Review for manuscript hess-8-7202-2011, entitled "Improving the characterization of initial condition for ensemble streamflow prediction using data assimilation", authored by C. M. DeChant and H. Moradkhani.

Summary:
The authors employ a particle filter to generate an ensemble of initial condition estimates of snow fields for the purpose of ensemble streamflow prediction. Ground-based SNOTEL measurements are assimilated by the particle filter into the SAC-SMA model. The results presented show improvement to streamflow predictions using the proposed framework, but that significant limitations exist particularly related to the representativeness of the SNOTEL measurements used during the assimilation. Overall, this manuscript should be of interest to both the snow data assimilation community as well as the ensemble streamflow forecasting community, but that the manuscript could benefit from a more careful description of the methods used as well as the interpretation of some of the results.

General Recommendation:
Publish with major revisions.

Manuscript Evaluation:

Principal Criteria:

Scientific Significance: Good (2)
Scientific Quality: Good (2)
Presentation Quality: Fair (3)

General Criteria:

1. Does the paper address relevant scientific questions within the scope of HESS?
   Yes.
2. Does the paper present novel concepts, ideas, tools, or data?
   Yes.
3. Are substantial conclusions reached?
   Yes, but the overall findings could benefit from a number of clarification (please see comments below).
4. Are the scientific methods and assumptions valid and clearly outlined?
   More or less (please see comments below).
5. Are the results sufficient to support the interpretations and conclusions?
   In general, yes. However, there are some conclusions made in which I was not fully convinced (please see comments below).
6. Is the description of experiments and calculations sufficiently complete and precise to allow their reproduction by fellow scientists (traceability of results)?
   No. A few items related to the methodology require some clarification

(please see comments below).

7. Do the authors give proper credit to related work and clearly indicate their own new/original contribution?
   Yes.
8. Does the title clearly reflect the contents of the paper?
   Yes.
9. Does the abstract provide a concise and complete summary?
   Yes.
10. Is the overall presentation well structured and clear?
    Yes, for the most part (please see comments below).
11. Is the language fluent and precise?
    Yes.
12. Are mathematical formulae, symbols, abbreviations, and units correctly defined and used?
    Yes, but could benefit from a bit of consistency.
13. Should any parts of the paper (text, formulae, figures, tables) be clarified, reduced, combined, or eliminated?
    No.
14. Are the number and quality of references appropriate?
    Yes.
15. Is the amount and quality of supplementary material appropriate?
    N/A.

Major Changes/Questions/Concerns:

1. P. 7211, L.15: What motivated the selection of these 15 particular station locations? Was it based on data availability? Do these 15 stations run the gambit of ecotones/climatologies pertinent to snow data assimilation? Please justify with a sentence or two.
2. P. 7213, L.15 and L.17: Should the subscript inside the model operator be “t-l”? Also, please try to be more consistent with your notation. For example, “i” appears in the subscript in some places and in the superscript in others. Please make your usage of the notation more consistent as to make it easier for the reader.
3. P. 7213, L.15: Does the model error have to be additive? What about multiplicative errors, e.g.? How might your formulation change using multiplicative errors?
4. P. 7213, L.15 and L.22: How did you define the model and measurement errors? Are they subjectively-chosen scalars? Are they a function of SWE such that the error increases with increasing snow amount? Also, what about representativeness errors (e.g. point-scale versus model-scale error)? Please include a few sentences as to how you selected these error quantities and the rationale for doing so.
5. P. 7214, L. 9: What is “R” in Equation 4? Also, what does the subscript “k+1” represent? Please define accordingly.
6. P. 7216, L. 1: Is “more accurate” the same as “less uncertainty”?
7. P. 7216, L. 14: Similar to Comment #2 above, please be more consistent

with your notation. “i” was once the replicate number and now it is the probability category. I understand these quantities are somewhat related, but there is no harm in using different notation for the two as to make the paper more reader-friendly.

8. P. 7220, L. 15: Why not apply larger measurement errors in the upper elevations? You clearly state why the performance of your method varies as a function of elevation as related to the SNOTEL locations, but you’ve done nothing to try and alleviate this shortcoming. Perhaps a more judicious use of error model(s) would be beneficial? In addition, please refer to Comment #4 shown above and the need to include some description of the error models you selected and your rationale for doing so.
9. P. 7221, L.3: I don’t believe the term “accurate” is best used here. Just because you have a relatively uniform rank histogram is no guarantee that your ensemble is accurate (see discussions by Talagrand as well as Hamill). Perhaps the term “consistent” is a better choice?
10. P. 7221, L. 18: Perhaps a synthetic study would be a worthwhile pursuit in a follow-on study?
11. P. 7221, L. 19-20: Again, I think a more effective error model might be beneficial here. The results in Figure 2 could be an excellent starting point for developing a more accurate/useful error model. It would be interesting to see a discussion on error model sensitivity and how accounting for the known deficiencies in the SNOTEL station locations could translate into improved particle filter performance.
12. P. 7228, Figure 3: Is the “absolute” part necessary? If the elevation difference is negative, it would benefit the reader to clearly see if the station locations are too low or two high for a given elevation band. Therefore, I suggest you make the very simple change of presenting average elevation difference rather than average “absolute” elevation difference.
13. P. 7230: The particle filter is effectively removing mass even though the same forcing is being applied. Is this solely a function of the location of the SNOTEL stations within a certain elevation band? Or are there more factors at play? Land DA systems are notorious for not conserving mass and it would behoove the reader to have you elaborate as to whether or not your system conserves mass. You hint at the benefits of the particle filter as related to its ability to conserve mass (P. 7213, L. 5-10), but I’m not convinced based on your results in Figure 4. Please comment on the issue of mass conservation with a particular emphasis on convincing the reader that mass is, in fact, being conserved and that the particle filter is really doing everything you claim it can do.
14. P. 7234: The x-axis is not a “Rank”, per se. Perhaps “Normalized Rank” would be more appropriate?
15. P. 7234: Is the ESP-DA rank histogram underdispersed? Overdispersed? Please elaborate with a sentence or two.

Minor Changes/Questions/Concerns:

1. P. 7214, L. 6: There is a noun missing in this sentence, but I'm not certain what it is. Perhaps you omitted the word "terms"?
2. P. 7214, L.15: This is an incomplete sentence as it is missing a verb. Should "equal" be changed to "are set equal"?
3. P. 7215, L. 20: This is a good point made in this sentence. Be certain to re-iterate that point elsewhere (e.g. Conclusions) to hammer that point home for the reader.
4. P. 7215, L. 26: I believe there are 10 ensemble replicates in the referenced figure rather than the 8 as specified?
5. P. 7220, L. 2: "Beginning" rather that "begging."