This manuscript describes the estimation methods for the atmospheric chemistry models generated by the current version of MechGen. As model generation must account for all reaction classes, the text covers a lot of topics, making this a very extensive document even with much of the discussion and tabulated material being presented in the supporting information. I support publication of the current work.

In reviewing this work, I tried to separate between textual/presentation issues, the quality of the estimation methods within the chosen representation, and the choice of which chemistry to include and how to estimate it. I find the quality of the presentation of a high standard, with only small suggestions for improvement. The estimation methods as presented work well, and represent the state of the current knowledge within the chosen representation; I have no suggestions for improvement there. There are two topics (aromatics, and CI) where I feel the discussion/representation strays too far from how recent literature understands the underlying reaction mechanism. For the CI, this does not affect the final estimations, while for aromatics I think the representation is more problematic (see below). Still, the manuscript describes the current state of the ever-evolving implementation of Mechgen very well, and this review is not the time to change how the mechanism generation is done. Many of the comments below are therefore mostly for additional literature data, or suggestions for future improvements rather than a critique of the current work, and need not be addressed other than perhaps an explicit mention in the text and/or SI to aid the reader (Note: I refer to own publications extensively below, not because they necessarily represent the best available but because it is easier for me to find suitable examples; citing other references may be more appropriate for this paper).

A general comment is that the text relies often on "common knowledge" and is occasionally very light on references to the relevant literature. I recognize that this manuscript is not the place for an extensive literature review, but it in many places it uses "we assume" / "we estimate" without any literature provided, while one or two well-chosen references could provide a more solid foundation with provenance for the reader.

Specific comments, main text:

Line 8: "O³P"
This is typically written as O(³P); perhaps its useful to use the classical notation instead of the abbreviation in the abstract.

Table 2, bottom, and a few other places: Criegee intermediates are not biradicals

Table 3:
- Not all notations used here can be found again in the supporting information (e.g. FRing(site) is missing in the caption of table A-42. It would be easier to link with the SI if all symbols here were present in the SI (most are).
- HCNbrs and nHCnbrs are not defined. I assume they mean hydrocarbon-neighbours and non-hydrocarbon-neighbours ?
- It is not clear what the non-neighbouring "beta-substituents" are in the case of a cycloaddition like for O3. Are O3P and O3 swapped?

Line 208, CI are represented as G(OO*)-O-G(O*): but there should be no central ether oxygen

Section 2.1.4 VOC + O3:
Designating all CI as excited when formed is not in agreement with the reaction mechanism: a fraction of stabilized CI are formed thermalized, especially for larger VOCs, with the remainder of the energy in the carbonyl fragment. There is no impact on the implementation, except that not separating between nascently thermalized CI, collisionally thermalized CI, and excited reactions of CI, requires a very different parametrization for the pressure-dependence. See also the related remarks below.

Section 2.1.5 O3P addition
Mention explicitly that addition on the central carbons of a conjugated system is assumed to be negligible, as opposed to cumulenes.

line 328 to 330: "The reaction with OH radicals ... were non-negligible"
This sentence might perhaps fit better on line 326, after "... do not form radicals".

line 482: "in direct sunlight"
For consistency, use "direct overhead sunlight"

line 510 and other places: "unpublished results"
Perhaps change to "Carter et al., unpublished results"

line 523: "and the rate constants are in cm3 molec-1 s-1"
These are unimolecular rates, remove "cm3 molec-1"

Line 524: "R = 1.99×10-3 deg K mole kcal-1"
Kelvin is not "degrees Kelvin" (as opposed to e.g. degrees Celsius)

Line 539: "cyclohexadieneones" -> "cyclohexadienones"

Line 548: "which would not be expected if this reaction did not occur".
This requires some references to recent literature, or an explanation

Line 647: "This is based on data of Slagle et al. (1984) for vinyl and Slagle et al. (1988) for methylvinyl and i-C4H9·. More recently, Matsugi and Miyoshi (2014) directly measured the yield of HCO radicals from vinyl"
Can the generalization of the data for small C2-C4 compounds to large molecules (e.g. C10) be relied upon? The rapid energy distribution over the entire molecule could lead to vinylperoxy stabilization. E.g. Sandia Nat. Lab. studied alkenylperoxy radicals made in this fashion (Meloni et al. 2007)

Section 3.2.1: "We expect that there should not be large entropy or energy barriers to these reactions, so MechGen assumes they all...
This is an example of where literature is available showing this expectation to be realistic and documented.
Miyoshi (2011) reports results of theoretical calculations of hydroperoxyalkyl rate constants that extrapolate to much lower rate constants at atmospheric temperatures. Vereecken and Nozière 2020 already mentioned that at least some of the rate coefficients (OOQO2 H-migration) by Miyoshi 2011 may not be very accurate. Other theoretical work yields rate coefficients that extrapolate to higher values, competitive to O2 addition. E.g. Ali et al. 2023.

and the third involving a different type of cyclization where a per oxy O transfers to the opposite double bond, forming an epoxide and an alkoxy radical, with branching ratio fEpox."

Recent literature indicates that this path is not important, and that epoxides are formed through epoxidation of beta-unsaturated alkoxy radicals. I have discussed this in detail in the past (e.g. Vereecken 2018, 2019 based on extensive theoretical literature and some experimental observations, and the non-contribution of this channel has been shown directly in recent experiments by Xu et al (2020). Postulating the direct formation of an epoxy-alkoxy radical skips intermediate steps for the epoxy-formation which are influenced by the reaction conditions (e.g. bimolecular reactions of the epoxy-RO2 intermediates after alkoxy epoxidation). Hence, attempting to shoehorn the observations onto this incorrect mechanism will make it hard in general to match the model to experimental data across a wider range of reaction conditions and for different aromatics, even when it appears to work for the limited observations where the parametrisation is based on. I understand that this representation of aromatic chemistry is common (e.g. Jenkin et al. 2018), but I strongly suggest adjusting the base reaction template for aromatic chemistry at some point to improve agreement with the available theoretical and experimental data (see e.g. Vereecken 2018 for an example mechanism).

Section 4.7 "Ring closure of unsaturated peroxy radicals"
- The fast ring closure reactions of mono-unsaturated RO2 (see Vereecken et al., 2021) could be an interesting addition. These reactions are predicted to be quite fast; experimental data should become available soon. This reaction class is probably more common than conjugated alkadiene RO2s.
- The current literature does not support formation of epoxy-alkoxy radicals in cyclic nor linear conjugated peroxy radicals (see comment on aromatics)
- Epoxidation is likely due to cyclization in beta-unsaturated alkoxy radicals (see also comment on alkoxy radicals epoxidation)

Section 5: "reactions of alkoxy radicals"
Epoxidation of beta-unsaturated alkoxy radicals could be an interesting addition. Theoretical work shows it is important in aromatic compounds (see above), and to some extent in the oxidation of multi-unsaturated species such as isoprene (e.g. Vereecken et al 2021, and Carlsson et al. 2023). It is an additional pathway for epoxide formation, in addition to the (higher barrier) beta-OOH alkyl radical decompositions (e.g. D'Ambro et al 2017, and IEPOX formation from isoprene)
In all cases they [CI] are expected to be formed in a vibrationally excited state, with the amount of excitation depending on the formation reaction, but sufficient to allow rapid interconversion between the syn and anti forms where applicable.

Configuration is assumed not to be a factor for the excited intermediates because they are assumed to rapidly interconvert.

This is not in agreement with the ozonolysis mechanism as derived in many experimental and theoretical studies. The syn/anti isomerisation TS has barriers of over 25 kcal/mol even in the most favorable cases and 30+ kcal/mol being more common (Vereecken et al. 2017, 2022 and literature therein), such that syn/anti isomerisation has reaction rates several orders of magnitude below the various exit channels accessible to all CI with barriers significantly below that (at the very least, 1.3 ring closure accessible to all CI with barrier ≤ ~20 kcal/mol even in the worst case and <15 kcal/mol being more common). These other channels will be taken in preference, at any level of excitation attainable from VOC ozonolysis or carbene+O2 reactions. For all practical purposes the syn/anti ratio is set directly at the dissociation of the POZ. Additionally, in POZ dissociation there is a fraction (small though it may be) of the CI that are formed stabilized, where "stabilized" means with an energy content comparable to or below the lowest exit channel barrier such that they will behave thermalized. The remaining energy from POZ dissociation is released towards the other fragment and the relative motion of the fragments. The yield of nascent stabilized CI depends on the size of the VOC, the relative size of the fragments with the concomitant probability distributions between the degrees of freedom, and the barriers for the available exit channels for the CI formed. The experimental yields of SCI as a function of pressure are difficult to reconcile with a generalized 100% excited CI formation especially for larger VOCs. Interesting here are the experiments by Drozd et al (2011), Newland et al (2020), and the predictive capabilities of the Newland et al SAR (2022) based on the energy distribution over the two fragments. Only for cyclic alkene ozonolysis and carbene+O2 reactions is it reasonable to assume 100% excited CI, as only a single product is formed.

In the end the Mechgen predictions for the product yields from VOC ozonolysis is not affected by how the CI formation and decomposition process is thought of, as the parametrization is designed to match the experimental observation irrespective of the underlying assumptions. Still, it might make sense to at least mention the possibility of a residual yield of stabilized CI even at low pressure, and state that syn/anti isomerisation has very high barriers compared to the other exit channels, making syn/anti isomerisation a negligible channel for any attainable excitation.

A recent update is available for oxygenated and unsaturated SCI (Vereecken et al. 2022)

Excited intermediates with zero or one substituent can undergo "hot acid" rearrangements where an H bonded to the Criegee group is involved in a rearrangement forming a highly excited acid, which then decomposes in several ways.

Decomposition of acids to anything other than CO, CO2 or ketene seems tenuous. as it is not supported by pyrolysis experiments of acids, not by theoretical literature which finds that all other pathways have too high a barrier to contribute significantly. It was a useful placeholder in early literature until better information was available, but other pathways have been found to explain some of the observed products. For example ethene ozonolysis (OH formation without a VHP channel) was often explained through hot acid decomposition, but seem more likely to be governed by ketohydroperoxide chemistry (see e.g. Pfeifle et al 2018 and later work on this channel). For some acid measurements it has been shown that they are probably secondary products (e.g. Orzechowska et al. 2005a/b).
Line 1515: "where an H bonded to the Criegee group" -> Perhaps say "an H-atom bonded to" for easier parsing (reads like H-bonded)?

Lines 1519-1526: ring closure of beta-carbonyl CI to a POZ
The middle structure in the scheme requires a double bond between the carbon atoms, and is equivalent to the POZ formed in an alkyne+O3 reaction. Vereecken et al. 2017, and Cremer et al 2001 calculated a large barrier (≥ 30 kcal/mol) for this ring closure, and Vereecken et al. 2017 state explicitly that ring closure is negligible against other CI reactions, making the O-rearrangement moot even for excited CI.

Specific comments, supporting information:

Page numbers are missing on some pages

page 7, notation for chemical species
This naming scheme allows for multiple different notations for the same molecule, similar to some other line notation (e.g. SMILES). This is flexible and works well with graph/connectivity implementations. However, for databases and search implementations it is useful to have a canonical notation that provides a well-described unique way to write a molecule, irrespective of how the structure is perceived. Is there a canonization algorithm?

page 17: "The product data for these reactions (Calvert et al., 2000) are consistent with addition to the primary radical occurring ~35% of the time."
O2-addition on allylic radicals is reversible, allowing re-equilibration between the various RO2 (see Peeters et al. 2009, 2014, Novelli et al. 2020, Medeiros et al. 2022). Assigning a single site-specificity without accounting for the redissociation does not allow a correct description of the RO2 isomeric distribution across all reaction conditions as the actual distribution depends on the reaction conditions. E.g. at high NO the product distribution is determined by the nascent RO2 distribution, while conditions allowing long-lived RO2 will be governed by an equilibrium RO2 distribution, which may not be the same as the nascent distribution (see e.g. isoprene).

page 17: "the non-oh OH substituted position" -> remove "oh"

page 18: cyclization of peroxy-substituted alkyl radicals
Recent theoretical work shows that the barrier height does not decrease smoothly with decreasing ring size (e.g. Ali and Saswathy, 2023)

page 20, Table A-9, entry for allylic radicals: the given ratio k(isom)/k(r+O2) seems not to take into account that the O2 addition on such radicals is reversible, and the "effective O2 addition rate" is thus also affected by the loss processes accessible after O2 addition.

page 27: "peroxy + peroxy reactions are generally not important loss processes"
This depends on the application. If SOA/aerosol formation is being modeled, dimer formation from RO2+RO2 is important even if it is a low-yield channel.

page 31: "unspeciated nitrate yields from formed in reactions " -> remove either "from" or "formed in"

page 37: "This is the either ratio" -> swap "the" and "either"

page 37: acylperoxy radicals
A more detailed set of corrections for acylperoxy radicals was recently given by Seal et al. 2023

page 39: "However, the three-parameter expression in Equation (III) is not well suited for general estimation purposes"
Perhaps mention that this is due to the strong cross-correlation between the three parameters when fitting

page 41: "relative to aliphatic unsubstituted radicals, which we presume refers to reactions of -C[OO·] radicals"
That is correct. See also Seal et al 2023 for more direct data and correction factors as a function of H-shift span.

page 41, caption table A-17, last line: "were derived" is very widely spaced.

page 47: "so that was taken as the actual lower limit for Ea0, and was used whenever the estimated values were lower than that."
This prohibits estimating relative rates between H-shift reactions within the same molecule, if more than 1 channel is "fast".

page 48: "Note that strain energies should or might be significantly different in reactions of cyclic peroxy radicals"
Vereecken et al. 2021 has rate coefficients for cycloperoxide-peroxy radicals. Irrespective of the presence of the oxygens in the ring backbone, the ring strain seems to prohibits significant H-migration in most of the configurations examined. Maybe this information could be of use, though that paper also makes the point that transference between the cyclic and acyclic cases is less than obvious.

page 52: "against the barrier heights calculated by Vereecken and Peeters (2009)"
There is an update for this SAR in Novelli et al. 2021

page 53, figure A-14:
This deviation is expected: the activation energy Ea contains both energetic factors (barrier height Eb) and temperature-dependent entropic factors. Deriving Ea solely from Eb is missing this aspect.

Page 56: "to the ketone group ketone products formed." -> correct sentence

Page 73 to 83: I'm not overly fond of using "Criegees" in formal text. If a short notation is needed, "CI" might be preferred.

Page 77: "excited CH3C[OO]" -> "excited CH3CH[OO]"

page 77 : "Note that the VHP reaction is not possible for tertiary substituents that do not have a hydrogens, so the excited radicals would only react via the hot acid reactions if possible, or only be stabilized if not, and the stabilized radical would not decompose even if it were in the syn configuration."
All CI have access to the at least the dioxirane route (1,3-ring clousre), as it only involves the [COO] moiety; only the rate is affected by the substituents.

page 77 : "do not have a hydrogens" -> "have a hydrogen" or "have hydrogens"
page 78: "One might expect that the extent of stabilization would increase with the size of the intermediate"
This was studied also by Newland et al 2020

page 82: "but Vereecken et al. (2017) apparently did not consider subsequent reactions of this excited species."
Vereecken et al. 2017 refers to the work of Kuwata et al. 2005, 2008, who studied this in detail. Since then there is also the experimental/theoretical work by Vansco et al. 2020, though these authors seem to ignore the epoxide route in favor of the H-migration route.

page 82: "we assume that this type of O-shift reaction is faster than any competing possible reactions for excited carbonyl-substituted Criegees"
This is in contradiction to the theoretical data, which finds unsurmountably high energy barriers for the ring closure to the POZ, compared to the other accessible reactions. (Cremer et al. 2001)

page 83: "the CIs. On the" -> missing space

page 109, caption Table A-42
This should explicitly mention the notation used in table 3 in the main paper. I assume this is "FHRing(site)". Please check whether all notations from table 3 are reused in the SI.

References


