Atmos. Chem. Phys. Discuss.,
https://doi.org/10.5194/acp-2019-298-RC3, 2019

Atmospheric Chemistry and Physics Discussions

ACPD

Interactive comment

Interactive comment on "Machine learning for observation bias correction with application to dust storm data assimilation" *by* J. Jianbing et al.

Anonymous Referee #3

Received and published: 14 June 2019

The aim of this work is to develop a methodology able to use PM10 as a tracer of dust with the last end of assimilating PM10 to improve the forecast of dust storms. The deterrent of direct assimilation of PM10 is that it does not only encompass dust but also other aerosol flavors that are more significant where anthropogenic sources of aerosols are present. To isolate dust and non-dust aerosols the idea is to consider non-dust as bias in the PM10 records. Then the problem is reduced to apply bias correction methods to PM10. Here the authors propose to use machine learning and chemical transport model for bias correction. Results show a more accurate forecast of dust storm if the PM10 assimilated is isolated form non-dust aerosols using the machine learning method.

The paper is well written, structured and results support the conclusion archived. The topic is innovative since it is one of the first works to apply machine learning to bias correction to isolate non-dust aerosols from PM10. In addition, the methodology developed here goes beyond to improve the forecast of dust storm and it can benefit the assimilation of other magnitudes. This study match with the ACP topics and its quality is close to its standards.

However, there are some general comments that need to be addressed or clarified along with the manuscript.

1) How this approach guarantees the "bias" in PM10 that you are correcting is non-dust and it is not a "real" bias present in your PM10 records?

2) How do you know a better performance of LSTM for bias correction (or dust isolation from PM10) and also better dust forecast if you are not comparing against dust records? I am aware of the lack of availability of them but at least it will be necessary to actually know that your PM10 without bias is dust by comparing with some station or a satellite image.

3) Also, the main caveat of the machine learning methods is the availability of a long series of records and the computational time. Do you think that this method will be able to correct the PM10 bias obtained a few hours (e.g. 12 h) before a dust storm and be ready to assimilate and perform the forecast a few hours later? The real applicability of LSTM is not clear. This point needs a little more discussion.

4) Finally, the conclusion about what method is better for bias correction looks like that is the same method that better "simulate" the non-dust aerosols. Then, the worst bias correction with LOTOS-EUROS model is just the worst performance in reproducing non-dust aerosols. This simplification is not straightforward needs to be clarified.

Minor comments

P4-L29. to large -> too large

P5-L2. Can easily applied -> can easily be applied

P7-L3. Following Jin et al., 2018

P9-L15. Anthropogenic emissions are from MEIC but it is not clear where natural emissions are obtained. Are these natural emissions; sea salt, biogenic emissions, and wildfires, but not dust?

P10-L11 How LSTM is used needs little more details. It is not clear why these elements of the input vector have been chosen. Do these observations correspond to blue dots in Fig. 1? Why do you use nearby sites for PM2.5 but not for the other species? Is the training period only past 18 h or from January 2013 to March 2015 as it is said in P11-L4? Are these data available hourly? P11-L23 What criteria do you use to consider that a series has high data missing?

P11-L4 I would like to see (or at least discuss) the sensitivity of the results to the training period. For instance, in the manuscript it is argued that from 01-2013 to 03-2015 the frequency of dust storm was low then it could be considered that all PM10 were non-dust. However, yb will be significantly larger during the days with a dust storm and this could affect the regression model. Have you checked if the regression model changes if you remove those days with dust storm?

P12-L12 In non-dust PM10 evaluation, how do you know that they are dust in Fig. 3? Is it only based on larger numbers?

P11-L1 How do you know LE/non-dust underestimate non-dust PM10? Is it based in Fig. 3 and lower values of PM10?

P11-L5 To show the better agreement of LSTM than LE/no-dust, could you give some number such as correlations?

P13-L25. Is a high level of pollution the reason why these sites are chosen to test the performance of LE/non-dust?

P20-L15 In addition to the orography, a reason why in Beijing the PM10 is underestimated could be related to higher PM10 non-dust values? Did you find any relationship between more polluted cities and worst performance of the bias correction?

In the section of the evaluation of the forecast skill, I miss comparing with "real" dust records instead of PM10.

Figures, in general, are clear and the captions properly describe them.

Fig.1 It is difficult to distinguish the star symbols which denote the cities location. Also, Is N=1352 number of stations used in LSTM?

Fig. 3 Use properly notation for float numbers (e.g. 2 10-6 instead of 0.00002).

Fig. 6 Unify the legends and try it does not overlap with the time series.

Interactive comment on Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2019-298, 2019.