Response to Reviewers' Comments on the original submission of

Assimilation of GPM-retrieved Ocean Surface Meteorology Data for Two Snowstorm Events during ICE-POP 2018

We would like to thank the reviewers and the editor for their thorough analysis of our manuscript and for suggesting changes that will help our paper to have a better quality. We have reproduced the reviewer's comments below — our responses appear in blue. The page and line numbers refer to the trackchanged manuscript.

Comments from Reviewer #1:

General comments

This is the second-round review of the manuscript. The authors have made efforts to address my previous concerns. Nevertheless, I am not satisfied with the revisions made so far. The revised manuscript is still lacking a clear goal, and too many unnecessary details are in fact very distractive. I still do not think this is a paper ready for publication in its current form. The detailed comments are listed below.

Thank you so much for the comments, and thank you so much for recognizing our effort in revising this paper. We have revised the objectives in the introduction section and the conclusion section. Please see Page 4-5 Line 128-132 and Page 17 Line 540-542.

Since the current paper is particularly written for the special issue "*Winter weather research in complex terrain during ICE-POP 2018 (International Collaborative Experiments for PyeongChang 2018 Olympic and Paralympic winter games)*" of GMD and the main purposes of the special issues are (https://gmd.copernicus.org/articles/special_issue10_1112.html):

"1) to document the scientific findings on the winter weather during the forecast demonstration project
2) to share scientific knowledge on processes of winter weathers that have been investigated with unprecedented dense observational networks,
3) to share current status and improved knowledge of forecasting of winter weathers, and
4) to document new retrieval and quality control methods of the operational and advanced instrument"

In this paper, we have introduced the surface meteorology data that has been retrieved from the GMP microwave observation. This dataset was created particularly in support of the ICE-POP 2018 field campaign. The main goals of this paper are to introduce this ICE POP dataset, to explore the application of this dataset through data assimilation, to demonstrate the influence of this dataset, and to examine whether or not the assimilation of this dataset is able to improve the forecasts of two heavy snowstorm that has been occurred during ICE-POP 2018.

In the revised paper, we have removed unimportant details as the reviewer suggested, yet kept other necessary details in order to support our key result and conclusions.

Major comments:

1. In summary, the manuscript basically shows that there are significant changes made to the background and the forecasts in the temperature, precipitation and moisture flux divergence can be improved. Also, it shows that the changes made to the background can be different between Feb and

Mar cases. Speculations are due to the different kinematic and thermodynamic processes involved. As a case study, we cannot draw any statistically robust conclusions from the results. Yet the manuscript failed to provide any physical link on how the changes to the background can lead to the changes in the forecasts. The authors should try identifying at least one major contribution of the DA towards the better predictions. For example, which one of the T/Q/WSD fields is more important to the better predictions? Can the increased/reduced temperature analysis trigger more/less convection-induced precipitations in certain areas, which leads to better precipitation forecasts? What exact kinematic or thermodynamic process is the key for the Feb and Mar differences? etc… Sensitivity experiments are strongly recommended to explore the impact of DA for your single case study.

We agree with the reviewer that better understanding of the physics on how the data assimilation influenced the winter storm forecast is very important and of great interest. Therefore, in our previous manuscript, we have added Fig. 10 to illustrate how the data assimilation influenced the low level (925 hPa) moisture flux convergence and moisture transport. We have found that the overestimate of precipitation in CTRL_Mar was improved in DA_Mar because of the weaker low level moisture convergence and moisture flux.

A thorough understanding of the physical mechanisms of the winter precipitation is a hot topic, and it is also one of the main research goals of ICE POP 2018 project. However, it requires a great amount of research and cannot be covered completely in the current paper. Forecast of winter precipitation in South Korea is particularly challenging due to its unique geographic features and large-scale circulation. Orographic effect, synoptic circulation, and air-sea interaction can all have important roles and interact with each other. Sensitivity studies and in-depth investigations on different dynamic, thermodynamic, and microphysics factors and processes (both large scale and local effect) need to be taken place to examine how, by how much, and for how long these factors can influence the winter precipitation and the model forecast. Some to the publications in this special issue tried to tackle this topic. For example, Gehring et al. (2020) indicated that a warm conveyor belt (WCB) had significant impact on the microphysics processes for the heavy snow storm on 28 February 2018. The WCB had an important role in generating supercooled liquid water and enhancing aggregation which led to rapid precipitation growth. Kim et al. (2021) showed that wind pattern and topography affect the microphysical processes for snowstorms over northeast part of south Korea. Strong wind shear and turbulence were the cause of the intense riming process that dominates over the mountainous region. For the coastal region, aggregation becomes more important than riming in generating precipitation.

On our end, further investigation will be performed to better understand how data assimilation modified these factors and processes which lead to a better forecast. These research activities require a great amount of work that will be conducted in our future studies related to this project.

Reference:

Kim, K., W. Bang, E.-C., Chang, F. J. Tapiador, C.-L. Tsai, E. Jung, and G. Lee, 2021: Impact of wind pattern and complex topography on snow microphysics during ICE-POP 2018. Atmos. Chem. Physi. 21, 11955–11978, https://doi.org/10.5194/acp-2021-128

Gehring, J., A. Oertel, É. Vignon, N. Jullien, N. Besic, and A. Berne, 2020: Microphysics and dynamics of snowfall associated with a warm conveyor belt over Korea. Atmos. Chem. Phys., 20, 7373–7392, https://doi.org/10.5194/acp-20-7373-2020.

2. L313-341: I see the authors have reduced the corresponding discussions. Yet I still don't see the

value of this O-B vs A-B comparison. Yes, DA can perform differently, and increment/influence can be large/small depends on the methodology/configurations. But in this current paper, the authors never compared any DA schemes other than a single scheme in two cases. I don't get what to expect from these comparisons. Again, it is a common knowledge that a successful DA is supposed to drag the analysis towards observations. Then, what else can we get from here? Are the increments balanced? Can the corrections be validated from other independent obs? I think my point is that you can use a few words/O-B/O-A statistics to support the idea that DA is working properly, but no need to use the figure and too much paragraph. Otherwise, try dig something new/unexpected.

Thank you for the comments and thanks the reviewer to agree with us on some of our points. In the revised manuscript, we tried our best to shorten this paragraph. Since the dataset contains 3 types of observation, we used a few sentences to describe the impact of each data type. We do want to keep the RMSD numbers at the end of the paragraph because this was recommended by reviewer #1 in the previous round of review.

Since this paper focuses on case studies for assimilating a new dataset with high resolution regional simulation, it is always a good idea to check the background, the observation innovation (O – B) and the analysis increment (A – B) and see whether or not the data assimilation is effective (similar plots can be found in numerous past studies such as Xiao et al. 2007, Bi et al. 2011, Chen et al. 2021, etc). The RMSD numbers can provide a general sense of how the model fields respond to data assimilation. However, the plots of O – B and A – B provide the horizontal distribution of the obs innovation and analysis increment. By comparing them, it is verified that the influence of the data is producing a reasonable spatial distribution.

Reference:

Xiao, Q., Kuo, Y., Sun, J., Lee, W., Barker, D. M., & Lim, E. (2007). An Approach of Radar Reflectivity Data Assimilation and Its Assessment with the Inland QPF of Typhoon Rusa (2002) at Landfall, *Journal of Applied Meteorology and Climatology*, *46*(1), 14-22.

Bi, L., Jung, J. A., Morgan, M. C., & Le Marshall, J. F. (2011). Assessment of Assimilating ASCAT Surface Wind Retrievals in the NCEP Global Data Assimilation System, *Monthly Weather Review*, *139*(11), 3405-3421.

Chen, S., Shih, C., Huang, C., & Teng, W. (2021). An Impact Study of GNSS RO Data on the Prediction of Typhoon Nepartak (2016) Using a Multiresolution Global Model with 3D-Hybrid Data Assimilation, *Weather and Forecasting*, *36*(3), 957-977.

3. L342-359: I don't think this exp design is a good approach to show how long or how much the DA impact lasts. First, CTRL_Mar is not initialized from the same background as DA_Mar, so the differences are not purely from DA. Second, since you are not verified against truth/nature run, the differences are not informative. In an extreme example, if the whole system is shifted from south to north (say, closer to the truth), yet you can still get the same average profile, and the difference won't grow if the parallel system moves similarly. Current configuration only shows that the DA forecast is being balanced after roughly 4~6 hours. Instead, you should use the statistics from the RMSE of Fig. 8 (or some other observation verifications) to do the impact study.
PS: I also still don’t understand why the authors want to use only the South Korean Surface Analysis dataset for verification. Isn’t it better to use the datasets with more confidence (e.g. examined before) for verification purposes?

We have updated Fig. 7 for a different time. In the current manuscript, the computation was conducted for the forecast after the 1st data assimilation cycle (0600 UTC 7 March 2018). DA_Mar assimilated both GPM-retrieved data and conventional data at 0600 UTC 7 March 2018 and CTRL_Mar assimilated only conventional data at 0600 UTC 7 March 2018. CTRL_Mar and DA_Mar used the same background field, so they should be a fair comparison.

Thank you so much for the suggestion of using the RMSD between the observational data and DA_Mar as in Fig. 8 to examine the magnitude and lasting period for data impact. However, this may not be a plausible method. Firstly, unlike OSE or OSSE, a real case study like this does not possess a nature run which is error free and has the same variables with perfect (or seemingly perfect) data. Secondly, the nature run of OSE or OSSE is at the same horizontal and vertical grids as the sensitivity experiment. However, in our case, the South Korean Surface Analysis that we used only covers continental South Korea (as shown in Fig. 8a). And the observational data that we have assimilated are over the ocean (as shown in Fig. 1). The analysis fields shown in Figs. 6 and 12 clearly indicated that most impact was found over ocean near the observed location and very slight impact was found over Korean Peninsula (another reason for keeping Figs. 6 and 12 in the paper). Therefore, if we use observational data over continental South Korea to do the evaluation on data impact, we will miss the vast area over the ocean where the majority impact occurred. This will likely to hold for the forecast within the first few hours of integration. That will not be a fair evaluation.

In this study, we choose to use the South Korean Surface Analysis dataset for independent verification because this dataset is in-situ data based on observations from the enhanced Automatic Weather Station network during the ICE POP 2018, hence have a better accuracy in snowfall measurement than the remote sensing observations. The dataset has been verified against the AWS observations over South Korean. According to Ryu et al. (2021), the methodology for this dataset has been examined by many past studies (e. g., Ozturk et al. 2014, Guo et al. 2011, Li and Parker 2008) which showed that this method may surpass the traditional limits of sampling theory and recover missing spatial or temporal data. Recent researches (e. g., Demissie et al. 2021; Chen et al. 2021) also adopted and cited Ryu et al. (2019).

Reference:

Chen, H., S. Sheng, C.-Y. Xu, Z. Li, W. Zhang, S. Wang, and S. Guo, 2021: A spatiotemporal estimation method for hourly rainfall based on F-SVD in the recommender system. Environmental Modelling & Software, 144, 105148, https://doi.org/10.1016/j.envsoft.2021.105148.

Demissie, T. A., and C. H. Sime, 2021: Assessment of the performance of CORDEX regional climate models in simulating rainfall and air temperature over southwest Ethiopia, Heliyon, 7, e07791. https://doi.org/10.1016/j.heliyon.2021.e07791.

Guo, D., H. L. X. Qu, and Y. Yao, 2011: Sparsity-based spatial interpolation in wireless sensor networks. J. Sensor, 11, 2385–2407.

Li, Y.Y., and L. Parker, 2008: Classification with missing data in wireless sensor network. IEEE SoutheastCon 2008, pp. 533–538.

Ozturk, S., T.-Y. Yu, L. Ding, R.D. Palmer, N. A. Gasperoni, 2014: Application of Compressive Sensing to Refractivity. IEEE Trans. Geosci. Remote Sens. 52, 2799–1809.

Ryu, S., J. J. Song, Y. Kim, S.-H. Jung, Y. Do, and G. Lee, 2021: Spatial Interpolation of Gauge Measured Rainfall Using Compressed Sensing, Asia-Pacific J. Atmos. Sci., 57, 331–345. https://doi.org/10.1007/s13143-020-00200-7.

4. L360-377, L380-390 & others: Throughout the manuscript, there are too many tedious fact descriptions and number counting without clear goals, which make the readers lose focus. Unless there is any physically important point, the detail listing is very distractive and unnecessary. For example, what's the major point here other than DA_Mar has less RMSE? Why are those spots in Taebaek Mountains or Sobaek Mountains so important? Can they physically influence the storm predictions somehow? You should really connect the details to specific improvements to make a meaningful point, otherwise, just summarize them with one or two sentences.

We have revised the discussion in L360-377, L380-390, and other places as well. Since most of the readers of GMD may live in countries other than South Korea, the inclusion of necessary knowledge about the South Korean cities and geography might be very helpful for them, so we kept some details that are important to support our conclusions.

5. L435-449: I still do not think it is a good way to discuss the quality of GPM precipitation data here. It is off topic, and you should evaluate this irrelevant (to the goal of this manuscript) dataset in another paper with more thorough analyses.

We have removed the discussion on the GPM IMERG data, please see Page 14, 16, and 17. The plots related to GPM observation have also been removed from Figs. 8 and 10.

Minor comments:

1. L115: 2m -> 2-m.
Corrected.

2. L119: Please try to reduce the use of however.
This has been changed to “on the other hand”.

3. L122: part -> partly
Corrected.

4. L301-302: How to deduce this implication from the figures or percentiles?
This was not deduced from the figure, this sentence has been revised.

5. L303-310: Did you assume the observations to be unbiased? And again, what do these numbers mean? Model is biased?
Yes, there was no bias correction adopted in the current data assimilation procedure. From the O – B scatterplots, it is shown that the departures between observation and model field are skewed to different extent. In future studies, we will examine the data for a longer period of time. When we have more robust conclusions about the bias of this GPM-retrieved surface observation data, we may consider adding a bias correction procedure before or during the minimization for a better data assimilation result (this point has been mentioned in the conclusion section on Page 18 Line 569-576).

6. L338-341: Directly calculate O-A and compare it with O-B will be more straightforward than A-B.
We would prefer to keep the A – B because it is an import data assimilation term, the so-called analysis increment.

Topical Editor decision: Reconsider after major revisions
by Chiel van Heerwaarden

Comments to the author:
While the paper has improved, the reviewer is not happy with the revisions. I have to agree with the reviewer that the paper still needs improvement in clarity and focus. I am willing to provide you this opportunity and advise you to take into account the suggestions of the reviewer in detail. Most important: please set the goal of the paper clearly, and try to remove irrelevant information.

Thank you so much for the comments and thank you so much for providing us the opportunity. In the updated manuscript, we made significant revisions on the research goal. Please see the introduction section (Page 4-5 Line 128-132) and the conclusion section (Page 17 Line 540-542).

We have also updated Figs. 8 and 10. In the new figures, we have removed the plots as well as the discussion related to GPM IMERG data quality (please see Pages 14, 16, and 17) as suggested by the reviewer.

We have updated Fig. 7 with the plots for the first data assimilation cycle to address the reviewer's concern.

We have also revised the result section from Page 10 to 17 to make the manuscript more concise.