Summary:

This manuscript presents ICESEE, a python-based, open-source data assimilation framework designed for coupling ensemble Kalman filter methods with ice sheet and Earth system models. The library aims to provide a modular and extensible framework for state and parameter estimation, with support for MPI-based parallelization, serial implementations of several EnKF variants, and containerized workflows. The authors demonstrate ICESEE through applications with Icepack and ISSM, representing two different coupling strategies. The manuscript also presents scaling tests for the ISSM coupling and argues that ICESEE can support scalable ensemble-based data assimilation for computationally intensive ice sheet simulations.

A general-purpose, open-source data assimilation framework for ice-sheet models would be valuable for the glaciology community, especially given the increasing use of ensemble methods for transient calibration, uncertainty quantification, and projection. The manuscript contains many useful implementation details and demonstrates substantial software-development effort. However, I think the current version needs revision before publication, particularly regarding the descriptions of EnKF equations and the presentation of test cases. Below, I present the main issues followed by more specific comments.

Main comments:

1. The mathematical formulation of the EnKF algorithms needs careful review and clarification (Section 2). For example, the manuscript appears to assign the dimension $n \times N_e$ to X_i^f, although X_i^f should represent a single ensemble member and therefore should have dimension n. The full ensemble matrix X^f would have dimension $n \times N_e$. This should be clarified throughout Section 2.

   The definition of the perturbed observation covariance is also ambiguous because D is described as containing columns d_i, although d_i denotes perturbed observations rather than zero-mean perturbations. In addition, the replacement of $HX'(X')^T H^T + \Upsilon\Upsilon^T$ by $(HX' + \Upsilon)(HX' + \Upsilon)$^T requires clarification because cross terms arise unless additional assumptions are made. Equation 10 is also unclear: if X^* is an $N_e \times N_e$ ensemble-space transformation, it should multiply the full ensemble matrix X^f, not a single state vector X_i^f. Similar dimensional issues appear in the deterministic filter equations. I recommend that the authors carefully revise Section 2, define the dimensions of all vectors and matrices, and ensure consistency with the cited EnKF formulations.

2. Section 2.3 describes a noise modeling strategy based on spatially correlated pseudo-random perturbations and observation-space perturbations constructed either from ensemble anomalies or pseudo-random fields. This may help maintain ensemble spread and reduce some artificial correlations. However, it is not clear how this strategy removes the need for localization. Localization addresses sampling-error-induced spurious correlations in the ensemble covariance, whereas the proposed noise modeling seems mainly controls ensemble spread and perturbation structure. The manuscript should

clarify whether the method truly replaces localization, approximates localization through prescribed spatial decorrelation of perturbations, or simply avoids explicit covariance localization in the presented experiments.

3. The current text in section 4 (application examples) sometimes makes it difficult to follow which parameters are directly updated by the EnKF, which variables are observed, and how the hybrid workflow improves results relative to EnKF-only approach. For the ISSM example, the authors should specify exactly which variables are assimilated and which variables are updated by the filter. If the hybrid workflow is a major contribution, the manuscript should include a comparison against a baseline EnKF without the inversion step.

Specific comments:

- Abstract "This matrix-free formulation eliminates the need for localization…". Please see the main comment above. I recommend revising this sentence. Avoiding explicit construction of the forecast covariance matrix is computationally valuable, but I am not convinced that it removes the effects of ensemble undersampling or spurious long-range correlations.

- Page 2 "A further major bottleneck identified in the literature…" Please add references for the literature here.

- Page 3 "Existing implementations are frequently tightly coupled to…". This sentence is not necessarily true since other DA software, such as DART, have already been coupled to many atmosphere/ocean models. Here, the contribution would be more specifically focused on ice sheet modeling and its coupling with inversion methods. Please revise this sentence.

- Page 4 "Stochastic ensemble filter, such as the EnKF,…" The acronym "EnKF" appears to have been used earlier as a more general ensemble Kalman filter.

- Page 7. 2.2.2. Ensemble Transform Kalman Filter (EnTKF). I think the acronym "ETKF" is more commonly used.

- Page 10. "In such cases, ICESEE uses the assimilated state variables as inputs to physic-based inversion procedures.." I am not convinced that inversion is necessarily a more physical approach than EnKF since inversion is also based on Bayesian approach with regularization.

- Page 11. "testing infrastructure" I do not see a clear description of this testing infrastructure in the manuscript.

- Page 14. "observation functions": What do "observation functions" refer to here?

- Page 16. “The figure also highlights how specific…” I recommend adding specific files and directory names associated with each workflow in Figure 3.

- Page 17. ISSM -> Ice-sheet and Sea-level System Model

- Page 22. “After about 60 years, the filter … constrained parameters that are not directly observed”: I think smb are observed according to the synthetic observation paragraph on page 21. Please see my main comment and specify which variables are observed and estimated.

- Figure 4: Could you also add changes in smb field?

- Page 22. “The ensemble spread around the mean remains nearly constant…”: In Figure 5, the ensemble spread appears to increase with time. It would also be helpful to include the observation uncertainty in Figure 5, so readers can compare the ensemble spread with the prescribed observational errors. I guess the initial spread is not enough or the initial mean is too far from the observation?

- Page 25. “to simulate the long-term evolution of a marine-terminating glacier in 3D”: What approximation is used for this test case? Is it a 3D model or 2D shelfy shallow approximation for this test case?

- Page 26. “exponential noise is added to the bed topography”: what does “exponential noise” mean here?

- Page 26. “ and a constant offset of 284 m is subtracted from”: Why is this offset applied?

- Page 26. “the bed topography conditioned on bed observations”: Why are bed observations assimilated if the initial bed is already derived using bed observations? I am also not sure why bed observations are assimilated only once. What happens if bed observations are not assimilated, or if they are assimilated multiple times?

- Page 26. “In addition, a uniform offset of 50 m …”: Why are these offsets applied?

- Page 26. “aircraft-based radar flight lines oriented along the x-direction”: is this the x-direction or the y-direction?

- Page 27. “the bed topography is observed directly only once at year 2”: Again, why is the bed topography assimilated only once?

- Page 28. “the grounding line position in the simulation with assimilation (solid line)”. I think this should be “without” assimilation.

- Page 29. “The bed corrections remain relatively small..” could you add changes in the bed variable in Figure 8 to support this?

- Page 30. “The inferred friction differences between the true and assimilated fields…”: is basal friction here inferred through inversion?

- Page 30. “The results show a large improvement….” This sentence is long and hard to follow. Please consider splitting it into two or three sentences. It would also be helpful to clarify whether the reduction in basal friction error is caused directly by the inversion step, while later improvements result from improved assimilated state variables being passed into subsequent inversions.

- Page 30. “Friction is set to zero beneath…” I thought the inversion was performed only on grounded ice.

- Page 32. Equation (20). Please check this equation. N_p is defined as “the number of ranks pinned to each ISSM instance,” which seems inconsistent because that quantity is already represented by model_nprocs. It may also be clearer to write the left-hand side as “total MPI ranks” rather than “ICESEE ranks,”

- Page 32. “This decline is primarily due to…”: What about the number of nodes? Can the increase in number of nodes also contribute to the decline in speed?

- Page 35. “leveraging physic-based cost function..”. Again, I do not think this method is more physical than EnKF DA.