Atmospheric Measurement Techniques Discussions

Atmos. Meas. Tech. Discuss., referee comment RC2
https://doi.org/10.5194/amt-2022-217-RC2, 2022

Comment on amt-2022-217

Anonymous Referee #2

Referee comment on "A cloud screening algorithm for ground-based sun photometry using all-sky images and deep transfer learning" by Eric A. Wendt et al., Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2022-217-RC2, 2022

Review of "A cloud screening algorithm for ground-based sun photometry using all-sky images and deep transfer learning", by Eric A. Wendt et al.

The manuscript introduces a new machine learning approach for cloud screening in sunphotometer measurements using an additional low-priced allsky camera. It describes in some detail the setup of camera system and machine learning approach. A training data set from three different camera systems is presented including a validation using parts of the data not used for training.

Major points of criticism:

- In my opinion there is a lack of motivation for the introduction of this system. It might be low cost, but still there is a need for additional instrumentation while cloud screening in sun photometer data is usually done using the sunphotometer data itself (by spatial of temporal variation tests, e.g., for the AERONET network).
- The manuscript only introduces a limited validation of the method and no comparison to established methods.
- The assumed better instrument independence of your approach, compared to standard methods, is at least questionable unless you show clear evidence. You are using compressed camera images of variable quality and finally state that this has clear effect yourself.
- At the same time, I doubt that the remaining presentation of the setup of a low-cost camera system from standard parts and the adjustment of an existing machine learning technique for general imagery (VGG-16, University of Oxford) to the allsky image cloud detection task is sufficient to justify a scientific publication in an atmospheric science publication like AMT.

In the present form I recommend the rejection of the manuscript. Resubmission after extension of the validation and comparison to other methods could be interesting.

More specific:

- Reading the first part of the introduction I already ask myself why you do not announce to compare your new method to a standard one based on the sunphotometer itself. An improvement could justify your approach.
- In line 47 you mention the "instrument-specific nature" of existing techniques, but in the following you will show that your method is very much instrument-specific itself. Compared to its dependence on camera system, the mentioned existing method is AOD based and thus should be – by design – instrument independent.
- From line 56-104 you describe a long list of existing machine learning approaches to analyze sky images for cloud classification and detection. I'm missing reasons why the community would consequently need your new technique. The reason that no other has been used in the context of cloud detection for sunphotometer purposes does not seem enough.
- Lines 124ff: You list the three camera systems' imagery you will use without stating the image format provided. For the WSISEG I could check that it is PNG format. Already this casts doubts on the instrument independence of your method as quite some processing and different types of compression happened to the data.
- Lines 156 ff: The lengthy description of the partial automation of the preparation of the "truth" training data is confusing. You should in the beginning of this paragraph what is "manual" and what "automation".
- Table1: What is a "sample"? One image, isn't it? This all sounds like a small data set. More problematic – a small data set with just 300 samples makes it impossible to compare to other methods validated in other specific or less specific situations.
- Table 2: Quite a part of the number in here does not seem "good". The problem is that you never stated what "sufficient" or "good" would be. And how limited other techniques are.
- Line 263: The statement "performed well relative to prior AOD screening algorithms" sounds very soft and is not corroborated by any shown data. Neither to your own "prior algorithms" nor to standard methods of the community (AERONET). The word "well" without any supporting data is used more often further down.
- In your limitations section you honestly state important points, but the manuscript does not provide the necessary cure or discussion. You are depending on image processing steps not within your control, i.e., camera configuration (white balance, contrast and color enhancement, compression, …) which makes your method instrument specific! And the selection of your small validation data set (e.g. without situations of high aerosol load) makes your scores hard to compare to other methods' results.