Atmos. Chem. Phys. Discuss.,
https://doi.org/10.5194/acp-2019-648-AC2, 2019

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

[journal abbreviation, manuscript]copernicus hyperref xcolor, colortbl

We appreciate the reviewer's time and effort to improve the clarity of this manuscript. It is our belief that the comments helped make our draft clear. Please find our point-by-point response in blue below.

General comments

The focus in the present study is to implement AOD derived from satellite observation over East Asian region with the Korean Geostationary Ocean Color Imager in attempt to improve air quality forecast. The preprocessed data were assimilated with three-dimensional variational data assimilation technique for the Weather Research and Forecasting model coupled with Chemistry. The impact of GOCI AOD on the air quality forecasting is examined by comparing the obtained results with AOD derived from MODIS observation as well as against in-situ PM2.5 at the surface. In the present study, the assimilation of purely surface PM2.5 concentrations systematically underestimates surface PM2.5 and prediction hold only for 6 hours. When the present GOCI AOD retrievals are assimilated with surface PM2.5 observations the forecasts are improved up to 24 h, with the most significant contributions to the prediction of heavy pollution events over South Korea.

The present study is very interesting and it is based on a comprehensive method, which is also very well described in the manuscript. In addition, the discussion of the results hold very well and this is also the case when uncertainties and limitations in the present study are discussed, as well as possible future improvements that could results in more realistic forecasts. However, there are two important questions or major comments that are in dispute and must be settled before this study can be accepted for publication in ACP.

Major comments

1. There is an issue when introducing information of GOCI AODs in the approach to simulate forecasts of air quality when the improvement of the latter is caused by an overestimation in GOCI AOD. For me this is not a correct approach and uncertain to rely on. One reason for the latter is that it seems not to be robust, considering that you will have differences in the statistics, (differences in the weights), when performing forecasts of air quality. The main reason for that is the cloudiness, thus, diverse cloudy conditions means differences in the availability/statistics in GOCI AOD from case to

case.

=> This statement is not concise, but we assume that the reviewer is concerned about two factors.

First, for the small data availability in the cloudy conditions, the total number of GOCI data used in the assimilation was at least more than 2,000 each cycle, as shown in Fig. 3 (with the right y-axis). As the number fluctuates with cycle, the data impact certainly changes (based on the differences between (o-b)'s and (o-a)'s in black), but (o-a)'s are always smaller than (o-b)'s, showing how robust our assimilation system is. Also, our conclusion is made not from a single case, but based on the one-month statistics.

In regards to the overestimation due to the assimilation of GOCI AODs, here is our response: Model trajectories are always deviated from the observed states and data assimilation is trying to pull the model states toward the observed information. If the model states were severely underestimated, they are drawn to observations to the extent the model is uncertain and the observations are trusted (based on their error statistics). Here the GOCI data tends to slightly overestimate surface PM2.5, trying to compensate for the systematic underestimation (which is a long-standing issue of the bulk GOCART aerosol scheme). But when assimilated with surface PM2.5 observations, the overestimation mostly disappears as both GOCI and surface data are affecting the model states together. The effect of GOCI data can be further adjusted through the observation error variance, but the observation error should not be adjusted for different experiments with different datasets. Moreover, this study is not meant to optimize the system for the particular case. As always, there might be a room for further optimizing the assimilation algorithm (such as improving observation error statistics or the observation operator) or further constraining the model states through other components like emission data, but that is not the focus of this study. The goal of this work is to examine the relative impact of the GOCI data in the same analysis and forecast system using the same forcing (e.g. emissions and boundary

conditions). Please note that our results were reliable and consistent throughout the month-long period.

2. There are issues with the language, which need to be improved. In the section Technical corrections below suggestions are given in an attempt to improve the language and clearness of the manuscript. However, my review and corrections of the manuscript concerning the language has only been carried out for the abstract and Introduction to show that the clearness of the text need to be improved. Therefore, I recommend that the full text needs an English proof-check.

=> This manuscript has been internally reviewed twice in our lab and proofread by another native English speaker. As we replied to the reviewer's technical corrections at the very bottom, most of the corrections the reviewer suggested are incorrect in English or distorted our points, if not irrelevant, in the manuscript. The reviewer raised most of the comments or questions regarding MODIS retrievals, which is not the focus of our study but included only for completeness, we thus wanted to stay focused on our goal and highlights of our work. However, we appreciate different views and tried our best to accommodate the reviewer's comments and reflect them in our manuscript unless we have a specific reason not to.

Specific comments

Page 5, Line 17: MODIS is not a satellite and which version is used here, 6.1?

=> We modified "MODIS and GOCI satellites" to "MODIS and GOCI sensors". And yes, version 6.1 was used.

Line 22: It is not correct to write that AOD measures something. This sentence need to be rewritten.

=> This is simply another way of expressing the definition.

In $https://neo.sci.gsfc.nasa.gov/view.php?datasetId=MODAL2\_M\_AER\_OD$, for example, aerosol optical thickness (or depth) is described as "a measure of how much light the airborne particles prevent from traveling through the atmosphere", which is consistent with our sentence "Aerosol optical depth (AOD) measures the amount of light extinction by aerosol scattering and absorption in the atmospheric column". As nothing is wrong with the expression, it is unchanged.

Page 7 Line 15. Which version of MODIS? Line 17. => V6.1

"Observation errors" are not the best name to use here and there is a later estimates from the MODIS aerosol team of the expected error in the MODIS retrievals of AOD. Suggestion: "The MODIS land and ocean retrievals give AOT at 550nm with expected error envelopes of AOT =±0.05± 0.15 *AOT (Levy et al., 2010) and AOT = +0.04+0.1 AOT = ?0.02??0.1AOT$(Levyetal., 2013), respectively, whicharisefromcombinederrorsinassumedbc$

=> In the data assimilation framework, errors are divided up into two major categories - background errors and observation errors. We appreciate the details of the reviewer's comments, but this study should read in the context of data assimilation, not for the retrieval of MODIS aerosol products. In other words, the errors should be defined either from the observation side or from the model side in the assimilation system, so we believe that it is more appropriate to describe the retrieval errors as observation errors. Also, the error estimate was first described in Remer et al. (2005), as stated in our draft, so we leave the reference as it is. But for better clarification, we modified the statement as below. "Following Remer et al. (2005), observation errors are specified as the retrieval errors: (0.03 + 0.05 * AOD) over ocean and (0.05 + 0.15 * AOD) over

land. They do not include the representativeness error and are slightly smaller than those for GOCI AOD, as described below."

Page 9 Line 1. And the sentence beginning with "Because the difference. . ..." If the difference is so small should you not then go for thinning?

=> As we stated in the second paragraph in page 8, it is common practice to go thinning satellite data to reduce the data volume in data assimilation. But based on our results, we decided to go for superobing instead of thinning, not because the forecast performance is much different (which is not the case), but mostly for the computational efficiency.

Lines 5 and 6. The word "validation" can be used when comparing satellite derived AOD against ground-based sun-photometer measurements. However, you cannot validate AOD obtained from passive remote sensing against AOD derived from observations with another satellite sensor used in passive remote sensing.

=> We do not want to argue about how others described their work. The term of "validation" was used in Choi et al. (2018) and we just adopted it here. Also, in a broad sense, the terminology of "validation" is commonly used when one data is evaluated against another independent observation in the data assimilation community, so we do not see it problematic. No changes.

Line 15. Concerning the sentence "When these different observation errors were applied to GOCI retrievals in the assimilation, the smallest error (ϵ_2) produced slightly better fits to observations specially for the high values (AOD > 2)....." This statement seems not hold, since ϵ_2 is not better than ϵ_1 over land for the situations with lower AOD.

=> We do not understand why our statement doesn't hold due to the case of lower AOD, which we did not even discuss here. We mentioned that the smallest error (ϵ_2) produced slightly better fits to observations for the high(!) AOD values. We also stated that such a result is not statistically significantly different, so we do not understand why the reviewer is arguing over the statement. No changes.

Line26. Concerning this statement "This is partly because AOD is not directly associated with surface PM2.5 and partly because large uncertainties in the forecast model and the emission forcing can dominate over the analysis error during the model integration." how about uncertainties may also be induced in the forecasts of PM2.5 since here we have to deal with ambient AOD while dry conditions for PM2.5?

=> The sentence "how about uncertainties may also be induced in the forecasts of PM2.5 since here we have to deal with ambient AOD while dry conditions for PM2.5?" doesn't make sense, so we are not sure what the reviewer's point is. But just for clarification, here is what we meant: Even if the high-volume AOD data are assimilated, if the information in AOD retrievals is not well matched with surface PM2.5, it may not contribute much to improving the forecast of surface PM2.5 concentrations. That's why it is important to examine the relationship between AOD and surface PM2.5.

Section 5 and page 15 Line 1 and first sentence. I think too strong positive words are used here when describing the GOCI AOD retrievals.

=> Which part is too strong? We concluded solely based on our results here. Please elaborate your comment or specify the statement that is considered to be unsuitable.

Line 15 and the sentence "However, the forecast error grew very quickly over the next 12 hours, underestimating PM2.5 at the surface, especially in the heavy pollution events where the forecast accuracy dropped from over 70% to $\sim$30% only in four

hours. Meanwhile, the GOCI AOD retrievals alone tended to overestimate surface PM2.5 but significantly contributed to improving air quality forecasts up to 24 h when assimilated with surface PM2.5 observations." Thus, the improvement increase for the wrong reason and one of the problems with that seems to be that the approach is not robust, considering that you will probably get variations in the amount of data/statistics you get from GOCI AOD when investigating different cases. This is because of cloudiness, which means no data when clouds are presented. Thus, the GOCI AOD statistics will varies between different forecasts investigated, thus the latter will be dependent on this.

=> We agree with the reviewer in that the results might vary between different cases to some extent. But we do not agree with the comment that the improvement was made for a wrong reason. During the one-month cycling, we could assimilate at least a couple of thousands of data points even in cloudy days, as shown in Fig. 3 where the right y-axis starts from 2,000. Also, if you compare the black dashed line (o-b GOCI) with the black solid line (o-a GOCI) in the figure, you can see the analysis produced much better fits to the observations, meaning that the analysis itself was done correctly. When the number of data in use gets reduced, the differences between (o-b)'s and (o-a)'s get reduced as well, implying that the effect of the data gets smaller, as expected. But (o-a)'s are consistently better than (o-b)'s throughout the cycles, showing how robust the system is. And even if the spatial coverage of GOCI data varies cycle to cycle, we still use more GOCI data than surface observations (which is typically around 1,000 at each cycle). Also note that we concluded about the data impact based on the month-long statistics, not on a single case.

Why is not the results obtained with MODIS included in the discussion of Section 5?

=> We chose to summarize our main points in the last section, rather than listing all the results. The impact of MODIS data has been examined in many previous studies, and we included MODIS data just for completeness, so nothing to emphasize on the

MODIS data here.

Figure 1 It is not clear how the two domains are connected to the solid box in the figure. Neither the text in the beginning Section 3.1 is clear about this.

=> We now modified the caption as “Two model domains - domain 1 (outer box) at 27-km resolution nested down to the inner box for domain 2 ” to add “(outer box)”.

The number of the in- situ stations at Korean peninsula that deliver PM2.4 data for the present study are so much more than is shown by the dots in this figure. This could be improved somewhat by reducing the size of the current Figure 1 (the right part/eastern part) and include instead at this place on the right an enlargement of Korean peninsula?

=> In fact, many Korean sites are tightly overlapped to each other. As such, zooming in the map does not help much to recognize individual sites. But Figure 13 (now as Figure 16) can give readers a sense of how they are distributed over South Korea. Since this study focused on the impact of GOCI data, not surface stations, we believe it is enough to show the entire model configuration in Fig.1. No changes.

Figure 2 It is not correct to write that AOD is retrieved at this time, since it is the observations that is carried out at this time and it is a very long way to come up with an estimate of AOD, for example you have to introduce a model that describe radiation transfer in the atmosphere. Change “retrieved” to “corresponding” in the first sentence of the figure caption to Figure 2. Describe in the figure caption the solid black box introduced.

=> This study is not meant for describing the retrieval process, but how the data is used in the assimilation cycle. The data was processed at the time and that’s how they are described and presented in Figure 2. Even in-situ measurements such as

radiosonde do not report the values at the exact time (depending on the vertical levels as it goes up), but in the data assimilation context, that’s how they are all described. Please note that we already illustrated the temporal distribution of the data in the last paragraph of page 7 (“In terms of temporal distribution,). In response to your last comment, though, we added one paragraph “Domain 2 is marked as a black box in each panel.” at the end of the caption.

Figure 3 What is it on the y-axes? Should it be AODo – AODa and AODo – AODb ?

=> We believe the reviewer already understood what we plotted based on our y-axis label (and our caption). Along with the main title “GOCI AOD”, it must be clear that our y-axis label “(o-a) and (o-b)” means “AODo-AODa and AODo-AODb”. No changes are made.

Figure 4 In the figure caption you have to refer to the body text about the three different types of observation errors. Take the color blind persons in consideration and use the three colors in combination with solid, dashed and dotted lines and separate land and ocean with heavy and normal lines, respectively.

=> This draft uses a lot of colors throughout the figures, and is not meant for color-blinded readers, unfortunately. But based on your comment, we added “The first two errors (ϵ_1 and ϵ_2) are described in equations (3) - (6) and the third error (ϵ_3) increases ϵ_2 by 20% everywhere.” in the caption.

Figure 5 R2 is squared correlation coefficient. => Corrected in figure and caption.

Figure 6 Keep the color for the lines but use solid and dashed lines. Why is not the results discussed more than for is included as phrase in the bracket? A suggestion,

skip the figure and write “(not shown)”.

=> Figure 6 is important in that it gives an overview of the entire month of interest, summarizing how much we can improve the forecasts with the assimilation of all the data considered for the whole period. Moreover, it nicely introduces the high pollution events we discuss later, so we should not omit it. And we decided to go with solid lines to better distinguish the forecasts from dots for observations. No changes are made.

Figure 7 It is a lot of space in the figure and therefore write the names of the species in all figures.

=> We decided to put the species name in the main title because the first panel does not have room for it due to the legend. As this figure has to take up the whole page (height-wise) anyway, we decided to keep the main title. No changes.

Figure 8 Write “Model levels” connected to the y-axes.

=> The caption already stated that it is the same as Figure 7. We tried to reserve more x-axis space to zoom in differences between the experiments here, dropping y-axis title intentionally. No changes made.

Figure 9 The title (above the figures) is problematic both considering the language and that it is actually not 100% monthly mean values that are presented. I suggest to remove it and change the figure caption to. “Horizontal distribution of analysis increments in PM2.5 (analysis-minus-background) at the lowest model level (k=1) in domain 1, averaged over the period 4 – 31 May 2016. Maximum and mean values corresponding to the domain in each experiment are shown in the upper right corner of each panel.” However, text describing the different figures are also needed in the figure caption. Since no alphabetic characters, a – d, have been included in these four figures then

you have to include the more complicated “upper left, upper right” etc. In addition the x-label text should be “dry PM2.5 [um m-3]”

=> We changed the main title to “Analysis increment”. As for the annotation of each panel, we believe the experiment name is the best way to describe each panel since this figure highlights differences between the experiments with different observations assimilated. We never use “upper left, upper right”, etc. here and go by the experiment name to be clear on which data we compare. As for the label bar title, PM2_5_DRY is the exact name of the model variable we read from the model output. The caption is now changed accordingly, as follows. “Horizontal distribution of analysis increments (analysis-minus-background) in PM2_5_DRY, the model variable corresponding to PM2.5, at the lowest level in domain 1, averaged over the period of May 4 - 31, 2016. Maximum and mean values of the domain in each experiment are shown in the upper right corner of each panel.”

The results presented in Figure 9 are somewhat difficult to understand, since including result of GOCI AOD means that the final scene (All) abrupt get higher PM2.5 values in the upper part of the figure (Figure 9d). Is this realistic? In addition, how could the MODIS AOD results that are so limited in available values/statistics for the investigation area contribute to an improved forecast? It seems also that the MODIS AODs are only available east of the Korean peninsula, while the aerosol sources are located in west, over China.

=> To determine if the analysis increment is realistic, we verified the analysis and the following forecasts with respect to observations, as shown in the following figures. In regards to the impact of MODIS AOD products, the data can affect 15 three-dimensional GOCART aerosol species in the model states and the impact of each data point can be extended to the neighboring area as specified in the background error covariance.

Figure 10 Writing “9-km simulations” is not clear, explain what it is. Suggestion for the figure caption “Root-mean-square-error (rmse, upper figure) and bias error (lower figure) obtained for forecasts, with respect to the investigation area of 9-km (domain 1), verified against surface in-situ PM2.5 from 361 stations in South Korea of the period 4 – 31 May 2016. Average values of the forecasts (24 hours with increment of 1 hour) is shown next to each experiment name, where also mean absolute error is presented.” Should the latter be presented in the upper figure instead? Please adjust the suggested figure text above if needed.

=> Based on the reviewer’s comment, the caption is changed as below. “A time series of root-mean-square-error (rmse; upper panel) and bias (lower panel) of the hourly forecasts from the 00 Z initialization for May 4 - 31, 2016. Different experiments in domain 2 are verified against the same surface PM2.5 observations from 361 stations in South Korea. An average of 0-24 h forecast errors is shown next to each experiment name. The mean absolute error (mae) over the 24-h forecasts is also shown in the lower panel.”

Figure 11 Suggestion “Figure 11. Same as Fig. 10, while here the results of forecast accuracy (%) for categorical forecasts are presented, subdivided according to classification of air quality in Tables 2 and 3.” Include “Model level” on the y-axes.

=> Figure 11 shows different statistics, not the model level, as shown in the main title. No changes made.

Technical corrections

It was very hard to read through all the comments because they are mostly incomplete and are not separated by lines. In many cases, the modifications that the reviewer suggested either do not flow well in our manuscript, or misrepresent our points, or are

simply wrong in grammar. The reviewer tries to change our manuscript line by line in his/her way, but we should ask for being respectful for the authors’ work. But here are our responses to the questions:

What is meant by “positive impact”?

=> The impact is considered to be positive when the following forecasts are improved with the reduction of forecast error (in terms of rmse, bias, and the categorical forecast accuracy).

Line 13 in abstract: The last part beginning with “with the most. . .” of this sentence is not clear.

=> We clearly demonstrated the most significant contributions to the prediction of heavy pollution events in Figure 11 where the assimilation of GOCI data produced the biggest improvement in b) high pollution accuracy.

Introduction Page 1 Line 21: “Surface concentrations” of what?

=> of chemical species. We believe this should be clear as the previous paragraph is immediately followed by this one.

Page 2, Line 4: What is meant by “these fast-varying complex mechanisms” or what is it pointed to?

=> All the mechanisms described in the previous paragraph, particularly the aerosol-meteorology interaction at short time scales. Again, this sentence is also connected with the paragraph right ahead.

Page 3, Line 10: Not clear what is meant by this “careful investigation of data characteristics”.

=> We meant by examining the data in various ways, particularly in preprocessing and the error characterization, as shown in figures 2-5.

The last part in this sentence is not clear “. . .. . .compared to that of other observations.” Line 13.

=> As described, compared to the impact of other observations on surface PM2.5 forecasts, we meant.