

## **Response to Anonymus Referee #1 for review on “Software and database structure to analyze the relationship between aerosol, clouds and precipitation: SAMAC” by Gagné et al.**

The authors would like to thank Anonymous Referee #1 for his or her comments on the significance of the work. We indeed had failed to convey what we had in mind for this piece of software and we believe the improvements made following these comments will help reach the relevant audience for this manuscript and better convince researchers to use SAMAC to compare their results with others'.

### Overview

*This study introduces a software program that researchers can use for the analysis of aerosol-cloud-precipitation data from airborne field campaigns. The topic of aerosol-cloud-precipitation interactions is a very important one and examining the data consistently and accurately across campaigns and research groups can be improved. I applaud the authors for trying to provide a framework for researchers to assist with their potential needs to do quick analysis in the field. The paper is written well, the title and abstract are appropriate, and the figures and tables are mostly clear. The major issue for this reviewer is that the scientific significance of the work needs to increase before publication can be recommended.*

The authors understand the concerns of this Reviewer and the comments above highlighted the lack of explanations regarding our intentions for the software. We emphasized these points in the new version of the text and we answer the Reviewer's specific concerns below.

### Specific comments:

*This reviewer prefers to be more convinced by the authors that there is great value in a tool like this in light of the fact that there are existing softwares and methods that are commonly and easily used already in the field.*

This is true and it was not our intention to replace existing software, methods and functions (or algorithms) with our own. In fact, SAMAC only provides basic analysis functions. What we aim to provide with SAMAC is a platform where all cloud-related data can be stored in one object, and which follows a flexible structure common to all clouds. Rather, we would like to have old and new algorithms added to SAMAC by and for the research community.

The idea behind this is that different research groups could put their data into the SAMAC structure and then analyze the data using SAMAC software. They could contribute their algorithms and use other research groups' algorithms to compare their different functions. They could also, of course, read and use other people's algorithms. This feature would allow researchers to save time, if they choose to use an existing algorithm, or to reduce chances of having error in their analysis code (by reviewing another group's code, or having their own code reviewed by others).

To clarify this, we added the following sentence in the abstract:

“Other researchers can readily use already submitted algorithms once their data is placed in the cloud structure provided, and they can contribute their own algorithms to the software for others to see and use.”

A paragraph was also added in the introduction:

“The creation and use of an open source standardized database structure and software would allow

researchers from different institutions to compare their measurements with those of others more easily. Moreover, a basic quantity, such as the concentration of below-cloud aerosols, could be calculated using exactly the same technique (algorithm), making the quantity more comparable across various clouds, campaigns, and in the literature. Such software could also improve analysis speed by producing basic plots and calculations so that the analyst can visualize multiple aspects of the measured clouds rapidly and then decide on the next analysis steps to take. Sharing algorithms between researchers also saves coding time to all participants in addition to providing a good basis for comparison.”

*I question how useful this tool can be for scientists who want to get deep into the analysis and produce publication-worthy figures and tables. Also, taking a generic approach to analyse aerosol-cloud-precipitation data arguably can also have the side-effect of stunting creativity in how people use data.*

As mentioned above, the functions provided with SAMAC are basic functions and it does not replace the work of scientists, quite the opposite. Scientists can use SAMAC to calculate some more standard or basic features either for one or for multiple clouds, and then use their own analysis tools to look into the deeper meaning of their data. Also, since users are encouraged to submit their own algorithms to SAMAC, we do not think they would feel limited to use only the functionalities included with the software.

Although users can add their code to SAMAC at any point in the publishing cycle, it is important to note that users are not obliged to submit their original algorithms to the public version of SAMAC, even if they have added it to their own version. Alternatively, a scientist could also publish his or her work first and submit their algorithms after publication. The licence governing SAMAC allows for maximum flexibility.

A new paragraph was added at the end of section 3:

“SAMAC falls under the protection of the GNU General Public License version 3. This license allows scientists to use the code in SAMAC, contribute new algorithms to the public version, but also allows them to add new algorithms to their own version without ever making that piece of code public, or they can also add it to SAMAC later, after their work has been published. It is the authors' view that the GNU General Public License version 3 allows for a maximum flexibility.”

*Field scientists have software packages and routines of their own that are efficient in terms of having the basic plots ready to go and also the ability to conduct more detailed analyses within the same framework, such as with IGOR software (amongst others probably) which is often used by aerosol researchers.*

That is also true, and we expect researchers to keep using whatever tools they have developed. The idea behind SAMAC -and we recognize we had not made that very clear in the AMTD version of this manuscript- is that researchers can compare their work rapidly with other groups' work. By putting their data into the SAMAC structure, they can have their data processed exactly the same way as another group's data. This would significantly improve reproducibility, and help solve questions such as “How did you get that kind of results, my data shows something completely different?”. We believe that the above-mentioned paragraph added to the introduction helps to clarify this point.

We have chosen to write SAMAC in a free, open-source and efficient programming language: Python. Proprietary software or programming languages such as IGOR require a licence that is often not transferable between institutions. Moreover, Python is a flexible programming language which is an advantage if users would like to contribute more advanced analysis software. Python developers are

also very active in scientific fields and keep adding functions to Python's scientific libraries such as Scipy/Numpy, Pandas, etc. This is detailed in section 3.

*I encourage the authors to do an improved job of convincing seasoned field data users as to why this software should be used and how this can benefit them beyond using the software packages that people have been using for years in the field quite efficiently already.*

The authors hope these new explanations have helped convince the Reviewer and the readers that SAMAC adds value to the already existing software. Again, we would like to emphasize that SAMAC is rather meant as a platform from which scientists can share, as a complement to the scientific literature.

We also addressed this problem with a modified paragraph in section 3:

“Because the software was designed to handle clouds flexibly (with different instrument types and number of instruments, different aircraft trajectories and series of manoeuvres, etc.), SAMAC is ideal for use and modification by other research groups to compare algorithms. Once the data is placed inside the cloud instance, the methods and functions can be applied right away. The use of SAMAC by other research groups should serve to increase comparability, test each others' algorithms on one's own data, and speed up data analysis by using other researchers' algorithms without a need to adapt the code. Moreover, shared open-source software makes data analysis more transparent and possible mistakes are more likely to be spotted and corrected (Challet & Le Du, 2003; Barnes, 2010; Merali, 2010).”

*The authors could also provide more motivation text about what target audiences would benefit most from this type of software since this reviewer's views are based on that of someone who has had extensive field experience; perhaps, the more relevant target audience are more inexperienced students?*

We believe SAMAC could be useful for anyone wanting to collaborate with other groups. It can also be used as an exploratory tool for more inexperienced students. We believe the added text in the abstract and introduction makes it more evident to the audience.

*To increase the value of this manuscript and software package, the issues above should be addressed which can help considerably.*

The authors tried to integrate these points into the manuscript. We believe these changes are likely to increase to user base who can also contribute to further developments of SAMAC. We would like to thank Anonymous Referee #1 for these comments.

#### Minor comments:

*Figure 6: y-axis units are wrong. LWP is generally  $g\ m^{-2}$ .*

Thank you. The correct units are now displayed in the figure as well as in the provided code.

*Page 3648, Line 22: “fulfil” spelled wrong*

This spelling is correct and recommended by the spelling software we used. It appears that two different spellings exist: <http://www.thefreedictionary.com/fulfil>.