**Atmospheric
Chemistry
and Physics
Discussions**

Interactive comment on "Investigation of the sources and processing of organic aerosol over the Central Mexican Plateau from aircraft measurements during MILAGRO" *by* P. F. DeCarlo et al.

P. F. DeCarlo et al.

peter.decarlo@psi.ch

Received and published: 19 May 2010

Anonymous Referee 2

R2.0) Overall this is an very good paper with much new information and many excellent insights. I have multiple concerns which may be viewed in apart as a statement about just how hard it is to determine mechanistic information from a complicated data set. Please try to address the concerns (General comments 1-6) to the extent that they can be addressed without greatly expanding the scope of the study. I recommend

publication after revisions.

We thank the reviewer for the thoughtful review and detailed comments. Below our responses follow each italicized original comment. We have attempted to clarify text which may have caused confusion, and to address the reviewer's comments in detail.

General Comments

R2.1) Sec 2.2.2 CO background I don't know if there is a satisfactory way to determine CO background. 55 or 60 ppb determined for RF3 or the entire C130 data set are, I assume, free tropospheric values. In Fig. S1-2 (a) is the C130 passing into and out of the free troposphere or are the $20\mu g/m^3$ organic plumes actually in the free troposphere? For measurements near Mexico City, higher values appear more appropriate, 90 ppb as used in RF12 up to the 125-130 used in other studies. The choice of background will not matter much in a concentrated urban plume but will make a significant difference in regional sampling. Fig. 4 shows that much of the data set has CO below 200 ppb on RF 12. CO values are hard to see on the RF3 plot, but appear to be even lower. My overall concern is applying somewhat perturbed background levels to air masses that might have zero organic at a higher CO level than used here. Comments?

The CO background values given are what we determined for the unperturbed CO levels in the clean air masses entering Central Mexico from the larger region. The values were determined and published in DeCarlo et al. (2008) using the regression of OA and CO with the constraint that CO values be less than 200ppbv. As a check background CO values were recalculated in this study using only data from plume edges, and similar values were found. A modeling study of MOZART-4 (Emmons et al. ACPD 2010 Figure 6) shows model derived CO concentrations which reproduce the trend of higher CO background value in RF12 compared to RF3. A comparison of the values in RF3 show that the numbers used here are lower than model predictions. This indicates that while there may be uncertainty associated with the CO values, assigning similar background values for both flights analyzed here would be incorrect.

R2.2) Can anything be said about the contributions of urban biomass burning (trash, cooking, and certain light industries) to BBOA. Yokelson et al (ACP, 2009 or 2010) reported emissions at these "non-traditional" source with implication that they contribute a large (perhaps, unrealistically large) fraction of OA. Also, as deGouw et al have shown in several locations, CH_3CN from nominally urban source regions scales with CO (but not nearly as steeply as in open fires). Yet, BBOA on RF12 is close to zero. This null result has implications which should be mentioned.

We do not have any additional information to constrain urban biomass burning, and the contributions of these sources would most likely be contained in the "urban" fraction of the source apportionment in the postprocessing data. This represents the first attempt at direct OA source apportionment from an aircraft to our knowledge, and our impression is that determining the relative contributions of different urban sources is better carried out with ground-based data inside the urban area, as e.g. done by Stone et al. (2008) or Aiken et al. (2009, 2010). Neither paper reported a large contribution from urban biomass burning, although this source needs further research as its chemical profiles are not well characterized. We are aware of at least a couple of research groups pursuing further work in that topic for Mexico City, and we hope that this will materialize into future publications.

R2.3) Sec 3.5.1 Ascribing 30% of CO to biomass burning on RF3. This may be correct but in my opinion it is not supported appropriately given the measurements available on the C-130. Pick a tracer that is emitted primarily from urban rather than biomass burning sources. Does the ratio of that tracer to CO change between RF3 and RF12? I don't off hand know emission factors relative to CO for fires, but I would consider, benzene, toluene, and NO_y. Another consideration is that emissions could have changed between RF3 (Monday) and RF12 (Saturday).

The 30% value comes from the Crounse et al. 2009 analysis which uses Acetylene as an urban tracer, and HCN as a biomass burning tracer. The methodology of that paper uses these tracers and appropriate cross terms to ascribe urban and BB contributions

of: CO, organic aerosol, NOy, benzene, aerosol light scattering. We use the 30% figure for the CO from the results of their analysis. We also agree that days of the week may have an influence in general, however we don't have enough information to address that possible effect explicitly in this paper.

The text has been modified to read:

"Results of the CO apportionment between urban and fire sources (Crounse et al. 2009) suggest that approximately 30% of the CO in the region could be attributed to biomass burning activities during RF3."

R2.4) Sec 3.5.2 Relation between BBOA and HCN It appear from the scatter plot in Fig. 4 that almost all of the data points have HCN below 1 ppb. Considering just these points, there is a good correlation between BBOA and CO on RF3 with slope (given in text) that is similar to what others have measured. Yet on RF12, there is almost no BBOA associated with 1 ppb HCN. Graphs are very small and hard to read in ACPD and with that caveat in mind, the RF12 time series of HCN has a background of about 500 ppt but I can't see a background on RF3. RF12 does show HCN "plumes" with increases from 500 to 1000 ppt coincident with other aerosol constituents. These features need to be explained. I presume that the C-130 had a PTR-MS that measured CH_3CN. Why is that not also used in identification of biomass burning? Some comments on the relative utility of these two tracers would be useful. CH_3CN is the more commonly used tracer, though this certainly reflects instrument availability. Perhaps CH_3CN could also supply confirming evidence for the near absence of biomass burning on RF12.

A discussion of both of these points is given in Appendix 2 (alternative HCN sources) and Appendix 3 (HCN and acetonitrile correlation) in Crounse et al. (2009). Briefly HCN and CH_3CN did correlate, however there were multiple instances when differences were observed. In RF12 the increase in HCN without subsequent OA increase is due to the Tula petrochemical complex.

The following text has been added to clarify this: "During RF12, narrow plumes with elevated HCN levels were intercepted. These plumes were not associated with BBOA but rather with the Tula petrochemical complex, see Fig SI-6."

RF12 also had significantly higher HCN background than many of the other flights as mentioned in Crounse et al. (2009). A figure has been added to the supplementary material showing the correlation of HCN with CO on RF 12 with points colored by the city portion and by flight segments in the vicinity of the Tula complex.

R2.5) Sec 3.5.4 Use of sulfate to apportion LV-OOA to open biomass burning I believe that using the difference in aerosol sulfate between two flights to apportion LV-OOA to biomass burning is more of a what if exercise than a result that follows from the data. I do not see any mechanistic reason why LV-OOA (in the absence of biomass burning) should be related to SO4 with the same proportionality constant on two flights. Locations of OA precursors and SO2 are different. Flight patterns shown in S1-1 are very different for the 2 flights. Perhaps more volcanic emissions were intercepted on

See response to R1.3.

R2.6) A2 Variation in the 4 factor solution This section was an eye-opener to me. Figures A1 and A2 are valuable, though in ACPD format most details are lost in A2. The family 3 solutions look very different than the solutions used in the paper that are from family 1. By the objective standards of Q and cumulative R^2, the family 1 solution is better but I don't know whether the differences are large enough to be a deciding factor. I am guessing that if a family 3 solution was used in this study, conclusions would be different. Even though the text states "that no solution is uniquely best and the variability within a general solution group should be considered some measure of the uncertainty of the final solution," this is unsettling and raises a number of questions: R2.6.1) Is the spread in PMF solution depending on seed typical? Is it the result of sampling a wide range of chemical conditions?

This question is difficult to answer until other researchers choose to apply similar anal-

ysis methods to their datasets. It is likely that the extremely wide range of sources and aging conditions played a role in the seed variability observed in this dataset.We hope that the analysis presented here helps to motivate similar AMS-PMF analysis in the future.

R2.6.2) Are the differences in r2 and Q really big enough to say that one solution is say 90% more probably correct than another? 3. If not, how is the family 1 solution justified?

Right now we cannot provide a quantitative probability of correctness in response to question 2. Future research specifically on those topics may be able to better answer what differences in R^2 and Q are needed for what level of confidence in the choice of a PMF solution. Compared to the solution chosen for this dataset there are seed solution in solution type 1 which have higher Total R^2 (but higher Q/Q_{exp} too), and there is a possibility that they may also be "valid" solutions, at which point the question becomes how much would this change the conclusions. Taking the variation within Solution type 1 (Figure A-1) there is a few percent trading of mass between different solutions within the family, and we can expect a similar effect of a few percent on the overall conclusions of the paper. Choices of other family solutions may have more variation compared to the chosen solution, but an in-depth analysis of those solutions may have yielded different post-processing methods or values, which would be very laborious, and is out of the scope of this study. The bottom line is that the solution chosen for the dataset is the best solution based on the objective metrics of total R^2 and Q/Q_{exp}, and variations in other similar solutions do not show major differences in mass apportionment.

R2.6.3) If you feel that these questions go beyond the material that belongs in this paper, a response only in the reply to reviewer would be appreciated.

see previous responses

R2.6.4) The BBOA to HCN slopes are very different for RF3 and RF12 as noted above.

It appears that this feature is robust to the selection of solution.

This is true, and underscores difference in the biomass burning sources between RF3 and RF12. This has also been discussed in Aiken et al. (2010), in relation to BBOA and CH3CN measured at a site in downtown Mexico City.

Specific Comments

R2.7) p2456, line 1-8 Not enough explanation for me to understand what is shown in Fig S1-2 and what it implies about overestimate of volatility.

This text has been modified to better explain the calculation of the blue line in this figure, and the revised text is quoted here:

"As a side point, the plume mixing edge data in Fig. SI-2 can also be used to provide a first-order evaluation of the degree of evaporation of the regional aged OA upon dilution. Using a CO concentration of 200ppbv and the calculated background value for RF 3 (60ppbv), one can calculate the CO concentration as a function of dilution ratio for the dilution of the 200ppbv plume with 60ppbv background air. A modeled OA concentration can also be determined as a function of dilution ratio assuming a starting concentration of $10\mu g/m^3$. The blue line in figure SI-2 the graph is the modeled volatile loss of OA from dilution for a combination of state-of-the art SOA models for Mexico City, as described in Dzepina et al. (2009). Most of the datapoints lie around the straight line and above the dilution line, which suggests that the volatility of the OA predicted by current models is overestimated, consistent with results of direct volatility measurements in Mexico City (Cappa and Jimenez, 2010; Huffman et al., 2009a)."

R2.8) p24599, line 13 and elsewhere. Unit in error matrix is Hz. I guess this is correct and standard cgs usage. It just sounds like an awkward replacement for what I think is counts per second.

This is a question of terminology, with counts per second being the same as Hz. However Hz is an SI unit (and not only a cgs one), see http://en.wikipedia.org/wiki/Hertz.

C2854

We have chosen to keep the usage of "Hz" as an SI unit, and also because we think that it is more common in the broader mass spectrometry area.

R2.9) p2459 line 17-20 Hard to follow. Construction of "joint data set" explained more clearly latter on in paper.

This text has been modified for clarification to read:

"The mass spectral datasets from RF3 and 12 were combined into a single dataset, with the mass spectral matrix (mass spectra vs time) from RF12 appended as rows to the mass spectral matrix from RF3. PMF2 was run on the combined dataset to force retrieved mass spectral profiles to be identical for both flights."

R2.10) p2460 line 11-13. Good point, that is often forgotten. Many other good points that I did not get around to calling out.

We appreciate the compliment.

R2.11) page 2472, Section 3.6 and Fig. 9 I found the discussion hard to follow. A more explicit description of plots in Fig. 9 would help. i.e. "post processed urban+non-BB regional OA (line 5). What trace should I be looking at? I know that this figure will be more legible in ACP format, but as it is I can't tell the difference between Crounse Traffic and Crounse Fire

The figure has been modified so that Crounse Traffic and Crounse Fire are more distinguishable. The legend for the figure has been modified to read:

"(a-d) show the RF3 and RF12 comparison to the Crounse et al. (2009) study for the urban OA factor and the fire OA factor. The postprocessed PMF components contribution to urban and fire sources are the solid colors stacked on top of each other. Parts e and f are scatter plots comparing the postprocessed PMF urban OA vs Tracer method urban OA and the postprocessed PMF fire OA vs the Tracer method fire OA respectively. The points are colored by flight."

C2855

R2.12) page 2473, line 21 What does subscript pp mean? "photochemical processing" is mentioned a few lines above but post processing makes more sense.

PP stands for postprocessing. These abbreviations have been added at the end of section 3.5 in the revised manuscript.

"Post processing of the PMF solution for RF3 will yield 2 sources $Urban_{pp}$ and $Fire_{pp}$, with the subscript indicating that they are the source attribution from the post-processed PMF solution."

R2.13) page 2477, line 8 family types What defines a family type?

Family types were "visually" determined by the first author based on mass spectral and timeseries features, and were a first attempt at grouping the seed solutions. This visual grouping was done prior to the statistical analysis. Solution characteristics such as the CHO family content in the HOA factor, how organic mass for known large fire plumes was distributed between factors, etc were used in the determination of family type. Note that the family type does not influence the choice of the PMF solution, other than serving to organize the appendix figures.

R2.14) page 2478, line 3 total R2 Is this equal to 10? 5 regressions times 2 flights?

Yes, the following text was modfied to clarify this for the reader: "... yielding the total R^2 value for that particular seed solution (maximum value of 10 from 5 regressions for each of the 2 flights)."

R2.15) page 2478, line 24 -28. take into account slope of regression when choosing a PMF solution. There implies that you know what the PMF solution should look like. Maybe solutions that are not within physical bounds could be eliminated – or solutions that conflict with some other aspect of data set.

To date, most AMS-PMF studies have relied solely on the R^2 as a metric for choosing a PMF solution. We are suggesting that in addition to the R^2 correlation, the slope also be considered in the choice of a solution.

R2.16) Figure 10. What is the relation between points on this graph and points in Fig. 4. The number of data points in Fig. 10 is much less than in Fig. 4. Also, urban CO in Fig. 10 is much lower than total CO in Fig. 4. Adding Fire CO and background to Urban points in Fig. 10 would not change this inequality.

CO urban and CO fire come from the Crounse et al. (2009) analysis as stated in the caption. This is the CO (above background) attributed to urban sources, and the CO (above background) associated with fire sources. The number of points is different because the Crounse et al. (2009) analysis utilized data from the Whole Air Sampler (WAS), which has many fewer points per flight and uses longer averaging which smooths out the highest values. To compare with this analysis AMS PMF results were averaged to the corresponding sampling times of the WAS.