Original reviewers comments are in *italic and black*, our answers are in blue.

1 Summary and Main Points

In the paper 'The future of Upernavik Isstrøm through ISMIP6 framework: Sensitivity analysis and Bayesian calibration of ensemble prediction', Jager and co-authors study several aspects associated with the evolution of the Upernavik Isstrøm Glacier, Greenland. Based on a statistical framework, numerical results obtained with the Elmer/Ice finite-element code, and observational data, they quantify the uncertainties associated with predictions of the future sea-level contribution. They improve the robustness of their analysis by considering a cross-validation step and by studying several weighting methods for assigning a likelihood score to the uncertain parameters.

I think that the paper will make a great addition to the scientific literature as it deals with an important topic, namely the quantification of uncertainties, and, more generally, the study of the methods that are used to produce such analyses. Nonetheless, I have a series of comments that I would like the authors to address prior to the publication of the manuscript. As described hereafter, those are mainly related to the form of the paper, rather than its scientific content.

2 General

My main comment is related to the way the paper is written. I have found the methods and results to be particularly interesting, but the style of the paper makes it quite difficult to grasp them efficiently. My main complaints are that the whole paper is very long (45 pages), that some parts are difficult to follow because of the lack of visual data, and that some subsections are particularly long. I would suggest the following changes:

• Streamlining the manuscript, in particular by focusing on the key points in each paragraph, and removing unnecessary repetitions.

• Focusing on the novel aspects of your study. To my understanding, these are the cross-validation (which I believe has not really been done previously in a glaciological context), and the use of different weighing methods.

• Adding figures/tables/schematics that allow to understand the content of the text in a visual and summarized way. For example, in Section 2, you could create a table with the different scenarios (Hpr, Cpr, Ppr), and for each of these scenarios specify the SMB that is used, as well as the front position and the uncertain parameters. For the observational data, you could create a table with the different types of observations that you have, their type, and where they come from.

• I wonder if the fparam weighing makes much sense, giving that this parameter is one of the uncertain parameters that are calibrated in the Bayesian process. Note that a classical way to favor specific values of fparam , given your knowledge of its importance, would be to change its prior distribution.

To reiterate, I find the paper to be both useful and significant. But I still think that it is important to improve its style, as it will help the audience to better understand the key points presented in the manuscript.

We thank the anonymous reviewer for their constructive review and positive comments. To improve the readability of the overall paper, we will make the following changes:

• We will try to clean up text that is redundant or heavy to read.

• We will change the structure of the headings in the results section to better highlight the two main parts of our study: the sensitivity analysis and the Bayesian calibration. We still think that these two parts are novel in our study, not only the novelties of the Bayesian calibration. No other study has explored the uncertainties of the ISMIP6 framework as exhaustively as we do here, with 3 SSPs and various RCMs.

• We will add two summarizing figures of the method. The first one will describe the different ensembles of the study to show their differences in forcing and summarize the different elements taken into account for the sensitivity analysis. The second one will summarize our methodology to produce robust Bayesian calibration. We hope this will help the readers and will also reduce the size of the text.

• We will add some justification for the use of the fparam weighting. It still makes sense for us to use it because it allows us to see the effect of the parameterisation developed in Jager et al. (2024) on the projections and shows that taking into account the effect of subglacial hydrology, at least in a parameterized way, significantly increases the projected mass loss of Upernavik Isstrøm. This also lets us evaluate this weighting against others to check if they can underscore the use of the parameterisation without the extensive detail used in the earlier study.

3 Specific comments

(1) [Line 24] It is a bit unclear at this stage what distinguishes the 'limited understanding of ice dynamics ' and 'uncertainties in Ice Sheet Models'. Maybe specify what you mean for the latter, e.g., 'uncertainties in the parameters of Ice Sheet Models'.

Yes it's quite similar, so we will change "ice dynamics" by "initial state".

(2) [Line 34] A paper that is missing for Antarctica is Bulthuis et al. (2019).

Agree, we will add it.

(3) [Line 63] 'initialisation' → 'initialization' as you use American English in your manuscript. Also check Lines 97, 158, 340, 546, 756, 774, 775, 792, and 864.

Thanks, we will correct it.

(4) [Line 65] The use of the active voice in this sentence is a bit weird here, given that the rest of the paragraph is written with the passive voice.

Agreed, we will change the sentence.

(5) [Line 168] I am guessing that the equation mentioned here should be Eq. 1, not Eq. 4.

Thanks, we will correct it.

(6) [Line 168] It is a bit confusing that the sensitivity indices Si are called 'first-order sensitivity indices' here, and not before. I would suggest discussing why the Si are called 'first-order indices', or directly mentioning Line 161 that the indices that you introduce are of first order. Otherwise, the

reader might wonder which indices you are talking about in this paragraph, as it is not clear that you are talking about the Si indices.

Agreed, we will change «the following indices» to «the first-order sensitivity indices» in the line 161 and change «Accurately computing sensitivity indices usually requires [...]» to «Accurately computing sensitivity indices of an order greater than one usually requires [...]» in the line 165.

(7) [Line 172] 'Y' needs to be written in italics (Y) here.

Thanks, we will correct it.

(8) [Equation 6] The first factor is not a prior distribution for the problem considered in the paper. Going back to Aschwanden and Brinkerhoff (2022), a possible name for this factor would be 'projection'.

The distinction between prior and posterior distributions (i.e., Bayes' theorem) appears later, implicitly, through the computation of the term $P(M|B)$ in equation (6). Specifically, Bayes' theorem writes

$$P(M|B)=\frac{P(B|M)P(M)}{P(B)}=\frac{P(B|M)P(M)}{\int P(B|M)P(M)\,dM}$$

where:

- *$P(M|B)$ is the posterior distribution;*
- *$P(B|M)$ is the likelihood distribution;*
- *$P(M)$ is the prior distribution.*

For all intents and purposes, you will find at the end of this review a few equations that show how, starting from (R1), I arrive at your equation (7).

We had a similar comment from the other reviewer and will relabel the two terms from "Prior" and "Posterior" to respectively "Model Prediction" and "Posterior Prediction."

(9) [Line 214] Ideally, you should define every variable that appear in the equations, so F_m^j, F_o^j, $Q^j_{m,i}$, … should be defined. To save space, it makes sense no to do so, but please at least mention in this paragraph that the subscript i is associated with the i-th member of the ensemble and that the superscript j is associated with the j-th observation.

Yes agreed, we will add the following sentence: «The subscript i is associated with the i-th member of the ensemble and the superscript j is associated with the j-th observation.»

(10) [Line 216] I am guessing there is an 'it' missing before 'is common' here.

Thanks, we will correct it.

(11) [Line 237] It really is a detail, but please avoid using fractions within the text. Instead, write the definition of wi as a full new equation, or write it as $w_i = P(B|M_i)/\sum^n_{k=1} P(B|M_k)$. Same comment for the factions that appear later on in the text.

Agreed, we will change it to a new full equation while we will change the writing of the other ones.

(12) [Equations (8)–(12)] I suggest removing equations (8)–(10), as these equations do not add much to the discussion, and might even appear unnecessarily technical. It seems to me that the

reader should be able to deduce from the Gaussian and independence assumptions that P(B |Mi) has the form shown in (11), which is quite standard.

Agreed, we will remove these equations.

(13) [Line 248] Technically, H is the measurement operator, not H(Mi) (which is the value taken by this operator when M = Mi).

Agreed, we will change this part of the sentence from «H(Mi) is the measurement operator corresponding to Qm,i» to «H is the measurement operator with H(Mi) corresponding to Qm,i»

(14) [Line 266] To be consistent, write f (RMSE, σ), not just f (RMSE).

We will delete this paragraph as it duplicate the one below.

(15) [Line 271] Please read again this paragraph, it seems that you repeat yourself.

Agreed. As said in the previous comment, we will delete the duplication.

(16) [Subsection 2.3.1] Overall, I think that this subsection is not structured in an efficient way: you first present the equations (12) and (13), corresponding to the 'classical' approach. Reading the beginning of this section, it seemed to me that you are going to use those expressions. But then you discuss limitations (which always is a real plus), and consequently modify you formulas. It might make more sense to directly state that while expressions (12) and (13) are the usual approach, you are not going to use them, and instead will use the formula (14) instead. On a related note, it is a bit surprising that you mention Line 249 that σ is the standard deviation of the observation error (while it is common, as you mention later, to include the model error in it). So maybe directly state the difficulty associated with σ, and that your equation (14) is a possible solution for it.

Thanks for the comment, we will restructure this section and rewrite partially some paragraphs for a smoother reading. We will start with the presentation of the different equations and mentioning than we will use equation 14. We will then explain the limitation of equation 13 and how the equation 14 allows us to bypass these limits.

About σ, we will add some elements in the paragraph dedicated to it. We will precise that it corresponds to the standard deviation of the observation error in equation 13, but in equation 14, it takes into account the structural error of the model as done in previous works (Murphy et al., 2009; Nias et al., 2019; Edwards et al., 2019).

(17) [Line 294] I wonder if the discussion of the assumptions that must be examined should not be included in the 'full-period weighting' item, Line 300.

Agreed, we will add it after the description of the full-period weighting.

(18) [Line 382] I do not agree with the contradiction indicated by the 'On the contrary' here: the fact that the sum of the first-order Sobol indices is greater than one does not prevent a substantial impact of specific parameter combinations. Furthermore, the fact that the sum of the first-order Sobol indices is smaller or equal than one does guarantee that the inputs are independent.

Agreed, it wasn't very clear that "on contrary" was there only to position the "smaller than 1". So, we will change the «On the contrary» to «otherwise».

(19) [Line 462] law → flaw .

Thanks, we will correct it.

(20) [Line 462] The fact that the priors and posteriors distributions are similar for several parameters is an important result. Maybe you could elaborate on that, both in terms of the interpretation that you give to this observation, and the conclusions that can be drawn for it.

We mention it briefly in the appendix B3: «In hindsight, our initial choice of distribution for these three parameters proves to be suitable due to the absence of significant changes observed in their posterior distributions.». We agree that it's an important result for our future perspectives but remains restricted to our study. Not everyone will use the same range of parameters because it can be specific to our catchment area. Moreover, you may need different parameters if you are not studying a tidewater glacier, you use an other ISM or you don't use the same framework than us.

(21) [Line 468] As before, this 'posterior ensemble' is a bit confusing as you are looking at the distribution of ice mass discharge, rather than the distributions of the inferred parameters (which have been analyzed in the previous subsection). Maybe use another name for this subsection, or precise in that name that you are going to talk about SLR predictions.

To clarify the overall structure, we have removed this sub-title (see answer to main comments).

(22) [Figure 5] This figure is difficult to read. Consider using brighter colors and larger labels.

Agreed, we will add hatch patterns and increase the label fontsize.

(23) [Line 553] It's → It is.

Thanks, we will correct it.

(24) [Line 595] dynamics modeling community → ice-sheet dynamics modeling community?

Thanks, we will correct it.

(25) [Line 660] it's → it is.

Thanks, we will correct it.

(26) [Line 746] Would that still be true if you looked beyond 2100? I have in mind the study of Brondex et al. (2017, 2019) which showed that the form of friction laws does have a strong impact (in particular, there is a significant difference between the regularized Coulomb and Budd laws).

To make clearer our message about the past reproduction of observations, we will add few words at the end of the sentence: «[...] in reproducing past acceleration of UI».

Otherwise, to answer the question, we think than our figure B1 gives some keys: the impact of the shape of the friction law does not have a major impact in 2100 unlike the use of the parameterisation developped in Jager et al. (2024). In our point of view, the main difference with Brondex et al. is this choice of proxy for the effective pressure N. In their case, they assume a perfect connection to the ocean with N proportional to the height above flotation, while Jager et al. (2024) and Joughin et al. (2019) both use a cut-off (the distance to the front in our case, the height above flotation in the case of Joughin). It means than far enough to the grounding line/the front position, the friction at the base is independent to the height above flotation/distance to the front regardless to the friction law used. On the contrary for Brondex et al., in case of Budd law, it will lead to a dependence to the effective pressure everywhere. However, this will not be the case for a regularised coulomb law because you will have some areas in a Weertman regime where friction only depends on the friction parameter A_s.

(27) [Line 789] Maybe add that 'SSA' also stands for Shallow-Shelf Approximation.

Agreed, we will add the following words: «also called Shallow-Shelf Approximation»

(28) [Lines 811, 843, 876] regularisation → regularization.

Thanks, we will correct it.

(29) [Line 840] Maybe add that this value of u0 is similar to the one chosen in Joughin et al. (2019).

Agreed, we will add the following words: «[...], which is similar to the median value of our previous study and to the one used in Joughin et al. (2019).»

(30) [Line 884] law → flaw .

Thanks, we will correct it.