Authors' response to comments by Referee #1

Black text: Reviewer's comment

Blue text: Authors' response

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

1. GENERAL

I discovered and read this paper with a great pleasure. Data assimilation of snow observation, in particular using the Particle Filter (PF) has been a growing topic in the snow/hydrology modelling community over the last five to ten years. However, this algorithm suffers from two strong limitations when it comes to large scale, data-scarce problems (which are the general case in the field). The first one is the curse of dimensionality, which can be solved by localizing the PF. But this solution generates a second strong limitation: individual members of localized PFs exhibit discontinuous spatial fields and noisy spatial correlation fields which are often detrimental for the PF performance itself.

Capitalizing on a localized PF variant using a spatial interpolation of the PF weights from Cantet et al., (2019), the authors efficiently introduce an approach coming from the hydrological modelling, the Schaake Shuffle to solve for both limitations: the localisation mitigates the curse of dimensionality, and the Schaake Shuffle is used to enforce the spatial structures of a deterministic run in the individual ensemble members in an elegant way.

Overall, I find that this paper is well within the scope of the journal, and has the potential to be a significant contribution to the snow/hydrology data assimilation community and beyond, because they address an important problem in a convincing and elegant way. I must admit, however, that I'm not satisfied with the theoretical justifications for the use of the interpolation of the weights within the PF from Cantet et al. (2019) (IDWPF), instead of the classical localised (LPF, e.g. Farchi and Bocquet, 2018). I would ask for more justifications, or a comparison between the LPF and the IDWPF.

The scientific quality of the writing is sometimes lacking rigor, especially in the Sections 1-3. I would ask for a significant effort on that. Nevertheless, I really appreciated the compactness of the paper and the efficiency of the results and discussion sections, which make a very clear and straightforward demonstration of the author's point.

To wrap up, I see a lot of potential in this paper, and despite my concerns, I am very confident that the authors will be able to address my comments in a revised version of the manuscript. Please find below some details on my main comments. I'm pleased to provide an annotated pdf version of the manuscript with comments and suggestions throughout.

We would like to thank Reviewer 1 for his detailed reading and his constructive commentaries and side notes. There is no doubt his review will help us improve the quality of our manuscript. We will try to address the different comments and to describe how this will affect the manuscript.

2. MAJOR comments

(1)

The authors an interpolation of the particle filter weights (IDWPF) from rather than a classical PF localisation (LPF) (see the review from Farchi et al., (2018)). I agree that the basic idea as formulated l 71-73 (a good-performing particle at a close location must also be good locally) resembles the theory. But in the classical LPF, based on Bayes theorem, and assuming equal prior weights, the posterior weights are computed by multiplying the likelihoods of the particles at the different (independent) observation sites (Eq. 27 from Farchi et al., (2018)) and then, normalising. Here, the IDWPF averages the normalised likelihoods of the particles at the different observation sites. There is a substantial conceptual difference here: averaging instead of multiplying. I suspect that this results in less sharp and potentially suboptimal (more conservative) PF analyses. For example, if a particle is given a zero likelihood at one location, the LPF will reject it, while there is still a chance for it to survive in the IDWPF.

Moreover, the arguments in Cantet et al., and the present manuscript used to justify the used of the IDWPF instead of the LPF failed to convince me: the LPF also proposes a ' tapering' method to smoothly reduce the influence of the observations with the distance (Eqs. 28-29 from Farchi et al., (2018)).

To wrap up, I'm not saying that the IDWPF method is wrong, and should be rejected. I can actually imagine that it could be more resilient to outliers in the observations, and its conservativeness could be an advantage. There may also be references in the literature to serve as base for the IDWPF. But in the present form, the justifications provided to substantially deviate from the main theory, the LPF, are too weak for me. I would suggest to make a considerable effort on justifications, or even to compare the IDWPF with the LPF. The latter would have the benefit of significantly increasing the potential impact of this paper thanks to the use of a more 'orthodox' method.

To help with the discussion, I'm pleased to provide a toy example comparing the IDWPF and the LPF in the form of a jupyter code attached or publicly available at:

https://github.com/bertrandcz/da_notebo

oks

We would like to underline that reviewer 1 summarized very accurately and very efficiently the main theoretical difference between the two types of particle filters (PF). We would just like to stress out that IDWPF is also based and Bayes' theorem. At gauged sites, the posterior weights are also derived by multiplying the prior weights by the likelihood of the particles. Then the posterior weights are interpolated in space to be applied to ungauged sites (i.e. grid points). In a case where a null likelihood is given to a particle at all observation sites, then the posterior weights would also be 0 at all observation sites (through multiplication), so this particle would be given 0 weight everywhere. Then, if a particle has a 0 likelihood at a site A, and a non zero likelihood at a site B, interpolated weights for ungauged sites would vary from 0 around point A to higher values around point B. This is a desirable behavior, as in this situation the particle would carry some useful information in the area around point B. In the case of the LPF, assuming points A and B are in the same block, the particle would be given a weight of zero for the whole block. It is right to say LPF is more selective (sharp) as it would discard a particle as soon as one observation indicates it is unlikely (whatever the other observations around might indicate), while IDWPF would tend to conserve a particle as long as a single observation finds it likely.

We also believe that block design for a territory as vast as our study area would have been challenging, due to very different density of observations in different regions. Interpolation appears to be more flexible to deal with this situation.

A whole paragraph has been added in the methodology section (l 196-205) to underline those differences. Also, we integrated the comments made by reviewer 1 in the margins of our submitted manuscript (supplementary document from reviewer 1). We believe those changes are adequate to answer reviewer 1's preoccupations. For the sake of clarity, specific answers are provided as direct response to each comment with line numbering in the attached pdf. All line numbers in our responses are related to the trackchange version of the revised manuscript.

We also agree that a detailed comparison of the two approaches with a common application would constitute a fascinating piece of work. Nevertheless, as the main purpose of our proposed manuscript is the introduction of the Schaake Shuffle to try to mediate the problem of spatial discontinuities in the spatialized PF, so we believe the comparison of IDWPF and LPF would be out of the scope of the paper. On this point, we agree with reviewer 2 and will consider the comparison for future studies. We explicitly suggested it as a work of interest in the conclusion (l 518-522).

(2)

The scientific quality and rigor of the writing is often not satisfactory in its present form, in particular in Secs. 1-3:

☐ Even though there is no doubt that the authors have a deep understanding of the PF and its terminology, there is sometimes a lack of rigor in the terminology and approximations that make several sentences turn wrong, and arguments fall short (e.g. l. 54, l.207-208, l. 218-219).

Those comments have been integrated (see l 70, 253-255 and 266-267 in the trackchange of the revised manuscript respectively)

☐ Even though the arguments are there, the logical formulation behind certain paragraphs is too loose to be convincing. I have no doubt that the authors can address that, but a considerable effort is required here.

All comments made by reviewer 1 in the margins of the initially submitted paper (supplementary document from reviewer 1) were treated and integrated. Detailed answers are provided as direct answer to the comments in the attached pdf. All line numbering in our answers refers to the trackchange version of the revised manuscript.

☐ There are some approximations in the description of the literature which may induce the reader into having misconception on the references (e.g. l. 69, l. 227-228), and change the conclusions of some paragraphs

See l 86-87 (former l 69). As suggested by one of the comments from reviewer 1, the paragraph l269-279 has been condensed to avoid repeating the introduction. The comment about this paragraph is no applicable anymore.

☐ the observation dataset and study area (Sec. 2) must be described with more details and rigor.
A new figure (Figure 2) has been integrated in the « Study area and data » section to help readers better appreciate the data set as well as SWE characteristics in Quebec.

I'm pleased to provide several suggestions on these points in the attached pdf.

Again, the authors want to thank reviewer 1 for his detailed comments in the attached pdf. The text has been modified to integrate those comments, especially the ones regarding the theoretical justification of our approach.

3. MINOR COMMENTS:

(1) With the Schaake shuffle as used here, the authors enforce the spatial distribution of individual particles to match those of the model, (instead of historical observations): but by doing so, don't we miss the opportunity to adjust the ensemble to the observed spatial structures? I'd be curious about overlaying Figs 3. and 6 with observed (in-situ) values to assess that.

Our main challenge in this project was to identify the observed spatial structured. Unfortunately, completely spatialised observations of SWE covering our study area is unavailable. The few available products are already the result of assimilation processes. Therefore, it was intended to estimate a spatial structure from the available local in-situ observations, but most of the data set comes from manual snow surveys. SWE at different observation sites is not measured at the same date and estimating the spatial structure using data collected over a couple of weeks is tricky, as the snowpack keeps evolving in between measurements. Hopefully, the deployment of automatic sensors, for instance gamma ray monitoring sensors, will help us solve this issue in the future.

Observed values have been overlain on Figure 4 (former Figure 3) only. We estimated it would not bring additional information to plot that on Figure 7 (former Figure 6) as it displays some individual particles. The point of this figure is to appreciate the spatial structure of the individual particles, not to evaluate the individual accuracy of each of them (it would have required to plot the 500 particles).

(2) The abstract is lacking of a general scientific context to start with. I think that the start is too technical for the scope of TC, the notion of 'particles' should be introduced, and given the level of technicity of the paper, a brief sentence describing the particle filter might be required in the abstract, in particular to make the need for a reordering possible to understand. Details on the ensemble construction might be appreciated also. Would benefit from a more rigorous description of the observations and validation data sets.

We agree with reviewer 1 on this point. The abstract has been reworked to start from a more general perspective and provide more theoretical insights. In particular, attention was given on methodology and data description. The abstract is now longer but will help less specialized readers to better understand the scope of the paper.

(3) When computing global metrics (Secs. 4.2 and 4.3, Fig 7,8 and 9), it could be fair and interesting to compare the assimilation products with their ensemble counterpart without assimilation, not only the deterministic run (called 'open loop' in the paper). Ensembles are often favored compared to deterministic runs in terms of RMSE and SWE, and it would enable the authors to put into perspective the impact of the assimilation in terms of spread-skill and CRPS.

This is an excellent point. We also believe an "ensemble open loop" simulation is a better (and fairer) benchmark. New simulations were ran and added in Figures 3, 4, 5, 6, 7, 8 and 10 of the revised manuscript. The associated comments have been updated. As expected, the ensemble open loop is characterized by a lower RMSE that its deterministic counterpart, but the spread of the members becomes very large at the end of the accumulation season.

References:

Cantet, P., Boucher, M.-A., Lachance-Coutier, S., Turcotte, R., and Fortin, V.: Using a particle filter to estimate the spatial distribution of the snowpack water equivalent, Journal of Hydrometeorology, 20, 577–594, https://doi.org/10.1175/JHM-D-18-0140.1, 2019.

Farchi, A. and Bocquet, M.: Review article: Comparison of local particle filters and new implementations, Nonlinear Processes in Geophysics, 25, 765–807, https://doi.org/https://doi.org/10.5194/npg-25-765-2018, 2018.

https://github.com/bertrandcz/da_notebooks/

Please also note the supplement to this comment: https://tc.copernicus.org/preprints/tc-2021-322/tc-2021-322-RC1-supplement.pdf

Authors' response to interactive comments by Anonymous Referee #2

Black text: Reviewer's comment

Blue text: Authors' response

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.

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.

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.

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 anew 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).

©c Author(s) 2021. CC BY 4.0 License.

Large-scale snow data assimilation using a spatialized particle filter: recovering the spatial structure of the particles

Jean Odry[1], Marie-Amélie Boucher[1], Simon Lachance-Cloutier[2], Richard Turcotte[2], and
Pierre-Yves St-Louis[2]

[1]Department of Civil and Building Engineering, Université de Sherbrooke, Sherbrooke, Canada
[2]Quebec Ministère de l'Environnement et de la Lutte contre les Changements Climatiques, Quebec, Canada

Correspondence: Jean Odry (jean.odry@gmail.com)

Abstract. The use of particle filters for data assimilation is increasingly popular because of its minimal assumptions. Nevertheless, implementing a particle filter over domains of large spatial dimensions remains challenging as the number of required particles rises exponentially as domain size increases. A common solution to overcome this issue is to localize the particle filter and consider a collection of local applications rather than a single regional one. Although this solution can solve the dimensionality limit, it can also create some spatial discontinuity inside the particles. This issue can become even more problematic when additional data is assimilated. The purpose of this study is to test the possibility of remedying the spatial discontinuities of the particles by locally reordering the particles.

We implement a spatialized particle filter to estimate the snow water equivalent (SWE) over a large territory in eastern Canada by assimilating local manual snow survey observations. We apply two reordering strategies based on 1) a simple ascending order sorting and 2) the Schaake Shuffle and evaluate their ability to maintain the spatial structure of the particles. To increase the amount of assimilated data, we investigate the inclusion of a second data set, in which SWE is indirectly estimated from snow depth. The two reordering solutions maintain the spatial structure of the individual particles throughout the winter season, which significantly reduces the random noise in the distribution of the particles and decreases the uncertainty associated with the estimation. The Schaake Shuffle proves to be a better tool for maintaining a realistic spatial structure for all particles, although we also found that sorting provides a simpler and satisfactory solution. The assimilation of the secondary data set improved SWE estimates in ungauged sites when compared with the open-loop model, but we noted no significant improvement when both snow courses and the SR50 data were assimilated.

1   Introduction

The accumulation and melting of snow dominate the hydrology of Nordic and mountainous regions (Doesken and Judson, 1997; Barnett et al., 2005; Hock et al., 2006). In these regions, accurate information about snow water equivalent (SWE) is crucial for streamflow forecasting (Li and Simonovic, 2002) and reservoir management (Schaefli et al., 2007). Over large

territories or catchments, the spatial distribution of SWE can be assessed using remote sensing (Goïta et al., 2003) or snow
modeling (Marks et al., 1999; Ohmura, 2001; Essery et al., 2013). The advantages of snow modeling include the possibility to
25 generalize over large territories and ensure a complete temporal coverage with a resolution determined by the user. However,
as with any kind of model, snow modeling is subject to uncertainty and inaccuracy stemming from errors in the inputs or the
model structure itself. Given that SWE is a cumulative variable, these errors obviously increase in importance throughout the
winter, making SWE estimates highly uncertain at the beginning of the melting season.

   Data assimilation can be used to limit the accumulation of uncertainty and error from input data and models. The general
30 idea is to correct the state variables of the model using the observations when they become available. The simplest method
is through direct insertion (e.g., Turcotte et al., 2007; Fletcher et al., 2012). In this approach, the value of any modeled state
variable of interest is discarded and replaced by the observed value for this variable, when available. Nonetheless, this sim-
ple method is limited because there is no consideration of the uncertainty of the observation itself (Slater and Clark, 2006).
Moreover, that the spatial representation of the observation is much lower than the model resolution is often ignored. Several
35 elaborate data assimilation schemes have been proposed in the literature to deal with these limitations, including optimal inter-
polation (Burgess and Webster, 1980), ensemble Kalman filtering (Evensen, 2003), and particle filtering (Gordon et al., 1993;
Moradkhani et al., 2005). The latter relies on fewer hypotheses than the other schemes in relation to the nonlinearity and non-
normality of the uncertainty distributions of the model's input variables, state variables, and outputs. The particle filter (PF) is
based on the simulation and weighting of different scenarios (particles) rather than modifying the simulated state variables of
40 the model; thus, the disruption of cross-correlations between state variables is avoided. These scenarios are generated within
the range of the uncertainty of the input data and/or the model structure. Assessing or defining the range of uncertainty for each
variable is not a trivial task, but it is also not unique to the particle filter (see, for instance, Thiboult and Anctil, 2015). The
particles are weighted using the observations when they become available. These weights can be used to build the distribution
of the variable of interest (model output).

45   A common limitation of the PF is the risk of particle degeneracy. After a certain number of time steps, the evolution of
most particles brings them so far from the observation that they are associated with a weight close to zero. This can be seen
as an impoverishment of the filter with fewer and fewer useful particles. At some point, a single particle concentrates all the
weight, and a distribution can no longer be built. This degeneracy can be avoided using some kind of particle resampling; the
particles with the lowest weights are discarded, whereas the better ones are duplicated. Several resampling algorithms have
50 been proposed in the literature, such as those of Gordon et al. (1993); Douc and Cappe (2005); Leisenring and Moradkhani
(2011).

   The PF has been successfully used for the local (onsite) assimilation of snow data (Leisenring and Moradkhani, 2011;
Magnusson et al., 2017). The application of PF over large territories, however, remains challenging. The number of particles
must grow exponentially with the dimension of the state variables to avoid filter degeneracy (Snyder et al., 2015). This "curse
55 of dimensionality" (Bengtsson et al., 2008) is a severe drawback for the application of the PF over large spatial domains.
Nonetheless, a correct assessment of the spatial distribution of SWE over large domains is often required for diverse applica-
tions, including flood forecasting in large catchments. Therefore, any data assimilation technique for snow that can be applied

to such a large spatial domain with satisfactory results is of very high practical importance for operational hydrology, as it could potentially improve flood forecasting and reservoir management.

60   Several procedures have been proposed to overcome the curse of dimensionality. A first approach, known as error inflation (Larue et al., 2018), consists of increasing the error associated with the observation to make the filter more tolerant. Cluzet et al. (2020) proposed a generalization of this approach to apply it over large domains. The most common approach to overcome the curse of dimensionality, however, is localization (Farchi and Bocquet, 2018). This family of methods assumes that points separated by a sufficiently large distance are independent. Then, given this independence, the assimilation is performed locally

65   or for a spatially limited block. In this way, the impact of each observation is limited in space, and the assimilation is performed on several subregions of lower spatial dimensions.

Recent work involving spatialization or the PF for assimilating SWE over large spatial domains includes that of Cluzet et al. (2020) who developed the k-local framework to select the appropriate data set to be assimilated in each localized block; this selection is based on the cross-correlation among observations. Cantet et al. (2019) proposed a spatialized particle filter using

70   the localization approach. They used spatially structured perturbations to generate the particles, computed the weights of the particles locally for each observation site, then interpolated the weights in space. The underlying assumption of their method is that for any spatially coherent particle (i.e., model simulation), sites that are close to each other in space should be characterized by similar weights. The method was successfully applied to the assimilation of SWE data from manual snow surveys in the snow module of the HYDROTEL model (Turcotte et al., 2007) over the province of Quebec in eastern Canada.

75   Although localization is an effective means of circumventing the curse of dimensionality, this approach can create physically unrealistic spatial discontinuities between the local blocks because of the local resampling and the resulting noise (Farchi and Bocquet, 2018). This can occur when a global particle is formed (post-resampling) with a given particle on block a and with a different particle on the adjacent block b. The apparition of these discontinuities can be mitigated by improving the resampling algorithm. For instance, Farchi and Bocquet (2018) proposed the use of the optimal transport theory to minimize the movements

80   of local particles during resampling. Nevertheless, such a technique cannot completely prevent the apparition of discontinuities, especially when resampling must be performed frequently. Some authors suggest smoothing out the discontinuities by space averaging the resampled fields (Penny and Miyoshi, 2016). This appears to be an effective but rather subjective way to remove the discontinuities, as this smoothing is not related to the underlying spatial structure of the physical variable. A solution for gluing the particles back together and solving the spatial discontinuities becomes necessary when the number of observation

85   sites or the temporal frequency of the observations increases. Whereas being able to assimilate more data is always appealing to better capture the spatial and temporal variability of the snowpack properties, a greater amount of resampling is expected when more observations are used to evaluate the particles of the PF. A greater number of resamplings is likely to create new spatial discontinuities. In a context where snow observations are increasingly automated, it becomes crucial to remedy the dimensionality curse.

90   In Cantet et al. (2019), only the data from manual snow surveys were assimilated in the snow model. The fact that this type of measurement is not available continuously in time reduces the risk of spatial discontinuities appearing because of resampling. However, a large number of SR50 sonic sensors, which measure snow depth at an hourly frequency, have been deployed across

Quebec since 2000. Furthermore, Odry et al. (2020) recently proposed a machine-learning model to evaluate SWE from snow depth, temperature, and precipitation indicators in Quebec. On the one hand, there is potential added value in including multiple types of snow data in the particle filter; on the other hand, however, including data that are available continuously with a very short time step increases the risk of spatial discontinuity appearing.

Here we present an innovative means of gluing the particles of a localized particle filter back together using the Schaake Shuffle (Clark et al., 2004). The Schaake Shuffle was introduced to rebuild the spatial structure in forecasted ensemble meteorological fields. It uses a set of spatial observations of the variable—where the spatial structure is not damaged—to reorganize the members of an ensemble cell by cell over a spatial grid. It is an actual shuffling of the members, and there is no alteration in the distribution of the local ensemble. Given the similarities to the spatial discontinuity problem in the localized particle filter, we hypothesize that the Schaake Shuffle can re-glue the resampled particles together and restore the spatial correlation, thereby making it possible to eliminate the curse of dimensionality by applying a localized particle filter. To rebuild the short range correlation among the particles using an alternative and simpler solution, we also implement a simple sorting, in ascending order, of the particles' SWE.

To verify this hypothesis, we propose using the spatial particle filter developed by Cantet et al. (2019) and implement the Schaake Shuffle after each resampling of the particles. For comparison purposes, we use the same study area and model as Cantet et al. (2019). The reference set of the Schaake Shuffle is simulated over a historical period using the open-loop configuration of the model, as no reliable spatial observation of snow is available for this region.

The availability of SR50 data provides a great opportunity to test both particle reordering when the amount of data is increased and the ability of our approach to assimilate other types of data having different uncertainties in the spatial particle filter (SPF). In this context, and as a secondary objective of this study, we also hypothesize that the assimilation of this secondary SWE data set improves the estimates of SWE compared with the open-loop simulation and that the addition of data in the assimilation scheme improves the overall quality of the SWE estimates in ungauged sites.

The next section describes the study area and the data. The methodology is outlined in a third section, and the results and related discussion are presented in the fourth part.

2 Experimental setup

2.1 Study area and snow data

Quebec is a province in eastern Canada. For most catchments, the hydrology of rivers within the province is largely influenced by the marked accumulation of snow during the winter, with an annual maximum peak flow occurring during the spring freshet. For this reason, the Quebec government has put much effort into monitoring and simulating SWE across the province. Our research form a part of this strategy.

Historically, SWE monitoring has occurred through manual snow surveys using standard Federal Snow Samplers. The measurements are undertaken manually at various sites along predetermined snow lines at differing temporal frequencies (monthly to bi-weekly). This sampling strategy provides information regarding SWE, snow depth, and snow bulk density.
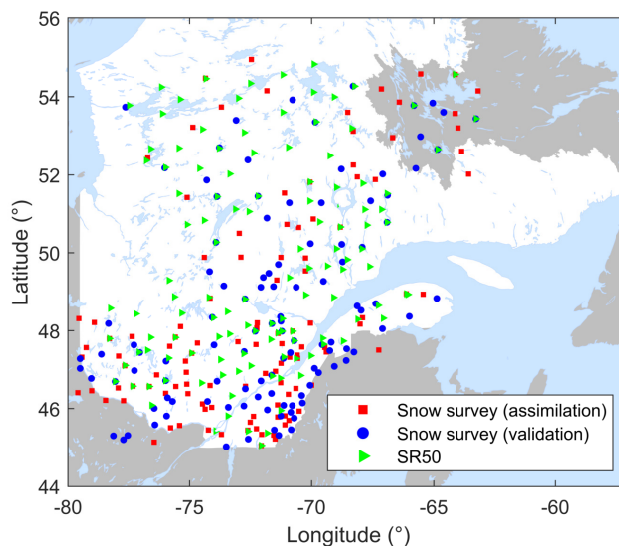
Figure 1. Snow observation sites in and around Quebec (Canada)

Since the beginning of the 2000s, automated snow sensors have been installed across Quebec. These automatic sonic sensors (SR50) measure snow depth continuously at an hourly time step. Although the temporal continuity of these snow depth series is a clear advantage, the absence of information regarding SWE is a drawback. This limitation is overcome by deriving daily time series of SWE from the SR50 data using an artificial neural network model developed for this purpose and trained and

130   validated against the manual snow survey data set (Odry et al., 2020). Because this is an indirect estimation of SWE, it is assumed that this second SWE data set is characterized by greater uncertainty than the data from the manual snow surveys. To be able to make use of both data sets, we focus on the 2011–2015 period for which both data sets have a relatively higher and more stable number of SR50 sensors.

     All snow data used in this research was provided by the Réseau de surveillance du climat du Québec (MELCC, 2019),

135   operated by the Quebec Ministère de l'Environnement et de la Lutte contre les Changements Climatiques (MELCC) and its partners from the private and public sectors inside the Réseau météorologique coopératif du Québec (RMCQ). The spatial distribution of observation sites is presented in Figure 1.

     This research focuses on a portion of the province located south of 53°N, where most of the population is concentrated. Northern Quebec differs markedly from the more southern region in terms of the amount of available snow data and snow

140   characteristics; therefore, its inclusion would require a separate study.

2.2   Snow model and meteorological forcing

We use the HYDROTEL snowpack model (HSM) in this study (Turcotte et al., 2007), which is used operationally for real-time SWE estimation at the MELCC. HSM is a temperature index–based model and therefore provides a simplified representation of the water and energy budgets of the snowpack. It is a rather simple snow model that only requires daily precipitation and

(cc) (i)

145   minimum and maximum temperatures as inputs. The implementation and parameterization of the model are those proposed by
Cantet et al. (2019). More detailed specifications about the model structure and physics can be found in Turcotte et al. (2007).

To force the HSM, daily grids of precipitation and minimum and maximum temperatures are used. They are produced by
the MELCC by kriging in situ measurements from local meteorological ground stations (Bergeron, 2017) on a regular 0.1°
resolution grid. Along with the interpolated variables, the corresponding kriging variances are also produced. They cover the
150   entire 1961–2018 period.

3 Methodology

3.1 Spatialized particle filter

The particle filter implemented in this study is that developed and tested by Cantet et al. (2019). This method was developed
to assimilate at-site snow data into a large-scale modeling framework. The main feature is to generate particles characterized
155   by a realistic temporal and spatial correlation so that each particle can be seen as a reasonable scenario. The observations are
then used to weight different particles when and where they are available. In this context, it is hypothesized that the weights
should be spatially correlated, given that the particles are also spatially correlated. In a second step, the weights estimated
at observation sites are interpolated in space. One of the most common limitations of particle filters is degeneracy. Filter
degeneracy occurs when most of the particles are characterized with a near-zero weighting, and a few particles concentrate all
160   the weight. In this case, the ensemble of particles has become too distant from the observations, and there is nothing left to
filter. To avoid filter degeneracy, Cantet et al. (2019) implemented the sampling importance resampling algorithm from Gordon
et al. (1993).

Following Cantet et al. (2019), we generated 500 particles using perturbations of the inputted meteorological data and the
predicted SWE. Perturbations of the precipitation and temperature inputs should represent the uncertainty associated with
165   the forcing. Given that these inputs are created through a spatial interpolation, they are perturbed using an additive noise
following a Gaussian distribution for which variance is set to the interpolation variance. The simulated SWE is perturbed
with a multiplicative uniform noise proposed by Clark et al. (2008) and used in Leisenring and Moradkhani (2011); this
perturbation process represents the structural uncertainty associated with the snow model. The temporal correlation of the
noise is maintained using the formulation (Evensen, 1994):

170    $$s_t = \alpha s_{t-1} + \sqrt{1-\alpha^2}\eta_t, \qquad \text{with} \qquad s_0 \sim \mathcal{N}(0,1), \tag{1}$$

where s_t is the random noise at time t, α represents the temporal correlation and is set at 0.95, and η_t is a random white noise
following a standard normal distribution.

The spatial correlation of the noise is built up using the exponential model (Uboldi et al., 2008):

$$r = \exp\frac{-d^2}{L^2}, \tag{2}$$

175   where r is the spatial correlation between two points separated by a distance d, and L is the range of the correlation, set at 200 km (Cantet et al., 2019).

In our study, local particle filters are implemented at the observation sites. The weight of a given particle at an observation site at a specific time when an observation is available depends on the likelihood of the measured SWE value given the SWE associated with this particle. This likelihood is represented by a Gaussian process:

180    $$y_t | x_t^i \sim \mathcal{N}(y_t - x_t^i, \sigma^2), \qquad (3)$$

where y_t represents the observed SWE at time t, x_t^i is the SWE simulated by the ith particle, and σ is the standard deviation of the Gaussian process and is related to the uncertainty associated with each type of SWE observation. For manual snow surveys, σ should represent both the uncertainty of the measurement itself and the error of spatial representation, i.e., how well a local measurement can represent the SWE averaged over a grid cell. σ is tuned to $0.1 * a * y_t + a$, with a set at 0.3 for the snow survey observations and to be determined for the SR50. In the case of SWE estimated from the SR50 snow depth measurements, σ should also represent the structural uncertainty of the artificial neural network model (Odry et al., 2020) used to estimate SWE. Moreover, although manual snow surveys average 10 measurements along a 100–300 m snow line, SR50 sensors provide a purely at-point measurement. As a consequence, the error for spatial representativity should be larger for SR50 than for manual snow surveys. The σ value associated with SR50 is unknown and must be estimated. The uncertainty

190   associated with the SR50s is also related to assimilation frequency. When the SR50 data are assimilated at a daily frequency, very similar values are presented several times to the filter, as there is a temporal correlation in the SWE. This would be equivalent to giving more weight to the SR50 observation. A means of alleviating this issue, at least in part, is to increase the uncertainty associated with the SR50; however, this approach can hide some temporal changes observed by the SR50. Here, we chose to assimilate the SR50 at a weekly frequency. In operational real-time simulations, the model is restarted each day from

195   seven days in the past to include all observations obtained over that one-week period. It would thus be possible to assimilate only the most recent SR50 observation and have an assimilation frequency close to weekly.

Under these hypotheses, when a given observation becomes available at time t, the weight x_t^i associated with the ith particle at the observation site can be updated as:

$$w_t^i = \frac{w_{t-1}^i p(y_t | x_t^i)}{\sum\limits_{i=1}^{N_P} w_{t-1}^i p(y_t | x_t^i)}, \qquad (4)$$

200   where N_P is the number of particles, set at 500. One can notice that the weight of a particle at a given time step depends on the weight at the previous time step. Therefore, the assimilation of continuous data with a time correlation can lead to assimilating the same information several times and to give too much weight to this information. As a consequence, the assimilation frequency should be investigated.

At a given time step, the weights are updated for any newly available observation using Equation 4. The weights at the

205   observation sites are spatially interpolated using a simple inverse distance-weighted method to estimate weights for each grid cell. This process makes it possible to include different observations obtained inside the same grid cell. The validity of the

interpolation of weights relies on the underlying hypothesis that a particle with a high likelihood at an observation site should be associated with a high likelihood for nearby locations.

Once the weights have been updated and interpolated, it is possible to investigate the risk of filter degeneracy. Degeneracy
210 occurs when a few particles concentrate all the weight. For each observation site and grid cell, N_t^{eff}, the effective sample size at time t is computed as:

$$N_t^{eff} = \frac{1}{\sum\limits_{i=1}^{N_P} (w_t^i)^2}. \tag{5}$$

If the effective sample size falls below the $0.8 N_P$ threshold (Magnusson et al., 2017), the particles are then locally resampled. It is important to highlight that whereas the particles are generated with a spatial structure, they are evaluated and potentially
215 resampled locally. The resampling is performed using the SIR algorithm (Gordon et al., 1993) and consists of deleting the particles with the lowest weights and duplicating those with the largest weights. The objective of this resampling is to obtain a set of particles associated with equal weights of $\frac{1}{N_P}$ that describe the same distribution as the weighted particles before resampling. Through this process, the filter retains a large number of meaningful particles, while the posterior distribution of SWE remains unchanged. Moreover, the SWE values and other state variables of the different particles are not modified.
220 Resampling affects only the weights, which makes it possible to maintain coherence between the different state variables.

Although the resampling aims to preserve the filter from degeneracy, it is also responsible for the risk of spatial discontinuity, as it is performed independently for each site and grid cell. Another approach would be to perform a global resampling; however, considering the large spatial dimension of the area, it is unlikely that only 500 particles would be sufficient to correctly describe snowpack variability over such a large territory. Such an approach would therefore require an exponential increase of
225 the number of particles (Snyder et al., 2015), i.e., the curse of dimensionality (Farchi and Bocquet, 2018). Another approach would be to decrease the amount of assimilated data or to increase the uncertainty associated with each observation (Larue et al., 2018; Cluzet et al., 2020). This solution may not be appropriate for a large territory having only a limited number of local observations, as it would deprive the assimilation scheme of precious information regarding the spatial diversity of SWE. Given these limitations, resampling is still performed locally (e.g., the localization technique), and the risk of discontinuity is
230 managed using a reordering of the particles.

3.2 Recovering the spatial structure by reordering the particles

We implemented two alternative procedures to maintain the spatial structure of each particle despite resampling: the Schaake Shuffle, proposed by Clark et al. (2004), and a simple sorting of the particles at each site.

The Schaake Shuffle is a reordering method that aims to recover the space–time structure in an ensemble of precipitation or
235 temperature forecast fields. In the original application of the Schaake Shuffle, the spatial structure of the ensemble members is lost during a downscaling procedure. The idea behind the Schaake Shuffle is to rank the members of the ensemble and match these ranks with those of past observations selected randomly from similar dates in the historical records. Although the Schaake Shuffle was not initially developed to be applied to the field of snow or data assimilation, its principle can easily be

==highlight== transferred from an ensemble of meteorological forecasts ==highlight== to a set of particles describing a snowpack. The only difficulty is to

240 find a replacement for the historical records, which are not available in our case and would typically never be available for any similar large-scale data assimilation problems.

For a given time and location, let $\mathbf{X}$ be the vector of length N_P of SWE values for all the particles. A vector of equal size $\mathbf{Y}$ is built from historical record at the same location and similar date. Considerations about the construction of $\mathbf{Y}$ are provided in the following sections. The vectors $\boldsymbol{\chi}$ and $\boldsymbol{\gamma}$ are the sorted versions of $\mathbf{X}$ and $\mathbf{Y}$, respectively:

245
$$\mathbf{X} = (x_1, x_2, ..., x_{N_P}), \tag{6}$$

$$\boldsymbol{\chi} = (x_{(1)}, x_{(2)}, ..., x_{(N_P)}), \qquad \text{where} \qquad x_{(1)} \leq x_{(2)} \leq ... \leq x_{(N_P)}, \tag{7}$$

$$\mathbf{Y} = (y_1, y_2, ..., y_{N_P}), \tag{8}$$

$$\boldsymbol{\gamma} = (y_{(1)}, y_{(2)}, ..., y_{(N_P)}), \qquad \text{where} \qquad y_{(1)} \leq y_{(2)} \leq ... \leq y_{(N_P)}. \tag{9}$$

Moreover, $\mathbf{B}$ is taken as the vector of indices describing the positions of the elements of γ in the original vector $\mathbf{Y}$. That is

250 to say the ith element of $\mathbf{B}$ describes where to find the ith element of γ in $\mathbf{Y}$. Finally, if $\mathbf{Z}$ is the reordered version of $\mathbf{X}$, then:

$$\mathbf{Z} = (z_1, z_2, ..., z_{N_P}), \qquad \text{with} \tag{10}$$

$$z_q = x_{(r)}, \tag{11}$$

$$q = \mathbf{B}[r], \qquad \text{and} \tag{12}$$

$$r = 1, ..., N_P. \tag{13}$$

255 This reordering must be applied at each site or grid cell. A more detailed procedure and a visual example are provided in (Clark et al., 2004).

==highlight== As it appears in equations 7–13, the main limitation to the application of the Schaake Shuffle is the availability of historical records for the construction of vector $\mathbf{Y}$. ==highlight== It is necessary to have records covering each site (observation sites and grid cells) and be of sufficient length to be able to draw N_P values from similar periods for each site. Given that an extensive data set does

260 not exist for SWE in Quebec, we propose to use a record of simulations using the HSM in its open-loop configuration. This selection ==highlight== rests ==highlight== on ==highlight== the assumption that the model without data assimilation can correctly estimate the spatial structure of SWE. One can notice that the actual SWE values in the historical records are not used in the Schaake Shuffle, only their respective rank is used to reorder the particles. ==highlight== In our application, the vector $\mathbf{Y}$ is constructed from a reference set of simulations produced with an open-loop configuration of the HSM for the years ==highlight== 1961 to 2004. ==highlight== Following Clark et al. (2004), only values associated

265 with a date within a window of ± 7 days around the assimilation date (but with different years) are used. Vector $\mathbf{Y}$ is built by randomly drawing 500 values from the 660 available records.

As a simpler alternative to the Schaake Shuffle, we also implemented a sorting of the particles. In this case, at each site, all 500 particles are sorted ==highlight== in ascending order. ==highlight== The idea is to rebuild the correlation between the sets of particles in neighboring

Table 1. Performance metrics

Metric	Focus	Perfect value	Formula
RMSE	Predictive accuracy	0	$RMSE = \sqrt{\frac{1}{n}\sum_{k=1}^{n}(\hat{y_k} - y_k)^2}$
MBE	Systematic bias	0	$MBE = \frac{1}{n}\sum_{k=1}^{n}(\hat{y_k} - y_k)$
CRPS	Global ensemble evaluation	0	$CRPS = \frac{1}{n}\sum_{k=1}^{n}\int_{-\infty}^{+\infty}(F_k^e(u) - F_k^o(u))^2 du$
Skill-to-spread ratio	Adequacy of spread and skill	1	$r = \dfrac{RMSE}{\sqrt{\frac{1}{n}\sum_{k=1}^{n}var_k}}$

n is the number of points, x_k and $\hat{x_k}$ are respectively the estimated and observed SWE values, var_k is the variance of the particles of the kth point, F_k^e is the cumulative distribution function described by the weighted particles of the kth point, and F_k^o is the Heaviside function associated with the observation.

sites—in the sense of the rank correlation. Although the sorting procedure is much simpler to implement and does not require

270   any reference record, it also removes the possibility to consider each particle as a reasonable snow scenario at a large scale. The sorting can be seen as a way of trading the large-scale structure for a short-scale one.

3.3   Validation procedure and metrics

To evaluate the impact of the different reordering strategies on the spatial structure of the particles and the final SWE estimation, we use experimental variograms. These plots provide the evolution of the variance between the sites as a function of the

275   separating distance. Because the variograms are only computed for a single date, the Pearson correlation between sets of particles is also used and computed at each time step so that the temporal evolution of the spatial structure can also be assessed. When investigating the spatial structure, all snow survey sites are used for assimilation.

To assess the ability of the overall method to estimate SWE, we compute two deterministic scores, the root mean squared error (RMSE) and the mean bias error (MBE), as well as two ensemble-based scores, the continuous ranked probability score

280   (CRPS, Matheson and Winkler, 1976) and the skill-to-spread ratio (Fortin et al., 2014). These metrics are summarized in Table 1. As the spatial particle filter intends to produce estimates at ungauged sites, only 50% of the manual snow survey sites are used for assimilation; the other half is kept as a validation subset. This sub-sampling strategy is presented in Figure 1. The 50% proportion is chosen so that approximately the same number of snow survey sites and SR50 are available for assimilation. All provided scores are computed only on the validation subset.
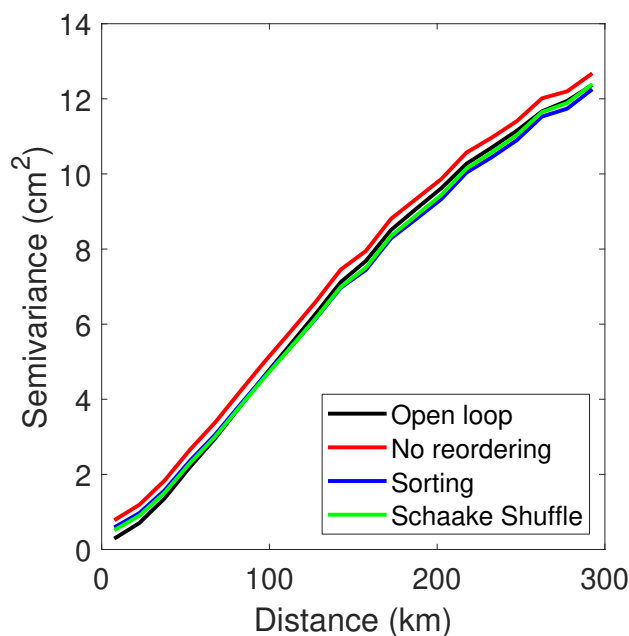
Figure 2. Variograms of the estimated snow water equivalent (SWE) on February, 28, 2012

4 Results and discussion

4.1 Spatial structure of the particles

The variograms of the estimated SWE for February 28th, 2012, for the different configurations (open loop, assimilation without reordering, assimilation with sorting of the particles, and assimilation with Schaake Shuffle) are provided in Figure 2. We selected February 28th because February is considered to be the last month of the accumulation season. The open-loop configuration corresponds to the deterministic simulation of HSM, whereas the three other curves correspond to the weighted average of the 500 particles. The four variograms are very similar, which demonstrates that the HSM in the open-loop configuration and the different assimilation strategies provide final SWE estimates with comparable spatial structures. In particular, the variograms have a semivariance below 1 cm^2 at a short-range. This value (nugget effect) is important, as it describes the spatial roughness or local noise. The presence of spatial discontinuities would create a nugget effect. It is interesting to note that the open-loop model is characterized by the lowest nugget effect (around 0 cm^2), which means an absence of local random noise, whereas the assimilation without reordering has the largest nugget effect, with a value around 1 cm^2. This large nugget effect can be a symptom of some noise affecting the correlation between neighboring sites.

The maps corresponding to the variograms in Figure 2 are provided in Figure 3. The four maps exhibit a similar spatial pattern, although the map from the open-loop simulation (panel a) is most distinct from the other ones. This was expected
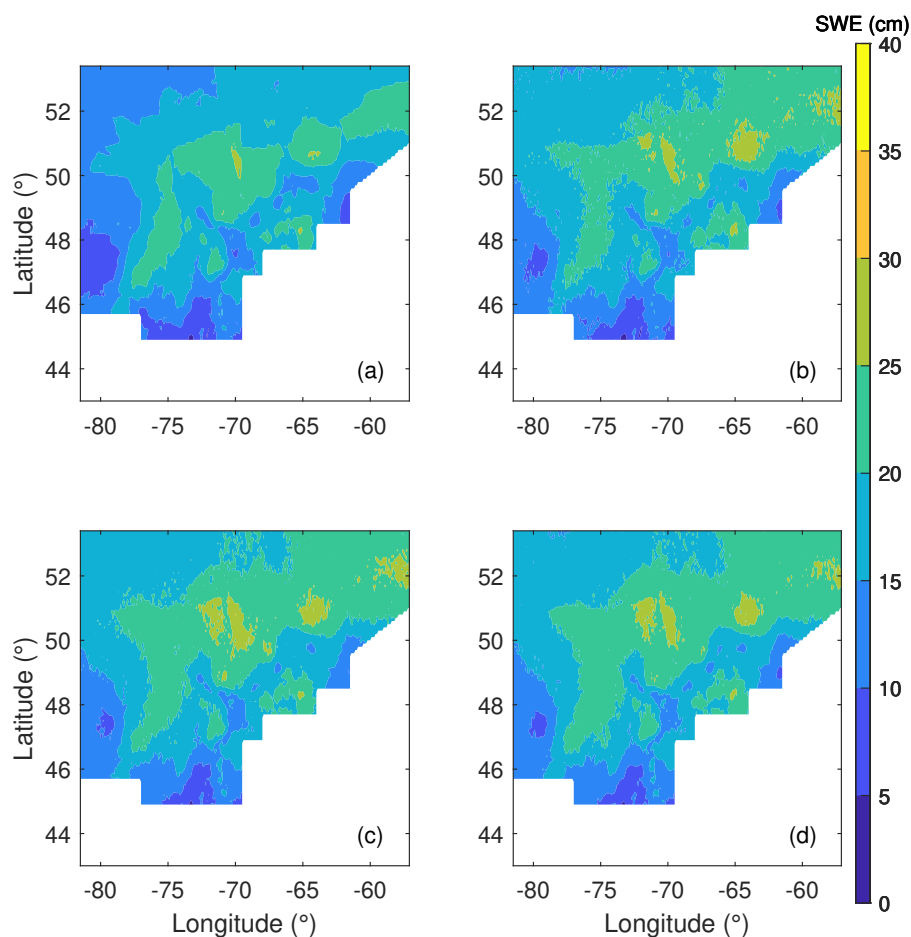
Figure 3. Estimated snow water equivalent (SWE) maps for February, 28, 2012; (a) open loop, (b) no reordering, (c) sorting, (d) Schaake Shuffle

300 because there is no assimilation involved in this simulation. The open-loop simulation also provides a smoother map compared with the other three, which is compatible with what is observed in the variograms.

Nevertheless, although it is important to assess the spatial structure of the final SWE estimates, the spatial particle filter relies on the assumption that the weights of the particles can be interpolated in space; therefore, the spatial structure of each particle must be preserved. To verify this, we assess the variograms of the 500 particles for the three reordering options (Figure

305 4). In Figure 4, the thin gray curves represent the variograms of the individual particles, whereas the bold colored lines are identical to those in Figure 2. Note that the vertical axes do not share the same scale. Moreover, the maps corresponding to two particles (number 250 and number 500) are provided in Figure 5 to illustrate the link between the variogram and spatial pattern. First, in the case where no reordering is used (Figure 4a), it is clear that the particles have a very different spatial structure from their weighted average. More specifically, all particles are characterized by a much greater nugget effect, ranging from
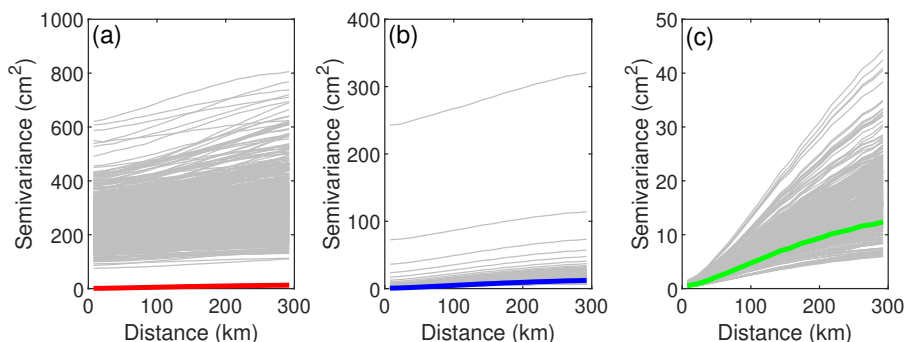
Figure 4. Variograms of the particles for the different reordering strategies for February, 28, 2012; (a) no reordering; (b) sorting; (c) Schaake Shuffle

approximately 100 cm^2 to around 600 cm^2; therefore, there is much short-range random variability. These nugget effects can be observed on Figure 5a and d, where the maps have a very rough aspect. That the individual particles exhibit a large nugget effect, whereas their weighted average does not exhibit this effect, indicates the presence of a large amount of random noise in each particle, which appears to be incompatible with the underlying assumption of the spatial particle filter.

Figure 4b presents the variograms of the individual particles when the sorting of the particles is used after each resampling. The sorting strongly attenuates the random noise in the particles. Most particles fall around the final estimate; nevertheless, some still exhibit a very large nugget effect. With the reordering by ascending order, we expected that some particles would show a disturbed spatial structure. Figure 5b and e provide more insight regarding this expectation. The two particles present very different patterns. This difference was expected, as one particle is located in the middle of the ensemble, whereas the other particle 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.

Reordering the particles with the Schaake Shuffle appears to be a better option to fix the spatial structure of the particles. Figure 4c illustrates that all particles have a near-zero nugget effect, and their variograms lie around that of the final estimate. This is compatible with the idea that the particles are alternative SWE scenarios, each providing a reasonable spatial structure. The same conclusions can be drawn from Figure 5c and f.

To assess the temporal evolution of the spatial structure, we refer to Figure 6, which shows the evolution of the correlation coefficient between the sets of particles at neighboring sites for winter 2011–2012. Here, the case of data assimilation without resampling is also plotted, allowing us to observe the impact of resampling and the spatial structure. Correlation decreases continuously with time for the cases without resampling (nor reordering) and without reordering. The main difference between the two options is a sudden drop of correlation values in March for the case of resampling without reordering. This drop is associated with a resampling. Whereas resampling can be associated with a large and immediate drop in correlation because of a disassociation of particles in the different sites, resampling cannot explain the continuous decrease in correlation. This
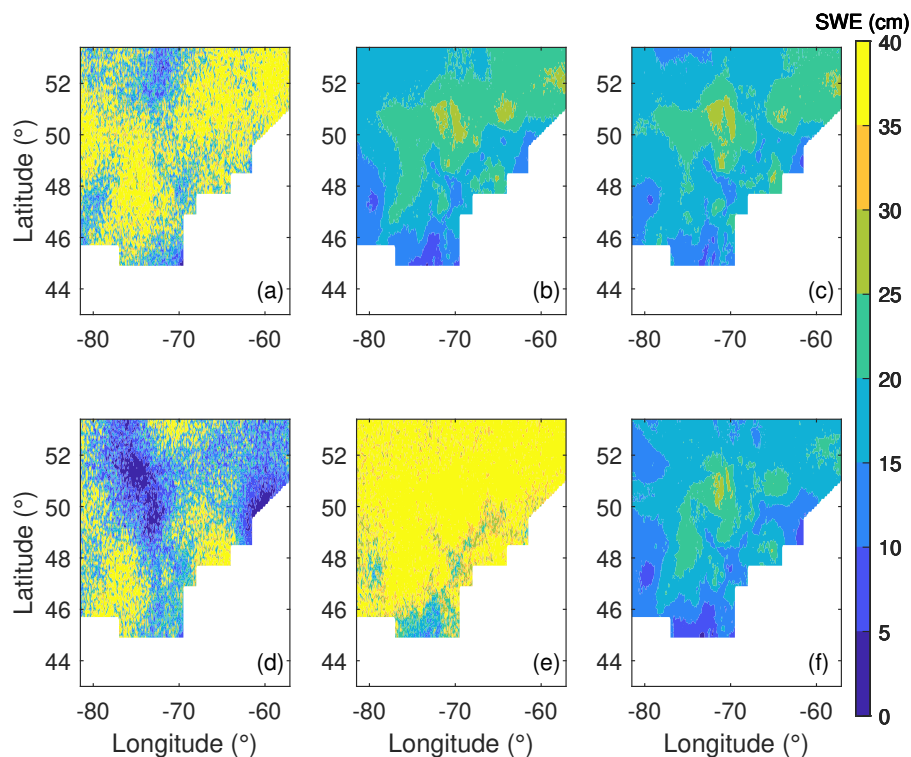
Figure 5. Snow water equivalent (SWE) for two particles for February, 28, 2012. Particle #250 (a–c) and Particle #500 (d–f); no reordering (a and d), sorting (b and e), and Schaake Shuffle (c and f)

decrease can be explained, however, by the cumulative nature of the snowpack and the perturbations. In this study and following Cantet et al. (2019), precipitation, temperature, and SWE itself are perturbed independently. The perturbations are generated

335   with a spatial structure, meaning that perturbations applied to neighboring sites are strongly correlated but are not identical. The simulated snowpack accumulates these differences throughout the winter, thus explaining the continuous decrease in correlation. Therefore, it appears that the assumption used to justify the interpolation of particle weights is not supported after a certain number of time steps without reordering, and resampling only aggravates the problem.

In Figure 6, the two reordering strategies maintain a very high correlation between neighboring sites. Some momentary

340   correlation losses are observed in April, which can be associated with the disappearing snow (some grid cells and particles having no snow).

As a consequence, it appears that both reordering strategies can maintain the spatial structure of the particles. In the present context, such a reordering is necessary to verify the underlying assumption of the spatial particle filter and makes it reasonable to interpolate the weights of the particle. From a more general point of view, reordering the particle appears to be an effective

345   way to deal with the spatial discontinuity created by the resampling in the case of localization of the particle filter. Reordering by sorting is easier to implement, as it does require an additional data set, but in this case, the individual particles cannot be

© ⊕ BY
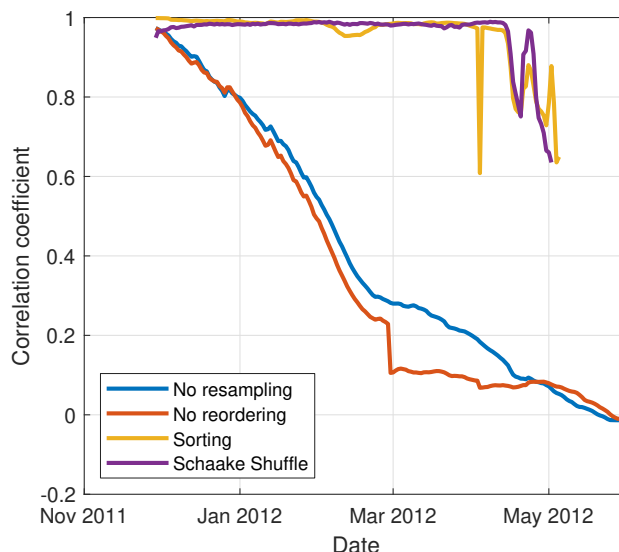
Figure 6. Evolution of the correlation between between two adjacent grid cells over the winter 2011–2012

seen as scenarios having a reasonable long-range correlation. In contrast, the Schaake Shuffle preserves the idea of each particle being a potential scenario. Because the SPF relies largely on the spatially structured perturbations and particles, the Shaake Shuffle appears to be a better solution to remain in a situation where the assumption necessary to interpolate the weight of the

350 particles remains true over the short and long range. Nonetheless, the Shaake Shuffle requires a long spatialized record to build the reference set. Here, we demonstrate that a historical open-loop simulation can be used rather than spatialized observations if we assume that the model can simulate a realistic spatial structure. In situations where a reference set cannot be acquired or created, the reordering by particle sorting can be an acceptable alternative to maintain the short-range correlation inside the particles.

355 **4.2 Snow water equivalent estimation**

In this and the following sections, all scores are computed solely on the validation subset (see Figure 1); the SWE data from the validation sites are never assimilated. The calculated scores intend to show the ability of the spatialized PF to estimate SWE in ungauged sites.

Figure 7 presents the cumulative distribution of the different performance metrics over the validation sites when only manual

360 snow surveys are assimilated. The data assimilation sensibly reduces the RMSE (panel a) and the MBE (panel b) relative to the open-loop simulation. Nonetheless, these two deterministic scores are very comparable for the three assimilation and reordering strategies. There is only a slight decrease in RMSE when using the Schaake Shuffle. Differences are more noticeable using the ensemble-based metrics. First and foremost, Figure 7c exhibits a large decrease in CRPS when using one of the reorderings. As the RMSE are comparable, this improvement of the CRPS is necessarily associated with a narrowing of the posterior
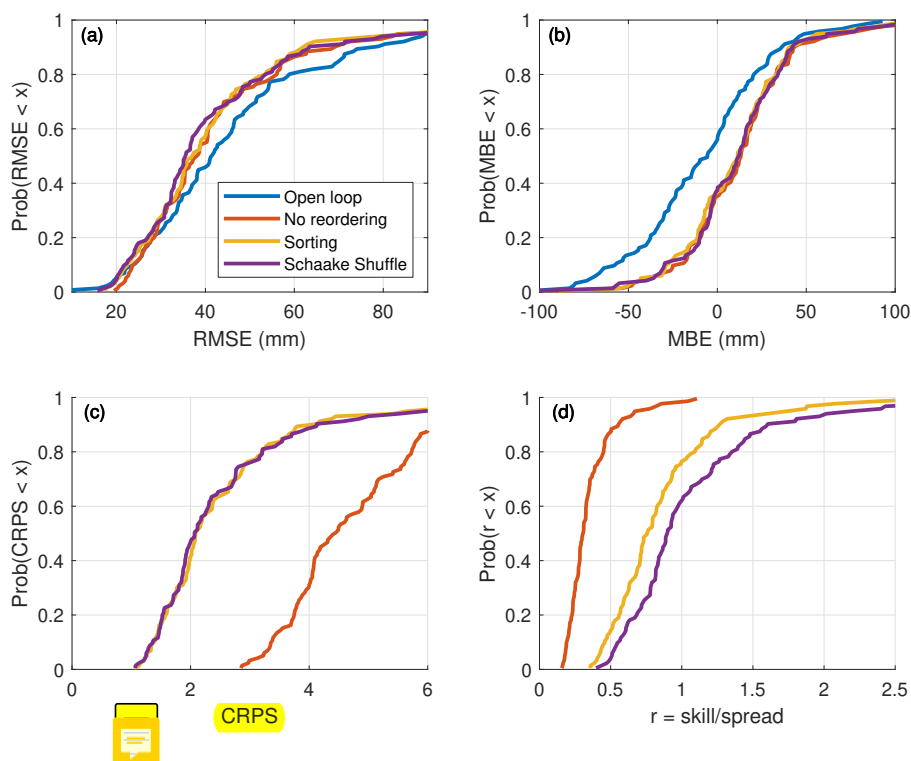
Figure 7. Cumulative distribution of the scores calculated for the 50% validation snow courses

365 distribution described by the particles. This result can be linked with the reduced amount of random noise associated with the reordering of the particles (see Section 4.1). The reduction of the ensemble spread is also apparent in Figure 7d. Because the skill (evaluated using RMSE) is not sensitive to the reordering, the increase in the skill-to-spread ratio can only be explained by a drop in terms of spread. The ratio is also systematically higher with the Schaake Shuffle than with the sorting, meaning that the Schaake Shuffle makes the spread of the ensemble of particles even narrower. In the case of the simple sorting, 80% of the

370 sites have a ratio below 1; thus, the particles are overspread. The situation improves slightly with use of the Schaake Shuffle.

4.3 Inclusion of the data from the SR50 sensors within the assimilation scheme

The first step before including the SR50 data is to estimate the uncertainty associated with the indirect observations. The total uncertainty is composed of the measurement uncertainty, the depth-to-SWE conversion error, and the spatial representation error. Here, the spatial representation error represents the difference between the local in situ observation and the simulated

375 value, which is averaged over a 0.1° resolution grid cell.

We assume that the uncertainty associated with the SR50 data follows $\sigma = 0.1 * a * y_t + a$, with the a-value larger than 3, as it is the value used for the snow survey data. Figure 8 displays the cumulative distribution of the performance metrics on validation sites when only the SR50 data is assimilated for different a values. The first observation is that the lowest (worst) scores are achieved for an a-value of 4. Higher (better) scores are attained at higher a-values, which confirms that the
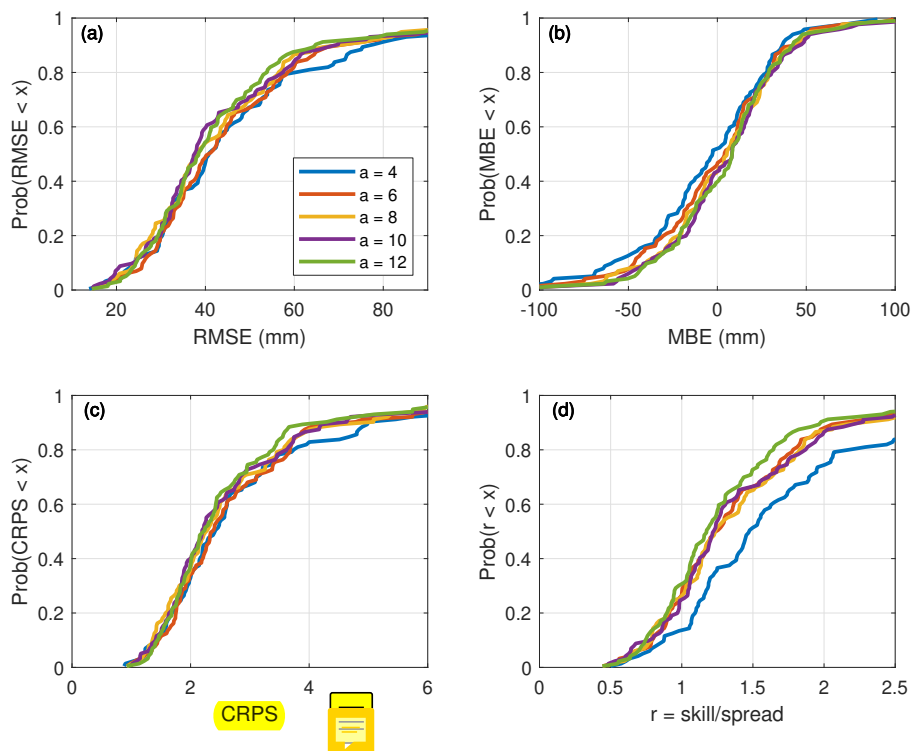
Figure 8. Cumulative distribution of the scores for the validation sites when assimilating SR50 data, according to the applied uncertainty model

380   uncertainty associated with the SR50 must be greater than that for the snow surveys. It also appears that if a is greater than 4, there are no large differences in the distribution of RMSE, MBE, or CRPS. Thus, the performances are not very sensitive to the a-value as long as it is sufficiently high (greater than 4). In Figure 8d, the skill-to-spread ratio increases with the a-values of the uncertainty associated with the observations from the SR50. This result indicates that greater uncertainty for these observations creates a larger spread of the particle, which is expected as a higher observation uncertainty provides greater flexibility to the particle around the observation and gives more weight to the model. Here, in the end and considering the observations above, we selected an a-value of 10 for the SR50 data. Given that the results are not very sensitive to this choice within this range of values, it is a judgment call based on the authors' expertise; nonetheless, a slightly higher or lower value could have been selected with minimal effect on the final result.

   Figure 9 displays the cumulative distributions of the performance metrics over the validation sites, according to the type
390   of data that is assimilated: snow surveys only, SR50 only, or both. First, all assimilation schemes significantly reduce the RMSE (panel a) for most sites. These results, similar to those of Cantet et al. (2019), demonstrate that the spatialized particle filter can assimilate local observations and use them to improve the global estimation of SWE. On panel b, the assimilation schemes increase the MBE values in all sites relative to those of the open-loop simulation, which confirms that assimilation tends to correct the underestimation of the open-loop simulation. When comparing the RMSE (panel a) and CRPS (panel c),

395   depending on the type of data that is assimilated, improved results are obtained when only the snow surveys are assimilated. The assimilation of SR50 observations only or of both types of observations simultaneously displays comparable performances, although still higher (better) than that of the open-loop configuration. Thus, in absence of snow surveys, it is possible and even desirable to assimilate estimated SWE data from SR50 observations to improve the simulation of SWE. Nonetheless, at this point, the validation process reaches its limitations. In our study, 50% of the snow survey sites are used in validation; therefore,

400   it is assumed that these observations represent the actual amount of SWE. Nevertheless, the model simulates SWE over a grid having a 0.1° resolution, thus local observations may not be accurate. Moreover, SR50 sensors and snow survey sites tend to not be installed in the same environment; SR50 sensors are generally installed as part of ground meteorological stations, which are located in a more open environments as per the World Meteorological Organization's guidelines, whereas snow surveys are almost always conducted in a forested environment. Consequently, the validation of snow survey sites only may produce a bias.

405   Finally, in Figure 9d, we note that the type of assimilated data has a large effect on the skill-to-spread ratio. The assimilation of SR50 greatly increases the ratio compared with the assimilation of snow surveys alone, whereas the assimilation of both types of data produces the highest values of skill-to-spread ratio. Given that the level of skill (RMSE in panel a) cannot explain this difference, the assimilation of SR50 only or of both types of data therefore decreases the particle spread in the ensemble. These results may be explained by the amount of assimilated data, as the more observations are assimilated, the more the particles are

410   constrained. The number of assimilated snow survey sites and SR50 are similar, but even with a weekly frequency, the number of observations is much higher for the SR50, which explains the decrease in spread. Obviously, the number of observations is even higher when assimilating both types of observations.

5   Conclusions

In this study, we propose an improvement of the spatial particle filter introduced by Cantet et al. (2019). The reordering of

415   particles using the Schaake Shuffle technique or a sorting procedure can help maintain the spatial structure of the particles within the ensemble. The Schaake Shuffle appears to be more effective in recreating the overall spatial structure of all particles. This solution makes it more reasonable to hypothesize that it is possible to spatially interpolate the weight of the particles; this is the basic assumption behind the spatialized particle filter. In cases where a sufficiently long reference record necessary to implement the Shaake Shuffle cannot be acquired, a reordering by particle sorting can provide a good alternative. In a more

420   general context, particle reordering is a good remedial solution to solve spatial discontinuity problems arising from resampling and localization when applying a particle filter to large spatial areas. Particle reordering may then be a new solution to the curse of dimensionality for particle filtering.

This research was also an opportunity to test the possibility of assimilating SWE estimates derived from automatic snow depth observations (SR50) and an ensemble of artificial neural networks (Odry et al., 2020). We demonstrated that the as-

425   similation of this additional data set alone outperforms the open-loop simulation. This observation confirms the relevance of this new SWE data set. However, the assimilation of both data from the SR50 and from snow surveys did not improve the simulation when compared with the assimilation of only the manual snow surveys. We attribute this result to the lower quality
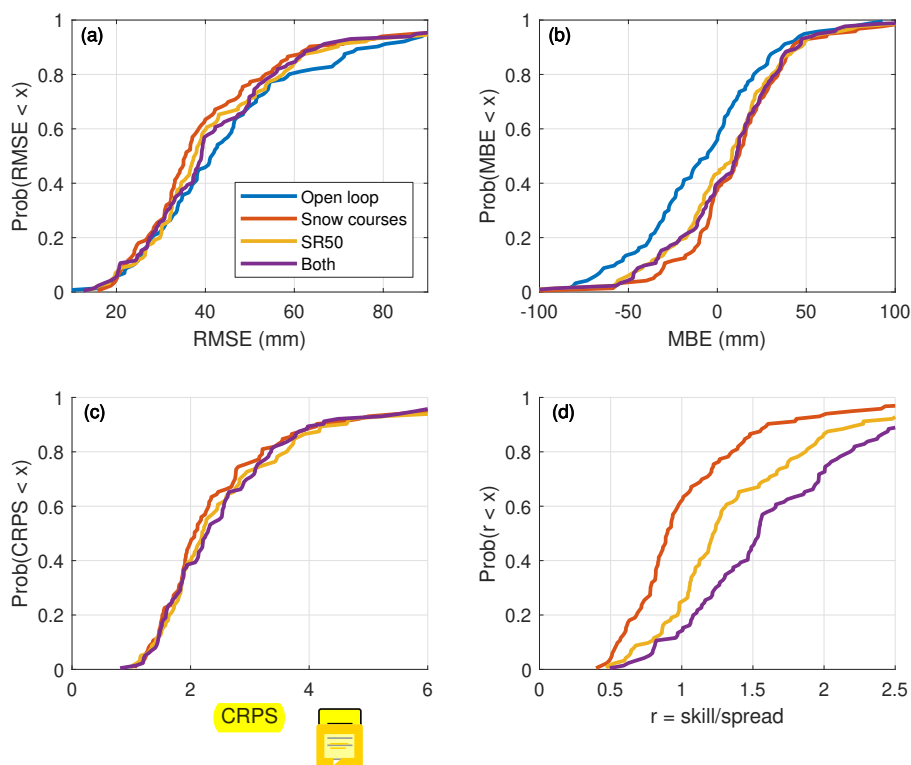
Figure 9. Cumulative distribution of the scores for the validation sites according to the nature of the assimilated data

of the SWE estimates from the SR50 and also to the validation process itself, which involves a portion of snow survey sites kept for validation. Nevertheless, the exercise demonstrated the possibility of assimilating different types of data together in the

430 spatial particle filter, the uncertainty associated with each kind of data being used to weight the relative influence of each type of observation. In this context, the current ongoing deployment in Quebec of automatic sensors capable of measuring SWE (rather than snow depth) by using natural gamma radiation constitutes a great opportunity. Not only could this third source of information be assimilated into the particle filter, but alternatively, it could also be used as an improved validation set, as these new gamma ray–based sensors provide sub-daily observations, and the added value of the SR50 in terms of temporal

435 representation may be better captured.

Code and data availability. The code for the SWE simulation using the spatialized PF with particle reordering is provided on Github https://github.com/TheDroplets/Snow_spatial_particle_filter (Odry, 2021). The data set used in this study, including the meteorological data and a part of the snow data that we are authorized to share can be accessed though the Harvard Dataverse, at https://dataverse.harvard.edu/dataverse/Odry_etal2021_largeScaleSpatialStructure. (Boucher, 2021a, b, c, d). There is a mandatory form to fill to access the data, and this

440 form can also be found in the same Dataverse.

Author contributions. Jean Odry performed all the computations, testing, and primary analysis and also wrote the first draft of the manuscript. Marie-Amélie Boucher initially proposed this study and supervised the work throughout the research and writing phases. Simon Lachance-Cloutier, Richard Turcotte, and Pierre-Yves St-Louis offered guidance, technical support, and expertise regarding data, snow models, and operational reality. They also participated in the editing of the manuscript.

445 *Competing interests.* The authors declare they have no competing interests.

Acknowledgements. The authors acknowledge the MELCC of Quebec for its financial support in the undertaking of this work. The authors thank Philippe Cantet for sharing the original code for the spatialized particle filter. Finally, they also want to thank the partners of the Réseau météorologique coopératif du Québec for sharing the snow data: Hydro-Québec, Rio Tinto, Ontario Power Generation, the Ministry of Northern Development, Mines, Natural Resources and Forestry of Ontario, Churchill Falls Corporation, Environment New Brunswick,
450 Maine Cooperative Snow Survey, Environment and Climate Change Canada, Quebec's Ministère des Forêts, de la Faune et des Parcs, and Quebec's Société de protection des forêts contre le feu.

References

Barnett, T. P., Adam, J. C., and Lettenmaier, D. P.: Potential impacts of a warming climate on water availability in snow-dominated regions, Nature, 438, 303, https://doi.org/10.1038/nature04141, 2005.

Bengtsson, T., Bickel, P., and Li, B.: Curse-of-dimensionality revisited: Collapse of the particle filter in very large scale systems, Institute of Mathematical Statistics, https://doi.org/10.1214/193940307000000518, 2008.

Bergeron, O.: Grilles climatiques quotidiennes du Programme de surveillance du climat du Québec, version 2 – Guide d'utilisation,, Tech. rep., Ministère du Développement durable, de l'Environnement et de la Lutte contre les changements climatiques, Direction du suivi de l'état de l'environnement,, Québec, 2017.

Boucher, M.-A.: Meteorological inputs, https://doi.org/10.7910/DVN/BXXRHL, 2021a.

Boucher, M.-A.: HYDROTEL Snow Model Parameters, https://doi.org/10.7910/DVN/RJSZIP, 2021b.

Boucher, M.-A.: Historical Snow Simulation (Open Loop), https://doi.org/10.7910/DVN/CJYMCV, 2021c.

Boucher, M.-A.: Snow Observation Data, https://doi.org/10.7910/DVN/NPB1JY, 2021d.

Burgess, T. M. and Webster, R.: Optimal Interpolation and Isarithmic Mapping of Soil Properties, Journal of Soil Science, 31, 315–331, https://doi.org/https://doi.org/10.1111/j.1365-2389.1980.tb02084.x, 1980.

Cantet, P., Boucher, M.-A., Lachance-Coutier, S., Turcotte, R., and Fortin, V.: Using a particle filter to estimate the spatial distribution of the snowpack water equivalent, Journal of Hydrometeorology, 20, 577–594, https://doi.org/10.1175/JHM-D-18-0140.1, 2019.

Clark, M., Gangopadhyay, S., Hay, L., Rajagopalan, B., and Wilby, R.: The Schaake Shuffle: A Method for Reconstructing Space–Time Variability in Forecasted Precipitation and Temperature Fields, Journal of Hydrometeorology, 5, 243–262, https://doi.org/10.1175/1525-7541(2004)005<0243:TSSAMF>2.0.CO;2, 2004.

Clark, M. P., Rupp, D. E., Woods, R. A., Zheng, X., Ibbitt, R. P., Slater, A. G., Schmidt, J., and Uddstrom, M. J.: Hydrological data assimilation with the ensemble Kalman filter: Use of streamflow observations to update states in a distributed hydrological model, Advances in Water Resources, 31, 1309–1324, https://doi.org/10.1016/j.advwatres.2008.06.005, 2008.

Cluzet, B., Lafaysse, M., Cosme, E., Albergel, C., Meunier, L.-F., and Dumont, M.: CrocO_v1.0 : a Particle Filter to assimilate snowpack observations in a spatialised framework, Geoscientific Model Development Discussions, pp. 1–36, https://doi.org/https://doi.org/10.5194/gmd-2020-130, 2020.

Doesken, N. J. and Judson, A.: The snow booklet: A guide to the science, climatology, and measurement of snow in the United States, Colorado State University Publications & Printing, 1997.

Douc, R. and Cappe, O.: Comparison of resampling schemes for particle filtering, in: ISPA 2005. Proceedings of the 4th International Symposium on Image and Signal Processing and Analysis, 2005., pp. 64–69, https://doi.org/10.1109/ISPA.2005.195385, iSSN: 1845-5921, 2005.

Essery, R., Morin, S., Lejeune, Y., and B Ménard, C.: A comparison of 1701 snow models using observations from an alpine site, Advances in Water Resources, 55, https://doi.org/10.1016/j.advwatres.2012.07.013, 2013.

Evensen, G.: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, Journal of Geophysical Research: Oceans, 99, 10 143–10 162, https://doi.org/https://doi.org/10.1029/94JC00572, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/94JC00572, 1994.

Evensen, G.: The Ensemble Kalman Filter: theoretical formulation and practical implementation, Ocean Dynamics, 53, 343–367, https://doi.org/10.1007/s10236-003-0036-9, 2003.

Farchi, A. and Bocquet, M.: Review article: Comparison of local particle filters and new implementations, Nonlinear Processes in Geophysics, 490   25, 765–807, https://doi.org/https://doi.org/10.5194/npg-25-765-2018, 2018.

Fletcher, S. J., Liston, G. E., Hiemstra, C. A., and Miller, S. D.: Assimilating MODIS and AMSR-E Snow Observations in a Snow Evolution Model, Journal of Hydrometeorology, 13, 1475–1492, https://doi.org/10.1175/JHM-D-11-082.1, 2012.

Fortin, V., Abaza, M., Anctil, F., and Turcotte, R.: Why Should Ensemble Spread Match the RMSE of the Ensemble Mean?, Journal of Hydrometeorology, 15, 1708–1713, https://doi.org/10.1175/JHM-D-14-0008.1, 2014.

495   Gordon, N. J., Salmond, D. J., and Smith, A. F. M.: Novel approach to nonlinear/non-Gaussian Bayesian state estimation, IEE Proceedings F (Radar and Signal Processing), 140, 107–113, https://doi.org/10.1049/ip-f-2.1993.0015, 1993.

Goïta, K., Walker, A. E., and Goodison, B. E.: Algorithm development for the estimation of snow water equivalent in the boreal forest using passive microwave data, International Journal of Remote Sensing, 24, 1097–1102, https://doi.org/10.1080/0143116021000044805, 2003.

Hock, R., Rees, G., Williams, M. W., and Ramirez, E.: Contribution from glaciers and snow cover to runoff from mountains in different 500   climates, Hydrological Processes, 20, 2089–2090, https://doi.org/10.1002/hyp.6206, 2006.

Larue, F., Royer, A., Sève, D. D., Roy, A., and Cosme, E.: Assimilation of passive microwave AMSR-2 satellite observations in a snowpack evolution model over northeastern Canada, Hydrology and Earth System Sciences, 22, 5711–5734, https://doi.org/https://doi.org/10.5194/hess-22-5711-2018, 2018.

Leisenring, M. and Moradkhani, H.: Snow water equivalent prediction using Bayesian data assimilation methods, Stochastic Environmental 505   Research and Risk Assessment, 25, 253–270, https://doi.org/10.1007/s00477-010-0445-5, 2011.

Li, L. and Simonovic, S. P.: System dynamics model for predicting floods from snowmelt in North American prairie watersheds, Hydrological Processes, 16, https://doi.org/10.1002/hyp.1064, 2002.

Magnusson, J., Winstral, A., Stordal, A. S., Essery, R., and Jonas, T.: Improving physically based snow simulations by assimilating snow depths using the particle filter, Water Resources Research, 53, 1125–1143, https://doi.org/10.1002/2016WR019092, 2017.

510   Marks, D., Domingo, J., Susong, D., Link, T., and Garen, D.: A spatially distributed energy balance snowmelt model for application in mountain basins, Hydrological Processes, 13, 1935–1959, https://doi.org/10.1002/(SICI)1099-1085(199909)13:12/13<1935::AID-HYP868>3.0.CO;2-C, 1999.

Matheson, J. E. and Winkler, R. L.: Scoring rules for continuous probability distributions, Management science, 22, 1087–1096, 1976.

MELCC: Données du réseau de surveillance du climat du Québec, Tech. rep., Direction générale du suivi de l'état de l'environnement, 515   Québec, 2019.

Moradkhani, H., Hsu, K.-L., Gupta, H., and Sorooshian, S.: Uncertainty assessment of hydrologic model states and parameters: Sequential data assimilation using the particle filter, Water Resources Research, 41, https://doi.org/https://doi.org/10.1029/2004WR003604, 2005.

Odry, J.: TheDroplets/Snow_spatial_particle_filter: First release of the codes for the spatial particle filter, https://doi.org/10.5281/zenodo.5531771, 2021.

520   Odry, J., Boucher, M. A., Cantet, P., Lachance-Cloutier, S., Turcotte, R., and St-Louis, P. Y.: Using artificial neural networks to estimate snow water equivalent from snow depth, Canadian Water Resources Journal / Revue canadienne des ressources hydriques, 0, 1–17, https://doi.org/10.1080/07011784.2020.1796817, 2020.

Ohmura, A.: Physical Basis for the Temperature-Based Melt-Index Method, Journal of Applied Meteorology and Climatology, 40, 753–761, https://doi.org/10.1175/1520-0450(2001)040<0753:PBFTTB>2.0.CO;2, 2001.

525   Penny, S. G. and Miyoshi, T.: A local particle filter for high-dimensional geophysical systems, Nonlinear Processes in Geophysics, 23, 391–405, https://doi.org/10.5194/npg-23-391-2016, 2016.

Schaefli, B., Hingray, B., and Musy, A.: Climate change and hydropower production in the Swiss Alps: quantification of potential impacts and related modelling uncertainties, Hydrology and Earth System Sciences, 11, 1191–1205, https://doi.org/https://doi.org/10.5194/hess-11-1191-2007, 2007.

530 Slater, A. G. and Clark, M. P.: Snow Data Assimilation via an Ensemble Kalman Filter, Journal of Hydrometeorology, 7, 478–493, https://doi.org/10.1175/JHM505.1, 2006.

Snyder, C., Bengtsson, T., and Morzfeld, M.: Performance Bounds for Particle Filters Using the Optimal Proposal, Monthly Weather Review, 143, 4750–4761, https://doi.org/10.1175/MWR-D-15-0144.1, 2015.

Thiboult, A. and Anctil, F.: On the difficulty to optimally implement the ensemble Kalman filter: An experiment based on many hydrological

535 models and catchments, Journal of Hydrology, 529, 1147–1160, 2015.

Turcotte, R., Fortin, L.-G., Fortin, V., Fortin, J.-P., and Villeneuve, J.-P.: Operational analysis of the spatial distribution and the temporal evolution of the snowpack water equivalent in southern Québec, Canada, Hydrology Research, 38, https://doi.org/10.2166/nh.2007.009, 2007.

Uboldi, F., Lussana, C., and Salvati, M.: Three-dimensional spatial interpolation of surface meteorological observations from

540 high-resolution local networks, Meteorological Applications, 15, 331–345, https://doi.org/https://doi.org/10.1002/met.76, _eprint: https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/met.76, 2008.