Atmos. Chem. Phys. Discuss.,
https://doi.org/10.5194/acp-2017-1189-AC2, 2018

Atmospheric Chemistry and Physics Discussions

Interactive comment on "Identification of new particle formation events with deep learning" *by* Jorma Joutsensaari et al.

Jorma Joutsensaari et al.

jorma.joutsensaari@uef.fi

Received and published: 18 June 2018

We thank the referee for valuable comments and suggestions to improve the manuscript (MS). We have considered the comments and will modify MS accordingly. Our detailed responses to the referee's comments are below.

Referee's comment 1:

You need to explain what happens when you have a different site: do you need to remake the training and testing with the new subset of images? Or can you use the developed recognition in this study without any changes? And what happens when you have a site that has completely different shape of the size distribution compared to San Petro Capofiume? Or size range of the size distribution? In other words, what are

your instructions and recommendations of how to proceed with your results and your methods when you do the same analysis at a different site? This needs to be clearly explained in abstract and conclusion sections. And there is no mentioning if you really recommend the method to be used already, or if you would like to develop it further before anyone should use it.

Authors' response:

We have not yet tested the method with other sites. Basically, "banana type" events, non-event days and bad data should be recognized from other site data if pictures are plotted roughly in a similar way (one-day plot, size ranges, axes and color map). The method analyses features from size distribution plots, which are quite similar in many cases in different sites. However, we still think that the CNN should be transfer learned again for new sites in order to get best results, especially if the shapes of size distributions are completely different (e.g. low tide events in Mace Head in Ireland or rush hour episodes in urban environments). In other words, once learned CNN can be used in other sites but more precise results will be got if CNN has been transfer learned with data from the same site. In general, we think that anyone can already use the method because the basic concept is efficient enough. However, classification accuracy can be improved by testing different parameters and optimizing set of classes.

We will discuss this in more detail in the abstract and conclusions sections in the revised version of MS.

Referee's comment 2:

In relation to this: If you have to remake the training at each measurement site, do you always need to train the dataset with 50 % of the pictures? Does this mean that you have to select 50 % of your data at a new site already classified manually to be able to do the classification with NN? If this is the case, it is a severe drawback of the method. If you have to classify 50 % of the days manually on each new site, then then there is little point of doing the NN classification. If this is the case, then please write it clearly

in the abstract and conclusions.

Authors' response:

In general, the percentage of labeled data is not the most important parameter, but the number of images in relation to the size of the CNN is the relevant one. For instance, if the CNN is small – small number of layers and small number of neuron per layer – then, the minimum amount of images in the training pool is small, however, if the CNN is large, then the training pool should be large too. Furthermore, the complexity of the problem to be solved affects the number of images needed.

As a rule of thumb, if the data are similar enough from one site to another – particle concentration, time scale, measurement device, etc. – and the data are depicted with the same color map and log scale, then the classifier can be readily used. The color map plays a role only if it is not an "optimal" map; for instance, if the colormap distorts the data – e.g. a small variation in number concentration creates a big difference in color – then choosing another colormap will modify the performance of the classifier.

In our case, we have totally ca. 2000 days for training in the 50 % case and, e.g., Class 1 has only ca. 130 days. If you merge some classes together (e.g. Classes 1 and 2) and you have well-classified data, less training data is needed. Generally, you need some data for training. Once you have trained CNN, you can used it for all new data from that site. Alternatively, simulated data could be used for training but we have not tested how well it works in practice.

We will discuss this in more detail in abstract and conclusions sections in the revised version of MS. In summary, we suggest that the CNN should likely be transfer-learned again for new sites to get the best results but in training, ca. 150 days per class should be enough to get a reasonable classification.

Referee's comment 3:

Abstract

The abstract is unusually long and has a very long description of the deep learning process. However, this is justified in this case, since aerosol researchers are normally not working with deep learning and a longer description is useful. So, the abstract should not be shortened.

Authors' response: We will only slightly modify abstract based on the Referee #1 comments.

Introduction and chapter 2.3

Referee's comment 4:

To be able to understand the NN methods, one way is to either, be very theoretically minded with an ability to understand abstract concepts and base your knowledge on this paper and other articles describing the methods, or you have to be practically oriented and learn by doing and be shown practical examples. As an extremely practically oriented person without an abstract mind set, I have no chance of understanding the methods based on reading. However, this does not automatically disqualify your text. After all, the abstract thinking might understand it. Hence, we have to accept that some people will understand the text, and some will not. Those that will not understand, will have to be learned by extensive simple examples, or by a teacher with a few practical examples, or at specialized workshops, and maybe with support from pedagogical video clips. Since it is not your task to develop extensive pedagogical descriptive examples (which is beyond the page limit of normal scientific papers), we have to accept this pedagogical problem and leave it as it is without further changes.

Authors' response:

We have recognized this problem how to describe the method in a simple enough way but simultaneously theoretically enough. Based on the Referee #1 comments, we will move some of the most theoretical parts of the method description to an appendix to make the text more concise. Very practical descriptions and some exam-

ples (videos and codes) can be found, e.g. from Mathworks (Matlab) web pages: https://se.mathworks.com/discovery/deep-learning.html. Furthermore, more detailed description can be found from textbooks of the subject (e.g. Buduma and Locascio, 2017; Duda et al., 2012, Ch. 6.2.), which we now cite in the text.

Introduction

Referee's comment 5:

Page 2, row 12: Please add that passing on the method of classifying new particle formation events to a second person(s) might lead to systematic bias. If the second person passes on the knowledge to a third person(s), the systematic bias could increase further, and so on. I have experienced this problem previously, and it is a serious problem with the manual classification, and gives further motivation to develop automatic methods.

Authors' response:

We will add to revised MS that passing on the manual classification method from researcher to researcher could lead an increasing systematic bias.

Referee's comment 6:

Page 2, rows 13-15: Wrong referencing to effects. That aerosols affect radiative balance does not automatically mean that they influence the climate via the direct and indirect effect. Please rephrase into something like this: ". . .radiative balance of the Earth and therefore the climate. They affect the climate directly by either scattering incoming solar radiation back to space or by absorbing it. Indirectly, aerosols affect the climate via their role in cloud formation as cloud condensation nuclei (CCN)."

Authors' response: We will change text accordingly.

Referee's comment 7:

Page 2, rows 18-19: Please add to the text that also the direct effect is leading to a

cooling.

Authors' response: We will change text as suggested.

Referee's comment 8:

Page 2, rows 19: Please avoid using everyday language like "some". Please write "part" instead.

Authors' response: We will change text as suggested.

Referee's comment 9:

Page 4, rows 30-31: "Therefore, the idea of reusing what is already known instead of re-learning from scratch every time a new class has come up.". This sentence is not grammatically correct.

Authors' response: This sentence have been deleted when we have reorganized MS.

Materials and Methods

Referee's comment 10:

Page 6, row 26: Please write that it is the "traditional method" in the title. Otherwise it can be confused with your new method.

Authors' response: We will change the title as suggested.

Referee's comment 11:

Page 9, row 28 until Page 10, row 2. You mention that you "also tested three different sets of particle size distribution images". The reader might understand that these 3 methods are additional ones to the original method, while I think you mean that these are all the three methods that you have. Please rephrase to make it clearer, maybe by avoiding using "also" in the first sentence.

Authors' response: We will change text as suggested.

Referee's comment 12:

Page 9, rows 12-13 and Page 10, row 6. Please refer to Table 1 when mentioning the training/testing procedure.

Authors' response: We will change text accordingly.

Results and discussion

Referee's comment 13:

Page 12, rows 24-25. Do you have statistics to support your claim? Authors' response:

We have not studied this by statistical analysis and it is just a general statement based on randomly selected days. It would be very time-consuming if we check all misclassified days manually. In general, a part of misclassifications is pure booking errors made by researchers (wrong class written to database) and thus human-made errors seem to be more common. In addition, classification can be easily vary from researcher to researcher.

Tables

Referee's comment 13:

Table 1 caption text does not make sense when reading for the first time without studying the manuscript in detail. Please explain shortly in the caption text what you mean with training and testing. I am aware that this is explained later in the text (chapter 2.3), but needs a short explanation also when you mention Table 1 for the first time. Alternatively, you can write in the Table caption that this training and testing is explained in chapter 2.3.

Authors' response: We will change the caption text as suggested.

Interactive comment on Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2017-1189, 2018.