

General Comments:

The Introduction should be expanded to include recent work on RCM emulators. For instance, the study by Bochow et al. (2025) develops an RCM emulator directly relevant to this work. A thorough comparison with such studies would better position the contribution of the manuscript and clarify its novelty.

- We thank the reviewer for this helpful suggestion. We have expanded the Introduction to include recent related work on RCM-based machine-learning approaches, specifically Bochow et al. (2025). We now clarify that Bochow et al. developed a physics-constrained ML framework for downscaling monthly Greenland surface mass balance and surface temperature fields from MAR, whereas our study focuses on the direct emulation of daily surface meltwater production and on its interpretable attribution using SHAP.

XGBoost is widely used but relatively old, and all cited references are more than five years old (Lines 85-100). Could the authors clarify the rationale for

selecting this specific model? Additionally, the manuscript would be strengthened by including comparisons with other commonly used machine learning

approaches, such as Support Vector Machine (SVM), and potentially with more advanced deep neural network models

- We thank the reviewer for this helpful comment. We have revised the manuscript to clarify the rationale for selecting XGBoost. Specifically, we now note that our emulator is trained on a large tabular regression dataset composed of surface energy balance, meteorological, and geographic predictors, for which gradient-boosted tree methods remain strong and computationally efficient baselines. We also clarify that interpretability is a central objective of this study, and that XGBoost is particularly advantageous in this context because it enables efficient SHAP-based attribution analysis (line 92).

Specific Comments:

Providing a conceptual overview diagram of the RCM emulator would greatly enhance the clarity and readability of the manuscript.

- This is now fig 1

Lines 29-30, 119, 191: The references are neither chronological nor alphabetically arranged by first author.

- References are now ordered alphabetically

Line 37: Remove the extra space before the semicolon.

- This has been removed

Line 43: Spell out "XGBoost" in full at first mention.

- This have been changed to eXtreme Gradient Boosting (XGBoost)

Lines 102-125: The attribution analysis using Shapley coefficients is a key novelty but currently quite dense and may be difficult for readers to follow. Consider breaking it into smaller paragraphs for readability. Additionally, including a schematic or figure illustrating how SHAP values are computed and interpreted would help readers unfamiliar with the method.

- We thank the reviewer for this suggestion. We have revised the SHAP/attribution section to improve readability by breaking the explanation into smaller paragraphs and streamlining the description of the method. In particular, we now separate (i) the motivation for using SHAP, (ii) the interpretation of Shapley values, and (iii) the specific attribution metrics used in this study. (section 2.3). Concerning the schematic: we would prefer not to add this figure as Shapley is widely used in the literature and there are several schemes that can be accessed with an easy search on the web. We feel that adding a single figure without having the proper space to contextualize would imbalance the paper.

Line 189: Clarify the exact years used for each data subset (training, validation, test).

- We thank the reviewer for this comment. We have clarified that the training, validation, and test subsets were obtained from a randomly shuffled split of the full dataset, rather than corresponding to specific years. The revised text now states the exact proportions used for training, validation, and testing.

Figure 2 shows linear correlations between predictors and predictand. However, the relationships among surface energy balance components are highly non-linear. Using simple linear correlation may underestimate these dependencies. Consider using more robust measures that capture non-linear relationships

- We thank the reviewer for this comment. We have clarified the text around Figure 2 to state that the correlation analysis was used to identify highly correlated predictors and reduce redundancy in the input feature set prior to training.

