Interactive comment on "Assessment of dicarbonyl contributions to secondary organic aerosols over China using RAMS-CMAQ" *by* Jialin Li et al.

Anonymous Referee #2

This manuscript simulated the SOA production by the aqueous uptake of dicarbonyls on cloud droplets (pathway M) and evaluated the impact on overall SOA concentrations over China. The authors also explored the impacts of increased aromatic emission and glyoxal yields from isoprene on the SOA production from pathway M.

While the subject is of interest to the community, the present paper is confusing in its methodology and conclusions. For example, the authors went to great length to prove that aromatic emissions may be underestimated, which certainly would have an impact on the simulated glyoxal concentration. But then the authors simulated Case 2 by tripling the aromatic emissions AND increasing the glyoxal yield from isoprene

by a factor of 5, without giving any rationale for the latter change. It is true that many models may have underestimated glyoxal yields from isoprene (Galloway et al. 2011), but the reader was under the impression that that was not the focus of the paper. Why then, did the author mix-up the two potential cause of glyoxal-underestimation in their study? Why not evaluate the impact of aromatic emissions first and then move on to the impact glyoxal yields from isoprene? I strongly suspect that the improvements in Case 2 relative to Case 1 was mostly due to the change in glyoxal yields. If the authors what to focus on this effect, they need to make the rationale clear from the introduction and methodology.

Also, many statements were not supported by the figures. For example, the authors stated that the simulated SOA in Case 2 increased by a factor of 2 relative to that in Case 1. This was not seen in Figure 4. Also, Case 0 (base case) in the subplots of Figure 4 were all different!

Overall, I recommend that the paper undergo major revision to clarify the confusing points.

Other major comments:S

lines 107-108: Is NCEP data not used as boundary condition to the RAMS model?

lines 246-247: "The O/P ratios are more concentrated for ARO1 than for ARO2, with both the mean and median O/P values being close to 2." Not sure what the authors meant by 'the O/P ratios' being more 'concentrated'. Did the authors mean that the O/P ratios for ARO1 were less variable than those for ARO2? Also, what does the variation refer to? Is it the hourly/diurnal variation? Or is it the day-to-day variation? Please also clarify in the caption of Figure 3.

LInes 253-254: Again, what did the authors mean by 'the ratios of ARO1 and ARO2 in episode 1 are more concentrated'? In general, 'concentrated' appears to not be the appropriate word in this context.

Lines 265-266: 'Case 0 is the base example'. Please clarify what this means? Do the authors mean that Case 0 was the simulation with only the three default SOA formation pathways in CMAQ (i.e., without the aqueous uptake production)?

Lines 269-270: What is the rationale for increasing the glyoxal yield from isoprene by 5? The authors did not mention this in the text up to this point. What was the original yield?

Lines 273-274: "The observed PM1/PM2.5 ratio of 0.77 (Xu et al., 2015) was used to convert the observed concentrations for comparison with the simulated results for PM2.5." Is this assumption appropriate for SOA? What are the evidences?

Lines 275-298 and Figure 4: The discussion here is confusing! The authors said that Case 0 (red lines in Fig 4) is the base case, which I assume is the simulated SOA from CMAQ from the simulation using the original emissions and using the standard SOA formation pathways. Why is it then, that the red lines in (a)(b) are different than those in (c)(d) and those in (e)(f)? Clearly there are some misunderstandings. Please clarify. Also, if the red lines should be the same in (a)(c)(e) and in (b)(d)(f), the subplots should be combined to make comparison easier.

lines 310-312: "When the impact of the underestimation of glyoxal is taken into consideration in case 2, the mean concentration of AORGCJ in both episodes increases by nearly a factor of 2 relative to case 1" I do not see this in Figure 5. Please clarify.

Table 2, Figure 4 and the related text: This tables shows that the simulated SOA is nearly doubled in Case 2 relative to Case 1 during episode 1. However this is not at all obvious in Figure 4. Please revise or explain why?

Figure 4: Missing the subplot labels (a), (b), ...(f). Also the colors of the lines in (c) and (d) are inconsistent and different from that in the legend.