Response to Reviewer #1:

First of all, we would like to thank you for your constructive comments and suggestions. Based on the comments from two reviewers and the executive editor, we have tried to address the issues raised by both reviewers in the revised manuscript. We would also like to respond to the comments point-by-point below. The reviewer's original comments are in italics. Our responses to reviewer's specific comments are put in a normal font. The added or modified parts in the revised manuscript are highlighted in a red color.

*Main comments*

*Comment) One additional CMAQ simulation and analysis that I would like to see is a comparison between using initial conditions from GOCI but without using kriging (single frame on the time of initialization without filling any gap) vs using kriging. This would help show how valuable is to enhance the GOCI data with kriging, which is one of the main novelties of this study. You could add some of these results to Fig 6.*

Reply) We agree with reviewer #1. We have carried out additional hindcast simulations, using initial conditions prepared by single frame of GOCI data. In these simulations, we did not fill any gap if there are not enough observations near model grids. Please, refer to p. 20, lines 454-460 and modified Fig. 6 in the revised manuscript.

*Comment) The results are shown for specific air pollution episodes. However, since the system is planned to be used operationally, it would also be useful to see some performance statistics for the less polluted conditions. That way it can be stated if this is a tool for episodes only or for any condition.*

Reply) We are happy to reply to this comment. We have also found that the STK method showed positive impacts on improving the performances of hindcast runs for less polluted conditions. We have tried to show the performances of the first and the second 24-h hindcast results, responding to a suggestion from reviewer #2. Please check out p. 8, lines 176-178; p. 17, lines 385-386; and pp. 24-25, lines 552-567 in the revised manuscript. Also, we added Fig. 10 into the revised manuscript.

*Comment) Since MODIS AOD is also an operational product and shows slightly better performance than GOCI, it would be nice to see additional sensitivity experiment where MODIS AOD is included into the kriging stage along with GOCI. Since the MODIS bias is very low, this could help with the systematic bias you get in your CMAQ results.*

Reply) We agree that the biases of GOCI data can be reduced with other observations such as MODIS data. However, unfortunately, the spatio-temporal kriging system used in this study can only treat the data on a regular spatial grid (e.g. the observation points are not changed with respect to time) collected by constant time interval. This is mainly due to the fact that the calculations of the semivariograms require some computational cost. For example, the MODIS case, where time and space of the observation are daily varied, needs additional time lags and fine spatial grids and/or some mathematical assumptions for the calculations of the semivariograms. In this context, we think that this technical issue would be a bit beyond the scope of our current work. We appreciate your comment. Please, refer to p. 26, lines 583-587 in the revised manuscript.

*Comment) When comparing GOCI vs MODIS (Fig 3) you are comparing a 6km vs a 10km retrieval, so resolution might play a role in the differences seen in performance and spatial*

*coverage. MODIS collection 6 provides 3km AOD, so it would be interesting to include this product as well to "bound" in some way the GOCI resolution.*

Reply) We changed Fig. 3(a), using MODIS Collection-6 3km AOD products. We also modified Fig. 3(c) to show the hourly coverages from 1 March to 31 May, 2012. Please, see pp. 9-10, lines 209-218 and also check out modified Fig. 3.

*Comment) In several places of the text the authors state that kriging is used instead of data assimilation (e.g., beginning of section 3.1 and section 2.3). However, the kriging is a data processing stage to enhance the GOCI data and is not related to data assimilation (i.e., combine model and observation). Actually, one could use the output from the kriging into a data assimilation system. What you are replacing by data assimilation is the way you change the model concentrations (e.g. assumption of a perfect observation, choice of observation operator and control variable). Please correct this throughout the text.*

Reply) We removed the sentences p. 5325, line 11 in the original manuscript and modified the relevant sentences. Please, see p. 17, lines 393-394.

*Other comments*

*Comment) Page 5319, Line 11. SeaWIFS is no longer operational since 2010 so I wouldn't include as a product that can be used*

Reply) We removed SeaWIFS from the sentence.

*Comments) Page 5322, Line 18. It is not clear what the numbers mean, hours? Configuration*

*index? Please rephrase*

Reply) We rephrased the sentence. Please, see p. 8, lines 170-172.

*Comment) Section 2.2.1. Please clarify that the GOCI vs MODIS comparison is done before kriging*

Reply) We modified the sentence in Sect. 2.2.1. Please, check out p. 9, lines 206-209.

*Comment) Fig. 3. How is spatial coverage from GOCI computed? Do you use a single GOCI time for a similar time of the Terra overpass? Or you average several GOCI frames? Please explain this in the text to better understand where the differences in spatial coverage come from.*

Reply) We changed the Fig. 3 to clarify how the spatial coverage from GOCI was computed, showing hourly-based spatial coverage. Please, check out the caption of Fig. 3.

*Comment) Section 2.3. The kringing fill the gaps of the GOCI data. Is the kriging able to fill all of the gaps? If not, how do you determine if there is enough data to fill a gap. If you don't fill a gap, then this column in the CMAQ is not modified, right?*

Reply) The ST-kriging methods can fill almost all of the gaps of GOCI data with some degree of uncertainties (e.g. the kriging variance explained in Sect. 3. 1). In this study, only the gaps having kriging variance (KV) less than 0.04 were used for preparing the initial conditions. Therefore, AOD columns having kriging variance greater than 0.04 were not used

in the initial conditions. Please, refer to p. 18, lines 419-412. We also added a figure in Appendix A, reflecting a suggestion from reviewer #2. Please, check out p. 29, lines 654-656 and Fig. A2.

*Comment) Page 5329, Lines 21-23. I'm having problems understanding this CV choice. You assume $SO_4$ and OA is the same only for the increment? Or you make $SO_4$ and OA be the same in the final model concentrations? You also mention something about the surface, so you don't do anything aloft? Please clarify*

Reply) First, the ratios of surface $SO_4^{2-}$ concentrations and surface OA concentrations were calculated at the selected model grid points where the ST-kriging AODs have the KVs less than 0.04. Then, OA concentrations through all model layers were multiplied by this ratio, and the surface OA concentrations were the same as the surface $SO_4^{2-}$ concentration. Because OA concentrations (from the surface to the top layer) were changed, initial AOD fields were also changed by Eq (1) in p. 14. Then, the differences between the ST-kriging AODs and modified background AODs (observational increments) were used to update the initial $SO_4^{2-}$ and OA concentrations corresponding to the contribution of $SO_4^{2-}$ and OA AODs to the modified initial AOD fields. Please, note that to prepare concentrations of CVs above surface, background modeling-derived vertical profiles and size distributions of aerosol species were used for converting 2-D AOD to 3-D aerosol mass concentrations in all of the STK cases. We rephrased pp. 16-17, lines 366-372.

*Comment) Page 5329, Lines 27. CV #4 balances $SO_4$, $NO_3$ and $NH_4$, but do you do anything with OA? Table 3 says you do change it.*

Reply) CV #4 also made equal changes in the $SO_4$ and OA concentrations at the surface layer, which are the same as CV #3. Please, see p. 16, lines 370-372.

*Comment) Page 5336, line 26-27. It can greatly influence composition but you show that for $PM_{10}$ and AOD differences between CVs are not large. Please rephrase this sentence.*

Reply) We rephrased the sentence. Please, see p. 25, 576-578.

Response to Reviewer #2

First of all, we would like to thank you for your constructive comments and suggestions. Based on the comments from two reviewers and the executive editor, we have tried to address the issues raised by both reviewers in the revised manuscript. We would also like to respond to the comments point-by-point below. The reviewer's original comments are in italics. Our responses to reviewer's specific comments are put in a normal font. The added or modified parts in the revised manuscript are highlighted in a red color.

*General comment*

*Comment) It is obvious that the model has the strict negative bias. The statistics analysis (Table C1630 4 and 5) shows large negative bias (MFB=-98.7—113.2% and NMB=-62.6—70.0%) in the noSTK case for AOD and $PM_{10}$. Figure 7 shows that the model forecast $PM_{10}$ concentrations less than half those of observations. The bias in the noSTK case is constant and so (negatively) large that this is not difficult that the STK cases obtain better scores (especially in short time forecast). In fact, negative biases are still found in the STK cases. I wonder if the application of the STK could show any modification in horizontal distribution (e.g., shape of $PM_{10}$ plume) or timing of peak of $PM_{10}$ and AOD.*

Reply) We have tried to address the issue of whether the applications of the STK cases could also improve the performances of hindcast runs for less polluted conditions where the negative biases were relatively small compared with those in high polluted conditions. We have also tested whether there are any improvements in horizontal distributions and/or timing of peaks of $PM_{10}$ and AOD. We found that there were positive effects not only on reducing negative biases but also on improving spatial distributions. Please see p. 8, lines 176-178; p. 17, lines 385-386; and pp. 24-25, lines 552-567. We also added Fig. 10 in the revised

manuscript and one table in the supplement (refer to Table S1)

*Comment) Other concern is the necessity of sophistication of model. The model bias rapidly negates the beneficial effect from modification of initial condition (i.e. the application of the STK), indicating that refinements of the model (e.g., emissions and parameterizations) might be much more effective to improve the $PM_{10}$ forecast than the replacement of initial condition.*

Reply) Yes, we completely agree with your opinion. The effects of the initial conditions cannot last long, which implies that both/either the CTM parameterizations and/or emissions is/are not currently perfect. We are working on these issues in separate frameworks. We believe that improving the parameterizations and/or emissions are beyond the scope of this manuscript. Please, consider positive impacts on improving the performances of hindcast runs for 48 hours. Please, refer to pp. 24-25, lines 552-567 in the revised manuscript

*Minor comments*

*Comment) 1. (P5317, L13) The STK can use observed data more than what? Please specify.*

Reply) We rephrased the sentence. Please, check out p. 2, lines 33-35 in the revised manuscript.

*Comment) 2. (P5318, L21) What was low (< 60%)? Which score? Please specify.*

Reply) We rephrased the sentence. Please, see pp. 3-4, lines 68-69.

*Comment) 3. (P5322, L10) The system is planned to be used as operational system. Is 12-*

*hour forecasting enough long for early warning of $PM_{10}$?*

Reply) We have been aware of this problem. In the revised manuscript, we showed the performances of hindcast runs for 48 hours. Please p. 8, lines 176-178; p. 17, lines 385-386; and pp. 24-25, lines 552-567. We also added Fig. 10 in the revised manuscript and one table in the supplement (refer to Table S1).

*Comment) 4. (P5323, L20) Did you use SSA and FMF in this study?*

Reply) We have not used SSA and FMF data from GOCI sensor due to their relatively large uncertainties. As soon as these data with high quality become available, we will try to include the SSA and FMF in our study.

*Comment) 5. (P5325, L11) Please describe the mathematical linkage. In this study, the kriging is used for temporal and spatial interpolation (compensation). Data assimilation techniques are based on statistical estimation and used for integrate model and observation.*

Reply) We corrected relevant sentences. Please, refer to p. 17, lines 393-394.

*Comment) 6. (P5325, L25) Did you use the STK to smooth existing data points? Did you apply the STK only to fill missing points?*

Reply) We constructed AOD fields at 12:00 LT, using the GOCI data at 09:30, 10:30, 11:30, and 13:30 LT over the model grid point where any data was not spatially coincided with GOCI observations. Therefore, all of ST-kriging AOD data were weighted averaged (or

smoothed) by surrounding observation data.

*Comment) 7. (P5326, L4) The STK method can use an AOD field at 12:00 derived from AOD fields observed at 9:30, 10:30, 11:30, 12:30 and 13:30. On the other hand, 3D-Var and OI are sequential technique. They can assimilate the AOD fields at observed times, sequentially. From this aspect, 3D-Var and OI can use more observation data points more than the STK method.*

Reply) We rephrased the sentences. Please, check out p. 12, lines 267-271.

*Comment) 8. (P5326, L1) 13:30 is better than 01:30.*

Reply) We corrected it.

*Comment) 9. (P5326, L10) "Uncertainties" is inadequate. "Biases" is proper.*

Reply) We changed it. Thank you for your comment.

*Comment) 10. (P5326, L15) The system is going to be used operationally. The light computational cost of the STK is one of advantages for operational forecasting. However, to calculate AOD field at 12:00 through the STK, we must wait for GOCI AOD data at 13:30. Is this OK for the schedule?*

Reply) One of the objectives in this study is to find the best combination for the PM forecast over Northeast Asia. Therefore, we chose to include GOCI data at 13:30 for the enhancement

of STK AOD quality. However, we completely agree that including GOCI data at 13:30 (or 12:30) should be considered carefully according to the available computational resources.

*Comment) 11. (P5328, L20) How did you modify vertical profiles of concentrations from the GOCI AOD?*

Reply) We rephrased relevant sentences to clarify the method for determining aerosol vertical profiles. Please, check out p. 16, lines 370-372.

*Comment) 12. (Section 3.1) Can you show AOD fields before and after the application of the STK? Exhibiting how the STK compensates missing regions due to clouds and high reflectance will reinforce readers' understanding. Other concern is if there enough GOCI data to fill missing points. Sometimes, we found large missing fields in satellite-measured AOT maps due to clouds. Can the STK method with five GOCI AOT fields fill the missing field completely? If not, how do you replace the initial condition of the forecasting.*

Reply) We added a figure showing the AOD fields before and after the application of the STK method. Please, check out Fig. A2 and p. 29, lines 654-656. We also added a sentence for clarifying how to fill out large missing fields. Please, see p. 18, lines 409-412.

*Comment) 13. (P5330, L19-21) Do these sentences mean that when the model has a considerable negative bias, the STK is favorable comparing with other DA methods?*

Reply) We changed the sentence in Sect. 3. 1. Please refer to p. 17, lines 393-394.

*Comment) 14. (P5333, L11-16) This result means that excepting sea-salt and BC from CVs leaded to better results? In the other words, including sea salt and BC in CVs degraded results?*

Reply) According to the background model simulations, the contributions of BC and sea-salt to AOD and $PM_{10}$ were less than 5 % over the domain. Therefore, we assumed that the effect of selecting BC and sea-salt for CVs would be relatively small. In particular, for high pollution episodes, the contributions of both species to total mass in Seoul Metropolitan Area (SMA) are minimal. However, we agree that it is necessary to consider those species as CVs in some episodes such as large biomass burning and in some areas nearby coastal regions.

Response to executive editor,

First of all, we would like to thank you for your constructive comments and suggestions. Based on the comments from two reviewers and the executive editor, we have tried to address the issues raised by both reviewers in the revised manuscript. We would also like to respond to the comments point-by-point below. The reviewer's original comments are in italics. Our responses to reviewer's specific comments are put in a normal font. The added or modified parts in the revised manuscript are highlighted in a red color.

*Comment) … please note that for your paper, the following requirements have not been met in the Discussions paper.*

*"– All papers must include a model name and version number (or other unique identifier) in the title."*

*Please add a name and a version number for the numerical system in your revised submission to GMD.*

Reply) We modified the title of the article as follows:

GIST-PM-Asia v1: Development of a numerical system to improve particulate matter forecasts in South Korea using geostationary satellite-retrieved aerosol optical data over Northeast Asia.

We also rephrased a sentence in Sect. 1. Please, check out p5, lines 103-104.