Earth Syst. Sci. Data Discuss.,
https://doi.org/10.5194/essd-2019-236-RC2, 2020

Open Access Earth System Science Data Discussions

Interactive comment on "Database for the kinetics of the gas-phase atmospheric reactions of organic compounds" *by* Max R. McGillen et al.

Anonymous Referee #2

Received and published: 3 March 2020

General comments

This paper reports and accompanies the release of a compilation of evaluated kinetic data for organic compounds, covering their reactions with OH, Cl, NO3 and O3. The resultant database is much larger than those for the same reaction classes in other reported evaluations, including rate coefficients for 2765 reactions of 1357 compounds. This is an exceptionally valuable resource for the atmospheric chemistry community, particularly for developers of structure activity relationships and explicit organic degradation mechanisms.

Although this is a very large database, the authors point out that there is the scope for further expansion, e.g. to include product channel branching ratios and reactions involving other oxidants. In addition, they also report the intention for the resource to be-

come a periodically updated living database. These are commendable aims, although only time will tell if they are practical and achievable with the resources available. As is inevitably the case with this type of activity, it will likely rely on the goodwill and independent efforts of the contributors, but hopefully publication of this work, followed by use and suitable citation of the database will help to stimulate such activity and support.

This paper acts as a reference document for both users and ongoing developers of this resource, and the publication of overview papers of this type is therefore important and necessary. In addition to providing a suitable citation that can acknowledge and reflect the considerable time and effort that goes into developing and maintaining a state-of-the-science resource such as this, it acts as a record of how the resource evolves. I recommend publication, although the authors should consider the mostly minor points below.

Specific comments

It is mentioned and demonstrated at several points that the evaluation is vastly larger than existing evaluations, such as those of IUPAC and JPL – and this is undoubtedly true. In the case of IUPAC (for example), I note that each evaluated reaction has a detailed data sheet, providing a summary of the reported measurements, and an explanation of the basis of the recommended rate coefficient. In some cases where there is only a single measurement, a recommendation is not made. Presumably, many of the enormous number of reactions covered by the present compilation have only single measurements, If so, can the authors confirm that these reactions are evaluated, and the reported measurement is not simply adopted because there is no conflicting information? Have any such reported measurements been excluded? It is noted that lines 180-190 summarise an evaluation procedure involving expert judgement bullet points and duplicate/triplicate checking. However, doing this for 1000s of individual cases with a single reported measurement is clearly very time consuming and not trivial. A little more information on the such practicalities, even if rather mundane, would be useful.

In addition, there is an inevitable inference that can be made that this work supersedes that of the other listed evaluations, and renders them redundant – even if that is not intended. This is because this work is reported to encompass everything in the other listed evaluations and much, much more. I suspect the authors do not desire this (given that there are contributors in common, and the goals and scopes of the activities are not the same), and this should probably be made clear.

Line 123: E and R require definitions here, rather than later (for E) and not at all for R.

Line 143: The kinetics of product formation from any reaction channel are determined by the overall rate coefficient for the reaction (i.e. product formation and reagent loss display the same time constant, even if that product channel does not account for the whole reaction). Indeed, many rate coefficients have been measured that way. Although the overall point being made in this sentence is valid, it is not a very clear way of making it.

Figure 5: The unknown unknowns presumably also include species that could be formed from the GECKO-A detailed protocol, but are ultimately excluded because of mechanism size limitation measures resulting from a simplification protocol (e.g. exclusion of minor reaction channels below a specified threshold). I think these are more correctly classed as “unexplored but potentially known unknowns”.

Line 37: I do not believe that alpha-pinene is mentioned in Aumont et al. (2005), and the potential number of species that can be generated is likely considerably more than 400,000 (given that Fig. 5 shows that a C7 1-alkene makes almost that many, with a 5% trimming threshold). Presumably, this number relates to the number of species generated in the version of GECKO-A you have run to produce Figure 5, including whatever simplification (trimming) measures are represented to limit mechanism size.

Interactive comment on Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2019-236, 2020.