Review of Sengupta et al.

The modeling of secondary organic aerosol (SOA) formation involves enormous complexity. Sengupta et al. used a perturbed parameter ensemble of 60 model simulations to evaluate six representative SOA precursors on aerosol number concentration and organic aerosol mass. The model results have been evaluated against measurements from sites across the globe. The authors determined the critical role of ELVOC and LVOC in new particle formation and growth, and the importance of LVOC and SVOC in simulating organic mass. They also found a high degree of equifinality in the SOA model. Parameter combinations that are best for aerosol number concentration are worst for organic aerosol mass. This manuscript provides a novel approach to evaluate effects of parameter combinations, rather than individual parameters, on both aerosol number and mass concentration that matter for climate impacts. I would recommend publication of this manuscript after addressing minor issues below.

1. A brief introduction on the meaning and climate implications of N3 and N50 should be added somewhere in the introduction (not later in Page 20).

2. Some more information about the GLOMAP model is needed in Section 2.1, e.g. what are the emission inventories used here? How does the model treat the removal of aerosols? What do anthropogenic VOCs include?

3. Table 2: B_ELVOC does not produce any SOA in the default scheme? Why?

4. Figure 2: Would you please say a bit more about how you design these 60 experiments? It may be clearer to include an appendix table to show the combination of yields you chose and explain why. At P14 L18 you said, “the total global production of SOA varies by only a factor of 4 and the lowest value is 220 Tg yr−1”. This really depends on the choice of yields combination, right? If you choose 0 for all yields, the SOA would be zero? And from Figure 2, I did not see an experiment that uses maximum scaling/yields for all sources. Would such an experiment show a much higher SOA production than 850 Tg/yr?

5. P13: Could you please provide more information on TSS? How did you calculate R and Ro? Is the normalized standard deviation calculated by all model vs observation monthly mean data across the globe? And could you give a more intuitive interpretation of the TSS values? Later in Figure 10-12 you defined “low/poor” versus “high/good” TSS, which should be explained here as well.

6. P14 L4: “The global mass of SOA produced in the model simulations ranges from 220 to 850 Tg yr−1”. Doesn’t this include the 107 Tg/yr of SOA produced by the default SOA scheme (based on Table 2)?

7. P19 L18: Why do high B_SVOC_I and A_LVOC suppress particle growth?

8. Caption of Figure 7 is wrong. N3 is not shown. Q2 and IQR not consistent.

9. Figure 7: Could the “global mean” OA be as high as 3-10 ug/m3? Did you do an area-averaged OA over the whole globe (including ocean), or the total land area, or averaged over the available sites?

10. P20 L13-14 (the first sentence actually) should refer to Figure 7.

11. P20 L21-22 Did Figure 7 include any information of nucleation? “Including new, more accurate nucleation pathways into models is unlikely to improve the model performance with respect to N50.” This sentence is confusing.

12. P20 The whole paragraph “OA concentrations in Figure 7 are found to be unrelated to B_ELVOC…” It could be helpful to include regression lines and some statistics (e.g. slope, significance, correlations) to support your statements here.

13. P21 “This is because A_LVOCs or A_SVOCs grow fewer particles than their biogenic counterparts and as a result changes in their concentrations have a lesser impact on simulated OA mass. ” Again, is such information included in Figure 7? Is it possible that it’s just because anthropogenic sources produce much less SOA than biogenic sources on the global scale?

14. P23 L24 “the model overestimates the strength of seasonal cycle”. This is only true for just a few sites, right?

15. Why not include color bars to show the TSS range in Figure 10-12?

16. P26 L17-26. These interpretations are very confusing, and Figure 10-12 are hard to understand. For example:

“where B_ELVOC yield is greater than 19.8%”: Does this refer to first column in Figure 10?

“B_LVOC is less than 113 Tg yr−1 (corresponding to a yield of 100% in Figure 10 first column, second row from top)”: I think this should refer to second column in Figure 10? The “first column, second row from top” refers to the scatter plot with B_ELVOC on the x-axis and B_SVOC_M on the y-axis. I don’t think this is the one you are referring to here.

“where the sum of anthropogenic LVOC and SVOC is greater than 127 Tg yr^{-1} (200% yield in Figure 10 fifth column, sixth row).” – There is no sixth row in Figure 10. Should be fifth row?

17. P26 Second paragraph about N50. Again, “Figure 11 first column, second row” refers to the scatter plot with B_ELVOC on the x-axis and B_SVOC_M on the y-axis, not consistent with what you are saying here.

18. Figure 10-12 may need to be rearranged for an easier interpretation. For example, both X and Y axis should include 6 ox-VOCs and in the same order, i.e. have B_ELVOC, B_LVOC, … A_SVOC on the X-axis from left to right, and on the Y-axis from bottom to top. Also please define “good” and “bad” TSS scores.

19. P31 Second line “… for all five ox-VOCs in 1-D”: should be “six ox-VOCs” here.

20. P34 I like the statements about B_LVOC, B_SVOC_I/M, A_SVOC which are much clearer than in the results section, but would you please refer back to the figures from which you draw those conclusions? In general Section 3.2 is hard to follow with quite a few typos/errors and Figure 10-12 are not easy to interpret. Please check through the manuscript for potential mismatches and typos.