Major comment:
The authors have made revisions that have improved the manuscript in response to the two reviewer comments. In addition to the specific comments below, the authors still need to clarify how the lidar data is better than ground, aircraft, or sonde data. There is no discussion in the abstract of why lidar data is needed, and in some of the discussion, it appears that surface data would suffice. I would support publication after the authors improve their discussion of how the ozone vertical structure information is beneficial over other data, and make that clear in the abstract and conclusions, and in their example discussions.

Specific comments:

Page 12, line 336 – The word 'background' in the ozone literature often refers to what ozone would be without anthropogenic influences. Consider using a different word to avoid confusion. The HLO curtain has lower free tropospheric ozone than the HMO profiles but higher surface ozone. So, it is likely that surface production is the most important factor here exacerbated by entrainment.

Also, the following sentence is confusing. Did you mean to say "low-level" twice? "For example, the HLO cluster reveals the specific case in which higher O_3 is captured early in the temporal profile in the low-level and translates to the higher O_3 captured in the low-level as well."

Line 354 – Are the cluster average temperatures only at the LIDAR locations? Or across the whole domain?

Line 366 – Please remove 'slightly', if it is not a real difference don't discuss.

Line 380 – There is higher ozone at the surface in the MCO and HLO curtains in Figure 4. Why is MDA8 ozone highest in HMO? If you calculate MDA8 ozone from Figure 4, do your results agree? Or are the regulatory sites missing these elevated ozone concentrations observed by the lidar?

Figure 7 – Some of the mid-level correlations don't look statistically significant at all.

Line 439 – Are these numbers percentages? Is +0.30 actually +30%? Clarify that these are mean normalized biases.

Line 447 – You just told us that GEOS-Chem underpredicts high concentrations, but then say this challenges the assumption that models struggle to capture extreme cases. This seems to be a contradiction. It could help to quantify whether GEOS-Chem underpredicts at a certain percentage, say the 90th, or 95th etc percentile. You could also consider adding a probability distribution function to clarify your points.

Line 504 – This statement again would definitely benefit from a probability distribution plot.

Line 530 – The conclusion about multi-day events is a great one to consider including in the abstract as an example of how lidar data can help models. However, it needs an explanation about why surface ozone data would not be sufficient. The discussion seems to be mainly about the surface, so it is unclear why the vertical structure is needed here.

Line 572 – It seems that the best use for lidar would be in simulating elevated surface ozone that appears to be from transport and entrainment. I am surprised there isn't more discussion of this application.

Supplement.
Missing a '.' In the paragraph before Table S1 between 'needed' and 'These'.