Author Comment on egusphere-2025-6276

Title: Revisiting the critical role of stabilized Criegee intermediates (sCIs) in sulfuric acid formation: coupling mechanistic updates with interpretable machine learning
Authors: Yuhuan Zhu et al.
Manuscript No.: egusphere-2025-6276
Correspondence to: Qiang Chen (chenqqh@lzu.edu.cn)

Dear Editor and Referees,
We sincerely thank the Editor and the Referees for their careful evaluation of our manuscript and for their constructive comments and suggestions. We especially appreciate the comments concerning the updates to the MCM v3.3.1 mechanism, particularly the selection and justification of reaction rate coefficients and kinetic parameters related to stabilized Criegee intermediates (sCIs), as well as the suggestions for improving the interpretation and presentation of the machine learning results. These comments have been very helpful in guiding us to further refine the mechanism updates, improve the clarity and interpretability of the machine learning results, and strengthen the overall scientific rigor of the manuscript. Specifically, we further revised the sCI-related gas-phase mechanism, expanded the AtChem-MCM simulations, reconstructed the XGBoost-based diagnostic framework, and reinterpreted the relative roles of sCI+SO_2 pathways in H_2SO_4 formation. These revisions have changed several quantitative results compared with the discussion manuscript, but the main qualitative conclusion remains unchanged: the revised mechanism still supports an important, regime-dependent role of sCI chemistry in atmospheric H_2SO_4 formation.
We have carefully considered all comments and revised the manuscript accordingly. Below, we provide a point-by-point response to each comment. For clarity, the Referees' comments are shown in black, our responses are shown in blue, and the corresponding revisions to the manuscript are summarized in red. Because these revisions required substantial new simulations and a near-complete rewriting of several sections, we are finalizing the revised manuscript and will submit it together with a complete "final author reply to the editor" as soon as permitted by the editor. In this public author comment, we provide detailed responses to the reviewers' major concerns and summarize the main changes and new results that will be incorporated into the revised manuscript.
Note: As the revised manuscript is still being finalized, the section, figure, table, and line numbers referred to in this public author comment should be regarded as provisional.

Reviewer #2

General Comment:
This study is relevant to the field of atmospheric chemistry and addresses an important topic: the contribution of the stabilized Criegee intermediates (sCI) reactions with SO_2 to the formation of H_2SO_4 in the atmosphere, which subsequent contributes to the

production of sulfate aerosols. The paper provides new findings coupling updates of the Master Chemical Mechanism (MCM), a benchmark mechanism within the atmospheric chemistry community, and machine learning technique. I recommend publication after addressing the comments below.

The text lacks clarity at times and there is need to provide more detail all throughout the manuscript, particularly about the application of machine learning method. The terms used in the description of the machine learning method should be clearly defined, given that the intended readership may not be familiar with this specialised terminology and the associated methodology. In general, the figure captions should be expanded to more clearly describe the plotted variables and symbols.

At the end of the Introduction the authors should state that they used field observations obtained at the Wuhai City Super Monitoring Station and explain why this site was chosen as representative for their study.

Author Response:

We sincerely thank the reviewer for the positive evaluation of the scientific relevance of our study and for the constructive suggestions regarding the clarity of the manuscript, the description of the machine learning methods, the figure captions, and the justification of the observational site. In the original manuscript, the role of machine learning and the connection between the different modelling steps were not explained with sufficient clarity. In the revised manuscript, we have therefore substantially rewritten the relevant methodological sections and revised the figure captions to make the analysis more transparent to readers who may not be familiar with machine learning terminology.

First, we have substantially revised and expanded Sect. 2.3, now entitled "**Machine learning surrogate and interpretation framework**", to clarify why machine learning was used and how each machine learning analysis serves a specific research objective. In the revised text, we now explicitly distinguish three related but different applications of machine learning in this study. The first XGBoost model, used in Sect. 3.1, was trained on AtChem scenario simulations to emulate the absolute sCI-driven SO2 oxidation rate, $L_{SO2,sCIs}$, and to diagnose how the updated MCM v3.3.1g mechanism changes the sensitivity of this rate to O_3, individual alkenes, relative humidity, and NO_2. The second XGBoost surrogate model, used in Sect. 3.2, was developed to predict μ_{sCIs}, the fractional contribution of the sCI pathway to total gas-phase SO_2 oxidation, under a broad range of atmospheric conditions. This model was further used to examine how the sensitivity of the total gas-phase SO_2 oxidation rate differs between low- and high-μ_{sCIs} regimes. The third machine learning analysis, used in Sect. 3.3, was based on ambient observations and was designed to examine whether the sensitivity relationships inferred from the AtChem-derived surrogate analysis are reflected in observed sulfate variability. This revision makes clear that machine learning was as an efficient surrogate and interpretation tool linked directly to the chemical questions addressed in each results section.

Second, we have added clearer definitions of the machine learning terminology used throughout the manuscript. In particular, we now define "features" as the input variables

used by the model and “target variable” as the output variable to be predicted. We also clarify that the feature variables differ among the three analyses. In Sect. 3.1, the features are the prescribed initial values of O_3, individual alkenes, RH, and NO_2 used as inputs for AtChem, while the target variable is the AtChem-simulated $L_{SO2,sCIs}$. In Sect. 3.2, the features are normalized scenario descriptors, including NO_x, NO_2 fraction, O_3, VOCs, alkene fraction (alkene%), RH, and the Aggregated Reactivity Index(ARI), while the target variables are diagnosed from the constrained AtChem simulations. In Sect. 3.3, the features are observed pollutant, meteorological, photolysis, and source-related variables, while the target variable is the observed sulfate concentration. We also define “hyperparameters” as model settings specified before training, such as learning rate, tree depth, number of trees, subsampling ratio, and regularization strength. The L1 and L2 penalties are now described as regularization terms used to reduce overfitting and improve model generalization.

Third, we have expanded the explanation of the interpretation methods, including SHAP, Sobol sensitivity analysis, and partial dependence plots. In the revised manuscript, SHAP values are defined as the contribution of each feature to an individual model prediction relative to the average model prediction. We now explicitly state that a positive SHAP value means that a feature increases the predicted target variable, whereas a negative SHAP value means that it suppresses the prediction. The mean absolute SHAP value, mean(|SHAP|), is used as a measure of global feature importance. We also clarify that partial dependence plots show the marginal response of the predicted target variable to one or two selected features while averaging over the remaining features. Sobol sensitivity analysis is now described as a variance-based global sensitivity method, in which the first-order index S_1 represents the independent contribution of a feature and the total-order index S_T includes both the independent contribution and all interaction effects.

Fourth, we have revised the captions of the machine-learning-related figures to define the plotted variables, axes, color scales, and symbols more explicitly. For example, in the SHAP summary plots, the revised captions now state that the x-axis represents the SHAP value, which indicates the direction and magnitude of the feature’s influence on the predicted target variable, while the color scale represents the actual feature value from low to high.

Finally, following the reviewer’s suggestion, we have revised the end of the Introduction to state explicitly that field observations from the Wuhai City Atmospheric Environment Super Monitoring Station were used in this study. We have also added a justification for selecting this site in Sect. 2.2.2(Observation data). Wuhai is a semi-arid coal-chemical industrial city in Inner Mongolia, China, characterized by relatively high SO_2 emissions, abundant anthropogenic VOCs including alkene precursors, frequent O_3 pollution events, and comparatively dry atmospheric conditions. These characteristics make Wuhai particularly suitable for examining gas-phase SO_2 oxidation under conditions where both sCI precursors, such as alkenes and O_3, and the sulfur precursor SO2 are present at elevated levels. The semi-arid environment is also relevant because lower humidity can reduce the competitive loss of sCIs to H_2O and water dimers, thereby providing favorable conditions for assessing when sCIs may effectively

compete with OH in SO_2 oxidation.
Overall, these revisions improve the clarity of the manuscript, define the machine-learning terminology more explicitly, better connect each modelling method to the corresponding scientific question, make the figure captions more informative, and provide a clearer rationale for the use of the Wuhai observations.

Specific Comments

Comment 1:

line 15: '…are recognised to be atmospheric intermediates responsible for the oxidation of sulfur dioxide' instead of '…are recognised to be one of the free radicals oxidated sulfur dioxide'. It is well known that Criegee intermediates have a zwitterionic character and, thus referring to them as 'intermediates' rather than 'free radicals' is more appropriate.

Response 1:

We thank the reviewer for this clarification and for correcting our terminology. sCIs are reactive zwitterionic intermediates ($R_1R_2C=O^+-O^-$), rather than traditional free radicals. We have therefore revised the description accordingly throughout the manuscript. We have revised and have been carefully reviewed and corrected similar inaccuracies throughout the manuscript.

Changes in Manuscript:

The sentence has been revised in Abstract.

Comment 2:

line 15: It is not clear what the word 'dominant' means in here. OH radicals are not the most abundant radicals in the atmosphere. The authors should clarify that the intended meaning is that OH is typically the dominant oxidant of SO2.

Response 2:

We thank the reviewer for this helpful comment. Here, we intended to indicate that OH is typically the dominant oxidant of SO_2, rather than the most abundant radical in the atmosphere.

Changes in Manuscript:

The sentence has been revised in Abstract.

Comment 3:

line 40: add the word 'Earth' before 'surface', i.e. 'near the Earth surface'.

Response 3:

We thank the reviewer for this suggestion. In the revised manuscript, the Introduction has been substantially reorganized, and the original sentence containing this wording has been removed. Nevertheless, we have carefully checked the manuscript and corrected similar expressions.

Changes in Manuscript:

The Introduction has been substantially revised, and the original sentence containing this expression has been removed.

Comment 4:

line 50: The reference 'Anon, n.d.' (line 423 in the reference list) should include the source of the pdf document, such as a webpage and the date the webpage was accessed by the authors.

Response 4:

We thank the reviewer for pointing out this reference issue. The incorrect citation "Anon, n.d." has been replaced with the appropriate reference to (Mauldin et al., 2012), and the full reference information has been corrected in the reference list. We have also carefully checked the manuscript for similar citation and reference errors and corrected them.

Changes in Manuscript:

Modified version: Multiple field observation studies conducted in the boreal forests in Finland (Mauldin et al., 2012), the SMEAR II and Hohenpeissenberg stations in Germany(Boy et al., 2013), and urban Beijing (Guo et al., 2021) have consistently supported the substantial role of sCIs in SO_2 oxidation and subsequent sulfuric acid (H_2SO_4) formation.

Comment 5:

line 54: the cited reference, Boy et al., 2013, is not in the list of references given at the end of the manuscript

Response 5:

We thank the reviewer for pointing this out. The missing reference, (Boy et al., 2013), has now been added to the reference list. We have also cross-checked the in-text citations and reference list throughout the manuscript to ensure consistency.

Changes in Manuscript:

Modified version: Multiple field observation studies conducted in the boreal forests in Finland (Mauldin et al., 2012), the SMEAR II and Hohenpeissenberg stations in Germany(Boy et al., 2013), and urban Beijing (Guo et al., 2021) have consistently supported the substantial role of sCIs in SO_2 oxidation and subsequent sulfuric acid (H_2SO_4) formation.

Comment 6:

line 58-59: lack of references for: 'Existing studies reported …environments'.

The authors should comment on a number of other relevant previous studies addressing the contribution of the sCI + SO2 reactions to the production of H2SO4 and sulfate aerosols in the atmosphere, such as: Mauldin Iii et al. Nature 2012; Kim, S et al. Environ. Sci. Technol. 2015; Kukui, A. Atmos. Chem. Phys. 2021; Sarwar, G.et al Atmos. Environ. 2014.

Response 6:

We thank the reviewer for pointing out these important and highly relevant studies. The original manuscript did not provide sufficient references to support the statement that the importance of sCI + SO_2 chemistry varies across atmospheric environments. In the revised Introduction, we have expanded the literature review to include both field-based and modeling studies addressing the contribution of sCIs to H_2SO_4 and sulfate

formation. Specifically, we have added (Mauldin et al., 2012), (Kim et al., 2015), and (Kukui et al., 2021) as field-based studies showing that sCIs can provide an additional, environment-dependent source of H_2SO_4 under VOC-rich or biogenically influenced conditions. (Mauldin et al., 2012) showed that OH oxidation alone could not fully explain the observed sulfuric acid budget in boreal forests, whereas inclusion of sCI chemistry improved the closure of H_2SO_4 formation. (Kim et al., 2015) evaluated sCI-derived H_2SO_4 production downwind of Dallas-Fort Worth and showed that the sCI pathway can provide an additional H_2SO_4 source, although it remains smaller than the OH pathway during midday. (Kukui et al., 2021) further estimated that sCIs contributed approximately 10% of daytime and 40% of nighttime H_2SO_4 formation at a Mediterranean site influenced by biogenic emissions. We have also added (Sarwar et al., 2014), which incorporated explicit SCI chemistry into the CMAQ model and demonstrated that SO_2 oxidation by sCIs can enhance summertime sulfate in biogenically active regions, with strong sensitivity to the assumed SCI+SO_2 and SCI + $H_2O/(H_2O)_2$ kinetics. These added studies further highlight the strong regional and environmental variability in the importance of sCI chemistry, and they reinforce the need to incorporate updated CI chemical kinetics when quantitatively evaluating the contribution of sCIs to SO_2 oxidation.

Changes in Manuscript:

The Introduction has been revised to include the suggested references.

Comment 7:

lines 63-64: The authors should change '…unimolecular decomposition as well' to '…unimolecular decomposition/isomerisation as well'

Response 7:

We thank the reviewer for pointing out this imprecise description of the reaction pathways. In the revised manuscript, the Introduction has been substantially reorganized, and the original sentence containing this expression has been removed. Nevertheless, we have carefully checked the manuscript and corrected similar expressions.

Changes in Manuscript:

The Introduction has been substantially revised, and the original sentence containing this expression has been removed.

Comment 8:

line 67 states that there is an 'intense competition' between H2O/(H2O)2 and SO2 for reaction with sCI'. The authors should clarify which type of Criegee intermediates they are referring to here as the intensity of this competition depends on the sCI structure; for example CH_2OO reacts much more rapidly with water vapour than other sCIs.

Response 8: We thank the reviewer for this important comment. In the revised manuscript, the Introduction has been substantially reorganized, and the original sentence containing this expression has been removed. Nevertheless, we have checked the manuscript carefully and specify the relevant sCI species where appropriate when discussing their reactions with $H_2O/(H_2O)_2$ and SO_2.

Changes in Manuscript:

The Introduction has been substantially revised, and the original sentence containing this expression has been removed.

Comment 9:

lines 70-71: It is not a direct reaction of VOCs with NOx producing O3. Therefore, please change the statement about the O3 formation in the troposphere at day time to one such as '…O3 is primarily formed through chemistry involving VOCs and NOx, where NO_2 produced in these reactions photolyses to generate ozone…'.

Response 9:

We thank the reviewer for catching this imprecise expression. O_3 is produced through the reaction of O_2 with O generated from NO_2 photolysis. This process is continuously driven by the complex photochemical cycles of its precursors, VOCs and NO_x. In the revised manuscript, the Introduction has been substantially reorganized, and the original sentence containing this expression has been removed. Nevertheless, we have carefully checked the manuscript and corrected similar expressions.

Changes in Manuscript:

The Introduction has been substantially revised, and the original sentence containing this expression has been removed.

Comment 10:

line 72: The O_3 photolysis in the presence of water vapour is an important source of OH at day time and should be mentioned: '…through the photolysis of precursors such as HONO and O_3 in the presence of water vapour…'

Response 10:

We thank the reviewer for this helpful suggestion. The photolysis of O_3 in the presence of water vapour is an important daytime source of OH and should be acknowledged when introducing major OH sources. In the revised manuscript, the Introduction has been substantially reorganized, and the original sentence containing this expression has been removed. Nevertheless, we have carefully checked the manuscript and corrected similar expressions.

Changes in Manuscript:

The Introduction has been substantially revised, and the original sentence containing this expression has been removed.

Comment 11:

lines 74-75: The statement 'OH and sCIs are intermediate products generated via different reaction pathways' is not completely true. The authors state just before this that alkene ozonolysis is an important source of OH at night time; the decomposition of sCI formed following alkene ozonolysis can be a significant source of OH.

Response 11:

We thank the reviewer for pointing out this inaccurate statement. OH and sCIs are not strictly generated through independent reaction pathways, because the decomposition of sCIs formed during alkene ozonolysis can also contribute to OH production. In the

revised manuscript, the Introduction has been substantially reorganized, and the original sentence containing this expression has been removed. Nevertheless, we have carefully checked the manuscript and revised related descriptions to better reflect the chemical coupling between OH and sCIs chemistry.

Changes in Manuscript:

The original sentence in the Introduction has been removed during revision. Modified version: Moreover, the production and loss processes of OH and sCIs are not mutually independent but are deeply coupled through the underlying reaction mechanisms (Zhu et al., 2023).

Comment 12:

line 85: Reference for MCM v3.3.1 is missing

Response 12:

We thank the reviewer for pointing this out. The citation to the MCM website has now been added to the manuscript.

Changes in Manuscript:

Modified version: As a prerequisite for all subsequent analyses, we first updated the gas-phase kinetics of Criegee intermediates in the Master Chemical Mechanism (MCM v3.3.1, via website: www.mcm.york.ac.uk) using the latest evaluated rate coefficients recommended by the International Union of Pure and Applied Chemistry (IUPAC), thereby reducing the propagation of mechanistic uncertainties into our conclusions.

Comment 13:

lines 106-107: The statement 'the concentration of bimolecular water is 10^4 times the concentration of H_2O in the atmosphere' is wrong. I think the authors meant the other way around. There is no explanation why [water monomer] was considered 10^4 times larger than [water dimer]. From the equilibrium constant for the dimerisation, K = [dimer]/[monomer]^2, it follows that at T = 25ºC [dimer] = 10^{-4} [monomer] if RH is around 13%. How much was the relative humidity and temperature at the observation site (Wuhai city)? A clear explanation about the choice of the 104 factor is needed, in both the manuscript and the supplement, where this factor is included in the rate coefficients for the sCI + water dimer reactions.

Response 13:

We thank the reviewer for this careful and important comment and for providing the precise thermodynamic context and guidance. First, we acknowledge that the relationship between water monomer and dimer concentrations was incorrectly stated in the original manuscript. In fact, in our study, $[H_2O]$ is considered to equal 1×10^4 $[H_2O]_2$.

The reviewer is absolutely correct that, strictly based on the equilibrium constant at 25 °C, a $[(H_2O)_2]/[H_2O]$ ratio of 10^{-4} corresponds to a relatively dry environment (e.g., RH around 13%). Previous studies have established that in typical Earth's atmospheric conditions, the relative concentration of water dimer to water monomer spans from 10^{-3} to 10^{-4} (Tretyakov et al., 2014). Even at a high relative humidity of 85% at 25 °C, the dimer concentration is roughly three orders of magnitude lower than the monomer

(Ryzhkov and Ariya, 2006). Regarding the meteorological conditions at our observation site: Wuhai city is located in an semi-arid region of Inner Mongolia, China. During our observation period, the average temperature was 28 ± 4 °C and the average relative humidity was relatively low, around 39 ± 14%.

However, regarding the valid question on the actual calculation of water dimer concentrations, treating this ratio as a constant introduces significant uncertainty. The actual dimer concentration scales quadratically with water monomer abundance and is highly temperature-dependent. To accurately represent the kinetics of the sCIs + $(H_2O)_2$ reaction in the box model, we employed a pre-equilibrium approximation. The dimer is assumed to be in steady-state equilibrium with the monomer. We rigorously parameterized the temperature-dependent equilibrium constant (K_{eq}^{D}) using high-accuracy thermochemical data from the Active Thermochemical Tables (ATcT) (Ruscic, 2013), allowing us to calculate an apparent third-order rate constant (k_{eff} = $k_{sCIs+(H2O)2} \cdot K_{eq}^{D}$). The complete mathematical derivation and parameter details have now been explicitly added to Section 2.1 of the revised manuscript.

Changes in Manuscript:

Modified version:

To accurately represent sCIs+$(H_2O)_2$ in MCM v3.3.1 within the box model, a pre-equilibrium approximation was employed. Given the rapid exchange between water monomers and dimers, the concentration of $(H_2O)_2$ is assumed to be in steady-state equilibrium with the water monomer:

$$K_{eq}^{D} = \frac{[(H_2O)_2]}{[H_2O]^2}$$

where K_{eq}^{D} is the temperature-dependent equilibrium constant(Scribano et al., 2006). The reaction of sCIs with water dimers was parameterised using an apparent third-order rate constant(Lade et al., 2024b), k_{eff}, such that the total reaction rate (r) is expressed as:

$$r = k_b[sCI][(H_2O)_2] = k_b \times K_{eq}^{D}[sCI][H_2O]^2 = k_{eff}[sCI][H_2O]^2$$

The apparent rate constant is defined as the product of the bimolecular rate constant for the sCIs+$(H_2O)_2$ reaction (k_b) and the equilibrium constant (K_{eq}^{D}). To ensure the highest thermodynamic accuracy in calculating K_{eq}^{D}, thermochemical data were retrieved from the Active Thermochemical Tables (ATcT). Standard Gibbs free energies of formation ($\Delta_r G_T^{\circ}$) for both the water monomer and the water dimer were extracted from Table 1 and Table 3 of (Ruscic, 2013). These discrete data points were then used to calculate the reaction Gibbs free energy ($\Delta_r G_T^{\circ}$) over the temperature range of 200-360 K, which covers the typical conditions of the troposphere. The temperature dependence of K_{eq}^{D} was then obtained via a linear regression of K_{eq}^{D} against $\frac{1}{T}$, resulting in the following Arrhenius-type parameterisation used in the model:

$$K_{eq}^{D}(T) = Aexp\left(\frac{B}{T}\right)$$

where the coefficients A and B represent the intercept and slope derived from the fitting procedure respectively, with A = 1.15×10^{-23} $cm^3 \cdot molecule^{-1}$ and B = 1549.32 K.

Comment 14:

Tables 1 and 2: Remove the word 'bimolecular' from the title as the tables show pseudo-first rate coefficients for the sCI decomposition/isomerisation too. Suggest adding notes under the tables showing the units of the rate coefficients (Please consult how tables are presented in other papers published in Atmos. Chem. Phys.) The errors associated with the rate coefficient values should be included, as well as the temperature, pressure, and a reference for MCM v3.3.1.

Response 14:

We thank the reviewer for this helpful suggestion. The tables have been comprehensively revised to include: (a) corrected titles; (b) units of rate coefficients in table notes; (c) associated uncertainties for all values; (d) temperature and pressure conditions; and (e) MCM v3.3.1 references.

Changes in Manuscript:

Tables 1 and 2 have been revised.

Comment 15:

line 119: Please add 'see Section 2.2.2 after '(…and PAN)'

Response 15:

We thank the reviewer for this suggestion. The manuscript has been revised accordingly.

Changes in Manuscript:

Modified version: In this mode, hourly observations(see Section 2.2.2) of trace gases (NO_x, CO, SO_2, HONO, NMHCs, and OVOCs), meteorological parameters (temperature, T; relative humidity, RH; and pressure, P), and photolysis frequencies act as time-varying constraints for the simulation.

Comment 16:

lines 122-133 (Observation data): The key observations time series as well as a chart showing the percentage contributions of the alkenes shown in Table 2 to the total alkene concentration during the campaign should be included in the supplement.

Were all the observations used in the present study? I suggest removal of the ones which were not used.

Response 16:

We thank the reviewer for this helpful suggestion. The observational constraints used in the simulations should be presented more clearly. In the revised Supplement, we have added a summary table of the major constrained variables during the target simulation period (1 June to 15 July 2021), including their mean, median, and standard deviation, as well as time-series figures showing the temporal variation of the major constrained species and meteorological parameters. We have further revised Section 2.2.2 (Observation data) to remove observations that were not used in the present study. In the original manuscript, this section partly described the measurement capability of the instruments, which may have caused ambiguity regarding which observations were actually used in the analysis. This has now been clarified.

Changes in Manuscript:

Modified version: "Hourly observations from the Wuhai City Atmospheric

Environment Super Monitoring Station were used for the observation-constrained simulations and observation-based machine learning analysis. The variables used in this study include trace gases ($PM_{2.5}$, CO, SO_2, NO_2, O_3, VOCs), meteorological parameters (WS, T, P, RH), photolysis frequencies, inorganic ions including SO_4^{2-} and NO_3^-, and Fe as an indicator of transition-metal-related aqueous-phase processes.". Variables measured at the station but not used in the present analysis have been removed from the revised description to avoid ambiguity. The Supplement now includes time series of the key constrained variables and a chart showing the relative contributions of the representative alkenes listed in Table 2 to total alkene abundance during the campaign.

Comment 17:

line 135, regarding 'We treated the AtChem inputs as features'. Is the meaning that part of the AtChem outputs represented input variables ('features') in the machine learning method? Please re-write the sentence to clarify. The term 'feature' should be explained here.

Response 17:

We thank the reviewer for pointing out this ambiguity. The original wording could be misread as implying that AtChem outputs were used as input variables in the machine-learning model. This was not our intention. In the revised manuscript, we now define the term "feature" at first use. In machine-learning terminology, a feature refers to an input predictor used by the model. In the AtChem-based machine learning analyses, the features are the prescribed chemical, meteorological, or scenario variables used as AtChem inputs or used to define the simulation scenarios, such as O_3, alkenes, NO_x, RH, temperature, and photolysis rates. The quantities calculated from the AtChem simulations, including $L_{SO2,sCIs}$, L_{SO2}, and μ_{sCIs}, are used as target variables rather than features. This revision clarifies that the machine learning model was trained to map prescribed AtChem input/scenario variables to AtChem-calculated diagnostic outputs, rather than treating AtChem outputs as input features.

Changes in Manuscript:

In the revised manuscript, we have added further clarifications in Section 2.3 (Machine learning surrogate and interpretation framework).

Comment 18:

line 142: Please explain what is meant by 'L1 and L2 penalties'

Response 18:

We thank the reviewer for this helpful suggestion. The terms "L1 and L2 penalties" were not sufficiently explained in the original manuscript. In the revised manuscript, we now define them when they are first mentioned. Specifically, we explain that L1 and L2 penalties are regularization terms added to the XGBoost objective function to reduce overfitting and improve model generalization. In XGBoost, the model prediction is obtained from an ensemble of decision trees, and the terminal leaves of these trees are assigned numerical weights. The L1 penalty penalizes the absolute magnitude of these leaf weights and can shrink small weights toward zero, thereby promoting a simpler and more sparse model. The L2 penalty penalizes the squared magnitude of the leaf

weights and discourages excessively large weights, thereby producing smoother and more stable predictions. Both penalties help prevent the model from fitting noise in the training data.

Changes in Manuscript:

We have incorporated a concise version of this explanation into Section 2.3 (Machine learning surrogate and interpretation framework). Modified version: "The L1 and L2 penalties are regularization terms used to reduce overfitting in XGBoost. The L1 penalty penalizes the absolute magnitude of the tree leaf weights and can promote sparsity, whereas the L2 penalty penalizes the squared magnitude of the leaf weights and discourages overly large weights, leading to smoother and more stable predictions."

Comment 19:

line 143: Please add a reference after Python xgboost library.

Response 19:

We thank the reviewer for this helpful suggestion, which improves the reproducibility of our study. We have updated the manuscript to explicitly cite both the foundational algorithm and the specific Python library used. The revised sentence now cites Chen and Guestrin (2016) and specifies the use of the XGBoost Python package (Version 2.1.3).

Changes in Manuscript:

Modified version: "In this study, the XGBoost model was implemented using the XGBoost Python package (Version 2.1.3; XGBoost Developers; https://github.com/dmlc/xgboost) ."

Comment 20:

line 144: The term 'hyperparameter' should be explained.

Response 20:

We thank the reviewer for this helpful suggestion. In the revised manuscript, we now explain that hyperparameters are model settings specified before training. Specifically, in XGBoost, hyperparameters control the learning process and model complexity. Examples include the learning rate, maximum tree depth, number of trees, subsampling ratio, and regularization strengths. These settings determine how fast the model learns, how complex each tree can be, how many trees are included in the ensemble, and how strongly overfitting is penalized. In this study, these hyperparameters were tuned using cross-validation to improve model performance and generalization.

Changes in Manuscript:

This definition has been added to Sect. 2.3. Modified version: "Hyperparameters refer to model settings specified before training which include the learning rate, maximum tree depth, number of trees, subsampling ratio, and regularization strengths in XGBoost. These hyperparameters control the learning process, model complexity, and the degree of regularization, and were tuned by cross-validation to reduce overfitting and improve generalization."

Comment 21:

line 152: Suggest to replace ‘xi and xc constitute…’ with ‘the sum of features xi and xc constitute…’

Response 21:

We thank the reviewer for this helpful wording suggestion. We have revised the sentence as suggested to make the notation clearer.

Changes in Manuscript:

This definition has been added to Sect. 2.3. Modified version: “Here, x_i denotes the feature or feature subset of interest, and x_c denotes all remaining complementary features. Together, x_i and x_c constitute the full set of input features used by the model.”.

Comment 22:

line 153: Add ‘(equation 1)’ at the end of the sentence, i.e. ‘…model (equation 1)’

Response 22:

We thank the reviewer for the careful reading. We have added "(equation 1)" at the end of the sentence as suggested.

Changes in Manuscript:

Modified version: “Here, x_i denotes the feature or feature subset of interest, and x_c denotes all remaining complementary features. Together, x_i and x_c constitute the full set of input features used by the model.”. The equation reference has also been added in the revised manuscript.

Comment 23:

line 154: What does E represent?

Response 23:

We thank the reviewer for pointing out this omission. In the Partial Dependence Plot (PDP) equation, E represents the expectation operator. Specifically, it denotes the average model prediction over the distribution of the complementary features x_c. In practical terms, the PDP is calculated by fixing the feature of interest x_i at a given value and averaging the model predictions over all samples of the remaining features x_c. We have added this explanation after the equation.

Changes in Manuscript:

This definition has been added to Sect. 2.3. Modified version: “E represents the mathematical expectation operator, which calculates the average model prediction over the marginal distribution of x_c.”

Comment 24:

line 161: The authors should provide examples of the specific fields they are referring to in: ‘…adopted across abroad range of fields’.

Response 24:

We thank the reviewer for this helpful suggestion. In the revised manuscript, we have rewritten this sentence to explicitly mention the fields in which SHAP has been applied.

Changes in Manuscript:

This definition has been added to Sect. 2.3. Modified version: “SHAP has been increasingly applied in environmental and atmospheric sciences, including urban

climate and remote-sensing studies that quantify the effects of urban morphology on land surface temperature (Chen et al., 2024), ecosystem-scale environmental response analysis (Yi and Wu, 2023), and atmospheric studies that separate aerosol effects from meteorological co-variability in cloud-water observations(Zhang et al., 2025).

Comment 25:

lines 175-176 states that 'Figures 1a and 1b displays the global SHAP values for each feature, ranked from top to bottom by their mean |SHAP| values.' However, the high to low axes in those figures are labelled 'feature value' instead of SHAP. The meaning of SHAP in the present study should be explained.

Response 25:

We thank the reviewer for pointing out this lack of clarity. To clarify, the SHAP summary plot (beeswarm plot) incorporates three distinct dimensions of information:

The Y-axis: The y-axis represents the individual features, which are arranged from top to bottom in descending order of their global importance (mean |SHAP|).

The X-axis: This represents the actual SHAP value, which shows the impact of a feature on the model's output for individual data points.

The Color Scale (High to Low): The color bar labeled "feature value" strictly represents the actual numerical magnitude of the input features (e.g., whether the initial concentration of O_3 or alkenes is high (red) or low (blue)).

Changes in Manuscript:

We have thoroughly revised the figure caption to ensure this multi-dimensional information of SHAP summary plot are clearly communicated to the readers.

Comment 26:

lines 200-201: The authors should provide references—such as Onel et al, Phys.Chem.Chem.Phys. 2021 and Lade et al. J. Phys. Chem. A 2024—that discuss the dominant removal of E-CH_3CHOO and CH_2OO by reaction with water vapour, in comparison with their losses via reaction with SO_2 under tropospherically relevant conditions.

Response 26:

We thank the reviewer for these relevant references. We have added (Onel et al., 2021) and (Lade et al., 2024a) to the discussion of the dominant removal pathways of E-CH_3CHOO and CH_2OO.

Changes in Manuscript:

We have added the citations for (Onel et al., 2021) and (Lade et al., 2024a) to Section 3.1.

Comment 27:

Figures 1(a) and 1(b): The meaning of the x axis labels are confusing and should be explained.

Response 27:

We thank the reviewer for pointing out this. The x-axis represents the SHAP value, which quantifies the contribution of each feature to the predicted $L_{SO2,sCIs}$ relative to the

mean model prediction. Positive SHAP values indicate that the feature increases the predicted SO_2 loss rate via sCIs, whereas negative values indicate a decrease.

Changes in Manuscript:

The revised caption now explains that the x-axis shows SHAP values and that the color scale represents feature values.

Comment 28:

Figure 2: The unit of SHAP value is mole cm^{-3} (unit of concentration) while in Figure 1 is mole cm^{-3} s^{-1} (unit of rate). Why is this difference? The authors should clarify why moles were used rather than number of molecules, the latter being more commonly used in atmospheric chemistry. What are D1(n=120) and D2(n=128) in the legend of top left figure? There are no units for 'chemical feature concentrations'. The authors should clarify what is meant by 'chemical feature concentrations' in both the main text and Figure 2 capture. Do these represent the initial alkene concentration inputs in the machine learning model?

Response 28:

We thank the reviewer for pointing out these ambiguities in Fig. 2. First, the SHAP values should have the same physical unit as the target variable of the XGBoost regression model. In this analysis, the target variable is the sCI-mediated SO_2 loss rate, $L_{SO2,sCIs}$; therefore, the SHAP values represent contributions to the predicted reaction rate relative to the model baseline. The original unit label "molecule cm^{-3}" was incorrect and has been corrected. To avoid overloading the figure, the revised axis is labeled "SHAP value (impact on $L_{SO2,sCIs}$)", and the caption states that the SHAP values correspond to the same unit as the target reaction rate, i.e., molecules cm^{-3} s^{-1}.

Second, in the manuscript, the intended unit was molecule (e.g., molecule·cm^{-3}·s^{-1}), but we incorrectly used "mole" as a shorthand for "molecule". We have now systematically checked and corrected all corresponding units throughout the manuscript and figures.

Third, the labels "D1" and "D2" in the original legend referred to the datasets generated using MCM v3.3.1 and MCM v3.3.1g, respectively, and the numbers in parentheses indicated the number of plotted samples. During the revision, this figure was removed and replaced with a revised analysis focused on the RH-dependent responses of key alkenes under MCM v3.3.1g mechanism. Therefore, the ambiguous D1/D2 notation no longer appears. The revised figure caption now explicitly defines the plotted variables, the model target, and the meaning of the SHAP axis.

Changes in Manuscript:

The original Fig. 2 has been replaced. The revised figure and caption now define the SHAP axis, the target variable $L_{SO2,sCIs}$, the meaning of the feature values, and the corrected unit convention.

Comment 29:

-line 235: The authors should specify which 'specific region' they are referring to

Response 29:

We thank the reviewer for this helpful suggestion. The "specific region" refers to the Wuhai City area in Inner Mongolia, China, where the monitoring station is located. We

have replaced the vague phrasing with the specific location name.

Changes in Manuscript:

The vague phrase "specific region" has been replaced by "Wuhai City, Inner Mongolia, China".

Comment 30:

lines 244 - 245 states: that 'the strongest pairs' are 'O3 × alkene%, O3 × VOCs, and VOCs × alkene%'. However, the Top 10 Interaction Effects in Figure 3b shows that the contribution of NOx × NO2% is larger than the contribution of VOCs × alkene%. Why NOx × NO2% is not listed in 'the strongest pairs'?

Response 30

We thank the reviewer for pointing out this inconsistency. The original wording did not accurately reflect the ranking shown in the ANOVA interaction plot. In the revised manuscript, the original ANOVA-based figure has been removed and replaced by Sobol sensitivity analysis, which more appropriately quantifies both first-order effects and interaction-related contributions. The revised daytime analysis shows that ARI, O_3, alkene%, and VOCs dominate the variance in μ_{sCIs}, with first-order Sobol indices(S_1) of approximately 0.18, 0.16, 0.15, and 0.11, respectively. The revised text therefore no longer makes the unsupported statement about the "strongest pairs" in the original ANOVA figure.

Changes in Manuscript:

The original ANOVA figure and the corresponding text describing the "strongest pairs" have been removed. The revised manuscript now reports Sobol first-order and total-order indices and explains the dominant controls on μ_{sCIs} using the updated sensitivity framework.

Comment 31:

Figure 3b: I recommend including error bars in the contribution values showed in the Main Effects Contribution and Top 10 Interaction Effects figures. The authors should describe more clearly the methodology used to generate the pie chart in in both the main text and Figure 3 capture. What is the meaning of the numbers on the right vertical axis of the figure in the bottom right corner? The legend includes only the text '(b) ANOVA effect analysis for all features', which is not sufficiently explanatory.

Response 31

We thank the reviewer for this helpful suggestion. The original Fig. 3b did not provide sufficient methodological information and that some graphical elements, including the contribution values and right-axis labels, were not clearly explained. In the revised manuscript, we have removed the ANOVA-based plot and replaced it with Sobol sensitivity analysis. This avoids the ambiguity associated with the previous pie/bar visualization and provides a clearer variance-based interpretation of main and interaction effects. The revised caption now defines S_1 as the independent contribution of each factor and S_T as the total contribution including interactions.

Changes in Manuscript:

The original ANOVA effect plot has been removed. The revised figure presents Sobol

sensitivity indices (S_1 and S_T), and the caption now explains the method, plotted quantities, and interpretation of first-order and total-order effects.

Comment 32:

line 258, regarding: 'The three factors with the largest main effects …in the order O3 > alkene% > VOCs.' This sentence refers to Figure 4 where the Main Effects Contribution plot shows that the order is O3 > VOCs > alkene%.

Response 32

We thank the reviewer for identifying this mismatch between the text and Fig. 4. The original sentence incorrectly described the order of the main effects. In the revised manuscript, the original nighttime ANOVA figure has been removed and replaced by Sobol sensitivity analysis. The revised nighttime results show that VOCs, alkene%, and O_3 are important controls, with first-order Sobol indices of approximately 0.13, 0.11, and 0.08, respectively. The revised text has therefore been rewritten according to the updated Sobol results rather than the previous ANOVA ranking.

Changes in Manuscript:

The inconsistent sentence and the original ANOVA figure have been removed. The revised nighttime discussion now reports Sobol sensitivity indices.

Comment 33:

line 258: The authors should explain why the NOx × NO2% interaction is not listed in 'the strongest pairs' because the Top 10 Interaction Effects plot shows that its contribution is the largest (see similar comment about Figure 3 above).

Response 33

We thank the reviewer for raising this point. In the original manuscript, VOCs, O_3, and alkene% were discussed separately from NO_x and NO_2%, and the importance of the NO_x×NO_2% interaction was not stated consistently with the ANOVA interaction ranking. We agree that this was an error in presentation. In the revised manuscript, the original ANOVA-based interaction plot has been removed and replaced by Sobol sensitivity analysis. The revised text now discusses interaction effects based on the difference between total-order (S_T) and first-order (S_1) Sobol indices.

Changes in Manuscript:

The original interaction ranking has been removed. The revised analysis now uses Sobol sensitivity indices to quantify independent and interaction-related contributions, and the discussion of NO_x and NO_2% has been rewritten accordingly.

Comment 34:

line 263: 'the promoting potential of O_3, VOCs, and alkene% on μsCIs% was unlocked' is confusing and should be clarified.

Response 34

We thank the reviewer for pointing out this unclear phrasing. The original sentence aimed to describe that under high NO_x/NO_2 conditions, increases in O_3, VOCs, and the fraction of alkenes lead to a larger increase in μ_{sCIs} (the relative contribution of sCIs to SO_2 oxidation). Mechanistically, at night, elevated NO_2 promotes OH termination,

reducing the OH-mediated SO_2 oxidation pathway. As a result, μ_{sCIs} becomes more sensitive to the availability of O_3, VOCs, and alkenes. We have revised the manuscript text.

Changes in Manuscript:

The "Summary and conclusion" has been substantially revised, and the original sentence containing this expression has been removed.

Comment 35:

Figure 4 capture: The word 'assessing' should be replaced by 'assessment of'. Regarding both Figures 3 and 4 captures: Please include how LSO2(g) was generated and what machine learning method was used to generate the plots in (a).

Response 35

We thank the reviewer for this suggestion. In the revised manuscript, the original Figs. 3 and 4 have been replaced because we expanded the simulation dataset and replaced the ANOVA-based interpretation with Sobol sensitivity analysis. The revised captions now explicitly state that μ_{sCIs} and L_{SO2} were diagnosed from AtChem simulations coupled with the updated MCM v3.3.1g mechanism. The XGBoost surrogate model was trained on these AtChem-derived outputs, and the PDPs show the predicted marginal responses from the trained model across the prescribed feature space. The revised figure captions also define each input feature, the target variable, and the meaning of S_1 and S_T in the Sobol panels.

Changes in Manuscript:

The revised captions now describe how the AtChem outputs were generated, which XGBoost surrogate model was used, and how the PDP and Sobol sensitivity panels should be interpreted.

Comment 36:

lines 273-284: The authors should clarify what is meant by low – high NO2%. The entire paragraph is somewhat confusing and should be reorganised for better clarity.

Response 36

We thank the reviewer for pointing out this lack of clarity. We thank the reviewer for pointing out the ambiguity in this paragraph. In the revised manuscript, we have clarified that "low NO_2%" and "high NO_2%" refer to the relative fraction of NO_2 within total NO_x, which modulates the partitioning between NO and NO_2. Specifically, low NO_2% indicates that most NO_x is present as NO, favoring radical propagation and OH regeneration, whereas high NO_2% indicates a higher proportion of NO_2, promoting OH termination. We have reorganized the paragraph to improve clarity and readability, explaining how NO_2% and NO_x alter the relative contribution of sCIs to SO_2 oxidation (μ_{sCIs}).

Changes in Manuscript:

The paragraph has been reorganized.

Comment 37:

Figure 5: There are no explanations for the numbers shown in any of the schematics (a-

f), making their meaning unclear. Please clarify in both the main text and the figure capture.

Response 37

We thank the reviewer for pointing out this lack of clarity. The numbers in the original Fig. 5 represent the reaction rate coefficients (in molecules·cm^{-3}·s^{-1}) for the key reaction pathways shown in the schematics. To avoid confusion, the revised manuscript now explicitly defines these numbers in both the main text and the figure caption. Where the figure was replaced during the structural revision, the new figure captions now define all plotted quantities and units.

Changes in Manuscript:

The relevant figure caption and main-text description have been revised to define the numbers and their units.

Comment 38:

Figure 6: What are the differences between SA-sCI and SA-sCIg and between SA-OH and SA-OHg? The main text should clearly state what SA-sCI and SA-OH represent and the figure capture should explain the meaning of all 4 notations (SA-sCI, SA-sCIg, SA-OH and SA-OHg) and which version of MCM corresponds to each of the plot line (black and purple).

Response 38

We thank the reviewer for pointing out this lack of clarity. In the original manuscript, SA-sCI and SA-OH referred to sulfuric acid production rates from the sCI- and OH-mediated oxidation pathways calculated with MCM v3.3.1, whereas SA-sCIg and SA-OHg referred to the corresponding rates calculated with the updated MCM v3.3.1g mechanism. We agree that this notation was confusing and that the comparison between the original and updated mechanisms in this later section distracted from the main purpose of the observational evaluation. In the revised manuscript, this figure has been removed. The revised observational section now focuses on (i) the comparison between simulated H_2SO_4-related diagnostics and observed SO_4^{2-}, and (ii) the comparison between AtChem-simulated μ_{sCIs} and the μ_{sCIs} predicted by the XGBoost surrogate model. Therefore, the unclear SA-sCI/SA-sCIg notation no longer appears in the revised figure.

Changes in Manuscript:

The original Fig. 6 and the SA-sCI/SA-sCIg notation have been removed. The revised figure now focuses on the observational evaluation of the updated mechanism and surrogate model, with all plotted variables defined explicitly in the caption.

Comment 39:

line 342: Please state the meaning of WS.

Response 39

We thank the reviewer for this suggestion. WS stands for wind speed (m/s). We clarified this in Section 2.2.2 (Observation data).

Comment 40:

line 348: The authors should clarify the rationale for including Fe in the features contributing to the sulfate formation.

Response 40

We thank the reviewer for this helpful suggestion. Fe was included as a feature because transition metals, especially Fe(III) and Mn(II), have been widely recognized as important indicators of aqueous-phase SO_2 oxidation processes in cloud/fog water and aerosol liquid water. In our XGBoost model, Fe served as an indicator to evaluate whether transition-metal-catalyzed processes were associated with sulfate formation during the study period. The SHAP results showed that Fe was associated with a negative effect on SO_4^{2-} during the 2021.06.01-2021.07.15 episode, which does not support a dominant role of aqueous-phase oxidation under these conditions. This has now been clarified in the 3.3 section.

Changes in Manuscript:

The rationale for including Fe has been added to Sect. 2.2.2 and Sect. 3.3, and Fe is now described as an indicator of transition-metal-related aqueous-phase oxidation.

Comment 41:

Figure 7(a): See comments on Figure 1a and b above.

Response 41

We have applied the same revisions as described for Figures 1a and 1b: improved readability, proper axis labels, and expanded figure captions.

Changes in Manuscript:

The figure has been revised with clearer axis labels, larger text, and an expanded caption.

Comment 42:

Figure 7(b): Explain 'high-sCIs and low-sCIs datasets'.

Response 42

We thank the reviewer for pointing out that the definitions of the high-sCIs and low-sCIs datasets were not sufficiently clear. In the revised manuscript, we have clarified that these datasets were defined according to the fractional contribution of the sCI pathway, μ_{sCIs}, simulated by AtChem coupled with the updated MCM v3.3.1g mechanism. Specifically, in Sect. 3.2, we calculated μ_{sCIs} for all designed simulation scenarios and used the median values of the scenario-based μ_{sCIs} distribution as the classification thresholds. The median threshold was 1.6% for daytime and 10% for nighttime. These thresholds were then applied consistently in Sect. 3.3 to classify the observation-constrained simulation period from 1 June to 15 July 2021. For this period, μ_{sCIs} was calculated at each time step from the AtChem-MCM v3.3.1g reaction rates, and daily mean values were calculated separately for daytime and nighttime. Daytime days with mean μ_{sCIs} above 1.6% were classified as the high-μ_{sCIs} dataset, while those below this threshold were classified as the low-μ_{sCIs} dataset. Similarly, nighttime periods with mean μ_{sCIs} above 10 % were classified as high-μ_{sCIs}, while those below this threshold were classified as low-μ_{sCIs}.

Changes in Manuscript:

We have revised the text and the caption these figures to define the high-μ_{sCIs} and low-

μ_{sCIs} datasets explicitly.

Comment 43:
line 354: A reference for the methodology used to generate comparative 'beeswarm' plots should be given.
Response 43
We thank the reviewer for this helpful suggestion. The comparative beeswarm plot used in this study is not a separate interpretation method; rather, it is a customized grouped visualization of standard SHAP values. To avoid confusion, we have revised the text and figure caption to explain how the comparative beeswarm plots were generated. Specifically, after calculating SHAP values for all samples, we divided the data into high- and low-μ_{sCIs} subsets. For each selected feature, the SHAP values from the two subsets were plotted on the same x-axis, with the high-μ_{sCIs} subset shown in the upper half of each feature row and the low-μ_{sCIs} subset shown in the lower half. The color scale represents the corresponding feature value from low to high, following the convention of standard SHAP beeswarm plots. Therefore, the plot is used only to visually compare the distribution and magnitude of feature contributions between the two regimes.
Changes in Manuscript:
Modified caption text: "The comparative beeswarm plots are grouped SHAP beeswarm visualizations. For each feature, points in the upper half of the row represent the high-μ_{sCIs} subset, whereas points in the lower half represent the low-μ_{sCIs} subset. The x-axis shows the SHAP value, indicating the contribution of the feature to the predicted SO_4^{2-}, and the color indicates the corresponding feature value from low to high."

Comment 44:
line 359: Replace the word 'indispensable' with 'significant'.
Response 44
We thank the reviewer for this helpful suggestion. We have carefully checked the manuscript and corrected similar expressions.
Changes in Manuscript:
The word "indispensable" has been replaced with "significant" and similar overstatements have been checked throughout the manuscript.

Comment 45:
line 378: Explain the word 'paradoxically'.
Response 45
We thank the reviewer for pointing this out. The word "paradoxically" was not sufficiently precise. What we intended to convey is that elevated NO_x decreased the overall SO_2 oxidation rate (H_2SO_4 formation) by reducing OH, whereas the sCIs pathway is much less affected. As a result, the fractional contribution of sCIs increases in relative terms, even though this does not imply an absolute enhancement of the sCIs pathway itself. We have revised the relevant text to make this point clearer.
Changes in Manuscript:

The “Summary and conclusion” has been substantially revised, and the original sentence containing this expression has been removed.

Comment 46:

line 383: Replace the word ‘indispensable’ with ‘important’.

Response 46

We thank the reviewer for this helpful suggestion. We have carefully checked the manuscript and corrected similar expressions.

Changes in Manuscript:

The “Summary and conclusion” has been substantially revised, and the original sentence containing this expression has been removed.

Comment 47:

lines 386-387 state: ‘While our box model simulations …they are limited by the exclusion of meteorological factors..’. Please clarify what meteorological factors were excluded as line 118 states: ‘the model was constrained by the observed meteorological parameters (T, RH and p)’.

Response 47

We thank the reviewer for identifying this ambiguity. Our original wording lacked precision. The AtChem box model was indeed constrained by observed meteorological parameters, including temperature, relative humidity, and pressure, as stated in Sect. 2.2.1. These variables were used to calculate water vapor concentrations and to account for their effects on gas-phase chemical reaction rates and photochemical conditions. What we intended to emphasize in the original sentence was not that meteorological variables were excluded, but that a zero-dimensional box model does not explicitly represent meteorology-driven physical and dynamical processes, such as horizontal advection, vertical mixing, boundary-layer evolution, dilution, and deposition. We have clarified this and revised the relevant text in the manuscript accordingly.

Changes in Manuscript:

Modified version: While our box-model simulations provide detailed mechanistic insight into gas-phase SO_2 oxidation, they remain limited by the zero-dimensional framework. Although the simulations were constrained by observed temperature, relative humidity, pressure, and photolysis frequencies, they do not explicitly represent meteorology-driven physical processes, such as horizontal transport, vertical mixing, boundary-layer evolution, turbulent dilution, and deposition. In addition, multiphase chemistry was not included. Future studies using three-dimensional chemical transport models are therefore needed to quantify the role of sCIs in complex polluted atmospheres more comprehensively.

References

Boy, M., Mogensen, D., Smolander, S., Zhou, L., Nieminen, T., Paasonen, P., Plass-Dülmer, C., Sipilä, M., Petäjä, T., Mauldin, L., Berresheim, H., and Kulmala, M.: Oxidation of SO2 by stabilized Criegee intermediate (sCI) radicals as a crucial source for atmospheric sulfuric acid concentrations, Atmos. Chem. Phys., 13, 3865–3879,

https://doi.org/10.5194/acp-13-3865-2013, 2013.

Chen, J., Wang, K., Du, P., Zang, Y., Zhang, P., Xia, J., Chen, C., and Yu, Z.: Quantifying the main and interactive effects of the dominant factors on the diurnal cycles of land surface temperature in typical urban functional zones, Sustainable Cities and Society, 114, 105727, https://doi.org/10.1016/j.scs.2024.105727, 2024.

Cox, R. A., Ammann, M., Crowley, J. N., Herrmann, H., Jenkin, M. E., McNeill, V. F., Mellouki, A., Troe, J., and Wallington, T. J.: Evaluated kinetic and photochemical data for atmospheric chemistry: Volume VII – Criegee intermediates, Atmos. Chem. Phys., 20, 13497–13519, https://doi.org/10.5194/acp-20-13497-2020, 2020.

Guo, Y., Yan, C., Li, C., Ma, W., Feng, Z., Zhou, Y., Lin, Z., Dada, L., Stolzenburg, D., Yin, R., Kontkanen, J., Daellenbach, K. R., Kangasluoma, J., Yao, L., Chu, B., Wang, Y., Cai, R., Bianchi, F., Liu, Y., and Kulmala, M.: Formation of nighttime sulfuric acid from the ozonolysis of alkenes in Beijing, Atmos. Chem. Phys., 21, 5499–5511, https://doi.org/10.5194/acp-21-5499-2021, 2021.

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

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

Kim, S., Guenther, A., Lefer, B., Flynn, J., Griffin, R., Rutter, A. P., Gong, L., and Cevik, B. K.: Potential Role of Stabilized Criegee Radicals in Sulfuric Acid Production in a High Biogenic VOC Environment, Environ. Sci. Technol., 49, 3383–3391, https://doi.org/10.1021/es505793t, 2015.

Kukui, A., Chartier, M., Wang, J., Chen, H., Dusanter, S., Sauvage, S., Michoud, V., Locoge, N., Gros, V., Bourrianne, T., Sellegri, K., and Pichon, J.-M.: Role of Criegee intermediates in the formation of sulfuric acid at a Mediterranean (Cape Corsica) site under influence of biogenic emissions, Atmos. Chem. Phys., 21, 13333–13351, https://doi.org/10.5194/acp-21-13333-2021, 2021.

Lade, R. E., Onel, L., Blitz, M. A., Seakins, P. W., and Stone, D.: Kinetics of the Gas-Phase Reactions of *syn* - and *anti* -CH_3 CHOO Criegee Intermediate Conformers with SO_2 as a Function of Temperature and Pressure, J. Phys. Chem. A, 128, 2815–2824, https://doi.org/10.1021/acs.jpca.4c00199, 2024a.

Lade, R. E., Blitz, M. A., Rowlinson, M., Evans, M. J., Seakins, P. W., and Stone, D.: Kinetics of the reactions of the Criegee intermediate CH_2 OO with water vapour: experimental measurements as a function of temperature and global atmospheric modelling, Environ. Sci.: Atmos., 4, 1294–1308, https://doi.org/10.1039/D4EA00097H,

2024b.

Mauldin, R. L., Berndt, T., Sipilä, M., Paasonen, P., Petäjä, T., Kim, S., Kurtén, T., Stratmann, F., Kerminen, V.-M., and Kulmala, M.: A new atmospherically relevant oxidant of sulphur dioxide, Nature, 488, 193–196, https://doi.org/10.1038/nature11278, 2012.

Onel, L., Lade, R., Mortiboy, J., Blitz, M. A., Seakins, P. W., Heard, D. E., and Stone, D.: Kinetics of the gas phase reaction of the Criegee intermediate CH_2 OO with SO_2 as a function of temperature, Phys. Chem. Chem. Phys., 23, 19415–19423, https://doi.org/10.1039/D1CP02932K, 2021.

Ruscic, B.: Active Thermochemical Tables: Water and Water Dimer, J. Phys. Chem. A, 117, 11940–11953, https://doi.org/10.1021/jp403197t, 2013.

Ryzhkov, A. B. and Ariya, P. A.: The importance of water clusters (H2O) (n= 2, … , 4) in the reaction of Criegee intermediate with water in the atmosphere, Chemical Physics Letters, 419, 479–485, https://doi.org/10.1016/j.cplett.2005.12.016, 2006.

Sarwar, G., Simon, H., Fahey, K., Mathur, R., Goliff, W. S., and Stockwell, W. R.: Impact of sulfur dioxide oxidation by Stabilized Criegee Intermediate on sulfate, Atmospheric Environment, 85, 204–214, https://doi.org/10.1016/j.atmosenv.2013.12.013, 2014.

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

Scribano, Y., Goldman, N., Saykally, R. J., and Leforestier, C.: Water Dimers in the Atmosphere III: Equilibrium Constant from a Flexible Potential, J. Phys. Chem. A, 110, 5411–5419, https://doi.org/10.1021/jp056759k, 2006.

Tretyakov, M. Y., Koshelev, M. A., Serov, E. A., Parshin, V. V., Odintsova, T. A., and Bubnov, G. M.: Water dimer and the atmospheric continuum, 2014.

WHO: WHO Global Air Quality Guidelines: Particulate Matter (PM2. 5 and PM10), Ozone, Nitrogen Dioxide, Sulfur Dioxide and Carbon Monoxide, 1st ed., World Health Organization, Geneva, 1 pp., 2021.

Yi, Z. and Wu, L.: Identification of factors influencing net primary productivity of terrestrial ecosystems based on interpretable machine learning --evidence from the county-level administrative districts in China, Journal of Environmental Management, 326, 116798, https://doi.org/10.1016/j.jenvman.2022.116798, 2023.

Zhang, H., Zhang, Y., Li, Z., and Zheng, Y.: Aerosol Influences on Cloud Water: Insights From ARM EPCAPE Observations With Explainable Machine Learning,

Geophysical Research Letters, 52, e2025GL115163, https://doi.org/10.1029/2025GL115163, 2025.

Zhu, T., Tang, M., Gao, M., Bi, X., Cao, J., Che, H., Chen, J., Ding, A., Fu, P., Gao, J., Gao, Y., Ge, M., Ge, X., Han, Z., He, H., Huang, R.-J., Huang, X., Liao, H., Liu, C., Liu, H., Liu, J., Liu, S. C., Lu, K., Ma, Q., Nie, W., Shao, M., Song, Y., Sun, Y., Tang, X., Wang, T., Wang, T., Wang, W., Wang, X., Wang, Z., Yin, Y., Zhang, Q., Zhang, W., Zhang, Y., Zhang, Y., Zhao, Y., Zheng, M., Zhu, B., and Zhu, J.: Recent Progress in Atmospheric Chemistry Research in China: Establishing a Theoretical Framework for the “Air Pollution Complex,” Adv. Atmos. Sci., 40, 1339–1361, https://doi.org/10.1007/s00376-023-2379-0, 2023.