) outcompeted IHN for NO, and thus the yield of IPC (Reaction R11) was changed from 0.3 to 0.05 from isoprene hydroxy peroxide (ISHP). This change came at the expense of $NO_x$ cycling under low-$NO_x$ conditions; however, it was observed that simulated $NO_x$ levels were largely the same between AMORE-Isoprene and the base CRACMM mechanism, suggesting that this adjustment would be a net benefit to the overall performance. Further discussion of IHN can be found in Sect. S7.

Additional reactions of the OH pathway with organic radicals (methyl radical, peroxyacetyl radical, and lumped terpene radicals) were added directly from the RACM2 mechanism. They were not identified as important by the AMORE mechanism, likely a result of the inputs chosen, but came at little additional computational cost because they did not require the addition of any intermediates. These added organic radical reactions allowed for product differentiation in environments where organic radical concentrations are significant.

Finally, the {$NO_3$, $HO_2$, $h\nu$} path was determined to be unnecessary due to the relatively small amount of flux carbon directed to it. Instead, the paths {$NO_3$, $HO_2$, $HO_2$} and {OH, $HO_2$, $h\nu$} were used in its place. The {$NO_3$, $HO_2$, $HO_2$} path was a variant on the existing {$NO_3$, $HO_2$} path. The {OH, $HO_2$, $h\nu$} path was added to represent any potential variation attributed to photolysis in the low-$NO_x$ regime. All of the above changes are shown in Fig. 2.

The stoichiometric coefficients of the products in the reduced mechanism were initially assigned based on the estimates given by the yield estimation algorithm and were then optimized manually. Notably, stoichiometric coefficients for oxidant and nitrogen oxide species as products, although clearly important, were not treated in the algorithm due to their implicit representation in the mechanism graph. Thus, while oxidants and nitrogen oxides were included in the algorithm as reactants, they were omitted as products, which reduced their overall accuracy. In particular, the relationship between $HO_x$, $NO_x$, and $O_3$ is very sensitive to changes in the isoprene mechanism and is important for determining yields of many other species. These oxidants and nitrogen oxides are of a high order of importance for mechanism accuracy, and so the manual adjustment of their presence in the mechanism was critical. Two clear examples of this are shown in Fig. 3. Prior to the adjustment shown in Fig. 3a, $HO_2$ was significantly reduced in the AMORE-Isoprene mechanism under low-$NO_x$ conditions. It was observed that Reaction (R4) (Table 4) involving the isoprene hydroxy peroxy radical (ISOP) was the main sink for $HO_2$ and that the accuracy was significantly increased by adding $HO_2$ to the product term. This original discrepancy likely reflects a cycling achieved by multiple reactions in the full mechanism that it was not possible to include in the reduced mechanism. Adding $HO_2$ to the product term was the only available way to have good agreement with the full mechanism. The reaction of isoprene hydroxy peroxy radical (ISOP) with NO was another reaction for which oxidant and nitrogen oxide cycling was very impactful. The addition of $NO_2$ and $HO_2$ to the products of Reaction (R5) (Table 4) was used to improve the accuracy of the AMORE-Isoprene mechanism. A demonstration of the adjustment improvements is shown in Fig. 3b. Further tests of the adjustments in Fig. 3 are given in Sect. S16.

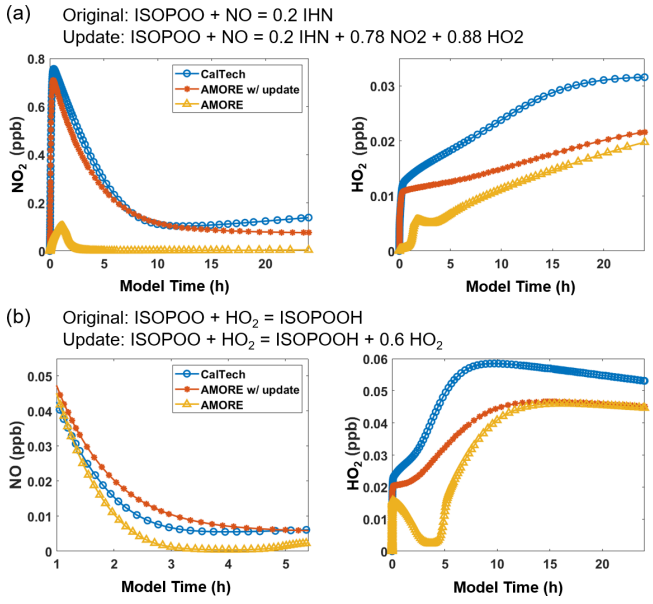


**Figure 3.** Box model simulations ($T = 292$ K, $p = 1000$ hPa) showing the improvement in performance of the AMORE mechanism for $HO_2$ and $NO_x$ after adding these species to the products of Reactions (R4) **(b)** and (R5) **(a)**. The original and updated reactions are shown above the plots. Inputs are **(a)** 200 ppb $H_2O_2$, 1 ppb NO, 10 ppb isoprene, and moderate-photolysis conditions (F0AM photolysis parameter = 1) and **(b)** 200 ppb $H_2O_2$, 1 ppb NO, 10 ppb isoprene, and high-photolysis conditions (F0AM photolysis parameter = 3.5).

All of the above adjustments were motivated by a clear improvement in mechanism performance that accompanied the change. See Sects. S8 and S9 and Tables S3 and S4 for a description of rate constants and species names.

All product stoichiometric coefficients were optimized for the accuracy of all priority species in a method similar to that shown for oxidants and nitrogen oxides in Fig. 3.

### 2.4.4 Mechanism error metric

For intercomparison of reduced mechanisms, full mechanisms, and experimental data, it was necessary to devise an accuracy metric based on the priority species and other measurable parameters. In the case of the isoprene mechanism, we focus on atmospheric oxidants and nitrogen oxides, organic aerosol, and other pollutants, namely formaldehyde and ozone.

In order to measure these parameters and create an accuracy metric, three steps were taken. The first was to define an error function for comparing the concentration of a species between two mechanisms in a box model simulation. The second was to determine the set of input conditions needed to capture the desired range of performance in the mechanism. The third step was to average errors across species and conditions in order to come up with a final metric. The error function for the comparison of concentration profiles of one species between two mechanisms formed the basis of the accuracy metric. The goal was to devise an error function which is bounded, so the natural choice was to normalize the error. In addition, for the purpose of averaging, the error function needed to always be positive to avoid canceling out errors. From this, an error metric was defined, as shown in Eq. (4):

$$E = \frac{\int_{t_0}^{t_f} \mathrm{abs}(T(t) - R(t))\mathrm{d}t}{\int_{t_0}^{t_f} \max(T(t), R(t))\mathrm{d}t}, \quad (4)$$

where $E$ is the error, $t_0$ is the initial time, $t_f$ is the final time, $T(t)$ is the concentration profile being tested, and $R(t)$ is the reference concentration profile. This concentration error metric ranges from zero to one, where zero is no error and one is infinite error. Figure S6 illustrates the behavior of the error metric for a sample set of profiles.

Although many important species are tracked in the isoprene mechanism, not all species contribute equally to observable parameters. A weighting scheme was devised to capture the relative importance of some species over others. The three main groupings that were included in the weighting scheme were oxidants and nitrogen oxides, priority pollutants, and isoprene SOA species. Each grouping was given an equal contribution to the overall error. The primary oxidant and nitrogen oxide species are OH, $HO_2$, NO, and $NO_2$. $NO_3$ is not involved in any significant cycles and is excluded from the oxidant and nitrogen oxide weighting scheme, but it still participates in the mechanism. The organic oxidants methyl radical and peroxyacetyl radical are of lesser importance than the primary oxidants and nitrogen oxides and are, thus, given a lower weighting. NO, $NO_2$, OH, and $HO_2$ are all given a 7 % weighting for the overall accuracy. The methyl radical and peroxyacetyl radical are given a weighting of 2.5 % each for a total of 33 % for oxidants and nitrogen oxides. Ozone and formaldehyde are classified as pollutants, and both are given a weighting of 17 % for a total weighting of 34 %. The formaldehyde error is multiplied by the fraction of the maximum formaldehyde concentration for a given input condition over the average maximum formaldehyde concentration over all input conditions. This gives formaldehyde more weighting as its relative concentration increases.

According to Bates and Jacob (2019), the average isoprene SOA contribution is divided up into 33 % IEPOX, 30 % isoprene nitrates, 30 % tetrafunctional isoprene compounds, 2.5 % glyoxal, and 4.5 % other. Most small isoprene mechanisms exclude tetrafunctional compounds, leaving IEPOX as a 50 % contribution, isoprene nitrates as a 45 % contribution, and glyoxal as a 4.5 % contribution. As with formaldehyde, each of these are scaled relative to their average maximum concentration. Thus, in our calculations SOA contributes 33 % to the total accuracy, with IEPOX contributing 16.5 %, isoprene nitrates contributing 15 %, and glyoxal contributing 1.5 %. Isoprene is omitted from the error metric, as its error is represented by the accuracy in the other parameters. Methyl

**Table 5.** Species used in the calculation of the mechanism error metric and their corresponding weight.

| Species | Fractional contribution |
|---|---|
| OH | 0.07 |
| $HO_2$ | 0.07 |
| NO | 0.07 |
| $NO_2$ | 0.07 |
| Methyl radical ($MO_2$) | 0.025 |
| Peroxyacetyl radical ($ACO_3$) | 0.025 |
| HCHO | 0.17 |
| $O_3$ | 0.17 |
| IEPOX | 0.165 |
| Isoprene nitrates | 0.15 |
| Glyoxal | 0.015 |

vinyl ketone, methacrolein, peroxyacetyl nitrate, and methyl glyoxal were omitted from the accuracy metric due to their relatively lower importance compared with the other species and owing to their coupling to species already present in the error metric. However, these four species are represented in AMORE-Isoprene, and information on the performance of each mechanism with respect to these species and all other important species can be found in Sect. S13 and Table S6. Table 5 shows each species and its contribution to the total error metric.

The error metric is calculated by running box model simulations of the Caltech mechanism and the test mechanism under all six conditions, calculating each individual species error and averaging them using the weights shown in Table 5, and then averaging between each of the six conditions to arrive at a single value. The error metric ranges from zero to one, with lower values corresponding to less error. This allows for the numerical comparison of various isoprene mechanisms to the Caltech full mechanism.

## 3 Results and discussion

The final AMORE-Isoprene mechanism consists of 9 species and 22 reactions. A full outline of the reactions is shown in Table 4. The nine isoprene species were isoprene (ISO), isoprene hydroxy peroxy radical (ISOP), isoprene hydroxy peroxide (ISHP), isoprene nitrooxy peroxy radicals ($INO_2$), isoprene hydroxy nitrates (IHN), the lumped species IPC and IPN, isoprene epoxydiol (IEPOX), and lumped multifunctional isoprene nitrates (ISON). IPC and IPN are named based on the reactions they participate in, but they have no true analogues in the full mechanism, as they are used primarily to expand the range of outputs and cycle oxidants and nitrogen oxides. In the following sections, AMORE-Isoprene's performance will be compared in box model simulations to the Caltech full mechanism (Sect. 3.1), compared to chamber data (Sect. 3.2), and compared to the CRACMM-baseline mechanism in CMAQ simulations (Sect. 3.3).

### 3.1 Ambient box model simulations

Using F0AM box model simulations and the error metric defined in Sect. 2.4.4, we were able to demonstrate the high accuracy of the AMORE-Isoprene mechanism relative to other mechanisms of similar size. Formaldehyde and $HO_2$ were chosen as exemplary species for visual comparison, as they demonstrate the high performance of AMORE-Isoprene relative to other isoprene mechanisms.

Figure 4 shows the concentration of $HO_2$ under the six conditions listed in Table 3. For $HO_2$, AMORE-Isoprene has stronger agreement with the Caltech full mechanism than the RACM2 isoprene mechanism. Under low-$NO_x$ conditions, the steady-state concentration of $HO_2$ was 0.054 ppb for the Caltech full mechanism, 0.045 ppb for the AMORE-Isoprene mechanism, 0.042 ppb for the CB6r3 mechanism, and 0.026 ppb for the RACM2 mechanism. Under high-$NO_x$ conditions, all mechanisms had similar concentrations of $HO_2$. Under high-$O_3$ conditions, the steady-state concentration of $HO_2$ was 0.05 ppb for the Caltech full mechanism, 0.04 ppb for AMORE-Isoprene, 0.04 ppb for CB6r3, and 0.02 ppb for RACM2. Under high-$NO_3$ concentrations, steady-state $HO_2$ concentrations were low for all mechanisms, and the peak concentration was 0.029 ppb for the Caltech full mechanism, 0.017 ppb for AMORE-Isoprene, 0.049 ppb for CB6r3, and 0.016 ppb for RACM2. In all cases, the steady-state and/or peak concentrations were closer to the Caltech full mechanism for AMORE-Isoprene than for RACM2. Simulated $HO_2$ concentration profiles are similar between AMORE-Isoprene and CB6r3, although AMORE-Isoprene tends to underestimate $HO_2$ under high-$NO_3$ conditions, whereas CB6r3 tends to overestimate $HO_2$ under the aforementioned conditions. As expected, the Caltech Reduced Plus mechanism has strong agreement with the Caltech full mechanism for $HO_2$.

Figure 5 shows the simulated concentration of formaldehyde under the six conditions listed in Table 3. For formaldehyde, AMORE-Isoprene consistently outperforms RACM2. For example, under low-$NO_x$ conditions, the peak formaldehyde concentration was just over 3 ppb for the Caltech full mechanism, 2.6 ppb for the AMORE-Isoprene mechanism, 2.1 ppb for the CB6r3 mechanism, and 0.6 ppb for the RACM2 mechanism. In addition, the mechanism performs similarly to CB6r3. The accuracy tends to be higher than CB6r3 for the first few hours of the simulation, with less accuracy later in the simulation due to a steeper decay of the formaldehyde concentration. This pattern can be seen clearly in Fig. 5a. These differences arise from the sources of formaldehyde, where AMORE-Isoprene produces more formaldehyde from first-generation oxidation intermediates, whereas CB6r3 has a higher contribution from later-generation oxidation intermediates. Formaldehyde is over-

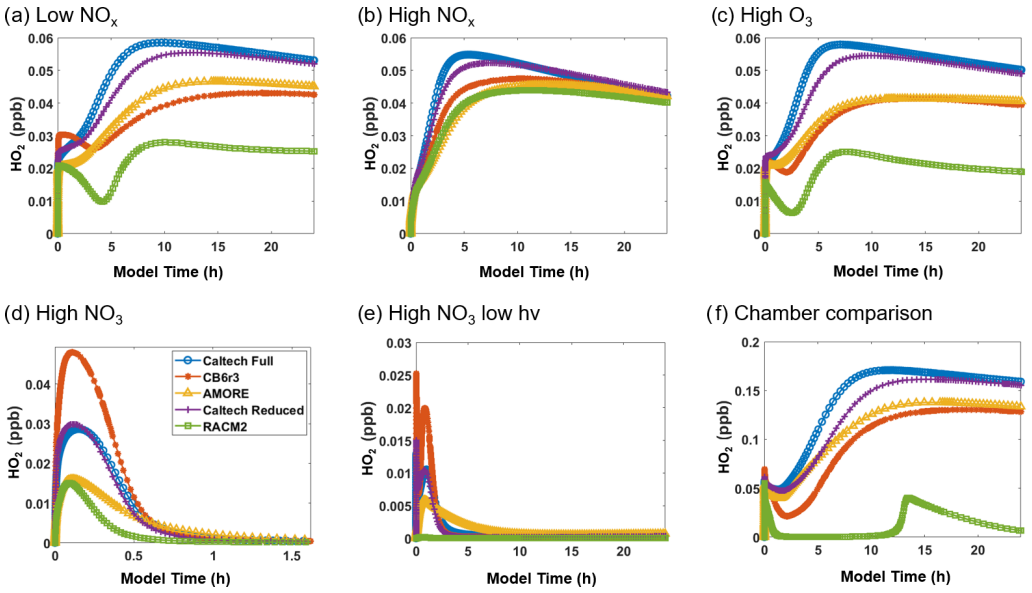


**Figure 4.** Box model predictions of $HO_2$ from multiple mechanisms (292 K and 1000 hPa) under the following conditions: **(a)** low $NO_x$; **(b)** high $NO_x$; **(c)** high $O_3$; **(d)** high $NO_3$; **(e)** high $NO_3$, low *hv*; and **(f)** Paulot et al. (2009) chamber.

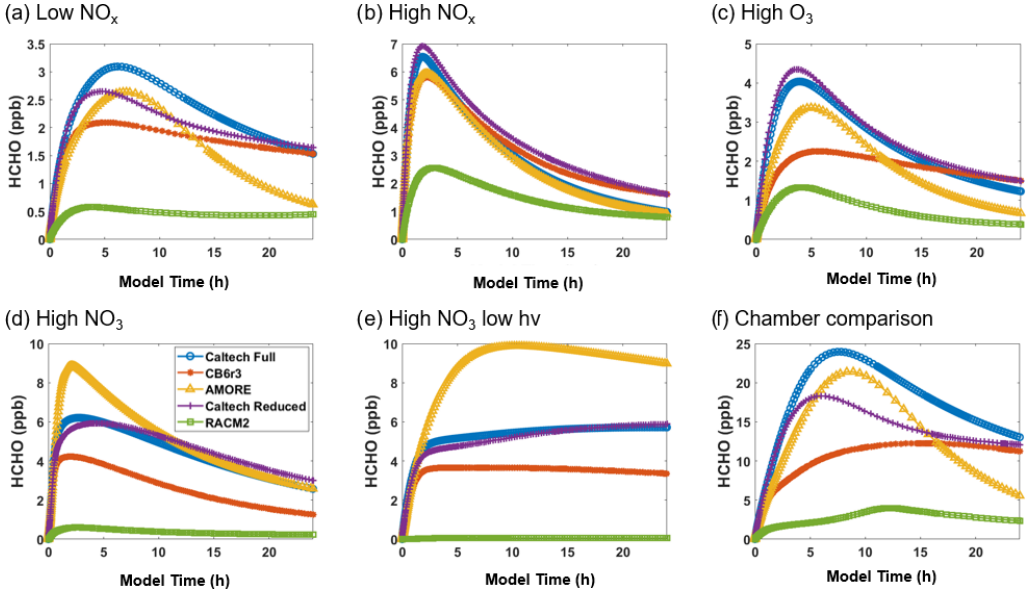


**Figure 5.** Box model predictions of formaldehyde from multiple mechanisms (292 K and 1000 hPa) under the following conditions: **(a)** low $NO_x$; **(b)** high $NO_x$; **(c)** high $O_3$; **(d)** high $NO_3$; **(e)** high $NO_3$, low *hv*; and **(f)** Paulot et al. (2009) chamber.

estimated under high-$NO_3$ conditions, although other small mechanisms significantly underestimate formaldehyde under the same conditions. This difference is primarily due to the slower rate of decay of formaldehyde under high-$NO_3$ conditions coupled with high initial production from the oxidation of isoprene. In contrast, the Caltech full mechanism displays a slower rate of formaldehyde production, as it is spread out over many more oxidation reactions at longer timescales. The other small reduced mechanisms (CB6r3 and RACM2) have lower formaldehyde production, due to fewer formaldehyde production pathways under high-$NO_3$ conditions, with RACM2 having over an order of magnitude lower formaldehyde concentrations.

Figure 6 shows the simulated concentration of the hydroxyl radical under the six conditions listed in Table 3. The AMORE-Isoprene mechanism performs similarly to other highly reduced mechanisms. As with other small mechanisms, AMORE-Isoprene is biased low compared with the

full Caltech mechanism. Under low-$NO_x$ conditions, the AMORE-Isoprene mechanism has near-equal behavior to CB6r3 and RACM2 at short time frames and has a more accurate steady-state value at longer time frames. Under high-$NO_x$ conditions, the RACM2 mechanism is the most accurate small mechanism, with AMORE-Isoprene having close but slightly lower OH concentrations. Under high-$O_3$ conditions, AMORE-Isoprene has the closest agreement with the Caltech full mechanism. The Caltech Reduced Plus mechanism has strong agreement with the full mechanism under all tested conditions, as would be expected. The main reason for the discrepancy in between AMORE-Isoprene and the Caltech full mechanism with respect to hydroxyl radical concentrations is that the Caltech full mechanism has a greater quantity of intermediate species which produce and consume the hydroxyl radical. This leads to slightly higher hydroxyl radical concentrations, and, given that the AMORE-Isoprene mechanism is a much smaller mechanism, there are limitations to the extent that this can be corrected. This is further evidenced by the fact that the other small mechanisms have similarly low-biased hydroxyl radical concentrations. Overall, the AMORE-Isoprene mechanism performs consistently well at predicting OH concentrations and is in line with similarly sized mechanisms in this regard.

In addition to these plots, a quantitative comparison was made between each of the mechanisms tested based on the overall mechanism error metric defined in Sect. 2.4.4. Figure 7 shows the mean accuracy of AMORE-Isoprene for a selection of species under each of the six simulation conditions. Lower values correspond to higher accuracy. The AMORE-Isoprene mechanism shows very high accuracy under all conditions, and it performed the best under high $NO_x$, high $O_3$, and in comparison to the chamber data from Paulot et al. (2009), which is relatively low with respect to $NO_x$, and OH oxidation dominates.

Table 6 shows the overall error of each reduced mechanism as defined in Sect. 2.4.4 with species weightings described in Table 5. The numerical errors shown in Table 6 represent weighted averages of the error across multiple priority species and the six conditions shown in Table 3. As a result, this error metric quantifies the overall performance of the mechanism in box model simulations, The results show that AMORE-Isoprene performs very well compared with mechanisms of a similar size. AMORE-Isoprene has an error of 0.17, which is much lower than that of CB6r3 (0.3) and RACM2 (0.44) but very close to that of Caltech Reduced Plus (0.13), the latter of which is a significantly larger mechanism. The main drivers of the low error for the AMORE-Isoprene mechanism are oxidant, nitrogen oxide, IEPOX, and formaldehyde concentrations. For example, the average error for IEPOX is 0.17 for AMORE-Isoprene compared with 0.27 for CB6r3 and 0.60 for RACM2. The average error for $NO_x$ species is 0.16 for AMORE-Isoprene, 0.29 for RACM2, and 0.35 for CB6r3. For $HO_x$ species, the average error is 0.25 for AMORE-Isoprene, 0.55 for RACM2, and 0.29 for CB6r3. For formaldehyde, the average error is 0.22 for AMORE-Isoprene, 0.79 for RACM2, and 0.3 for CB6r3. The error information for each mechanism can be found in Sect. S13 and Table S6. Additional box model plots can be found in Figs. S14–S19. These results validate the AMORE reduction process as a useful method of mechanism reduction and demonstrate that small mechanisms can retain significant accuracy compared to a much larger reference mechanism.

**Table 6.** Total error (individual species error described in Eq. 4, species weighting shown in Table 5, and further discussion in Sect. 2.4.4) and mechanism size for four reduced isoprene mechanisms with the Caltech full mechanism as a basis of comparison. Individual species error shown averaged over the six tested conditions.

| | AMORE | Caltech Reduced Plus | RACM2 | CB6r3 |
|---|---|---|---|---|
| Species | 12 | 131 | 9 | 10 |
| Reactions | 22 | 220 | 12 | 17 |
| Total error | 0.17 | 0.13 | 0.44 | 0.3 |
| O3 | 0.12 | 0.02 | 0.15 | 0.12 |
| NO | 0.12 | 0.06 | 0.22 | 0.28 |
| NO2 | 0.19 | 0.08 | 0.38 | 0.42 |
| HO | 0.20 | 0.17 | 0.44 | 0.30 |
| HO2 | 0.29 | 0.11 | 0.67 | 0.29 |
| NO3 | 0.36 | 0.09 | 0.47 | 0.25 |
| ISOP | 0.14 | 0.06 | 0.18 | 0.11 |
| IEPOX | 0.17 | 0.12 | 0.60 | 0.27 |
| HCHO | 0.22 | 0.11 | 0.79 | 0.30 |
| MO2 | 0.53 | 0.20 | 0.59 | 0.56 |
| ACO3 | 0.56 | 0.27 | 0.72 | 0.44 |
| PAN | 0.53 | 0.21 | 0.85 | 0.52 |
| ISOPN | 0.43 | 0.26 | 0.61 | 0.77 |
| GLY | 0.64 | 0.60 | 0.86 | 0.57 |
| MGLY | 0.63 | 0.14 | 0.79 | 0.23 |

A complete species list can be found in Table S1 in the Supplement.

### 3.2 Chamber box model simulations

In order to determine the accuracy of the Caltech full mechanism, which was augmented in this work, chamber data were used for comparison. These data come from Paulot et al. (2009) and contain concentration profiles for isoprene, isoprene hydroxy peroxides (ISHP), and IEPOX. The conditions of the chamber study were replicated using the F0AM box model to determine the accuracy of the Caltech full mechanism and the reduced isoprene mechanisms. As expected, the Caltech full mechanism matched the concentrations of all measured species from the chamber study.

Figure 8 shows the results for IEPOX. The AMORE-Isoprene mechanism is in good qualitative and quantitative agreement with the Caltech full mechanism and the Paulot et al. (2009) chamber data concentration profile for IEPOX. The peak IEPOX concentration is roughly 27 ppb according to the chamber data, compared to 25 ppb for the Caltech full mechanism, 25 ppb for the AMORE-Isoprene mech-

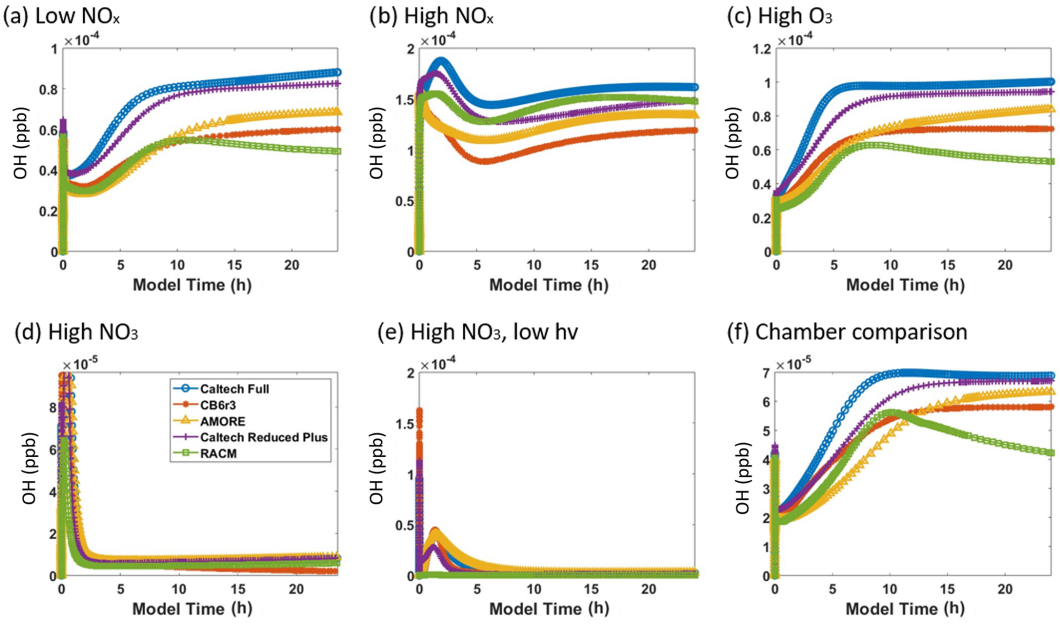


**Figure 6.** Box model predictions of the hydroxyl radical from multiple mechanisms (292 K and 1000 hPa) under the following conditions: **(a)** low $NO_x$; **(b)** high $NO_x$; **(c)** high $O_3$; **(d)** high $NO_3$; **(e)** high $NO_3$, low $h\nu$; and **(f)** Paulot et al. (2009) chamber.

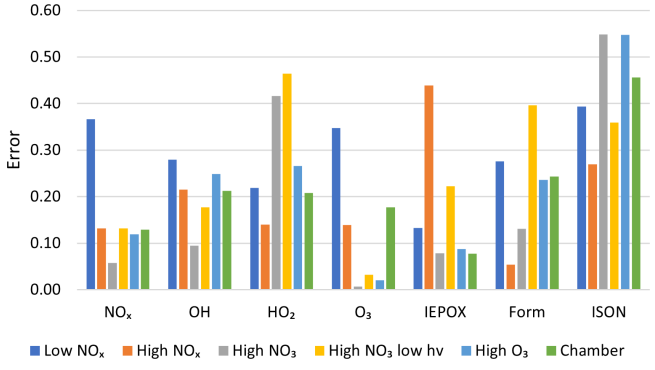


**Figure 7.** Measured error (Eq. 4) of the AMORE-Isoprene mechanism under six different conditions for seven select species groups. Errors are averaged between species for multiple-species groups.

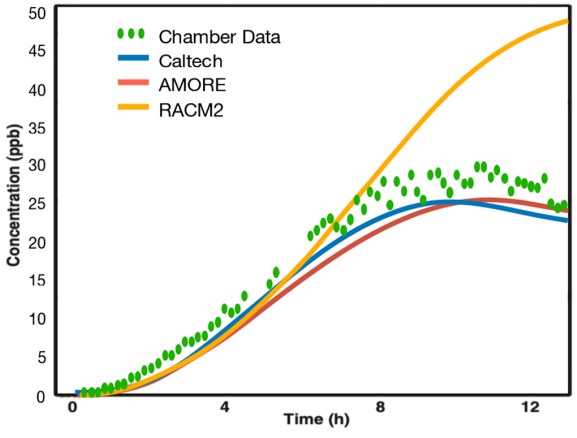


**Figure 8.** IEPOX concentration comparison between chamber data from Paulot et al. (2009) and F0AM box model simulations using the reported chamber conditions. The Caltech full mechanism closely matches the measured values, as does the AMORE-Isoprene mechanism.

anism, and 50 ppb for the RACM2 mechanism. In addition, the timing of the peak matches closely with the chamber data. The chamber data IEPOX peak occurs at around 10.5 h, compared to 9.5 h for the Caltech full mechanism, 10.5 h for the AMORE-Isoprene mechanism, and over 12 h for the RACM2 mechanism. Model evaluation with chamber data is particularly important for this species because (a) it is a key species for SOA formation from isoprene and (b) relatively few ambient measurements of IEPOX exist for validation.

### 3.3 CMAQ Testing

When AMORE-Isoprene is included in the CRACMM scheme of CMAQ (CRACMM1AMORE), improvements in $O_3$ and formaldehyde bias are observed compared with AQS ambient observations. Figure 9 shows the bias of CRACMM-baseline and AMORE-Isoprene compared with AQS data. Both formaldehyde and ozone observations were underestimated by CMAQ, particularly at higher concentrations of each species. AMORE-Isoprene predicted higher ozone and formaldehyde than CRACMM-baseline, thereby reducing the bias for both species.

The mean bias of formaldehyde decreased by 0.27 ppb with the implementation of AMORE-Isoprene, for all

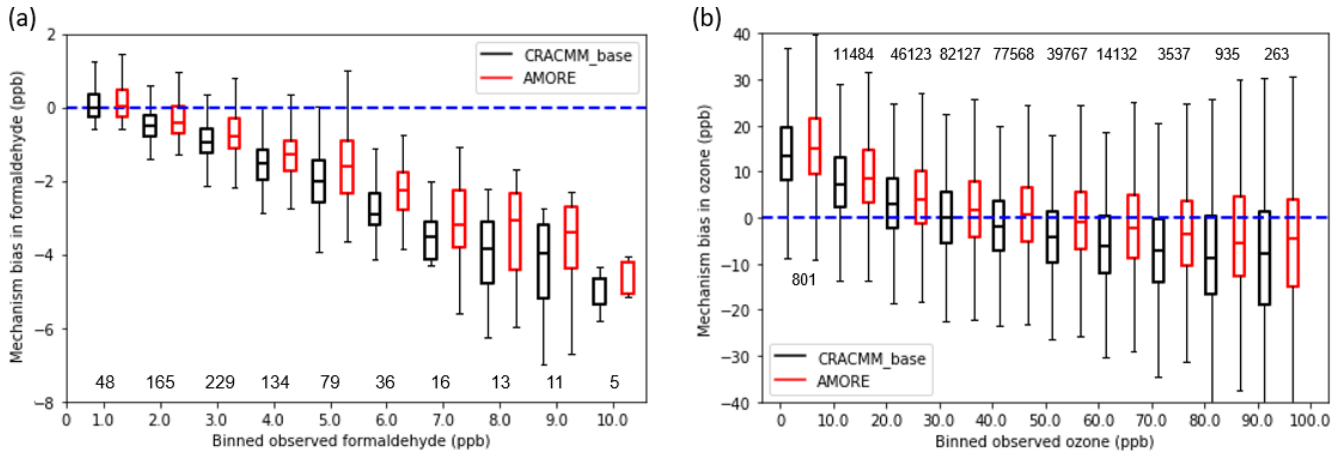


**Figure 9.** Binned mean bias of AMORE-Isoprene and baseline CRACMM for **(a)** formaldehyde and **(b)** ozone compared with AQS data for the northeastern US during summer 2018. Numbers for box plots indicate the number of data points in each observed range.

formaldehyde concentrations. The CRACMM-baseline simulation tended to underestimate formaldehyde, especially at increasing formaldehyde concentrations. In all concentrations, AMORE-Isoprene increased the simulated formaldehyde concentration by roughly 0.25 ppb. This is in line with the box model simulations shown in Fig. 5, where AMORE-Isoprene (yellow) had consistently higher formaldehyde concentrations than the RACM2 mechanism (green), which is used in the CRACMM-baseline isoprene mechanism, under a wide range of conditions.

The mean bias of ozone for concentrations above 50 ppb during the daytime was decreased by 3.4 ppb with the implementation of AMORE-Isoprene. AMORE-Isoprene slightly increased the bias at low ozone concentrations, as ozone is overpredicted at low concentrations. At higher concentrations, where health implications are presumably more serious, AMORE-Isoprene yielded significantly higher accuracy. AMORE-Isoprene generally tended to increase ozone concentrations by roughly 2–3 ppb for all ozone concentrations. AMORE-Isoprene tends to have higher ozone concentrations and better agreement with the Caltech full mechanism than RACM2 in low-$NO_x$ box model simulations. Thus, the difference may be attributable a higher prevalence of low-$NO_x$ conditions.

The CMAQ implementation also included heterogeneous chemistry for IEPOX and first-generation isoprene organic nitrates as well as deposition for all species. These processes, while not included in our box models, did not significantly impact the overall performance of the mechanism, as organic carbon (OC) values were similar between AMORE-Isoprene and the base CRACMM1 mechanism (see Sect. S15). No significant changes were identified for other observed species such as $NO_y$, and isoprene (Sect. S15). The CMAQ model runtime did not increase substantially with AMORE-Isoprene.

## 4 Conclusions

We have developed a new reduced isoprene oxidation mechanism for application in large-scale atmospheric models, using a novel, semiautomated, graph-theory-based approach. Rigorous testing has demonstrated that the AMORE-Isoprene mechanism's performance is very good for its size, with improved accuracy compared with CB6r3 and RACM2.

A small, accurate isoprene oxidation mechanism would improve the performance of many large-scale models, as we have demonstrated with CMAQ-CRACMM1AMORE simulations, where there was a noticeable improvement in both ozone and formaldehyde bias. In the future, we plan additional testing of AMORE-Isoprene in other chemical transport models to characterize the impacts of this mechanism more broadly.

During the algorithmic and manual adjustment process, several useful concepts were developed. First, for a small number of desired measurable outputs, small mechanisms can reach high levels of accuracy if properly structured and optimized. Second, the optimization of oxidants and nitrogen oxides, which are highly coupled to the isoprene mechanism, takes precedence over the optimization of other species, as inaccuracies in coupled species ultimately propagate to uncoupled species. In addition, the observation that methods reliant on removing aspects of the full mechanism would not work for this application was very important. The path-based approach that we have developed to "summarize" the mechanism may be a more sensible starting point for the reduction of other atmospheric reaction networks as well.

The AMORE-Isoprene mechanism demonstrates that there is a significant potential advantage in the use of algorithms for model reduction. Additional development, informed by the experiences of this study, is underway to more fully automate the model reduction process and further reduce the need for manual adjustments. Future work will ex-

tend this work to application to the reduction of a wide range of atmospheric chemical mechanisms in addition to the isoprene oxidation mechanism.

*Code and data availability.* CMAQv5.3.3 is available at https://github.com/USEPA/CMAQ (last access: 16 March 2023) and is archived at https://doi.org/10.5281/zenodo.5213949 (US EPA Office of Research and Development, 2021). The exact CMAQ code used in this work and the CMAQ output are available at https://doi.org/10.23719/1527975 (Pye, 2023).

Code and data for the AMORE algorithm are available at https://github.com/fcw2110/AMORE_supplementary_files (last access: 16 March 2023) and are archived at https://doi.org/10.5281/zenodo.7106505 (fcw2110, 2022).

*Supplement.* The supplement related to this article is available online at: https://doi.org/10.5194/gmd-16-1801-2023-supplement.

*Author contributions.* All authors contributed to writing the manuscript. FW, SS, and VFM developed the model reduction algorithm. FW, BP, SS, HOTP, and VFM developed the reduced model. FW, BP, and HOTP performed simulations. The Community Regional Atmospheric Chemistry Multiphase Mechanism was conceived by HOTP. Research was conceived by FW, VFM, SS, DMW, DKH, and AMF.

*Competing interests.* At least one of the (co-)authors is a member of the editorial board of *Geoscientific Model Development*. The peer-review process was guided by an independent editor, and the authors also have no other competing interests to declare.

*Disclaimer.* The US Environmental Protection Agency (EPA) does not endorse any products or commercial services mentioned in this publication. All opinions expressed in this paper are the author's and do not necessarily reflect the policies and views of US EPA, DOE, or ORAU/ORISE.

Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

*Acknowledgements.* The authors would like to thank Kelvin Bates and Glenn Wolfe for helpful discussions. We are also grateful to Jon Pleim, Ana Torres-Vazquez, and Christine Allen for assistance with CMAQ simulation inputs.

*Financial support.* This publication was developed under assistance agreement no. 84001301 awarded by the US EPA to V. Faye McNeill, Daniel M. Westervelt, Daven K. Henze, and Arlene M. Fiore. This work was supported in part by the EPA Office of Research and Development. This research was also supported in part by an appointment to the EPA Research Participation Program administered by the Oak Ridge Institute for Science and Education (ORISE) through an interagency agreement between the US Department of Energy (DOE) and the US EPA. ORISE is managed by ORAU under DOE contract no. DE-SC0014664.

*Review statement.* This paper was edited by Rolf Sander and reviewed by two anonymous referees.

## References

Appel, K. W., Gilliam, R. C., Davis, N., Zubrow, A., and Howard, S. C.: Overview of the atmospheric model evaluation tool (AMET) v1.1 for evaluating meteorological and air quality models, Environ. Modell. Softw., 26, 434–443, https://doi.org/10.1016/j.envsoft.2010.09.007, 2011.

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, https://doi.org/10.5194/gmd-14-2867-2021, 2021.

Aumont, B., Szopa, S., and Madronich, S.: Modelling the evolution of organic carbon during its gas-phase tropospheric oxidation: development of an explicit model based on a self generating approach, Atmos. Chem. Phys., 5, 2497–2517, https://doi.org/10.5194/acp-5-2497-2005, 2005.

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, https://doi.org/10.5194/gmd-9-2191-2016, 2016.

Bates, K. H. and Jacob, D. J.: A new model mechanism for atmospheric oxidation of isoprene: global effects on oxidants, nitrogen oxides, organic products, and secondary organic aerosol, Atmos. Chem. Phys., 19, 9613–9640, https://doi.org/10.5194/acp-19-9613-2019, 2019.

Butler, T. M., Taraborrelli, D., Brühl, C., Fischer, H., Harder, H., Martinez, M., Williams, J., Lawrence, M. G., and Lelieveld, J.: Improved simulation of isoprene oxidation chemistry with the ECHAM5/MESSy chemistry-climate model: lessons from the GABRIEL airborne field campaign, Atmos. Chem. Phys., 8, 4529–4546, https://doi.org/10.5194/acp-8-4529-2008, 2008.

Carlton, A. G., Wiedinmyer, C., and Kroll, J. H.: A review of Secondary Organic Aerosol (SOA) formation from isoprene, Atmos. Chem. Phys., 9, 4987–5005, https://doi.org/10.5194/acp-9-4987-2009, 2009.

Emery, C., Jung, J., Koo, B., and Yarwood, G.: Improvements to CAMx snow cover treatments and Carbon Bond chemical mechanism for winter ozone, Final Report for Utah Department of Environmental Quality, Division of Air Quality, Salt Lake City, UT, August 2015, http://www.camx.com/files/udaq_snowchem_final_6aug15.pdf (last access: 17 March 2023), 2015.

EPA: Technical support document EPA's 2014 national air toxics assessment, https://www.epa.gov/sites/default/files/2018-09/

documents/2014_nata_technical_support_document.pdf (last access: 17 March 2023), 2018.

Farmer, D. K., Matsunaga, A., Docherty, K. S., Surratt, J. D., Seinfeld, J. H., Ziemann, P. J., and Jimenez, J. L.: Response of an aerosol mass spectrometer to organonitrates and organosulfates and implications for atmospheric chemistry, P. Natl. Acad. Sci. USA, 107, 6670–6675, https://doi.org/10.1073/pnas.0912340107, 2010.

fcw2110: fcw2110/AMORE_supplementary_files: AMORE_supplementary_files (amore_isoprene), Zenodo [data set and code], https://doi.org/10.5281/zenodo.7106505, 2022.

Fiore, A., Oberman, J., Lin, M., Zhang, L., Clifton, O., Jacob, D., Naik, V., Horowitz, L., Pinto, J., and Milly, G.: Estimating North American background ozone in U.S. surface air with two independent global models: Variability, uncertainties, and recommendations, Atmos. Environ., 96, 284–300, https://doi.org/10.1016/j.atmosenv.2014.07.045, 2014.

Fiore, A. M., Horowitz, L. W., Purves, D. W., Levy II, H., Evans, M. J., Wang, Y., Li, Q., and Yantosca, R. M.: Evaluating the contribution of changes in isoprene emissions to surface ozone trends over the eastern United States, J. Geophys. Res.-Atmos., 110, D12303, https://doi.org/10.1029/2004JD005485, 2005.

Fiore, A. M., Levy II, H., and Jaffe, D. A.: North American isoprene influence on intercontinental ozone pollution, Atmos. Chem. Phys., 11, 1697–1710, https://doi.org/10.5194/acp-11-1697-2011, 2011.

Fu, T.-M., Jacob, D. J., and Heald, C. L.: Aqueous-phase reactive uptake of dicarbonyls as a source of organic aerosol over eastern North America, Atmos. Environ., 43, 1814–1822, https://doi.org/10.1016/j.atmosenv.2008.12.029, 2009.

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

Guo, J. J., Fiore, A. M., Murray, L. T., Jaffe, D. A., Schnell, J. L., Moore, C. T., and Milly, G. P.: Average versus high surface ozone levels over the continental USA: model bias, background influences, and interannual variability, Atmos. Chem. Phys., 18, 12123–12140, https://doi.org/10.5194/acp-18-12123-2018, 2018.

Henze, D. K. and Seinfeld, J. H.: Global secondary organic aerosol from isoprene oxidation, Geophys. Res. Lett., 33, L09812, https://doi.org/10.1029/2006GL025976, 2006.

Jenkin, M., Khan, M., Shallcross, D., Bergström, R., Simpson, D., Murphy, K., and Rickard, A.: The CRI v2.2 reduced degradation scheme for isoprene, Atmos. Environ., 212, 172–182, https://doi.org/10.1016/j.atmosenv.2019.05.055, 2019.

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, https://doi.org/10.5194/acp-15-11433-2015, 2015.

Kelp, M. M., Jacob, D. J., Lin, H., and Sulprizio, M. P.: An Online-Learned Neural Network Chemical Solver for Stable Long-Term Global Simulations of Atmospheric Chemistry, J. Adv. Model. Earth Sy., 14, e2021MS002926, https://doi.org/10.1029/2021MS002926, 2022.

Kroll, J. H., Ng, N. L., Murphy, S. M., Flagan, R. C., and Seinfeld, J. H.: Secondary Organic Aerosol Formation from Isoprene Photooxidation, Environ. Sci. Technol., 40, 1869–1877, https://doi.org/10.1021/es0524301, pMID: 16570610, 2006.

Lin, H., Long, M. S., Sander, R., Sandu, A., Yantosca, R. M., Estrada, L. A., Shen, L., and Jacob, D. J.: An adaptive auto-reduction solver for speeding up integration of chemical kinetics in atmospheric chemistry models: Implementation and evaluation in the Kinetic Pre-Processor (KPP) version 3.0.0, J. Adv. Model. Earth Sy., 15, e2022MS003293, https://doi.org/10.1029/2022MS003293, 2023.

Liu, J., D'Ambro, E. L., Lee, B. H., Lopez-Hilfiker, F. D., Zaveri, R. A., Rivera-Rios, J. C., Keutsch, F. N., Iyer, S., Kurten, T., Zhang, Z., Gold, A., Surratt, J. D., Shilling, J. E., and Thornton, J. A.: Efficient Isoprene Secondary Organic Aerosol Formation from a Non-IEPOX Pathway, Environ. Sci. Technol., 50, 9872–9880, https://doi.org/10.1021/acs.est.6b01872, 2016.

Lu, T. and Law, C. K.: A directed relation graph method for mechanism reduction, P. Combust. Inst., 30, 1333–1341, https://doi.org/10.1016/j.proci.2004.08.145, 2005.

Lu, T., Ju, Y., and Law, C. K.: Complex CSP for chemistry reduction and analysis, Combust. Flame, 126, 1445–1455, https://doi.org/10.1016/S0010-2180(01)00252-8, 2001.

Mansouri, K., Grulke, C. M., Judson, R. S., and Williams, A. J.: OPERA models for predicting physicochemical properties and environmental fate endpoints, J. Cheminformatics, 10, 10, https://doi.org/10.1186/s13321-018-0263-1, 2018.

Massias, A., Diamantis, D., Mastorakos, E., and Goussis, D.: An algorithm for the construction of global reduced mechanisms with CSP data, Combust. Flame, 117, 685–708, https://doi.org/10.1016/S0010-2180(98)00132-1, 1999.

Muñoz, A.: Isoprene+NO+hydrogen peroxide + OH – Gas-phase oxidation – product study, Tech. rep., CEAM, https://data.eurochamp.org/data-access/chamber-experiments/25320384-8599-4bc9-a35d-993a77cec7c (last access: 1 March 2023), 2021a.

Muñoz, A.: Isoprene+ozone+carbon monoxide + O3 – Gas-phase oxidation – product study, Tech. rep., CEAM, https://data.eurochamp.org/data-access/chamber-experiments/25320384-8599-4bc9-a35d-993a77cec7c (last access: 1 March 2023), 2021b.

Muñoz, A. and Gómez-Alvarez, E.: Overview Of The Eurochamp Database Of European Atmosphere Simulation Chambers, in: Simulation and Assessment of Chemical Processes in a Multiphase Environment, Springer Netherlands, 61–70, ISBN 9781402088469, 2008.

Nikolaou, Z. M., Chen, J.-Y., Proestos, Y., Lelieveld, J., and Sander, R.: Accelerating simulations using REDCHEM_v0.0 for atmospheric chemistry mechanism reduction, Geosci. Model Dev., 11, 3391–3407, https://doi.org/10.5194/gmd-11-3391-2018, 2018.

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, https://doi.org/10.5194/gmd-3-243-2010, 2010.

Paulot, F., Crounse, J. D., Kjaergaard, H. G., Kürten, A., St. Clair, J. M., Seinfeld, J. H., and Wennberg, P. O.: Unexpected Epoxide Formation in the Gas-Phase Photooxidation of Isoprene, Science, 325, 730–733, https://doi.org/10.1126/science.1172910, 2009.

Pepiot-Desjardins, P. and Pitsch, H.: An efficient error-propagation-based reduction method for large chem-

ical kinetic mechanisms, Combust. Flame, 154, 67–81, https://doi.org/10.1016/j.combustflame.2007.10.020, 2008.

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 Northeast U.S. surface ozone predictions to the representation of atmospheric chemistry in CRACMMv1.0, EGUsphere [preprint], https://doi.org/10.5194/egusphere-2023-288, 2023.

Pleim, J. E., Ran, L., Appel, W., Shephard, M. W., and Cady-Pereira, K.: New Bidirectional Ammonia Flux Model in an Air Quality Model Coupled With an Agricultural Model, J. Adv. Model. Earth Sy., 11, 2934–2957, https://doi.org/10.1029/2019MS001728, 2019.

Pye, H. O. T.: Data for AMORE-Isoprene v1.0: A new reduced mechanism for gas-phase isoprene oxidation, EPA Environmental Dataset Gateway [data set], https://doi.org/10.23719/1527975, 2023.

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, Environ. Sci. Technol., 47, 11056–11064, https://doi.org/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, https://doi.org/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) version 1.0, Atmos. Chem. Phys. Discuss. [preprint], https://doi.org/10.5194/acp-2022-695, in review, 2022.

Ratkiewicz, A. and Truong, T. N.: Application of Chemical Graph Theory for Automated Mechanism Generation, J. Chem. Inf. Comp. Sci., 43, 36–44, https://doi.org/10.1021/ci020297f, 2003.

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, https://doi.org/10.5194/acp-13-9695-2013, 2013.

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), Environ. Sci. Technol., 50, 12356–12364, https://doi.org/10.1021/acs.est.6b04752, 2016.

Schwantes, R. H., Teng, A. P., Nguyen, T. B., Coggon, M. M., Crounse, J. D., St. Clair, J. M., Zhang, X., Schilling, K. A., Seinfeld, J. H., and Wennberg, P. O.: Isoprene $NO_3$ Oxidation Products from the $RO_2$ + $HO_2$ Pathway, J. Phys. Chem. A, 119, 10158–10171, https://doi.org/10.1021/acs.jpca.5b06355, 2015.

Seltzer, K. M., Pennington, E., Rao, V., Murphy, B. N., Strum, M., Isaacs, K. K., and Pye, H. O. T.: Reactive organic carbon emissions from volatile chemical products, Atmos. Chem. Phys., 21, 5079–5100, https://doi.org/10.5194/acp-21-5079-2021, 2021.

Silva, S. J., Burrows, S. M., Evans, M. J., and Halappanavar, M.: A Graph Theoretical Intercomparison of Atmospheric Chemical Mechanisms, Geophys. Res. Lett., 48, e2020GL090481, https://doi.org/10.1029/2020GL090481, 2021.

Sturm, O.: Advecting Superspecies: Reduced order modeling of organic aerosols in LOTOS-EUROS using machine learning, Master's thesis, Delft Netherlands, http://resolver.tudelft.nl/uuid:2c3be50e-5340-4495-a0b7-1670db9be329 (last access: 17 March 2023), 2021.

Sun, W., Chen, Z., Gou, X., and Ju, Y.: A path flux analysis method for the reduction of detailed chemical kinetic mechanisms, Combus. Flame, 157, 1298–1307, https://doi.org/10.1016/j.combustflame.2010.03.006, 2010.

Teng, A. P., Crounse, J. D., and Wennberg, P. O.: Isoprene Peroxy Radical Dynamics, J. Ame. Chem. Soc., 139, 5367–5377, https://doi.org/10.1021/jacs.6b12838, 2017.

Tomlin, A. S., Pilling, M. J., Turányi, T., Merkin, J. H., and Brindley, J.: Mechanism reduction for the oscillatory oxidation of hydrogen: Sensitivity and quasi-steady-state analyses, Combust. Flame, 91, 107–130, https://doi.org/10.1016/0010-2180(92)90094-6, 1992.

Tomlin, A. S., Turányi, T., and Pilling, M. J.: Chapter 4 Mathematical tools for the construction, investigation and reduction of combustion mechanisms, in: Low-Temperature Combustion and Autoignition, edited by: Pilling, M., vol. 35 of Comprehensive Chemical Kinetics, Elsevier, 293–437, https://doi.org/10.1016/S0069-8040(97)80019-2, 1997.

Torres-Vazquez, A., Pleim, J., Gilliam, R., and Pouliot, G.: Performance Evaluation of the Meteorology and Air Quality Conditions From Multiscale WRF-CMAQ Simulations for the Long Island Sound Tropospheric Ozone Study (LISTOS), J. Geophys. Res.-Atmos., 127, e2021JD035890, https://doi.org/10.1029/2021JD035890, 2022.

Travis, K. R., Judd, L. M., Crawford, J. H., Chen, G., Szykman, J., Whitehill, A., Valin, L. C., Spinei, E., Janz, S., Nowlan, C. R., Kwon, H.-A., Fried, A., and Walega, J.: Can Column Formaldehyde Observations Inform Air Quality Monitoring Strategies for Ozone and Related Photochemical Oxidants?, J. Geophys. Res.-Atmos., 127, e2022JD036638, https://doi.org/10.1029/2022JD036638, 2022.

US EPA Office of Research and Development: CMAQ (5.3.3), Zenodo [code], https://doi.org/10.5281/zenodo.5213949, 2021.

Vasquez, K. T., Crounse, J. D., Schulze, B. C., Bates, K. H., Teng, A. P., Xu, L., Allen, H. M., and Wennberg, P. O.: Rapid hydrolysis of tertiary isoprene nitrate efficiently removes $NO_x$ from the atmosphere, P. Natl. Acad. Sci. USA, 117, 33011–33016, https://doi.org/10.1073/pnas.2017442117, 2020.

Watson, L., Shallcross, D., Utembe, S., and Jenkin, M.: A Common Representative Intermediates (CRI) mechanism for VOC degradation. Part 2: Gas phase mechanism reduction, Atmos. Environ., 42, 7196–7204, https://doi.org/10.1016/j.atmosenv.2008.07.034, 2008.

Wei, J. and Kuo, J. C. W.: Lumping Analysis in Monomolecular Reaction Systems. Analysis of the Exactly Lumpable System, Ind. Eng. Chem. Fund., 8, 114–123, https://doi.org/10.1021/i160029a019, 1969.

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, Chem. Rev., 118, 3337–3390, https://doi.org/10.1021/acs.chemrev.7b00439, pMID: 29522327, 2018.

Whitehouse, L. E., Tomlin, A. S., and Pilling, M. J.: Systematic reduction of complex tropospheric chemical mechanisms, Part I: sensitivity and time-scale analyses, Atmos. Chem. Phys., 4, 2025–2056, https://doi.org/10.5194/acp-4-2025-2004, 2004a.

Whitehouse, L. E., Tomlin, A. S., and Pilling, M. J.: Systematic reduction of complex tropospheric chemical mechanisms, Part II: Lumping using a time-scale based approach, Atmos. Chem. Phys., 4, 2057–2081, https://doi.org/10.5194/acp-4-2057-2004, 2004b.

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, https://doi.org/10.5194/gmd-9-3309-2016, 2016.

Xia, A. G., Michelangeli, D. V., and Makar, P. A.: Mechanism reduction for the formation of secondary organic aerosol for integration into a 3-dimensional regional air quality model: $\alpha$-pinene oxidation system, Atmos. Chem. Phys., 9, 4341–4362, https://doi.org/10.5194/acp-9-4341-2009, 2009.

Yarwood, G., Rao, S., Yocke, M., and Whitten, G.: Updates to the carbon bond chemical mechanism: CB05, Final report to the US EPA, RT-0400675, 8, 246 pp., 2005.

Yarwood, G., Jung, J., Whitten, G. Z., Heo, G., Mellberg, J., and Estes, M.: Updates to the Carbon Bond mechanism for version 6 (CB6), in: 9th Annual CMAS Conference, Chapel Hill, NC, 11–13, https://cmascenter.org/conference/2010/abstracts/emery_updates_carbon_2010.pdf (last access: 17 March 2023), 2010.

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, Environ. Sci. Technol., 51, 5650–5657, https://doi.org/10.1021/acs.est.7b01356, 2017.