



EGUsphere, referee comment RC1
<https://doi.org/10.5194/egusphere-2022-245-RC1>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on egusphere-2022-245

William Carter (Referee)

Referee comment on "GENERator of reduced Organic Aerosol mechanism (GENOA v1.0): an automatic generation tool of semi-explicit mechanisms" by Zhizhao Wang et al., EGU sphere, <https://doi.org/10.5194/egusphere-2022-245-RC1>, 2022

This paper discusses procedures developed to reduce the size of large gas/aerosol mechanisms to greatly reduced mechanisms that give predictions of secondary organic aerosol (SOA) that agree with those of the larger mechanism in a selected set of environmental conditions to within a specified tolerance. Because of the complexity of the atmospheric reactions of organic compounds, it is necessary to use reduced mechanisms in practical airshed model applications, but most reduced mechanisms used in airshed model applications were developed primarily focused on accurate ozone predictions. However, SOA predictions are much more affected by the chemical complexity of organic reactions than predictions of ozone, and developing methods to reduce mechanisms for without significantly affecting SOA predictions is an important research priority in atmospheric chemical mechanism development.

In fact, work with GECKO suggest that multi-generation mechanisms may be necessary for reliable SOA predictions, and even the MCM, which greatly lumps reactions of 2nd and higher generation products, may be too reduced for this application. However, the use of MCM as an example is sufficient to illustrate the method, and as discussed below MCM may be about the largest mechanism that could be reduced using the method discussed in this work, given current computer capabilities. In any case, without some way to reduce these huge mechanisms, we have no choice to continue to rely on the empirical and parameterized SOA models that are adjusted to fit SOA yields measured environmental chamber data without consideration of the actual chemistry and how different chemical conditions in the atmosphere affect SOA yields. If we had suitable reduction methods for SOA predictions, then use of SOA models that are based on actual chemistry might become practical.

The method discussed here involves use of a 3D grid model representing a large continental domain and various seasons, with SOA calculated using the large mechanism to be reduced as the starting point, and then uses sets of 0D scenarios derived from selected grid cells and times during the grid model simulation in an algorithm to develop the reduced mechanism. This has the disadvantage that it requires a full 3D calculation

with the full mechanism, which may be possible for MCM, but not for huge GECKO-like mechanisms that are probably what are really required. It may be possible to revise or supplement this method so that OD calculations with selected scenarios may be sufficient to serve as the standard, but that is not discussed in this work. However, the method discussed here is a useful starting point, and is worth publishing for this reason.

The focus of this paper is a specific software package developed for mechanism reduction, which they call GENOA, though the main interest of this paper from a scientific perspective is the method itself. The name of the software is somewhat misleading because it is not actually a mechanism generator, but instead is a method to reduce existing mechanisms, potentially including those developed by actual mechanism generators such as GECKO. This could potentially be a useful tool for the atmospheric chemistry research community, if suitably documented and made publicly available. The fact that they include a users manual along with the Supplementary Information suggests that this is the intention of the authors. If it has sufficient flexibility, it could possibly be used for other criteria besides SOA predictions and perhaps even for other applications besides atmospheric modeling, such as, for example, combustion modeling or reducing large liquid-phase mechanisms. However, the users manual does not contain sufficient information to actually run the model for general applications, other than duplicating the results given in this paper. In particular, it would need information on how to interface this with output of existing 3D air quality models.

It is worth pointing out that the reduced mechanisms developed using this are strictly speaking reliable to give predictions to within the desired tolerances only for the airshed conditions used in its development. They developed this example using the conditions of all of Europe for a whole year, but it may be more practical to use this to optimize mechanisms for specific urban scenarios, for use in regulatory modeling. To be useful for this, the users manual would need to be improved so it can be used with other 3D models and modeling scenarios.

Although this paper is reasonably well written, it does have areas where improvements are needed before it is accepted for publication, and I have some suggestions. These are given below in approximate order of importance.

It looks relative changes in SOA concentrations are used as the criteria to test a reduction approach. This means that 5% error in a grid cell where almost no SOA is formed is given equal weights to 5% error in a grid cell with high SOA, where the model prediction is relatively more important. Wouldn't absolute error be a better criterion, or at least among the criteria employed? Shouldn't there be a cutoff to remove cells with very low SOA, or was this incorporated implicitly by the choice of testing scenarios?

The 3D model simulations of continental Europe employing the chemistry-transport model

CHIMERE was used as the standard against which the reductions were compared. However, CHIMERE mechanism is not exactly the same as MCM, and I could not find an indication of whether the "CHIMERE surrogate for sesquiterpene" (line 198) is exactly the same as the MCM β -caryophyllene mechanism as used in this work. If that is the case, it should be stated explicitly when CHIMERE is first mentioned, since I couldn't find such a statement in Lanzafame et al (2022). If it is not the case, then justification needs to be given as to whether this is an appropriate standard against which to test the reductions.

It is not clear to me whether the concentrations of the inorganic species such as OH, O₃, HO₂, etc were constrained to be exactly what was calculated by CHIMERE in the 0D models for the selected locations, or if they were calculated using the mechanisms using boundary or initial conditions somehow obtained from CHIMERE. If the former (which I presume to be the case) this should be stated explicitly, and if the latter than more detail about the inputs to the 0D models need to be provided, if only in an Appendix.

The computation of stoichiometric coefficients and rate constants for new lumped species (as shown in Figure 2) based on concentrations and lifetimes need to be discussed. The use of "average produced concentrations from five-day 0D simulations" needs more discussion than as a footnote in the table. Presumably the [HO₂], [NO], etc. concentrations used to compute the lifetimes are also averaged? What are the 5 0D simulations they use to compute the average? The later discussion indicates they use 8 scenarios for initial training, and more than that when they are close to being finished.

The maps indicate that the reduction errors are the greatest in Southern Europe than in the North. Is that because there is more secondary SOA predicted for the South, which should be more photochemically reactive? It might be useful to show a relationship between SOA level predicted and the reduction error for the various grid cells or scenarios.

On line 169 they state that "jumping" is restricted to cases where the intermediate to be removed results in the formation of only a single compound. However, this is not the case for the examples they show on Table 4, where each intermediate forms more than one compound. (Reactions 14 and 15 in Table 5 reflect merged competing reactions, and are not explicit.) In fact, any rapidly reacting intermediates that have several reactions that form different products, can be "jumped" without affecting simulation results as long as their reactions are either unimolecular or with O₂, so lifetimes would not vary with conditions. This is the case almost all of the alkoxy radicals in explicit atmospheric mechanisms. This should be pointed out in the discussion. The main approximation with "jumping" compounds with several reactions would that it would not deal with temperature effects if some of the competing reactions have different temperature dependences in the full model. Whether this is acceptable for their continental model could be determined by their algorithm.

The example of lumping they present in Table 2 gives a rather trivial reduction example that should not really require an algorithm. The chemical lifetimes of the 3 peroxy radical species being lumped are the same for the major loss processes, with the only difference being the RO₂+RO₂ reactions that make almost no contribution (see Figure 2), so the

"weighting ratios" should actually be very close the stoichiometric coefficients in the explicit mechanism (reaction "0" in Table 1), and the simulation results should also be the very close. An additional example where the lumping may not be as obvious, and which may potentially be more dependent on environmental conditions, should also be given. This would aid the discussion of the lumping approach and give the reader a better understanding of the lumping approach and its general utility and potential drawbacks. Or is the kind of peroxy radical lumping illustrated in Table 2 the only type of lumping they actually did? If so, this should be stated.

I don't understand the difference between "removing species" and "removing elementary-like reactions". Strictly speaking, all species that are removed should be done by renormalizing yields from competing pathways if the reactions are lumped, which is the same as converting them to "elementary-like reactions", and then removing the less important reactions. Without renormalization of the other branching ratios, the removal would cause an unnecessary loss of mass or, worse, an artificial radical sink. An artificial radical sink may not affect SOA predictions if the concentrations of the inorganics are constrained in the test simulations (and thus would pass the tests in the reduction algorithm discussed here if that were the case), but the reduced mechanism would be unsuitable for predicting radical levels if used in a full model. Even small artificial radical sinks may have non-negligible effects on the radical chain reactions involved in many photooxidation mechanisms.

Section 2.2.3 describes the 12,159 conditions used for the testing datasets. It should be stated in that section that Figure D1 has a map of these locations. What are the 4,717 conditions mapped on Figure 6?

Are any of the lumping criteria listed around Lines 134-143 options that can be varied by the user? Define "extremely short lifetimes". Do they allow lumping compounds that react with different atmospheric species or types of reaction (e.g., with OH, O₃, NO, etc, by photolysis, or unimolecular), but happen to have comparable lifetimes for most scenarios?

It probably would be appropriate to point out that the MCM mechanism they use as the starting point has already been reduced to some extent. It already "jumps" carbon-centered radicals by replacing them by the peroxy radical formed when they react. It also combines elementary reactions with the same reactants into combined reactions with non-integer stoichiometric coefficients. Removing species from such reactions should not be done without renormalizing coefficients representing the competing processes, as discussed above.

Figure 2 and Table A1 use " R_{RNO_3} " to indicate the fraction of peroxy radicals reacting with NO₃. The term " R_{NO_3} " should be used instead, to be consistent with the terminology for the other peroxy loss processes. I found this confusing when I first saw this.

Suggestions:

It would be better if the four items in the reduction strategy were discussed in the order they are actually applied in the algorithm. "Lumping" and "replacing" is discussed first, but actually "removing" and "jumping" is done first, which is appropriate.

It would be useful to include a table of options or parameters that control the reduction process, indicating which can be changed by the user, and the value(s) of the parameters used in the examples in this work. This would be particularly relevant if the intention is to make this software available for use by other researchers.

The geographical and meteorological conditions for the 8 OD scenarios used for the training dataset are summarized on Table 5, and a map of their locations is shown on Figure C1 in an Appendix. The map showing the locations is useful and should be in the main paper, and if the points are shown along with the "errors" on Figure 6, it would take no added journal space and also indicate whether there is an relationship between the errors and the locations of the training scenarios. It also would be helpful for Table 5 to have columns showing the amounts of SOA calculated, the major daytime loss process for β -caryophyllene, and at least the day and night fractions that peroxy radicals react with NO. These are the most important chemical differences in the scenarios, and including this on the table would give this information in a more compact and perhaps clearer manner than on Figure 2. Figure 2 is also useful, but it is difficult to read and could be formatted better.

It would also be useful to include a map showing locations where the 150 testing scenarios were derived. This may cause too much clutter if added to Figure 6, but should at least be in the Appendix. There should also be a map showing the SOA levels calculated by the CHIMERE model for the European domain. Perhaps it could be combined with the figure in the Appendix showing the locations of the testing scenarios, though it may be useful enough to be in the main text.

Figure 4 would be more useful if it included some brief indication of what kinds of compounds the model species refer to.