Atmos. Chem. Phys. Discuss.,
https://doi.org/10.5194/acp-2019-298-RC2, 2019

Atmospheric Chemistry and Physics Discussions

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

Anonymous Referee #2

Received and published: 13 June 2019

This article looks at the effect of applying a bias correction term to PM10 observations in order to assimilate them in to a dust model as 'dust' observations. The bias correction is needed since PM10 observations account not just for dust, but also other types of aerosol. The paper compares two different bias correction methods to just assimilating the PM10 observations directly. The first bias correction method is to use a CTM to calculate the non-dust part of the PM10 observations and the second more novel method is a machine learning approach.

The paper is in general well written with clear figures, although there are a few grammatical English errors (those that I could easily comment on are listed under typos below). The structure is straight forward and the logic easy to follow. It is also fairly comprehensive and so provides an in-depth examination of the two different bias cor-

rection methods. However, it does not feel that this is too much information but rather provides context as to why the assimilation results are seen.

I have two main comments to make on the article. The first is that it wasn't clear to me, from reading through the article, how such a bias correction would be applied in a real life scenario. There is some comment on this in the summary and conclusions but it would be interesting and very relevant to understand how either bias correction scheme might be applied in practice. It is not obvious to me how this would be possible. My second comment is that I sometimes struggled to understand the detail of what had actually been done in each of the two schemes and this is addressed in the more specific comments below:

Comments:

Pg. 1, line 20. I don't believe that the conclusion can be drawn in general that the best results are obtained when using a machine learning model. This is true in this article using this CTM, but the results that show an under-estimation of the non-dust PM10 from the CTM would, to me, provide evidence that the CTM is a slightly flawed model for this aspect and hence why the machine learning approach performs better. It would be enough to change this sentence to 'The best results are obtained when using the machine learning model...'

Pg. 7, line 7. What do you mean to release the efforts in updating the tangent linear model? To me this would imply the effort to keep the tangent linear model up to date with the full non-linear model, but I'm assuming that what you really mean is to reduce the time spent in calculating the tangent linear model?

Pg. 9, Section 3.1. I assume this chemical transport model was chosen because it produces a full operational forecast over the modeling domain? Does this therefore make the application of the bias correction possible for a real life scenario? If true, then this would provide a good explanation for why this model is used to evaluate between the two different schemes when it is clearly unable to estimate the non-dust PM10

observations.

Pg. 10, line 10. Are the input observations of PM25, SO2, NO2, O3 and CO all coming from the same station as the PM10 observations used to compare the output? Is this why the nearby sites of PM25 are included as a separate line. Does the machine learning model therefore come up with a different value of PM10 for each station separately or is all the data included together so that given the input observations from any station, a generic PM10 output would be generated?

Pg. 10, eqn 8. What does the 'm' stand for in this equation. Is it observation across the full domain and time period or just observations at one station over the time period?

Pg. 11, line 22-line 28. The description in this paragraph needs further clarification. I believe that you are subtracting the 1-hr forecast made from April 15, 19:00 (so valid at 20:00?) from all the PM10 observations between 8:00 and 19:00 for assimilation. Or is it the one hour forecast valid at each specific time of the observation? In which case does each observation from each station require different 18hour input? Why is the 1-hr forecast chosen rather than the actual value coming from the machine-learning method? How is a forecast made from a machine learning method?

Similarly, is the 12-hr forecast from April 15, 19:00 (so valid at April 16, 07:00) added to all the forecast dust PM10 values to compare to the observations?

Pg. 9-11, Section 3.2. How would such a machine learning algorithm be used in practice? Would it need the constant evaluation of running 18hrs of data? Or if a dust storm was forecast, would the calculations be switched on. At what point would the PM10 observations be assimilated. Or is this envisaged mode for a reanalysis style product of PM10 observations?

Pg. 12, line 9-10. You state that the according to the LOTOS-EUROS bias corrected observations, the dust storm seems to have already reached central China which was in reality not the case? How do you know this? Nothing from this figure provides

evidence of what the reality was? It would be worth referencing how you know this to be true.

Pg. 13, line 17-19. This to me provides evidence that the LOTOS-EUROS model is not doing a good job of predicting the non-dust PM10 and so is never going to match the performance of the machine learning algorithm. Is there a reason for choosing this model or not exploring CTMs that may provide a better match?

Pg. 18, line 20. Again, you state that the LOTOS-EUROS observations over-estimate the observations and yet I can't see anything in Figure 8 that shows the observations, so that we know this.

Pg. 18, line 31-33. It is interesting to me that although the peak is a better match with the machine learning algorithm, the forecast is actually slightly better with the LOTOS-EUROS model. Do you have any feeling for why this may be?

Typos:

Pg. 1, line 13. I think this should read 'The latter is trained by learning using two years of historical samples'.

Pg. 2, line 9. Should be 'progress has also been made'

Pg. 2, line 19. Should be 'A wide variety of data assimilation techniques have been used...'

Pg. 2, line 28. Should be 'In the presence of biases...'

Pg. 10, line 25. Should be 'Earlier studies showed that the...'

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