Review of Hu et al. "Characterization of a real-time tracer for Isoprene Epoxydiols-derived SOA from aerosol mass spectrometer measurements."

General Comments

In this manuscript, the authors analyze AMS data from multiple sources to determine whether m/z 82 and, more specifically the C5H6O fragment, are robust AMS tracers for SOA formed from isoprene epoxydiols (IEPOX). The main goals of the manuscript are to determine the range of f_{C5H6O} in ambient IEPOX PMF factors, determine the background f_{C5H6O} in different ambient environments, and determine whether SOA generated from monoterpenes (MT) and analyzed with the AMS have significant signal at the C5H6O fragment. The authors quantify the f_{C5H6O} for areas influenced by biomass burning and urban pollution, areas with heavy MT emissions, and areas with strong isoprene emissions. They present a method for estimating the SOA mass formed from IEPOX chemistry from f_{C5H6O} and compare this method to PMF results from the SOAS campaign. They argue that IEPOX SOA mass estimated from f_{C5H6O} should be within a factor of 2 of that determined by a more rigorous PMF analysis.

This manuscript will primarily be of interest to AMS users and less so to general readers of ACP. Nevertheless, the AMS is probably widespread enough to justify publication in ACP rather than a more specialized journal (e.g., AMT). In general, the conclusions are well-supported and the analysis seems to be carefully done and robust. There are however, several things that the authors should do before publication to improve the clarity and readability of the manuscript. First, many of the key figures are essentially illegible because a large amount of data is placed in multi-panel figures that end up being too small to read. Figures 3 and 5 are particularly bad though several others would also benefit from being larger and from multiple datasets being differentiated from one another more clearly. Second, there are several places where the authors could be more clear/specific in their writing. For example, when they refer to IEPOX SOA, it often isn't clear whether they are talking about the PMF factor or the general concept of SOA formed from IEPOX. As another example it isn't clear how exactly all the average f_{C5H6O} values were calculated. There are a relatively large number of typos and grammatical mistakes and, while it was generally clear what the authors meant, it would be good if the authors gave the manuscript a more careful read before final publication. After these corrections are made, the manuscript should be publishable in ACP.

Major Specific Comments

P 11227, lines 1-2 and Page11244 , lines 19-21. Can you be more specific about how you arrive at your conclusion that the IEPOX SOA estimate from f_{C5H6O} will be accurate to within a factor of 2? Have you used your estimation method on more than the SOAS datasets to estimate the accuracy? As you mention, the SOAS data probably represents a best case scenario. Can you explain the bounds of when your f_{C5H6O} estimation method can/can't or should/shouldn't be used? Is there a lower limit on f_{C5H6O} (relative to the total organic particle mass) below which the

estimation method is no longer accurate? In general it would be a benefit to the AMS community if you can explain the limits and bounds of your estimation method more clearly.

Abstract, lines 18-19 and several other places in paper (e.g., p 11243 lines 13-16). Several times in the manuscript, the authors compare f_{C5H6O} of a bulk OA sample (for example, monoterpene SOA) to the f_{C5H6O} found for the PMF factor attributed to IEPOX SOA. They authors do this to illustrate that f_{C5H6O} is enhanced in IEPOX SOA and presumably to imply that the IEPOX f_{C5H6O} signal is enhanced relative to other potential interferences (i.e., monoterpene SOA). However, it isn't really relevant to compare the f_{C5H6O} of a PMF factor that is ~15% of the total OA to the f_{C5H6O} for the entire OA sample. To me this is misleading. From reading the manuscript, it seems the background f_{C5H6O} is a minimum of 2 per mil and up to 4 per mil for areas of high MT emissions. The f_{C5H6O} for all OA seems to be 5-6 per mil in many areas heavily influenced by isoprene emissions. So the f_{C5H6O} "signal" from IEPOX SOA relative to the background f_{C5H6O} "noise" isn't very elevated in most areas. Borneo (and perhaps the Amazon) seems to be an exception. Can the authors comment more on this issue?

Abstract lines 15-20 and through paper. Please explain how you are weighting the average f_{C5H6O} when combining data from many different studies, if at all. I can easily imagine that high frequency data from one study would completely overwhelm the average because of the larger number of points. As an example, aircraft data are recorded generally at 0.1 – 1 Hz, but ground data are typically averaged over significantly longer timescales. How do you treat this?

Related to this point, in the abstract, you list the average f_{C5H6O} for MT influenced airmasses as 3.1 per mil. The Rocky Mountain data average is 3.7 per mil, the DC3 data influenced by MT emissions average 4.1 per mil and the boreal forest data average 2.5 per mil. From these values, it seems like the f_{C5H6O} for MT background should be a little higher than 3.1. How do you calculate the 3.1 number given in the abstract?

Through paper: There are a large number of unpublished studies cited in this manuscript. 11 cited referenced are unpublished; 6 are under review (i.e., discussion manuscripts) and 5 are "in preparation". Some of the "in preparation" datasets, primarily from PMF analysis of field data, are used in the manuscript. To me this seems unusual because there has been no peer-review of this data and insufficient details are provided in the manuscript to assess the data quality. I was unable to find ACP's policy on this, so I leave it to the editor to decide if this is an issue or not. The "in preparation" data are used heavily in the figures and it is difficult to say whether the authors would have come to the same conclusions or whether their conclusions would have been as robust, if this data were to be excluded.

Page 11233, line6-8. f_{C5H6O} has a very specific meaning as does f_{82}. They are not the same. I find it highly objectionable that UMR f_{82} data are included in the f_{C5H6O} average and labeled as f_{C5H6O}. Further, it isn't clear which datasets were analyzed for f_{C5H6O} and which for f_{82}. Please either remove the f_{82} from the f_{C5H6O} average or call the average f_{82}.

Figure 3: This figure is generally illegible, with the legends particularly so. Please revise. What are the arrows pointing to on the right Y axes?

Figure 4. It looks like, if a PDF of f_{C5H6O} for monoterpene SOA were placed on this figure, it would be very similar to the PDFs of the isoprene influenced field data. Can you also include the PDF for monoterpene SOA in the figure? Doesn't this argue that there is in fact a very significant contribution of monoterpene SOA to f_{C5H6O}? It looks like the Borneo data PDF is significantly higher in f_{C5H6O} than the others field data PDFs and the monoterpene lab SOA PDF. This figure seems to suggest that the "interference" from monoterpene SOA could be worse than the authors argue. If the Borneo data are excluded from the average f_{C5H6O}, does it change significantly?

Figure 4: It is very difficult to distinguish the colors of many of the lines from one another because of the color choices and size of the figure. Please revise.

It isn't clear what the arrow pointing to the right Y axis is meant to indicate.

Figure 5: The figure is generally illegible due to size and the amount of information on the figure. The symbols are indistinguishable from one another and the legend is impossible to read. I can't make out any of the numbered points aside from 1, 2, and 13. Please revise.

It isn't clear what the pink arrow in the middle of the figure is meant to indicate.

Minor Comments and Technical Corrections

Through paper: The authors often use the term average when the text seems to indicate they really mean mode (based on a vertical line drawn to the mode in most figures). Please clarify when/if you mean average and when/if you mean mode. This is relevant because few of the PDFs appear to be normally distributed.

Abstract and through paper. It would be helpful to define the per mil symbol the first time in is introduced.

Page 11226, line 4-5. What other low NO oxidation pathways would produce IEPOX-SOA? This is alluded to several times, but never defined. Do you mean IEPOX-SOA the PMF factor or do you mean SOA produced from IEPOX? It is confusing at times to discern whether the authors are talking about SOA formed from IEPOX (a mix of some known and some unknown organics produced by a specific process) or the PMF factor attributed to IEPOX SOA (an output of PMF). This is one clear case.

Page 11226, line 9-11. Consider revising this sentence for clarity.

Page 11228, lines 22-25. Conversion of IEPOX to IOPOX-SOA requires gas-to-particle partitioning. I think you mean non-reactive partitioning here, but please clarify.

Page 11230, lines 8-9. What is the rationale for including the polluted Amazon site in the "strongly influenced by isoprene" category as opposed to "strongly influenced by urban emissions" category? To me "polluted" in this context means influenced by Manaus emission. It would be good to clarify why the data were place in one category rather than the other.

Page 11230, lines 14-15. Many of campaigns actually haven't been described in the literature and the referenced are listed as "in preparation" (see related comment in major comments section).

Page 11232, lines 26-28. Revise this sentence for clarity.

Page 11235, line 14. Add "that" between conditions and are.

Page 11240, line 4. Revise "Amazon forest down Manaus campaigns".

Page 11240, line 12. Revise "have low f_{C5H6O} are"

Figure5, page 112340, lines 14-19. I'm struggling to see how the points group into a triangle. The points don't seem to group into any shape at all. Please clarify.

Page 11240, line 16-18. Revise for clarity.

Page 11243, line 13, Revise "An alternative estimate as f_{C5H6O} from area"

Section 3.10 and Figure 8. In the preceding section (3.9), you present two alternative expressions for estimating f_{C5H6O} background for MT influenced areas. Which expression was used in Figure 8?