Response to reviewers for "Application of the Statistical Oxidation Model (SOM) to Secondary Organic Aerosol Formation from Photooxidation of C_{12} Alkanes" by C. D. Cappa, X. Zhang, C.L. Loza, J.S. Craven, L.D. Yee and J.H. Seinfeld.

We thank both reviewers for their comments, which have helped to improve the paper. Our specific responses can be found below, with reviewer comments in black and our responses in red.

Response to reviewer #1:

Do the authors have any insight regarding the impact of reacted hydrocarbon levels and/or SOA mass loading on the fitted parameters? Given that the rates of oxidation will be different in the gas and condensed phases (or particle surface), as well as the resulting ΔLVP and O:C ratios, it would be instructive to understand how much the parameterizations change as conditions in the chamber approach those in the ambient atmosphere, changing the distribution of products in the gas and condensed phases.

It is difficult to answer this question without having experimental data to which the model can be fit. Ideally, the best-fit parameters would be independent of the hydrocarbon or oxidant levels because the SOM is capturing the time-dependent formation of gas and particle phase species in a general manner. The resulting model aerosol *properties* (e.g. O:C or mass) would, however, depend on the specific conditions, but the best-fit *parameters* should not. It is nonetheless instructive to investigate how the O:C depends on the amount of initial hydrocarbon for fixed model parameters. Consider, for example, the low-NOx dodecane system. In the model, after 30 hours of reaction for the base (experimental) conditions, where $[HC]_0$ = 33.5 ppb (=233 μg m^{-3}), the O:C = 0.44 (here, $P_{frag} = (O:C)^{mfrag}$ has been used). If we decrease the model $[HC]_0$ by an order of magnitude to 3.35 ppb (=23.3 μg m^{-3}), the calculated O:C increases to 0.49. Alternatively, if we increase the model $[HC]_0$ by an order of magnitude to 335 ppb (= 2330 μg m^{-3}), the calculated O:C decreases to 0.3. Thus, we see that under more atmospherically relevant conditions (e.g. $[HC]_0$ < 1000 μg m^{-3}), the calculated O:C increases, but really only exhibits a weak dependence on the initial hydrocarbon concentration, indicating relatively small (although noticeable) changes to composition.

It is suggested that the authors cite Lim and Ziemann (ES&T, 2009) in which it was demonstrated, in accord with the subject manuscript, that cyclic alkanes have higher SOA yields than linear and branched alkanes.

Lim and Ziemann are now cited: "Aerosol mass yields for the two ring-containing compounds are larger than for the non-ring containing compounds (Table 1), consistent with previous results from Lim and Ziemann (2009)."

The legend in Fig. 1/Fig. 7 is somewhat difficult to follow. It is suggested the authors consider including both the solid and dashed model lines (it does not seem like much, if any, additional space would be required).

The legend has been changed to make it clearer with both solid and dashed model lines.

In Fig. 2, ΔIVP should be ΔLVP.

This has been fixed in Fig. 2, as well as Figs. 8 and 9.

Response to Reviewer #2

The quantitative data used in Fig. 2 are already given in Table 2 (p. 27100). Figure 2 (p. 27102) is therefore useless and can be removed. Moreover, the histograms in Fig. 2 can also be found in Fig. 8 and 9.

The histograms in Fig. 2 are different than in Fig. 8 and 9. Figs 8 and 9 show the results after constraining the fitting for low and high NOx conditions, respectively, while Fig. 2 shows the overall optimized fit parameters for both low and high NOx conditions on the same graph. Although we agree that there is some redundancy in having Fig. 2 and Table 2, we also believe that there is value in showing the results graphically so that some readers are able to more easily visualize the differences that were determined between the different systems. We have therefore retained Fig. 2.

Figure 6 (p. 27106) is difficult to read. Fig. 6A includes a very large number of time profiles (more than 30). The "orange cluster" (Nc=8-12, No=3) seems to lump species having very distinct behaviors at the beginning of the simulation. Is there a good reason to lump in that "cluster" the Nc=12 species (i.e. functionalization pathway) with the Nc =8-11 (fragmentation pathways). Furthermore, the only species that can be unambiguously identified in Fig. 6B is the (Nc=12, No=2) species. It would be helpful to add more information in Fig. 6B to identify the major contributors to the SOA mass.

We have now split Figure 6 into two figures. What was formerly Figure 6a is now part of a figure (Fig. 6) in which the different clusters are separated into different panels so that they are easier to see. What was formerly Figure 6b is now part of a figure (Fig. 7) that shows "snapshots" of the relative abundance of the various SOM species, shown on a carbon/oxygen grid (as in Figure 4), to indicate more clearly which species are the major contributors to the SOA mass.

The main reason that the Nc=12 (No=3) species is clustered with the Nc=8-11 (No=3) species is that these species all seem to exhibit similar time-dependent behavior. Admittedly, this is based on visual inspection and thus is a judgment, but we believe it to be reasonable. We note, however, that it is not all Nc=8-11 species, but specifically those that contain 3 oxygen atoms. Thus, they represent a combination of functionalization and fragmentation, which is the likely reason for the similarity with the Nc=12, No=3 species.

These new figures are shown below.

P27090 (section 5.1). The Fig. 6A suggests a direct link between "cluster" and generation number. If this is correct, it would be helpful to examine and discuss this link in section 5.1.

In the absence of fragmentation, we agree that this would be true. However because fragmentation "messes up" the relationship between generation number and oxygen addition, it is difficult to establish clear links. Nonetheless, there does appear to be, at times, a link between "cluster" and number of oxygen atoms added. We have added brief discussion to section 5.1 regarding this link.

P27093, line 18-21. It is shown that a reasonable model/measurement agreement can be obtained even if one of the parameters (here cfrag) is constrained. Therefore, I expect that various combinations of the

6 adjustable parameters could provide simulation results fitting the experiments in a reasonable way (at least within the uncertainties of the experiments). Can such distinct sets of parameters be identified or, in other words, can the uncertainties be quantified for the 6 values provided in each optimized set? This might be a critical point to interpret the physical meaning of the 6 parameters obtained after the regression.

The reviewer raises an important concern regarding the uniqueness of the derived parameters in terms of how it influences the interpretation of the physical meaning of these parameters. This is discussed near the end of Section 5.2. It is our experience that the parameters that exhibit the greatest correlation – and thus likely uncertainty – are the ΔLVP and fragmentation parameters. However, it is at this point in time difficult for us to *quantitatively* assess the uncertainties in the specific derived parameter values. An additional way in which we have tested the robustness of the fit parameters is to initialize the fit with different parameters (such as much lower ΔLVP) and see what result is obtained. In general, we find that either the fit algorithm yields similar best-fit parameters as reported in the manuscript or that the fit algorithm "gets lost" and is incapable of finding a solution altogether. Although qualitative, we believe that these type of assessments provide confidence in the approximate magnitude of the best-fit parameters. For example, if we obtain ΔLVP = 2 as a best fit, then we can be confident that the value is >1.5 and likely >1.8 (or so), allowing for distinction between broad compound classes (e.g. ketones vs. alcohols). It is for this reason that we tend to focus on the broad differences in the derived parameters when making physical interpretations, as we believe these to be reasonably robust. Future efforts may attempt to utilize more advanced fitting algorithms (such as genetic algorithms) that differently sample the solution space available and may allow for more quantitative assessment of the parameter uncertainties, but this is not possible for the current study.

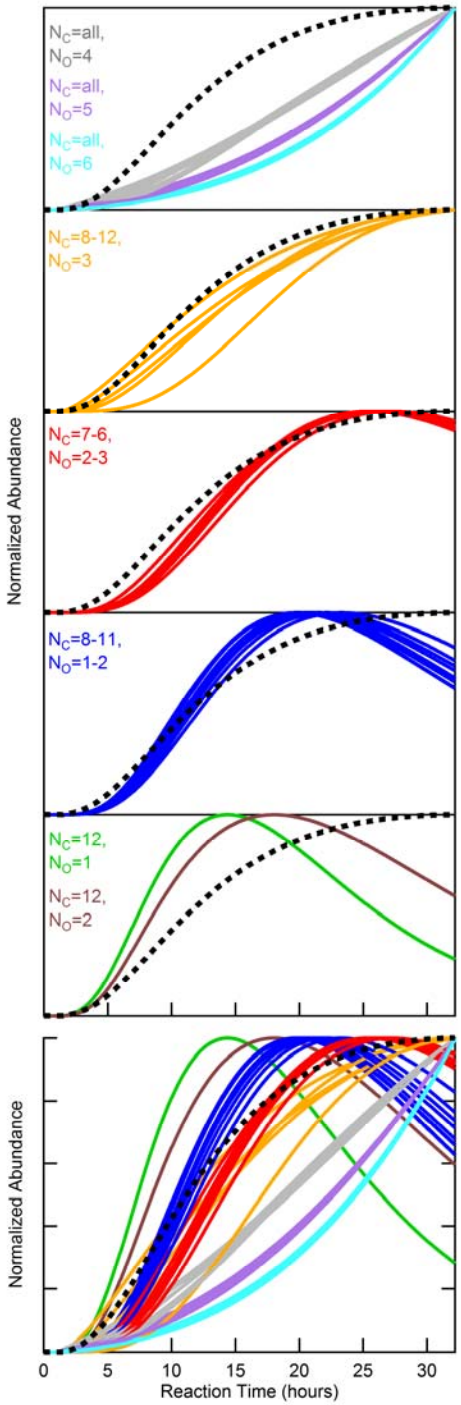

Figure 6. Simulated time-evolution of each SOM species (i.e. N_C/N_O pair) during the low-NO_x photooxidation of dodecane. Species that exhibit a similar temporal dependence have been grouped into "clusters" and are shown in individual panels with the same color; the N_C and N_O of the species that belong to each "cluster" for each panel are indicated as labels. Each species profile has been normalized to its maximum concentration. For reference, the dashed black trace in each panel shows the time-evolution of the normalized total SOA mass. The bottom panel shows all traces together so that the time-evolution of the various clusters can be compared.

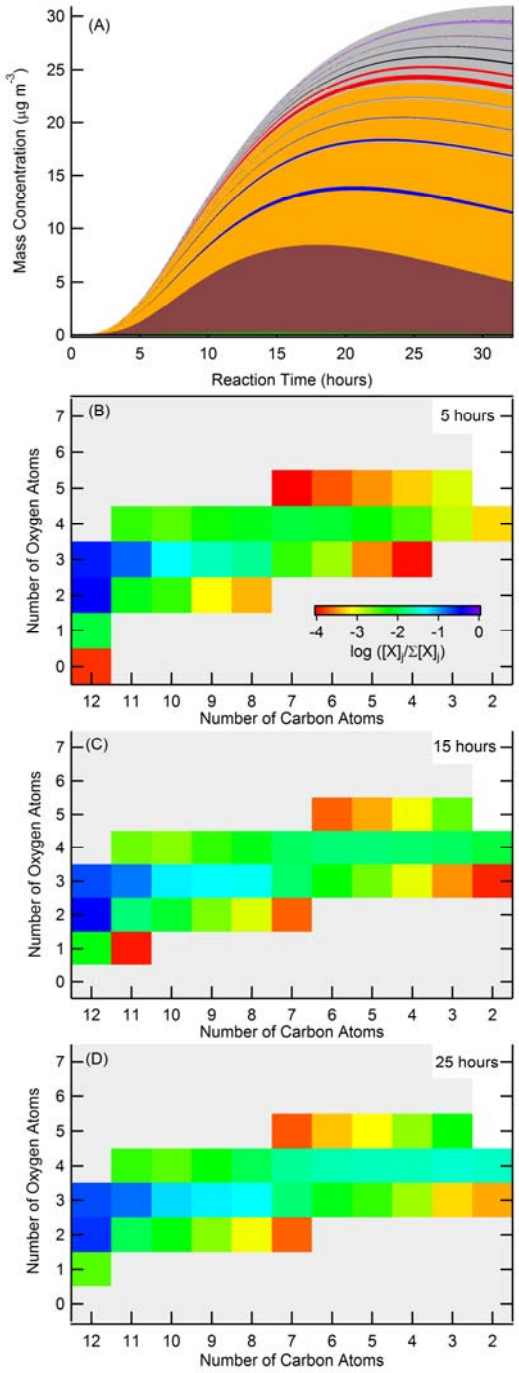

Figure 7. (A) Simulated mass concentrations of each SOM species, with each species stacked on the previous. The colors correspond to those in Figure and indicate species of a particular „cluster“. Panels (B), (C) and (D) show snapshots at 5, 15 and 25 hours of reaction, respectively, of the normalized mass concentration of all SOM species on an oxygen/carbon grid, with concentration indicated by the color of each cell in the grid (see legend). For each snapshot, the mass concentration of each species is normalized by the total SOA mass concentration as in Fig. 6; species with a normalized abundance $<10^{-4}$ are shown as gray and species with O:C > 2 are shown as white.