EGUsphere, referee comment RC2
https://doi.org/10.5194/egusphere-2022-699-RC2, 2022

Comment on egusphere-2022-699

Anonymous Referee #2

Referee comment on "The Stochastic Ice-Sheet and Sea-Level System Model v1.0 (StISSM v1.0)" by Vincent Verjans et al., EGUsphere,
https://doi.org/10.5194/egusphere-2022-699-RC2, 2022

The goal of the paper is to present the recently developed Stochastic Ice-Sheet and Sea-Level System Model (StISSM), which adds stochastic parametrizations and tools to handle ensemble runs to the ISSM model. Available parametrizations are based on Gaussian noise, added at each (stochastic) time-step, and autoregressive processes, which can be used to represent surface mass balance and ocean forcing. The model capabilities are showcased targeting idealized experiments (including MISMIP+) and the Greenland ice sheet (GrIS).

Accounting for uncertainty in ice-sheet modeling is of paramount importance and tools like the one presented here are important and worthy of publication. However, I find the exposition hard to follow. Details about the stochastic model are at times buried in the presentation of the numerical experiments and some important details are missing. In my understanding the StISSM provides two stochastic processes in time (potentially different in each sub-domain): an autoregressive process and an one based on Gaussian noise. It also allows to run ensemble members in parallel, although it is not clear how the computational resources (cores, memory) are used. Statistical quantities (e.g. moments, p-values) are computed and reported in the Results section, but it is not clear whether these are computed directly by StISSM or how the data from the ensembles are collected. Finally, how is the stochastic layer implemented? How is it coupled to ISSM? Is it a driver to ISSM? Is it implemented in C++, Python or other languages?

A good part of the paper is devoted to using the StISSM for applying different stochastic parametrizations to two synthetic ice problems and a real ice sheet (Greenland). The numerical experiments are well thought out but I don't think it is particularly useful to target three different applications. I think it would have been better to target only one application (maybe a glacier) and show the effect of different choices of the parametrizations on the glacier evolution and mass balance. Targeting different (and complex) problems makes it harder to understand the impact of different parametrizations, without adding much in terms of explaining or demonstrating the stochastic model. Moreover, given that the main novelty introduced by StISSM is that it provides parametrizations, I would have expected more emphasis on 1) why to choose a specific stochastic (e.g autoregressive) processes for modeling, e.g., the surface mass balance (SMB), 2) how the stochastic process compares, in a statistical sense, with available time-series of SMB, 3) what is the impact of using a first-order versus an higher-order autoregressive process, and so on.

I'm not asking the authors to completely change the numerical experiments, but I would encourage them to significantly improve the presentation, clearly exposing the additional capabilities introduced by the StISSM, adding some important software and computational details and expanding on how to choose (calibrate?) the stochastic parametrizations for different forcing/parameters. Some additional details of the numerical experiments, that are important for completeness and reproducibility but that are not essential for explaining the new capabilities of StISSM ,could be moved to the appendix.

Here are some additional comments:

- Eq. (1). Is the Gaussian noise uniform in space even if the mean value is not? Please specify this in the text and discuss this choice. Please specify that the "mean" is intended in time, not in space (if I understand correctly)
- line 153: would it make sense to have different stochastic time steps for different parameters?
- line 164: OK, so the spatial stochasticity is introduced at the subdomain level. This should have been explained before, in the introduction and, in more detail in section 2.1 where it should be explained that eq. (1) is at the subdomain level.
- eq. (7): I do not fully understand the purpose of the intercept and trend terms. Also, what is the choice for beta_0 and beta_1 in the numerical experiments in sections 3?
- line 235: can you detail how you manage resources (nodes, cores, memory) when you run in parallel multiple members of the ensemble (each of the members might need to be distributed on several ranks). Do you use any strategy to reduce I/O and storage when running large ensembles? Any strategy to monitor the runs (e.g. what happens if a few of the 500 simulations in the ensemble fail?)
- eq (8) and (11): this is very minor, but I think that the use of squared terms "C_W^2" and "C_B^2" is poor notation. I know it is somewhat common, but it is misleading because it sort of implies that C_W and C_B have some physical meaning. Using the square to denote positive quantities (if that's the reason for the square) is hardly defensible because there are a lot of other physical variables (e.g. thickness) or coefficients (flow factor) that are positive (or nonnegative) and they are not denoted with a square of some other quantity. I would suggest dropping the square and using directly the coefficient C_W and C_B.
- Sections 3 and 4: The rigid separation of the "Model experiments" and "Results" sections makes it harder to follow the exposition. I think that the Results part should follow the Model Experiments part for each of the three examples.