Referee #2

In general, the manuscript is well written with clear objectives, meticulous methods, and results. The study introduced a novel parallelization method to accelerate the SnowModel and apply it to simulations on a larger scale, which carries significant scientific significance. However, I am concerned that the scientific reproducibility and presentation quality of this manuscript should be improved before any publication with standards expected for GMD. Below, I will provide detailed comments on each section:

Thank you for your support and suggestions for improvement.

Section 2: While it briefly introduced SnowModel and the authors' motivation for parallelization, I suggest separating the introduction to SnowModel into its own section and incorporating schematic diagrams of the model's structure. These diagrams would assist readers in understanding the parallelization strategies discussed in Section 3.3, and the "Parallelization Motivation" could be a subsection within Section 2.2.

Thank you for your comment. We included a diagram that reflects the important submodules of SnowModel. We were a bit confused about the other suggestions. Currently, Section 2 provides an introduction to SnowModel with a page of text describing the model. Then Section 2.1 provides motivation for its parallelization. Then the text moves onto Section 3 (i.e. the Methods). Are you suggesting splitting up Section 2 into Section 2.1 (SnowModel) and Section 2.2 (Parallel Motivation) because currently Section 2.2 does not exist. After reading over both referee's comments regarding confusion about the structure of the paper, we propose the following changes.

1. Introduction
2. Background
   - 2.1. SnowModel
   - 2.2. Coarray Fortran
   - 2.3. Model Domains
   - 2.4. Parallelization Motivation
3. Methods
   - 3.1. Parallel Implementation
     - 3.1.1. Partitioning Algorithm
     - 3.1.2. Non-trivial Parallelization
       - 3.1.2.1. Wind and Solar Radiation Models
       - 3.1.2.2. Snow Redistribution
     - 3.1.3. File I/O
       - 3.1.3.1. Parallel Inputs
       - 3.1.3.2. Parallel Outputs
   - 3.2. Simulation Experiments

```
       3.2.1.   Parallel SnowModel Validation
       3.2.2.   Parallel SnowModel Performance
          3.2.2.1.   Strong Scaling
          3.2.2.2.   Code Profiling
          3.2.2.3.   CONUS Simulations
4.   Results
   4.1.   Parallel SnowModel Validation
   4.2.   Parallel SnowModel Performance
      4.2.1.   Strong Scaling
      4.2.2.   Code Profiling
      4.2.3.   CONUS Simulations
5.   Discussion
6.   Results

……
```

Section 3: This section provides a wealth of code examples and diagrams that effectively elucidate the parallelization methods. The readers with some programming background can easily grasp the details of the parallelization techniques. However, the Section 3 delves excessively into minutiae, potentially causing readers to become lost in the details. Consider shortening this section, focusing on key aspects.

Thank you for your comment. Section 3 was significantly simplified by deleting extra content relating to CAF syntax and algorithms used in some of the non-trivial parallelization techniques in hopes to not lose the reader, while still providing information relevant to its parallelization. Additionally, as discussed above, we added a methodology section for the simulation experiments in Section 3.

Section 4: The results presented in this section are somewhat confusing, raising concerns about the scientific quality and reproducibility of the study. Firstly, there is an overabundance of content related to model setup and evaluation metrics, which should not be presented as results. Furthermore, compared to Section 4.2, Sections 4.1 and 4.3 provide insufficient results, with a suspicion of excessive elaboration to magnify their importance.

Thank you for your comment. We moved the description of the experiments and evaluation metrics to Section 3.2. We are not sure what is meant by “insufficient results”. As discussed in response to referee #1’s comments, we have expanded the timing of the validation and performance experiments to make the results more meaningful. Additionally, we significantly simplified the text within the results

sections in hopes of not providing excessive elaboration. If there is anything we missed here, please let us know.

In **Section 4.1**, the description of the model setup occupies a disproportionate amount of space. The data provided to support validation conclusions are overly simplistic, such as "All variables across all processes produced RMSE values of 10^{-6}" (Lines 341-342). I would like to see more detailed model comparisons, preferably presented in graphical form. Otherwise, consider merging this section with others.

Thank you for your comment. We changed the text in Lines 341-342 to the following.

Comparing the serial output of each of the seventeen selected variables (see Appendix B for a list of those variables) to those of each experiment conducted with a different number of processes produced RMSE values of 10^{-6}

Additionally, we don't feel a graphical representation would be appropriate when output results are identical. We could provide an image of distributed SWE from a serial and parallel simulation on April 1st and then show the difference. However, if the difference is effectively zero everywhere, then it doesn't make for a very interesting visualization.

In **Section 4.2**, the authors present code profiling and speedup plots for three different stages, but I couldn't discern specific differences between "Distributed High Sync" and "Distributed Low Sync." I attempted to find an explanation in Section 3.4 but failed. Without a more detailed explanation, readers will struggle to understand the scientific significance of these results. For instance, it would be helpful to clarify what code optimizations improved process communication and reduced wait times.

Thank you for your comment. We moved the methodology of this section to be within Section 3.2. Additionally, we included a more succinct description of the different versions [i.e. "Distributed High Sync", etc]. We think that will make the methods and results pertinent to this section much clearer.

Section 4.3 displays spatial results and time series of SWE, but it lacks information on how other snow properties performed. To convincingly demonstrate that Parallel SnowModel successfully simulates distributed snow over CONUS, it is essential to provide additional output results for different variables.

Thank you for your comment. SnowModel is primarily used to simulate SWE. We will explore graphics that also look into snow density,

sublimation, and melt. However, SnowModel does contain many other hydrologic variables of interest.

Section 6: This section extensively references the work of others and highlights the relevance of this study to their work. However, I believe this content would be better placed within the Discussion section. The Conclusions section should provide a comprehensive summary of the study's work and results, offer conclusive remarks, and state the research's significance without excessive referencing.

Thank you. We switched content from the Conclusions and Discussion Sections in response to this comment.

In conclusion, the manuscript requires further improvement to meet the publication requirements of the journal, particularly regarding scientific quality and presentation quality. I therefore conclude with a **major revision** and hope that the revised manuscript will address the above-mentioned issues.

Thank you referee #2 for your comments. We are undergoing major revisions to the structure of the manuscript in an attempt to enhance the scientific and presentation quality.