Reviewer 3

This manuscript introduces a drastically reduced isoprene oxidation mechanism "SICMA" for use in chemical transport modeling to minimize the computational cost of chemistry steps within the model. The authors specify a five-reaction, three-isoprene-related-species format for a reaction network meant to mimic the effect of isoprene photooxidation on formaldehyde, OH, and HO2. They then optimize tunable parameters within this framework (including some reaction rates and product yields) to minimize differences between SICMA and a detailed mechanism of concentrations of HCHO, OH, and HO2 using a simple box model setup. Next, they incorporate SICMA into the global MAGRITTE model and show that the reduced mechanism adequately represents outcomes of interest over major isoprene-emitting regions. Finally, they use SICMA to perform a model inversion, deriving isoprene emissions from measured TROPOMI satellite HCHO columns, and find minimal overall differences from an equivalent inversion with complex chemistry.

Overall, the manuscript is very well-written, and the utility of SICMA is made fairly obvious for this specific application of deriving biogenic VOC emissions from formaldehyde columns -- a process that requires substantial computational capacity and for which SICMA is specifically optimized. I recommend the manuscript be published in GMD following minor revisions to address the points raised below. I'd also like to note that my area of expertise does not include the emission inversion process, so hopefully another reviewer is better able to assess those sections.

We thank the anonymous referee for their positive feedback and helpful comments.

Based on the recommendations of the reviewers, we have made various changes to the chemical mechanism and the manuscript. We first briefly summarize the most important changes to SICMA.

In SICMA, we have introduced an additional reaction: ISOP + O_3. This brings the total to six reactions. The ozonolysis reaction is introduced to improve the nighttime chemistry in the model. During the day, this reaction has a near-negligible impact, but it is dominant during the night and significantly affects the diurnal cycles of isoprene, HCHO, and the nighttime HOx concentrations. Additionally, we have made the following changes to the parameter optimization algorithm:

- We now optimize parameters for HOx recycling in the isomerization reaction (R4).
- We increased box-model simulation period to 5 days in order to improve sensitivity of parameter optimization for the IOX reaction rate (k_5).
- We changed the fixed diurnal isoprene concentration profile to a diurnal isoprene emission profile to better account for the feedback of OH on isoprene.

Based on this new reaction scheme and optimization setup, we have derived a new set of parameters, which we implemented and tested in the MAGRITTE CTM. This new version of SICMA leads to improved diurnal cycles in the CTM and a smaller bias in HCHO and isoprene compared to the reference chemistry. The new results have been incorporated into the revised version of the paper.

Below, we address the different remarks of the referee (in black). Our response is formulated in blue font.

General Comments

First, I believe that the title and abstract moderately over-sell the utility of SICMA; it's very specifically designed to optimize formaldehyde, and does an excellent job as demonstrated here for inversions from formaldehyde observations, but specifically does not optimize NOx or ozone or allow for any isoprene (or monoterpene) derived SOA formation, which severely limits its utility for many other applications in global chemistry transport modeling. I'd recommend specifically referencing formaldehyde, or perhaps formaldehyde and HOx together, in the title to clarify.

Thanks for your comment. The mechanism is specifically designed for BVOC emission inversions using HCHO columns and lacks several key reactions needed for modelling NOy long-range transport (e.g. through organic nitrates), ozone production, and SOA. We therefore made the title of the manuscript more specific: “A simplified isoprene oxidation mechanism for fast formaldehyde-based emission inversion of isoprene”. The revised title reflects more accurately both the intended application and the current limitations of SICMA. We have also made changes in the abstract and throughout the text to make this point more clear.

Second and perhaps most importantly, I find the choice of combining isoprene with monoterpenes in a 10:1 ratio for the mechanism simplification and subsequent inversion very strange and highly limiting. if i'm understanding it right, you've optimized not an isoprene mechanism, but a fixed, specific isoprene + monoterpene combination that therefore doesn't really let you back out isoprene emissions precisely, and doesn't allow for flexibility in the spatial and temporal distribution of the isoprene-to-monoterpene ratio. Would it have been severely limiting to optimize a separate monoterpene + OH --> x*HCHO reaction (or a few reactions to get the NOx dependence right), or to consider the monoterpene-derived formaldehyde fixed and only optimize the mechanism and perform an inversion for isoprene?

I hesitate to suggest that the authors reformulate their mechanism or rerun any simulations, but at least acknowledging this detail of the mechanism in the abstract and discussing the implications for inversions and for forward modeling in the conclusions (beyond the cursory acknowledgment of the "crude" accounting on L 326) seems necessary to again avoid overselling the utility of this mechanism. The prospects for improving monoterpene representation in future versions of SICMA should also be discussed.

The SICMA mechanism indeed assumes a monoterpene emission source equal to 10% of the isoprene molar flux. This has been made clearer throughout the text. Including monoterpene species into the mechanism would have been possible, although not straightforward, at the cost of several additional species and reactions. However, SICMA was designed to be as compact as possible, with the use case of an adjoint-based isoprene emission inversion using HCHO data. For this use case, it was desirable to have a minimum number of reactions, with a focus on isoprene, which has been the topic of intensive research in the last decades and for which kinetic uncertainties are lower than for the monoterpenes. Though currently one of the major sources of discrepancy is the variation of the monoterpene-to-isoprene emission ratio, we argue that the main goal of isoprene emission inversion, focused on tropical ecosystems, can be achieved with this crude representation of monoterpenes.

Nonetheless, the addition of monoterpene chemistry to SICMA could broaden the applicability of the mechanism. We have added a note on the prospect of extending SICMA to monoterpenes in the concluding section.

Further more minor concerns and comments are provided (mostly with line numbers refering to the original manuscript preprint) below:

• Can you provide quantitative evidence of the reduction in computational cost with the SICMA mechanism relative to MAGRITTE for forward and inverse global modeling runs?

In our CTM simulations, we have retained the full set of VOC species and reactions to ensure an identical treatment of anthropogenic and pyrogenic VOC chemistry in both simulations, since these VOC categories share many species with isoprene and monoterpenes. Consequently, the reduction in computational cost of the simplified isoprene mechanism cannot be isolated directly from the computation time of the forward and inverse modelling runs.

Instead, we have calculated the reduction in computational time in the box model simulations between the MAGRITTE and SICMA chemistry. Here, we find a reduction of 83% in computational cost for a simulation of 14 days. We have added a comment on the difference in computational time in the box model simulations in Sect. 4.

• L 139-140: if both HO2 and isomerization contribute in the same environments, why is it necessary to separate them out into two reactions? Are their contributions spatially and temporally heterogeneous enough for this to matter?

The HO_2 and isomerization reactions can contribute similarly in some environments, yet different in others. This depends on the local HO_2 concentrations, which in turn depend on many environmental factors like NO_x and ozone concentrations, temperature, radiation, and humidity. In high-NOx conditions, the ISOPO2 + NO pathway generally dominates, whereas in low-NOx conditions, the reaction with HO_2 and the isomerization can both be important.

Our experience in testing a version of SICMA without the isomerization reaction in the CTM showed poor results, particularly for HOx and NOx concentrations. The isomerization is also an important sink for the isoprene peroxy radical ISOPO2, which would otherwise always be forced to react with either HO_2 or NO. By including the isomerization reaction, the fate of ISOPO2 is much closer to that of the analogous compounds in the MAGRITTEv1.2 chemistry.

• L 146-147 & 155-156: Isomerization being applied to all RO2, instead of just the fraction of RO2 radicals able to undergo 1,6 H-shifts, seems like it should introduce a lot of potential for deviation from the main mechanism. Is this satisfactorily ameliorated by fitting the other parameters? Why not scale the isomerization rate to the fractional abundance of the isomers that are able to undergo the 1,6 H shift? And similarly, why not optimize the yields of OH and HO2 in these reactions? Are they really so well-constrained with actual experimental evidence that they merit this fixed treatment?

The isomerization reaction rate already represents a bulk rate considering the fraction of peroxy radicals able to undergo the 1,6 H-shift, as is suggested in your comment. The MAGRITTEv1.1 mechanism itself is

not explicit (which would require far too many species for a 3-dimensional model) and incorporates bulk isomerization rates for two different pools of peroxy radicals, as described in detail in Müller et al. (2019). Based on the (bulk) isomerization reactions in the MAGRITTE chemistry, SICMA remains an even more condensed representation.

We thank the reviewer for the excellent suggestion to optimize the OH and HO_2 yields in these reactions. The HOx coefficients in reaction R4 were initially fixed to limit the number of free parameters in SICMA, leading to inaccuracies. In the revised version of SICMA, we optimize the HOx parameters of the isomerization reaction.

• L 163: This fixed value for O3 seems relatively low, as though it might underestimate the contribution of ozonolysis in many actual ambient conditions, and the fixed value for water vapor seem would also impose a limitation on the actual variability of formaldehyde yields from that ozonolysis. Can you estimate how much potential error this introduces to the mechanism by excluding an ozonolysis reaction, and how big a deal it would be if your simulations designed for mechanism optimization has higher ozone concentrations?

Thanks for this comment. We note that ozone is not fixed in the box model runs. Its initial concentration reflects low-NOx tropical environments. We have included the evolution of ozone concentration in Fig. 2. Despite the low initial ozone, conditions in the box model lead to a strong variation of ozone throughout the 5-day simulation. Nonetheless, we have tested varying the initial ozone concentration and humidity in the model, and found negligible impact on the results. Furthermore, as discussed above, we have included the ozonolysis reaction in the revised SICMA. It has a significant impact on nighttime HCHO production, HOx generation, and isoprene concentration. Although its reaction rate is relatively low, this reaction is important for the diurnal abundances of the relevant compounds.

• L 167: Imposing a peak of 1 ppb of isoprene and continuous emissions means that your mechanism is optimized for early-generation, short-term formaldehyde production; could this be why you have lower formaldehyde columns over the ocean in comparison to MAGRITTE (fig 4)? Could this be fixed by optimizing against a time series that allows for longer durations, and either (a) adding a third generation, or (b) allowing the IOX + OH reaction to produce a small amount of IOX, or of another formaldehyde precursor already contained in the mechanism?

The HCHO column underestimation over the ocean is likely not related to the long-range transport of VOC intermediates. Instead, its major cause is the absence of long-lived organic nitrates, which can release NOx in remote regions. This then leads to higher OH and more background HCHO production through methane oxidation in MAGRITTEv1.2 compared to SICMA, as already discussed in Sect. 5.2.

We thank the reviewer for the helpful comment on the fixed isoprene profile in the box model runs. We updated the box model setup, which now includes a diurnal isoprene emission source and a 4.5-day simulation period. This captures both the early HCHO generation and the delayed production through the formation of IOX. We found that the optimized parameter set of SICMA leads to a good agreement with MAGRITTE after 7 days, suggesting that the delayed production is sufficiently well captured. A third-generation product could make slight improvements, but was not considered given the already good performance of the current mechanism. Regenerating IOX is an interesting idea, though we have not tested this in the mechanism.

• L 197-202: Can you comment / speculate on what causes these differences, and potentially how to fix them in future SICMA versions? It seems in particular that a midday difference would be problematic for comparison with a midday satellite...

The nighttime differences in OH and HO_2 were mainly caused by the absence of the ozonolysis reaction. After including this reaction in the revised SICMA version, the nighttime OH and HO_2 concentrations have improved considerably (see Fig. 2 of the revised paper). There is still a small midday difference for HCHO in the box models, with higher HCHO concentrations in SICMA. This is likely due to a higher noontime HCHO production in combination with lower morning and evening HCHO production.

It is unclear what the exact reasons are for this discrepancy, but it is likely related to the simplistic representation of IOX, with only one reaction (IOX + OH) and a single rate (k_5). In MAGRITTEv1.2, isoprene oxidation leads to a long series of oxidation steps involving many intermediate species which can react in various ways and at different rates. A different evolution of HCHO production is therefore to be expected.

• Fig 3: Why does Figure 3 show the best HCHO agreement midday if that's when Figure 2 showed the worst agreement?

With the inclusion of the ozonolysis reaction, the diurnal cycle of SICMA is improved, especially during nighttime. However, we still see a larger nighttime production of HCHO in MAGRITTEv1.2, and in the absence of a strong HCHO sink, the HCHO difference builds up until the end of the night. During daytime, both HCHO production and loss rates increase substantially, primarily due to the stronger photochemical activity and increased OH concentrations. The resulting faster chemistry gradually reduces the HCHO concentration difference inherited from nighttime, and the two mechanisms converge toward similar daytime HCHO concentrations. Consequently, the agreement between SICMA and MAGRITTEv1.2 is best around midday, despite the larger differences that developed during the preceding night.

• L 220-221: Is it important to also evaluate regions beyond the emission hotspots? Genuine question -- maybe if you're just trying to optimize total gross emissions, the answer's no, but it also seems like the potential for the reduced mechanism to deviate from the main one would increase in places either downwind from emissions, where the temporal evolution of the oxidation chain matters, or where you have a mixture of different emission sources

Currently, we have indeed evaluated the diurnal cycles only in emission hotspots. However, Figs. 4 and 5 show global maps of the differences in the CTM simulations, where the differences downwind can be observed. Since these are relative differences, large impacts can be seen in more pristine environments. Figure 5 suggests that the long-range transport of NOx reservoirs impacts chemical composition downwind.