This study describes the production of an aerosol reanalysis for the period 2011-2015 through the assimilation of quality assured MODIS observations into the MASINGAR mk-2 global aerosol transport model. The data assimilation scheme used is a 2D-Var method. The paper is well written. Authors describe with clarity the different components of the system: the aerosol transport model, the assimilation method, the observations and the observation operator. The quality of assimilation system is evaluated through internal checks based on analysis and first guess departures from assimilated observations, while the quality of the reanalysis product is evaluated with independent observations, and compared to a control experiment which has been run without data assimilation. I think that the paper is suitable for publication after addressing the following comments.

We thank a reviewer for careful reading our manuscript and for giving useful comments. We have deliberately and considered your comments. We believe that we have made adequate corrections and answers to your comments. In revised manuscript, the changes are highlighted by yellow markers.

Specific comments:

1) I am concerned about your background error covariance (BEC) matrix, and it would be good if could add some more comments about them:

a) Can you please justify your choice of using in equation 26 a normalized temporal standard deviation in AOD? Other studies have estimated background error covariances using independent observations or difference of forecasts at different lead times.

We got an idea from the NMC method for the estimate of BECs in the current version. In the NMC, variations of forecasts at different lead times are used. In this study, we used variations of forecasts at different times ($\hat{\sigma}_{\mathrm{FR}}$), because, for the reanalysis use, we do not perform the forecasts with different lead times. However, comparing with variations of forecasts with different lead times, distribution of variation forecasts at different times becomes to be smoothen and makes shape of aerosol plumes obscured. This is the reason that we used the estimate expressed by Eq. (26). We added the following text and use “the horizontal structure of AOD field at the assimilation time” instead of “the flow-dependent structure of AOD” to convey our intentions more correctly.

“The fraction on the right-hand side of Eq. (26) indicates that the standard deviation is normalized by the mean value. The standard deviation derived from 31-day AODs leads to relatively smooth variation fields. To introduce the horizontal structure of AOD field at the assimilation time into the background error covariance, the normalized standard deviation is multiplied by the forecast AOD.”

b) Can you please justify your choice of expressing the flow-dependent component of the BECs by the forecast AOD (P12, L11-12), and comment on possible drawbacks?

Thank you for your question. Please refer the previous reply (reply to the comment #1a) for question about “flow-dependent”, and I would discuss about the drawbacks in the reply to the next comment (reply to the comment #1c).

c) Your background error covariances are quite large, in particular (by construction) for high simulated AOD values. Are you not this way over-fitting too much the assimilated observations?

Large BEC for high simulated AOD values might be a common problem in the aerosol assimilation even with the NMC and ensemble based-methods, because high simulated AOD often leads larger variation (spread).

The chi-square test and the sensitivity experiment in which the background error covariances were uniformly decreased by 60% shows that although the background and observation errors were persistently overestimated, the balance between background and observation errors was well-balanced and stable (see reply to the comment #6). The sanity test with the MODIS AOD and examination for the first guess (FG) confirm that the analyses caused few over-fitting. We added the texts in the revised manuscript as follows:

“Both RA and FG from the additional experiment obtained worse scores in the validations with MODIS and AERONET AODs than the standard experiment. For the χ^2 value, the additional experiment shows much larger variation (standard deviation). These results imply that although there were the persistent overestimates of background and observation errors, they were well-balanced and stable in the standard experiment.”

One of possible drawbacks for the estimate of BECs is the analyses at megacities. Please refer reply to the next comment (comment #1d) for this point.

In the current version of JRAero, the background error was in proportion to the forecast AOD (i.e., first guess AOD; see Eq. (26)). This means that the background errors where the model did not predict aerosols became so small. Therefore, in that situation, the assimilation could not reproduce the major aerosol event because of the small background error (also see reply to the comment #5 of reviewer #4). This should be another possible drawback. We added discussion about this situation in Section 5 (Future direction) as follows:

"In the current version, the background error was in proportion to the forecast AOD (Eq. (26)), and became small where the model did not predict aerosols. Therefore, the analysis could not reproduce aerosol events that satellites observed but the model failed to predict (e.g., dust storms and biomass burning) due to the small background error. The ensemble-based estimate of the background error considering uncertainty in emissions will bring better analysis for this situation."

d) Studies by Rubin and colleagues have shown the importance of flow-dependent BECs in aerosol data assimilation, and how ensemble methods can best estimate the temporal evolution of the background error covariances. As you say in the section on "Future directions", you plan to use better BECs in the future, and I think that you should mention this point also when you describe your assimilation method, in the conclusions, and in the abstract.

Thank you for your suggestion. We added the following text and use "the horizontal structure of AOD field at the assimilation time" instead of "the flow-dependent structure of AOD" to convey our intentions more correctly, as answered in the previous comment.

Better estimate of BECs is our important direction to next version of the analysis and ensemble-based methods are quite powerful to obtain the flow-dependent BECs. We added the following texts that mention that the ensemble-based estimate of background error covariance has the possibility to overcome problems included in the current version of the reanalysis referring the most recent results by Rubin et al. (2017) in Section 3.3.2 and Section 5 (also see reply to comment #10 of reviewer #4).

Section 3.3.2: "The probability of a successful retrieval can be reduced during high-AOD events (Lynch et al., 2016); thus, fewer available satellite observations over megacities during high-AOD events may also account for the negative biases in RA AODs. Rubin et al. (2017) applied an ensemble-based assimilation method to NAAPS and found that flow-dependent error covariances estimated by ensemble simulations utilized the AERONET AOD efficiently and brought better analyses at Beijing and Kanpur. Sophistication of the background error covariance and assimilation of additional observations have the potential to improve the analyses at the megacities."

Future directions: "At megacity and mountain sites, assimilation provided limited improvement, and positive and negative biases remained in the reanalysis. A plausible reason is the coarse model resolution, which is insufficient to resolve high-AOD events around megacities and local terrain effects in mountainous areas. Therefore, we plan to rerun the reanalysis with a finer resolution and check the performance of the model. Re-examination of the background error by an ensemble-based method (Yumimoto et al., 2016) and assimilation of additional observations (e.g., the AERONET AOD) have the potential to improve the analyses at the megacity sites (Rubin et al., 2017)."

Rubin, J. I., Reid, J. S., Hansen, J. A., Anderson, J. L., Holben, B. N., Lynch, P., Westphal, D. L. and Zhang, J.: Assimilation of AERONET and MODIS AOT observations using Variational and Ensemble Data Assimilation Methods and Its Impact on Aerosol Forecasting Skill, J. Geophys. Res. Atmos., doi:10.1002/2016JD026067, 2017.

This paper constitutes a comprehensive report on the current version of JRAero. From the validation studies, we found that the estimate of BECs seems to be large and should be improved. However, that is beyond the scope of this paper. Therefore, we added more detailed discussion about the update of the BECs in the Evaluation results and Future directions rather than the Conclusions and the Abstract.

2) Could you show or comments on whether your analysis is smooth in space and time throughout the day?

a) Given that observations are really assimilated in a given location at most once a day, your reanalysis could have jumps from the one time step where observations are present to the others. If so, you could mention in your conclusions, and in the future outlook, that observations in a future reanalysis should actually have a good coverage. Regarding this issue, Lunch et al. (2016) showed the importance of model tuning in particular when there are areas not covered by assimilated observations, which are therefore highly impacted by the model first guess.

Your suggestion is important with aspect of temporal continuity of reanalysis product (the non-diagonal BECs prevent the reanalysis from spatial jumps caused by the spatial limited MODIS AOD due to clouds and snows). We added the following texts in the revised manuscript.

Future direction: "The limited temporal coverage of MODIS AOD (once a day) might cause temporal jumps or discontinuities in the reanalysis. Lunch et al. (2016) suggested that the importance of model tuning in particular when there are areas not covered by assimilated observations. To obtain a better coverage of assimilation data is important for the future development. Smoother techniques (e.g., 4-dimensional variational method (Yumimoto and Takemura, 2013; 2015) and ensemble Kalman smoother (Schutgens et al. (2012)) are also useful."

b) When using localization it is important that analyses in neighbour regions share assimilated observations. How do you deal with this aspect when you divide the model space in local regions?

The local region is set to each model grid. The analysis increment of AOD ($\delta\boldsymbol{\tau}^a$) at each model grid is solved independently with observations (more correctly, innovations ($\boldsymbol{\tau}^o - \mathbf{H}_l\boldsymbol{\tau}^f$)) included in the local region (see Eq. (19)). Also see reply to comment #3 by the reviewer #1.

We modified the texts as follows:

"We introduced a localization technique used in LETKF to the system that divides the model space into local regions using a prescribed localization scale. The localization technique solves the analysis increment of AOD at each model grid with observations included in the local region independently (see Eq. (19)), reduces spurious error covariance with distance and enables parallel processing to be used to reduce computational cost."

3) It might be good to stress a bit more that assimilated MODIS observations are not an independent set of observations to validate the reanalysis (e.g., in the abstract P1, L12-15, in the introduction P4, L14-15, . . .), but they can be used to perform valuable sanity checks on the assimilation system.

Thank you for your suggestion.

We added the texts in the revised manuscript.

Abstract: "Comparisons with MODIS AODs showed that the reanalysis showed much better agreement than the free run (without assimilation) of the aerosol model and improved under- and overestimation in the free run, thus confirming the sanity of the data assimilation system."

Introduction: "Section 3 focuses on the sanity check and evaluation of the reanalysis product with MODIS AOD and independent observation data."

Tittle of the Section 3.2: "Sanity check by MODIS AODs"

4) P12, L27: Can you justify the use of a diagonal observation error covariance matrix?

Ideally, we should utilize correlation in the observation error covariance. However, accurate estimate of the correlation remains challenging (even in NWP assimilation). We need to put a lot of effort into this problem both for production of reanalysis and aerosol forecasting.

We add some texts in the revised manuscript.

"We assumed that the observation error covariance matrix (**R**) was diagonal because of the difficulty of accurate estimate of the correlation and assigned uncertainty of AOD provided by the NRL-UMD MODIS AOD product to the diagonal component of the observation error covariance matrix. The uncertainty includes empirical estimation of observation error and representativeness error based on variability of the L2 dataset (Zhang and Reid, 2006)."

5) P13, L18: Don't you think that it would be more correct to use in the validation AERONET AODs which are the closest in time to model results, without doing any averaging of them?

The best way is to compare observed AOD with modeled AOD at the observed time. However, that is unfeasible, because the temporal resolution of the AERONET AOD is much finer than that of model output. Additionally, the AERONET AODs include fine temporal variations (less than 1 hour) that the global model cannot reproduce. The aim of the reanalysis is to provide 6-hourly, daily and monthly aerosol fields, not reproduce the fine temporal variations. Therefore, we averaged the AERONET AODs into 1-hour bin, and compared with the model results.

6) P14, L23-27: What about the accuracy of the analysis when decreasing the BECs? Rather than persistent overestimation of the observation error, could not be that your current BECs simply do not describe well the structures of your background model errors?

Both RA (analyzed) and FG (first guess) from the sensitivity experiment in which the background error covariances were uniformly decreased by 60% shows worse agreement with the MODIS AOD than the main experiment. RMSD (correlation coefficient) for RA and FG during 2011–2012 was degraded from 0.05 (0.93) and 0.098 (0.73) to 0.07 (0.87) and 0.10 (0.71), respectively. The independent validation with the AERONET AOD also shows that the sensitivity experiment obtained worse scores comparing with the main experiment. For chi-square value, the sensitivity experiment shows larger variation compared with the main experiment (0.027 versus 0.0059). These results imply that although the background and observation errors were persistently overestimated, the balance between background and observation errors was well-balanced and stable.

"Both RA and FG from the sensitivity experiment obtained worse scores in the validations by MODIS and AERONET AODs than the main experiment. For the χ^2 value, the sensitivity experiment shows much larger variation (standard deviation) compared with the main experiment. These results imply that although there are the persistent overestimates of background and observation errors, they were well-balanced and stable in the main experiment."

7) Could you also add in Figure 6 the analysis increments (analysis minus first guess) and comment on them? This would allow you to identify better local systematic corrections made by the observations and hence discuss model bias, while the differences of Figure 6c (analysis minus free run) are affected also by corrections that might happen somewhere else and are transported.

We added increments (analysis minus first guess) as Figure 6d, and re-labeled analysis minus free run as difference (Fig. 6c). Comments on them were added in the manuscript.

"Figure 6d exhibits the distribution of the increment (RA AOD minus FG AOD) that is derived from the 5-year average of the modifications by the assimilations (i.e., $\delta\boldsymbol{\tau}^{a}$ in Eq. (19)). The increment shows lower amplitudes and different distributions in several regions (particularly in the downwind regions of aerosol sources) compared with the difference (Fig. 6c). It is because that the difference is affected by transport of the modifications after the assimilations and the increment (FG AOD) takes into account accumulation of previous assimilations. The effect by the accumulation also appears as much better statistics of FG AOD (e.g., lower root mean square error (RMSD) and mean fractional bias (MFB)) comparing with FR AOD (Fig. 7)."

8) P15, L26 and P19, L24: I think that you can only verify with a certain degree of independence your 6h forecast, and not a forecast up to 24h. Don't you produce a 6h forecast from each analysis step?

We employed the MODIS observation as the assimilation data. MODIS is onboard Low Earth Orbit (LEO) satellites and make observation on the same region once a day. What we want to mean by "short-term (6–24 hour) forecasting" is a short-term model simulation after the analysis to the next analysis (when the next MODIS observation is available) in a certain region. We corrected the text to "the short-term forecasting until the next analysis (until the next MODIS AOD is available)".

9) As you show in P15, L34 and P16, L1, the mean bias and the MFB can have a different sign. Therefore in the discussions in section 3.3.3 (You mean Section 3.3.2?) should you not add the value of the mean bias, or use the wording "mean fractional bias" when commenting on the MFB values at the various stations?

We used “a positive MFB value” and “a negative MFB value” in the discussions in Section 3.3.2 when commenting on the MFB values in the revised manuscript.

10) The first few months of simulations (clearly a spin up period for the data assimilation) could have been removed when estimating the statistics reported in the various tables...

As mentioned in the Subsection 2.4, we made the 3-month spin-up (October–December 2010) of free run and then performed the 5-year reanalysis. The spin-up period was excluded from the validations. However, to make a more perfect validation, we also should make few months of spin-up with data assimilation. In the next version of the reanalysis, we will introduce a spin-up with data assimilation (e.g., few months of spin-up by free run, few months of spin-up with data assimilation, and then reanalysis period).

Technical corrections:

We thank a reviewer for careful reading our manuscript and for giving useful correction. We corrected in the revised manuscript.

1) P3, L33: The NASA aerosol reanalysis is called Modern Era Retrospective analysis for Research and Applications Aerosol Reanalysis (MERRAero)

Now, aerosol variables are included in additional file collections in MERRA-2. (https://disc.gsfc.nasa.gov/datareleases/merra_2_data_release)

2) P5, L7: Please change the position of the word “respectively” to avoid ambiguity

3) P7, L5: Please change “land-cover factors vegetation” with “land-cover factors for vegetation”

4) P7, L8-L13: Please change “grid(s)” with “grid cell(s)”

5) P7, L21: Please add the air density and the gravitational constant which are also present in equation 9

6) P7: Please use a consistent symbol for particle diameter d_s or D_s between equation 9 and 10

7) P11, L3: Please remove one of the K in equation 20