Response to Comments on the Manuscript:

"Investigating the performance of Genetic Particle Filter in snow data assimilation across snow climates"

(HESS-2022-350)

March 26, 2023

The authors gratefully acknowledge the editors and the anonymous reviewers for their constructive comments. We have made a comprehensive revision of our previous manuscript. The main modifications have been highlighted in yellow color in order to highlight the issues raised by the editors and anonymous reviewers, and are summarized as follows:

For more details, please refer to the item-by-item response. Thank you for your time.

Response to comments by Reviewer #1:

We would like to take this opportunity to gratefully thank the reviewer for his/her constructive comments and recommendations for improving the paper. An item-by-item, point-by-point response to the interesting comments raised by the reviewer follows.

The topic addressed is within the scope of HESS. The manuscript is generally well organized and results are clearly presented. This manuscript investigated the potential of GPF used as a snow data assimilation scheme across different snow climates, the results presented in this manuscript will help develop new data assimilation scheme and improve the simulation accuracy of land surface model that leads to improve weather and climate prediction. In my opinion, this manuscript could be accepted for publication in HESS after the following comments are addressed.

Comments:

Line 106: "Above studies" may need some recent references.
Reply: Thanks for your sincere and constructive suggestions. We have included recent literature in this section.

Line 245: "The number of particles was set to 100" have been expressed in line 225, I suggest deleting one.
Reply: Thanks for your sincere and constructive suggestions. We have deleted the sentence in this line.

Line 250: the variance scaling factor of the temperature was set to 2.0, why this value was chosen, 3.0 or other value can be used here?

Reply: Thanks for your sincere and constructive suggestions. The variance scaling factor of the temperature was referenced the method of Lei et al. (2014), and the value was obtained through repeated attempts and experiments.

Line 259: What does the "SD" is refer to? Do you mean SD is the abbreviation of snow depth?

Reply: Thanks for your sincere and constructive suggestions. Here the SD is the abbreviation of snow depth and we have remarked in the manuscript.

The abstract should provide some numerical values from the performance metrics of the results.

Reply: Thanks for your sincere and constructive suggestions. We have provided some numerical values from the performance metrics of the results in the abstract and rewrite the abstract seriously.

Line 236: Except for the air temperature and precipitation can be perturbed, whether other meteorological forcing variable can be perturbed, such as relative humidity and wind speed? As far as I know, the wind speed has great impact on the distribution of snow.

Reply: Thank you for your sincere and constructive suggestions. Snow depth is primarily influenced by precipitation and temperature, with snow being highly sensitive to air temperature. Additionally, wind speed and other meteorological forcing variables such as longwave radiation and shortwave radiation have a significant impact on the distribution of snow. Although our study focuses on the performance of the genetic particle filter at the site scale, we acknowledge that the distribution of snow at the regional scale was not taken into consideration. Therefore, we only perturbed the air temperature and precipitation.

The English writing has to be polished.

Reply: Thanks for your sincere and constructive suggestions. We have proof-read the whole manuscript to fix the language and grammar issues, and an English writing service has been purchased for this manuscript.

Response to comments by Reviewer #2：

Snow depth simulation is difficult in land surface models and the data assimilation for snow depth is of importance for cold regions hydrology and energy balance modeling. This manuscript tried to propose a new data assimilation scheme for the land surface model. I read it with high interest but did not find out the logic of this manuscript. Therefore, I cannot give too positive evaluation at the current stage, I would suggest a thorough revision before it can be considered for publication in the journals. first, the language is quite poor and the writing is difficult to understand especially in the introduction and results and discussion parts. I can't understand the importance of this work except from my own understanding of the cold regions modeling and data assimilation. Second, some of the references are not shown in the references part even they are put in the main text. This is awful and I feel that the preparation for this work was not serious and also not strictly following the journal's rules. Besides, the results and discussion are quite awful in writing, as I can't find out the useful information from this work concluded by the authors. This is a quite pity issue even the meaningful work was conducted. Given the above mentioned issues, I feel that the detailed comments are not necessary if the authors don't make a thorough revision on the whole story telling logic. Therefore, I suggest a rejection this turn and a chance for resubmission with a clear outline that focuses on the most interesting part of the

work would be a good suggestion from my side. Sorry for being not too positive this time given the current version of the manuscript.

Reply: Thanks for your sincere and constructive suggestions. The reviewer has pointed out the deficiency of this manuscript, and we sincerely agree the comments. We do our best to revise this manuscript and examine the language seriously, and we also re-considered the structure of this manuscript. An English writing service has been purchased for this manuscript, and we look forward to your positive comments this time.

Response to comments by Reviewer #3：

Summary: The study is focused on the simulation of snow depth with a simple energy balance model with assimilation with observations. The data assimilation scheme is based on a genetic particle filter algorithm. The main conclusion is that this algorithm performs well for snow depth simulations.

Recommendation: the manuscript is in general terms poorly written. The English needs would need extensive copy-editing, and this deficiency often hinders the understanding of technical and scientific aspects of the study. In addition, the manuscript does not include important information that is critical to understand what has been done. One is that there is no description of the genetic particle filter algorithm itself, which is surprising. Unfortunately, I cannot recommend the publication of the manuscript. A revision would entail rewriting the manuscript almost completely.

Main points
1) **The presentation is poor. The English needs very extensive revisions; acronyms are not defined (for instance SD, which I interpret to be snow depth! SWE and GPF are not defined either). The discussion is restricted to the own results and does not place the results in the framework of previous studies (what has been learned, what is the novelty?)**

Reply: Thanks for your sincere and constructive suggestions. We have thorough revised the manuscript according to the reviewer's suggestion.

2) **The destitution of important technical aspects is missing. The genetic algorithm is essentially not described the paragraph starting in line 218 is so obscure that essentially nothing can be understood. The reader is not informed of many technical aspects. What are the 'particles'? how are they genetically generated? what are the crossover and mutation operators? why the genetic algorithm would improve on the deficiencies of other particle filters?**

Reply: Thanks for your sincere and constructive suggestions. We have introduced the Genetic Particle Filter algorithm in detail in the revised manuscript, including the selection, crossover and mutation process for the particles. The random samples propagating in state space are used to approximate the probability density function of state variable, in this case, the integral operation was replaced with the sample mean to obtain the minimum variance estimate of the state. And the sample members are called "particles" in particle filter. The particles were generated by forcing the model operator with the perturbed meteorological forcing data. The crossover operators are below equations:

$$x'_m = \alpha x_m + (1-\beta) x_n \tag{1}$$

$$x'_n = \beta x_n + (1-\alpha) x_m \tag{2}$$

Where α and β are the empirical crossover coefficient, $\alpha = 0.45$, $\beta = 0.55$ in this paper. The mutation operator is,

$$x_k^{'} = x_k + \eta * Uniform \qquad （3）$$

where the *Uniform* represents random of uniform distribution and η is the empirical coefficient which was set to 0.01 in this paper. The problem of particle degradation solved by conventional resampling methods like multinominal resampling and systematic resampling always results in particle impoverishment, the diversity of particles will be greatly enhanced using the genetic algorithm in particle resampling. We have revised the manuscript seriously according to the reviewer's suggestion.

3) **I kept wondering of the utility and meaning of some of the mathematical assumptions. For instance, equation 9 seems to be unnecessary complicated. The distribution of the random noise w is just uniform in (-2,2), so there is no need for the additional complexity of equation 9. Also, why would the temperature errors be uniformly distributed? why between -2 and 2 and which units represent those numbers (I guess C?). This is an example of a problem that goes through the whole manuscript.**

Reply: Thank you for your sincere and constructive suggestions. We used the perturbation method of meteorological forcing data in this study as described by Lei et al. (2014), which has been proven to be an effective method in data assimilation. In this study, the temperature unit is Kelvin, and we assumed the temperature error to be uniformly distributed. The variance scaling factor of the temperature was set to 2.0. We completely agree with the reviewer's opinion that equation 9 is unnecessarily complicated. If the reviewer and editor believe it to be unnecessary, we can certainly remove it. We have thoroughly revised the manuscript according to the reviewer's suggestions and appreciate the valuable feedback. Thank you.

Lei, F. N., Huang, C. L., Shen, H. F., et al. (2014), Improving the estimation of hydrological states in the SWAT model via the ensemble Kalman smoother: Synthetic experiments for the Heihe River Basin in northwest China, Advances in Water Resources, 67: 32-45.

Response to comments by Reviewer #4：

This paper optimizes the snow depth data simulated by the model by using the snow depth data observed at the station. The main questions are as follows:

1. **The title of the article is not accurate, and the purpose of assimilation cannot be obtained. The meaning of "snow" is too broad, so it needs to be specific;**

Reply: Thanks for your sincere and constructive suggestions. In this study, the Genetic Particle Filter was used as a data assimilation scheme to improve the simulation accuracy of snow depth by Noah-MP model, as the reviewer said, the meaning of "snow" is too broad in the title, and we have revised the title to "A genetic particle filter scheme for univariate data assimilation into Noah-MP model across snow climates", and we hope this title can highlight the theme of this manuscript.

2. **Is the Ws in NOAH-MP optimized by snow depth observation? Ws is the snow water equivalent. Is there any difference between the two?**

Reply: Thanks for your sincere and constructive suggestions. The snow water equivalent was calculated by snow depth and snow density in Noah-MP model, and we have noticed the snow water equivalent was optimized by snow depth observation. However, it is worth noting that most of the sites selected for this study do not have snow water equivalent observation data. Therefore, we were unable to include the assimilation results of snow water equivalent in the manuscript.

3. The observation operator and model operator mentioned in Flow Chart 1 are not seen in the text, and need to be clarified;

Reply: Thanks for your sincere and constructive suggestions. In this study, the Noah-MP model is the model operator mentioned in Flow Chart 1. In this study, the station observation data was assimilated into the Noah-MP model and there is no remote sensing observation data assimilated into the model. Therefore, the observation operator is an identity matrix in this study, and we have stated in section 2.

4. The optimized state variables are not seen in the paper. If only the simulation results of NOAH-MP are corrected, which cannot be regarded as assimilation;

Reply: Thanks for your sincere and constructive suggestions. The original objective of this study was to evaluate the performance of the genetic particle filter in different snow climates. To accomplish this, we assimilated observed snow depth at the point-scale into the Noah-MP model. As a result of the assimilation process, the snow simulation process was optimized, and we observed updates to all snow variables. However, we would like to note that due to the lack of snow water equivalent observation data at most of the sites, we were only able to present the assimilation results of snow depth in the manuscript.

5. The calculation formula of fsnow, g needs to be given. Snow depth is also used in the calculation formula of fsnow, g. If snow depth is assimilated, has this been considered? If regional assimilation is carried out, how is fsnow, g calculated?

Reply: Thanks for your sincere and constructive suggestions. The main objective of this study is to evaluate the performance of the Genetic Particle Filter across different snow climates, and to assess its feasibility as a point-scale snow data assimilation scheme. Through our study, we discovered that the fsnow variable can also be updated when observed snow depth is assimilated. Our next step is to expand this study to the regional level, where we will assimilate snow albedo products into the land surface model using the genetic particle filter to improve fsnow. In this study, our focus was on investigating the feasibility of GPF as a snow data assimilation scheme and evaluating its performance across different snow climates. We hope that the conclusions drawn from this manuscript will be a valuable reference for our upcoming regional snow data assimilation experiment.

6. The introduction of assimilation process in this paper is not complete and detailed enough, and needs to be further improved.

Reply: Thanks for your sincere and constructive suggestions. We have thoroughly revised the manuscript and provided a more detailed introduction to the assimilation process, as suggested by the reviewer.

Response to comments by Reviewer #5：

The novelty of this research is not clear. If the genetic particle filter is a new method, then its advantages against other methods should be investigated directly. If the "across snow

climates" is a novelty, then I do not think the eight stations can represent snow climates considering the highly heterogeneous snow distribution. Finally, "the higher assimilating frequency, the higher simulation performance" is not surprising. Thus, I suggest the authors to clarify the novelty more clearly, so as to help the reader get the importance of this work. The method needs further introduction. Although the genetic particle filter data assimilation scheme is introduced in section 2.3, I am still confused that which variable you assimilate into the land surface model? If you only assimilate the snow depth, then how do you deal with other snow variables (e.g., snow water equivalent, snow density and snow age)? Also, how do you deal with the potential inconsistency between snow and ground temperature (for example, when the model shows no snow and the ground temperature is above zero, then how to assimilate the observed snow depth)? Actually, I am very concerned about the assimilation and evaluation. In figure 3, it seems you assimilate the observation every 5 days and then evaluate the model simulation at the same step? If this is the case, then will the direct insertion method show higher performance than the genetic particle filter?

Reply: Thanks for your sincere and constructive suggestions. We have thoroughly revised the manuscript and provided a detailed introduction to the methodology. The primary aim of this study is to evaluate the feasibility of using the Genetic Particle Filter as a snow data assimilation scheme in point-scale and to examine its performance across different snow climates. To achieve this objective, we assimilated observed snow depth at point-scale into the Noah-MP model. Initially, the snow cover was divided into multiple layers based on the updated snow depth, and the snow water equivalent was updated using corresponding proportions calculated by the updated snow depth. Next, the snow depth was treated as an independent variable to calculate other snow variables, such as snow cover fraction, etc. During the snow period, we found that it could be divided into snow accumulation period, snow stable period, and snow ablation period. We only assimilated observed snow depth values that were above zero, that is, we did not assimilate observations when the model showed no snow. In this study, we assimilated snow depth observations every 5 days, while the temporal resolution of simulation results and snow depth observations used for evaluation was one day. Although we initially used the direct insertion method to improve the simulation results, we found that its performance was worse than the particle filter and could lead to a model shock during the assimilation period.

The spatial difference. It seems the spatial difference among different stations is not strong and few information can be get (except the robust of the result, may be). Some insightful analysis on the spatial difference may help improve the manuscript. The writing needs careful proofreading.

Reply: Thanks for your sincere and constructive suggestions. We have thorough revised the manuscript according to the reviewer's suggestion, and an English writing service has been purchased for this manuscript, thank you very much.

For example:
1. **L48: "succeeds in catching snow dynamics is" may be "succeed in catching snow dynamics is"**

Reply: Thanks for your sincere and constructive suggestions. We have revised in the manuscript.

2. **L51: "is aimed at investigating ... and obtain the ..." may be "is aimed at investigating ... and obtaining the ..."**

Reply: Thanks for your sincere and constructive suggestions. We have revised in the manuscript.

3. **L60: "However, this method possible result in ..." may be "However, this method possible results in ...".**
Reply: Thanks for your sincere and constructive suggestions. We have revised in the manuscript.

4. **L68: "this method does not require a model a model linearization." what do you mean?**
Reply: Thanks for your sincere and constructive suggestions. The Kalman Filter is a useful tool for linear systems, but it requires a model linearization for a nonlinear system. To overcome this limitation, the Monte Carlo approach was employed to approximate error estimates based on an ensemble of model simulations in the Ensemble Kalman Filter (EnKF). Consequently, the EnKF can be used in nonlinear systems without requiring model linearization.

Response to comments by Reviewer #6:
In this manuscript, You et al. set up a particle filtering framework using the Genetic Algorithm to avoid particle filter-typical degeneracy and sample impoverishment issues. They apply this framework to snow depth measuring sites in different climatological regions, expecting to learn about particle filter performance at each of the sites. They analyze the assimilated snow depth with respect to the suitability of their particle filtering algorithm for application in different snow climates, the influence of the used particle number on performance metrics, and the influence of assimilation window length on the performance. The manuscript is structurally well-organized. The topic is in general very interesting and the effort to push the field of data assimilation forward is very much appreciated. However, in my opinion there are some significant weaknesses in this manuscript:

- a lacking motivation of the research question(s') relevance
Reply: Thanks for your sincere and constructive suggestions. The original goal of this study was to investigate the performance of the Genetic Particle Filter as a snow data assimilation scheme across various snow climates. We attempted to resample the particles using a genetic algorithm. Following the reviewer's suggestions, we have thoroughly revised the manuscript.

- a superficial description of the measurement sites and their properties, making a meaningful interpretation of the results with respect to the research question(s) difficult.
Reply: Thanks for your sincere and constructive suggestions. We explored the sensitivity of snow simulations to parameterization schemes within the Noah-MP model using the same dataset of eight measurement sites as in our previous work (You et al., 2020). A detailed description, including the location map of the measurement sites, can be found in that paper. Therefore, in this manuscript, we have only briefly described the measurement sites. However, if the reviewer and editor feel that a more detailed description is necessary, we are happy to add it at any time. We have thoroughly analyzed the experimental results, as per the reviewer's suggestions. Thank you very much.

- a lacking presentation of the Genetic Algorithm and stressing why this method is the most suitable for the analysis
Reply: Thanks for your sincere and constructive suggestions. The degeneracy and sample impoverishment issues are common faults of particle filter, here, the genetic algorithm was used in resampling process and was expected to effectively mitigate these issues. We have made a detailed

description of the Genetic Particle Filter in the manuscript according to the reviewer's suggestion, thank you very much.

- a superficial interpretation of the results, in particular with respect to the overarching hypothesis (different filter performance in different "snow climates")

Reply: Thanks for your sincere and constructive suggestions. We have thorough revised the manuscript and presented a detailed discussion in the manuscript according to the reviewer's suggestion.

- issues with the used literature in the References section and in general a rather scarce literature selection. Difficulties in the use of English, which makes some sections of the manuscript hard to understand

Reply: Thanks for your sincere and constructive suggestions. We have completed the references and thorough revised the manuscript according to the reviewer's suggestion. Additionally, an English writing service has been purchased for this manuscript. Thank you very much.

- an intransparent (or simply not listed?) choice of model parameter values and meteorological values to perturb; unclear or not explained error distribution choices.

Reply: Thanks for your sincere and constructive suggestions. In this study, the model parameters were obtained from the look-up table within the Noah-MP model, based on the soil and vegetation type of the sites. We did not perturb the model parameters but rather perturbed the meteorological forcing to produce ensembles. The perturbation method and the error distribution choices were presented in the manuscript.

- a results and discussion section that partly loses contact with the research questions

Reply: Thanks for your sincere and constructive suggestions. We acknowledge that the discussion section in the previous version lacked connection with the research questions, and we have thoroughly revised the manuscript based on the reviewer's suggestions.

If this manuscript is accepted for a major revision process, it should be largely rewritten and then undergo line-by-line comments in a second review iteration. The focus should first be on the following aspects:

- reworking the manuscript research questions (is it about the filter performance in different climates as the title suggests or about the three questions formulated at the end of the introduction, or both?)

Reply: Thanks for your sincere and constructive suggestions. We have thorough reworked the manuscript according to the reviewer's suggestion.

- a more comprehensive literature review on the technical literature regarding the research questions

Reply: Thanks for your sincere and constructive suggestions. We have thorough revised the manuscript according to the reviewer's suggestion.

- a more detailed description of the used particle filter method and why this filter is chosen to be the most suitable to answer the research question

Reply: Thanks for your sincere and constructive suggestions. We have provided a comprehensive explanation of the genetic particle filter process in our manuscript and compared its assimilation results to those obtained using particle filters with generic resampling methods in section 3.5.

- a more critical questioning of the results, in particular with respect to the 100-particle threshold (e.g. why in Fig. 7 a minimum exists at 100 particles).

Reply: Thanks for your sincere and constructive suggestions. In this study, we performed a sensitivity analysis of the Genetic Particle Filter (GPF) to particle number at eight sites. Our findings show that the root mean square error (RMSE) of the assimilation results decreased as the particle number increased. We observed that the RMSE reached a minimum value when the particle number was equal to 100. However, increasing the number of particles beyond 100 led to an increase in RMSE. In other words, 100 particles were sufficient to represent the model state. Excessive particles led to a poor assimilation result and increased computational burden.

Response to comments by Editor：

Dear authors,

This manuscript uses field data to improve snow modelling through the genetic particle filter algorithm. The topic is highly relevant and fits nicely within this SI. However, the manuscript requires significant improvements in order to be considered for possible publication in HESS. The Introduction section should be shorted, more concise and better highlight the research question and the need for this study, and use more updated references. The study sites must be better described, namely regarding the snow differences between the sites, so that we can understand the real application of your proposed method. The methodology requires relevant improvements to describe the genetic particle filter algorithm, and all the mathematical assumptions performed. The Results must be clearly presented and discussed. Discussion should clearly present the advantages of the proposed method comparing with others, and the limitations linked with the assumption performed. It is also important to compare the results with those from previous studies, and bring more references into this section. The Conclusions section must clearly identify the novelty and main messages of this study, and clearly identify why the proposed method is better than the available ones. Language editing is also required. More detailed comments have been provided by reviewers and must be considered in improving the manuscript. There was one late reviewer whose comments are provided bellow and should be also considered.

Reply: Thanks for your sincere and constructive suggestions. As you said, some significant weaknesses in this manuscript and we have thorough revised the manuscript according to the Editor and Reviewer's suggestion, thank you very much. For the language editing, we have purchased an English writing service for this manuscript. And we look forward to receiving your positive feedback this time. Thank you very much.