

Referee comment: Joint Bayesian Calibration of Frontal Ablation and Surface Mass Balance in Global Glacier Models

The paper describes a coupled model approach to study marine-terminating glaciers. The workflow joins three established models for climatic mass balance and ice dynamics to estimate mass balances and frontal ablation. The model framework makes use of four tunable parameters, which are tuned in a thorough Bayesian calibration.

The manuscript is well structured and written. The individual model components are only briefly described in the manuscript, as they are well documented in the original model papers. However, there are a few important assumptions that should be stated within the manuscript (see specific comments). The overall presentation and figure quality are good, with a few minor adjustments needed.

General remarks:

Terminology: The title refers to *surface mass balance*, whereas the manuscript predominantly uses *climatic mass balance* for the PyGEM component. I assume that these are used interchangeably here, since the PyGEM formulation described includes accumulation, surface melt/ablation, and refreezing, while internal and basal mass-balance components are neglected. If this is the intended meaning, I suggest stating this explicitly when climatic mass balance is first introduced, to avoid ambiguity.

Ensemble: Make sure that it is clear throughout when “ensemble” is mentioned whether the context refers to the Bayesian ensemble or the ensemble of climate models (for example, Fig. 1). Especially in the figures, I would rather have it stated one more time than not. In the manuscript text it is usually clear.

Specific remarks:

Line 77ff: Detailed information can be found in the referenced papers, but it is not clear a) if the temperature bias correction is acting on the threshold for solid/liquid precipitation, refreezing, and melting, or only on some of these processes. Based on the references, if this is true, change it to the “climate/ERA5” model temperature bias. This makes it clear where in the model chain the bias correction is applied.

b) State that the degree-day factors of ice and firn are dependent on the one for snow, and therefore only the snow degree-day factor is calibrated. In the current manuscript, the reader could easily believe that the other two are kept constant. (Section 3.1 in Rounce et al. 2020a)

Lines 125ff: A small flowchart would increase understandability and serve as a quick reference during the reading process.

Section 2.3: Moving the extended description of your Bayesian calibration method to the supplement is good. Just make sure that all main assumptions of your method are given in the main manuscript. “For simplicity we have here made the typical strong

constraint assumption (see Evensen et al., 2022) that the forward model is perfect other than errors in the parameters.” This may need to be stated in the main text prior to/around line 144.

Line 250: State which three variables are meant; on first read, it was not immediately clear to me, and I had to go to the figure and back to the lines above.

Figure 3 and corresponding text: There are outliers with much worse metrics. Maybe the plot could be altered to include symbols/color shading for each glacier, so one could see if it is the same ones that deviate strongly from the 1:1 line in 3a, 3c, and 3e. If so, only highlighting those could also be an option. (Do we know why?) The manuscript text need not contain a full discussion of those, but it would be possible to discuss whether the “bad” performance occurs for each variable or only one.

Section 5, Lines 211ff/Section 6.2: Please clarify whether the parameters obtained from the Bayesian calibration using ERA5 are also used for the simulations driven by the individual climate models, or whether the parameters are recalibrated separately for each climate model. In particular, I assume that the calibrated temperature bias correction (T_{bias}) and precipitation correction factor (k_p) obtained with ERA5 are applied unchanged to all climate models. If this is the case, this assumption should be stated explicitly.

Using the same bias-correction parameters for all climate models and ERA5 is a simplification, since the magnitude and structure of model-specific biases may differ between the climate models. However, given that the primary purpose of the paper is to demonstrate the Bayesian calibration methodology and that the projections are presented mainly as a proof of concept, I consider this assumption reasonable. Nevertheless, it should be clearly stated and discussed as a limitation.

Figure 4: If this figure contains glaciers which are excluded later in the process (L 312 and L 322) for the analysis, highlight those in the figure. It could well be that they are not calculated at all and are therefore not shown here anyway, but I cannot fully figure this out. Does the cumulative mass change (Fig. 5) include the unconverged ones? Clarify which glaciers are excluded where and why more clearly.

Why are you plotting only the length change per glacier (Fig. 4), while climatic mass balance and frontal ablation are only shown cumulatively (Fig. 5)?

Section 6.3: Having a substantial number of surge-type glaciers in the dataset makes the entire approach less robust in its findings. Can you comment and add to what degree surge processes are occurring during the “training period” beyond the statement that they “showed limited influence”? Add a discussion about how, if any surges occur during the 19 years, the terminus movement actually impacts the calibration approach.