Atmospheric
Chemistry
and Physics
Discussions

Interactive comment on "Improving air quality forecasting with the assimilation of GOCI AOD retrievals during the KORUS-AQ period" *by* Soyoung Ha et al.

Soyoung Ha et al.

syha@ucar.edu

Received and published: 22 December 2019

xcolor, colortbl

We appreciate the reviewer's valuable comments. It is our belief that they greatly helped improving our manuscript. In response to your major concern, we now added three more figures (as new figures 12 - 14) and one more table (Table 4) for additional verification against independent observations during the KORUS-AQ field campaign. Please find our point-by-point response in blue below.

C1

Anonymous Referee 2:

Received and published: 24 October 2019

The manuscript presents a study assimilating ground-based observations and satellite based retrievals to improve PM2.5 forecasts. The topic is relevant and in the scope of the Journal. There have been several previous studies about assimilating geostationary satellite retrievals, especially those from GOCI. This study builds over them and, in my mind, has a few additional contributions: The first and most important is that assimilating GOCI by itself doesn't seem to improve the forecasts for the study period, sometimes even making it worse. The authors only find improvements when they assimilate both surface and satellite data. Other contributions include separating assimilation performance by pollution regimes, and showing sensitivity studies on how to represent the error on the observations and how to aggregate them to speed up the assimilation algorithm without worsening performance. These represent good contributions to the field and would grant publishing in ACP.

=> The most important result of this study is summarized in Fig. 11 where the assimilation of GOCI retrievals was shown to be particularly helpful in improving the forecast accuracy in high pollution events and keeping the positive impact for 24 h. Without the help of GOCI retrievals, the assimilation of surface PM alone was not effective after 6 h, especially in predicting high pollution events. Hence, we disagree with your first and the most important finding from our study that assimilating GOCI alone doesn't improve the forecasts and sometimes makes them worse.
In response to your comments and for clarification, however, we added new figures 12 -14 where 0-23h forecasts from each experiment are verified against independent observations to demonstrate the positive impact of GOCI without the help of surface PM observations. With respect to both total AOD at 500 nm from AERONET sites (Fig. 13) and to surface PM2.5 concentrations from the stations operated by NIER during

the field campaign (Fig. 14), the assimilation of GOCI alone mostly outperforms PM and MODIS experiments. Details on the new figures are now added as the last two paragraph in section 4.2.

However, I think the paper needs a lot more work before it's ready for publication. In terms of the science, I think they need to do more work on understanding why GOCI makes the assimilation worse and why does it get better when including the surface observation. The study period coincides with a major field campaign where additional airborne and ground-based observation were made to try to tackle these questions. The authors also talk about the operational forecasts, so it would be good if they could include the performance of this system to understand if the assimilation efforts can help improve the current system. Also, it needs major improvements in the English. I suggest the authors to find support by a native English speaker. A few major and minor comments below.

=> Again, GOCI did not make the assimilation worse. As noted in lines 27 – 28 in page 11 in the original version, the assimilation of AOD retrievals (either GOCI or MODIS) alone does not improve the surface analysis. The main reason for that is that AOD retrievals are not directly associated with surface PM2.5 concentrations. This is already discussed along with the imperfection of the observation operator and the numerical modeling system in the last paragraph of section 3.2.2 (Page 9: lines 25 - 29) and lines 12 - 15 in Page 11. Moreover, in the 3DVAR assimilation, the model estimates of AOD are strongly constrained by the model error structure of each aerosol species both horizontally and vertically. As such, the most challenging part with the real data assimilation is to improve subsequent forecasts *per se*. Please note that this is not an idealized study using a simple model or simulated observations. It is not practically feasible to isolate numerous factors in the study where real observations are assimilated using the real (e.g. non-idealized) system. However, based on your comments,

we decided to add one more paragraph in the last section. Please find the second paragraph in section 5 that summarizes our discussion on this matter.

As for the operational forecasts, it is correct that this study was motivated by an effort to improve the current operational forecasts, but because they use a different chemical transport model with no data assimilation, it is not relevant to directly compare to them. Also, the focus of this study is to examine the impact of GOCI retrievals (which are not included in the operational forecasts at NIER), so we did not include them here.

In regards to English, this manuscript was internally reviewed twice and already proofread by another native English speaker. Without exact lines or paragraphs specified, it is hard to figure out where and how we need to make major improvements. But we made some corrections while editing section 4.2 to add new figures and table. For instance, we changed page 11, line 28 as "the analysis error is smaller than those in other experiments" by as "those in". Also, in respect of reviewer's comment, we've gone through another round of proof-reading and made necessary corrections throughout the manuscript once again. Thanks for your suggestion.

Comments by line (<page> <lines>):

2 3-8. Provide references to these statements

=> This article is now added to our reference: Chang, L.-S., Cho, A., Park, H., Nam, K., Kim, D., Hong, J.-H., and Song, C.-K.: Human-model hybrid Korean air quality forecasting system, Journal of the Air and Waste Management Association, 66, 896–911, https://doi.org/10.1080/10962247.2016.1206995, 2016.

5 21-22. But you mention above that Saide et al. (2014) used MOSAIC within GSI. For for Page 15, lines 1-5.

=> We now added "publicly" before "available ∼".

6 6. By doing cycles every 6 hours you are not taking advantage of the hourly time resolution of GOCI data

=> Agree. We may need to increase the cycling frequency to 3 hourly or hourly in the near future. We now added one statement "This configuration was chosen in the limitation of computational resources, but the use of higher resolutions both in time and space might be desirable to further improve forecast skills in the future." in lines 12 - 14 in page 5 (Section 3.1).

Figure 1. Why show observations for a given time? Why not show maybe an average of the period analized?

=> Figure 1 simply shows the model domain with the observing network. No changes are made.

Eqns 3-6. Please explain why there are more than one error equation and when would you use which

=> Please check the paragraph for the equations (Page 9; lines 4-16) once again as we already explained that the retrieval error was estimated (by Choi et al. 2018) differently depending on which data was used for the verification of the retrievals. And due to representativeness error in the model, the observation error can be adjusted in the assimilation. The following statement (lines 17-22 in the same page 9) already described what we used (ϵ_2) and why we used it.

Figure 5 does not contributes to much information so I would drop it along with the discussion about it

=> Figure 5 is answering part of your question on why the assimilation of GOCI alone does not seem to improve forecasts. We keep it.

Figures 8 nd 7. You could model vertical distribution and impact after assimilation using airborne data and surface lidars deployed as part of KORUS-AQ

=> Not clear on your point here. Figures 7 and 8 show how the model responded to the assimilation of observations in use. This analysis is needed to understand how our assimilation worked in the model space. It has nothing to do with verification.

14 4-9. Be more specific here, mention the approached that you used of smoothing observations instead of thinning

=> We now added at the end of line 9 "We averaged all the pixels over each grid box at 27-km resolution (e.g. superobing) instead of thinning them randomly, for instance."

15 6-11. I don't agree with these statements. There were a couple of flights during the hazy period you study that could be useful. You are showing that GOCI only degrades performance while including surface improves, so you should be changing the vertical resolution through the data assimilation. KORUS-AQ had airborne and ground based lidars that you can evaluate against to assess this. You can also evaluate the model ability to represent aerosol composition, both from supersites observations and aircrafts. You could also be including the ground based (AERONET-DRAGON) and airborne (4STAR) AOD data. There were even PM monitors in different ships that you could also use for evaluation.

=> We now added three new figures (figures 12 – 14) and one more table (Table 4) for more verification. Please find the corresponding statement in the last part of section 4.2.

Minor Edits (<page> <lines>): A few corrections but in general the manuscript doesn't read well for English

1 2: ". . . every day for the last decade, providing . . ."

=> Changed, as suggested.

1 4: "assimilated to make systematic improvements on air quality forecasting in South"

=> We prefer our original statement. No changes.

1 19: I would change "complications" by "uncertainties"

=> Changed.

8 2-4: This sentence is very confusion, I would just erase it and keep the last sentence of the paragraph

=> This statement is related to time windowing, which is an important aspect of 3DVAR since it counts all the observations within the window as available at the same time. And weighting the observations based on their report time can affect the analysis quality. We leave it as it is.

9 14: ". . . account for representativeness error, we also tested with . . ."

=> We do not see any difference by changing "tried" with "tested". No changes.

9 24: "particularly for pollution events"

=> "polluted" is now changed to "pollution", as suggested.

9 32: "the two observations types. . .."

=> We believe "two observation types" is correct, not "two observations types". No changes.

13 27. This sentence is not clear, why do you mean by power instability?

=> That's what we were told. But we omit "due to the power instability" now.

13 33-34: . . ., but higher levels of pollution in SMA are not simulated either.

=> Changed as suggested. Thank you.