Response to Reviewer 1

We thank the reviewer for the insightful comments and suggestions, which helped to improve the quality of this article. We have addressed the reviewer's concern by making the following modifications to the paper.

General comments:

Line 122: briefly explain what the 3-cluster solution is, and link to Figure 1a & sect 3.1.

We changed this sentence to:

"We classified the air masses that arrived in Berkeley during the fire period into 3 clusters based on their origins, as shown in Figure 1a, and discussed in Section 3.1."

Lines 246-251: Smoke age is estimated for cluster 3; can it be estimated and provided for clusters 1 and 2?

We added the estimates into this paragraph, now it reads:

"The mean backward trajectory of each cluster is also shown on Fig. 1a. Plumes in cluster 1 arrived from the northeast at relatively low speed. **Their plume ages were quite variable, ranging from around 5 to 12 hours.** They mainly picked up smoke from the Atlas Fire. Plumes in cluster 2 originated from the west coast. They picked up smoke from the wildfires (mainly Sonoma County fires) and then transported it to the Bay Area. **The plume ages were estimated to be between 6 to 10 hours.** Smoke in cluster 2 is expected to be more aged than smoke in cluster 3. Plumes in cluster 3 traveled 3-5 hours from the fires to the UCB campus, as estimated from HYSPLIT. They mainly transported smoke from the Sonoma County Fires to the Bay Area."

I would note where the subscript '74' comes from in Eq 2 when introducing the equation, I had to search for that (explained several sections earlier).

We replace the 74 with *n*, and mention *n* = 74 here. Now it reads:

"The concentration timelines were first normalized to prevent all the high (or low) concentration compounds getting clustered into the same factor. Then the Euclidean distance between each pair of normalized timelines (e.g. compound concentration vectors u and v) is calculated by:

$$d_{uv} = \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + ... + (u_n - v_n)^2} \qquad (2)$$

where u_i and v_i are the normalized concentrations of compound u and v at time step i, ***n* is the number of data points (*n* = 74 in this case)** respectively."

I recommend briefly explaining what the dendrogram means in Fig 7a.

As suggested, we add this short description to it, now it reads:

Figure 7a displays the dendrogram of the factors. **The dendrogram shows the hierarchical relationship between factors. The lower the distance is, the more correlated two factors are.**

For instance, Factor 4 and Factor 5 have high correlation, but they are not well correlated with Factor 6.

Sect 3.5: it may be worth pointing out that rapid chemistry and SOA formation may occur before the time of the first set of BAAQMD measurements. Could the authors note the estimated physical distance of the monitors closest to the fires, and provide the range of physical distances in monitors?

We added this after the first sentence of this paragraph:

"The Napa measurement station is around 5 km and 10 km from the perimeters of the Atlas Fire and the Nuns Fire, respectively. Therefore, the plumes captured at the Napa station already had ages of at least 10-20 minutes and likely longer. In this period, rapid chemistry and SOA formation might have taken place. The Vallejo measurement station is 24 km from the Napa station. Measurement stations in Berkeley, San Francisco and Oakland are 50-70 km away from the Napa station. The farthest measurement stations in San Jose are more than 110 km away from the Napa station."

We also added a scale to the map of the air quality measurement stations in Figure S3 in the Supplement.

Figures/tables:

Figure 3 & 8: consider marking the different periods discussed in sect 3.2 (plume, background, BB influenced) on the figure to increase its usefulness. (I personally find time series figures that have the background of the plot shaded different colors to indicate different periods to be particularly useful. For example, see figures 1-2 of Mattila et al., 2020)

We thank the reviewer for suggesting this. There are already multiple colors on this plot. We considered this suggestion, but decided that adding shading would make the plot too crowded. Unlike Mattila et al. (2020), we classified the plume/BB influenced/background periods only based on the total quantified OA of each sample, which has already been shown by the height of each bar. We think the readers will be able to find the plumes periods. With utmost respect to the reviewer's comment, we plan to keep the figure in its original format.

Figure 5 & 9: what are the vertical dashed lines? Are these marking specific plume periods? Please specify in the figure caption. Again, marking the different periods can make these types of figures more easy to digest. (See comment for Figure 3 & 8)

We changed Figure 5 as suggested by the reviewer, as below. However, in Figure 9, we have already used shading to indicate the solar radiation. Adding another shading may overcomplicate the figure. We thus leave this figure as is. We added "The vertical reference lines mark peaks of fire plumes." to the caption of Figure 9.

Figure 5 now appears as follows:
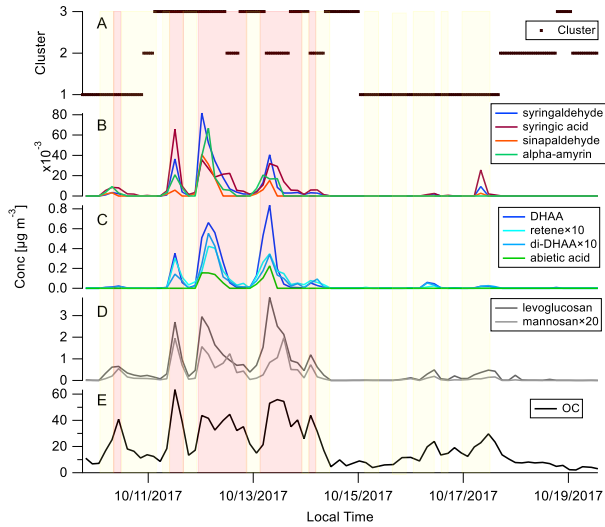
Figure 5: Time series of **(a)** backward trajectory cluster **(b)** hardwood primary BBOA markers and **(c)** softwood primary BBOA markers **(d)** general primary BBOA markers and **(e)** OC. Light red shadings show the plume periods, and light yellow shadings show the BB influence periods as described in Section 3.2.

Figure 7b: the colors between 3 and 4 and between 1 and 7 are difficult to distinguish. I recommend making all of the colors contrast more here; could also consider using different shapes to help with the contrast.

We changed those colors as suggested by the reviewer. We matched the colors in Figure 7 and 8. Now Figure 7 appears as:
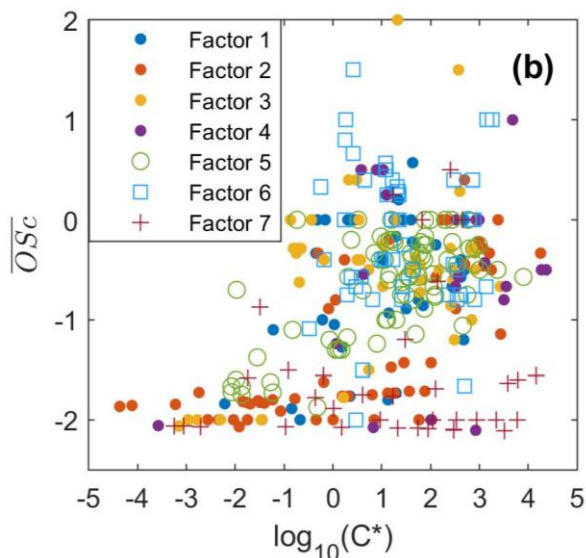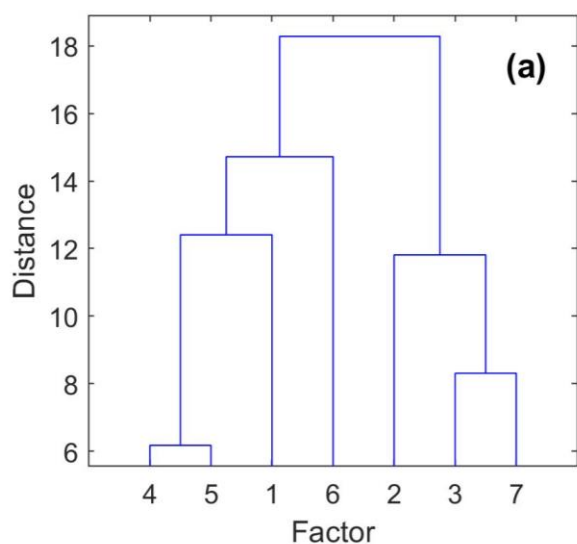
Technical comments:

There are a number of and minor grammatical errors (primarily in the form of word omissions; e.g. lines 32, 43-44, 223) throughout the paper.

We rewrite the sentences as follows:

Line 32: These compounds have high $\overline{OS}_c$, and **they** are also semivolatile.

Line 43-44: Many gas phase volatile organic compounds (VOCs**), intermediate-volatile** and semi-volatile organic compounds (I/SVOCs) are also emitted in biomass fires (Grieshop et al., 2009a, 2009b; May et al., 2013).

Line 223: The concentration timelines were first normalized to prevent all the high (or low) concentration compounds **from being** clustered into the same factor.

Line 257: 'The Oct 11 nighttime plume': this is referred to the Oct 12 plume in Fig 2. Please make sure how each plume is referred to (Oct 11, Oct 12, etc) is consistent throughout the paper.

We changed the description to "**the October 11 night to October 12 early morning plume**", in both the writing and the caption of the figure. We also changed the first sentence in section 3.5 in the same way.

Line 433-434: "The levels of PM2.5 at the sites considered here reached peak 0-5 hours after the Napa site." This sentence seems to be missing a word, such as "reached their peak"

We thank the reviewer for pointing this out. We made this mistake 3 times in the article and we have all of them fixed.