GECKO-A v1.0: Exploring VOC Oxidation Trajectories Through Comparison with the Master Chemical Mechanism

Bernard Aumont[1], Richard Valorso[1], Andrew Rickard[2,3], Mike Jenkin[4], Julia Lee-Taylor[5], John Orlando[5], Sasha Madronich[5], Luc Vereecken[6], Marie Camredon[1]

[1] Univ Paris Est Creteil and Université Paris Cité, CNRS, LISA, F-94010 Créteil, France
[2] Wolfson Atmospheric Chemistry Laboratories, Department of Chemistry, University of York, York, UK
[3] National Centre for Atmospheric Science, Department of Chemistry, University of York, York, UK
[4] Atmospheric Chemistry Services, Okehampton, Devon, EX20 4QB, UK
[5] Atmospheric Chemistry Observations and Modeling laboratory, NSF National Center for Atmospheric Research, Boulder, CO, 80307
[6] Institute of Climate and Energy Systems ICE-3: Troposphere, Forschungszentrum Jülich GmbH, 52425 Jülich, Germany

Correspondence to: Bernard Aumont (bernard.aumont@lisa.ipsl.fr), Richard Valorso (richard.valorso@lisa.ipsl.fr)

Abstract. Numerical models are crucial tools for understanding complex chemical systems such as the atmosphere, and their sensitivity across a range of conditions. In atmospheric chemistry models, reaction mechanisms are used to represent chemical transformations and define the underlying system of equations. Building highly explicit mechanisms that capture the full complexity of organic oxidation occurring in the atmosphere remains challenging owing to the large number of intermediates involved, the breadth of reaction pathways, and the limited availability of reliable kinetic and thermodynamic data. The Generator for Explicit Chemistry and Kinetics of Organics in the Atmosphere (GECKO-A) was developed to address these limitations by enabling the systematic construction of near-explicit mechanisms. Here, we present its first open-source release (v1.0), which incorporates updated chemical protocols and structure-activity relationships, together with its companion box model for mechanism integration. GECKO-A's performance is ~~evaluated~~benchmarked through systematic comparisons with the Master Chemical Mechanism (MCM v3.3.1), based on simulations of the oxidation of five representative hydrocarbons (butane, octane, dodecane, toluene, and α-pinene) under environmental conditions ranging from urban to remote. The two approaches yield similar oxidation pathways for small and structurally simple compounds. However, differences increase with the size and complexity of the carbon backbone. In particular, the simplifications inherent to the MCM tend to limit the formation of multifunctional products and promote earlier fragmentation, resulting in notable discrepancies in the predicted volatility of secondary organic carbon and, consequently, in secondary organic aerosol yields. These discrepancies stem from how each mechanism treats multigenerational oxidation. The two approaches remain similar for early-generation products, but diverge as later generations become predominant. This divergence also extends to the water solubility of secondary organic carbon, and to its partitioning into aqueous condensed phases. GECKO-A also enables mechanism generation at different levels of chemical resolution. A reduced mechanism, denoted red-GECKO-A, is compared to the high-resolution mechanism used throughout this study. It closely reproduces the simulated oxidative trajectories for most parent compounds and scenarios. Discrepancies emerge for octane and dodecane under NOx-poor conditions.

1 Introduction

Primary volatile organic compounds (VOCs) emitted into the atmosphere from both biogenic and anthropogenic sources are removed through oxidation processes. These transformations lead to the formation of secondary pollutants, including ozone, oxygenated compounds and secondary organic aerosol, with important implications for air quality, visibility, and climate (e.g., Hallquist et al., 2009; Monks et al., 2015; Pye et al., 2023). Atmospheric oxidation proceeds through successive steps involving a large number of intermediate species (e.g., Atkinson and Arey, 2003; Calvert et al., 2015; Goldstein and Galbally, 2007), ultimately leading to the removal of organic carbon via deposition or its conversion to CO and CO_2 (e.g., Hallquist et al., 2009). This system is inherently complex due to the diversity of emitted compounds, the influence of environmental conditions on reaction pathways, and the interplay between gas-phase chemistry and processes occurring in condensed phases, both particulate and aqueous (e.g., Seinfeld and Pandis, 2016). Improving our understanding of these processes requires models that explicitly represent multiphase oxidation mechanisms and can be evaluated against observations (e.g., Ervens et al., 2024; Heald and Kroll, 2020).

In atmospheric models, chemical transformations are described through reaction mechanisms. These mechanisms consist of sets of reactions that define the system of ordinary differential equations to be solved to describe the time-dependent transformations of the species involved. Developing explicit mechanisms, which aim to represent all relevant physicochemical steps in VOC oxidation, remains highly challenging. This difficulty arises from both the large number of species and reactions involved and the need to assign reliable kinetic and thermodynamic parameters to each process (Ervens et al., 2024; Vereecken et al., 2018).

The Master Chemical Mechanism (MCM) is a semi-explicit mechanism developed over several decades to describe the gas-phase oxidation of VOCs (Jenkin et al., 1997, 2015; Saunders et al., 2003). Originally designed to model multi-day ozone formation over NW Europe, it has become a widely used reference within the atmospheric chemistry community. However, its manual construction is extremely labor-intensive and limits the ability to incorporate updates or perform systematic sensitivity analyses. To overcome these limitations, automated mechanism generators have been developed. Among these, two tools currently under active development are the SAPRC atmospheric chemical mechanism generation system (MechGen) (Carter et al., 2025a, b) and the Generator for Explicit Chemistry and Kinetics of Organics in the Atmosphere (GECKO-A) (Aumont et al., 2005; Camredon et al., 2007).

In this paper, we present the latest version of the GECKO-A generator, released as open-source software (GECKO-A v1.0) (Aumont et al., 2025a), along with its companion box model for mechanism integration (Aumont et al., 2025b). Initially designed to simulate the gas-phase oxidation of small (C<8) aliphatic and acyclic hydrocarbons~~,~~ (Aumont et al., 2005), GECKO-A has progressively evolved to include phase partitioning processes, mechanism reduction strategies, especially to deal with VOCs having a large carbon backbone, and the treatment of additional families of compounds such as monoterpenes and monocyclic aromatics~~.~~ (Valorso et al., 2011; Jenkin et al., 2018b). A chronology of updates to the structure-activity relationships (SARs) used in GECKO-A is given in Sect. 2.1.

The GECKO-A generator and its box model are coded in Fortran 90. The code has recently been refactored into modular components, each dedicated to specific chemical processes or technical functions such as input-output handling. The codebase

increasingly embraces modern software development practices, including version control with Git, and is compatible with major Fortran compilers such as Intel ifort and GNU gfortran. It has been tested across multiple high-performance computing platforms, ensuring robustness and reproducibility.

GECKO-A v1.0 and its companion box model are then applied to explore oxidation trajectories under different environmental conditions for a selected set of hydrocarbons (butane, octane, dodecane, toluene, and α-pinene). The chemical processes represented in GECKO-A v1.0 are conceptually similar to those implemented in the latest version of the MCM (version 3.3.1, Jenkin et al., 2015). A systematic comparison of oxidation trajectories obtained with both mechanisms is therefore conducted to analyze their respective strengths and limitations across different atmospheric conditions. In this comparison, the MCM is used as a benchmark. When the simulated trajectories diverge, the aim is to determine the origin of the discrepancy. It may reflect the added value of GECKO-A's higher-resolution explicit treatment relative to the MCM's necessary simplifications. It may instead point to limitations specific to GECKO-A. This distinction guides the interpretation of results throughout the paper.

A set of parameters is available to control the chemical resolution of the generated mechanisms. A parameter set with more stringent reduction values, denoted red-GECKO-A, is thus used to generate lower-resolution mechanisms. These mechanisms are compared with the high-resolution GECKO-A mechanisms in Sect. 6. This comparison aims to illustrate, through this example, the reduction capabilities offered by GECKO-A and the magnitude of the discrepancies that such a reduction can introduce relative to the high-resolution mechanism.

2 The GECKO-A tool

For an organic precursor (or a list of organic precursors) provided as input, the GECKO-A tool generates as outputs (1) a detailed/explicit gaseous chemical mechanism until the full oxidation of the carbon, (2) organic physicochemical properties for mass transfer of the precursor as well as for the secondary organic products formed during the oxidation and (3) their gas to particle and/or gas to wall dynamic partitioning. The general principles underlying the generation of oxidation mechanisms for organic compounds are described in Aumont et al. (2005) as well as summarized in several review articles (e.g., Ervens et al., 2024; Vereecken et al., 2018). In general, mechanism generators can be viewed as computational tools that reproduce the process of chemical mechanism development as performed by chemists. Two main steps can be identified: (1) listing all possible reaction pathways based on the structure of the organic compound, and (2) determining rate constants and reaction products using structure-activity relationships (SARs) when experimental data ~~is~~are not available. This procedure is repeated for each newly formed species. A mechanism construction protocol is defined beforehand to systematically specify the reaction pathways to be considered for each possible organic compound and to select optional mechanism size limitation strategies. GECKO-A is a computer program that automatically generates chemical mechanisms based on this predefined protocol.

The GECKO-A mechanism generation system handles only the chemistry of species containing at least two carbon atoms. GECKO-A uses a predefined set of reactions and reaction rates for inorganic species and C_1 hydrocarbons, available in the mch_singlec.dat (for methane chemistry) and mch_inorg.dat files in the DATA directory of the distributed version (bibliographic references are included within these files). Inorganic species are not covered by the SARs described below. For C_1 species, it is well established that their reactivity is unique and typically does not follow the reactivity trends of larger

organic species. Predefining the chemistry for these critical species avoids biasing the mechanism through inappropriate application of SARs. Less common C_1 species formed during the oxidation of more complex compounds (e.g., aromatics) are automatically added only if they are produced during mechanism generation. Similarly, whenever experimental kinetic or product data are available for larger organic species, they are given priority over SAR-based estimates during mechanism generation. These data are compiled in dedicated files within the DATA directory of the distributed version, organized by reaction type (e.g., VOC+OH, VOC+NO_3, ...). As for the C_1 and inorganic chemistry files described above, each of these files also includes the corresponding literature references.

2.1 Gaseous mechanism construction protocol and SARs

The first gas-phase protocol implemented in GECKO-A is described in Aumont et al. (2005). Since then, the model has been extended to handle cyclic species (Valorso et al., 2011) and monocyclic aromatic compounds (Jenkin et al., 2018b). GECKO-A can process a wide range of hydrocarbons, including alkanes and alkenes with linear, branched, or cyclic structures, as well as monocyclic aromatics. For these carbon backbones, additional functional groups can be represented, including aldehyde, ketone, alcohol, hydroperoxide, ether, nitrate, peroxynitrate, nitro, carboxylic acid, peracid, and peroxyacyl nitrate moieties. However, the current version does not account for polycyclic aromatic compounds, reduced nitrogen species such as amines, or more generally organic compounds containing atoms other than carbon, hydrogen, oxygen, and nitrogen.

As GECKO-A evolved, the SARs used were regularly updated. The current version uses radical chemistry SARs of Vereecken ~~et al.~~and Peeters (2009, 2010) for alkoxy radicals, the SARs of Jenkin et al. (2019) for peroxy radicals, and the SARs of Newland et al. (2022) for Criegee intermediates. The reactivity of VOCs with oxidants has also been updated, in particular for reactions with OH (Jenkin et al., 2018a, ~~2018b~~b), NO_3 (Kerdouci et al., 2014), and O_3 (Jenkin et al., 2020). For thermodynamic properties relevant to phase partitioning, saturation vapor pressures p_{sat} can be estimated using any of 3 methods: Myrdal and Yalkowsky (1997), SIMPOL-1 (Pankow and Asher, 2008) or Nannoolal et al. (2004, 2008), as described by Valorso et al. (2011). Effective Henry's law constants, H, are estimated using the GROMHE method (Raventos-Duran et al., 2010), as described in Mouchel-Vallon et al. (2013, 2017). Finally, gas-phase diffusion coefficients, D, required to represent mass transfer between gas and condensed phases are estimated using the Fuller method (Poling et al., 2001), with diffusion volumes provided in the GECKO-A output files.

Updates to the GECKO-A tool have focused on multigenerational processes. However, the current version of GECKO-A does not yet include several key reactions that are now well-~~characterized~~documented in the literature. Major limitations include the absence of H-shift isomerization reactions involving RO_2 radicals~~, as well as~~ and ring-closure (cyclization) reactions for unsaturated RO_2 radicals (e.g., Bianchi et al., 2019; Crounse et al., 2013; Vereecken~~, 2019~~ and Nozière, 2020; Vereecken and Peeters, 2004; Vereecken et al., 2021a), as well as ring-closure reactions of unsaturated alkoxy radicals (e.g., Carlsson et al., 2023; Vereecken and Peeters, 2004, 2012; Vereecken et al., 2021b). These processes play an important role in the formation of low-volatility species and, consequently, in the production of secondary organic aerosols (SOA). ~~These~~The unimolecular reactions of peroxy radicals (autoxidation) are not implemented in this version but are currently under development, and their integration into GECKO-A is planned for the next release ~~of the tool (see also Sect. 6).~~(see also Sect. 7). Ring-closure reactions of unsaturated alkoxy radicals would first require a dedicated structure-activity relationship, which is not currently available. Additional updates are required to incorporate an isoprene chemical mechanism that aligns with current knowledge. For this species, it is now possible to develop near-explicit mechanisms based on theoretical and experimental data (Wennberg et al.,

2018). In the current version, the MCM v3.3.1 isoprene scheme is provided (Jenkin et al., 2015) and is directly applied to all C_4 and C_5 species derived from isoprene, while other species are generated using GECKO-A's standard protocol. For this version of GECKO-A, using the MCM-based approach for isoprene is recommended over the automatically generated mechanism ~~(see~~. Its activation is described in the ~~online technical~~user documentation ~~for further details~~of GECKO-A v1.0 (Aumont et al., 2026). Improved incorporation of similar chemistry for other conjugated systems, such as reversible O_2 addition to allylic radicals, is also under development. Finally, photolysis reactions have not been updated since the first version of GECKO-A (Aumont et al., 2005) and require a comprehensive revision to reflect recent advances in this field.

2.2 Mechanism configuration

Several parameters control how the program selects chemical pathways during mechanism generation. These parameters are directly accessible and modifiable by the user. They are primarily used to reduce the size of the chemical scheme. These parameters are briefly presented below and summarized in Table 1.

Within the GECKO-A tool, a generation is defined as the set of non-radical species produced from a given compound. The parent hydrocarbon is associated with generation zero. Species of the first generation thus constitute the pool of non-radical species produced from the parent compound, and so on for subsequent generations. The algorithm operates by processing generations sequentially. The consumption pathways of organic compounds from generation n+1 are thus generated only once all reactions associated with species from generation n are completed. A parameter (denoted maxgen) allows the user to limit the number of generations considered for mechanism construction. The chemical losses of species X from the final generation are simply transcribed in the mechanism as X + Ox → n XCLOSS, where Ox is the oxidant driving the loss, the stoichiometric coefficient n is the number of carbon atoms in species X and XCLOSS is a tracking species used to conserve carbon mass and to verify that the induced losses remain negligible during simulations.

A key mechanism size-limiting parameter is based on the volatility of non-radical species. Under typical atmospheric and laboratory conditions, the characteristic time for partitioning between the gas phase and condensed phases is fast compared to the characteristic time for gas-phase oxidation (e.g., Alfarra et al., 2023; Bertrand et al., 2018; La et al., 2016; Nah et al., 2016). Low-volatility non-radical species are thus rapidly sequestered in the condensed phase, and their loss via gas-phase oxidation becomes negligible, rendering their description in the mechanism unnecessary. The generation of the gas-phase oxidation mechanism is therefore halted for non-radical species whose vapor pressure falls below a threshold value (denoted critvp). To verify this condition, p_{sat} is calculated at a reference temperature TK, whose value is directly modifiable by the user.

Additional parameters are used to filter species produced with low yields. A first filter acts directly at the level of each generated reaction to select which products to retain. If the branching ratio BR associated with a reaction product falls below a threshold (denoted brcut), its formation is ignored and the other reaction parameters, i.e. total reaction rate constant and/or product distribution, are rescaled accordingly. Two other parameters are based on the maximum formation yield Y_{max} of species from the parent compound. This yield is calculated for each generated species, independent of chemical conditions (NO_x level, oxidant ratios, etc.). The calculation relies solely on unavoidable reaction branching (oxidant attack sites, per-reaction product yields, etc.) and is performed at reference temperature TK (see above). This choice reflects a design principle of GECKO-A. Mechanisms are intended to be generated once and applied across a wide range of conditions, rather than tailored to a specific environment. For peroxy radical reactions with NO or HO_2, the calculation assumes 100% conversion of

Table 1: User-configurable parameters (see text) controlling mechanism generation in GECKO-A. "Open source" refers to the default parameter values distributed with the public release (Sect. 2.4). "GECKO-A" refers to the high-resolution configuration adopted for this study (Sect. 4.1) and used throughout Sect. 5. "red-GECKO-A" refers to the reduced configuration introduced in Sect. 4.1 and used in Sect. 6.

Parameter [(a)]	Description	Open source [(b)]	GECKO-A	red-GECKO-A
maxgen	Maximum number of oxidation generations	1	20	5
critvp	p_{sat} cutoff ($\log_{10}$ atm) for gas-phase reactivity	-13	-13	-13
TK	Reference temperature (K)	298	298	298
brcut	Minimum branching ratio to retain a product	5×10^{-2}	5×10^{-3}	5×10^{-2}
rxloss	Minimum yield to assign explicit products	1×10^{-10}	1×10^{-6}	1×10^{-10}
yldcut	Minimum branching ratio×yield to retain a reaction	1×10^{-3}	1×10^{-5}	1×10^{-3}
isomerfg	Enable surrogate (isomer) lumping	True	False	True
highnoxfg	Restrict RO_2 reactions to RO_2+NO only	False	False	False
rx_ro2_multiclass	RO_2+RO_2 treated as 9 vs. 1 classes	True	False	False
g2pfg	Enable gas–particle mass transfer	True	True	True
g2wfg	Enable gas–wall mass transfer	False	False	False

(a): Parameters are set in the namelist file gecko.nml, in the sections "thresholds" (critvp, brcut, rxloss, yldcut), "env_cond" (TK), "reductions" (maxgen, isomerfg, highnoxfg, rx_ro2_multiclass) and "process" (g2pfg, g2wfg).
(b): The values provided in the open-source software (Aumont et al., 2025a) are set to minimize the size of the generated mechanism for initial exploratory use without requiring substantial computational resources. The exception is rxloss, which is set low to preserve the carbon mass balance in simple test cases. These values are intended to be adjusted according to user needs.

the parent radical through the considered pathway. Reactions with other peroxy radicals are instead assumed to convert only a small, fixed a priori fraction of the parent radical. This avoids biasing mechanism generation toward pathways that are significant only under specific conditions. Such pathways would otherwise appear more important than they typically are. This affects both the selection of species for isomer substitution and the decision to keep species and reactions in the mechanism (see below). A critical threshold (denoted rxloss) is defined below which stable (i.e., non-radical) species are not assigned explicit reaction products, i.e., if Y_{max} < rxloss. The loss of species X meeting this criterion is then represented as described previously for final-generation species, using the tracking species XCLOSS. Furthermore, the contribution of a reaction to the overall oxidation budget also depends on the combined value of the reactant's Y_{max} and the branching ratio of the reaction under consideration. A critical threshold (denoted yldcut) is defined to filter reactions whose effect is negligible. This filter applies to reactions of radical species. A reaction is ignored during mechanism generation when the condition $BR \times Y_{max} <$ yldcut is satisfied.

The use of surrogate species is available as an optional strategy to reduce the size of the chemical mechanism. Substitution of a newly generated species by another species already included in the mechanism is permissible if the two are positional isomers, that is, if they share the same molecular formula and the same set of functional groups. Substitution is further restricted to non-radical species with more than four carbon atoms and beyond the second generation only. When multiple substitutes are possible, a priority scheme is applied to identify the most appropriate surrogate~~.~~, as described in Sect. S1. These priorities are based on the structure of the carbon skeleton, the vicinity of functional groups, and bond conjugation. If more than one

Table 1: User-configurable parameters (see text) controlling mechanism generation in GECKO-A.

	Open source [(a)]	This work
critvp [(b)]	−13	−13
brcut	5×10^{-2}	5×10^{-3}
yldcut	1×10^{-3}	1×10^{-5}
rxloss	1×10^{-10}	1×10^{-6}
maxgen	1	20
TK	298	298
isomerfg	True	False
highnoxfg	False	False
rx_ro2_multiclass	True	False
g2pfg	True	True
g2wfg	False	False

(a): The values provided in the open-source software (Aumont et al., 2025a) are set to minimize the size of the generated mechanism for initial exploratory use without requiring substantial computational resources. These values are intended to be adjusted according to user needs.
(b): The threshold vapor pressure is given in $\log_{10}$ (atm).

substitute remains after the priority list is exhausted, the surrogate with the highest Y_{max} value is selected. A parameter (isomerfg) allows the user to enable or disable the use of surrogate species during mechanism generation.

The branching ratios resulting from bimolecular reactions of peroxy radicals depend directly on NO_x concentrations. For NO_x-rich environments (for example, urban areas or certain laboratory experiments), the RO_2 + NO reaction dominates over other bimolecular reactions which can then be ignored. A parameter (highnoxfg) allows the user to limit mechanism generation to only RO_2 + NO reactions. Similarly, a parameter (rx_ro2_multiclass) allows the user to retain either a single class or nine distinct classes of peroxy radicals to represent RO_2 + RO_2 reactions in the mechanism, as described in Jenkin et al. (2019).

Dynamic phase transfer of a species is represented as pseudo-unimolecular reactions, in which the first letter of each species identifier designates the phases involved. In addition to the gas phase (code letter 'G'), partitioning with two condensed phases can be considered: a particulate phase (code letter 'A') and reactor walls (code letter 'W'). For example, the absorption of any species named X between the gas and particulate phases is transcribed as the pseudo-reaction GX → AX, and its desorption is represented by the reverse pseudo-reaction AX → GX. The user can enable or disable the writing of gas-to-particle and gas-to-wall mass transfer reactions via the parameters g2pfg and g2wfg, respectively.

2.3 Mechanism generation: size, complexity, and reduction strategies

The size of the generated mechanism depends on the organic precursor provided as input and the values assigned to the parameters listed in Table 1. Figure 1a shows the evolution of the number of generated species (N_S) for *n*-alkanes as a function of the carbon atom number of the parent compound, using the set ofGECKO-A parameters reported in Table 1 for this study.. For short carbon chains, N_S increases exponentially with the number of carbon atoms, as described in Aumont et al. (2005). For larger species, the exponential increase in gas-phase products becomes limited by the conversion to non-volatile species, which are assumed to partition to aerosols rather than undergoing additional reactions in the gas-phase. With the exclusion of non-volatile species ($p_{sat} < 10^{-13}$ atm), the growth rate for the *n*-alkane series decreases beyond C_8, and N_S reaches a plateau

around C_{16} – C_{20}, before decreasing for longer chains. Chemical functionalization pathways reach the p_{sat} threshold more rapidly for longer chains, thereby reducing the average number of generations required to complete the mechanism — and, consequently, limiting N_S. For a given number of carbon atoms, the size of the mechanism generated also depends significantly on the complexity of the carbon skeleton (branching, unsaturation, cyclic structure). For example, N_S increases by a factor of 17 from n-heptane to toluene and by a factor of around 3 from decane to α-pinene.

Figure 1a also shows the reduction in mechanism size achieved through the use of surrogate species. This reduction is moderate for short-chain organic species, due to the limited availability of positional isomers meeting the substitution criteria. The number of available isomers increases with the size of the carbon skeleton, thereby enhancing the effectiveness of this reduction approach. For the *n*-alkane series, the reduction is only about 30% for butane but typically exceeds a factor of 10 beyond C_{10} species. Figure 1a also shows the size reduction achieved using the parameter combination for red-GECKO-A (Table 1, Sect. 4.1). For the *n*-alkane series, this reduction reaches about 50% for butane and typically exceeds a factor of 50 beyond C_{10} species, a substantially larger reduction than that achieved through isomer lumping alone.

Figure 1b shows the increase in mechanism size with the maximum number of generations (maxgen) considered for three parent hydrocarbons: dodecane, toluene, and α-pinene. The growth follows an exponential trend in the early generations, with species multiplying from one generation to the next. For these parent hydrocarbons, N_S increases by roughly an order of magnitude per generation. The growth rate progressively decreases with the number of generations. This inflection arises because progressive functionalization of species with increasing generation number lowers their vapor pressure below the critvp threshold, and because newly produced species can already be present in the mechanism as products of earlier generations. N_S approaches a plateau around eight generations and increases only marginally beyond that point.

The increase in mechanism size with decreasing vapor pressure threshold (critvp) is shown in Fig. 1c for dodecane, toluene, and α-pinene. This parameter strongly determines the size of the generated mechanism and must be adjusted by the user to suit the simulation objectives. For most applications, the acceptable range for the critical pressure threshold typically lies between 10^{-13} and 10^{-10} atm, below which species are predominantly sequestered in the condensed phase, and above which species remain sufficiently volatile to require consideration of their gas-phase oxidation under most conditions. Over this range, the mechanism size typically increases by a factor of 1.5 when the vapor pressure threshold decreases by one log unit.

The computational time for generating mechanisms for the *n*-alkane series is reported in Fig. 1d. For smaller mechanisms (up to a few hundred thousand species), computational time is roughly proportional to the number of generated species (approximately 5–6 minutes per hundred thousand species). Beyond a few million species, the time typically increases quadratically with the number of generated species.

2.4 GECKO-A open source, website and documentation

The GECKO-A tool was released as open-source software in 2025 (Aumont et al~~.~~., 2025a). The code is publicly available on GitLab (https://gitlab.in2p3.fr/ipsl/lisa/geckoa/public), together with a wiki for technical documentation. A permanent archived version is also accessible via Zenodo (https://doi.org/10.5281/zenodo.~~15309904~~15309905). Model parameters can be configured through a Fortran namelist, allowing users to adjust the key settings described in Sect. 2.2. Table 1 shows the parameter values provided with the open-source version. These default values are designed to generate compact chemical

mechanisms that remain computationally efficient for box-model simulations and initial exploratory analyses. These settings should, however, be adapted by the user to better match specific research objectives. In addition to the parameters summarized in Table 1, the namelist can include further options, allowing users to activate additional, optional reaction pathways that are not enabled by default. This includes, for example, the heterogeneous formation of cyclic hemiacetals and dihydrofurans (DHFs) from 1,4-hydroxycarbonyls (Lim and Ziemann, 2009). This option is activated with the dhffg flag, to be added in the "process" section of the namelist (see the user documentation of GECKO-A v1.0 in Aumont et al., 2026).

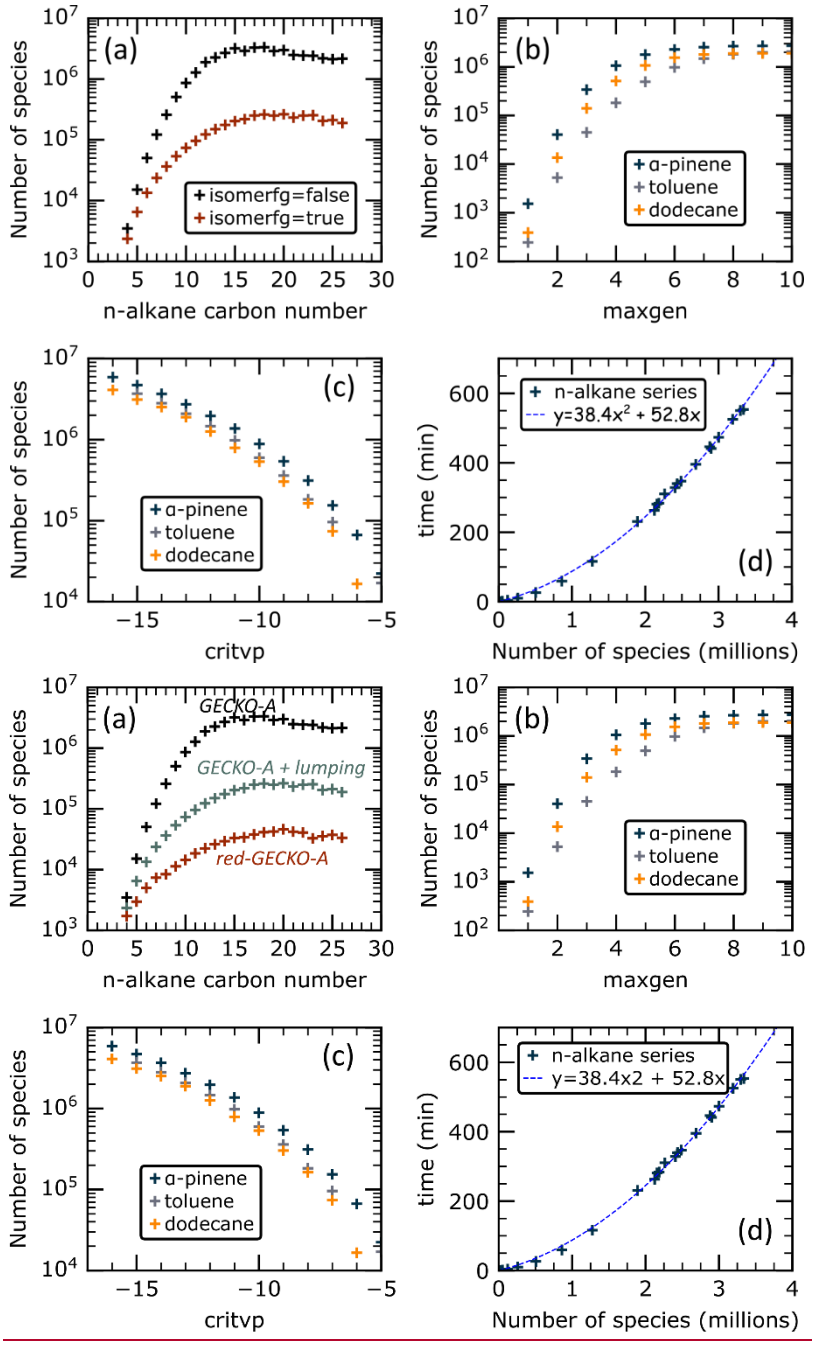


Figure 1: Number of species generated in GECKO-A mechanisms for the *n*-alkane series, for GECKO-A, GECKO-A with isomer lumping, and ~~without the use of surrogate species (~~red-GECKO-A (Table 1; panel a~~),~~); for α-pinene, toluene, and dodecane as a function of the maximum number of generations maxgen (panel b) and the vapor pressure threshold critvp, in log10(atm) (panel c). Indicative computational time for mechanism generation of the *n*-alkane series is shown in panel d (Intel Xeon Gold 6238R, 2.20 GHz). Unless

otherwise specified (panel a), mechanisms are generated using the GECKO-A parameter values ~~adopted for this work (see~~ (Table 1), except for the parameter examined in each panel.

In parallel, a dedicated website has been developed to provide, for a given organic compound, its kinetic constants with atmospheric oxidants, its physicochemical properties relevant to phase partitioning, and a partial chemical mechanism limited to the first 1000 reactions (~~www.~~geckoa.lisa.u-pec.fr, last access January 2026). This limit is intentionally imposed to maintain reasonable computational load and responsiveness on the server, as the platform is intended primarily to illustrate the capabilities of GECKO-A rather than for operational use. Each user request triggers a real-time execution of GECKO-A to generate the corresponding outputs. The GECKO-A Online version is identical to the open-source release.

GECKO-A produces two mechanism files: a standard version and an annotated version. The standard file contains the chemical scheme in a format directly usable for atmospheric simulations with the GECKO-A box model. The annotated version tags each reaction with one to three alphanumeric codes, which are linked to a dedicated dictionary file (references.dat). This dictionary provides detailed metadata, including the structure-activity relationships used, references to experimental data, reaction types, and other protocol-specific information. The annotations ensure full traceability of the methodology underlying each reaction. The GECKO-A website provides examples of these reaction codes and their associated metadata, with both elements explicitly included for each reaction displayed.

3 The GECKO-A box model

The GECKO-A box model is a zero-dimensional kinetic model that enables time-dependent simulations of chemical species concentration evolution, based on user-defined initial concentrations and environmental conditions. The model accounts for species sources and sinks as defined by the underlying chemical mechanism. While several box models exist in the literature, a distinctive feature of the GECKO-A box model is the two-step solver implemented (Verwer, 1994; Verwer et al., 1996), a Jacobian-free iterative method designed for nonlinear systems. This approach has proven particularly well-suited for solving the ordinary differential equations (ODEs) derived from mechanisms involving millions of species and reactions, such as those generated by GECKO-A (e.g., Aumont et al., 2005).

3.1 Box model description

The evolution of chemical species concentrations in the box model is described by a system of ODEs of the form:

$$\frac{dC_{X,\varphi}}{dt} = P_{X,\varphi} - L_{X,\varphi}$$

Here $C_{X,\varphi}$ is the concentration of species X in phase φ (gas, particle and/or chamber walls), while $P_{X,\varphi}$ and $L_{X,\varphi}$ represents its production and loss terms, respectively. These terms are defined by the input chemical mechanism, which includes gas-phase reactions and dynamic phase partitioning. Additionally, non-species-specific losses (e.g., first-order dilution or particle loss rates) can be directly incorporated into the box model through namelist parameters (see below).

As mentioned in Sect. 2.2, a given species X in the gas and particle phase is represented using two distinct variables GX and AX, respectively. The gas–particle mass transfer of the species X is simply described as two unimolecular reactions (La et al., 2016; McVay et al., 2016):

GX (gas) → AX (particle)	k_{gp}
AX (particle) → GX (gas)	k_{pg}

where k_{gp} and k_{pg} denote the uptake and evaporation first order rate constants, respectively. The mass transfer coefficient k_{gp} is calculated using the Fuchs–Sutugin formulation (Seinfeld and Pandis, 2016):

$$k_{gp} = 4\pi D r_p \frac{0.75\alpha(1+K_n)}{0.75\alpha + 0.283\alpha K_n + K_n + K_n^2} C_p$$

where D is the gas-phase diffusion coefficient of the species, α is the accommodation coefficient (default value set to 1), r_p is the particle radius, C_p is the particle number concentration, and K_n is the Knudsen number ($K_n = \lambda / r_p$, with λ the mean free path). The diffusion coefficient of each species is estimated using the Fuller structure–property relationship (Poling et al., 2001), and the mean free path is evaluated as $\lambda = 3D/c$, where c is the mean molecular speed calculated from kinetic gas theory. The desorption rate constant k_{pg} is derived from the equilibrium constant K_{eq} ($K_{eq} = k_{gp}/k_{pg}$). Phase equilibrium is described by an absorptive partitioning process following Raoult's law (Pankow, 1994):

$$K_{eq} = \frac{RT}{p_{sat}} \times \frac{C_{OA}}{M_{OA}\gamma_{OA}}$$

where C_{OA} is the organic aerosol mass concentration, M_{OA} is the mean molar mass of the organic aerosol, and γ_{OA} is the activity coefficient in the particle phase (default value set to 1).

Similar to gas-particle partitioning, gas-wall mass transfer for a given species X is represented by two pseudo-first-order reactions (La et al., 2016):

GX (gas) → WX (wall)	k_{gw}
WX (wall) → GX (gas)	k_{wg}

where WX represents species X sorbed onto the wall (e.g., a Teflon film), and k_{gw} and k_{wg} denote the first-order uptake and evaporation rate coefficients, respectively. The same uptake coefficient k_{gw} is applied to all organic species, including the parent hydrocarbon, with a default value corresponding to a gas-to-wall lifetime of 15 minutes (La et al., 2016; Matsunaga and Ziemann ~~et al.,~~ 2010). This value is user-modifiable. The equilibrium constant $K_{eq,gw}$ which constrains k_{wg} is computed using the parametrization from Matsunaga and Ziemann (2010), as detailed in La et al. (2016):

$$K_{eq,gw} = \frac{k_{gw}}{k_{wg}} = \frac{RT}{p_{sat}} \times \frac{C_w}{M_w \gamma_w}$$

In this expression, C_w and M_w represent the equivalent organic aerosol mass concentration and equivalent mean molar weight associated with the wall, respectively, while γ_w is the activity coefficient of the species on the wall. The ratio $C_w/M_w\gamma_w$ is specific to each species and empirically derived from chamber observations (Matsunaga and Ziemann, 2010; Yeh and Ziemann, 2015). The value of this ratio is provided as auxiliary information alongside each wall ↔ gas mass transfer pseudo-

reaction pair. In the current version of the model, GECKO-A provides two default values for this ratio during mechanism generation: 20 and 120 μmol m^{-3}, for hydrocarbons and oxygenated organic compounds, respectively. The values provided here define a generic framework for gas-wall mass transfer within the GECKO-A code structure. These default parameterizations are intended as a practical starting point and should be adjusted by the user to the specific chamber and compounds under consideration, for example by substituting alternative published parameterizations such as that of Krechmer et al. (2016), as illustrated by Peng et al. (2021).

By default, this version of the box model sets the physical parameters (temperature, pressure, relative humidity) to the values provided in the input file. Routines with clearly defined input/output variables are available, enabling users to easily integrate custom parameterizations with just a few lines of code. This allows for dynamic modulation of these parameters over time to meet specific simulation needs. Photolysis rate coefficients are provided for clear-sky conditions at varying solar zenith angles (SZAs). These coefficients were calculated using the radiative transfer model TUV v5.4 (Madronich and Flocke, 1997). The default vertical profiles provided in TUV v5.4 for calculating spectral actinic fluxes and photolysis rate coefficients at 0.5 km above sea level were used (as provided by the defin2 input file, https://www2.acom.ucar.edu/modeling/tuv-download, last access: 4 May 2026) and correspond to typical continental values (ozone column of 300 DU). The box model simulations incorporated the rate coefficients either by fixing the SZA (constant light intensity) or by modeling its diurnal variation for a specific date and location. Photolysis rate coefficients computed for black light chamber lights are also available. The box model does not represent particle microphysical processes (e.g. nucleation). Simulations including gas-to-particle mass transfer must therefore specify an initial seed concentration of non-volatile inorganic and/or organic compounds, defined by the user in the input file. Examples of box model applications include evaluation against laboratory chamber ~~experiments~~data (e.g., La et al., 2016; McVay et al., 2016; Peng et al., 2021; Valorso et al., 2011) as well as simulations of realistic, complex VOC mixtures in urban and forested environments and outflow plumes (e.g., Lee-Taylor et al., 2011, 2015; Mouchel-Vallon et al., 2020).

3.2 Optimizing Runtime

The most computationally demanding step in the box model is the iterative calculation of production and loss terms for each of the N_S species, which determines their rate of change. Since these calculations are independent for each species within a time step, this part of the code has been parallelized in the GECKO-A box model using the OpenMP library. To assess the efficiency of this parallelization, benchmark tests were performed to evaluate the simulation runtime as a function of the number of cores used. These tests were conducted for the remote, continental, and urban scenarios described in Sect. 4.3, using α-pinene as the parent compound. The results, illustrating the sensitivity of runtime to core count, are presented in Fig. 2. The parallelization achieved a speedup factor of approximately 3.5 when using 12 cores, with minimal further improvement observed beyond this threshold.

3.3 GECKO-A box model open source and documentation

The GECKO-A box model was released as open-source software in 2025 (Aumont et al., 2025b), alongside the GECKO-A mechanism generator. A permanent archived version is available on Zenodo (https://doi.org/10.5281/zenodo.~~15310152~~15310153), while the latest code and technical documentation are hosted on GitLab (https://gitlab.in2p3.fr/ipsl/lisa/geckoa/public), including a dedicated wiki for user guidance. Chemical mechanisms generated

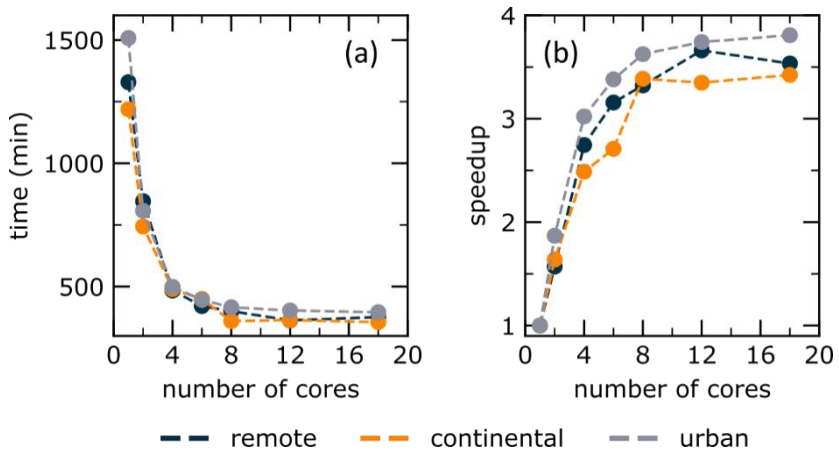


Figure 2: Computational time for box model simulations as a function of the number of cores (panel a) and speedup achieved through parallelization (panel b). Benchmarks are performed for the remote, continental, and urban scenarios (5 simulated days – see text) using the α-pinene mechanism (2.7×10^6 species, 6.7×10^6 reactions). Simulation times are indicative and correspond to one node of the LISA cluster.

by GECKO-A, along with their associated property files, can be directly integrated into the box model. Simulation parameters, such as start/stop times, number of time steps, and solver tolerance, first-order dilution or particle loss rates, etc are configured via a Fortran namelist. Additional fine-tuning parameters (e.g., default gas-to-wall lifetime, accommodation coefficient, etc) are accessible within a dedicated Fortran routine (parameter_mod.f90) and must be modified prior to compilation, as needed. Simulation results are output in NetCDF format, a machine-independent and widely adopted standard in atmospheric sciences (https://www.unidata.ucar.edu/software/netcdf). This format ensures compatibility with modern analysis tools (e.g., Python, R) and simplifies data sharing. ~~A~~An example Python script for extracting data from GECKO-A simulation output files is ~~available on~~provided with the box ~~model's GitLab wiki (https://gitlab.in2p3.fr/ipsl/lisa/geckoa/public/boxmodel4gecko/-/wikis/home/output-files~~model (see the user documentation of the box model v1.0 in Aumont et al., 2026).

4 Model setup and scenarios

A series of simulations was conducted to ~~evaluate~~assess the performance of GECKO-A v1.0. These simulations aimed to explore the oxidative trajectories generated by GECKO-A and compare them against a reference mechanism, MCM (v3.3.1), to assess differences in mechanistic behavior.

For this comparison, a set of parent hydrocarbons available in the MCM v3.3.1 database was selected. Three linear alkanes of increasing chain length—butane, octane, and dodecane—were chosen to investigate how carbon backbone size influences the convergence or divergence of the two mechanisms. Dodecane, the largest alkane available in MCM v3.3.1, was included to represent longer-chain compounds. Additionally, toluene and α-pinene were incorporated to represent diverse molecular structures commonly found in atmospheric emissions, including cyclic, olefinic, and aromatic compounds, as well as to represent both anthropogenic and biogenic sources.

This study thus systematically compares the oxidation of these five structurally diverse parent hydrocarbons using both GECKO-A and MCM mechanisms across contrasting environmental scenarios, ranging from urban to remote conditions. In the following section, we outline the mechanism setup and the simulation scenarios used to explore oxidative trajectories and perform the mechanism comparison.

4.1 GECKO-A setup

The configuration used to generate the GECKO-A mechanisms for the five organic precursors is summarized in Table 1 and described below. As the objective of this study is to examine in detail the oxidative trajectories produced by GECKO-A, the generation parameters were selected to minimize the impact of mechanism reduction. ~~In particular,~~These mechanisms, generated at high resolution, are simply referred to as "GECKO-A" hereafter and correspond to the "GECKO-A" column of Table 1. In particular, these mechanisms were generated without isomer lumping, in contrast to the default GECKO-A configuration. In addition, no restriction was imposed on the number of generations during mechanism construction. One exception to this approach concerns the treatment of RO_2+RO_2 reactions. The nine peroxy radical classes of Jenkin et al. (2019) were not distinguished (rx_ro2_multiclass set to False), for reasons detailed in Sect. 4.3. To limit mechanism size, the gas-phase reactivity of species with saturation vapor pressure p_{sat} below 10^{-13} atm was neglected (see critvp in Table 1), as these compounds are expected to reside predominantly in the condensed phase (see Sect. 5.2). Sensitivity tests using a lower p_{sat} threshold showed no significant impact on the results presented here.

The branching ratio threshold brcut, used to filter minor reaction pathways, was set to 0.5%. The threshold rxloss, below which stable non-radical species are not assigned explicit reaction products, was set to 10^{-6}, and the yield cutoff yldcut was set to 10^{-5}. For the conditions examined in this study, further reduction of these thresholds did not significantly affect the simulation metrics, while carbon loss remained below 1% at the end of the simulations (see Sect. 5.3).

In addition to the mechanism generated at high resolution, a second mechanism was generated for each of the five parent compounds, using more stringent reduction parameters. This parameter set, denoted red-GECKO-A, retains all default parameter values of the open-source version of GECKO-A (Table 1), except for two parameters, maxgen and rx_ro2_multiclass. For simplicity, GECKO-A and red-GECKO-A refer hereafter both to these parameter sets and to the mechanisms generated with them. The red-GECKO-A configuration is intended to illustrate the reduction capabilities offered by GECKO-A (Sect. 6), rather than to systematically explore the effect of each reduction parameter on the simulated oxidative trajectories.

Results obtained with the GECKO-A mechanism (Sect. 5.9) show that the contribution of species beyond the fourth generation becomes marginal, regardless of the parent compound or scenario considered (Fig. 17). maxgen was therefore set to 5 for red-GECKO-A. This value limits the carbon losses associated with mechanism truncation, while substantially reducing its size (Fig. 1b). This number of generations is commonly used in previous applications of GECKO-A (e.g., Isaacman-VanWertz and Aumont, 2021; La et al., 2016; Lee-Taylor et al., 2011). rx_ro2_multiclass, in turn, was kept at the value retained for the GECKO-A mechanism (False), rather than reverted to the open-source default value (see Sect. 4.3). The other parameters were kept at their open-source default values. This includes rxloss (10^{-10}), even though this value is lower than in GECKO-A (10^{-6}). This difference has little effect on the generated mechanisms, since maxgen, brcut, and yldcut already limit the formation of minor stable species.

Non-radical species were represented in both the gas and particulate phases, with chemistry occurring exclusively in the gas phase. Prescribed concentrations of pre-existing organic aerosol were used for each scenario (see Sect. 4.3). Saturation vapor pressures were estimated using the method of Nannoolal et al. (2004, 2008). All mechanisms and associated thermodynamic properties were generated using the open-source version of GECKO-A described in Sect. 2.

4.2 The Master Chemical Mechanism (MCM)

The MCM ~~mechanism~~ is developed manually following standardized protocols to address missing data (Bloss et al., ~~2005~~2005a; Jenkin et al., 1997, 2015; Saunders et al., 2003). The current MCM (version 3.3.1) covers the degradation of 143 primary VOCs, including carbon skeletons up to C_{12} as well as one C_{15} sesquiterpene. In order to maintain a manageable mechanism size while retaining essential features, systematic simplifications are applied. These include neglecting pathways contributing less than 5% to VOC + oxidant reactions, ignoring pathways with likely negligible contributions under typical tropospheric conditions, lumping isomers by considering only one representative oxidation pathway for alkanes with $C_{>7}$, and assuming that "minor products" (e.g., nitrates and hydroperoxides) react to form only species already present in the mechanism. Additionally, peroxy radical recombination reactions are parameterized using a single RO_2 surrogate species representing the sum of all peroxy radicals, and considering only one reaction pathway for RO_2 radicals formed from minor products. These simplifications are described in detail in Jenkin et al. (1997) and Saunders et al. (2003). Incorporating updates into a manually built mechanism is also difficult, as already noted in the introduction (Sect. 1). In practice, updates are not applied as a single, mechanism-wide revision each time a new SAR becomes available. They are instead introduced on a compound-by-compound basis, documented in dedicated publications (e.g., Jenkin et al., 2015, for isoprene). As a result, MCM v3.3.1 combines chemistry from different periods, with updates varying by compound family. In addition, some reactions or rate coefficients were adjusted to reproduce chamber observations (e.g., Bloss et al., 2005a, b; Jenkin et al., 2012; Jenkin et al., 2015). The MCM therefore reflects observational knowledge beyond its construction rules, which reinforces its value as a benchmark.

The MCM extraction was performed here from the online MCM platform (https://mcm.york.ac.uk/MCM/; last access: 10 April 2025) for all available non-sulfur and non-halogenated organic species. This extraction includes their reactions with associated kinetic parameters, and the absorption cross sections and quantum yields required to calculate photolysis rate coefficients. However, the inorganic reaction subset was excluded. Instead, the GECKO-A inorganic mechanism was used for all simulations to maintain identical buffered conditions across scenarios with both mechanisms (see the following section). The resulting extracted MCM organic mechanism comprises 16,603 reactions and 5,603 species. Of these, 184, 313, 411, 283, and 305 species are specifically involved in the degradation of butane, octane, dodecane, toluene, and α-pinene, respectively.

Thermodynamic properties (p_{sat}, H, D) were calculated for all non-radical species using the estimation methods implemented in the GECKO-A tool. Gas–particle partitioning reactions were incorporated into the mechanism following the procedures described above for the GECKO-A mechanism. Photolysis rate coefficients were computed using the TUV-5.4 model under the same physical conditions as those applied to the GECKO-A mechanisms. The MCM mechanism was formatted to comply with the file standards adopted in the GECKO-A tools, enabling its use in the box model described in Sect. 3.

4.3 Scenario

Table 2: Chemical forcing applied to the different scenarios.

	Remote	Remote Continental	Continental	Polluted Continental	Urban
NO_x (ppb)	0.010	0.025	0.50	2.0	20
O_3 (ppb)	40	40	40	40	40
CH_4 (ppb)	1750	1750	1750	1750	1750
CO (ppb)	120	120	150	200	300
pHCHO (ppb) [a]	0	0	2	5	10
pRH reactivity with OH (s^{-1}) [b]	0	1	6	9	13
Org. Aerosols ($\mu g/m^3$)	10	10	10	10	10

(a): pHCHO is a proxy used to regulate the initiation of radical chains (see text).
(b): pRH is a volatile organic compound (VOC) proxy used to regulate the chemistry of the HO_x/NO_x system (see text). Its reactivity with the OH radical is adjusted according to the considered scenario.

Several environmental scenarios were designed to compare the oxidative trajectories simulated by the GECKO-A and MCM mechanisms. The fate of gas-phase organic compounds strongly depends on the chemical regime considered. The scenarios therefore span a range of conditions, from NO_x-saturated regimes typical of urban environments to NO_x-limited regimes characteristic of regions remote from anthropogenic sources. Five scenarios were developed and are hereafter referred to as "urban", "polluted continental", "continental", "remote continental", and "remote". Their purpose is not to reproduce any specific location, but rather to represent chemical regimes typically encountered in the lower troposphere. These scenarios are based on fixed concentrations of key species such as NO_x, O_3, CH_4, and CO (see Table 2).

A proxy hydrocarbon, denoted pRH below, is introduced to regulate HO_x chemistry and to provide a chemical background within which each parent hydrocarbon will be studied. A canonical chemical mechanism is associated with this proxy. This mechanism is restricted to the following transformations:

$$pRH + OH \rightarrow pRO_2$$
$$pRO_2 + NO \rightarrow NO_2 + HO_2$$
$$pRO_2 + HO_2 \rightarrow pROOH$$
$$pRO_2 + pRO_2 \rightarrow 0.74\ HO_2$$
$$pROOH + h\nu \rightarrow OH + HO_2$$
$$pROOH + OH \rightarrow OH$$
$$pROOH + OH \rightarrow pRO_2$$

The rate constants for reactions involving pRO_2 and pROOH are based on those of CH_3O_2 and CH_3OOH, respectively. The reactivity of pRH with OH is adjusted according to the scenario considered (see Table 2). Similarly, a proxy for organic species undergoing photolysis, denoted pHCHO, is introduced to regulate the initiation of radical chains. The chemistry of this proxy is directly based on that of HCHO, and its concentration is scenario-specific (see Table 2).

Table 2: Chemical forcing applied to the different scenarios.

	Remote	Remote Continental	Continental	Polluted Continental	Urban
NO_x (ppb)	0.010	0.025	0.50	2.0	20
O_3 (ppb)	40	40	40	40	40
CH_4 (ppb)	1750	1750	1750	1750	1750
CO (ppb)	120	120	150	200	300
pHCHO (ppb) [(a)]	0	0	2	5	10
pRH reactivity with OH (s^{-1}) [(b)]	0	1	6	9	13
Org. Aerosols ($\mu g/m^3$)	10	10	10	10	10

(a): pHCHO is a proxy used to regulate the initiation of radical chains (see text).
(b): pRH is a volatile organic compound (VOC) proxy used to regulate the chemistry of the HO_x/NO_x system (see text). Its reactivity with the OH radical is adjusted according to the considered scenario.

All simulations using the GECKO-A and MCM mechanisms employ the same reaction scheme for inorganic chemistry. A loss term, analogous to dry deposition, is applied to terminal species (HNO_3, H_2O_2) to prevent their accumulation over time. Without this term, their concentrations would increase anomalously, ultimately leading to an overestimation of their contribution to the NO_x/HO_x budget. Likewise, all simulations use the same mechanism for fixed carbon species (CH_4, CO, pHCHO, pRH). These carbon species (and their reaction products) are tagged to distinguish them from species generated by the GECKO-A and MCM oxidation mechanisms during mass balance calculations.

The temperature is set to 298 K and relative humidity to 70%. The organic aerosol (OA) mass concentration is $C_{OA} = 10$ µg m^{-3}. The particle size distribution is monodisperse with a radius $r_p = 0.1$ µm. Particles are assumed non-volatile and composed of molecules with an average molar mass $M_{OA} = 200$ g mol^{-1}. Radiation inputs correspond to a constant solar zenith angle of 45° under the conditions described in Sect. 3.1.

Each scenario is integrated over time until a steady state is reached for all unconstrained species. The hydrocarbon under investigation is not included during this spin-up phase. This steady state is then used as the initial condition for simulations exploring the oxidative trajectories of the parent hydrocarbons. The steady-state concentrations of selected radicals are reported in Table 3. The magnitude of the branching ratios for the reactions RO_2+ HO_2, RO_2 + NO, and RO_2 + ~~RO_2~~XO_2 is also provided for reference in Table 3. These ratios are calculated from the steady-state concentrations of HO_2, NO, and RO_2, using the kinetic constants from Jenkin et al. (2019) for a hypothetical long-chain RO_2 radical belonging to the 7th RO_2 class for RO_2 + RO_2 reactions (i.e., primary or secondary unsubstituted RO_2 radicals). For the scenarios considered in this study, the RO_2+XO_2 branching ratio does not exceed 13%. Moreover, XO_2 is dominated by pRO_2, which originates from the pRH proxy rather than from the parent compound under study. This explains the choice of rx_ro2_multiclass set to False for mechanism generation. Distinguishing the nine RO_2 classes of Jenkin et al. (2019) for this minor pathway, already dominated by a fixed species, would not introduce any notable change in the simulated oxidative trajectories.

Table 3: Steady-State concentrations of selected species and organic peroxy (RO_2) branching ratios across different atmospheric scenarios.

	Remote	Remote Continental	Continental	Polluted Continental	Urban
NO concentration [(a)]	4.3E+07	1.0E+08	2.1E+09	8.4E+09	1.1E+11
OH concentration [(a)]	2.9E+06	1.6E+06	1.5E+06	3.2E+06	2.4E+06
HO_2 concentration [(a)]	3.9E+08	2.5E+08	4.6E+08	5.1E+08	5.4E+07
XO_2 concentration [(a)(c)]	4.9E+08	7.8E+08	4.6E+08	3.9E+08	3.3E+07
RO_2 + NO branch. ratio [(b)]	0.04	0.13	0.64	0.86	1.00
RO_2 + HO_2 branch. ratio [(b)]	0.90	0.74	0.34	0.13	0.00
RO_2 + XO_2 branch. ratio [(b)(c)]	0.06	0.13	0.02	0.01	0.00

(a): Concentrations are expressed in molecules cm^{-3}.
(b): The RO_2 branching ratios are calculated using the following bimolecular rate constants (in cm^3 $molecule^{-1}$ s^{-1}): $k_{RO2+NO} = 9.0 \times 10^{-12}$, $k_{RO2+HO2} = 2.2 \times 10^{-11}$, $k_{RO2+XO2} = 1.2 \times 10^{-12}$. These constants are based on Jenkin et al. (2019). The ratios are provided as indicative values, assuming RO_2 is a long-carbon-chain radical with multiple O and N functional groups, representative of the 7th RO_2 group (see Jenkin et al., 2019).
(c): XO_2 represents the sum of CH_3O_2 and pRO_2, with CH_3O_2 and pRO_2 originating from the fixed species CH_4 and pRH, respectively.

Simulations were then performed for five days under these steady state scenarios, using the MCM and GECKO-A mechanisms, for the five organic precursors. Simulations are initialized with a mixing ratio of 1 pptC for the parent compound under scrutiny, an amount chosen to be sufficiently small to avoid perturbing the chemical scenario or altering the buffered conditions, regardless of the hydrocarbon or scenario considered. Consequently, the mechanisms are compared under identical conditions in terms of solver settings, physical parameters (temperature, relative humidity, radiation), and chemistry (buffered conditions with identical steady-state levels of inorganic species and organic aerosol mass). Any differences observed between MCM and GECKO-A in mass balances stem solely from the oxidative pathways described by each mechanism, without feedback loops caused by changes in the simulated scenarios.

~~5 Results~~

5 GECKO-A oxidative trajectories, benchmarked against MCM

Each result presented below illustrates the oxidative trajectory simulated by GECKO-A for a given hydrocarbon and scenario, with the MCM used throughout as a benchmark to assess this trajectory against an established reference mechanism.

5.1 Gas- and particle-phase carbon budgets

Figure 3 shows the temporal evolution of gaseous and particulate organic carbon simulated with MCM and GECKO-A for various hydrocarbons (butane, octane, dodecane, toluene, and α-pinene) under the environmental scenarios described in Sect. 4.3. Results are expressed in pptC. Since each simulation is initialized with 1 pptC of parent compound (see above), the scale can also be viewed as the yield relative to initial organic carbon on a carbon atom basis. The missing fraction corresponds to inorganic carbon (CO + CO_2).

As expected, first-generation oxidation products of parent VOCs form mainly gaseous organic compounds (secondary organic gases, denoted SOG hereafter). Multigenerational oxidation generates progressively lower volatility species that condense onto the pre-existing organic aerosol (Sect. 5.2). The particulate organic compound (SOA) yield increases with carbon chain length and from urban to remote scenarios (Fig. 3). For butane, this yield remains negligible in all scenarios with both mechanisms. In contrast, it exceeds 60% for dodecane and α-pinene with GECKO-A under remote conditions.

Excellent agreement is observed between the two mechanisms for butane across all scenarios. For the alkane series, discrepancies increase with parent compound carbon chain length. For SOG, both mechanisms simulate similar trends, but diverge increasingly from urban to remote conditions. Results differ strongly for SOA: with MCM, octane produces almost no SOA, whereas with GECKO-A the yield increases from a few percent under urban conditions to about 20% under remote conditions. Dodecane shows a similar trend, with significant SOA yield with the MCM from the continental scenario ~~onward~~ (8%), ~~reaching~~increasing to 50% under remote conditions. Maximum SOA yields simulated with GECKO-A increase from 25% under urban conditions to 70% under remote conditions.

Similar trends are observed for toluene and α-pinene. MCM simulates no substantial SOA production for toluene, whereas GECKO-A simulates yields from a few percent under urban conditions to 42% under remote conditions. For α-pinene, SOA production with MCM does not exceed 5% under any conditions, whereas with GECKO-A it ranges from 4% under urban conditions to 65% under remote conditions. This SOA underestimation by MCM leads to SOG overestimation compared to GECKO-A.

Figure 3 reveals marked disparities in gas/particle partitioning of secondary organic compounds, particularly for hydrocarbons with long carbon chains or complex structures. The oxidative pathways simulated by the two mechanisms thus produce secondary species with distinct volatility distributions, which are analyzed in the following section.

Finally, Fig. 3 shows that the polluted continental and remote continental scenarios yield intermediate results between urban and continental, and between continental and remote, respectively, without notable singularities. To simplify the figures, results from these intermediate scenarios are omitted in the following sections.

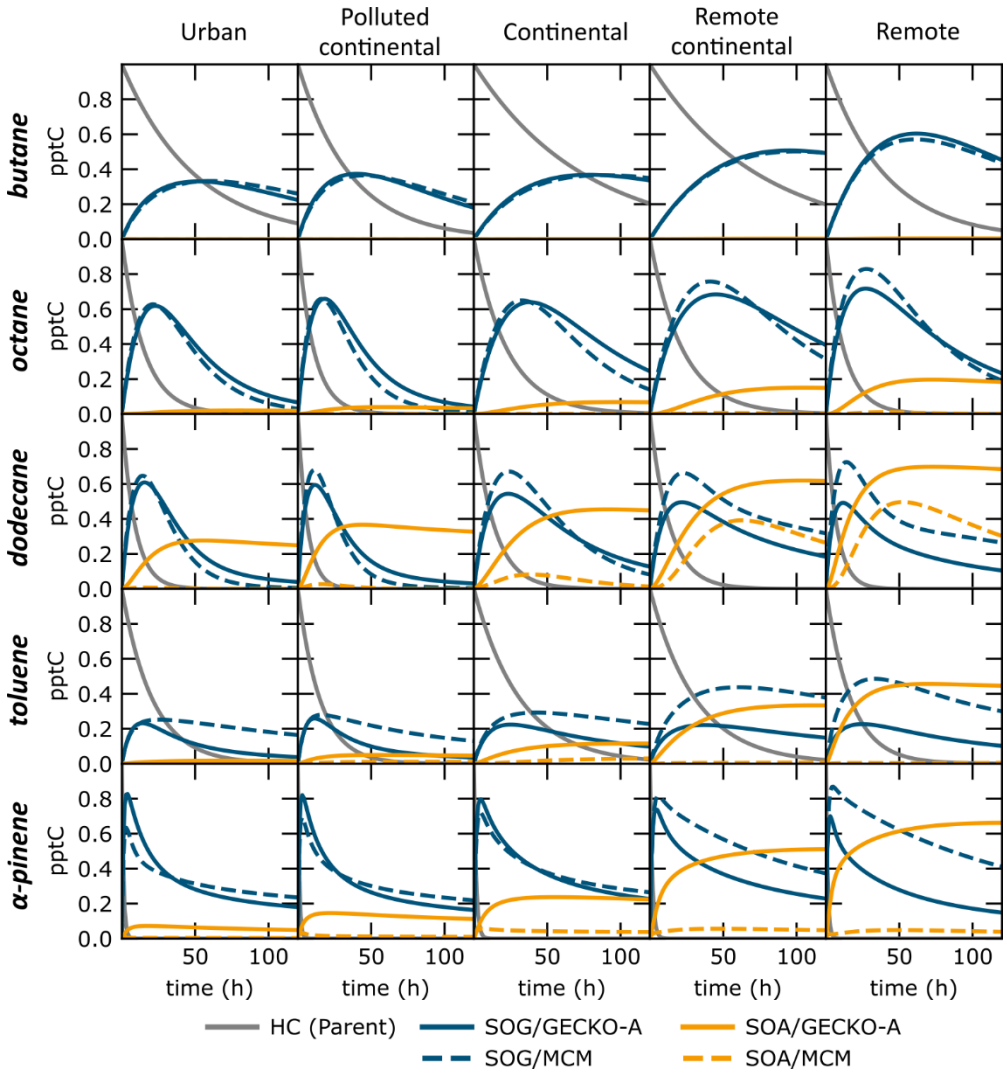


Figure 3: Temporal evolution of the organic carbon budget for various parent hydrocarbons (HC) under different environmental scenarios. Left to right panels: urban, polluted continental, continental, remote continental, and remote conditions. Top to bottom panels: butane, octane, dodecane, toluene, and α-pinene as parent HC. ~~Grey~~Gray lines: decay of the parent hydrocarbon; blue lines: secondary organic gases (SOG); orange lines: secondary organic aerosols (SOA). Solid lines: GECKO-A simulation results; dashed lines: MCM simulation results.

5.2 Volatility distribution

Figure 4 shows the temporal evolution of secondary organic carbon by volatility class for different parent hydrocarbons and scenarios. The secondary organic carbon (SOC) is defined here as the sum of gaseous organic species (SOG) and particulate organic species (SOA), expressed on a carbon-atom basis. Volatility classes are defined at 298 K as: volatile organic compounds (VOC, $p_{sat} > 10^{-4}$ atm), intermediate-volatility organic compounds (IVOC, $10^{-7} < p_{sat} < 10^{-4}$ atm), semi-volatile organic compounds (SVOC, $10^{-10} < p_{sat} < 10^{-7}$ atm), low-volatility organic compounds (LVOC, $10^{-12} < p_{sat} <$

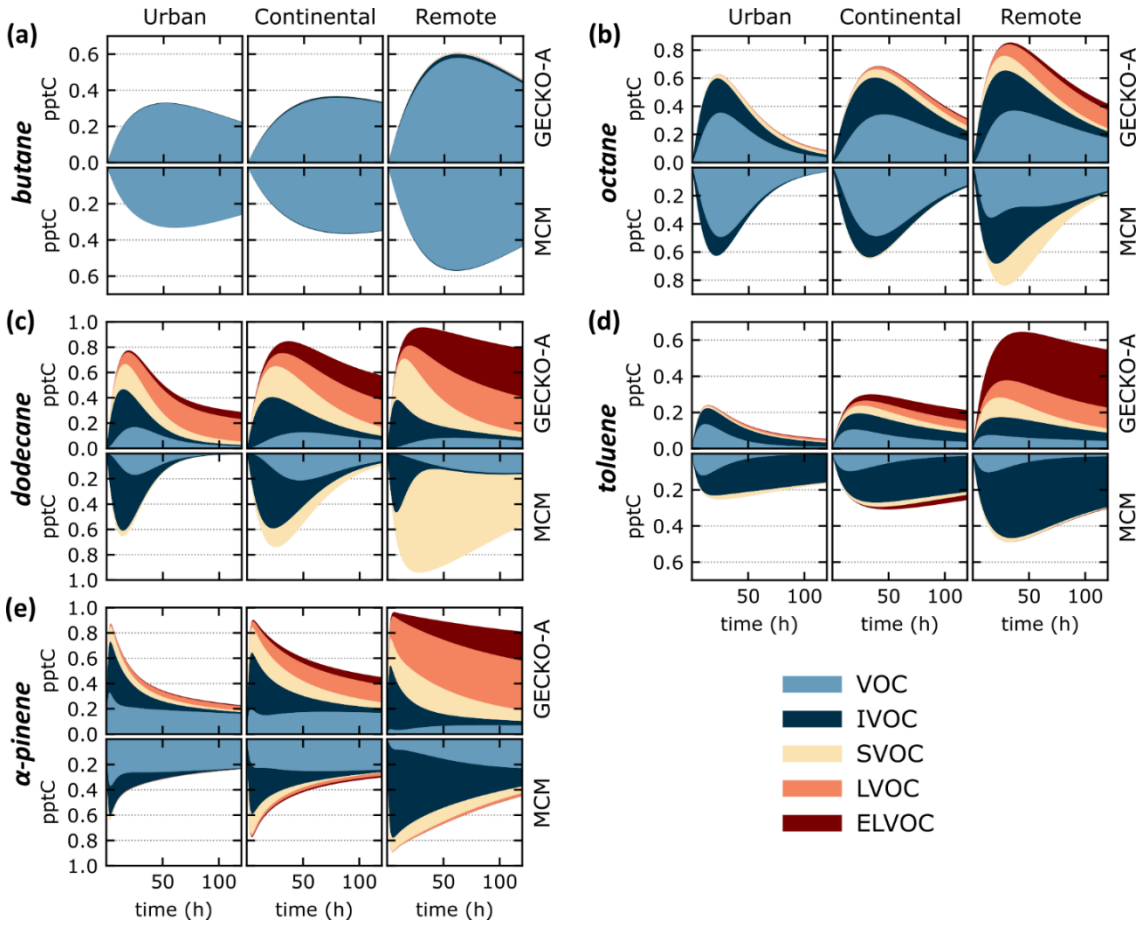


Figure 4: Temporal evolution of secondary organic carbon (SOC) by volatility group for various parent hydrocarbons under different environmental scenarios. Panels (a–e): butane, octane, dodecane, toluene, and α-pinene as parent hydrocarbons, respectively. From left to right in each panel: urban, continental, and remote scenarios. Top and bottom lines in each panel: GECKO-A and MCM simulation results, respectively. Color code: light blue (VOC), dark blue (IVOC), yellow (SVOC), orange (LVOC), red (ELVOC).

10^{-10} atm), and extremely low-volatility organic compounds (ELVOC, $p_{sat} < 10^{-12}$ atm). The distribution shifts toward lower volatilities as parent compound carbon chain length increases and from urban to remote scenarios.

Except for butane, the distributions simulated with GECKO-A and MCM exhibit distinct characteristics. The two mechanisms agree well at the beginning of oxidation, suggesting similar first-generation species volatilities. As oxidation proceeds, GECKO-A produces species of increasingly lower volatility. LVOC and ELVOC formation thus becomes the dominant fraction of secondary organic carbon with GECKO-A for dodecane, toluene, and α-pinene, particularly under continental and remote scenarios. In contrast, MCM does not produce significant amounts of LVOC and ELVOC, and the secondary organic carbon distribution is dominated by SVOC and IVOC.

At equilibrium, gas/particle partitioning of a species follows $K_{eq} = C_{part}/C_{gas}$, where C_{part} and C_{gas} are molecular concentrations per unit air volume and K_{eq} is the equilibrium constant (Sect. 3.1). The particulate fraction ζ_{part} of a species

is given by $\zeta_{part} = K_{eq}/(1 + K_{eq})$. Under the imposed conditions (Sect. 4.3), equal partitioning between the gas and condensed organic phases ($K_{eq} = 1$) is obtained for $p_{sat}^{g:o} = 1.2 \times 10^{-9}$ atm, where $p_{sat}^{g:o}$ represents the vapor pressure corresponding to 50:50 gas–organic particle partitioning (see also Sect. 5.10). Thus, at equilibrium, LVOC exhibit $\zeta_{part} > 0.92$ and ELVOC $\zeta_{part} > 0.999$, whereas IVOC remain almost exclusively in the gas phase ($\zeta_{part} < 0.01$). The characteristic timescale for gas-particle equilibrium depends on mass transfer constants k_{gp} and k_{pg}: $\tau_{eq} = 1/(k_{pg} + k_{gp})$. For the conditions simulated here, τ_{eq} is typically shorter than a few minutes, significantly shorter than the characteristic timescale of gas-phase oxidation. Dynamic gas–particle mass transfer (Sect. ~~4.1)~~3.1) therefore leads to phase partitioning close to equilibrium at any given time.

The lack of substantial LVOC and ELVOC production with MCM explains the low SOA production compared to GECKO-A. The oxidative pathways simulated by the two mechanisms progressively diverge during oxidation. These results suggest that MCM favors carbon skeleton fragmentation pathways over functionalization, ~~which is favored by~~whereas GECKO-A favors the latter. Functionalization and fragmentation are examined in the following sections.

5.3 Distribution of carbon by chain length

Parent compounds (primary carbon) are completely consumed by the end of the simulation, except for butane (Fig. 3). Figure 5 presents the time evolution of secondary carbon for different parent hydrocarbons and scenarios, with species grouped by number of carbon atoms. A significant fraction of the initial carbon is converted into CO and CO_2, reported as inorganic carbon (IC) in Fig. 5. Figure 5 shows a carbon loss for MCM simulations. This deficit typically reaches 20% at simulation end but exceeds 40% for octane (Fig. 5), mainly due to missing CO_2 representation in the reaction products.

The evolutions simulated with MCM and GECKO-A share common features. First-generation species typically preserve the parent carbon backbone. Fragmentation increases as oxidation progresses, leading to formation of organic species with shorter carbon chains and eventually CO and CO_2.

For butane oxidation, carbon skeleton size distributions simulated with GECKO-A and MCM are very similar across all scenarios. Divergences between the two mechanisms increase with parent compound size. Initially similar distributions progressively diverge as oxidation proceeds. Generally, GECKO-A simulations retain a substantial fraction of carbon in organic form at simulation end, mainly as compounds preserving the initial carbon skeleton. In contrast, the organic fraction simulated with MCM is lower, with smaller contribution from species preserving the initial carbon backbone. For a given parent compound, divergences typically increase from urban to remote scenarios.

For example, for dodecane under urban conditions, oxidation to CO and CO_2 is almost complete with MCM by 75 hours, whereas the organic fraction remains at 20% at simulation end with GECKO-A, mainly as C_{12} compounds. Under remote conditions, C_{12} compounds represent more than 60% at simulation end with GECKO-A, compared to only 35% with MCM. These trends are observed for other parent compounds, with varying magnitudes depending on their nature (Fig. 5).

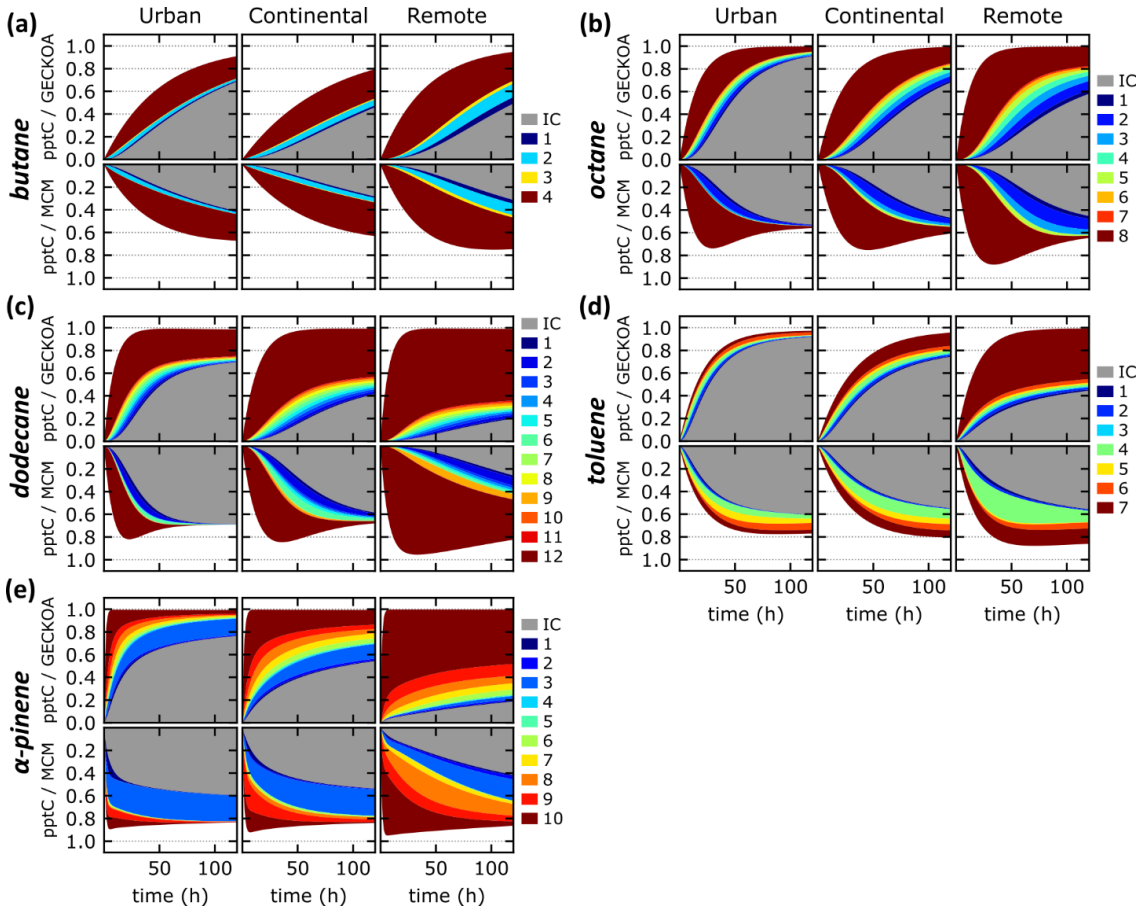


Figure 5: Temporal evolution of secondary carbon summed across gas and particle phases and grouped by carbon chain length for various parent hydrocarbons under different environmental scenarios. Panels (a–e): butane, octane, dodecane, toluene, and α-pinene as parent hydrocarbons, respectively. From left to right in each panel: urban, continental, and remote scenarios. Top and bottom rows in each panel: GECKO-A and MCM simulation results, respectively. The color gradient indicates the number of carbon atoms in the species, with dark red representing organic species with carbon chain lengths equal to the parent hydrocarbon, dark blue representing organic C_1 species, and intermediate colors for chain lengths in between. The gray color is for secondary inorganic carbon (IC = CO + CO_2).

5.4 Distribution by number of functional groups

Figure 6 shows the mixing ratio evolution of secondary species classified by the number of functional groups on their carbon skeleton. The counting is performed only for oxygenated moieties (as presented in Sect. 2.1 and listed in Fig. 7) and excludes structural features of the carbon backbone, such as C=C bonds or cyclic structures. The first generation of species is mainly composed of mono- and bifunctional species. More highly functionalized species form later from multigenerational oxidation.

For butane, the evolutions simulated with GECKO-A and MCM are similar. However, distributions by number of functional groups diverge considerably between the two mechanisms as parent compound carbon chain length and structural complexity

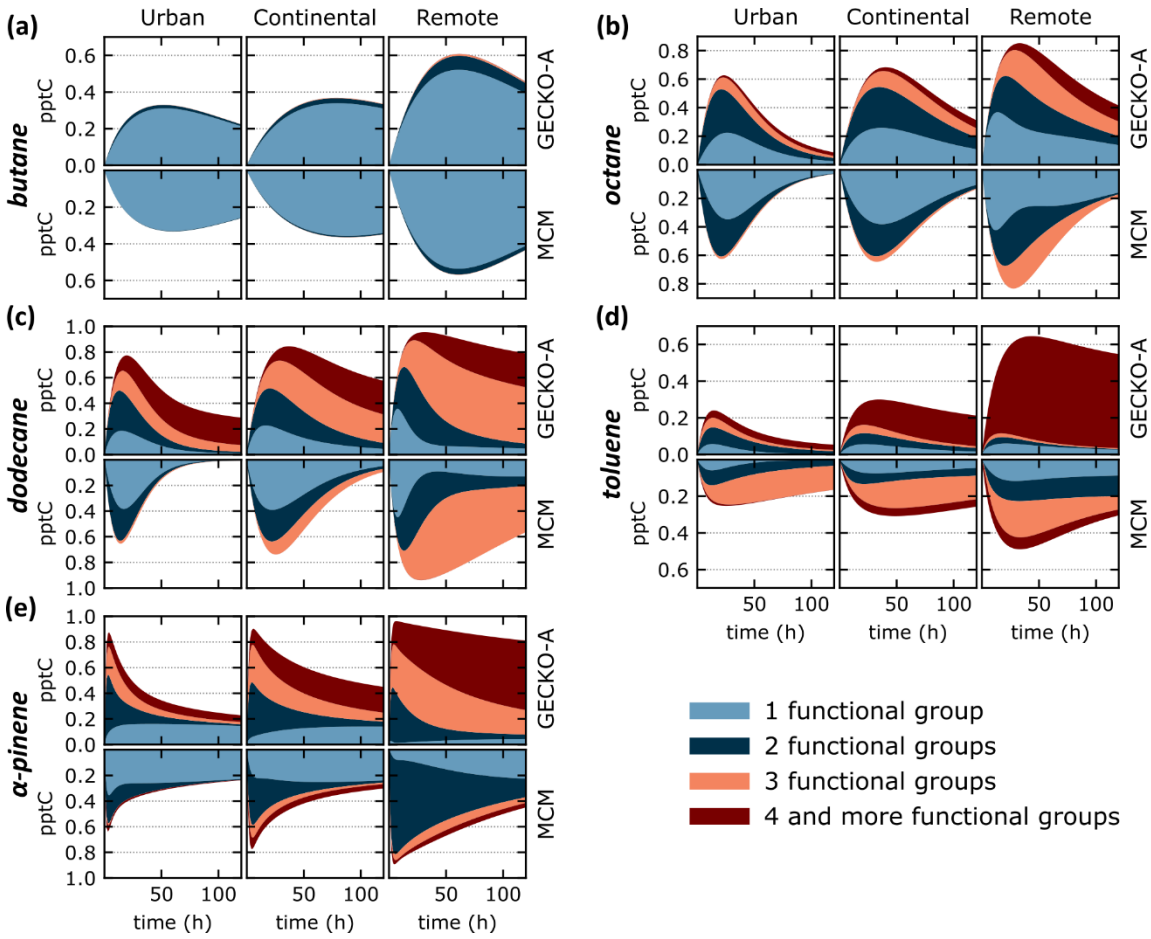


Figure 6: Temporal evolution of the mixing ratio (pptC) of secondary organic species summed across gas and particle phases and grouped by the number of functional groups on the carbon backbone for various parent hydrocarbons under different environmental scenarios. Panels (a–e): butane, octane, dodecane, toluene, and α-pinene as parent hydrocarbons, respectively. From left to right in each panel: urban, continental, and remote scenarios. Top and bottom rows in each panel: GECKO-A and MCM simulation results, respectively. The color indicates the number of functional groups borne by the carbon backbone of each species.

increase. The major difference concerns production of species bearing more than three functional groups: their contribution is negligible or low with MCM, whereas they become a major component at simulation end with GECKO-A (Fig. 6).

The oxidative pathways simulated with MCM and GECKO-A therefore differ in the degrees of functionalization and fragmentation of the parent compound carbon skeleton. MCM favors fragmentation of multifunctional species, thereby limiting low-volatility species formation (Sect. 5.2). In contrast, GECKO-A allows more extensive functionalization of the carbon skeleton before fragmentation. This process leads to formation of very low-volatility species (LVOC and ELVOC; Sect. 5.2), which generally carry four or more functional groups and retain a carbon skeleton close to that of the parent compound. This behavior explains the ~~low~~higher SOA yields simulated with ~~MCM~~GECKO-A compared to ~~GECKO-A~~MCM.

5.5 Organic carbon functionalization

Figure 7 shows the temporal evolution of the ratio between the number of functional groups and the number of carbon atoms for the secondary organic carbon (SOC). For a given functional group g (e.g., alcohol, nitrate, carboxylic acid, etc.), this ratio is computed at each time step as follows:

$$r_{g:c} = \frac{\sum_{i=1}^{N} n_{g,i} \times C_i}{\sum_{i=1}^{N} n_{c,i} \times C_i}$$

where $r_{g:c}$ is the number of functional groups g per carbon atom, N is the total number of organic compounds in the mechanism, C_i is the concentration of species i, and $n_{g,i}$ and $n_{c,i}$ are the number of functional groups g and the number of carbon atoms in species i, respectively. On this basis, a degree of organic carbon functionalization, denoted d_f, can be defined as: $d_f = \sum r_{g:c}$, where the sum is taken over all functional groups g and across both gas and particle phases.

The temporal evolution of d_f simulated with MCM and GECKO-A is similar. In all scenarios and for the different parent compounds, d_f exhibits a characteristic behavior: it increases progressively as oxidation proceeds, reaching typical values on the order of 40%. Toluene, however, shows substantially higher d_f values, reaching about 60% with MCM and exceeding 80% with GECKO-A.

The same functional groups are produced by MCM and GECKO-A, most often in comparable proportions. This similarity reflects the use of closely related protocols describing the oxidation steps and their dependence on environmental conditions. Nevertheless, some distinctive features can be identified. The ratio of aldehyde and ketone functions per carbon atom tends to be higher with MCM than with GECKO-A for linear and cyclic compounds, respectively. Conversely, nitrate functions (under urban scenarios) and hydroperoxide functions (under remote scenarios) exhibit higher ratios with GECKO-A than with MCM.

As discussed previously (Sect. 5.3), MCM tends to promote stronger fragmentation of the carbon backbone than GECKO-A. In both mechanisms, C–C bond scission most often occurs at the level of an alkoxy radical (RO·). This reaction produces a carbonyl species and an alkyl-type radical. When the latter is functionalized, it typically evolves through decomposition of the functional group. For example, in the case of an alkoxy radical bearing a nitrate group on the α-carbon, decomposition leads to the reaction >C(ONO_2)–C(O·)< → >C=O + >C=O + NO_2. Similar reactions occur for alkoxy radicals functionalized on the α-carbon by other groups, in particular hydroperoxides. The higher abundance of aldehyde and ketone functions and the lower abundance of nitrate and hydroperoxide functions in MCM compared to GECKO-A are therefore consistent with the stronger fragmentation promoted by MCM relative to GECKO-A.

In addition, in order to limit the diversity of secondary species and reduce the overall size of the mechanism, the MCM applies simplified treatments to the reactions of nitrate- and hydroperoxide-containing compounds with OH. In this mechanism, these reactions are often restricted to the direct conversion of nitrate or hydroperoxide functionalities into carbonyl groups, without considering other potential reactive sites within the molecule. This simplification leads to a progressive under-representation of nitrate and hydroperoxide functionalities in favor of aldehydes and ketones as oxidation proceeds. This aspect is discussed in more detail in the next section.

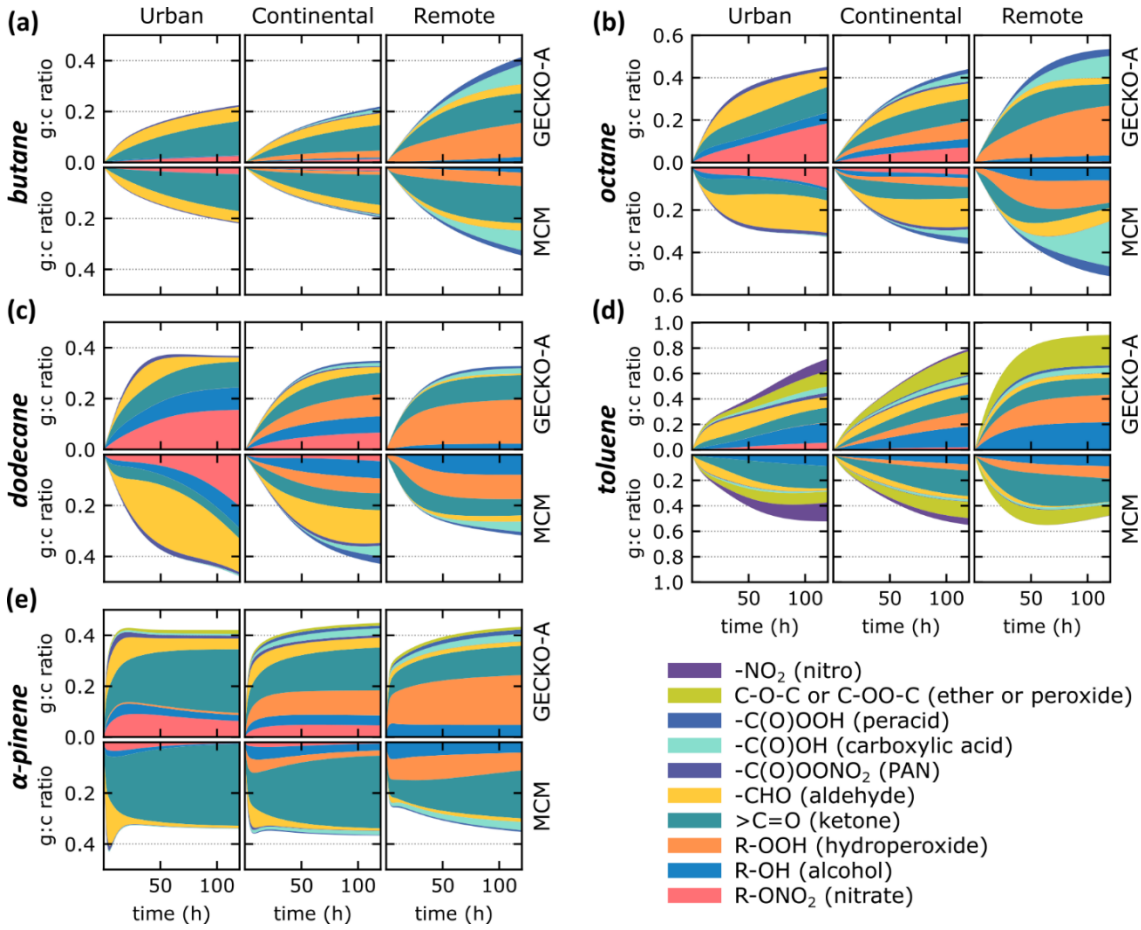


Figure 7: Temporal evolution of the functional group to carbon ratio $r_{g:c}$,for various parent hydrocarbons under different environmental scenarios. Panels (a–e) correspond to butane, octane, dodecane, toluene, and α-pinene, respectively. From left to right in each panel: urban, continental, and remote scenarios. The top and bottom rows show results from GECKO-A and MCM simulations, respectively. Colors indicate the contribution of each organic functional group to the overall degree of organic carbon functionalization assessed across gas and particle phases (see text).

5.6 Bulk gas phase metrics

5.6.a OH reactivity

To compare the gas-phase reactivity induced by SOG during oxidation, a normalized index of reactivity with OH is introduced. This index, hereafter referred to as I_{OHR}, is defined as:

$$I_{OHR} = \frac{\sum_{i=1}^{N} k_i \times [\mathrm{SOG}_i]}{k_{\mathrm{HC}} \times [\mathrm{HC}]_0}$$

where N is the number of SOG in the mechanism, k_i is the OH rate constant for the reaction of SOG species i, $[\mathrm{SOG}_i]$ is the concentration of SOG species i at time t, k_{HC} is the OH rate constant of the parent hydrocarbon, and $[\mathrm{HC}]_0$ is the initial

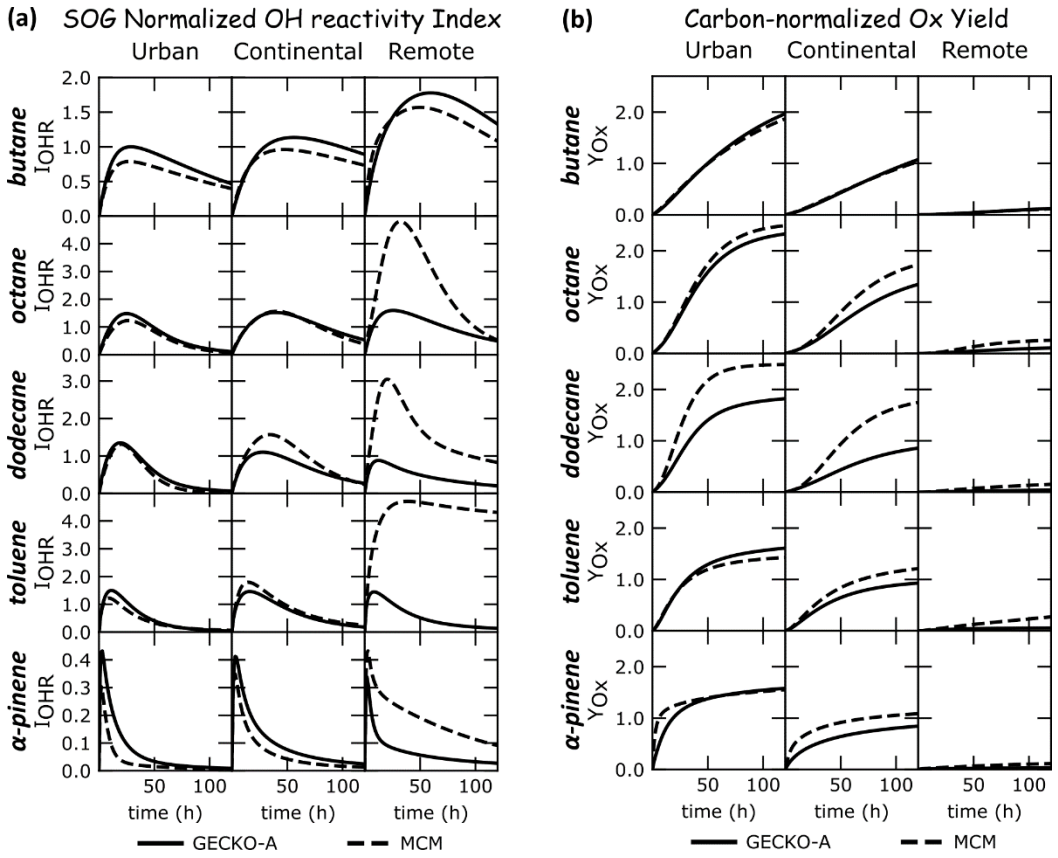


Figure 8: Temporal evolution of the SOG normalized OH reactivity index (panel a) and the carbon-normalized O_x yield (panel b) for various parent hydrocarbons under different environmental scenarios. From left to right: urban, continental, and remote scenarios. From top to bottom: butane, octane, dodecane, toluene, and α-pinene as parent compounds. Solid lines: GECKO-A simulation results; dashed lines: MCM simulation results.

concentration of the parent hydrocarbon. When $I_{OHR} = 1$, the bulk OH reactivity of the SOG pool is equal to the initial OH reactivity of the parent compound.

Figure 8a shows the temporal evolution of I_{OHR} for various parent hydrocarbons and environmental scenarios. I_{OHR} follows a characteristic profile: it initially increases as SOG are formed, reaches a maximum, and then decreases as SOG are further oxidized into CO and CO_2. The maximum values of I_{OHR} (denoted I_{OHR}^{max}) vary with the scenario and depend on the reactivity of the parent hydrocarbon. In GECKO-A simulations, I_{OHR}^{max} typically ranges between 1 and 2 for alkanes and toluene, and is around 0.4 for α-pinene.

The temporal profiles of I_{OHR} obtained with GECKO-A and MCM are in good agreement for the urban and continental scenarios. In contrast, except for butane, I_{OHR} values diverge markedly between the two mechanisms under remote conditions for all species. For example, for toluene, I_{OHR}^{max} reaches 4.7 with MCM, compared to 1.4 with GECKO-A. Discrepancies of similar magnitude are observed for octane and dodecane (see Fig. 8a). For α-pinene, remote scenario I_{OHR} values are higher with MCM than with GECKO-A, with differences generally exceeding a factor of two after a few hours of simulation.

These differences are partly explained by the SOG mixing ratios simulated with MCM and GECKO-A (see Fig. 3). With MCM, secondary organic carbon remains predominantly in the gas phase as SOG, whereas with GECKO-A it is partitioned between SOA and SOG (see Sect. 5.1). The lower SOG concentrations simulated with GECKO-A therefore explain, to first order, the lower I_{OHR} values compared to MCM.

Some specific cases in the remote scenario deserve closer examination. For octane, MCM and GECKO-A simulate very similar SOG mixing ratios (see Fig. 3), which cannot account for the large differences in reactivity, with I_{OHR}^{max} values of 4.8 with MCM and 1.6 with GECKO-A. With MCM, two species dominate the contribution to I_{OHR}: 5-hydroperoxyoctan-3-one (C_2H_5–CO–CH_2–CH(OOH)–C_3H_7, denoted CO3C85OOH in MCM) and 2-hydroperoxy-5-hydroxyoctan-3-one (CH_3–CH(OOH)–CO–CH_2–CH(OH)–C_3H_7, denoted C81OOH). Under NO_x-free conditions, the OH-initiated oxidation of octane in MCM follows a linear pathway, producing the following successive generations of stable species:

$$\text{n-}C_8H_{18} \rightarrow C_2H_5\text{–CH(OOH)–}C_5H_{11} \rightarrow C_2H_5\text{–CO–}C_5H_{11} \rightarrow C_2H_5\text{–CO–}CH_2\text{–CH(OOH)–}C_3H_7 \text{ (CO3C85OOH).}$$

CO3C85OOH is thus a third-generation product. Its OH reaction rate constant in MCM is $k = 8.7 \times 10^{-11}$ cm^3 molecule^{-1} s^{-1}, corresponding to a lifetime of 1.1 h under remote conditions. Reaction with OH produces the peroxy radical C_2H_5–CO–CH_2–CH(OO·)–C_3H_7 (CO3C85O2), which subsequently reacts with HO_2 to regenerate CO3C85OOH:

$$\text{CO3C85OOH} + OH \rightarrow \text{CO3C85O2} + H_2O$$
$$\text{CO3C85O2} + HO_2 \rightarrow \text{CO3C85OOH} + O_2$$

The CO3C85OOH/CO3C85O2 pair acts as a catalytic cycle for the net reaction $OH + HO_2 \rightarrow H_2O + O_2$, although its impact is not quantified under the buffered conditions of the simulated scenarios. Due to this cyclic behavior, the reaction of CO3C85OOH with OH does not represent an effective sink for this species, enabling its accumulation despite its high reactivity toward OH. This behavior results from the simplifications introduced in the MCM to limit mechanism size (see Sect. 4.2), whereby so-called minor products such as hydroperoxides are often represented through simplified degradation pathways that regenerate existing species. At the time of I_{OHR}^{max}, CO3C85OOH accounts for 32% of SOG carbon and has a normalized OH reactivity of 2.6. In MCM, its effective loss is slow, dominated by photolysis (lifetime of 83 h) and by reaction of CO3C85O2 with NO (branching ratio of about 4% under remote conditions, corresponding to an effective lifetime of 27 h), leading to the formation of C81OOH.

Similarly, the reaction of C81OOH with OH is fast (lifetime of 1.8 h in MCM) but does not lead to net loss in the absence of NO_x, as it produces a peroxy radical (C81O2) that reacts with HO_2 to regenerate C81OOH. At the I_{OHR}^{max} peak, C81OOH represents 24% of SOG carbon and has a normalized OH reactivity of 1.2. Together, CO3C85OOH and C81OOH account for 79% of I_{OHR}^{max}.

This behavior is not observed with GECKO-A, where the OH reactivity of the different reactive sites within each molecule is explicitly accounted for, thereby avoiding the accumulation of species through such cyclic processes. Under remote conditions, the sum of the normalized reactivities of the ten main contributors accounts for only 0.48 out of a total I_{OHR}^{max} of 1.6, indicating that SOG reactivity is distributed over a broad range of species.

Oxidation of the other hydrocarbons under NO_x-free remote conditions with MCM leads to similar biases in gas-phase OH reactivity. For example, dodecane oxidation follows a scheme analogous to that of octane. The third-generation species 5-hydroperoxydodecan-3-one (C_2H_5–CO–CH_2–CH(OOH)–C_7H_{15}, denoted C121OOH in MCM) is formed and exhibits an OH reactivity comparable to that of CO3C85OOH, leading to an equivalent catalytic cycle between C121OOH and C121O2:

$$\text{C121OOH} + \text{OH} \rightarrow \text{C121O2} + H_2O$$
$$\text{C121O2} + HO_2 \rightarrow \text{C121OOH} + O_2$$

At the I_{OHR}^{max} peak, C121OOH has a normalized OH reactivity of 2.4 and alone accounts for 80% of I_{OHR}^{max} (3.0). Effective loss of C121OOH, through photolysis or reaction of C121O2 with NO, leads to the formation of CH_3–CH(OOH)–CO–CH_2–CH(OH)–C_7H_{15} (C122OOH), which is the only substantial contributor to SOA. Its contribution to I_{OHR}^{max} is modest (6%), but this species becomes a major contributor to I_{OHR} after about 50 h of simulation, as it progressively evaporates from the particle phase (see Fig. 8).

An identical catalytic cycle is also observed during α-pinene oxidation under NO_x-free conditions. The species 2-hydroperoxypinan-3-ol (denoted APINAOOH in MCM) is a first-generation product, formed with a yield of 57% following OH addition to one of the C=C bond carbons and subsequent formation of a hydroperoxide function on the other carbon. Reaction of APINAOOH with OH produces 2-peroxypinan-3-ol (APINAO2), which regenerates APINAOOH via reaction with HO_2. As a result, the effective consumption of APINAOOH in MCM is slow and is controlled mainly by APINAOOH photolysis and by reaction of APINAO2 with NO. Under remote conditions, the effective lifetime of APINAOOH exceeds 50 h, making this species the dominant contributor to I_{OHR}^{max}.

Finally, a cyclic process also arises during toluene oxidation with MCM under NO_x-free conditions. It is driven by the chemistry of methylcatechol (MCATECHOL), a second-generation product of OH-initiated toluene oxidation, formed with a yield of 13%. The chemical mechanism of MCATECHOL in MCM is detailed in Fig. 9. Oxidation leads to the formation of a hydroperoxide (MCATEC1OOH), which, as in the previous cases, enters a catalytic cycle with its corresponding peroxy radical:

$$\text{MCATEC1OOH} + \text{OH} \rightarrow \text{MCATEC1O2} + H_2O$$
$$\text{MCATEC1O2} + HO_2 \rightarrow \text{MCATEC1OOH} + O_2$$

In MCM, the rate constant for the reaction MCATEC1OOH + OH is close to the collision limit, resulting in a short lifetime of approximately 0.5 h for MCATEC1OOH. Effective loss of the MCATEC1OOH / MCATEC1O2 pair leads to the formation of MCATEC1O (see Fig. 9), which in turn regenerates MCATEC1OOH. In the absence of NO_x, no pathway results in an effective loss of the MCATEC1OOH / MCATEC1O2 / MCATEC1O pool. The only effective sink is the reaction of MCATEC1O with NO_2 (see Fig. 9). Under remote-scenario conditions, the branching ratio of this pathway is particularly small, below 0.02%, making the effective loss of MCATEC1OOH extremely slow. At the time of the I_{OHR}^{max} peak, the normalized OH reactivity of MCATEC1OOH is 3.6, accounting for 77% of I_{OHR}^{max}, well above the values obtained with GECKO-A (see Fig. 8). Similar anomalously high persistence of certain products in the MCM has previously been reported for oxidation pathways derived from m-xylene under NO_x-free conditions (Peng et al., 2021), arising from identical cyclic processes to those observed here for methylcatechol.

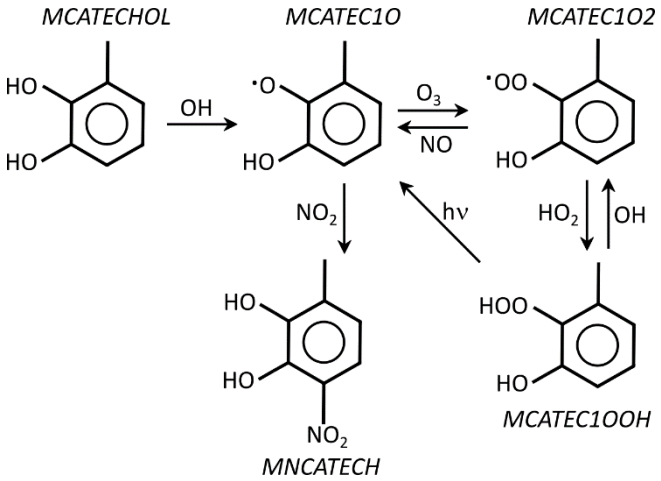


Figure 9: Evolution of methyl catechol (MCATECHOL) by reaction with OH in the MCM mechanism.

The absence of mechanisms leading to an effective loss of certain SOG under NO_x-free conditions in MCM thus explains the high I_{OHR} values observed in the remote scenario. This behavior is directly related to the development choices made during the manual construction of the MCM, where the mechanism was strategically simplified in a framework primarily aimed at investigating multi-day ozone formation over northwestern Europe. The anthropogenic origins of aromatics and alkanes make their oxidation under low-NO_x conditions unlikely in the atmosphere. For these species, the remote scenario is therefore likely unrepresentative of their typical oxidation conditions. In contrast, NO_x-free conditions can occasionally occur in certain reactor environments. In such cases, alternative reaction pathways for specific species must be added to the MCM to prevent their anomalous accumulation (Peng et al., 2021).

5.6.b O_x production

To quantify O_x production (O_3 + NO_2), whose concentrations are constrained in the scenarios (see Sect. 4.3), a tracer species, denoted POx, is introduced into the mechanism. This tracer quantifies the conversion of NO to NO_2 by peroxy radicals (RO_2, HO_2) formed during oxidation of the parent compound. POx has no loss pathways and therefore accumulates as oxidation proceeds, representing the total flux through these conversion channels. To compare O_x production between MCM and GECKO-A for the different parent hydrocarbons, a carbon-normalized O_x yield (denoted Y_{Ox}) is defined as:

$$Y_{Ox} = \frac{[\mathrm{POx}]_t}{n_{\mathrm{c,HC}} \times [\mathrm{HC}]_0}$$

where $[\mathrm{POx}]_t$ is the cumulative concentration of O_x produced at time t, $n_{\mathrm{c,HC}}$ is the number of carbon atoms in the parent compound, and $[\mathrm{HC}]_0$ is the initial concentration of the parent hydrocarbon.

Figure 8b shows the temporal evolution of Y_{Ox} for the different hydrocarbons and scenarios. Y_{Ox} follows a characteristic trend: it increases progressively as oxidation proceeds and reaches a plateau once oxidation is complete. For a given hydrocarbon, Y_{Ox} decreases from the urban to the remote scenario, since O_x production depends directly on NO concentrations. Under NO_x-saturated conditions, the branching ratio for peroxy radical reactions with NO approaches unity, maximizing Y_{Ox}. In this

regime, oxidation of a carbon atom with an oxidation state of −II (e.g., a methylene group, $-CH_2-$) to +II (a carbonyl group, −CO−) produces 2 O_x, while complete oxidation to CO_2 produces 3 O_x.

The urban scenario lies in a NO_x-saturated regime (see Table 3). After 120 h of simulation, Y_{Ox} values obtained with GECKO-A under this scenario are around 2.0 for alkanes and about 1.5 for toluene and α-pinene. These values are close to the theoretical maxima and are consistent with the high degree of oxidation reported in previous sections, such as the strong functionalization of organic carbon (see Fig. 7) and the substantial contribution of CO and CO_2 to the carbon budget (see Fig. 5). In the continental scenario, Y_{Ox} is approximately a factor of two lower than in the urban scenario, while it becomes nearly negligible in the remote scenario.

Agreement between MCM and GECKO-A for Y_{Ox} is excellent for butane under all scenarios, and remains satisfactory for all hydrocarbons in the urban scenario. The main divergence is observed for dodecane: at the end of the simulation, Y_{Ox} reaches 1.8 with GECKO-A and 2.5 with MCM. This difference is explained by the sequestration of organic carbon into SOA, which removes it from further gas-phase oxidation. On a carbon basis, the SOA yield reaches 25% with GECKO-A but is negligible with MCM (see Fig. 3). The formation of low-volatility species (LVOC and ELVOC), which constitute a substantial fraction of SOA (see Fig. 4), therefore ~~slow~~slows down the oxidation of organic carbon. This effect is even more pronounced in the continental scenario, where differences in SOA yield increase further: SOA reaches 45% at the end of the simulation with GECKO-A, compared to only 1% with MCM (see Fig. 3). The resulting difference in Y_{Ox} then reaches a factor of two (0.86 with GECKO-A versus 1.7 with MCM). A similar mechanism explains the divergence in Y_{Ox} for α-pinene in the continental scenario.

Overall, MCM and GECKO-A yield comparable O_x production potentials for precursors with a low SOA formation potential. In contrast, SOA formation inhibits gas-phase oxidation and, consequently, O_x production in NO_x-rich environments. Differences in Y_{Ox} therefore increase as gas-phase oxidation progresses, i.e., as more low-volatility organic species contributing to SOA formation are produced.

5.7 Shared and mechanism-specific species

In the previous sections, the comparison focused on the bulk properties and overall composition of organic matter. The objective here is to compare the two mechanisms at the molecular scale. However, MCM, which is manually developed, relies on strategic simplifications intended to limit its size (see Sect. 4.2). For example, the six distinct RO_2 radicals produced by the reaction of dodecane with OH are lumped into a single surrogate species, with the peroxy group positioned on the third carbon atom. As a result, first-generation isomers are represented by a single generic isomer that encompasses all possible positional isomers. The six distinct organic nitrates formed via RO_2 + NO reactions are for example represented solely by 3-nitrooxydodecane. This lumping approach is applied generation after generation, reducing the complexity of the mechanism at the expense of a progressive loss of information regarding the position of functional groups along the carbon skeleton and, consequently, regarding the specific reactivity of each species.

In contrast, the GECKO-A configuration used here for the mechanism generation explicitly treats all reactions without lumping (see Sect. 4.1). A direct molecular-level comparison is therefore not straightforward, as the nature of the species differs between the mechanisms: they are often surrogate species in MCM but always explicit in GECKO-A. To enable a consistent

comparison, the concentrations of all positional isomers are summed for each mechanism, thereby grouping all positional isomers into a single effective concentration. As already mentioned in Sect. ~~3.1,~~2.2, two species are considered positional isomers if they share the same molecular formula and the same set of functional groups. Hereafter, all positional isomers are referred to as an isomer group.

The GECKO-A mechanism includes a large number of organic species that contribute negligibly to the overall mass budget. A filtering criterion is therefore applied to identify non radical species of interest. To identify species with a non-negligible contribution, a carbon-integrated abundance (CIA) is defined as:

$$CIA_i = \frac{n_{c,i} \int C_i(t)dt}{n_{c,HC} \int [HC]_0 dt}$$

where $n_{c,i}$ is the number of carbon atoms in isomer group i, and $C_i(t)$ is its concentration at time t. Time integration, spanning from the start to the end of the simulation, captures both the abundance and persistence of isomer groups during oxidation. Thus, CIA represents the fraction of the initial carbon budget that is associated, over the course of the simulation, with a given isomer group. An isomer group is retained if CIA exceeds a threshold of 1×10^{-4} and only non-radical species are examined.

Figure 10 shows the number of isomer groups contributing to the organic carbon budget in the continental scenario, which was selected because it spans a wide diversity of species formed under both NO_x-saturated and NO_x-limited conditions. The corresponding results for the urban and remote scenarios are provided in the Supplementary Information (Fig. ~~S1~~S2 and ~~S2~~S3). The sets displayed in Fig. 10 distinguish between isomer groups that are exclusive to GECKO-A, exclusive to the MCM, or common to both mechanisms.

As expected, the number of isomer groups increases with the size and structural complexity of the parent hydrocarbon. Accordingly, this number rises from 75 (34) for butane to 675 (66) for α-pinene in GECKO-A (MCM).

With GECKO-A, and with the exception of butane, the number of isomer groups increases with the degree of functionalization of the carbon skeleton, from monofunctional to polyfunctional species (defined here as species bearing more than two functional groups). For example, in the case of dodecane, the numbers of mono-, bi-, and polyfunctional isomer groups are 68, 149, and 249, respectively. This trend is not systematically observed in the MCM, where bifunctional isomer groups often dominate. For dodecane, for instance, 31 bifunctional isomer groups are identified, compared to only 18 groups bearing 3 or more functional groups. This behavior reflects the tendency of the MCM to favor fragmentation over functionalization, in contrast to GECKO-A, as already discussed in Sect. 5.3 and 5.4.

Most monofunctional isomer groups represented in the MCM are also present in GECKO-A. This also holds for the bi- and polyfunctional isomer groups formed during alkane oxidation. For example, in the case of octane, only 3 isomer groups out of a total of 52 in the MCM are not identified in GECKO-A. The discrepancies become more pronounced for structurally complex precursors such as α-pinene: in this case, 20 isomer groups out of 66 in the MCM are not found in GECKO-A, including 4 bifunctional groups and 16 polyfunctional groups.

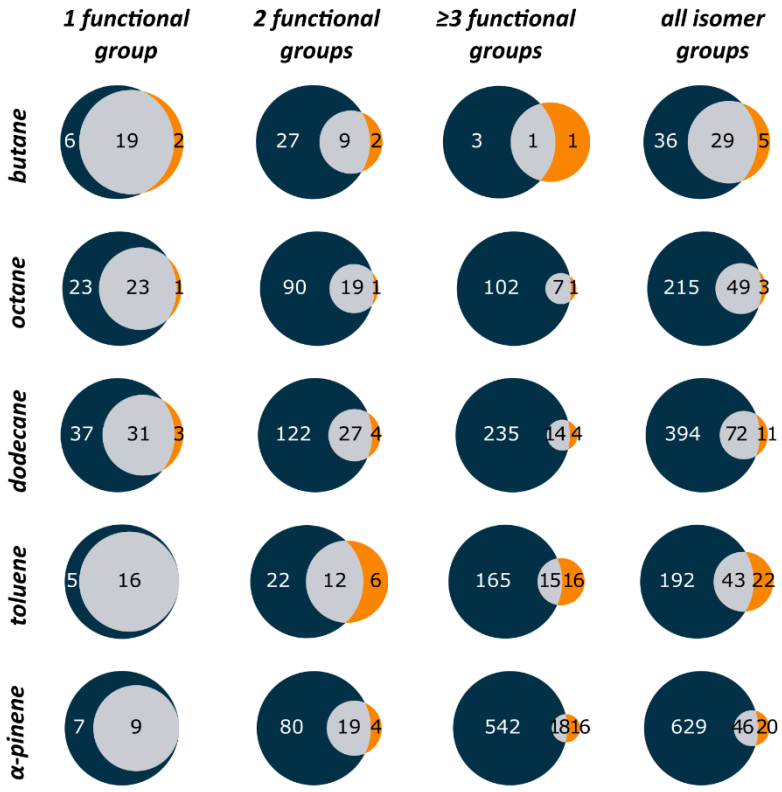


Figure 10: Number of non-radical isomer groups contributing to the carbon budget in the continental scenario, partitioned by presence in GECKO-A only, MCM only, or both mechanisms. Left to right panels: monofunctional isomers, bifunctional isomers, polyfunctional isomers (more than two functional groups) and all isomer groups. Top to bottom panels: butane, octane, dodecane, toluene, and α-pinene as parent compounds. Blue: isomer groups present only in GECKO-A; orange: isomer groups present only in MCM; gray: isomer groups common to both mechanisms.

The fact that the majority of MCM isomer groups are encompassed within those present in GECKO-A suggests that the chemistry represented by the MCM is qualitatively captured by GECKO-A. This is particularly true for the early generations of oxidation, for which the mono- and bifunctional isomer groups in GECKO-A largely overlap with those of the MCM. However, the substantially larger number of isomer groups in GECKO-A indicates that the use of surrogate species in the MCM alone cannot account for the differences observed between the two mechanisms. The greater diversity of isomer groups in GECKO-A reflects a more comprehensive and detailed chemical description, especially for polyfunctional species formed through functionalization pathways—processes that are largely simplified or neglected in the MCM.

5.8 Major isomer groups

To identify the dominant isomer groups in each mechanism and assess their consistency, all isomer groups were ranked according to their CIA values. The ranking was established using the combined set of CIA values from the MCM and

GECKO-A. Specifically, for each isomer group, the maximum CIA value obtained across the two mechanisms was retained, and the groups were ranked in descending order based on this value.

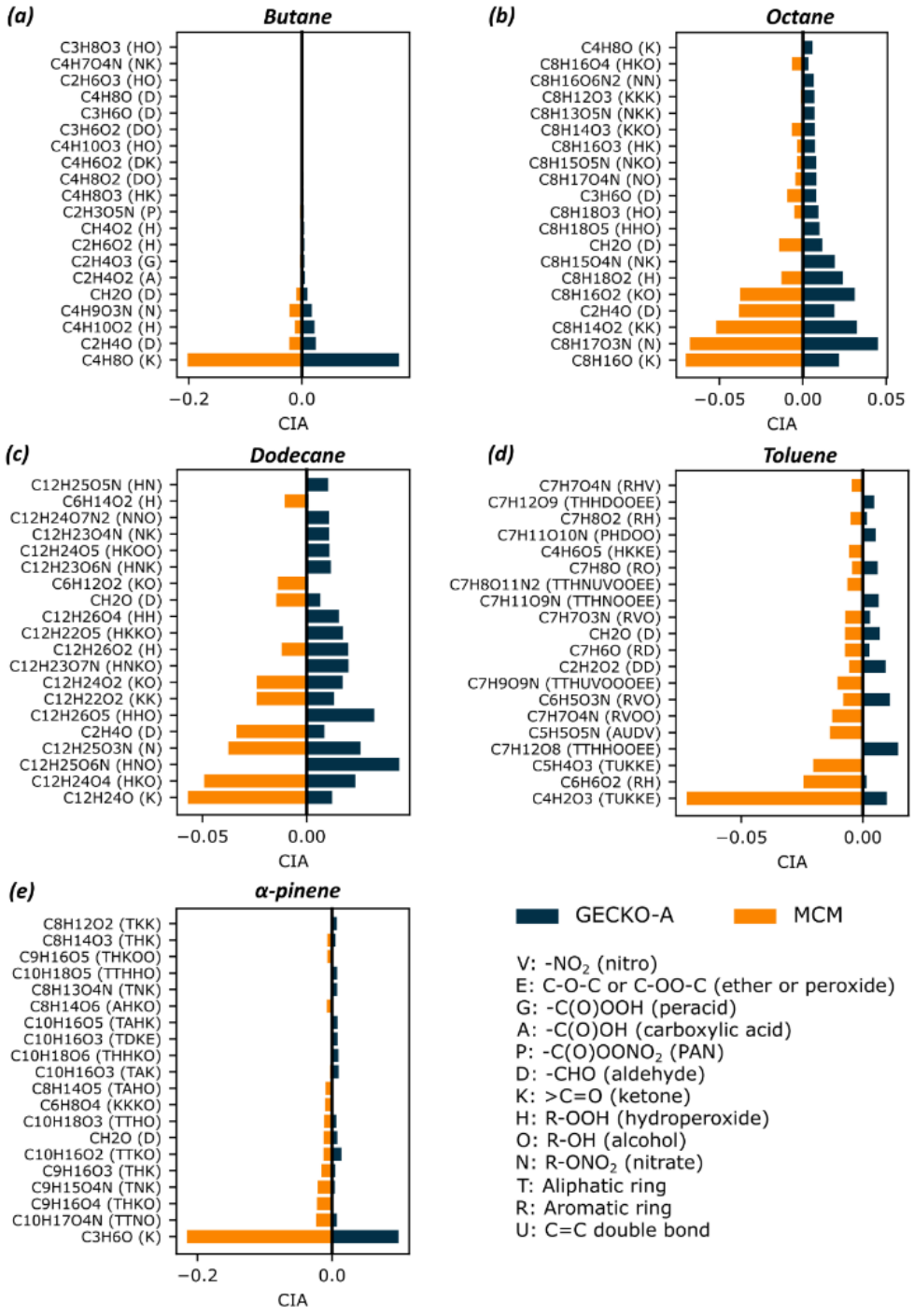
Figure 11 shows the top 20 isomer groups identified for the continental scenario and for the different parent hydrocarbons. The results are displayed as mirrored bar plots, comparing the CIA values obtained with the MCM and GECKO-A for each group. Each isomer group is identified by its molecular formula and by the nature of the functional groups present on the carbon skeleton (see Fig. 11). The groups are ordered by CIA, without distinction between mechanisms. This representation allows the dominant groups to be readily identified and their consistency between the two mechanisms to be evaluated.

For alkanes, good agreement is observed between the two mechanisms for the major isomer groups, with CIA values of comparable magnitude in the MCM and GECKO-A. A major difference, however, lies in the presence—exclusively in GECKO-A—of dominant isomer groups that are absent from the MCM. These groups mainly correspond to difunctional nitrate-containing species and to polyfunctional compounds. These discrepancies stem from the design choices made during the manual development of the MCM, which favors reaction pathways leading to the removal of such functional groups in order to limit their propagation across successive generations and thereby constrain the complexity of secondary products.

A similar behavior is observed for hydroperoxides. In the MCM, two hydroperoxides are produced as first-generation products from octane and dodecane oxidation: an alkyl hydroperoxide and a 1,4-hydroxy hydroperoxide. As for their nitrate counterparts, subsequent reactions lead to the removal of the hydroperoxide functionality, yielding a ketone and a 1,4-hydroxy ketone, respectively. In contrast, the explicit treatment of all reactive sites in GECKO-A favors the persistence of hydroperoxide functionalities over multiple generations. Consequently, polyfunctional species bearing one or more hydroperoxide groups form major isomer groups in GECKO-A (see Fig. 11).

Discrepancies among the dominant isomer groups are more pronounced for toluene, reflecting substantially different reaction mechanisms for these compounds. In the case of toluene, the MCM treatment of benzaldehyde—a first-generation product—leads to the accumulation of phenyl hydroperoxide (C6H5OOH in the MCM) through a cyclic process similar to that described for methylcatechol (Fig. 9). This mechanism, detailed in Fig. 12, involves phenylperoxy (C6H5O2) and phenoxy (C6H5O) radicals. The effective loss of C6H5OOH mainly proceeds via the reaction of the phenoxy radical with NO_2, in competition with its reaction with O_3 (Fig. 12). Under continental scenario conditions, the reaction with O_3 is strongly favored (branching ratio ≈ 95%), resulting in an anomalously slow effective loss of phenyl hydroperoxide in the MCM (effective lifetime of ~90 h). Consequently, this compound ranks as the second most important species according to CIA in the MCM (Fig. 11~~).~~d).

Maleic anhydride ($C_4H_2O_3$; Fig. 11~~)~~d) emerges as a major contributor to the carbon budget in both toluene mechanisms. In both cases, it originates from the evolution of unsaturated 1,4-dicarbonyl compounds, which are major first-generation products formed during aromatic ring opening. However, the integrated abundance of maleic anhydride is substantially higher in the MCM than in GECKO-A. This discrepancy is primarily explained by differences in the loss rate constants used in the two mechanisms. The rate constant adopted in GECKO-A, estimated from the SAR of Jenkin et al. (2018a), is approximately an order of magnitude larger than that used in the MCM (1.7×10^{-11} vs. 1.4×10^{-12} cm^3 $molecule^{-1}$ s^{-1}). As a result, under continental conditions, maleic anhydride is the most abundant secondary organic species in the MCM (CIA = 0.072), whereas its abundance is more moderate in GECKO-A (CIA = 0.010).


(a)
Butane
C3H8O3 (HO)
C4H7O4N (NK)
C2H6O3 (HO)
C4H8O (D)
C3H6O (D)
C3H6O2 (DO)
C4H10O3 (HO)
C4H6O2 (DK)
C4H8O2 (DO)
C4H8O3 (HK)
C2H3O5N (P)
CH4O2 (H)
C2H6O2 (H)
C2H4O3 (G)
C2H4O2 (A)
CH2O (D)
C4H9O3N (N)
C4H10O2 (H)
C2H4O (D)
C4H8O (K)
−0.2
0.0
CIA
(b)
Octane
C4H8O (K)
C8H16O4 (HKO)
C8H16O6N2 (NN)
C8H12O3 (KKK)
C8H13O5N (NKK)
C8H14O3 (KKO)
C8H16O3 (HK)
C8H15O5N (NKO)
C8H17O4N (NO)
C3H6O (D)
C8H18O3 (HO)
C8H18O5 (HHO)
CH2O (D)
C8H15O4N (NK)
C8H18O2 (H)
C8H16O2 (KO)
C2H4O (D)
C8H14O2 (KK)
C8H17O3N (N)
C8H16O (K)
−0.05
0.00
0.05
CIA
(c)
Dodecane
C12H25O5N (HN)
C6H14O2 (H)
C12H24O7N2 (NNO)
C12H23O4N (NK)
C12H24O5 (HKOO)
C12H23O6N (HNK)
C6H12O2 (KO)
CH2O (D)
C12H26O4 (HH)
C12H22O5 (HKKO)
C12H26O2 (H)
C12H23O7N (HNKO)
C12H24O2 (KO)
C12H22O2 (KK)
C12H26O5 (HHO)
C2H4O (D)
C12H25O3N (N)
C12H25O6N (HNO)
C12H24O4 (HKO)
C12H24O (K)
−0.05
0.00
CIA
(d)
Toluene
C7H7O4N (RHV)
C7H12O9 (THHDOOEE)
C7H8O2 (RH)
C7H11O10N (PHDOO)
C4H6O5 (HKKE)
C7H8O (RO)
C7H8O11N2 (TTHNUVOOEE)
C7H11O9N (TTHNOOEE)
C7H7O3N (RVO)
CH2O (D)
C7H6O (RD)
C2H2O2 (DD)
C7H9O9N (TTHUVOOOEE)
C6H5O3N (RVO)
C7H7O4N (RVOO)
C5H5O5N (AUDV)
C7H12O8 (TTHHOOEE)
C5H4O3 (TUKKE)
C6H6O2 (RH)
C4H2O3 (TUKKE)
−0.05
0.00
CIA
(e)
α-pinene
C8H12O2 (TKK)
C8H14O3 (THK)
C9H16O5 (THKOO)
C10H18O5 (TTHHO)
C8H13O4N (TNK)
C8H14O6 (AHKO)
C10H16O5 (TAHK)
C10H16O3 (TDKE)
C10H18O6 (THHKO)
C10H16O3 (TAK)
C8H14O5 (TAHO)
C6H8O4 (KKKO)
C10H18O3 (TTHO)
CH2O (D)
C10H16O2 (TTKO)
C9H16O3 (THK)
C9H15O4N (TNK)
C9H16O4 (THKO)
C10H17O4N (TTNO)
C3H6O (K)
−0.2
0.0
CIA
GECKO-A
MCM
V: -NO2 (nitro)
E: C-O-C or C-OO-C (ether or peroxide)
G: -C(O)OOH (peracid)
A: -C(O)OH (carboxylic acid)
P: -C(O)OONO2 (PAN)
D: -CHO (aldehyde)
K: >C=O (ketone)
H: R-OOH (hydroperoxide)
O: R-OH (alcohol)
N: R-ONO2 (nitrate)
T: Aliphatic ring
R: Aromatic ring
U: C=C double bond

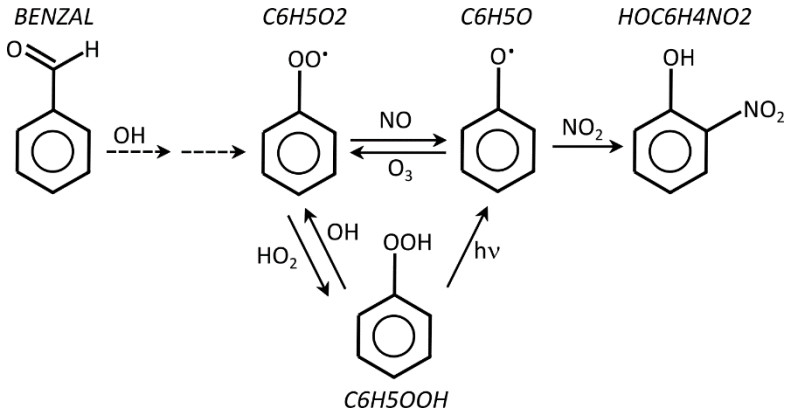


Figure 12: Evolution of benzaldehyde (BENZAL) by reaction with OH in the MCM mechanism.

Figure 11: Major isomer groups contributing to the all-phases carbon budget in the continental scenario. Panels (a–e) correspond to butane, octane, dodecane, toluene, and α-pinene, respectively. Isomer groups are ranked by decreasing carbon-integrated abundance (CIA, see text) without mechanism distinction. Labels report the molecular formula of each group, and codes in parentheses indicate the functional groups on the carbon skeleton as specified in the legend. Left bars (orange): MCM; right bars (blue): GECKO-A.

Nevertheless, available evidence suggests that the OH reactivity of maleic anhydride is overestimated in both mechanisms. A recent kinetic database recommends a rate constant of 3.9×10^{-13} cm^3 $molecule^{-1}$ s^{-1} (https://data.eurochamp.org/sar-database/; last access May 7, 2026), nearly an order of magnitude lower than the MCM value. A sensitivity test performed by imposing this reference value in GECKO-A results in maleic anhydride becoming the most abundant organic species, with a CIA of 0.080. This highlights current limitations of structure–activity relationships (SARs) for estimating rate constants of polyfunctional molecules in mechanism generators, as discussed e.g. in Ervens et al. (2024), McGillen et al. (2020) and Vereecken et al. (2018). In addition, the formation pathways of maleic anhydride differ between the two mechanisms. In the MCM, it is produced via cyclization of unsaturated carbonyl acyloxy radicals of the form –CO–C=C–C(O)O·, generated by photolysis and by reactions of unsaturated 1,4-dicarbonyls with OH and NO_3. In GECKO-A, by contrast, maleic anhydride is formed exclusively through photolysis of unsaturated 1,4-dicarbonyls, as described in Aumont et al. (2005). The protocol currently implemented in GECKO-A for the chemistry of these compounds appears no longer fully consistent with the current state of knowledge and will need to be revised in future versions.

Moreover, the oxidation of toluene with GECKO-A leads to the formation of a bicyclic compound featuring a peroxide bridge and two alcohol and hydroperoxide functionalities ($C_7H_{12}O_8$; Fig. 13). This second-generation species, whose formation is illustrated in Fig. 13, has extremely low ~~volatily~~ volatility (ELVOC) and becomes sequestered in the particle phase. Its subsequent oxidation is therefore negligible, and the compound accumulates over the course of toluene oxidation. Under continental conditions, $C_7H_{12}O_8$ is the most abundant species in GECKO-A (CIA = 0.015). An isomer of its precursor ($C_7H_{10}O_5$) is also produced in the MCM (denoted TLBIPEROOH). However, ~~its~~the oxidation of this isomer by OH leads to the removal of the hydroperoxide moiety, forming a ketone, in a manner similar to that discussed above for alkanes. Subsequent oxidation in the MCM promotes molecular fragmentation, thereby bypassing the formation of highly oxidized isomers.

Finally, under continental conditions, α-pinene is consumed by O_3 and OH in proportions of 55% and 45%, respectively. Note that the OH + α-pinene reactions are predefined in GECKO-A, as described in McVay et al. (2016). Among secondary organic

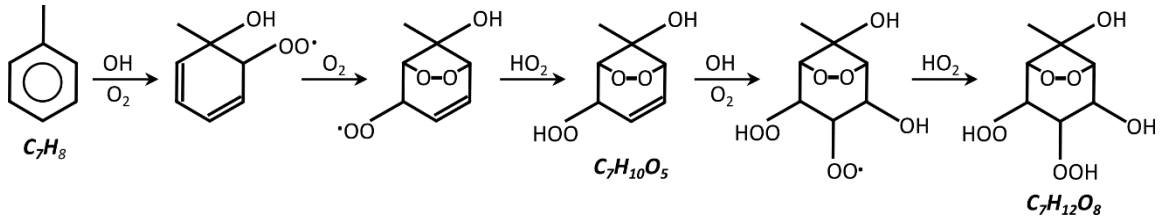


Figure 13: Typical reaction pathway from toluene leading to ELVOC formation under low-NO_x conditions in GECKO-A.

species, acetone dominates the organic carbon budget, with CIA values of 0.21 and 0.10 in the MCM and GECKO-A, respectively. Its formation results from multiple reaction pathways. The comparable CIA values indicate a reasonable level of agreement between the two mechanisms regarding the overall oxidative processing.

Both mechanisms agree on the dominant groups of simple difunctional species formed via OH addition to the C=C double bond while preserving the bicyclic C_{10} structure of α-pinene. This includes, for example, hydroxynitrates (TTNO), hydroxyhydroperoxides (TTHO), and hydroxyketones (TTKO; see Fig. 11~~).~~e).

In contrast, the two mechanisms diverge markedly in their treatment of ozonolysis. In the MCM, ozonolysis predominantly leads to the formation of C_9 species (or smaller carbon skeletons), some of which are weakly reactive and rank among the five species with the highest CIA values. In GECKO-A, a substantial fraction of ozonolysis products retains a C_{10} skeleton, resulting in the formation of carboxylic acids and formates that appear among the ten species with the highest CIA values.

With the exception of acetone, CIA values remain low in the α-pinene mechanisms. This reflects the large diversity of oxidation products, each formed with a low yield and often exhibiting high reactivity. As a consequence, the CIA metric appears to be of limited relevance for assessing convergence or divergence between the two mechanisms in this case. Nevertheless, inspection of Fig. 11e shows that the MCM tends to produce, among its most abundant species, compounds with carbon skeletons smaller than C_{10}, consistent with the trends highlighted in Fig. 5. In contrast, GECKO-A retains a significant fraction of secondary organic carbon in the form of C_{10} compounds, including polyfunctional species that rank among the most abundant and contribute to SOA formation, in agreement with the results shown in Fig. 4 and 5.

5.9 Complexity of secondary organic carbon

The isomer-group analysis provides only a partial assessment of the diversity of species involved in the secondary organic carbon mass budget. To refine this analysis, all organic species were ranked in decreasing order of their CIA values, irrespective of phase. Figure 14 shows the cumulative contribution of the main contributors (grouped in sets of 10, 100, and 1000 species) to the SOC mass budget for GECKO-A and the MCM under the different scenarios.

As expected, the MCM SOC budget is largely confined to a limited number of species. For example, in the oxidation of dodecane, the 10 most abundant species account for 88%, 68%, and 90% of the total SOC budget in the urban, continental, and remote scenarios, respectively (see Fig. 14). Similar trends are observed for the other parent hydrocarbons: in all scenarios, the 100 most abundant species account for essentially the entire SOC budget in the MCM.

In contrast, although the 10 dominant species also explain most of the SOC budget for butane in GECKO-A, their relative contribution decreases sharply with increasing carbon chain length of the parent hydrocarbon. For dodecane in GECKO-A,

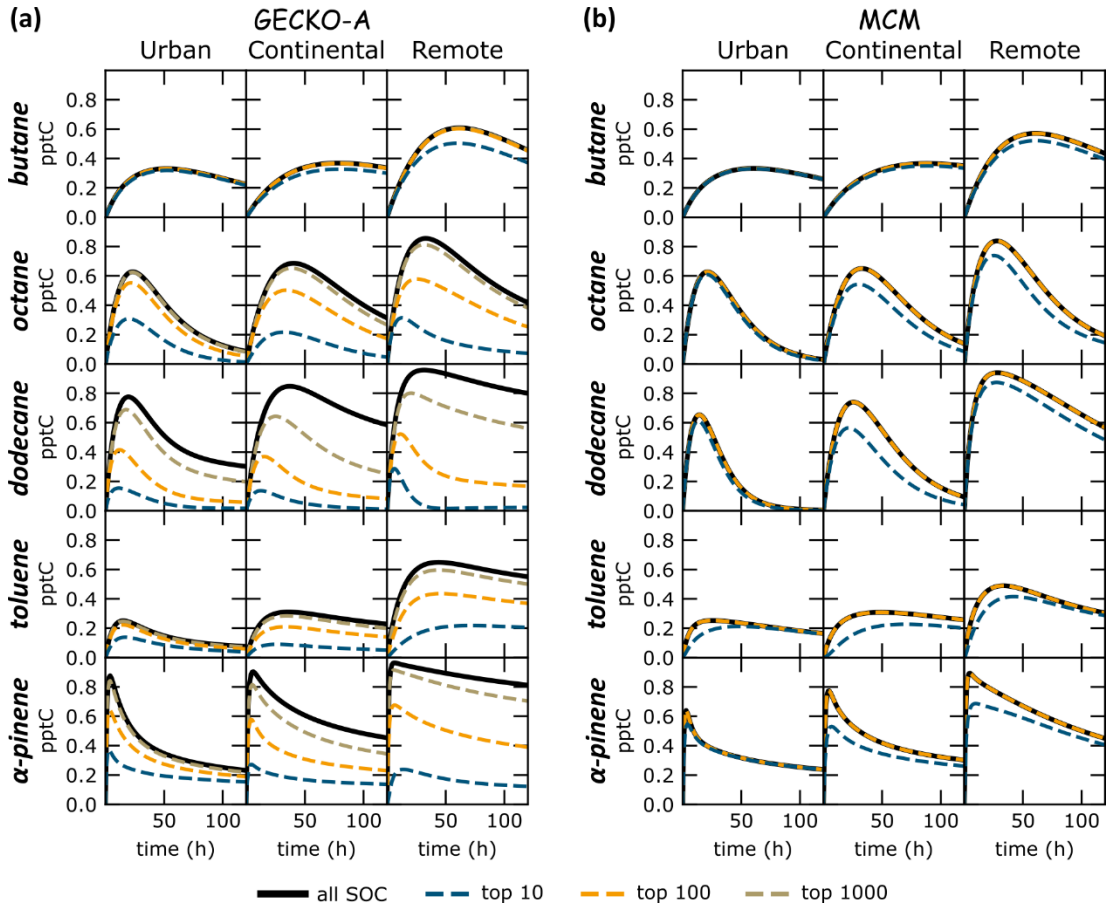


Figure 14: Temporal evolution of cumulative secondary organic carbon (SOC) contributions in GECKO-A (panel a) and MCM (panel b) simulations for various parent hydrocarbons under different environmental scenarios. Columns (left to right): urban, continental, and remote scenarios. Rows (top to bottom): butane, octane, dodecane, toluene, and α-pinene as parent compounds. Black solid lines: total SOC; blue dashed lines: cumulative contribution of the top 10 species by carbon-integrated abundance; orange dotted lines: top 100 species; ~~grey~~gray dotted lines: top 1000 species.

the 10 most abundant species contribute only 12%, 8%, and 6% of the total SOC budget in the urban, ~~remote, and~~ continental, and remote scenarios, respectively. Even when the 1000 most abundant species are included, the budget remains incomplete, with cumulative contributions of 79%, 62%, and 77% for these same scenarios. This highlights the pronounced chemical diversity represented in GECKO-A, where a large number of species, each present at low abundance, collectively contribute to the overall SOC budget.

It is important to note that the composition of the dominant-species pool varies across scenarios. Species diversity contributing to SOC is particularly high in the continental scenario, which combines the formation of compounds typical of NO_x-rich environments (e.g., organic nitrates) and NO_x-poor environments (e.g., hydroperoxides). Consequently, the fraction of the SOC budget explained by a given number of species is generally lower in the continental scenario, requiring the inclusion of a larger number of compounds to achieve comparable coverage.

Mechanisms cannot be reduced solely to species that contribute substantially to the SOC mass budget. Numerous reactive intermediates—although present at low concentrations and contributing negligibly to the mass budget due to their high reactivity (e.g., radicals)—play key roles in oxidation pathways and influence branching ratios under varying environmental conditions. Nevertheless, the number of species that contribute significantly to the SOC mass budget remains approximately three orders of magnitude lower than the total number of species generated by GECKO-A (Fig. 1). Figure ~~S3~~17 shows the cumulative contribution of species formed in successive oxidation generations and reveals that most of the SOC ~~mass~~budget is accounted for by species up to the fourth generation, regardless of the parent hydrocarbon or scenario considered. This contribution, however, varies by parent compound. For butane, the carbon budget is almost entirely carried by first- and second-generation species. For dodecane, species from generations three and four make an increasing contribution to this budget as oxidation proceeds (Fig. 17). These findings suggest significant potential for mechanism simplification in GECKO-A without major loss of chemical information. Systematic and automated reduction approaches, based on tools such as GENOA (Wang et al., 2022, 2023, 2026), are currently under development to reduce the size of GECKO-A-generated mechanisms. Sect. 6 illustrates the reduction capabilities currently available in GECKO-A, but a detailed treatment of GECKO-A mechanism reduction is beyond the scope of the present study.

5.10 Volatility versus water solubility of organic carbon

Simulated differences in the composition of secondary organic carbon directly affect the evolution of its thermodynamic properties during oxidation, as well as its gas–particle partitioning and associated processes such as dry and wet deposition or multiphase oxidation. To examine divergences between the GECKO-A and MCM mechanisms, Fig. 15 presents the distribution of species as a bubble plot in a two-dimensional space defined by saturated vapor pressure p_{sat} and effective Henry's law constant H. These distributions are shown for different parent compounds (butane, octane, dodecane, toluene, and α-pinene) under the continental scenario after 72 h of simulation. Similar results are obtained for the other scenarios (not shown).

Thermodynamic properties (p_{sat} and H) were estimated using structure–property relationships (see Sect. 4.1 and 4.2). Because several positional isomers often share identical values of p_{sat} and H, they were grouped in Fig. 15 to avoid overlapping bubbles. A filter was also applied to exclude species with negligible concentrations.

Phase-dominance boundaries (gas, organic particulate, or aqueous) are displayed in the background of Fig. 15 as a fixed reference framework, assuming thermodynamic equilibrium. These boundaries are calculated for prescribed particulate phases: a purely organic phase (mean molar mass $M_{OA} = 200$ g mol^{-1}) and a purely aqueous phase, each with a mass concentration of 10 µg m^{-3}. The logarithmic axes of p_{sat} and H in Fig. 15 thus define a chemical partitioning space. A similar framework and its underlying assumptions are discussed in detail in Wania et al. (2015). The phase-separation lines are briefly summarized below.

The partitioning equations of a species between the gas and condensed (organic or aqueous) phases are provided in Sect. S5. The vapor pressure corresponding to equal partitioning (50:50) between the gas and organic condensed phase is given by

$$p_{sat}^{(g:o)} = RTC_{OA}/M_{OA}$$

Under the present conditions, $p_{sat}^{(g:o)} = 1.2 \times 10^{-9}$ atm (see Sect. 5.2). Similarly, equal partitioning between the gas and aqueous phases is obtained for (see Sect. S5)

$$H^{(g:w)} = \frac{1}{RT\omega_L}$$

where ω_L is the liquid water mixing ratio (volume of liquid water per volume of air). With $\omega_L = 10^{-11}$ (corresponding to 10 µg m^{-3}), $H^{(g:w)} = 4.1 \times 10^9$ M atm^{-1}. Finally, the relationship between Henry's law constant and vapor pressure for 50:50 partitioning between aqueous and organic condensed phases is of the form

$$H^{(w:o)} = \frac{a}{p_{sat}}$$

with $a = 10^{-9} C_{OA}/\omega_L M_{OA}$ (see Sect. S5). Under the selected conditions, $a = 5.0$ M. In logarithmic space, $\log(H^{(w:o)})$ as a function of $\log(p_{sat})$ is therefore a straight line with slope −1.

Denoting f_o, f_w, and f_g as the fractions of a species in the organic, aqueous, and gas phases, respectively, the boundary lines in chemical partitioning space are given by the following relationships (see Sect. S5):

$$H = H^{(g:w)} \cdot \frac{(1-f_o)}{f_o} \cdot \frac{p_{sat}^{(g:o)}}{p_{sat}} - H^{(g:w)}$$

$$H = H^{(g:w)} \cdot \frac{f_w}{1-f_w} \cdot \left(1 + \frac{p_{sat}^{(g:o)}}{p_{sat}}\right)$$

$$H = \frac{1-f_g}{f_g} \cdot H^{(g:w)} - \frac{p_{sat}^{(g:o)}}{p_{sat}} \cdot H^{(g:w)}$$

For each phase i, boundary lines are shown in Fig. 15 for $f_i = 0.50$, 0.90, and 0.99.

The domains in Fig. 15 are only intended to map species according to their relative affinity for a given phase. They provide only a projection framework, constructed under the assumptions that the organic and aqueous condensed phases are completely immiscible and that absorption into these phases is ideal, following Raoult's and Henry's laws, respectively. The limitations of this representation are discussed in Wania et al. (2015). Partial miscibility and liquid-liquid phase separation, non-ideality, salting out, condensed phase composition, phase state and dynamic mass-transfer effects must naturally be considered in any comprehensive application (Perraud et al., 2012; Pye et al., 2017; Schervish et al., 2024; Shiraiwa et al., 2013; Shiraiwa and Seinfeld, 2012; Shrivastava et al., 2017; Wania et al., 2015; Zuend and Seinfeld, 2012), but this lies beyond the scope of the present study.

As oxidation proceeds, oxygenation of the carbon backbone increases. This decreases volatility and enhances water solubility. Accordingly, Fig. 15 shows that species are distributed mainly along an axis of decreasing solubility as volatility increases, as discussed in Hodzic et al. (2014), Isaacman-VanWertz and Aumont (2021), Lannuque et al. (2018), Wania et al. (2015).

As expected, Fig. 15 indicates that species in the MCM remain too volatile to exhibit significant affinity for the organic phase, consistent with the findings of Sect. 5.1. Their water solubility is also insufficient for substantial uptake into an aqueous particulate phase. The affinity of MCM-produced species clearly remains toward the gas phase, irrespective of the parent compound.

~~In contrast, except for butane, GECKO-A mechanisms lead to the formation of low-volatility species, as discussed previously. However, their properties indicate that they are not only weakly volatile but also highly water soluble. For many of them, affinity for an aqueous phase clearly exceeds that for an organic phase. The box-model configuration described in Sect. 3.1, which considers partitioning only into an organic particulate phase, therefore appears as a first-order approximation requiring refinement. For typical atmospheric applications, partitioning into condensed phases in its various forms should be considered (e.g. Couvidat et al., 2013; Jo et al., 2024; Kim et al., 2019; Pye et al. 2017; Shrivastava et al., 2017; Tilgner et al., 2021).~~

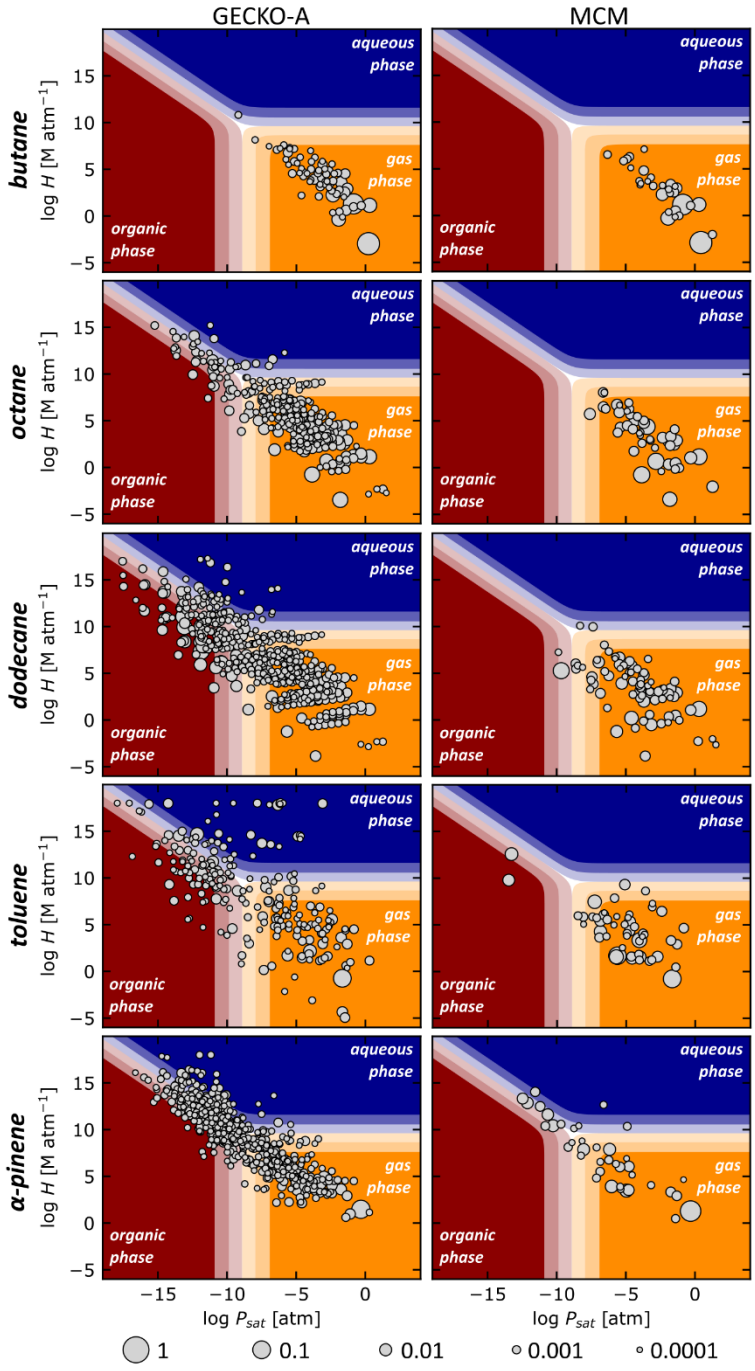


Figure 15: Distribution of species from GECKO-A (left column) and MCM (right column) as a function of their saturation vapor pressure and effective Henry's law constant, for the continental scenario after 72 hours of simulation. From top to bottom: butane, octane, dodecane, toluene, and α-pinene as parent compounds. Colored domains indicate the dominant phase of each species at thermodynamic equilibrium: dark, intermediate, and light shades representing partitioning greater than 99%, 90%, and 50% in the respective phase.

Domain boundaries are calculated for a mass concentration of 10 µg m^{-3} in both organic particulate and aqueous phases. Bubble size represents the relative carbon abundance of each species normalized to the initial parent compound.

In contrast, except for butane, GECKO-A mechanisms lead to the formation of low-volatility species, as discussed previously. However, their properties indicate that they are not only weakly volatile but also highly water-soluble. For many of them, affinity for an aqueous phase clearly exceeds that for an organic phase. The box-model configuration described in Sect. 3.1, which considers partitioning only into an organic particulate phase, therefore appears as a first-order approximation requiring refinement. For typical atmospheric applications, partitioning into various forms of condensed phases should be considered (e.g. Couvidat et al., 2013; Jo et al., 2024; Kim et al., 2019; Pye et al., 2017; Shrivastava et al., 2017; Tilgner et al., 2021).

Liquid water content can vary by several orders of magnitude under atmospheric conditions; for example, it typically ranges from 10^{-8} to 10^{-6} under cloud conditions (e.g. Seinfeld and Pandis, 2016). Increasing liquid water content shifts the aqueous-affinity domain toward lower H values in chemical partitioning space. The same species distributions shown in Fig. 15 are therefore projected onto a partitioning space representative of cloud liquid water content in Fig. S4. Under these conditions, MCM species still show a clear preference for the gas phase with a weak affinity for the aqueous phase for most parent compounds—, with the exception of α-pinene. In contrast, GECKO-A species exhibit a large affinity for the aqueous phase for all parent compounds.

~~**6 Conclusion**~~

6 red-GECKO-A versus GECKO-A

The effect of chemical resolution on the simulated oxidative trajectories is examined by comparing simulation results obtained with GECKO-A and red-GECKO-A (Sect. 4.1), for the five parent compounds under the environmental scenarios described in Sect. 4.3. This comparison follows the same protocol used in Sect. 5 to compare GECKO-A and MCM, with red-GECKO-A simply substituted for MCM. The trajectories obtained with GECKO-A are therefore identical to those already presented in Sect. 5.

The aim is not to systematically explore the effect of each reduction parameter on these simulated trajectories. The aim is rather to illustrate the reduction capabilities offered by GECKO-A and to highlight the magnitude and nature of the discrepancies that such a reduction can introduce relative to the high-resolution mechanism. The organic carbon budget is shown in Fig. 16, along with its breakdown by oxidation generation in Fig. 17. The remaining comparisons (volatility distribution, carbon chain length distribution, distribution by number of functional groups, degree of functionalization, bulk reactivity metrics, isomer group diversity, major isomer groups, cumulative contribution to the carbon budget, and volatility-solubility distribution) are provided in the Supplement.

The carbon budgets simulated with GECKO-A and red-GECKO-A are in excellent agreement for butane, toluene, and α-pinene in all scenarios (Fig. 16). However, substantial discrepancies are highlighted for octane and dodecane. For both compounds, red-GECKO-A underestimates the SOA yield and overestimates the SOG mixing ratio. Under remote conditions, for example, the SOA yield at the end of the simulation decreases from 0.68 to 0.45 for dodecane, and from 0.18 to 0.07 for

octane. These discrepancies increase progressively as NOx levels decrease, becoming noticeable from the continental scenario onward.

These discrepancies are in part the result of carbon loss associated with the truncation of the red-GECKO-A mechanism at 5 generations (Sect. 4.1), rather than only a simple redistribution between the gas and particulate phases. This loss is accounted for by the XCLOSS tracking species (Sect. 2.2). This loss remains small for butane and α-pinene (≤0.002 in all scenarios). It is larger for octane and toluene, reaching 0.058 and 0.030, respectively, under remote conditions. Dodecane shows the largest carbon losses, reaching 0.074 under urban conditions and 0.171 under remote conditions. These losses appear more limited under NOx-rich to moderately NOx-rich regimes than under NOx-poor regimes. The deviation between GECKO-A and red-GECKO-A cannot, however, be exclusively attributed to this carbon loss. The isomerfg parameter also affects the partitioning of carbon among species, and therefore potentially the volatility distribution of species from one generation to the next.

Figure 17 shows that the reduction has little effect on the generation-resolved distribution of secondary organic carbon for butane, toluene, and α-pinene, in all scenarios. Octane and dodecane, in contrast, show more pronounced discrepancies. These discrepancies are limited to specific generations rather than spread evenly across them.

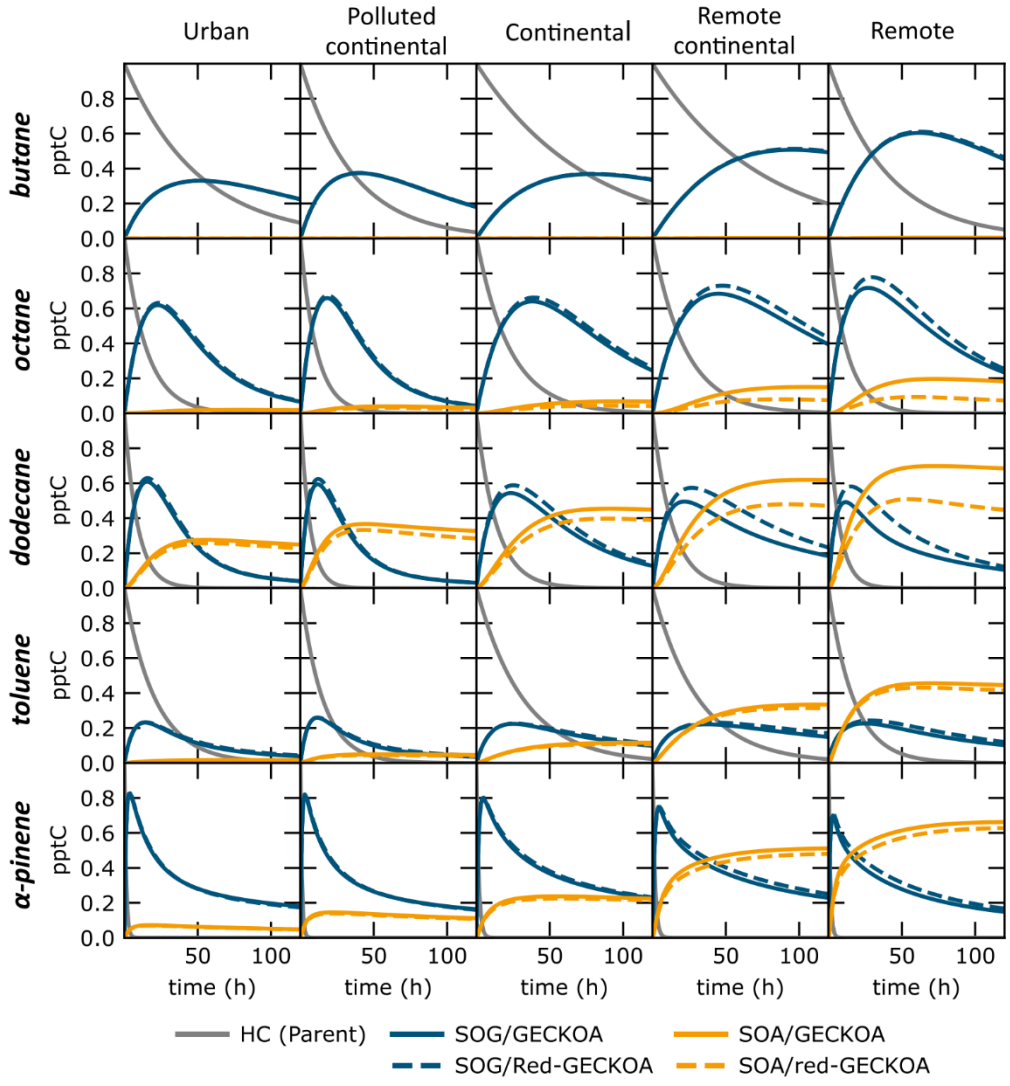


Figure 16: Temporal evolution of the organic carbon budget for various parent hydrocarbons (HC) under different environmental scenarios. Left to right panels: urban, polluted continental, continental, remote continental, and remote conditions. Top to bottom panels: butane, octane, dodecane, toluene, and α-pinene as parent HC. Gray lines: decay of the parent hydrocarbon; blue lines: secondary organic gases (SOG); orange lines: secondary organic aerosols (SOA). Solid lines: GECKO-A simulation results; dashed lines: red-GECKO-A simulation results.

In particular, the contribution of the third generation differs markedly between the two mechanisms for dodecane. This difference increases from NOx-rich to NOx-poor conditions. Under remote conditions, for example, the third generation accounts for 0.42 of the initial carbon with GECKO-A, compared with 0.20 with red-GECKO-A. The gas-particle partitioning of this generation differs substantially between the two mechanisms (not shown). At the end of the simulation, third-generation species in the particulate phase account for 0.38 of the initial carbon under remote conditions with GECKO-A, compared with only 0.17 with red-GECKO-A. In contrast, the gas-phase contribution of the third generation remains similar in the two mechanisms (about 0.03 in both cases). Octane shows a similar pattern under remote conditions, but the effect is weaker. The precise origin of this phase-partitioning discrepancy remains to be explored, but isomerfg appears to be a plausible candidate.

This parameter only applies beyond the second generation, which is exactly where the discrepancy appears. maxgen, by contrast, cannot structurally be responsible before the fifth generation. Other parameters also differ between the two mechanisms (brcut, yldcut, rxloss), but their respective contributions cannot be isolated without dedicated sensitivity tests. This question will need to be revisited once autoxidation is implemented in GECKO-A. This development is expected to substantially affect the sequestration of carbon into the particle phase from the earliest generations (Sect. 7).

Figure S9 shows that the functional group to carbon ratio ($r_{g:c}$) remains nearly identical between GECKO-A and red-GECKO-A for all five compounds, regardless of the scenario considered. The degree of functionalization d_f never differs by more than 0.04 between the two mechanisms. This similarity is expected, since the chemical protocol of red-GECKO-A is strictly identical to that of GECKO-A. The isomer groups dominating the organic carbon budget are also identical between the two mechanisms (Fig. S12). Fewer species are also sufficient to account for the secondary organic carbon budget with red-GECKO-A (Fig. S13). This result is likewise expected, since red-GECKO-A itself generates a substantially smaller number of species (Sect. 2.3). For dodecane under urban conditions, for example, the 10 most contributing species account for 0.36 of the total SOC budget with red-GECKO-A, compared with 0.12 with GECKO-A. Finally, the volatility-solubility distribution (Fig. S14) remains qualitatively similar between the two mechanisms for all five compounds.

Octane and dodecane depart from this agreement for several other compositional descriptors, especially under remote conditions. The ELVOC+LVOC fraction decreases with red-GECKO-A for both compounds (Fig. S6). For dodecane, that fraction decreases from 0.66 to 0.42 at the end of the simulation. For octane, it decreases from 0.17 to 0.07 over the same period. The inorganic carbon fraction also increases with red-GECKO-A (Fig. S7), from 0.20 to 0.26 for dodecane, and from 0.58 to 0.61 for octane. The distribution by number of functional groups (Fig. S8) is also modified. For dodecane, the discrepancy is almost entirely due to trifunctional species, whose contribution decreases from 0.44 to 0.20 of the initial carbon under remote conditions, while the contribution of species with 4 or more functional groups remains unchanged (0.27 in both mechanisms). For octane, trifunctional species and species with four or more functional groups contribute equally to the discrepancy, each decreasing from about 0.11 to 0.05 under the same scenario. I_{OHR}^{max} (Fig. S10a) differs moderately for dodecane under remote conditions, with a relative difference of 16%. This is not the case for octane, whose I_{OHR}^{max} remains close between the two mechanisms. The diversity of isomer groups also decreases for both compounds (Fig. S11). The number of groups decreases from 264 to 231 for octane and from 466 to 387 for dodecane, a decrease mainly involving polyfunctional species.

Butane, toluene, and α-pinene remain in good agreement for all the metrics and descriptors examined. The diversity of isomer groups decreases only slightly for toluene and α-pinene, from 235 to 221 groups and from 675 to 652 groups, respectively (Fig. S11). Y_{Ox} (Fig. S10b) shows no such compound- or scenario-specific pattern. The largest absolute difference in the entire dataset actually concerns toluene under urban conditions, where Y_{Ox} decreases from 1.61 with GECKO-A to 1.55 with red-GECKO-A, a difference of only 0.06.

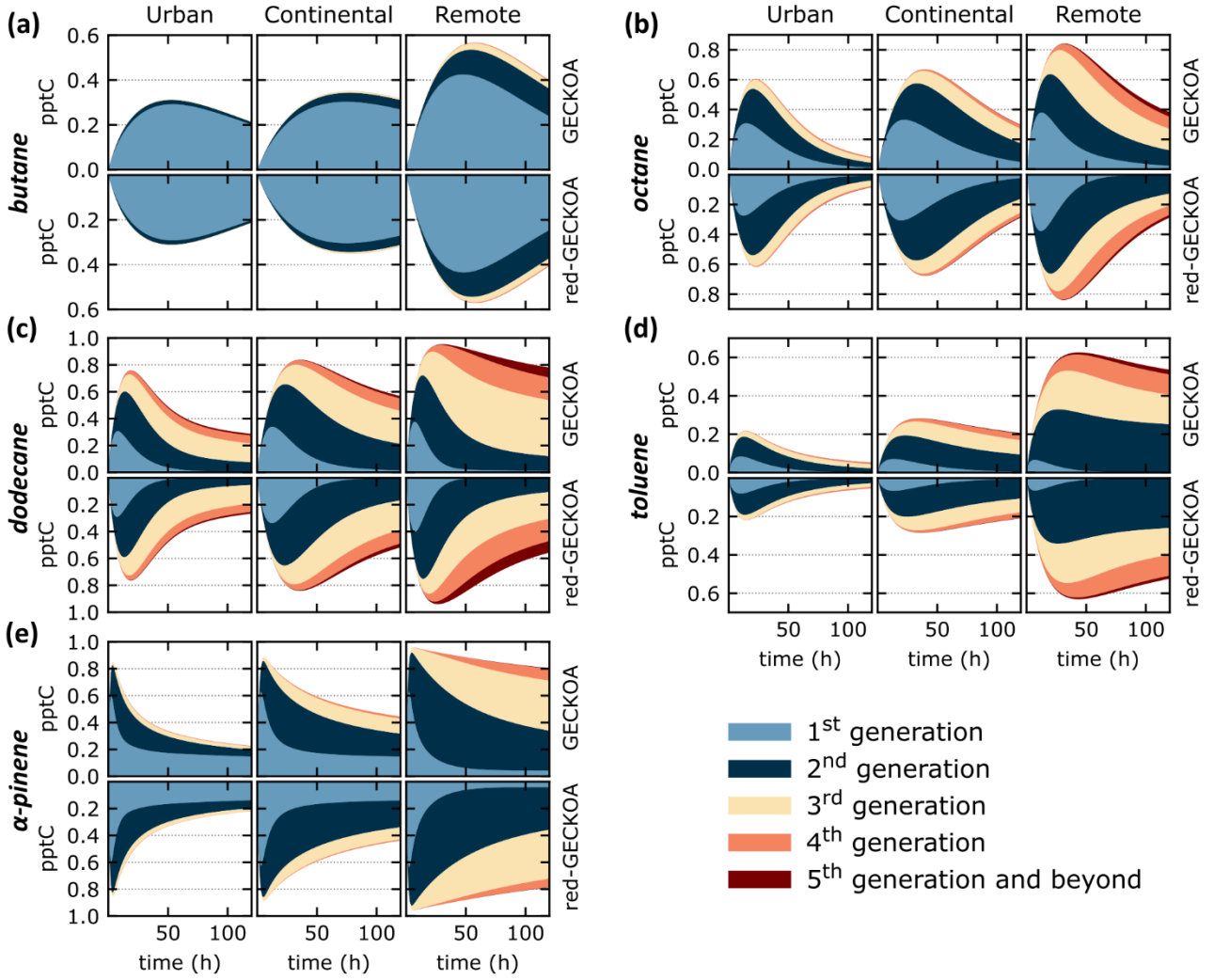


Figure 17: Time evolution of the mixing ratio (pptC) of secondary organic carbon (SOC) grouped by generation for various parent hydrocarbons under different environmental scenarios. Panels (a–e): butane, octane, dodecane, toluene, and α-pinene as parent hydrocarbons, respectively. Top and bottom rows in each panel: GECKO-A and red-GECKO-A simulation results, respectively. From left to right in each panel: urban, continental, and remote scenarios. The color indicates the generation of the species.

These results show that red-GECKO-A faithfully reproduces the trajectories simulated with GECKO-A for three of the five parent compounds, across all the metrics and descriptors examined. For octane and dodecane, discrepancies appear under NOx-poor conditions. These discrepancies mainly affect the partitioning of secondary organic carbon between the gas and particulate phases, and its distribution across oxidation generations. The functionalization chemistry itself, and the identity of the dominant isomer groups, remain largely unaffected by the reduction. The example of red-GECKO-A illustrates how the practical consequences of a reduction strategy vary with the parent compound and the chemical regime considered. GECKO-A v1.0 thus offers a set of adjustable parameters to configure the generation of reduced mechanisms. The configuration used to generate red-GECKO-A remains operational, provided its consequences are evaluated for the compound and conditions of interest, and its parameters are adjusted to the needs of the intended application.

7 Conclusions

This study presents the ~~latest~~ version 1.0 of the GECKO-A model and its companion box model. The GECKO-A code has undergone a complete overhaul, including a modular redesign of its architecture, an updated protocol for mechanism generation, and the open-source release of both the GECKO-A generator and the box model (Aumont et al., 2025a, ~~2025b~~b). In addition, the box model has been parallelized and optimized to run on multiple high-performance computing environments.

To ~~evaluate~~assess the performance and capabilities of this new GECKO-A version, the oxidative trajectories of five parent hydrocarbons (butane, octane, dodecane, toluene, and α-pinene) were examined using their MCM and GECKO-A mechanisms under a range of idealized scenarios spanning environments from urban to remote. Both mechanisms rely on similar protocols to represent the key steps of atmospheric gas-phase oxidation of organic compounds and their dependence on environmental conditions. As a result, the simulated distribution of organic moieties is broadly similar between the two mechanisms, regardless of the environment or the parent hydrocarbon considered.

The simulated oxidative trajectories are comparable for hydrocarbons with small and simple carbon skeletons, such as butane. However, divergences become more pronounced as the size and complexity of the carbon backbone increase. This behavior reflects the inherent challenges associated with the manual development of mechanisms in the MCM. For simple and small carbon structures, a near-explicit development is feasible, and the mechanisms generated automatically by GECKO-A do not differ substantially from those in the MCM. However, the manual construction of explicit mechanisms becomes impractical for structures containing more than five carbon atoms (see Fig. 1), thus requiring the introduction of strategic simplifications and reductions during mechanism development that limit the number of reactions but retain the essential features for which the mechanism is specifically designed (e.g., multi-day ozone formation over NW Europe, in the case of the MCM).

The simplifications and reductions applied in the development of the MCM appear to limit the formation of multifunctional species and to favor early fragmentation of the carbon backbone. Although the trajectories remain similar at the beginning of the oxidation process, when first-generation species dominate, they diverge significantly afterward, when species from successive generations become predominant. The impact on bulk gas-phase metrics (such as O_x production and OH reactivity) remains moderate in NO_x-rich environments, but the discrepancies between GECKO-A and the MCM increase as conditions become more NO_x-limited.

The simulated trajectories also differ markedly in terms of the volatility and aqueous solubility of secondary organic carbon. With the MCM, the yield of secondary organic aerosol (SOA) remains low, regardless of the parent hydrocarbon or environment. In contrast, with GECKO-A, SOA yield generally increases with the size and complexity of the parent compound's carbon backbone and when moving from urban to remote environments. The MCM has primarily been developed to represent VOC gas-phase oxidation for the purpose of simulating gas-phase $HO_x/NO_x/O_x$ chemistry, and therefore the strategic simplifications implemented in the MCM are not specifically designed to capture the gas-phase formation of low-volatility species that play a key role in SOA production under atmospheric conditions.

On the other hand, the explicit generation approach used here to produce mechanisms with GECKO-A leads to a multitude of species whose cumulative yields remain negligible under all atmospheric conditions. Significant mechanism reductions

therefore appear feasible without substantial loss of chemical information. Some approaches are already implemented in the GECKO-A tool and are available for routine use, such as isomer lumping or the elimination of species and reactions falling below a critical threshold. This potential is illustrated by red-GECKO-A (Sect. 6), a reduced configuration that closely reproduces the oxidative trajectories simulated with GECKO-A for most parent compounds and scenarios. However, substantial discrepancies were identified for octane and dodecane under NOx-poor conditions. These discrepancies affect the partitioning of secondary organic carbon between the gas and particulate phases, and its distribution across oxidation generations. Reduction parameters should therefore be evaluated and adjusted to the compound and conditions of interest. A specific tool, GENOA, is also under development to reduce the size of GECKO-A-generated mechanisms to a level enabling their integration into chemistry-transport models (Wang et al., 2026). More broadly, explicit multigenerational mechanisms such as GECKO-A or MechGen (Carter et al., 2025a) are essential to our fundamental understanding of the chemistry driving SOA and water-soluble organic carbon formation. Other, necessarily simplified approaches remain essential for predictive modeling, but explicit mechanisms currently provide a benchmark for this purpose. This benchmark is traceable to current chemical understanding (Ervens et al., 2024).

The versions of MCM and GECKO-A used in this study consider only the progressive multigenerational oxidation of organic compounds. However, it is now established that certain compounds also undergo rapid oxidation on short timescales via successive unimolecular reactions involving H-shift isomerization of RO_2 and ring closure for unsaturated RO_2 (Bianchi et al., 2019; Crounse et al., 2013; Praske et al., 2018). This process, termed autoxidation, is now recognized as playing a key role in the production of highly oxygenated molecules (HOM) of very low volatility and contributing to SOA formation from the first generation of species (Pye et al., 2019; Rissanen et al., 2014). Bimolecular recombination reactions of RO_2 also lead to the formation of low-volatility dimers, which play a key role in the nucleation and growth of organic particles (Berndt et al., 2018; Ehn et al., 2014; Mayhew et al., 2025). A Peroxy Radical Autoxidation Mechanism (PRAM) has been developed on a theoretical and empirical basis to simulate HOM production during the oxidation of several terpenes with OH and O_3, and has been coupled with the MCM (Roldin et al., 2019). The omission of these processes in the GECKO-A tool represents a major limitation for modeling secondary organic aerosol (SOA) formation. A recent systematic comparison between GECKO-A and the SAPRC MechGen generator identified this omission as a major source of the discrepancies found between the two mechanism generation systems, particularly for volatility, solubility, and functional group distributions (Lee-Taylor et al., 2026). Implementing autoxidation could also facilitate the generation of reduced mechanisms. Earlier sequestration of organic carbon into the particle phase indeed limits the need to describe oxidation over several generations to produce low-volatility species, thereby making reduced mechanisms lighter. Autoxidation could thus narrow the SOA discrepancies between reduced and high-resolution configurations of GECKO-A, such as those found between red-GECKO-A and GECKO-A (Sect. 6).

SARs have been proposed to estimate the rate constants for RO_2 isomerization and ring closure reactions (Franzon et al., 2025; Vereecken and Nozière, 2020; Vereecken et al., ~~2020, 2021~~2021a). The H-shift SARs are already exploited in the MechGen mechanism generator (Carter et al., ~~2025~~2025a, b). Developments are underway to integrate these autoxidation processes into GECKO-A in a systematic manner based on SARs, and they will be available in the next version of GECKO-A. Exploratory work on dimer formation from a set of RO_2 generated by GECKO-A has also been undertaken (Franzon et al., 2024), paving the way for future implementation of oligomer formation in GECKO-A.

Code availability

The source codes of GECKO-A v1.0 and its companion box model v1.0 are permanently archived on Zenodo at https://doi.org/10.5281/zenodo.~~15309904~~15309905 (Aumont et al., 2025a) and https://doi.org/10.5281/zenodo.~~15310152~~15310153 (Aumont et al., 2025b), respectively. Both codes are also publicly available on GitLab (https://gitlab.in2p3.fr/ipsl/lisa/geckoa/public~~), where~~). Dedicated wiki pages there provide technical documentation and usage instructions. A frozen snapshot of this documentation, corresponding specifically to GECKO-A v1.0 and its box model, is included in the archive referenced in the Data availability section.

Data availability

The data supporting this study are permanently archived on Zenodo: https://doi.org/10.5281/zenodo.23034925 (Aumont et al., 2026). This archive includes the chemical mechanisms generated with GECKO-A v1.0 for this study, using both the GECKO-A and red-GECKO-A parameter sets (Table 1). It also includes the MCM v3.3.1 mechanism reformatted to comply with the file standards adopted by the GECKO-A tools, and all model outputs in NetCDF format~~ are permanently archived on Zenodo: https://doi.org/10.5281/zenodo.20367277 (Aumont et al., 2026)~~. Finally, it includes a frozen snapshot of the user documentation of GECKO-A v1.0 and its box model.

Author contribution

B.A., R.V., and M.C. designed the study, performed the simulations, and wrote the manuscript. B.A., R.V., M.C., J.L-T and S.M. developed the GECKO-A v1.0 code. All authors contributed to the development of the protocol implemented in GECKO-A for mechanism generation and participated in the review and editing of the manuscript.

Competing interests

The contact author has declared that none of the authors has any competing interests.

Acknowledgements

We gratefully acknowledge Lauri Franzon for helpful discussions related to this work and Camille Mouchel-Vallon for early work on box model parallelization. We also thank the many contributors who have participated in the development of the GECKO-A project over the past two decades.

Financial Support

This work was supported by the MAGNIFY project (Mechanisms for Atmospheric chemistry: GeneratioN, Interpretation and FidelitY), with funding from the French National Research Agency (ANR, grant no. ANR-14-CE01-0010) and the UKRI Natural Environment Research Council (NERC, grant no. NE/M013448/1). Additional funding was provided by the European Commission's EUROCHAMP-2020 project (grant no. 730997) and the UK National Centre for Atmospheric Sciences (NCAS) Air Pollution Theme. This work also received partial support from the ANR under the project "Atmospheric Fate of the First-Generation oxidation products of Biogenic Volatile Organic Compounds" (grant no. ANR-22CE01-0015-01). NSF NCAR is operated by the University Corporation for Atmospheric Research under the sponsorship of the National Science Foundation.

References

Alfarra, R., Camredon, M., Cazaunau, M., Doussin, J.-F., Fuchs, H., Jorga, S., McFiggans, G., Newland, M. J., Pandis, S., Rickard, A. R., and Saathoff, H.: Physical and Chemical Characterization of the Chamber, in: A Practical Guide to Atmospheric Simulation Chambers, edited by: Doussin, J.-F., Fuchs, H., Kiendler-Scharr, A., Seakins, P., and Wenger, J., Springer International Publishing, Cham, 73–111, https://doi.org/10.1007/978-3-031-22277-1_2, 2023.

Atkinson, R. and Arey, J.: Atmospheric Degradation of Volatile Organic Compounds, Chem. Rev., 103, 4605–4638, https://doi.org/10.1021/cr0206420, 2003.

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.

Aumont, B., Camredon, M., Lee-Taylor, J., and Valorso, R.: GECKO-A: Generator for Explicit Chemistry and Kinetics of Organics in the Atmosphere, https://doi.org/10.5281/zenodo.15309905, 2025a.

Aumont, B., Camredon, M., Lee-Taylor, J., and Valorso, R.: Box Model for GECKO-A, https://doi.org/10.5281/zenodo.15310153, 2025b.

Aumont, B., Valorso, R., Rickard, A., Jenkin, M., Lee-Taylor, J., Orlando, J., Madronich, S., Vereecken, L., and Camredon, M.: Data for "GECKO-A v1.0: Exploring VOC Oxidation Trajectories Through Comparison with the MCM",", https://doi.org/10.5281/zenodo.~~20367277~~23034925, 2026.

Berndt, T., Mentler, B., Scholz, W., Fischer, L., Herrmann, H., Kulmala, M., and Hansel, A.: Accretion Product Formation from Ozonolysis and OH Radical Reaction of α-Pinene: Mechanistic Insight and the Influence of Isoprene and Ethylene, Environ. Sci. Technol., 52, 11069–11077, https://doi.org/10.1021/acs.est.8b02210, 2018.

Bertrand, A., Stefenelli, G., Pieber, S. M., Bruns, E. A., Temime-Roussel, B., Slowik, J. G., Wortham, H., Prévôt, A. S. H., El Haddad, I., and Marchand, N.: Influence of the vapor wall loss on the degradation rate constants in chamber experiments of levoglucosan and other biomass burning markers, Atmospheric Chemistry and Physics, 18, 10915–10930, https://doi.org/10.5194/acp-18-10915-2018, 2018.

Bianchi, F., Kurtén, T., Riva, M., Mohr, C., Rissanen, M. P., Roldin, P., Berndt, T., Crounse, J. D., Wennberg, P. O., Mentel, T. F., Wildt, J., Junninen, H., Jokinen, T., Kulmala, M., Worsnop, D. R., Thornton, J. A., Donahue, N., Kjaergaard, H. G., and Ehn, M.: Highly Oxygenated Organic Molecules (HOM) from Gas-Phase Autoxidation Involving Peroxy Radicals: A Key Contributor to Atmospheric Aerosol, Chem. Rev., 119, 3472–3509, https://doi.org/10.1021/acs.chemrev.8b00395, 2019.

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

Bloss, C., Wagner, V., Bonzanini, A., Jenkin, M. E., Wirtz, K., Martin-Reviejo, M., and Pilling, M. J.: Evaluation of detailed aromatic mechanisms (MCMv3 and MCMv3.1) against environmental chamber data, Atmospheric Chemistry and Physics, 5, 623–639, https://doi.org/10.5194/acp-5-623-2005, 2005b.

Calvert, J. G., Orlando, J. J., Stockwell, W. R., and Wallington, T. J.: The Mechanisms of Reactions Influencing Atmospheric Ozone, Oxford University Press, https://doi.org/10.1093/oso/9780190233020.001.0001, 2015.

Camredon, M., Aumont, B., Lee-Taylor, J., and Madronich, S.: The SOA/VOC/NOx system: an explicit model of secondary organic aerosol formation, Atmos. Chem. Phys., 7, 5599–5610, https://doi.org/10.5194/acp-7-5599-2007, 2007.

Carlsson, P. T. M., Vereecken, L., Novelli, A., Bernard, F., Brown, S. S., Brownwood, B., Cho, C., Crowley, J. N., Dewald, P., Edwards, P. M., Friedrich, N., Fry, J. L., Hallquist, M., Hantschke, L., Hohaus, T., Kang, S., Liebmann, J., Mayhew, A. W., Mentel, T., Reimer, D., Rohrer, F., Shenolikar, J., Tillmann, R., Tsiligiannis, E., Wu, R., Wahner, A., Kiendler-Scharr, A., and Fuchs, H.: Comparison of isoprene chemical mechanisms under atmospheric night-time conditions in chamber experiments: evidence of hydroperoxy aldehydes and epoxy products from NO3 oxidation, Atmospheric Chemistry and Physics, 23, 3147–3180, https://doi.org/10.5194/acp-23-3147-2023, 2023.

Carter, W. P. L., Jiang, J., Orlando, J. J., and Barsanti, K. C.: Derivation of atmospheric reaction mechanisms for volatile organic compounds by the SAPRC mechanism generation system (MechGen), Atmospheric Chemistry and Physics, 25, 199–242, https://doi.org/10.5194/acp-25-199-2025, 2025a.

Carter, W. P. L., Jiang, J., Wang, Z., and Barsanti, K. C.: The SAPRC atmospheric chemical mechanism generation system (MechGen), Geoscientific Model Development, 18, 8461–8483, https://doi.org/10.5194/gmd-18-8461-2025, 2025b.

Couvidat, F., Sartelet, K., and Seigneur, C.: Investigating the Impact of Aqueous-Phase Chemistry and Wet Deposition on Organic Aerosol Formation Using a Molecular Surrogate Modeling Approach, Environ. Sci. Technol., 47, 914–922, https://doi.org/10.1021/es3034318, 2013.

Crounse, J. D., Nielsen, L. B., Jørgensen, S., Kjaergaard, H. G., and Wennberg, P. O.: Autoxidation of Organic Compounds in the Atmosphere, J. Phys. Chem. Lett., 4, 3513–3520, https://doi.org/10.1021/jz4019207, 2013.

Ehn, M., Thornton, J. A., Kleist, E., Sipilä, M., Junninen, H., Pullinen, I., Springer, M., Rubach, F., Tillmann, R., Lee, B., Lopez-Hilfiker, F., Andres, S., Acir, I.-H., Rissanen, M., Jokinen, T., Schobesberger, S., Kangasluoma, J., Kontkanen, J., Nieminen, T., Kurtén, T., Nielsen, L. B., Jørgensen, S., Kjaergaard, H. G., Canagaratna, M., Maso, M. D., Berndt, T., Petäjä, T., Wahner, A., Kerminen, V.-M., Kulmala, M., Worsnop, D. R., Wildt, J., and Mentel, T. F.: A large source of low-volatility secondary organic aerosol, Nature, 506, 476–479, https://doi.org/10.1038/nature13032, 2014.

Ervens, B., Rickard, A., Aumont, B., Carter, W. P. L., McGillen, M., Mellouki, A., Orlando, J., Picquet-Varrault, B., Seakins, P., Stockwell, W. R., Vereecken, L., and Wallington, T. J.: Opinion: Challenges and needs of tropospheric chemical mechanism development, Atmospheric Chemistry and Physics, 24, 13317–13339, https://doi.org/10.5194/acp-24-13317-2024, 2024.

Franzon, L., Camredon, M., Valorso, R., Aumont, B., and ~~Kurten~~Kurtén, T.: Ether and ester formation from peroxy radical recombination: a qualitative reaction channel analysis, Atmos. Chem. Phys., 24, 11679–11699, https://doi.org/10.5194/acp-24-11679-2024, 2024.

Franzon, L., Savolainen, A., Iyer, S., Rissanen, M., and Kurtén, T.: Rapid unimolecular reactions of acyl peroxy radicals: extending the structure–activity relationships, Phys. Chem. Chem. Phys., 27, 12198–12210, https://doi.org/10.1039/D5CP01175B, 2025.

Goldstein, A. H. and Galbally, I. E.: Known and Unexplored Organic Constituents in the Earth's Atmosphere, Environ. Sci. Technol., 41, 1514–1521, https://doi.org/10.1021/es072476p, 2007.

Hallquist, M., Wenger, J. C., Baltensperger, U., Rudich, Y., Simpson, D., Claeys, M., Dommen, J., Donahue, N. M., George, C., Goldstein, A. H., Hamilton, J. F., Herrmann, H., Hoffmann, T., Iinuma, Y., Jang, M., Jenkin, M. E., Jimenez, J. L., Kiendler-

Scharr, A., Maenhaut, W., McFiggans, G., Mentel, T. F., Monod, A., Prévôt, A. S. H., Seinfeld, J. H., Surratt, J. D., Szmigielski, R., and Wildt, J.: The formation, properties and impact of secondary organic aerosol: current and emerging issues, Atmospheric Chemistry and Physics, 9, 5155–5236, https://doi.org/10.5194/acp-9-5155-2009, 2009.

Heald, C. L. and Kroll, J. H.: The fuel of atmospheric chemistry: Toward a complete description of reactive organic carbon, Science Advances, 6, eaay8967, https://doi.org/10.1126/sciadv.aay8967, 2020.

Hodzic, A., Aumont, B., Knote, C., Lee-Taylor, J., Madronich, S., and Tyndall, G.: Volatility dependence of Henry's law constants of condensable organics: Application to estimate depositional loss of secondary organic aerosols, Geophys. Res. Lett., 41, 4795–4804, https://doi.org/10.1002/2014GL060649, 2014.

Isaacman-VanWertz, G. and Aumont, B.: Impact of organic molecular structure on the estimation of atmospherically relevant physicochemical parameters, Atmos. Chem. Phys., 21, 6541–6563, https://doi.org/10.5194/acp-21-6541-2021, 2021.

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

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

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.

Jenkin, M. E., Valorso, R., Aumont, B., Rickard, A. R., and Wallington, T. J.: Estimation of rate coefficients and branching ratios for gas-phase reactions of OH with aliphatic organic compounds for use in automated mechanism construction, Atmos. Chem. Phys., 18, 9297–9328, https://doi.org/10.5194/acp-18-9297-2018, 2018a.

Jenkin, M. E., Valorso, R., Aumont, B., Rickard, A. R., and Wallington, T. J.: Estimation of rate coefficients and branching ratios for gas-phase reactions of OH with aromatic organic compounds for use in automated mechanism construction, Atmos. Chem. Phys., 18, 9329–9349, https://doi.org/10.5194/acp-18-9329-2018, 2018b.

Jenkin, M. E., Valorso, R., Aumont, B., and Rickard, A. R.: Estimation of rate coefficients and branching ratios for reactions of organic peroxy radicals for use in automated mechanism construction, Atmos. Chem. Phys., 19, 7691–7717, https://doi.org/10.5194/acp-19-7691-2019, 2019.

Jenkin, M. E., Valorso, R., Aumont, B., Newland, M. J., and Rickard, A. R.: Estimation of rate coefficients for the reactions of O3 with unsaturated organic compounds for use in automated mechanism construction, Atmos. Chem. Phys., 20, 12921–12937, https://doi.org/10.5194/acp-20-12921-2020, 2020.

Jo, Y., Jang, M., Han, S., Madhu, A., Koo, B., Jia, Y., Yu, Z., Kim, S., and Park, J.: CAMx–UNIPAR simulation of secondary organic aerosol mass formed from multiphase reactions of hydrocarbons under the Central Valley urban atmospheres of California, Atmospheric Chemistry and Physics, 24, 487–508, https://doi.org/10.5194/acp-24-487-2024, 2024.

Kerdouci, J., Picquet-Varrault, B., and Doussin, J.-F.: Structure–activity relationship for the gas-phase reactions of NO3 radical with organic compounds: Update and extension to aldehydes, Atmospheric Environment, 84, 363–372, https://doi.org/10.1016/j.atmosenv.2013.11.024, 2014.

Kim, Y., Sartelet, K., and Couvidat, F.: Modeling the effect of non-ideality, dynamic mass transfer and viscosity on SOA formation in a 3-D air quality model, Atmospheric Chemistry and Physics, 19, 1241–1261, https://doi.org/10.5194/acp-19-1241-2019, 2019.

Krechmer, J. E., Pagonis, D., Ziemann, P. J., and Jimenez, J. L.: Quantification of Gas-Wall Partitioning in Teflon Environmental Chambers Using Rapid Bursts of Low-Volatility Oxidized Species Generated in Situ, Environ. Sci. Technol., 50, 5757–5765, https://doi.org/10.1021/acs.est.6b00606, 2016.

La, Y. S., Camredon, M., Ziemann, P. J., Valorso, R., Matsunaga, A., Lannuque, V., Lee-Taylor, J., Hodzic, A., Madronich, S., and Aumont, B.: Impact of chamber wall loss of gaseous organic compounds on secondary organic aerosol formation: explicit modeling of SOA formation from alkane and alkene oxidation, Atmos. Chem. Phys., 16, 1417–1431, https://doi.org/10.5194/acp-16-1417-2016, 2016.

Lannuque, V., Camredon, M., Couvidat, F., Hodzic, A., Valorso, R., Madronich, S., Bessagnet, B., and Aumont, B.: Exploration of the influence of environmental conditions on secondary organic aerosol formation and organic species properties using explicit simulations: development of the VBS-GECKO parameterization, Atmos. Chem. Phys., 18, 13411–13428, https://doi.org/10.5194/acp-18-13411-2018, 2018.

Lee-Taylor, J., Madronich, S., Aumont, B., Baker, A., Camredon, M., Hodzic, A., Tyndall, G. S., Apel, E., and Zaveri, R. A.: Explicit modeling of organic chemistry and secondary organic aerosol partitioning for Mexico City and its outflow plume, Atmos. Chem. Phys., 11, 13219–13241, https://doi.org/10.5194/acp-11-13219-2011, 2011.

Lee-Taylor, J., Hodzic, A., Madronich, S., Aumont, B., Camredon, M., and Valorso, R.: Multiday production of condensing organic aerosol mass in urban and forest outflow, Atmos. Chem. Phys., 15, 595–615, https://doi.org/10.5194/acp-15-595-2015, 2015.

Lee-Taylor, J., Carter, W. P. L., Aumont, B., Orlando, J. J., Wang, Z., Camredon, M., Valorso, R., and Barsanti, K. C.: Comparison of the GECKO-A and SAPRC MechGen Atmospheric Chemical Mechanism Generators, EGUsphere, 1–37, https://doi.org/10.5194/egusphere-2026-4908, 2026.

Lim, Y. B. and Ziemann, P. J.: Kinetics of the heterogeneous conversion of 1,4-hydroxycarbonyls to cyclic hemiacetals and dihydrofurans on organic aerosol particles, Phys. Chem. Chem. Phys., 11, 8029–8039, https://doi.org/10.1039/b904333k, 2009.

Madronich, S. and Flocke, S.: Theoretical Estimation of Biologically Effective UV Radiation at the Earth's Surface, in: Solar Ultraviolet Radiation, 23–48, https://doi.org/10.1007/978-3-662-03375-3_3, 1997.

Matsunaga, A. and Ziemann, P. J.: Gas-Wall Partitioning of Organic Compounds in a Teflon Film Chamber and Potential Effects on Reaction Product and Aerosol Yield Measurements, Aerosol Science and Technology, 44, 881–892, https://doi.org/10.1080/02786826.2010.501044, 2010.

Mayhew, A. W., Franzon, L., Bates, K. H., Kurtén, T., Lopez-Hilfiker, F. D., Mohr, C., Rickard, A. R., Thornton, J. A., and Haskins, J. D.: The global importance of gas-phase peroxy radical accretion reactions for secondary organic aerosol loading, Atmospheric Chemistry and Physics, 25, 17027–17046, https://doi.org/10.5194/acp-25-17027-2025, 2025.

McGillen, M. R., Carter, W. P. L., Mellouki, A., Orlando, J. J., Picquet-Varrault, B., and Wallington, T. J.: Database for the kinetics of the gas-phase atmospheric reactions of organic compounds, Earth System Science Data, 12, 1203–1216, https://doi.org/10.5194/essd-12-1203-2020, 2020.

McVay, R. C., Zhang, X., Aumont, B., Valorso, R., Camredon, M., La, Y. S., Wennberg, P. O., and Seinfeld, J. H.: SOA formation from the photooxidation of α-pinene: systematic exploration of the simulation of chamber data, Atmospheric Chemistry and Physics, 16, 2785–2802, https://doi.org/10.5194/acp-16-2785-2016, 2016.

Monks, P. S., Archibald, A. T., Colette, A., Cooper, O., Coyle, M., Derwent, R., Fowler, D., Granier, C., Law, K. S., Mills, G. E., Stevenson, D. S., Tarasova, O., Thouret, V., von Schneidemesser, E., Sommariva, R., Wild, O., and Williams, M. L.: Tropospheric ozone and its precursors from the urban to the global scale from air quality to short-lived climate forcer, Atmospheric Chemistry and Physics, 15, 8889–8973, https://doi.org/10.5194/acp-15-8889-2015, 2015.

Mouchel-Vallon, C., Bräuer, P., Camredon, M., Valorso, R., Madronich, S., Herrmann, H., and Aumont, B.: Explicit modeling of volatile organic compounds partitioning in the atmospheric aqueous phase, Atmos. Chem. Phys., 13, 1023–1037, https://doi.org/10.5194/acp-13-1023-2013, 2013.

Mouchel-Vallon, C., Deguillaume, L., Monod, A., Perroux, H., Rose, C., Ghigo, G., Long, Y., Leriche, M., Aumont, B., Patryl, L., Armand, P., and Chaumerliac, N.: CLEPS 1.0: A new protocol for cloud aqueous phase oxidation of VOC mechanisms, Geosci. Model Dev., 10, 1339–1362, https://doi.org/10.5194/gmd-10-1339-2017, 2017.

Mouchel-Vallon, C., Lee-Taylor, J., Hodzic, A., Artaxo, P., Aumont, B., Camredon, M., Gurarie, D., Jimenez, J.-L., Lenschow, D. H., Martin, S. T., Nascimento, J., Orlando, J. J., Palm, B. B., Shilling, J. E., Shrivastava, M., and Madronich, S.: Exploration of oxidative chemistry and secondary organic aerosol formation in the Amazon during the wet season: explicit modeling of the Manaus urban plume with GECKO-A, Atmos. Chem. Phys., 20, 5995–6014, https://doi.org/10.5194/acp-20-5995-2020, 2020.

Myrdal, P. B. and Yalkowsky, S. H.: Estimating Pure Component Vapor Pressures of Complex Organic Molecules, Ind. Eng. Chem. Res., 36, 2494–2499, https://doi.org/10.1021/ie950242l, 1997.

Nah, T., McVay, R. C., Zhang, X., Boyd, C. M., Seinfeld, J. H., and Ng, N. L.: Influence of seed aerosol surface area and oxidation rate on vapor wall deposition and SOA mass yields: a case study with α-pinene ozonolysis, Atmospheric Chemistry and Physics, 16, 9361–9379, https://doi.org/10.5194/acp-16-9361-2016, 2016.

Nannoolal, Y., Rarey, J., Ramjugernath, D., and Cordes, W.: Estimation of pure component properties: Part 1. Estimation of the normal boiling point of non-electrolyte organic compounds via group contributions and group interactions, Fluid Phase Equilibria, 226, 45–63, https://doi.org/10.1016/j.fluid.2004.09.001, 2004.

Nannoolal, Y., Rarey, J., and Ramjugernath, D.: Estimation of pure component properties: Part 3. Estimation of the vapor pressure of non-electrolyte organic compounds via group contributions and group interactions, Fluid Phase Equilibria, 269, 117–133, https://doi.org/10.1016/j.fluid.2008.04.020, 2008.

Newland, M. J., Mouchel-Vallon, C., Valorso, R., Aumont, B., Vereecken, L., Jenkin, M. E., and Rickard, A. R.: Estimation of mechanistic parameters in the gas-phase reactions of ozone with alkenes for use in automated mechanism construction, Atmos. Chem. Phys., 22, 6167–6195, https://doi.org/10.5194/acp-22-6167-2022, 2022.

Pankow, J. F.: An absorption model of gas/particle partitioning of organic compounds in the atmosphere, Atmospheric Environment, 28, 185–188, https://doi.org/10.1016/1352-2310(94)90093-0, 1994.

Pankow, J. F. and Asher, W. E.: SIMPOL.1: a simple group contribution method for predicting vapor pressures and enthalpies of vaporization of multifunctional organic compounds, Atmospheric Chemistry and Physics, 8, 2773–2796, https://doi.org/10.5194/acp-8-2773-2008, 2008.

Peng, Z., Lee-Taylor, J., Stark, H., Orlando, J. J., Aumont, B., and Jimenez, J. L.: Evolution of OH reactivity in NO-free volatile organic compound photooxidation investigated by the fully explicit GECKO-A model, Atmos. Chem. Phys., 21, 14649–14669, https://doi.org/10.5194/acp-21-14649-2021, 2021.

Perraud, V., Bruns, E. A., Ezell, M. J., Johnson, S. N., Yu, Y., Alexander, M. L., Zelenyuk, A., Imre, D., Chang, W. L., Dabdub, D., Pankow, J. F., and Finlayson-Pitts, B. J.: Nonequilibrium atmospheric secondary organic aerosol formation and growth, Proceedings of the National Academy of Sciences, 109, 2836–2841, https://doi.org/10.1073/pnas.1119909109, 2012.

Poling, B. E., Prausnitz, J. M., and O'Connell, J. P.: The Properties of Gases and Liquids, McGraw Hill, New York, 768 pp., 2001.

Praske, E., Otkjær, R. V., Crounse, J. D., Hethcox, J. C., Stoltz, B. M., Kjaergaard, H. G., and Wennberg, P. O.: Atmospheric autoxidation is increasingly important in urban and suburban North America, Proceedings of the National Academy of Sciences, 115, 64–69, https://doi.org/10.1073/pnas.1715540115, 2018.

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, Atmospheric Chemistry and Physics, 17, 343–369, https://doi.org/10.5194/acp-17-343-2017, 2017.

Pye, H. O. T., D'Ambro, E. L., Lee, B. H., Schobesberger, S., Takeuchi, M., Zhao, Y., Lopez-Hilfiker, F., Liu, J., Shilling, J. E., Xing, J., Mathur, R., Middlebrook, A. M., Liao, J., Welti, A., Graus, M., Warneke, C., de Gouw, J. A., Holloway, J. S., Ryerson, T. B., Pollack, I. B., and Thornton, J. A.: Anthropogenic enhancements to production of highly oxygenated molecules from autoxidation, Proceedings of the National Academy of Sciences, 116, 6641–6646, https://doi.org/10.1073/pnas.1810774116, 2019.

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), Atmospheric Chemistry and Physics, 23, 5043–5099, https://doi.org/10.5194/acp-23-5043-2023, 2023.

Raventos-Duran, T., Camredon, M., Valorso, R., Mouchel-Vallon, C., and Aumont, B.: Structure-activity relationships to estimate the effective Henry's law constants of organics of atmospheric interest, Atmos. Chem. Phys., 10, 7643–7654, https://doi.org/10.5194/acp-10-7643-2010, 2010.

Rissanen, M. P., Kurtén, T., Sipilä, M., Thornton, J. A., Kangasluoma, J., Sarnela, N., Junninen, H., Jørgensen, S., Schallhart, S., Kajos, M. K., Taipale, R., Springer, M., Mentel, T. F., Ruuskanen, T., Petäjä, T., Worsnop, D. R., Kjaergaard, H. G., and

Ehn, M.: The Formation of Highly Oxidized Multifunctional Products in the Ozonolysis of Cyclohexene, J. Am. Chem. Soc., 136, 15596–15606, https://doi.org/10.1021/ja507146s, 2014.

Roldin, P., Ehn, M., Kurtén, T., Olenius, T., Rissanen, M. P., Sarnela, N., Elm, J., Rantala, P., Hao, L., Hyttinen, N., Heikkinen, L., Worsnop, D. R., Pichelstorfer, L., Xavier, C., Clusius, P., Öström, E., Petäjä, T., Kulmala, M., Vehkamäki, H., Virtanen, A., Riipinen, I., and Boy, M.: The role of highly oxygenated organic molecules in the Boreal aerosol-cloud-climate system, Nat Commun, 10, 4370, https://doi.org/10.1038/s41467-019-12338-8, 2019.

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

Schervish, M., Donahue, N. M., and Shiraiwa, M.: Effects of volatility, viscosity, and non-ideality on particle–particle mixing timescales of secondary organic aerosols, Aerosol Science and Technology, 58, 411–426, https://doi.org/10.1080/02786826.2023.2256827, 2024.

Seinfeld, J. H. and Pandis, S. N.: Atmospheric chemistry and physics: from air pollution to climate change, Third edition., John Wiley & Sons, Hoboken, New Jersey, 2016.

Shiraiwa, M. and Seinfeld, J. H.: Equilibration timescale of atmospheric secondary organic aerosol partitioning, Geophysical Research Letters, 39, https://doi.org/10.1029/2012GL054008, 2012.

Shiraiwa, M., Zuend, A., Bertram, A. K., and Seinfeld, J. H.: Gas–particle partitioning of atmospheric aerosols: interplay of physical state, non-ideal mixing and morphology, Phys. Chem. Chem. Phys., 15, 11441–11453, https://doi.org/10.1039/C3CP51595H, 2013.

Shrivastava, M., Cappa, C. D., Fan, J., Goldstein, A. H., Guenther, A. B., Jimenez, J. L., Kuang, C., Laskin, A., Martin, S. T., Ng, N. L., Petaja, T., Pierce, J. R., Rasch, P. J., Roldin, P., Seinfeld, J. H., Shilling, J., Smith, J. N., Thornton, J. A., Volkamer, R., Wang, J., Worsnop, D. R., Zaveri, R. A., Zelenyuk, A., and Zhang, Q.: Recent advances in understanding secondary organic aerosol: Implications for global climate forcing, Reviews of Geophysics, 55, 509–559, https://doi.org/10.1002/2016RG000540, 2017.

Tilgner, A., Schaefer, T., Alexander, B., Barth, M., Collett Jr., J. L., Fahey, K. M., Nenes, A., Pye, H. O. T., Herrmann, H., and McNeill, V. F.: Acidity and the multiphase chemistry of atmospheric aqueous particles and clouds, Atmospheric Chemistry and Physics, 21, 13483–13536, https://doi.org/10.5194/acp-21-13483-2021, 2021.

Valorso, R., Aumont, B., Camredon, M., Raventos-Duran, T., Mouchel-Vallon, C., Ng, N. L., Seinfeld, J. H., Lee-Taylor, J., and Madronich, S.: Explicit modelling of SOA formation from α-pinene photooxidation: sensitivity to vapour pressure estimation, Atmos. Chem. Phys., 11, 6895–6910, https://doi.org/10.5194/acp-11-6895-2011, 2011.

Vereecken, L~~.: Reaction Mechanisms for the Atmospheric Oxidation of Monocyclic Aromatic Compounds, in: Advances in Atmospheric Chemistry, vol. 2, World Scientific, 377–527, https://doi.org/10.1142/9789813271838_0006, 2019.~~

~~Vereecken, L~~. and Nozière, B.: H migration in peroxy radicals under atmospheric conditions, Atmospheric Chemistry and Physics, 20, 7429–7458, https://doi.org/10.5194/acp-20-7429-2020, 2020.

Vereecken, L. and Peeters, J.: Nontraditional (Per)oxy Ring-Closure Paths in the Atmospheric Oxidation of Isoprene and Monoterpenes, J. Phys. Chem. A, 108, 5197–5204, https://doi.org/10.1021/jp049219g, 2004.

Vereecken, L. and Peeters, J.: Decomposition of substituted alkoxy radicals—part I: a generalized structure–activity relationship for reaction barrier heights, Phys. Chem. Chem. Phys., 11, 9062–9074, https://doi.org/10.1039/B909712K, 2009.

Vereecken, L. and Peeters, J.: A structure–activity relationship for the rate coefficient of H-migration in substituted alkoxy radicals, Phys. Chem. Chem. Phys., 12, 12608–12620, https://doi.org/10.1039/C0CP00387E, 2010.

Vereecken, L. and Peeters, J.: A theoretical study of the OH-initiated gas-phase oxidation mechanism of β-pinene (C10H16): first generation products, Phys. Chem. Chem. Phys., 14, 3802–3815, https://doi.org/10.1039/c2cp23711c, 2012.

Vereecken, L., Aumont, B., Barnes, I., Bozzelli, J. W., Goldman, M. J., Green, W.H., Madronich, S., McGillen, M. R., Mellouki, A., Orlando, J. J., Picquet-Varrault, B., Rickard, A. R., Stockwell, W. R., Wallington, T. J., and Carter, W. ~~p.~~ P.L.: Perspective on Mechanism Development and Structure-Activity Relationships for Gas-Phase Atmospheric Chemistry, International Journal of Chemical Kinetics, 50, 435–469, https://doi.org/10.1002/kin.21172, 2018.

Vereecken, L., Vu, G., Wahner, A., Kiendler-Scharr, A., and Nguyen, H. M. T.: A structure activity relationship for ring closure reactions in unsaturated alkylperoxy radicals, Phys. Chem. Chem. Phys., 23, 16564–16576, https://doi.org/10.1039/D1CP02758A, ~~2021~~2021a.

Vereecken, L., Carlsson, P. T. M., Novelli, A., Bernard, F., Brown, S. S., Cho, C., Crowley, J. N., Fuchs, H., Mellouki, W., Reimer, D., Shenolikar, J., Tillmann, R., Zhou, L., Kiendler-Scharr, A., and Wahner, A.: Theoretical and experimental study of peroxy and alkoxy radicals in the NO3-initiated oxidation of isoprene, Phys. Chem. Chem. Phys., 23, 5496–5515, https://doi.org/10.1039/d0cp06267g, 2021b.

Verwer, J. G.: Gauss–Seidel Iteration for Stiff ODES from Chemical Kinetics, SIAM J. Sci. Comput., 15, 1243–1250, https://doi.org/10.1137/0915076, 1994.

Verwer, J. G., Blom, J. G., van Loon, M., and Spee, E. J.: A comparison of stiff ODE solvers for atmospheric chemistry problems, Atmospheric Environment, 30, 49–58, https://doi.org/10.1016/1352-2310(95)00283-5, 1996.

Wang, Z., Couvidat, F., and Sartelet, K.: GENerator of reduced Organic Aerosol mechanism (GENOA v1.0): an automatic generation tool of semi-explicit mechanisms, Geoscientific Model Development, 15, 8957–8982, https://doi.org/10.5194/gmd-15-8957-2022, 2022.

Wang, Z., Couvidat, F., and Sartelet, K.: Implementation of a parallel reduction algorithm in the GENerator of reduced Organic Aerosol mechanisms (GENOA v2.0): Application to multiple monoterpene aerosol precursors, Journal of Aerosol Science, 174, 106248, https://doi.org/10.1016/j.jaerosci.2023.106248, 2023.

Wang, Z., Carter, W. P. L., Lee-Taylor, J., Orlando, J., Ye, Q., Valorso, R., Camredon, M., Aumont, B., and Barsanti, K.: GENOA v3: A flexible framework for reduction and exploration of highly detailed chemical mechanisms, EGUsphere, 1–69, https://doi.org/10.5194/egusphere-2026-1921, 2026.

Wania, F., Lei, Y. D., Wang, C., Abbatt, J. P. D., and Goss, K.-U.: Using the chemical equilibrium partitioning space to explore factors influencing the phase distribution of compounds involved in secondary organic aerosol formation, Atmospheric Chemistry and Physics, 15, 3395–3412, https://doi.org/10.5194/acp-15-3395-2015, 2015.

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, 2018.

Yeh, G. K. and Ziemann, P. J.: Gas-Wall Partitioning of Oxygenated Organic Compounds: Measurements, Structure–Activity Relationships, and Correlation with Gas Chromatographic Retention Factor, Aerosol Science and Technology, 49, 727–738, https://doi.org/10.1080/02786826.2015.1068427, 2015.

Zuend, A. and Seinfeld, J. H.: Modeling the gas-particle partitioning of secondary organic aerosol: the importance of liquid-liquid phase separation, Atmospheric Chemistry and Physics, 12, 3857–3882, https://doi.org/10.5194/acp-12-3857-2012, 2012.