Comment on acp-2021-124
Anonymous Referee #1

The authors report 2-year near-continuous measurements of volatile organic compounds at the Maïdo Observatory using a quadrupole proton transfer reaction mass spectrometer. Timeseries during the whole period and diel profiles for i) different seasons and ii) periods when the wind originated from the East or the West are presented for 6 VOCs (CH$_3$CN, C$_5$H$_8$, Iox, C$_6$H$_6$, C$_8$H$_{10}$, and DMS) as well as for meteorological parameters like temperature, total solar radiation, and direct radiation. Positive Matrix Factorization (PMF) is used that includes 11 VOCs and CO to separate their observed mixing ratios to different pollution sources including background, primary biogenic, secondary biogenic, and anthropogenic. Finally, the authors use FLEXPART-AROME back trajectory calculations to obtain more information on the source of the PMF factors. Overall, I find this study valuable. My main recommendation is that the authors provide more details on their modeling efforts that I found hard to follow without referring to their previous publications. Also, highlighted below are some suggestions to improve the discussion of the paper.

Specific comments

I would recommend that the authors avoid unnecessary abbreviations that are also not frequently used including FT, NWP, ACs, RTs, DISP, BS, AC, and more... I think using the original wording would improve a lot the readability of the paper since I found myself many times trying to find what the abbreviations meant.

Line 108: 81 is the fragment of monoterpenes while 137 is the original monoterpenes m/z. E/N is high enough for the instrument to possibly mostly see the signal in 81 but that might be something to mention here.

Line 111-113: I am surprised to see 93 in there which I would expect to be most affected by the toluene signal which should not be long-lived and therefore not have high background. Any ideas why?

Line 117-118: You are still expecting MVK and MACR. Since this instrument cannot separate the contribution of the different compounds I would promote changing the naming to something like e.g., secondary oxidation products. This would be something to change throughout the paper.
Line 129: A graph in the SI to support this comparison would be nice to have.

Line 131: A sentence or two to briefly describe the indirect measurements would be informative here.

Line 135: Discussion of the overall measurement uncertainties based on the calibrations and humidity correction would be nice to have in a table. For Iox I would expect the calibration factor for MVK compared to ISOPOOH to be drastically different. It would be nice to add some sentences regarding the uncertainty for accurate measurements for this m/z.

Section 2.3.2: I find this section hard to read and challenging to follow the details of the approach used in this study without reading the cited publications. I feel this could be improved if the authors elaborated a bit more on the model they use and assume that readers haven’t read and don’t have to read in detail the mentioned publications. Introduction to what each model does in detail and the benefits of combining the two models would be great. Characteristic examples that I found hard to follow were line 173-175, line 183-185, line 195-197.

Line 193-194: Is this done assuming an emission rate? How is this derived?

Line 232: Are the authors capable of proving this is only secondary biogenic and not in general secondary? If not, I would recommend broadening to just secondary.

Section 3.1.1-3.1.5: In the section naming, I suggest providing the chemical formula and possible compound name as a more precise representation of what can be measured with a PTR-MS. Also, ratios of individual compounds, especially of anthropogenic nature, to CO might be informative in identifying pollution sources. Did the authors check these ratios and compare them to other emission sources? Comparison to other studies and inventories to improve the discussion of each section would be a great addition here.

Section 3.1.3: I would consider changing this name to secondary oxidation products and moving this section to 3.1.5 for the reader to have a better understanding of the primary emission trends first before discussing secondary sources. Also, this section is focusing the discussion on secondary biogenic sources alone and completely dismisses all possible sources of MVK and MACR. Although biogenic oxidation may contribute to the signal, anthropogenic sources in the island could impact the observed signal and trends. For example, I find it interesting that the trends of Iox are matching the trends of C8H10. Further discussion on this would clarify more the impact of different pollution sources on these trends.

Line 355-357: Do you mean that random hours of the day were chosen to reduce the dataset length for 3 different PMF inputs to reduce the length of the data to 1/3? If so I would recommend rephrasing especially since the reader cannot tell much by Figure 8. For Figure 8 I would also consider providing all timeseries together in one panel and zooming in to one specific plume (e.g., August-September-October) using a different subpanel to highlight the differences in the data chosen per run.

Line 387: Please elaborate a bit more on why these compounds are present in the BG factor. Is this something expected? I would think so but a discussion here would be great.

Section 3.2.2: Do the authors expect urban and industrial emissions from human activity from the island itself to play no role in the observed trends? I think a discussion on the local vs. long-range emissions detection would be valuable here. For example, when the back trajectory analysis indicates sources originating from the island's industrial or urban areas, does the anthropogenic PMF factor increase? Graphs that highlight that would be
great. Reading through the paper I see that Figure 15 already does that to some extent. Wouldn’t this therefore further supports the influence of local sources?

Section 3.2.3: How confident are the authors that the anthropogenic and biogenic emissions are fully disentangled? Based on their diurnal patterns I would expect that the PMF has a hard time separating co-emitted sources that are both expected to increase midday. Could that be the reason for the increased contribution of C_8H_{10} here?

Section 3.2.4: I think that this factor is not discussed enough and by the current discussion the naming should change to secondary oxidation rather than biogenic. Also, it would be great if the authors could discuss the contribution of other compounds in this factor including CH_3CN, C_5H_8, and MEK. Is this related to the challenges of the PMF separating different sources?

Section 3.3: I had a hard time following this section mostly because I don’t understand how this model works. It would be great if a more detailed discussion of the model was included. If limited in space this could also be thrown in the supplementary material of this paper. Correlations of this model to PMF do not look great and a discussion on the reasons why could be further investigated.

Section 3.3.1: How were the SRRs categorized in the model?

Line 526-530: The model and PMF have a weak correlation. Are these statements only based on Figure 15? What is the value of section 3.3.1 in the paper?

Figure 2, 5, 10: It would be a nice addition to generate monthly timeseries that would improve the discussions. Either the daily or monthly measurement figures could then be moved to the supplement. I think this will also improve the discussion of the figures since it is currently based on monthly trends.

Figure 2: It would be great if the seasons could be added to the figure at the top and also the abbreviations mentioned in the caption. Many readers will first look at the figures and it would be easier to follow if the abbreviations are repeated in the caption. Also, I would strongly recommend that the authors change the x-axis to months instead of days of the year since the discussion is anyways referring to months.

**Technical comments:**

Line 28: Since there are numerous publications on this matter I would recommend adding: “... (e.g., Jerrett et al., 2009).”

Line 171, 179, 180, etc.: In general there are a lot of abbreviations through the text. Many are used sparsely. I suggest avoiding abbreviations when not needed to improve the readability of the paper.

Line 242: Delete extra bracket.

Line 354: DISP and BS refer to one word each so I find no reason to generate more abbreviations here. Just use displacement and bootstrapping.

Line 367: delete double dot.

Line 377: delete “again”.

Line 381: change to “... it is more likely...”
Figure 3: I would suggest authors change the daylight background to an orange shaded area instead of lines since the figure is hard to look at with all these lines. Rather than that, I like this figure! Same comment for Figure 6.