

Response letter to RC2

Dear RC2,

We thank you for the review comments, we believe that the manuscript is strengthened after addressing them. In this letter the original comments are found with our response in red color.

General comments

This manuscript presents a regional reanalysis system for snow water equivalent. It is named CARRA-Land-Pv1 and is based on a multi-layer snow model (ISBA, with 12 vertical layers) and a Local Ensemble Transform Kalman Filter to assimilate snow depth in the model. This new reanalysis is evaluated against a reference analysis, CARRA, which also used ISBA but as part of a larger model that also includes all the other components of a numerical weather prediction model. Therefore, CARRA and the new CARRA-Land-Pv1 proposed by the authors share the same snow model and the same meteorological forcings (generated by CARRA) but they differ in 1) the assimilation method itself and 2) the observation sets that are assimilated (partially, as some observation stations are used by both systems). The two systems are also compared against observations from six completely independent snow pillow stations that are not assimilated in CARRA nor CARRA-Land-Pv1. The authors show that, in general, their proposed new reanalysis outperforms CARRA.

I believe the topic is of great interest to the readers of The Cryosphere. However, there are a few methodological aspects that need to be clarified, and I also found minor linguistic issues that need to be corrected. Specifically, details regarding the perturbation of meteorological forcings need to be added, and the LENKF method needs to be explained better. Finally, I could not find the resolution of the grid anywhere. I apologise if I have missed it, but if it is not mentioned, it should.

Indeed the grid resolution was missing in the original manuscript. We add this in the revised version. Note also that the configuration of ISBA is different in CARRA (single layer) and CARRA_Land_Pv1 (multi-layer).

Specific comments

1. Introduction: too much emphasis on mountains

We think that mountains are particularly important in terms of seasonal snow as they receive more snow than lower-laying regions and are subject to larger uncertainties in snow estimation products.

2. Line 118: The terms in equation (1) are not adequately defined, as it is not indicated what « x » stands for. You should take what is written at line 159 (« where x represents the

ensemble control vector (...) a and b indicate analysis and background (...) » and place it on line 118 instead.

We thank the reviewer for pointing out this missing detail, we add definitions of x^a and x^b (these are not necessary ensembles at this point).

3. Lines 143-148: You mention on several occasions that the state vector has high dimensionality. It would be good to provide the reader with numbers regarding what size is small, what size is average, and what size is large. In addition, I think a more detailed explanation of the concept of localization is needed. In particular, how do you ensure that neighboring points remain correlated (preserving the spatial structure of snow depth and SWE) if data assimilation is performed independently point by point?

We agree that the wording is vague, we include an example of a typical number and how it is reduced using the local filter. We also improve the description of the LETKF and how information is distributed spatially.

4. Figure 1: I cannot understand this figure. I have read the text several times, I still don't understand the remapping. I understand that the purpose is to better represent the spatial uncertainty of precipitation, but I don't understand the method and Figure 1. Is it possible to modify Figure 1 to make it clearer? Maybe provide an example with numbers and/or real grid cells? The Appendix also did not help.

We understand that Figure 1 could be difficult to interpret, we will thus suggest the following figures to replace it.

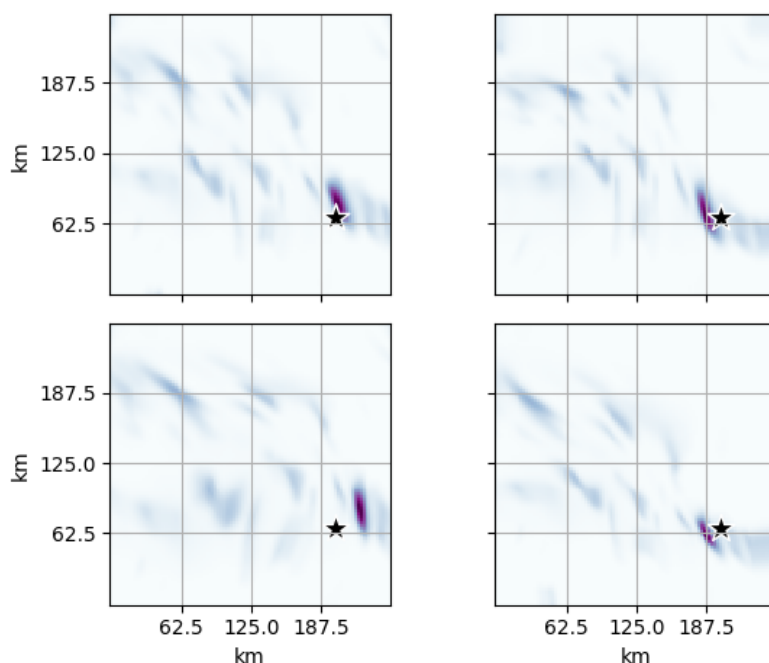


Fig 1R a) Precipitation field on model resolution with a small (3 member) ensemble. Upper left panel shows precipitation rate from the CARRA dataset, upper right and lower panels

show ensemble realizations after remapping is applied. The star marks a reference point, for example an observation site.

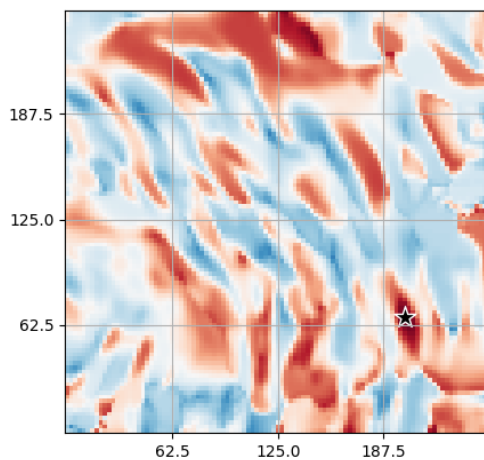


Fig 1R b) Colors indicate ensemble correlations between the values in the point marked by star and at the respective pixel. Red colors indicate positive correlations. If more snow than the model was measured at the marker, snow would be increased/decreased at the red/blue areas. Note that localization is not applied in this figure.

We hope these figures are more intuitive as they represent model derived precipitation. We also clarify the text describing the method and extend the appendix with more details.

5. Section 2.4: Please provide a table that indicates the range of perturbations used for each variable, as well as any relevant equations.

We include perturbation parameters and extend the description of the perturbation methods.

6. Line 310: You find the ensemble to be only marginally better than the average of all its members. However, I'm wondering about the potential users of that new reanalyse. Do they want only deterministic estimates of snow depth, or is there value for them in having access to the full ensemble and information about the uncertainty? I think this should be discussed.

We thank the reviewer for this comment, and believe this point is not covered sufficiently. We extend the discussion to include the aspect of ensemble uncertainty. Specifically we discuss the fact that the CRPS is better relative to the MAE of the ensemble mean for the OBS-ONLY-CARRA observations. This suggests that the ensemble is quite narrow and depends more on observation error than background error. On the other hand, away from observations, the ensemble has more spread indicating a less certain estimate.

7. Line 316: I'm not sure « climatology » is the right word for geographically close stations. I understand that they might have differences in their observation records, but can we really talk about different climatology?

We replaced "climatology" with "local conditions and systematic differences"

Linguistic comments

8. Abstract, lines 2-4: « merge information from the two sources » is not clear, as the « two sources » could be understood as either « prediction systems » and « historical reanalysis » or « observations » and « physical laws in models ». The reader has to logically deduce that you are referring to the latter, but it is not clear from the way the sentence is written. I suggest reformulating.

We replace “information from the two sources “ with ”information from model estimates and observations“ to clarify.

9. Abstract line 11: Is it possible to replace « relatively large » by something more precise?

We agree that “relatively large” is not precise and believe “multivariate” is a better word here.

10. Line 28, remove the comma between « models » and « is »

Done

11. Line 155: This group of equations should be numbered

Done

12. Line 167, there is an « s » missing in « represents»

Done

13. Lines 382-383: The sentence « In emission modelling, the use of statistical models for subgrid processes has been used to encounter similar problems » Needs to be reformulated. I suggest « In emission modelling, statistical models for subgrid processes have been used to counter similar problems

We adopt the suggested sentence.

14. Line 325: the word « stations » is missing between « observation » and « available »

Add “stations”