Response to Referee Suzanne Crumeyrolle for review on "Software and database structure to analyze the relationship between aerosol, clouds and precipitation: SAMAC" by Gagne et al.

Overview

This paper is describing a new software program for aerosol-cloud-precipitation data analysis. This software can be use on the field to check the quality of the data or later on to produce high quality figure for publications. This software was developed with a free and open source programming language, which makes it even more valuable. As the harmonization of airborne data is challenging, the use of this software by the entire community would be a first step.
The paper is well written and describe the software in an efficient way.

The authors would like to thank Referee Crumeyrolle for carefully testing the software. We realize this is an important investment of time and we appreciate her feedback on her user experience.

The structure of the software has been modified to facilitate collaboration and contributions by members of the scientific community. These modifications had bearing on some of the points raised by this Referee. We answer to the specific comments below.

Specific major comments:

This software does not include the different corrections usually applied to the aerosol/cloud/precipitation data. Indeed, most of the aerosol measurements are corrected to take into account different artefacts like coincidences, truncation angles etc. . . These corrections are most of the time well described in the literature (Bond et al., 1998; Anderson and Ogren, 1998 ; Virkkula et al. (2010) ; Wiedensohler et al. (2012), etc. . .) and most of the times are similar from one data set to another. However, slight differences may exist and/or coding errors remain possible. Thus, this method should involve three steps : (1) integration of the raw data, (2) correction and (3) analyse of these corrected data. Of course, the codes to correct the data might beneficiate of the collaborative work from the entire community but it would be nice to, at least, mention it in the manuscript.

SAMAC indeed does not include any data pre-processing algorithms, it focuses on the cleaning of processed data and their analysis. There is already an effort to compare different pre-processing algorithms, namely through Workshops on Data Analysis of Cloud Microphysical Measurements which will hopefully ensure that optimal algorithms are used with a minimum of coding errors.

The authors believe that the different data processing functions are beyond the scope of this work. Of course, SAMAC being open-source and free, it could serve as a basis for, or could integrate such pre-processing software if members of the community would be willing to work on it.

We added a paragraph at the end of section 4, citing the references given by Referee Crumeyrolle:
"SAMAC does not include any data pre-processing algorithms i.e. instrument-specific corrections applied to raw data (e.g. Bond et al., 1998; Anderson and Ogren, 1998; Virkkula et al., 2010 and Wiedensohler et al., 2012). There already exists such pre-processing software for airborne measurements, for example, EGADS (EUFAR General Airborne Data-processing Software, https://code.google.com/p/eufar-egads/). There are already efforts to harmonize these processes, notably through the Workshops on Data Analysis of Cloud Microphysical Measurements. The implementation of such algorithms to SAMAC is out of the scope of this paper, but users are very

welcome to add them to SAMAC if they so wish."

It is well known that the aerosol-cloud interactions highly depend on the cloud characteristics and on the atmospheric dynamics at cloud boundaries. In SAMAC, the type (cumulus, strato-cumulus. . .) of cloud are described but how do you take into account the cloud macroscopic properties ?

SAMAC only provides basic functionalities and was designed to accommodate more. A user can add a description field using *c.describe* and choosing the "add" option. It could also be added to the describe method for all users if a field was used commonly enough. We mention the possibility of adding fields in section 5.2:
"We used the method *describe* to enter general information on the cloud using existing keys and adding new ones."

Specific minor comments:

P3653, Line 7: what do you mean when you said 'to compare large amounts of clouds with different characteristics'.

That phrase was indeed not very clear. We re-wrote that sentence.
"As discussed earlier, in order to use SAMAC as a comparison platform for different clouds with wide-ranging characteristics, we need a standard but flexible data structure that can accommodate many situations."

P3655, Line 12: 'The complete format description and an example of software guiding . . .'. The format of the pdf file (AllDocumentation.pdf), which can be download with the software, make it difficult to read and to follow. It needs to be improved. Moreover, the wiki pages are helpful and a link is needed before the conclusions.

Following the restructuring of the software, the "AllDocumentation.pdf" file has been removed, and we rather link to the wiki pages which are indeed much better presented and easier to read and navigate. The link was added in section 3: "Documentation material is available on the wiki pages (book icon, https://github.com/StephGagne/SAMAC/wiki). It includes a description of the SAMAC cloud object structure, a description of the methods and functions associated with cloud objects, and a guide on how to create and populate a cloud object. It also includes an example cloud object with which to explore SAMAC." and we refer to the wiki pages throughout the text.

P3657, Line 28: The 'average_value' function : It is not clear what this function is doing. I tried this function and I compared it with results from excel and MATLAB (which are similar), It didn't match.

Average_value was not a function but a variable into which the answer of the operation on the right of the equation would be stored. The function, numpy.mean(), is from the Python Numpy library and is meant to calculate the mean of the vector inside the parentheses. The authors are not sure where the inconsistency occurred. The authors tried using the "average" function in a spreadsheet and got the same answer both in Python and in the spreadsheet. If the problem persists, more details will be needed to be able to find the reason for this discrepancy.

Line 22: The function 'MyFirstCloud.plotavsd' is not working, at least not with the SOLAS cloud object. According to the figure 1, I specify the instrument name as a FSSP124 but it didn't work.

Yes, this is caused by the fact that the cloud used for producing figures and the cloud provided as an example with the software are not the same clouds, as Referee Crumeyrolle also mentioned in a comment below. The instrument names in the provided example cloud are *RainMeasurer* , *AerosolMeasurer* and *CDMeasurer*.

Figure 4a: The color code that appears on the bottom figure4a corresponds to the different periods (above cloud, below cloud etc.). A color scale may be added to clarify this as well as a color scale for the time of flight on the top figure4a.

The authors agree with the referee. This figure, however, is already quite full and the authors have not yet thought of a satisfying way to add these features without overcrowding the figure. It might be noted that the colour coding of the different legs should have already been explained in the help string of the function. This was corrected and can be accessed by typing *help(samac.overview)*.

Figure 4b: The downloaded data are not corresponding to those plotted in the manuscript. Indeed, the flight trajectory is located near the French coasts.

This is true. The authors used real measurements for the figures in the paper because we think they are more interesting for the readers: the values are closer to real measurements and the features are real. The data provided in the example cloud is not from real measurements because we could not provide these data due to ownership issues. The example cloud has simpler data, takes less memory, and presents a simple dataset to familiarize users with SAMAC.

We added a comment in section 3:
“It also includes an example cloud object with which to explore SAMAC. This example cloud, despite similarities, is not the same cloud that is presented in this manuscript.”