Atmos. Meas. Tech. Discuss.,
doi:10.5194/amt-2020-368-AC2, 2020

Atmospheric Measurement Techniques Discussions

Open Access

EGU

Interactive comment on "Reducing cloud contamination in AOD measurements" *by* Verena Schenzinger and Axel Kreuter

Verena Schenzinger and Axel Kreuter

verena.schenzinger@i-med.ac.at

Received and published: 16 December 2020

General comments 1. The introduction needs to be improved, with more detailed description of the GAW-PFR network and related references. Other cloud flagging algorithms could be mentioned, as well as the main principles of them (temporal AOD variations, spectral changes/Angstrom exponent, etc.). The AOD spectral derivatives and curvature used later, suggest including reference to O'Neill papers on this matter. Overall, references are missing to support certain statements throughout the text (see specific comments).

We will add references as suggested by Referee 1 regarding other cloud flagging algorithms.

Printer-friendly version

Discussion paper

2. The "multiplet" method is not at all described (lines 60-65). First, the authors should provide how many measurements in what time interval are analyzed. And they should also provide the threshold value for cloud flagging. Second, the Smirnov (2000) algorithm is much more complex than the "triplet" criterion. Several other steps are given(on the second temporal AOD derivative, on the AOD standard deviation for the entire day, etc.). The description provided by the authors is too simplified. You need to be much more specific and detailed.

We will go into more detail about Smirnov (2000) algorithm. However, the details are best looked up in the original paper. While there are indeed more criteria than the multiplet, the latter is responsible for most removed data at our site. We will add the exact values on how many points are flagged by the different criteria.

3. The applicability of the method is highly questionable, for two reasons. First, the threshold value used to flag the data (0.012 for d20) seems to be based on the analysis of few days only. This is not acceptable in my view. Many circumstances, simply other AOD level or other aerosol type, can change the variability of the AOD or the Angstrom exponent. This is not considered at all in the discussion. I would recommend analyzing the four variables and the d20 distance in different situations (high and low AOD, different aerosol types, season, etc.) to derive a more robust criterion.

This is a justified concern which we share, but which is at the very root of the problem: to find the best threshold in a continuous distribution of d20 values (not a binary one clearly distinguishing between cloudy and clear) such that as many cloudy and as little cloud free values as possible are flagged. We used a number of clear sky days (about 150) as a starting point for a reference value and indeed fine tuned / checked according to benchmark days for different situations (high and low AOD,

Printer-friendly version

Discussion paper

different aerosol types, season, etc.), see figures 3 and A1. Even a human observer will have to decide at which point in the transition he deems the sun obstructed when a cloud is moving. Furthermore, the real positives/negatives, especially in this grey zone, are unknown – this is the reason this kind of algorithm is needed in the first place. While we would also prefer a more objective and deterministic way of choosing the threshold, we believe that in principle there is no real alternative here to improve the threshold determination.

Second, this method is restricted to one site. It performs apparently well in Innsbruck, but the authors should not pretend that it can be expanded as is to other sites (or instruments). We have no idea how it would perform in an Amazonian site with highly variable biomass burning aerosol, for instance. Claiming that the method can improve the Smirnov (2000) algorithm (line 145) in cloud flagging is too much to say. Suggesting that it can work as good as Giles (2018) with less input information (line 146), is also too much to say. Those algorithms have been tested over a huge database and thoroughly describe the difficulties and the compromise that needs to be taken in an operational algorithm.

There is no reason why the algorithm should not perform at sites with similar aerosol conditions (Europe/North America). For more dissimilar sites it depends whether the density of points in clear conditions is dissimilar enough from cloudy conditions. The volcano example is the closest we'd have to such conditions and does not suggest that the algorithm cannot be applied here. We do not claim to improve Smirnov, but employ a completely different algorithm to solve an open issue in the time-series analysis, which is detection of thin clouds. Aureole scan provides admittedly more information, but is not available everywhere, and takes a lot of time. No comparison with the performance of Giles (2018) can be done, and we did not intend to suggest a better performance. Nevertheless, it was indeed only applied to Innsbruck data, so we will formulate the conclusions regarding the generalizability more carefully.

4. The validation of the method needs independent data, such as all-sky camera, lidar, ceilometer, solar radiation or even a human observer. The all-sky camera is mentioned in the paper but nothing is shown and no systematic analysis seems to be used for analyzing the performance of the clustering routine. The fact that more data points are removed is not sufficient. This can be easily achieved by any algorithm by using stricter thresholds.

The problem with all the suggested independent data is that the frequency of measurements is much lower, and they do not necessarily measure in the same path as the sun photometer. Furthermore, there are uncertainties in distinguishing cloud-contaminated from cloud-free datapoints as well, so it cannot serve as reference in a systematic analysis. However, clouds have clear spectral signature (as seen in the Angstrom parameter plots) which can be determined with sun photometer data. We will add sky camera pictures for the example in Figure 2.
Yes, stricter thresholds will lead to removal of more points, shifting the mistakes from false negatives to false positives, but our algorithm does not merely remove more, but also different datapoints, as seen in Figure 4. Again, while some datapoints can be identified as clear/cloudy without a doubt, there are many in the grey transition zone of a cloud forming or passing in the path. Where the cut is set exactly in the continuous transition depends on the kind of error one prefers to avoid.

Specific comments
Line 22: good place to add more information and references about GAW network (Wehrli, Kazadzis), and maybe Skynet cloud-screening (Takamura).
We will add this accordingly.

Line 23: do you think your algorithm could be used for lunar measurements too? This is emerging technique and worth mentioning, since your algorithm does not require

Printer-friendly version

Discussion paper

any additional measurement, only AOD.
That is an interesting point. There is no obvious reason it should not work. However, it sounds like even more of a stretch than just applying it to different measurement sites.

Line 25: what happens if there are cirrus or contrails during most of the day? What happens if there is a change of aerosol type during the day?
We will add some detail.

Line 51: I guess you use all 4 wavelengths, but please specify.
We will change the text accordingly.

Line 52: this a good place to cite O'Neill papers on AOD spectral derivatives.
The references will be added.

Line 61: as mentioned in the general comments, Smirnov's paper includes a long list of criteria. Please be more specific.
Yes, see also general comment 2

Line 63: please provide the time interval for the multiplet and total number of measurements. The PFR raw acquisition time is 2 seconds if I remember correctly. Do you use 1 minute averages, standard deviations, how do you look at fluctuations over de 5 minutes period? Please provide details, including the threshold used for cloud flagging.
Yes, see also general comment 2
We do not use a minute average, but value at acquisition time with a frequency of 1 minute. We will add this.

Line 74: the division by 10 to make the parameters comparable in magnitude is somewhat arbitrary. This is fine if it works, but couldn't you try some kind of mathematical normalization?
There are two common normalization strategies:
1) By mean and standard deviation
Given that AOD (nor the other parameters) is not by any means normally distributed,

this is hardly applicable. There is not even consensus whether to use arithmetic or geometric means when it comes to averaging AOD over 1 day/month. . .
2) Mapping onto a [0,1] interval
This would kind of defeat the purpose of finding outliers with our method as it changes the density of the data. Of course it can be done, when using it on the whole time-series – ultimately this means multiplying by a fixed factor (to have intervals of the same length, it is not necessary to shift the lowest value to 0 as only distances are analyzed), which is what we already are doing. This can be done, but should not significantly alter the outcome.
The division by 10 can be seen as a weighting function rather than a normalization of the different dimensions.

Line 74-75: the sentence about the finite difference units is hard to understand. Please rewrite.
We will change the text accordingly.

Line 80: do you mean high AOD or high number of data?
High number of data, we will clarify the text.

Lines 81-84: the procedure to derive the d20 threshold seems arbitrary ("...is further fine tuned using days on which the Multiplet routine fails...") and is not explained at all. The paper must describe this method in detail so that others can reproduce it. Moreover, as explained in the general comment 3, few clear days are in my view not sufficient to derive the threshold. A robust statistics over a large sample of data would be desirable. The description should also reveal the difficulties (too strict threshold removes too many good data, too loose threshold allows too many cloud-contaminated data...and so on). And such analysis requires ancillary information to assess what data points are clear and which are cloud contaminated.
See general comment 2 and 4. We will add some comment on false positives/negatives.

Line 94: the airmass limitation was not mentioned until this point. It is one of the criteria used in Smirnov's algorithm. As explained above, you need to be more specific and detailed in describing the algorithms.
See general comment 2

Figures 2 and 3: please enlarge fonts in the axe's labels.
We will change the figures accordingly.

Figure 3 caption: better say "solid line" than "black line" (the lines are actually in white-color)
We will change the text accordingly.

Line 110: maybe add reference about Saharan dust over Austria
Reference will be added.

Line 114: add reference about Eija volcanic eruption over Austria.
Reference will be added.

Line 118: what criteria or ancillary data are used for a manual screening? Note that readers need to know how to reproduce your results.
As it is inspected by a human, the criteria cannot be put down as specific numbers.

Line 121: maybe add reference about sunshine hours
Yes, we will add a reference. It is only a plausibility argument, though.

Line 127: wrong subscript, the longest wavelength is 870nm, I guess.
Yes, thank you. We will correct the error.

Line 137: 0.001 is the uncertainty is for the reference instruments if Langleys are made at Mauna Loa (Toledano 2018; Kazadzis 2018). For a field instrument (side-to-sidecalibration), the uncertainty is rather 0.5-1% in calibration constant or 0.005-0.01 in AOD.
We will change the text accordingly

Line 145: the data shown in the paper do not support this strong assessment in the conclusions.
See general comment 3

Line 162: the sentence about the results in high latitude sites is speculative and not supported by any data in the manuscript.
See general comment 3 We will reformulate.

Line 165: please rewrite the sentence, it's hard to understand.
Ok, we will rephrase.

Line 170: "association": do you mean correlation, covariance, ..?
Mutual conditional information (details in reference)

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2020-368, 2020.