Response to Reviewer 1

Black: reviewer comments (in the report or in the attached files)

Blue: authors response (the line numbers refers to the trackchange version)

I would like to thank the authors for the substantial modifications to the manuscript that address all my major comments in a convincing way. I believe that the proposed modifications improve the scientific quality and the impact of the paper, in particular:
- the inclusion of the openloop run in the results
- improved figs 4 and 6 of the revised manuscript.

I no longer have major comments, and I am confident that the following minor points can be addressed without considerable effort. I am pleased to provided an annotated version of the revised manuscript with further details on the following points, and additional technical comments.
(1) l79: description of the k-local methodology is wrong.
(2) the inclusion of the openloop in the results is really an improvement, but I'm a bit surprised by the high spatial granularity of the openloop compared to the spatial correlation of the applied perturbations. The explanations l. 370-375 are not fully convincing.
(3) The rigor of the writing has been substantially improved by the revisions, but I find that the style could be improved towards more concision/structure (e.g. very long Sec 4.1 could be splitted, Sec 4.3 more concise).

I find a lack of references to the snow data assimilation literature in the discussion, a point that I omitted in my first revision, so I will not point this as a reason for minor revisions. The authors might take this as an advice to increase the impact of their conclusions by framing them in a more general setting.

The authors would like to thank Reviewer 2 once again for his precious and precise comments. We tried to integrate all those comments in the new version of the paper. In the following, we go through each comment you made into the annotated pdf file. We have copied those comments here and addressed them one by one.

L73, fig 6 caption, 353, 398, 399, 430, 477, 478 : typos

All the typos have been corrected, thank you for kindly pointing them out.

L79 and comment (1) : Thanks for the reformulation, but this is wrong again, and this must be fixed. I acknowledge that my comment was not clear enough. I also acknowledge that my own paper (Cluzet et al., GMD, 2020) is not clear enough on that, and I want to apologize for that.

Anyway let's hope some formulas will make things clearer:
-> Considered point p1, ensemble E(p1)=[E(1,p1),E(2,p1),...,E(Nens,p1)]

-> Point where we have an observation p2, E(p2)=[E(1,p2),E(2,p2),...,E(Nens,p2)]
-> Correlation(p1,p2)= Pearson(E(p1), E(p2))

```
k-local algo:
for pj in observed_points
   compute C(j) = Correlation(p1,pj)
end for
selection_points = location_of_the_k_maximal_values_of_C(j).
```

=> selection_points is the set of locations where we have an observation and which exhibit the highest ensemble correlations with the considered location.

The sentence has been edited to better summarize the actual algorithm (l79-83)

Figure 4 : Thanks for the modifications on this figure. consider removing one plot (deterministic or openloop) to enable enlarging figures for a better readability.

The plots on Figure 4 were enlarged. It was possible to do so by changing the display rather than losing a panel. Keeping the five panels makes it easier to associates the maps on this figure with the variograms on the previous one (figure 3) and the different results for the five cases.

Figure 5 : What do the solid lines represent on this figure?

The caption of this figure has been modified and now reads: « Variograms of the particles for the different reordering strategies for February, 28, 2012; (a) ensemble open loop; (b) no reordering;(c) sorting; (d) Schaake Shuffle. Solid lines represent the variograms of the estimated SWE as displayed on Figure 3. »

L343 : At this stage, we don' t know what are the two particles. Please introduce them briefly in the text, not only in the figure legend.

The following sentence has bee added on line 360 To introduce the notion of particles: « The maps represents the state of two individual particles (#250 and #500). The two particles present very different patterns. This difference was expected, as particle #250 is located in the middle of the ensemble, whereas particle #500 gathers all the highest SWE values. Such differences illustrate that with sorting, the individual particles cannot be considered as potential SWE map scenarios. The sorting can rebuild the short-range correlation, which is necessary for the spatial particle filter, but the sorting can jeopardize the long-range correlation and the general spatial pattern. »

L344 : as particle #250 (it' s not obvious to the reader that the # of the particle refers to the position within the ensemble)

Figure 6 : Thanks for adding the openloop to this plot. But this raises a question here, that to my opinion, must be commented:

Since there is no resampling in the openloop, why is there such a spatial noise in Figs (a,b)? The granularity seems but more fine than the depicted 200km correlation radius (Eq. 2). Please comment.

This was changed as recommended

Fig 6 displays the behaviour of individual particles, not the weighted mean of the ensemble. As to the fine granularity, our belief is that even if the perturbations (on precipitation, temperature and SWE) are spatially correlated, they still introduce small random differences between neighbouring cells at each time step. SWE being a cumulative variable, those differences just cumulate throughout the winter. This effect disappears when averaging the particle, as it can be seen on fig 2, 4 ans especially 5. Fig 6 displays SWE at the end of February (close to the end of the accumulation season). A sentence has been added l 355-358.

L353-354 See my comment on Fig. 6. And explanation for the origin of the noise in the openloop is required.

See previous comment.

L360 You might consider replacing by simply 'the assimilation dates'.

This has been modified as suggested.

L366 : late February?

Yes, it has been corrected.

L397-398 : The decrease in MBE with respect to the deterministic run is also present without data assimilation (in the openloop) so this is not strictly thanks to DA.

You are absolutely right. This behaviour has to be associated to the deterministic vs ensemble configuration. The sentence (415-416) has been edited and now reads: « The shift from deterministic to ensemble open loop configuration as well as the data assimilation sensibly reduce the RMSE (panel a) and the MBE (panel b) …. »

L 399-400 it is not clear whether this difference is really significant (use ' seems' or other word).

This sentence has been modified and now reads (l418) as: « … the RMSE values seems higher than those obtained using the assimilation. »

L 399-400 Same here, not sure whether this is significant (and not the case for high RMSE values (>60mm))

This sentence was also modified and now reads (lines ???-???) as: « There seems to be only a slight decrease in RMSE when using the Schaake Shuffle. »

L403-410 : This explanation is welcome, but not here, as it cuts the flow. Consider moving to 3.3

The CRPS description has been moved to section 3.3, l 302-309

L415-416 : more common formulation: 'the ensemble is over-dispersive'

The term « overspread » was replaced by « over-dispersive »

L419 : There is a problem with this sentence.

The sentence has been edited and now reads (l 442-444) as: « According to those results, the simple sorting of the particle appears to be as effective as the more elaborate Schaake Shuffle to maintain the spatial structure of the particle. »

Authors’ response to interactive comments by Anonymous Referee #2

Black text: Reviewer’s comment

Blue text: Authors’ response after the 1st round of comments

Green text: Authors’ response after the 2nd round of comments (the line numbers refer to the trackchange version)

In the author’s answers the line numbers refer to the trackchange version of the revised manuscript.

This manuscript develops and evaluates methods to maintain appropriate spatial correlation when using a particle filter to estimate SWE for a model of southern Quebec. The application is clearly defined and the problems that can arise when using a particle filter with affordable size are described. As noted in the manuscript, it is well-known that particle filters can diverge and that the number of particles needed to avoid this divergence increases at least exponentially with the number of spatial degrees of freedom in the model. While the number of degrees of freedom in the SWE model being used is not explicitly investigated, indirect evidence is provided that 500 particles is insufficient for the application at hand.

The authors describe the problems that arise from the limited number of particles in terms of an inappropriate ‘scrambling’ of the resampled particles at adjacent grid points. Small differences in the impact of observations can lead to a resampled particle with quite different values at adjacent gridpoints that are believed to be strongly correlated for example. To address this, the authors propose to ‘reorder’ the association between the values and the particle index at each gridpoint. Their control procedure simply sorts the particles at each grid point so that particle 1 is associated with the smallest value of SWE for each model point. The new method proposed, referred to here as a Schaake Shuffle, uses a reference set of ‘particles’ for the model points, in this case generated by sampling periodically from a free run of the SWE model. This reference distribution might be referred to as a ‘climatological sample’ in some other earth system assimilation applications. This reference distribution provides information about the correlation structure of the free model. As a hypothetical example, it could include information that when SWE is higher in western Quebec it is usually lower in eastern Quebec. This type of information could be reflected in the particle filter assimilation after the use of the Schaake Shuffle reordering.

The authors would like the thank Reviewer 2 for their careful consideration of our work and mindful remarks and critics. We believe the integration of those comments have improved the overall quality of the manuscript. We describe below how those comments have been integrated into the revised version of the manuscript.

Since we did not have a detailed revision from Reviewer 2 for this round, in this second round we tried to rework with the same initial comments. Nevertheless, the authors may require more specifics to better understand the suggestions from Reviewer 2.

A number of metrics are used to assess different aspects of a basic particle filter, the naïve sorting, the Shuffle, and the free run (open loop). No error estimates are provided for most of the results so it is difficult to assess the significance, and this is something that should at least be discussed if it cannot be formally addressed. For instance, it is difficult to assess if the 4 different curves in Figure 2 are meaningfully different. It appears that the basic particle filter is an outlier, while the other three are indistinguishable, but appearances can be deceiving. It is even more difficult to assess the significance of differences in Figure 3. In this case, maybe the open loop is an outlier, but I have no idea at all whether there are any meaningful differences in the other three plots. Any guidance the authors could provide would be helpful. At this point, I would be forced to conclude that there is no evidence that the 3 particle filter methods produce significantly different estimates of pointwise SWE.

The purpose of the paper is the investigate the potential of particle reordering in a spatialized version of the particle filter. Some of the figures intend to exhibit the similarities and differences between the different implementations, especially in term of performances. Some other figures are used to show that the reordering does not completely perturbate the assimilation process. Former Figures 2 and 3 (now 3 and 4) belong in the second category. Those figures investigate the spatial structure of the finale SWE estimate (weighted average of the particles). So the similarities between the variograms is a desired behavior. Figure 3 (former Figure 2) is included to demonstrate that particle reordering does not perturbate the overall spatial structure which is made up by the snow model from the meteorological forcings. Two sentences were added l321-223 to emphasize this point: « This similarity is a desired behavior, as the spatial structure should be guided by the snow model and its meteorological forcings. Especially, Figure 2 demonstrate that particle reordering does not affect the overall spatial structure of the final SWE estimate. »

Figure 4 (former Figure 3) provides a more visual representation of the spatial structure. An additional discussion on the similarities between the maps was already added in the previous round of review (l332-339).

The scores provided in Figure 9 (former Figure 8) are error estimates. Nevertheless, the distributions of the scores summarize a lot of information as they aggregate errors calculated over time and space.

There clearly are meaningful differences between the 3 filter methods in some of the subsequent figures. Not surprisingly, the variograms in figure 4 are very different.

However, more evidence about which is better could be provided in the discussion. I suspect that a solid argument could be made that the Shuffle results are probably better, but this requires knowing something about the correct answer and the authors should try to discuss how that could be known with some additional clarity. Figure 6, perhaps the most important in the manuscript, clearly shows a difference in the correlations between the base particle filter and the two correction methods. This is important since the thesis was that the correlations were damaged by the particle filter. However, no solid evidence is provided of what the answer should be for this application. I believe that the correlation scales are probably much larger than the base case, but I am not convinced that they are long as indicated by the sort and Shuffle. Again, if the authors could provide some information about what the right answer is believed to be it could strengthen the argument for using one of the new methods.

As a matter of fact the right answer is not known. If a spatial observation of SWE over such a large domain existed, then it would not be necessary to use modelling and data assimilation. All currently existing spatialized SWE products rely on some kind of data assimilation (sometimes without even using in situ observations). That is the reason why we chose to evaluate our approach using cross-validation over the in situ SWE dataset. As shown on Figure 8 (former Figure7) the particle reordering helps to slightly decrease the RMSE of SWE estimation in validation sites (which data has not been assimilated) compared to an assimilation without reodering. However, the impact of reordering is clearly significant in terms of reducing (improving) the CRPS, which means that the uncertainty estimation provided by the particle dispersion is much better with reordering. This can be explained by the reduction of short-range noise visible on the different maps and variograms. Reducing the uncertainty associated with estimations at new sites (again, the data from validation sites is not assimilated) is already a great achievement as this kind of SWE estimation is aimed to be used in real time for flood forecasting. So, it appears that particle reordering helps reducing the range in which we believe the actual SWE to be for ungauged sites, which is a great help for forecasters.

Finally, the difference between the Schaake Shuffle and the naïve sorting is not clear. The only clear disadvantage of the naïve sorting is that it produces some absurd particles (as seen on Figure 6 panel f). Nevertheless it appears to be a good and simple fix to reduce short-range noise and maintain a coherent spatial structure in the particle filter.

The paragraph l439-444 is used to draw those conclusions. Table 2 was also added to provide a more numerical reading of Figure 8. The beginning of the conclusion was also reworked to better underline the advantages of particle reordering in general, before commenting on the differences between the two proposed reordering techniques.

In summary, the manuscript is very clear in its description of the application, the challenges to the particle filter, and the description of the new methods. It is less clear in providing evidence about the efficacy of the new methods. It is my somewhat uninformed opinion that the Shuffle has some nice features, but stronger evidence of this would be a nice addition.

We understand that most of the comments from Reviewer 2 are articulated around the description and comments related to the different figures in the results section. In particular, it is true that in most figures, several alternatives are characterized by similar distributions of scores (curves) or spatial patterns (maps). The purpose of some of the figures was to show that the Schaake Shuffle affects (and improves) the spatial structure of individual particles, but not the overall behavior of the filter. The reordering affects specifically the spread of the particles rather than their central behavior. Consequently, it is absolutely normal (and expected) that some maps or curves are similar. To help the reader, we modified the descriptions of the figures mentioned by Reviewer 2 and took time to describe whether or not it was expected to notice differences between the different alternatives, and why it was expected. See l361-365, 391-396, 398-402 and 455-461

As an end note, I would suggest that state-of-the-art ensemble filters, or localized particle filters that make use of some of the advantages of ensemble filters, could be a competitive alternative for this application. Ensemble Kalman filters derive much of their power by being able to approximate the most important covarying directions in model phase space which is what the particle filter is unable to do. Localizing the ensemble filter can result in high-quality assimilated estimates of covariance with ensembles much smaller than 500 members. For instance, work by Zhang https://doi.org/10.1002/2015JD024248 and references cited therein report on ensemble Kalman filter data assimilation using multiple types of observations in a comprehensive land surface model. Work by Poterjoy documents the power of localized particle filters using a theoretically-supported approach that could extend Zhang's results to deal better with the bounded nature of SWE, https://doi.org/10.1175/MWR-D-15-0163.1

Work by Anderson extends ensemble filters to bounded quantities like SWE while retaining the high-quality covariance estimates from localized ensemble filters https://doi.org/10.1175/MWR-D-19-0307.1

The authors might want to evaluate the efficacy of some of these methods for the Quebec SWE problem in their future work.

A more in-depth comparison of the difference between our particle filter and more standard ones has also been mentioned by Reviewer 1, so this point is discussed more in depth in our answer to Reviewer 1. And a new paragraph was integrated in the methodology section (lines 196-205). We want to thank Reviewer 2 for their bibliographical suggestions. The final suggestion by Reviewer 2 will also be taken into account for future work and is mentioned in the final portion of the conclusion (lines 535-537).

The reference from Poterjoy (2016) was also added in the introduction l79-83.