

Geosci. Model Dev. Discuss., referee comment RC2
<https://doi.org/10.5194/gmd-2022-30-RC2>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on gmd-2022-30

Anonymous Referee #2

Referee comment on "A preliminary evaluation of FY-4A visible radiance data assimilation by the WRF (ARW v4.1.1)/DART (Manhattan release v9.8.0)-RTTOV (v12.3) system for a tropical storm case" by Yongbo Zhou et al., Geosci. Model Dev. Discuss., <https://doi.org/10.5194/gmd-2022-30-RC2>, 2022

Review of "A preliminary evaluation of WRF (ARW v4.1.1)/DART (Manhattan release v9.8.0)-RTTOV (v12.3) in assimilating satellite visible radiance data for a cyclone case" by Zhou et. al, 2022, submitted

General comments

This paper deals with the data assimilation of visible satellite radiances (also referred to as reflectance). Such observations are relatively new to the numerical weather prediction (NWP) community, since the absence of fast and accurate forward operators made their operational exploitation inconceivable for many decades.

In a set-up using the WRF model, the DART data assimilation framework providing EAKF and RHF filters plus the RTTOV-DOM forward operator, both numerical data assimilation cycle experiments and single observation experiments are conducted in an OSSE framework.

Thereby, relevant aspects related to reflectance data assimilation are shown and outlined, including non-linearity, non-Gaussianity, observation weight related to thinning length scales and update frequencies. Further, limits of reflectance data assimilation are discussed, e.g. the lack of vertical height information of the observed clouds and ambiguities in cloud phase and particle size distribution. While generally interesting, from what I can see, most of these aspects have already been discussed by Scheck et. al, 2020 in the context of the COSMO + KENDA system. I therefore strongly advice that the authors include a detailed discussion of how their findings relate to the previous study, to which extent they confirm or contradict previous findings and which parts of their analysis are uniquely novel.

While key sensitivities of the data assimilation cycle are discussed and evaluated very detailed with respect to analysis verification, it would be of great practical relevance to understand also how such sensitivities relate to the forecast quality and the forecast error growth. The assumption that a better analysis leads to a better forecast is by no means trivial particularly when dealing with cloud variables whose properties violate the mathematical assumptions of filter algorithms (linearity and Gaussianity), and which are prone to model biases and compensating model errors of the NWP model. Further, the analysis ensemble mean, which is mostly verified in this study, is not a physically consistent state so that it is not obvious to which extent the analysis error reduction related to the ensemble mean is beneficial for the accuracy of the individual forecast ensemble members which are initialised from the respective analysed model states. I therefore suggest to add results related to forecast verification (i.e., forecast verification of experiments 1-6). A discussion of the sensitivities of forecast quality and error growth for cloud variables, but also for other model parameters like temperature and humidity would add significant value to the publication and could provide guidance to colleagues for preparing visible radiance data assimilation also in an operational context.

While this manuscript contains some very interesting material, to be suitable for publication it requires some substantial changes. Please find below.

- The key research questions should be stated more clearly in the paper overview and at the beginning of each paragraph. It should be motivated why the different investigations are done and which question is followed therein.
- Could you motivate why you do have chosen to assimilate the observations in physical variables of radiance rather than reflectance which makes it much more easy to estimate how optically thick the clouds are? And which seems to be much more convenient in the data assimilation community?
- The fundamentally new methods and findings for the research community should be pointed out in a more precise way
- Forecasts should be added to the cycled data assimilation experiments to show the sensitivities of both analysis and forecast error to the update frequencies, thinning length scales, chosen DA filter and so forth
- The figures have to be revised fundamentally. Please increase the picture resolution and enlarge lines, labels and axes. It is very hard to differentiate between the triangles, diamonds and squares even when viewing the figures with a large zoom factor.
- The questions below should be answered in the text
- Please consider the comments below to improve the text

Specific comments

- Title: Code versions do not have to be part of the title of a scientific paper, suggestion: "A preliminary evaluation of visible radiance data assimilation for a cyclone case"

Sec 1.

- Please add some comments on challenges and potential of all-sky data assimilation
- Please elaborate bit more on why you think it is interesting to assimilate visible satellite radiances. Which forecast impact do you expect? What is different from IR or MW all-sky data assimilation?
- Please note and correct in the text: RTTOV is no forward operator, but rather a collection of forward operators (radiative transfer package)
- The goal of the publication and its value to the scientific community has to be stated clearly at the end of the introduction section

Sec 2.1

- Are the grid spacings between nature run, control run and DA runs equivalent?
- Why do you use higher resolved LBCs for the control / Da experiment than for the nature run / truth? An OSSE should represent the difference between real atmosphere (which is much higher resolved than a forecast model) and a forecast model. Thus, I would rather use the higher resolved LBCs for nature run / truth. Do you conduct short-range forecasts as well? Do the LBCs introduce the cyclone to the model domain across the lateral boundaries or does it fully develop within the model domain? If the first is the case this may be problematic.
- Why do you use the ensemble mean of the nature run as truth? The resulting model state is physically inconsistent between the variables
- How do you exactly produce the synthetic observations? What is the role of the 2km-original AGRI observations? Do you use them as observation locations for the synthetic observations? How do you assign observations and model grid points? Do you first assign model grid points to observation locations (which will lead to one grid point being assigned to many observations at a 15 km – 2km scale difference) and then apply thinning at the observation locations? Do you interpolate and how or do you use nearest-neighbour? Is it right that you simulate the observations based on the truth which is equal to ensemble mean? The next reasonable step would be to perturb the synthetic observations based on an estimated observation error distribution. Do you do that and how?
- Please explain the set-up of the OSSE in much more detail following the questions posed above. At the current state, it is hard to understand and not reproducible.
- Please explain in the text if you assimilate any other observations or only visible radiances

Sec 2.2

- For better understanding of the sensitivities of visible radiances on model variables please add a discussion of the subgrid-scale cloud variables which are presumably input to the forward operator. How is that realized in your system? How are subgrid-scale clouds parameterized? This may be important to discuss and understand the potential detrimental impact on non-cloud prognostic variables

- Please define “cloud water path” in the text (appears for the first time in Figure 2)

Sec 2.3.2

- It will help the reader if you explain in more detail the goal of the pointwise DA experiments. What do you want to show here?
- What is the meteorological situation at the 4 points that you have chosen in the domain? Please motivate why you have chosen exactly these points
- Would it be an idea to refer to the pointwise data assimilation experiments as single observation experiments? This is a term that seems to be more convenient in the data assimilation community
- Please clarify that the cycled experiments are also run in OSSE set-up. This does not seem to become clear in a moment.
Do you also run forecasts or only DA experiments?
VIS data assimilation strongly interacts with forecasts, so I would ask you to verify forecasts as well to show if the DA is successful in terms of forecast impact
- Table1: Please explain the variable names or rather write the physical variable names, e.g. QICE=cloud ice mixing ratio.

Sec 3.1.

- I would suggest to refer to that kind of experiment as “single observation experiment”
- Please explain what you mean by “cloud water path”. Is that only vertically integrated liquid water? Or liquid or ice water? Or the sum of the two? May I ask why you do not show the ensemble distribution of cloud water and cloud ice in the left-hand plots since you compare to them on the right-hand side?
- Please motivate more for the reader why you assess and show that kind of experiment. What is your goal with that? Please make that very clear to explain the key issue with ambiguities in visible radiances, i.e. total water mass, cloud phase, effective radii, vertical position, multiple layers with different phases
- In Figure 3, do you work in observation space on the left-hand side and in model space on the right-hand side? So you try to figure out to what degree the model variables are improved if the analysis is drawn towards the observation in observation space? Could you clarify that in the text, please?
- I could not distinguish between the lines in Figure 3. Please draw fatter lines, fatter axes, fatter labels. Use different colours rather than symbols because it is unfortunately really hard to distinguish between them. It is confusing that the diamond sign means “truth” on the left and “analysis” on the right
- Since GMD is a journal dealing with models and the key goal of data assimilation is better meteorological forecasting skill it will add great value to the paper if you discuss not only the statistical properties of the single-obs experiments, but also explain the meteorological situations: e.g. in Figure b1/b2) in the truth run we have an optically thick water cloud. However, the model shows an ice cloud lying over a water cloud. Data assimilation draws radiance towards the truth and is well able to enhance the water cloud. However, the false alarm ice cloud is also enhanced, since a) VIS observations can only constrain vertical integrals, b) there seems to exist spurious

correlation of the ice clouds in the background ensemble with the observation and we cannot vertically localize that due to missing information on the cloud top height and vertical cloud extent. Clarify that VIS observations are sensitive to the cloud water mass in the column and the particle size distribution. Ice clouds consist of few big particles and are typically much more optically thin than water clouds that consist of many small particles. Try to explain a bit more the microphysical connection between clouds and radiance to the reader

- Please add this kind of meteorological detail to clarify the potential and limits of the VIS data assimilation in view of specific meteorological cloud situations
- In Scheck et. al 2020 such kind of case study has been performed as well. Please reference and compare your discussion to the results found there

Sec 3.1

- Do you assimilate any observations additional to radiance observations here?
- Do you assess first guesses or analyses or forecasts here?
- Please improve Figure 7. It is very hard to distinguish the symbols. Use fatter labels and axes and fatter lines and maybe different colors for the different experiments
- Clarify if the cycled experiments are OSSE experiments or if you assimilate original satellite observations
- You should add that the no-clouds situation in the ensemble is referred to as "zero-spread" problem
- You state that RMSE, MAE etc. measure different aspects of accuracy than FSS. Please explain which aspects and why you want to assess them
- To which degree do the results found for the six different experiments hold for forecast impact?
- Please replace "fake" correlations by "spurious" correlations
- Please elaborate a bit on why less observations and less frequent update intervals may lead to better overall forecasting skill
- Do you have suggestions why VIS DA may have detrimental impact on dynamic and thermodynamic prognostic variables? What is the role of the NWP model in here? What is the role of subgrid-scale clouds? Could you add a plot illustrating the detrimental impact please? Please discuss this issue in a bit more depth and to debate potential fixes
- I am unhappy with your term "thermodynamic" variables. Wind is no thermodynamic variable. Maybe you could refer to the variables you want to address as "non-cloud" variables
- You state that VIS radiance data do not have an apparent dependence on "thermodynamic" variables. I do not agree with that. Clouds are advected by wind fields so that cloud position error is correlated with wind field errors. Clouds depend on temperature and humidity. Subgrid clouds are typically parameterized in terms of grid-scale humidity fields.

Sec 3.2.1

- At first glance it has been unclear to me which kind of model state you verify here. Is it

linear analyses, nonlinear analyses, forecasts, first guesses? Please clarify both in the text and in the figures

- Please explain why it is interesting to assess the temporal evolution of MPI / MPE of effective radius
- Please use the term "false alarm" clouds instead of "fake" clouds
- Please replace "updated in negative ways" by "analysis increments with wrong signs / wrong magnitudes"
- Does the underestimation of the effective radius come along with an overestimation of radiance, i.e. a positive radiance bias?
- Is the effective radius input to the RTTOV-DOM forward operator? Could you motivate why you show Figures 9 and 10, please? What is shown in Figure 9?
- Since you seem to assimilate the VIS observations over high terrains in China – do you have any quality control included that rejects observations that may be mixed up with snow or ice?
- How do you set the observation error? Next to the number of observations assimilated which is determined by thinning length scales and update frequencies you can control the weight of the observations by choosing a larger observation error. It would be nice if you included that in the discussion and potentially even in your experiments

Sec 4.

- What is the main message that you want to present to the reader related to your discussion on observation rejection?
- Please revise Figure 11. It is very hard to distinguish between the triangles, squares etc. and to recognize them. Maybe you do not have to display every time step in the plot
- I'm wondering why the departure between first guess and observations increases over time. In a healthy DA system, the average first guess error tends to decrease with increasing number of DA cycles. Do you have any explanation for that?
- Why do you use quality control to control the number of assimilated observations? In my view, quality control should sort out erroneous or non-representative observations. You should rather control the number of observations by horizontal localization, thinning and superobbing as well as observation error
- If I understand correctly, the outlier threshold acts on the first guess departure of the ensemble mean. Large first guess departures typically occur when clouds are missing or you have location error of clouds which tends to happen quite often for clouds- and precipitation-sensitive observations. In my opinion, being able to correct for such location errors or false alarms in the analysis is of particular importance. Why do you choose to sort out these observations? Would it be possible to inflate observation error in those cases rather than not assimilating the observations at all?

Sec 5.

- Updating only cloud variables for NWP forecasts is not practical for operational NWP. Do you have other suggestions to deal with potential detrimental effects on forecast skill of temperature, humidity, wind etc.?

- Please replace “modelling experiments” by “data assimilation experiments”. As far as I understood you do not show forecasts or their verification.
- Please replace “cloud simulations” by “cloud states”. You mostly look at analysis ensemble means which is not the same as a “simulation” or a “free model state”
- If you verify analyses the term “cloud forecasting skills” may be inappropriate
- Why would increased model grid spacing lead to a more nonlinear relationship between radiance and LWP / IWP? Or do you mean nonlinearity due to resolved convective processes? Please clarify
- Please elaborate a bit more on ideas for further research based on your found results in the conclusion. What did you learn and what do you suggest to deal with the found problems? What do you suggest for operational VIS radiance data assimilation?

Spelling, grammar, typos

- L-8: *there are great potentials in assimilating*
change to: *there is great potential related to assimilating ...*
- L-31: unique cloud information complementing the one contained in IR and MW data
- L-40: direct data assimilation critically depends on observation operators
- L-47: Method for Fast Satellite Image Synthesis (MFASIS)
- L-53: single-scattering method for SW radiative processes
- L-59, 61, 64: kill superfluous “the”: assimilated GOES-9 VIS radiance, is ensemble-based methods
- L-93: revise end of sentence “A nature run is..”
- L-140: Other parameters not explicitly mentioned are set to default values.
- L-186: “To demonstrate the basic ability of the DA scheme.”; it is unclear what you mean by that. Do you mean “to demonstrate the basic technical functionality of assimilating visible radiance data by employing EAKF”?
- 281: double “the”
- L-302: could clearly suppress false alarm clouds
- L-324: through spurious correlation between VIS radiance
- L-339: At the initial cycling step, convective initiation occurred in the nature run
- L-348: non-cloud state variables obtain analysis increments with wrong sign such that analysis error is increased compared to first guess error
- L-363: false alarm clouds
- L-364: much closer
- L-396: by the DART system
- L-408: by the detrimental effects on analysis error of the non-cloud ...
- L-420: such as the Atmospheric Motion Vector
- L-426: life cycle, i.e. the intensification and decay processes of a cyclone
- L-428: the adjustment of CWP
- In general: replace thermodynamic by non-cloud