

## Overview

Molecular composition of clouds: a comparison between samples collected at tropical (Réunion Island, France) and mid-north (puy de Dôme, France) latitudes.

Pailler et al.

## General Comments

The purpose of this paper is to describe the molecular composition of dissolved organic matter in cloud water at a novel site in Reunion Island and compare it to Puy de Dome in France using primarily FT-ICR MS analysis. The samples are also compared to previous studies and use various metrics to evaluate the composition of the samples for comparison.

Overall, I feel this is a good paper that lays good groundwork for the analysis of cloud water in remote areas that have not previously been investigated with this type of analysis. There are some things that I am interested in and things that should be addressed before full publication, but they are relatively minor and should not hinder its publication in my view.

## Specific Comments

1. Line 177: MFAssignR also incorporates H<sub>2</sub>O, CH<sub>2</sub>O, and O homologous series for formula extension.

A citation of the package on GitHub, or the manuscript itself (Schum et al. Env. Res. 2020) would be a good addition to this section as well.

2. Line 274-275: Is there an explanation for why 22/10/2019 has so few MF compared to 8/10/2019? Or maybe why 8/10/2021 has so many more than the rest of the samples? It seems like the DOC is pretty similar between them, with the main differences coming from the inorganic ions. Do you think it is related to the actual sample itself, or to the blank subtraction method? Conservative blank subtraction is a good choice, but I am curious what the formula numbers looked like prior to blank subtraction and whether they were more similar at that point.

3. Lines 303-304: You mention that the average OSC is similar between PUY and REU autumn samples, while this can definitely just be a coincidence (considering the different sources and conditions) I was curious if you looked into the molecular formulas to see what sort of differences occurred in them. For example is the OSC heavily influenced in both cases by a common set of molecular formulas (even if they are different molecules) or are there really no similarities at all, they just happen to average out to the same OSC?

4. Lines 362-364: If I am understanding correctly, the general percentage of formulas in each classification is similar between REU and PUY, which seems reasonable, I am still curious about the specific differences between the molecules in one sample or another in a more comprehensive view. Do the formulas in each classification match each other between the different sites or are they largely different? For example, for the LipidC classification, are the formulas found at REU and PUY 90% common, 70%, 50%, less? I think it could be interesting to see if the detailed composition of these samples is very different or the same, since it may say something about the cloud processing results. The "averages" are very useful, but as you have mentioned, even the same formula doesn't necessarily mean

the same molecule, so if a set of molecular formulas are in a particular classification, they may not be similar in any other way, or they could be very similar and highlight that cloud processing brings organic matter to a similar specific result.

5. Lines 370: While the FT-ICR is very well suited and effective for this work, the lack of structural information is a shortcoming as noted here, is there any interest in doing LC or fragmentation analysis in the future for these samples or others?

6. Lines 378-384: You are taking appropriate caution in classifying these molecules as one specific class or another with the database, but I was curious whether if you took a few of the formulas that you have classified as “prenol lipids” for example and just looked for any molecule matching that formula (in other databases or the search engine of your choice) if you could get any other classification?

7. Lines 426-428: I do not quite understand this sentence. Are the measured concentrations for alpha pinene 0.5, 71.5, and 2 for R8, R9, and R10B, while the beta pinene concentrations were 39.9 and 1.3 for R8, R9, and R10B, or are the detection limits for alpha pinene 39.9 and for beta pinene they are 1.3? I think the sentence could be restructured for clarity.

8. Lines 457: Does this mean that the organosulfate intensity was low in all samples (REU and PUY) with the exception of PUY 8/10/2021, or are you just comparing PUY 08/10/2021 to other PUY samples?

Additionally, you explain the higher occurrence of limonene organosulfates at REU by the increased emission of limonene at the site, which makes sense, but does that imply that the organosulfate formation from limonene is a faster process than the oxidation of pinene? My understanding of the reason given for the relative lack of pinene oxidation products is that the emissions were too fresh to have oxidized yet. Is the organosulfate a primary oxidation product like C<sub>8</sub>H<sub>12</sub>O<sub>5</sub>? Or is the explanation that there is more limonene emissions relative to the pinenes?

9. Lines 465: What were the N and S beta caryophallene formulas? Is there any way to know that the formulas are N or S caryophyllene molecules other than matching the formulas? While presence of their emission sources on the coast may explain the N and S beta caryophyllene, why would there be no CHO oxidation products? Are the N and S reactions that much more favorable than the O oxidation? Or is the concentration of N and S so overwhelming that the O oxidation doesn't really occur, relative to N and S?

10. Line 524: According to the classification you say that 50% of the molecules observed are reduced, is the explanation that the organic matter in the clouds is fairly fresh and hasn't had a chance to oxidize more completely yet?

### Technical Corrections

11. Line 25: Somewhat contradictory statements, can consider changing the language a bit to get to the assumed intended meaning.

12. Line 179: It may be more consistent and precise to say “same mass” instead of “same peak”, since the parenthetical on line 180 says “unique mass”.

13. Lines 323: Should probably change “is” to “are”

14. Lines 520: Instead of “emitted” you should probably say something like “developed” or “produced”. Overall the language in this manuscript is very good, but there are few minor things, like this and the comments for lines 426-428 that could be adjusted.