

A transfer learning approach for simulation of Greenland Ice Sheet meltwater runoff

Maya Lillian Maciel-Seidman¹, Jonathan Charles Ryan², Sarah Eileen Esenther³, Asa Kristina Rennermalm⁴, Laurence Smith⁵, Xavier Fettweis⁶

1. Duke University

2. Duke University

3. Brown University

4. Rutgers, The State University of New Jersey

5. Brown University

6. University of Liège

Abstract

Surface melt and runoff are responsible for much of the Greenland Ice Sheet's mass loss over the last few decades. However, complex hydrological processes such as refreezing, percolation, and lateral flow make it challenging to accurately simulate meltwater runoff with physically-based models. In this study, we present a new approach for predicting daily meltwater runoff on the Greenland Ice Sheet using the principles of transfer learning. We first develop deep learning models to emulate the regional climate model Modèle Atmosphérique Régional (MAR). We then fine-tune these emulators with in-situ observations of meltwater runoff from three catchments on the Greenland Ice Sheet. We find that the inclusion of observational data during the fine-tuning stage yields a 59% reduction in meltwater runoff error rate relative to emulators that are only trained using climate model data. Across all three catchments, the average error rate of the MAR emulators was $\pm 19\%$, while the average error rate of the TL models was $\pm 12\%$. We also find that models developed using transfer learning can generalize to unseen catchments, provided that they are trained using in-situ observations from the same region. Finally, we find that the transfer learning models can generalize to unseen years and fill gaps in proglacial discharge measurements, although this works best when they are trained on many (i.e. ≥ 2) years of observational data. Overall, our study demonstrates the advantages of conducting transfer learning by fine-tuning deep learning emulators with in-situ observations to yield more accurate predictions of Greenland Ice Sheet meltwater runoff.

Keywords

deep learning, meltwater runoff, supraglacial hydrology, Greenland Ice Sheet, ice sheet processes

A transfer learning approach for simulation of Greenland Ice Sheet meltwater runoff

M.L. Maciel-Seidman¹, J.C. Ryan¹, S.E. Esenther², Å.K. Rennermalm³, L.C. Smith^{2,4}, X. Fettweis⁵

¹Division of Earth and Climate Sciences, Nicholas School of the Environment, Duke University, Durham, NC, USA

²Department of Earth, Environmental and Planetary Sciences, Brown University, Providence, RI, USA 02912

³Department of Geography, Rutgers, The State University of New Jersey, New Brunswick, NJ, USA 08901

⁴Institute at Brown for Environment and Society, Brown University, Providence, RI, USA 02912

⁵Laboratory of Climatology, SPHERES Research Unit, Geography Department, University of Liège, Liège, Belgium

13

Corresponding author: Maya Maciel-Seidman (mlm211@duke.edu)

Key Points:

- We develop a transfer learning approach that fine-tunes deep learning emulators for more accurate prediction of ice sheet meltwater runoff
- We find that, relative to emulators, transfer learning effectively reduces error rate in total annual meltwater runoff simulation
- We evaluate the ability of our approach to generalize spatially and temporally to fill data gaps in meltwater runoff observations

Abstract

Surface melt and runoff are responsible for much of the Greenland Ice Sheet's mass loss over the last few decades. However, complex hydrological processes such as refreezing, percolation, and lateral flow make it challenging to accurately simulate meltwater runoff with physically-based models. In this study, we present a new approach for predicting daily meltwater runoff on the Greenland Ice Sheet using the principles of transfer learning. We first develop deep learning models to emulate the regional climate model Modèle Atmosphérique Régional (MAR). We then fine-tune these emulators with in-situ observations of meltwater runoff from three catchments on the Greenland Ice Sheet. We find that the inclusion of observational data during the fine-tuning stage yields a 10–59% reduction in meltwater runoff error rate relative to emulators that are only trained using climate model data. Across all three catchments, the average error rate of the MAR emulators was $\pm 19\%$, while the average error rate of the TL models was $\pm 12\%$. We also find that models developed using transfer learning can generalize to unseen catchments, provided that they are trained using in-situ observations from the same region. Finally, we find that the transfer learning models can generalize to unseen years and fill gaps in proglacial discharge measurements, although this works best when they are trained on many (i.e. >2) years of observational data. Overall, our study demonstrates the advantages of conducting transfer learning by fine-tuning deep learning emulators with in-situ observations to yield more accurate predictions of Greenland Ice Sheet meltwater runoff.

Plain Language Summary

Meltwater runoff from the Greenland Ice Sheet has increased since the 1990s, contributing to rising sea levels. Currently most projections of meltwater runoff are determined using physically-based regional climate models. However, these models have been found to be biased in their predictions of meltwater runoff. In this study, we develop an alternative approach for predicting Greenland Ice Sheet meltwater runoff using the principles of transfer learning models. These models first learn general relationships between atmospheric and surface variables and meltwater runoff as predicted by a regional climate model. Then, we mitigate biases in these models by allowing them to additionally learn from direct measurements of meltwater runoff. Our transfer learning models yield more accurate predictions of meltwater runoff, as compared to models that only learn from regional climate models. Our transfer learning models can additionally predict meltwater runoff in unseen areas of the ice sheet and unseen years, and can fill gaps in records of meltwater runoff measurements. Our study therefore shows that transfer learning models may improve climate model predictions of meltwater runoff, allowing us to better understand Greenland Ice Sheet contributions to global sea level rise.

1 Introduction

Greenland Ice Sheet mass balance has become increasingly negative over the last few decades, contributing roughly 0.66 mm, annually, to global sea-level rise between 2002–2017 (Otosaka et al., 2023). A large component of this mass loss is attributed to meltwater runoff which was 21% higher between 2011–2020 than the long-term (1980–2010) mean (Slater et al., 2021) in response to warmer air temperatures (Noël et al., 2021), melt-albedo feedbacks (Ryan, 2024),

and reductions in firn pore space (Machguth et al., 2016). Enhanced meltwater runoff not only contributes to sea-level rise, but also to the weakening of the Atlantic Meridional Overturning Circulation (AMOC) (Rahmstorf et al., 2015; Swingedouw et al., 2022). Accurate projections of meltwater runoff are therefore crucial for quantifying the contribution of the Greenland Ice Sheet to global sea-level rise, determining the response of ice sheet surface mass balance to climate change, and investigating ocean circulation patterns that influence the Earth's climate system.

Current approaches to determine meltwater runoff from the Greenland Ice Sheet rely heavily on physically-based regional climate models (RCMs). Polar RCMs including RACMO, HIRHAM, and MAR (Christensen et al., 2007; Timmermans et al., 2026), as well as global reanalyses such as MERRA-2 (Lim et al., 2016) are widely used to study ice sheet-atmosphere interactions and Greenland Ice Sheet processes. Surface melt and runoff outputs from these RCMs have been employed for investigating supraglacial flow routing (Gleason et al., 2021; Hill & Dow, 2023; Yang et al., 2020), characterizing subglacial hydrological systems (Karlsson et al., 2021; Mankoff & Tulaczyk, 2017; Yang et al., 2020; Young et al., 2022), and modeling submarine melt at tidewater glacier calving fronts (Cook et al., 2022; Slater & Straneo, 2022; Xu et al., 2012). However, despite their wide usage, accurately simulating meltwater runoff with physically-based models is challenging due to complex processes including retention, refreezing, and percolation. Incomplete parameterization of these processes may cause biases in RCM predictions of meltwater runoff (Cooper et al., 2025). Evidence for model uncertainties can be found in intercomparison studies which demonstrate that estimates of meltwater runoff from RACMO, HIRHAM, and MAR range from 302 to 491 Gt yr⁻¹ for the 1980–2012 period (Fettweis et al., 2020) and by 964 to 1735 Gt yr⁻¹ by 2100 (Glaude et al., 2024). Likewise, in-situ observations of meltwater runoff measured at a catchment in the southwestern ablation zone during July 2015 and 2016 show that RCM runoff ranges from 7% lower to 58% higher than measured (Cooper et al., 2025; Smith et al., 2017). Finally, remote sensing observations of supraglacial lake filling rates have demonstrated that RCMs overestimate runoff by +106 to +123% in the same region (Yang et al., 2019). These issues motivate the exploration of alternative approaches for predicting Greenland Ice Sheet meltwater runoff.

One alternative strategy for predicting meltwater runoff is machine learning, in which algorithms learn relationships between inputs (e.g. climate) and outputs (e.g. meltwater runoff) without requiring explicit physical parameterization. The advantage of this approach is that it allows complex processes (e.g. refreezing, retention etc.) to be captured implicitly from in-situ data alone (Rosier et al., 2023). Another advantage of machine learning models is their relatively low computational costs. While RCMs are computationally expensive and require extensive spin-up times, machine learning models have faster training speeds, enabling

uncertainties to be explored in greater detail. Machine learning may therefore offer a viable and efficient alternative to process-based models for simulating meltwater runoff.

Recent studies have applied machine learning to emulate surface melt across the Greenland Ice Sheet (Schlager et al., 2026; Tedesco et al., 2026). Schlager et al. (2026), for example, trained neural network models using atmospheric (e.g. downward shortwave radiation) and surface (e.g. albedo) variables from HIRHAM5 to predict HIRHAM5 surface melt, while Tedesco et al. (2026) trained an XGBoost model using similar variables from MAR to predict MAR surface melt. Both studies developed RCM “emulators” capable of accurate and computationally efficient surface melt prediction. However, since emulators are trained exclusively on RCM outputs, their accuracy is inherently limited by the RCM uncertainties from which they were trained. Emulators alone are therefore insufficient for reducing uncertainty in meltwater runoff projections from the Greenland Ice Sheet.

Reducing the dependence on biased training data requires more advanced machine learning strategies. One such approach, termed transfer learning (TL) (Bozinovski & Fulgosi, 1976), uses a two-phase process to train models. In the first phase, a model is pre-trained on data from one source task, and in the second phase it is fine-tuned with data from a target task (Han et al., 2021). TL is especially useful when there exists a dense dataset for the pre-training stage, but a sparse dataset for the fine-tuning stage (Hosna et al., 2022). This approach may be especially relevant for predicting meltwater runoff since a model could be pre-trained with large amounts of RCM outputs and fine-tuned with sparse in-situ measurements of meltwater runoff which are more accurate. In this case, the source task is learning the general relationships between RCM outputs of atmospheric variables and meltwater runoff, and the target task is predicting actual observations of meltwater runoff. Conducting TL with in-situ observations may therefore provide a viable strategy for reducing systematic biases in machine learning models that currently rely heavily on RCMs (Eyring et al., 2024). However, despite the promises of transfer learning, it has yet to be applied for predicting meltwater runoff from the Greenland Ice Sheet.

In this study we investigate the ability of TL to accurately predict daily meltwater runoff from the Greenland Ice Sheet at the catchment scale. To do this, we pre-train a deep neural network on daily atmospheric outputs from the polar RCM *Modèle Atmosphérique Régional* (MAR) version 3.14. Next, we conduct TL by fine-tuning the deep neural network with direct observations of meltwater runoff from the southwest (Rio Behar Internally Drained Catchment and Akuliarusiaruup Kuua River) and northwest (Minturn River) Greenland. We then evaluate the potential of our approach in three ways. First, we assess the ability of our TL models to generalize across the Greenland Ice Sheet by training with in-situ data from one catchment and predicting meltwater runoff at two unseen catchments. Second, we assess the ability of our TL

models to generalize temporally by withholding years of training data and predicting those withheld years. Third, we assess the ability of our TL models to fill gaps in daily resolution meltwater runoff data. Overall, our study advances the use of deep TL for predicting Greenland Ice Sheet meltwater runoff, presenting opportunities to investigate sources of bias in physically-based models and predicting meltwater runoff from other ungauged catchments.

2 Data and Methods

We carried out our TL approach in two separate phases (Figure 1). We first developed deep learning models to emulate MAR version 3.14 (hereafter referred to as “MAR emulators”) for three catchments on the Greenland Ice Sheet: Rio Behar, Akuliarusiarsuup Kuua (AK4), and Minturn). We then fine-tuned our MAR emulators with in-situ observations of meltwater runoff from rivers that drain the three catchments (hereafter referred to as “TL models”). In the subsequent analysis, we investigate the spatial and temporal generalizability of these TL models for predicting meltwater runoff at a catchment-scale.

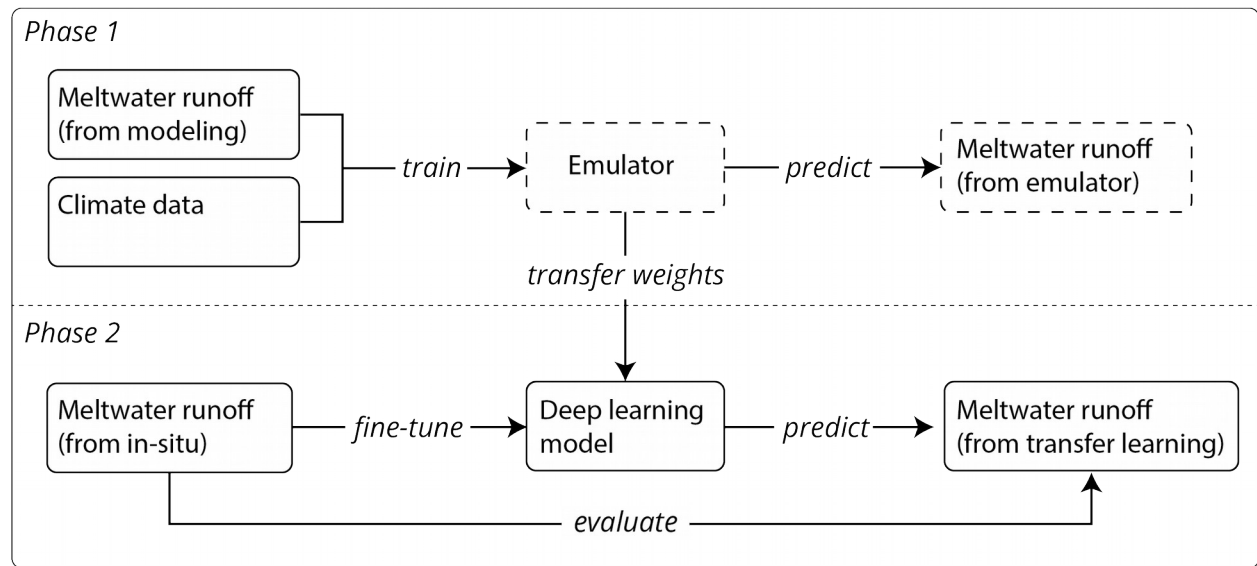


Figure 1. Conceptual overview of modeling approach employed in this study. Dashed boxes indicate intermediate products. Previous studies have carried out RCM emulation (Phase 1). Our study advances these studies by applying transfer learning (Phase 2) to improve the accuracy of meltwater runoff predictions.

2.1 Data sources

2.1.1 MAR data

To pre-train our MAR emulators, we used daily outputs from MAR version 3.14 (hereafter referred to as MAR) for the years 2000–2024 at a 10 km spatial resolution. MAR was forced with European Centre for Medium-Range Weather Forecasts Reanalysis v5 (ERA5) global

reanalysis and is fully coupled with the Soil Ice Snow Vegetation Atmosphere Transfer (SISVAT) model (Fettweis et al., 2017; Hersbach et al., 2020; Ridder & Gallée, 1998) allowing to explicitly simulate meltwater production, retention and percolation. Runoff occurs when there is bare ice at the surface or when the snowpack is liquid water saturated. However, lake/river hydrology is not taken into account and the spatial resolution used here ($10 \times 10 \text{ km}^2$) disallows to resolve small ice caps or glaciers tongues. MAR has been extensively used to provide meltwater runoff present day hindcast and future projections over the Greenland Ice Sheet (e.g. Fettweis et al., 2017, 2020; Lefebvre et al., 2005; Rae et al., 2012).

2.1.2 In-situ meltwater runoff

We used daily discharge data from three supraglacial rivers to develop our transfer learning approach: Rio Behar, AK4, and Minturn (Figure 2). We selected these records because their catchments are 1) largely glaciated (75–90%) and 2) have been accurately delineated by previous studies (Esenther et al., 2026; Rennermalm et al., 2012, 2017; Smith et al., 2017, 2021). Other rivers in Greenland have been gauged (Bahbah et al., 2025; Kondo, 2020; Tedstone, 2017) but were deemed inappropriate for this particular study either because their catchments are less glaciated (i.e. $<25\%$) or because of large catchment uncertainty (e.g. Watson River) (Mankoff et al., 2020). In-situ meltwater runoff was measured at Rio Behar catchment (67.04° , -49.02°) from July 20–23, 2015 and July 5–13, 2016 by an Acoustic Doppler Current Profiler (ADCP) suspended across a supraglacial river immediately upstream of the catchment's terminal moulin (Smith et al., 2017, 2021). Its 65 km^2 catchment (65 km^2 and 100% glaciated) was delineated from high-resolution images and digital elevation models of the ice surface from DigitalGlobe WorldView-1 and WorldView-2 satellites (Smith et al., 2017). At the AK4 catchment, in-situ measurements of proglacial river stage were collected by a Solinst Levelogger at station AK-004-001 (67.15° , -50.11°) along the Akuliarusiarsuup Kuua River's northern tributary from June 2008–December 2016 (Rennermalm et al., 2012, 2017). These stage measurements were converted to discharge with a rating curve that was constructed from 36 observations of in-situ discharge observations that were collected between 2008–2012 (Rennermalm et al., 2012, 2017). The AK4 catchment (34 km^2 and 90% glaciated) was delineated by a previous study using a DEM (Mankoff, 2025). In-situ proglacial river stage measurements were collected for the Minturn catchment (78.59° , -68.99°) by a Sutron Constant Flow bubbler from July 2019–December 2022 (Esenther et al., 2023). The Minturn ($3,721 \text{ km}^2$ and 75% glaciated) catchment boundaries were delineated from ArcticDEM (Goldstein et al., 2023) with small corrections along surface river networks made using RiverNet/SegConnector algorithm maps of supraglacial stream networks derived from WorldView satellite images (Esenther et al., 2026).

198

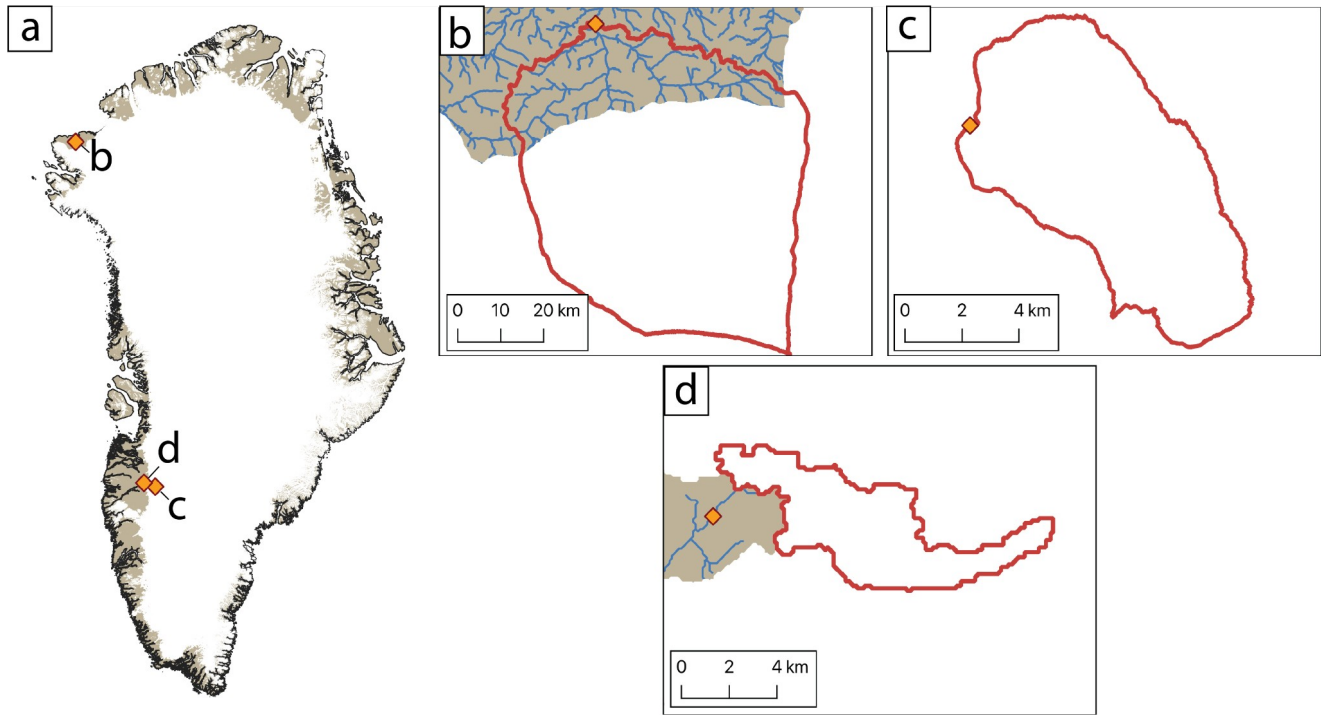


Figure 2. a) Locations of RB, AK4, and Minturn study catchments on the Greenland Ice Sheet. b) Zoomed inset showing Minturn catchment. c) Zoomed inset showing Rio Behar catchment. d) Zoomed inset showing AK4 catchment. Red lines represent catchment delineations (Esenther et al., 2026; Goldstein et al., 2023; Smith et al., 2017), blue lines represent proglacial streams (Mankoff, 2025), white background represents ice sheet (Citterio & Ahlstrøm, 2025), beige background represents land (Gerrish, 2020), and orange diamonds represent locations of the in-situ gauges (Esenther et al., 2023; Rennermalm et al., 2017; Smith et al., 2017, 2021).

2.2 Methods

2.2.1 Pre-Processing

To derive modeled climate data and meltwater runoff for each catchment, we intersected gridded 10 km daily MAR outputs with our delineated catchment polygons. Since catchment polygons often contain grid cells that are only 1) partially glaciated and 2) partially intersect MAR grid cells, we weighted the climate and meltwater runoff data before computing catchment means (Fig. S1). To do this, we first computed the ice sheet fraction of each MAR grid cell using the MAR ice sheet mask, which represents ice sheet area as percentage of total grid cell area (Fig. S2). Next, we computed the fractional coverage of MAR grid cells intersecting each catchment. We multiplied these two weights to produce the effective glaciated fraction of each MAR grid cell within each catchment. We then computed weighted-area catchment means of climate variables for the 2000–2024 period at a daily resolution.

2.2.2 Model architecture of catchment-scale MAR emulators

We selected a multilayer perceptron (MLP) (Rosenblatt, 1962) as the underlying model architecture for our MAR emulators and TL models. MLPs are feedforward deep neural networks consisting of an input layer, multiple hidden layers, and an output layer, in which each layer contains a set of neurons (Martinez-Ramon et al., 2024; Rosenblatt, 1962). Our MLPs consist of 4 dense hidden layers, each containing 64 nodes that are activated by Rectified Linear Unit (ReLU) functions. In between each Dense hidden layer is a dropout layer with a dropout rate of 0.2. Training was performed using the “Adam” algorithm (Kingma & Ba, 2017) with a learning rate of 0.001 for 1,000 epochs with a batch size of 32. We implemented early stopping with a patience of 20 epochs, which stops training before 1,000 epochs if the MLP’s loss stops improving and remains unimproved for 20 epochs. We used Mean Squared Error (MSE) as the MLP’s loss function. All MLPs were developed using TensorFlow 2.21.0 and Keras 3.13.2 and trained on an iMac with an Apple M4 chip and 32 GB RAM using 25 years (2000–2024) of catchment area-weighted mean MAR outputs ($n=9132$). A separate MAR emulator and TL model were developed for each catchment.

2.2.3 Systematic feature selection

We selected features for training our MAR emulators and TL models using a systematic feature selection approach (Bolón-Canedo & Alonso-Betanzos, 2019). To do this, we initially examined the explained variance (R^2) between MAR meltwater runoff and atmospheric and surface outputs that previous studies have identified as important predictors (Schlager et al., 2026; Tedesco et al., 2026) (Fig. S3). We then performed systematic feature selection on different predictor sets, i.e., different combinations of MAR atmospheric outputs, containing MAR atmospheric outputs with the highest R^2 values (see Table 1). For each predictor set, we produced an ensemble of 100 MLPs trained with MAR outputs for the years 2000–2024. For each ensemble, we calculated the ensemble Bayesian Information Criterion (BIC) which is a measure of model performance that accounts for model complexity by penalizing models based on the number of parameters (Schwarz, 1978). We defined the final predictor set as the model with the lowest BIC, allowing us to choose predictor sets that balance both model error and complexity (Neath & Cavanaugh, 2012).

2.2.4 Analysis of transfer learning models

We evaluated the ability of our TL models to generalize runoff predictions in three separate ways: 1) spatially (i.e. to unseen catchments), 2) temporally (i.e. to unseen years), and 3) ability to perform time-series gap-filling (e.g. caused by stream gauge failure). To evaluate spatial generalization, we used a TL model trained on one catchment to predict meltwater runoff in another. We then repeated this step for all combinations of catchments for all years. To evaluate temporal generalization, we performed Leave-One-Out Cross-Validation (LOOCV)

(Stone, 1974) for each catchment, withholding each year from fine-tuning and predicting meltwater runoff for the held-out year, with an ensemble of 100 predictions made per year. We evaluated the error rate of the spatial and temporal generalizability experiments using Equation 1:

$$Error\ rate = \left(\frac{|Q_{i-situ} - Q_{model}|}{Q_{i-situ}} \right) \quad (1)$$

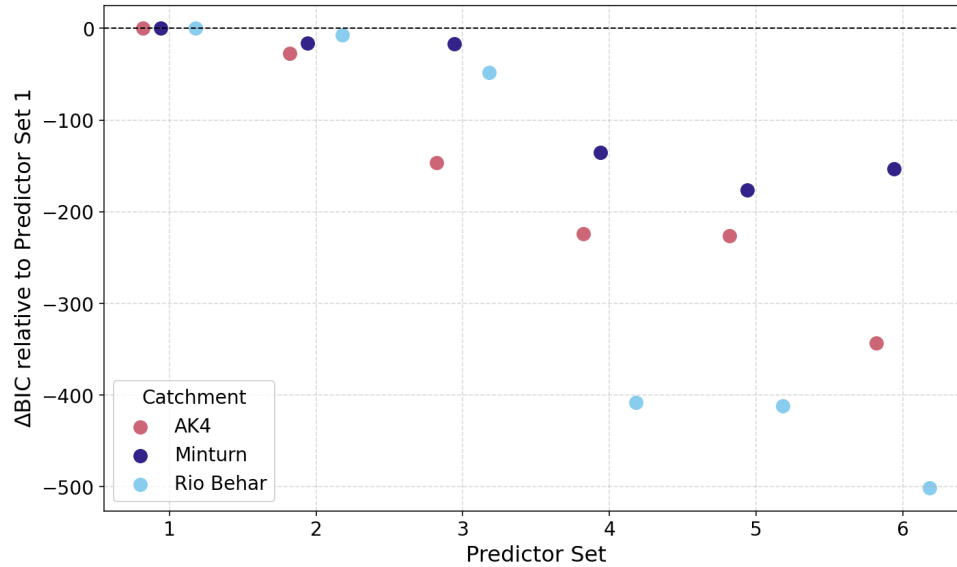
where Q is total annual meltwater runoff. We express the result as a percentage where low values indicate high accuracy and vice versa. To evaluate gap-filling performance, we randomly removed one week of daily runoff (two days for Rio Behar due to low number of in-situ observations) during June–August for each year for each catchment. We then predicted the missing values using both our TL models and linear interpolation (i.e. a more commonly used gap-filling method). We repeated this 1,000 times for each year for each catchment and compared methods using the ensemble mean normalized MSE (NMSE), which is MSE divided by the catchment's mean meltwater runoff.

3 Results

3.1 Optimal predictors

We find that the most effective predictor set for catchment-scale meltwater runoff is air temperature, ice surface temperature, albedo, and shortwave downward radiation (Table 1). For three of the three catchments (Rio Behar, AK4, and Minturn) this predictor set has the lowest BIC of the six predictor sets tested (Figure 3). This predictor set is therefore used for all the MAR emulators in the rest of the study.

MAR output	Predictor set
Air temperature	1
Air temperature & ice surface temperature	2
Air temperature & shortwave downward	3
Air temperature, ice surface temperature & albedo	4
Air temperature & albedo	5
Air temperature, ice surface temperature, shortwave downward & albedo	6

Table 1. Predictor sets created from different combinations of MAR output variables.**Figure****3.**

Change in ensemble Bayesian Information Criterion (BIC) for different predictor sets for each catchment, relative to the air temperature predictor set. Each ensemble contained 100 MLPs for each predictor set at each catchment, trained with 25 years of MAR outputs and predicting a random year. Variables in each predictor set can be found in Table 1.

3.2 Performance of MAR emulators relative to MAR

We find that our MAR emulators are able to accurately reproduce MAR meltwater runoff. The coefficients of determination (R^2) between MAR 2000–2024 daily meltwater runoff and MAR emulator-predicted 2000–2024 daily meltwater runoff range from 0.90 at the Minturn catchment to 0.96 at the AK4 catchment (Figure 4). The Root Mean Square Error (RMSE) between MAR daily meltwater runoff and MAR emulator-predicted daily meltwater runoff ranges from ± 0.81 mmWE at the Minturn catchment to ± 3.12 mmWE at the AK4 catchment (Figure 4). Additionally, the relationship between MAR daily meltwater runoff and MAR emulator-predicted daily meltwater runoff is statistically significant with $p < 0.001$ at all three catchments (Figure 4). These high R^2 values, low RMSE values, and statistically significant p-values demonstrate that our MAR emulators successfully learned the relationships between MAR outputs and MAR meltwater runoff. These values are also similar to two previous studies who found that machine learning models are able to simulate surface melt with high accuracy (Schlager et al., 2026; Tedesco et al., 2026).

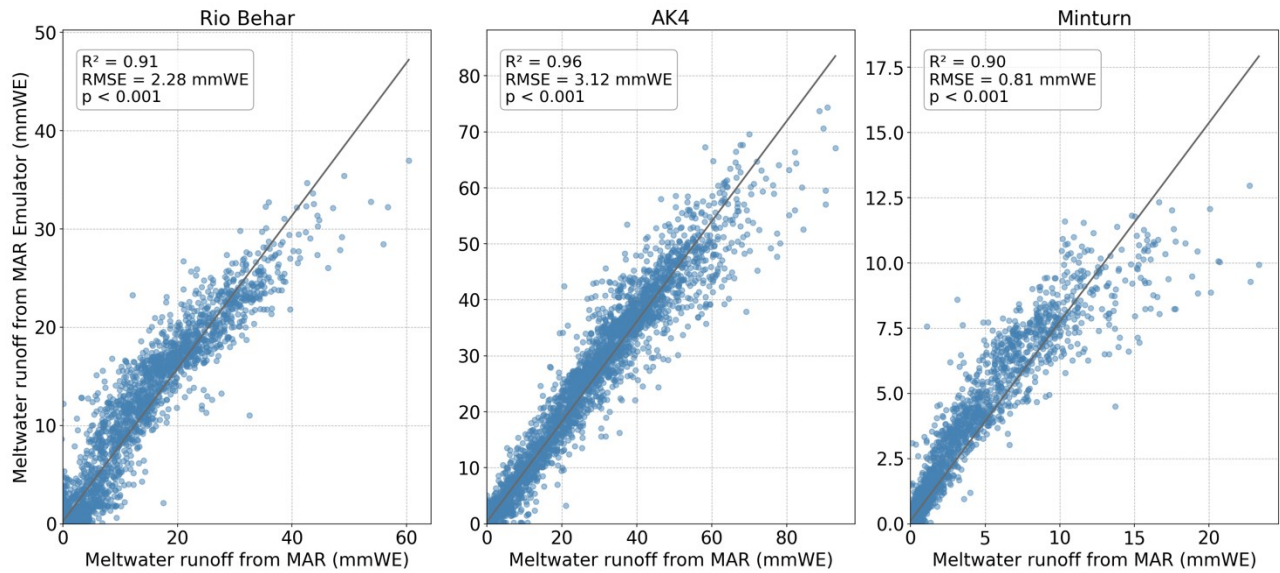


Figure 4. Comparison of daily meltwater runoff between MAR and our MAR emulators for the three study catchments (2000–2024).

3.3 Performance of TL models relative to MAR emulators

Relative to the MAR emulators, we find that the TL models improve predictions of cumulative meltwater runoff but that the magnitude of improvement varies across catchments. Across all three catchments, the average error rate of the MAR emulators was $\pm 19\%$, while the average error rate of the TL models was $\pm 12\%$ (Tables S4–S6). Conducting TL therefore yielded a 36.8% reduction in error rate relative to the MAR emulators (Figure 5). However, we find that improvement in error rate varies substantially between catchments. Conducting TL at the Rio Behar catchment decreased average model error rate from $\pm 20\%$ to $\pm 17\%$ for July 20–23, 2015 and July 5–13, 2016 (Figure 6, Table S4). At the AK4 catchment, conducting TL decreased average model error rate from $\pm 27\%$ to $\pm 11\%$ across all years predicted during the 2008–2016 period (Figure 7, Table S5). At the Minturn catchment conducting TL decreased average model error rate from $\pm 10\%$ to $\pm 9\%$ across all years predicted during the 2019–2022 period (Figure 8, Table S6). Therefore, overall, we find that our TL approach is able to provide more accurate predictions of meltwater runoff than emulation alone. Although we note that the error rate improvement varies substantially across catchments (Figure 5). Causes for this variation are explored in Section 4.

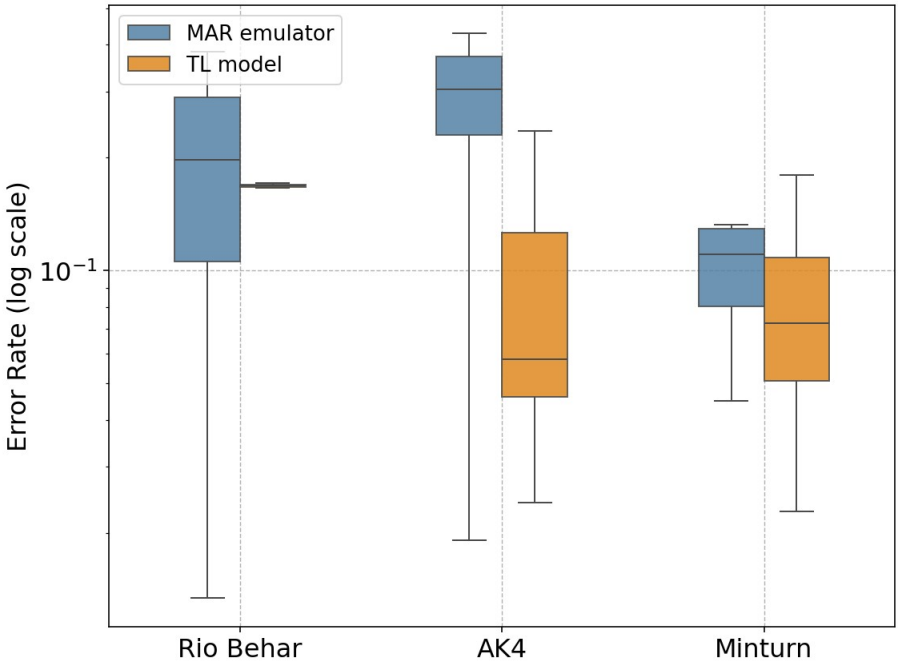


Figure 5. Catchment-scale MAR emulator and TL model error rate for all years predicted. Error rate is presented on a log scale to better visualize the effect of TL on error rate as error rate ranges multiple orders of magnitude between catchments.

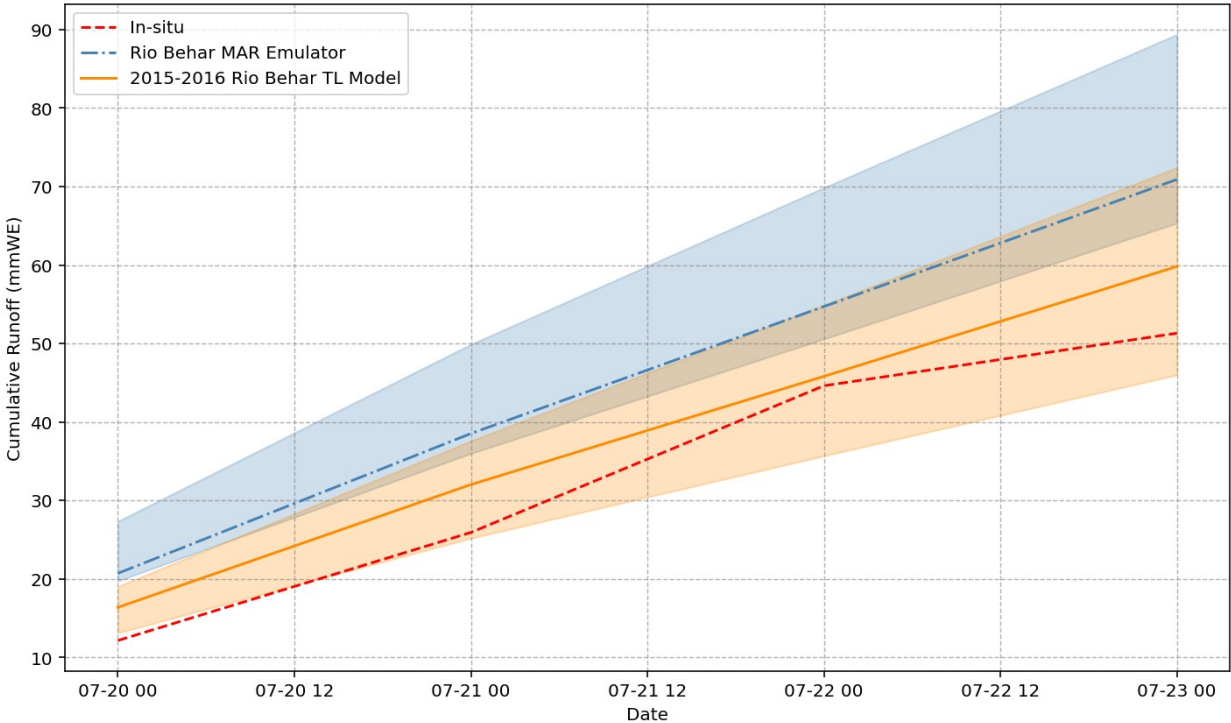


Figure 6. MAR emulator and TL model predictions of cumulative meltwater runoff compared with in-situ measurements at Rio Behar for July 20-23, 2015. Blue and orange lines with shaded bounds depict cumulative meltwater runoff predictions from the MAR emulator and TL model, respectively. The dashed red line depicts in-situ cumulative meltwater runoff. Shaded regions represent the 5–95% predictive interval from 100 Monte Carlo dropout realizations of the MAR emulator and TL model, indicating their respective predictive uncertainties.

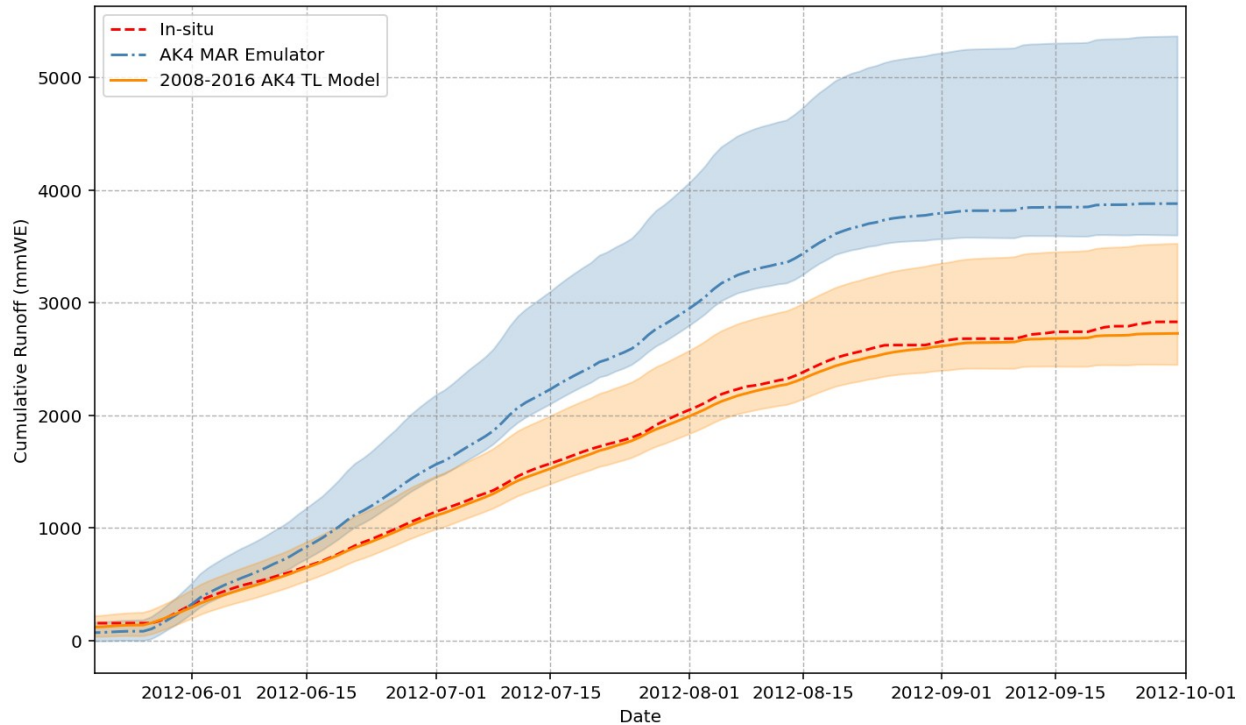


Figure 7. MAR emulator and TL model predictions of cumulative meltwater runoff compared with in-situ measurements at AK4 for 2012. Blue and orange lines with shaded bounds depict cumulative meltwater runoff predictions from the MAR emulator and TL model, respectively. The dashed red line depicts in-situ cumulative meltwater runoff. Shaded regions represent the 5–95% predictive interval from 100 Monte Carlo dropout realizations of the MAR emulator and TL model, indicating their respective predictive uncertainties.

360

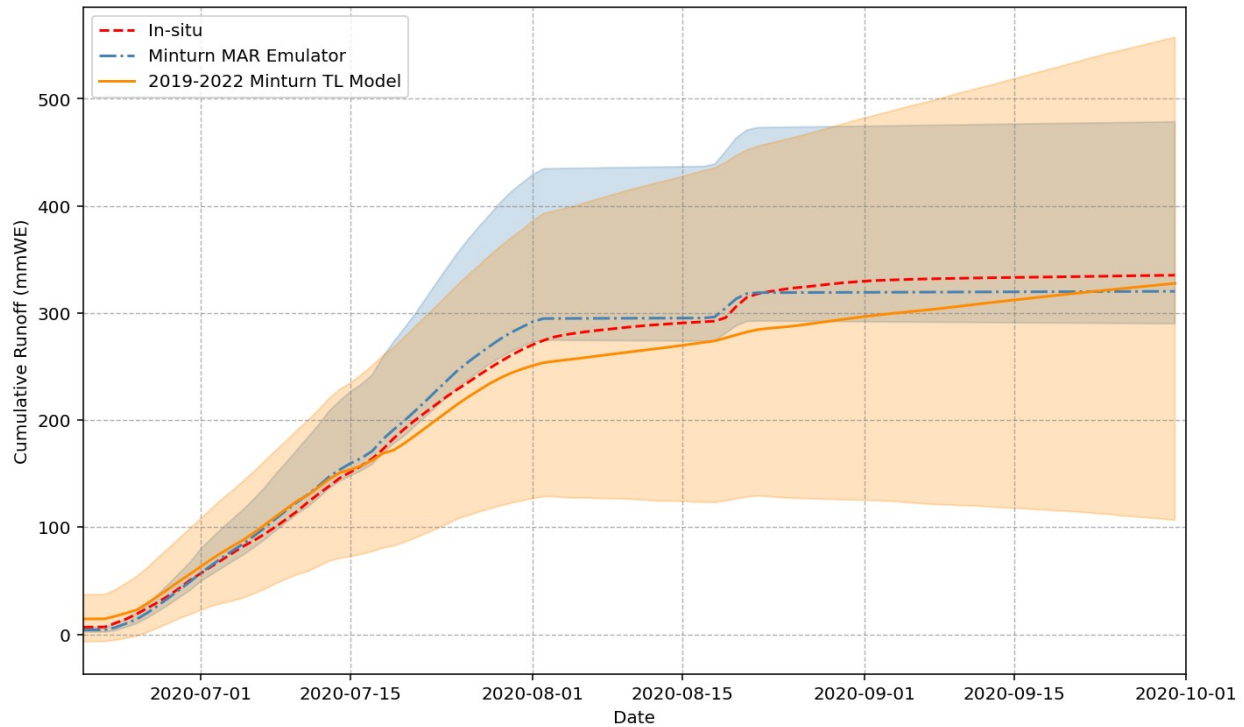


Figure 8. MAR emulator and TL model predictions of cumulative meltwater runoff compared with in-situ measurements at Minturn for 2020. Blue and orange lines with shaded bounds depict cumulative meltwater runoff predictions from the MAR emulator and TL model, respectively. The dashed red line depicts in-situ cumulative meltwater runoff. Shaded regions represent the 5–95% predictive interval from 100 Monte Carlo dropout realizations of the MAR emulator and TL model, indicating their respective predictive uncertainties.

3.4 Ability of TL models to generalize spatially

We find that our TL models have potential for accurately predicting cumulative meltwater runoff in unseen catchments, provided that the catchments are situated in similar regions. The TL models fine-tuned with in-situ data from catchments in southwest Greenland exhibit some generalizability, despite differences in catchment sizes (Figure 9). The Rio Behar TL model yields a mean error rate of $\pm 33\%$ when predicting AK4, compared to the AK4 TL model's mean error rate of $\pm 10\%$ (Figure 9). Likewise, the AK4 TL model yields a mean error rate of $\pm 26\%$ when predicting Rio Behar, compared to the Rio Behar TL model's mean error rate of $\pm 17\%$. Interestingly, the TL model for the northwestern catchment (Minturn) is able to generalize to the southwestern catchments with accuracy similar to that in the region in which they were trained (Figure 9). The Minturn TL model predicts Rio Behar and AK4 with error rates of $\pm 78\%$ and $\pm 76\%$, respectively (Figure 9). On the other hand, the TL models trained in southwestern catchments have a higher error rate when used to predict the northwestern catchment. The Rio

Behar TL model yields a mean error rate of $\pm 105\%$ and the AK4 TL model yields a mean error rate of $\pm 161\%$ when applied to Minturn (Figure 9). We find that it is possible that the TL models fine-tuned with in-situ data from southwestern catchments generalize well to the other southwestern catchment because they are in the same region. However, to determine if this is also true for northwestern catchments, we would need in-situ data from a catchment in the same region as Minturn.

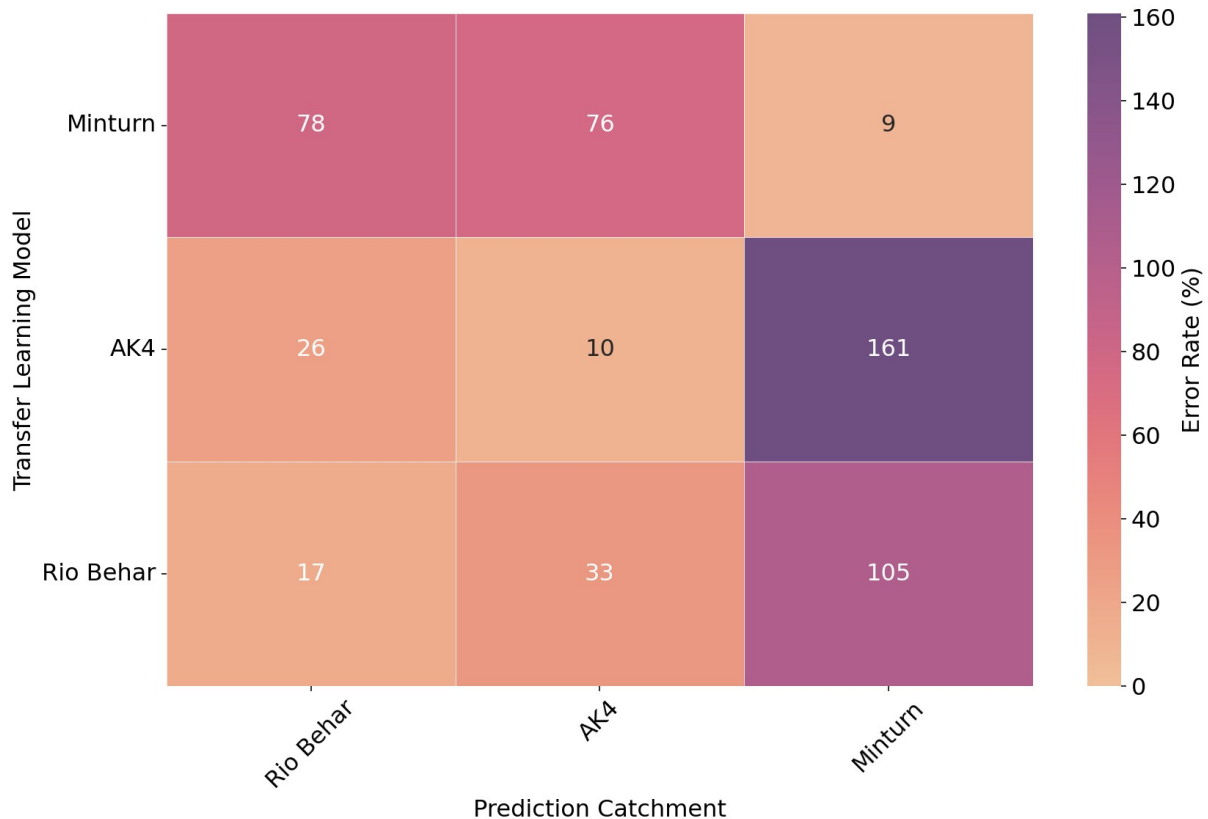


Figure 9. Generalization error of TL models across catchments. Values in each cell represent the mean error rate across all years at the prediction site.

3.5 Ability of TL models to generalize temporally

We find that our TL models are able to accurately predict meltwater runoff in years of unseen data, given that they are fine-tuned with sufficient amounts of in-situ data. The AK4 TL model, which was fine-tuned with nine years of in-situ data, predicts unseen melt years with a mean error rate of $\pm 17\%$ which is comparable to the AK4 TL model fine-tuned with all years of in-situ data ($\pm 14\%$; Figure 10a). This indicates that nine years provides sufficient amount of variation for our TL models to generalize across a range of melt conditions, including the 2012 extreme melt event (Nghiem et al., 2012) (Figure 7). The Minturn TL model, which is fine-tuned with four

years of in-situ data, also generalizes well, though the mean error rate for predicting unseen melt years ($\pm 4\%$) is lower than that of the Minturn TL model fine-tuned with all years of in-situ data ($\pm 6\%$; Figure 10b). On the other hand, we find that the TL model fine-tuned with only two years of in-situ data at Rio Behar generalizes less well to unseen data as its mean error rate is higher when predicting unseen years than at other catchments. The Rio Behar TL model predicts unseen melt years with a mean error rate of $\pm 23\%$, which is higher than the Rio Behar TL model fine-tuned with all years of $\pm 5\%$ (Figure 10c). Our findings therefore indicate that fine-tuning with a greater range of years increases the ability of our TL models to generalize temporally, producing lower and more comparable error rates between models fine-tuned with all years of in-situ data and those which were not fine-tuned with in-situ data from withheld years.

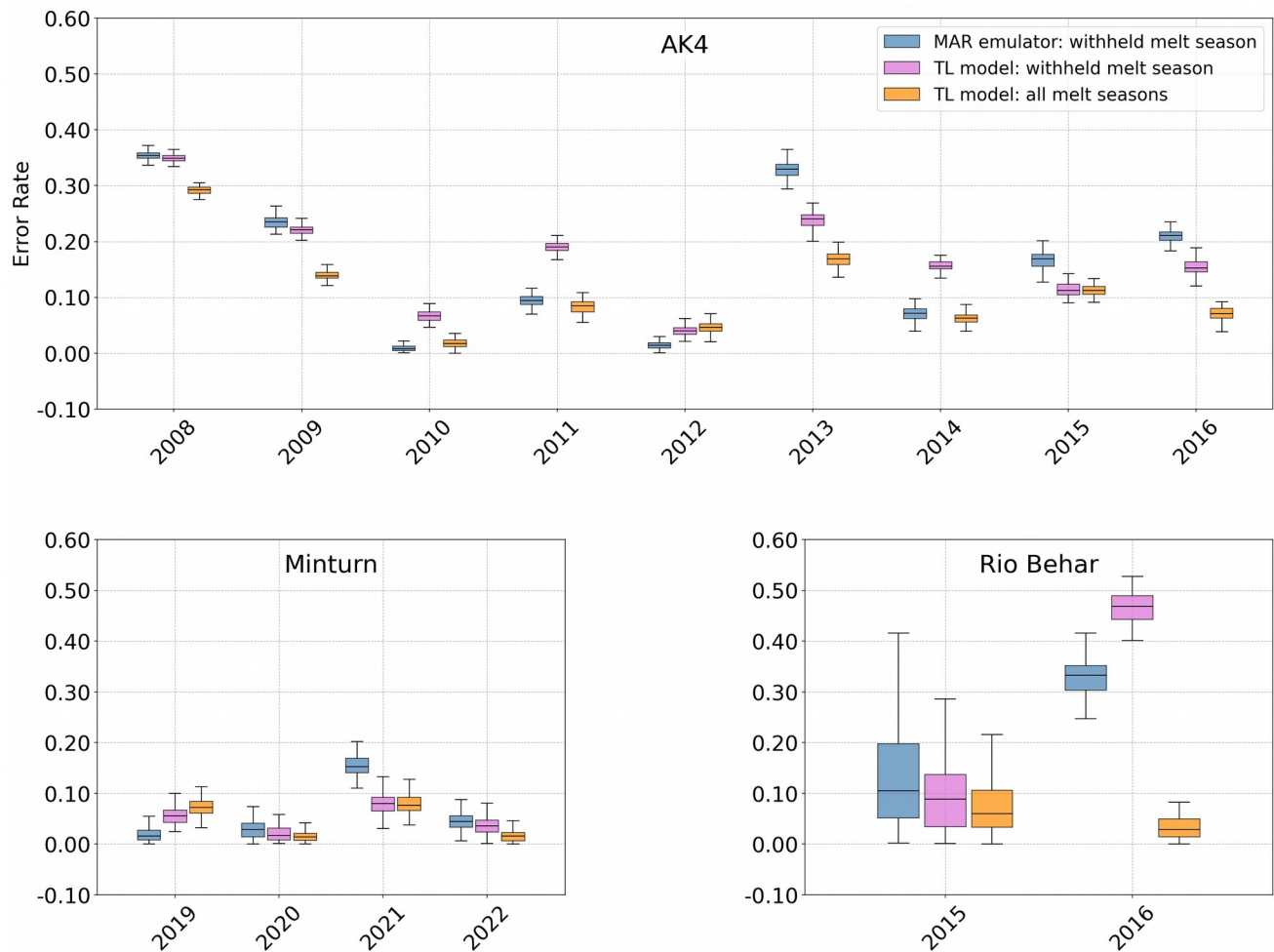


Figure 10. Error rate of MAR emulators pre-trained with in-situ data from all melt seasons except the withheld predicted year (blue), TL models fine-tuned with in-situ data from all years (orange) and TL models fine-tuned with in-situ data from all years except the withheld

predicted year (pink). 100 predictions were made by each model for each year at each catchment.

3.6 Filling temporal gaps in runoff time series with TL models

We find that our TL models have potential for gap-filling in-situ river discharge data at some catchments. For AK4, the TL model ensemble mean NMSE for filling random data gaps is 2.12 while the ensemble mean NMSE of linear interpolation is 7.45 (Table 2). Likewise, at Rio Behar, the TL model ensemble mean NMSE is 2.30 and the linear interpolation ensemble mean NMSE is 3.16 (Table 2). However, at Minturn, we find that linear interpolation is more effective at filling random gaps in data than our TL models. We speculate that the high performance of our TL model can be attributed to the larger amount of in-situ data for fine-tuning (nine years) which is more than double that used to fine-tune the TL models at Minturn, which performed worse. Larger amounts of training data, however, cannot explain the high performance of the Rio Behar TL model relative to linear interpolation because this model was only fine-tuned with 13 days of in-situ data. The better performance of the TL model may be because the data gaps for this catchment were only two days. Our findings indicate that at least nine melt seasons of in-situ data may be needed to fine-tune TL models to accurately fill gaps in daily meltwater runoff since our TL models fine-tuned with less than nine melt seasons of in-situ data were unable to outperform linear interpolation.

Catchment	Linear interpolation ensemble mean NMSE	TL model ensemble mean NMSE
Rio Behar	3.16	2.30
AK4	7.45	2.12
Minturn	0.82	0.99

Table 2. Ensemble mean NMSE values for TL models and linear interpolation filling random gaps in daily meltwater runoff time series. An ensemble of 1,000 random gaps was run for each melt year at each catchment. Gaps were week-long for AK4 and Minturn, and 2 days long at Rio Behar, due to less data.

4 Discussion

4.1 Improved meltwater runoff prediction using transfer learning

We developed a deep transfer learning approach for predicting meltwater runoff at three catchments on the Greenland Ice Sheet. We find that, by fine-tuning RCM emulators using in-situ measurements of meltwater runoff, we can simulate total annual meltwater runoff with a mean error rate of $\pm 15\%$ across three catchments in Southwest and Northwest Greenland. Our transfer learning approach, which incorporates training data from both RCMs and in-situ measurements, therefore offers a viable approach for accurately simulating meltwater runoff

from the Greenland Ice Sheet that advances previous studies which have relied on RCM outputs alone. We also demonstrate, using several experiments, that our TL models have the potential to generalize spatially and are able to generalize temporally. When predicting unseen catchments, our TL models fine-tuned with in-situ data from southwestern catchments generalized spatially to the other southwestern catchment (Figure 9). With in-situ data from more catchments, we could determine if our TL models have the highest accuracy when they generalize spatially at the regional scale. Our TL models can generalize temporally to predict unseen years of meltwater with lower error rates than those of our MAR emulators (Figure 10). Our findings therefore indicate that transfer learning is an effective approach for predicting meltwater runoff in nearby ungauged catchments or extending time-series of meltwater runoff in catchments that are already gauged.

Our analysis advances the use of machine learning for predicting surface melt and runoff from the Greenland Ice Sheet. The RCM emulators developed by our study differ from two other recently developed emulators (Schlager et al., 2026; Tedesco et al., 2026) in two primary ways. First, our emulators predict meltwater runoff, as opposed to surface melt. We predict runoff because it is directly measured by in-situ river gauges that we used to fine-tune our emulators, and because runoff directly contributes to global sea level. In some ways, this made our task more complex because meltwater runoff not only depends on the amount of energy absorbed at the ice sheet surface but also on several other processes such as retention and refreezing, which determine how much meltwater flows into supraglacial rivers. Uncertainties in modeled meltwater runoff are therefore likely to be higher than modeled surface melt. Having said that, our emulators have comparable accuracy ($R^2 = 0.88\text{--}0.96$) to the surface melt emulators developed by Schlager et al. (2026) and Tedesco et al. (2026). Transfer learning could be applied to surface melt emulators, but it would require fine-tuning with a different source of in-situ measurements such as surface height measurements from ablation stakes or sonic ranging sensors mounted on automated weather stations (Chandler et al., 2021; Fausto et al., 2021; Hermann et al., 2018). Similar transfer learning methods could then be applied to improve emulators of surface melt developed by previous studies.

The second way our study differs from two previous studies is that we predict at the catchment-scale rather than major drainage basins or across elevation bands (Schlager et al., 2026; Tedesco et al., 2026). Since we fine-tune with in-situ river gauges, our approach is limited to the catchment-scale. An advantage of this scale is that our predictions could be directly incorporated into existing freshwater flux datasets which quantify Greenland Ice Sheet meltwater runoff into the ocean at specific outlets (e.g. Igneczi & Bamber, 2025; Mankoff et al., 2020). Our approach could therefore be used to investigate how much meltwater runoff is flowing into the marine environment and how this is changing in different parts of the ice sheet,

as well as to extend or fill data gaps in existing streamflow records. The disadvantage is that our approach cannot be used to predict meltwater runoff over large areas or in a gridded format. This makes our meltwater runoff predictions somewhat limited since we cannot investigate elevational changes in meltwater runoff or incorporate meltwater runoff predictions in sheet models. Further research will be required to develop a transfer learning approach that can produce meltwater runoff predictions as a gridded product.

Our findings demonstrate the need for records of meltwater runoff from many catchments across the Greenland Ice Sheet. We find that the magnitude of improvement depends on the amount of in-situ data. In our evaluation we found that conducting TL with nine years of in-situ data yielded greater improvement in model accuracy than conducting TL with in-situ data from only two or four years. Leveraging longer records of in-situ data captures wider variation in volumes of meltwater runoff and provides more context for TL models to be able to generalize temporally and predict meltwater runoff from unseen years. This finding implies that longer in-situ meltwater runoff records are necessary to properly fine-tune and validate catchment-scale TL models. Likewise, we also find that spatially extensive in-situ data are needed to ensure cross-catchment generalizability of TL models. To generalize across catchments and predict catchment-scale meltwater runoff across the Greenland Ice Sheet, a diverse set of in-situ meltwater runoff records are needed from variable catchments to fine-tune TL models.

4.2 Emulating meltwater runoff on the Greenland Ice Sheet

We found that certain catchments were easier to emulate than others. For example, we find that our MAR emulator performed better for the Rio Behar and AK4 catchments than for Minturn. We speculate that the Rio Behar and AK4 are easier to predict ($R^2 = 0.91$ and 0.96 , respectively) because their catchments are relatively small (65 km^2 and 34 km^2 , respectively) and predominantly consist of glacier ice. Most surface melt is therefore efficiently routed through the catchment with lag times much less than a day (e.g. 5.5 hrs for Rio Behar and 5 hrs for AK4) (Smith et al., 2017, 2021). Daily meltwater runoff can therefore be accurately predicted using daily atmospheric variables along with surface albedo and ice surface temperature. We speculate that the slightly worse performance of our MAR emulator in Minturn ($R^2 = 0.90$) is because this catchment contains substantial amounts of firn and snow. Meltwater generated in these catchments may therefore be more likely to infiltrate into pore spaces than in the predominantly glacier ice catchments of Rio Behar and AK4 (Harper et al., 2012). Meltwater runoff in these catchments may therefore depend not only on instantaneous surface energy balance, but also firn storage capacity. To account for this, future studies should consider including features that represent antecedent conditions such firn air content in their predictor sets. Alternatively, studies could explore the inclusion of lagged variables (e.g. subsurface

storage) that account for delays between peak melt and runoff caused by infiltration in firn or snow.

We find that the best performing emulators for predicting meltwater runoff are those trained with predictor sets containing albedo. In all three study catchments, we found that the inclusion of albedo as a predictor greatly increased model accuracy (Fig. 3). This is consistent with previous studies which found that inclusion of albedo is critical for RCM emulators to predict meltwater production and surface melt (Schlager et al., 2026; Tedesco et al., 2026). Since Greenland is generally cold, shortwave energy absorption is a major source of energy to the ice sheet surface in the ablation zone, which could explain why albedo is so important for surface melting and runoff. Additionally, albedo may encode information about surface conditions as glacier ice has a lower capacity to store meltwater than firn and snow, leading to higher meltwater runoff in bare-ice catchments (Ryan et al., 2019). We also find that ice surface temperature is an important variable for predicting meltwater runoff since snow and ice cannot melt until they reach 0°C. Advantageously, Greenland Ice Sheet albedo and ice surface temperature observations are readily available from satellite remote sensing (Fang et al., 2023; Hall et al., 2008; Karagali et al., 2022; Ryan et al., 2026; Zhang et al., 2026). As such, future studies may be able to predict meltwater runoff with even higher accuracy by incorporating satellite observations of albedo and ice surface temperature.

4.3 Applications of TL in hydrology

While our application of TL is novel for the Greenland Ice Sheet, this approach is increasingly used in terrestrial hydrology where studies have found that TL models consistently outperform deep learning models which were trained solely on sparse data from target basins. In terrestrial hydrology, a primary goal is to pre-train deep learning models with in-situ hydrological records from well-monitored regions, then conduct TL using sparse records to improve streamflow predictions in poorly-monitored regions (Echternacht et al., 2025; Ma et al., 2021; Ougahi & Rowan, 2026). This is similar to our approach, which pre-trains Greenland Ice Sheet runoff models with dense records (MAR meltwater runoff) then fine-tunes them with sparse records (in-situ meltwater runoff). However, terrestrial hydrology TL studies have the advantage of being able to leverage in-situ datasets during both the pre-training and fine-tuning stages as there are hundreds if not thousands of gauged catchments in the continental Europe and USA (Coxon et al., 2020; Höge et al., 2023; Loritz et al., 2024; Newman et al., 2022). In Greenland, few catchments are gauged (Chen et al., 2025; Mankoff et al., 2020), so we pre-trained our TL models with RCM outputs to make up for this scarcity of in-situ runoff records.

5 Conclusions

In this study, we developed catchment-scale MAR emulators and transfer learning models that we trained and tested with in-situ runoff records from three catchments in southwest and northwest Greenland. We also tested the ability of our TL models to generalize spatially and temporally and fill data gaps in in-situ records. Our TL models are more accurate than our MAR emulators because they remove inherent biases present in the MAR emulators by fine-tuning deep learning models with in-situ observations of meltwater runoff. Furthermore, our TL models show potential for generalizing spatially to other catchments in the same region as the catchment used in the fine-tuning stage. When generalizing temporally to unseen melt seasons, our TL models exhibit higher accuracy than our MAR emulators. Our TL models also generalize temporally with higher accuracy when they are fine-tuned with more years of in-situ meltwater runoff data. At catchments where meltwater runoff has been measured in-situ for many years, our TL models are better at gap-filling than conventional methods. As such, our TL models have implications for predicting meltwater runoff at ungauged catchments, extending temporal records of meltwater runoff data in gauged catchments, and filling gaps in meltwater runoff data in gauged catchments with missing data.

Acknowledgments

This research was supported by NASA Cryospheric Science award #80NSSC25K7364 managed by Dr. Thorsten Markus. In-situ runoff measurements at Rio Behar and Minturn River were supported previously by NASA Cryospheric Science awards NNX14AH93G and 80NSSC19K0942, respectively. We thank Matthew Cooper for providing the Rio Behar catchment delineation.

Open Research

The code and data are currently available at: <https://doi.org/10.5281/zenodo.21649502>

Conflict of Interest Disclosure

The authors declare there are no conflicts of interest for this manuscript.

References

- Bahbah, R., Kjeldsen, K. K., & van As, D. (2025). *Watson River Discharge* (Version 6) [Dataset]. GEUS Dataverse. <https://doi.org/10.22008/FK2/XEHYCM>
- Bolón-Canedo, V., & Alonso-Betanzos, A. (2019). Ensembles for feature selection: A review and future trends. *Information Fusion*, 52, 1–12. <https://doi.org/10.1016/j.inffus.2018.11.008>
- Bozinovski, S., & Fulgosi, A. (1976). The influence of pattern similarity and transfer learning upon training of a base perceptron b2. *Proceedings of Symposium Informatica*, 3, 121. <https://cir.nii.ac.jp/crid/1370306904409027332>
- Chandler, D. M., Wadham, J., Nienow, P., Doyle, S. H., Tedstone, A., Telling, J., Hawkings, J., Alcock, J., Linhoff, B., & Hubbard, A. L. (2021). *Ice ablation record from the Greenland Ice Sheet measured in spring/summer 2012* [Dataset]. PANGAEA. <https://doi.org/10.1594/PANGAEA.926842>

- Chen, D., Yang, K., Man, M., Huang, C., Wang, Y., Yi, X., & Zhu, Y. (2025). Monitoring northern Greenland proglacial river discharge from space. *Remote Sensing of Environment*, 317, 114529. <https://doi.org/10.1016/j.rse.2024.114529>
- Christensen, O. B., Drews, M., Christensen, J. H., Dethloff, K., Ketelsen, K., Hebestadt, I., & Rinke, A. (2007). Technical report 06-17 The HIRHAM Regional Climate Model Version 5 (β). *Technical Report*.
- Citterio, M., & Ahlstrøm, A. P. (2025). *PROMICE-1980s Ice Mask* [Dataset]. GEUS Dataverse. <https://doi.org/10.22008/FK2/PRWITW>
- Cook, S. J., Christoffersen, P., & Todd, J. (2022). A fully-coupled 3D model of a large Greenlandic outlet glacier with evolving subglacial hydrology, frontal plume melting and calving. *Journal of Glaciology*, 68(269), 486–502. <https://doi.org/10.1017/jog.2021.109>
- Cooper, M. G., Smith, L. C., Rennermalm, Å. K., Ryan, J. C., Pitcher, L. H., Liston, G. E., Miège, C., Cooley, S. W., & van As, D. (2025). Greenland ice sheet runoff reduced by meltwater refreezing in bare ice. *Nature Communications*, 16(1), 8273. <https://doi.org/10.1038/s41467-025-62281-0>
- Coxon, G., Addor, N., Bloomfield, J. P., Freer, J., Fry, M., Hannaford, J., Howden, N. J. K., Lane, R., Lewis, M., Robinson, E. L., Wagener, T., & Woods, R. (2020). CAMELS-GB: Hydrometeorological time series and landscape attributes for 671 catchments in Great Britain. *Earth System Science Data*, 12(4), 2459–2483. <https://doi.org/10.5194/essd-12-2459-2020>
- Echternacht, H., Campos, L., de Martinho, A. D., Souza, D. P. M. de, de Santis, R. B., Gontijo, T. S., Bodini, M., Gorgoglione, A., Saporetti, C. M., & Goliatt, L. (2025). Convolutional neural networks with transfer learning for natural river flow prediction in ungauged basins. *Scientific Reports*, 15(1), 23873. <https://doi.org/10.1038/s41598-025-07088-1>
- Esenther, S. E., Lee, M. N., Esenther, B. G., & Smith, L. C. (2026). Automated Mapping of Supraglacial Stream Networks on the Greenland Ice Sheet Using Dual U-Net Convolutional Neural Networks. *IEEE Geoscience and Remote Sensing Letters*, 23, 1–5. <https://doi.org/10.1109/LGRS.2026.3663855>
- Esenther, S. E., Smith, L. C., LeWinter, A., Pitcher, L. H., Overstreet, B. T., Kehl, A., Onclin, C., Goldstein, S., & Ryan, J. C. (2023). New proglacial meteorology and river stage observations from Inglefield Land and Pituffik, NW Greenland. *Geoscientific Instrumentation, Methods and Data Systems*, 12(2), 215–230. <https://doi.org/10.5194/gi-12-215-2023>
- Eyring, V., Gentine, P., Camps-Valls, G., Lawrence, D. M., & Reichstein, M. (2024). AI-empowered next-generation multiscale climate modelling for mitigation and adaptation. *Nature Geoscience*, 17(10), 963–971. <https://doi.org/10.1038/s41561-024-01527-w>
- Fang, Z., Wang, N., Wu, Y., & Zhang, Y. (2023). Greenland-Ice-Sheet Surface Temperature and Melt Extent from 2000 to 2020 and Implications for Mass Balance. *Remote Sensing*, 15(4), 1149. <https://doi.org/10.3390/rs15041149>
- Fausto, R. S., van As, D., Mankoff, K. D., Vandecrux, B., Citterio, M., Ahlstrøm, A. P., Andersen, S. B., Colgan, W., Karlsson, N. B., Kjeldsen, K. K., Korsgaard, N. J., Larsen, S. H., Nielsen, S., Pedersen, A. Ø., Shields, C. L., Solgaard, A. M., & Box, J. E. (2021). Programme for Monitoring of the Greenland Ice Sheet (PROMICE) automatic weather station data. *Earth System Science Data*, 13(8), 3819–3845. <https://doi.org/10.5194/essd-13-3819-2021>
- Fettweis, X., Box, J. E., Agosta, C., Amory, C., Kittel, C., Lang, C., van As, D., Machguth, H., & Gallée, H. (2017). Reconstructions of the 1900–2015 Greenland ice sheet surface mass balance using the regional climate MAR model. *The Cryosphere*, 11(2), 1015–1033. <https://doi.org/10.5194/tc-11-1015-2017>
- Fettweis, X., Hofer, S., Krebs-Kanzow, U., Amory, C., Aoki, T., Berends, C. J., Born, A., Box, J. E., Delhasse, A., Fujita, K., Gierz, P., Goelzer, H., Hanna, E., Hashimoto, A., Huybrechts, P., Kapsch, M.-L., King, M. D., Kittel, C., Lang, C., ... Zolles, T. (2020). GrSMBMIP: Intercomparison of the modelled 1980–2012 surface mass balance over the Greenland Ice Sheet. *The Cryosphere*, 14(11), 3935–3958. <https://doi.org/10.5194/tc-14-3935-2020>
- Gerrish, L. (2020). *The coastline of Kalaallit Nunaat/ Greenland available as a shapefile and geopackage, covering the main land and islands, with glacier fronts updated as of 2017*. (Version 1.0, p. 2 files, 5.26 MB) [Application/octet-stream,application/zip]. UK Polar Data Centre, Natural Environment Research Council, UK Research & Innovation. <https://doi.org/10.5285/8CECDE06-8474-4B58-A9CB-B820FA4C9429>
- Glaude, Q., Noel, B., Olesen, M., Van den Broeke, M., van de Berg, W. J., Mottram, R., Hansen, N., Delhasse, A., Amory, C., Kittel, C., Goelzer, H., & Fettweis, X. (2024). A Factor Two Difference in 21st-Century Greenland Ice Sheet Surface Mass Balance Projections From Three Regional Climate Models Under a Strong Warming Scenario (SSP5-8.5). *Geophysical Research Letters*, 51(22), e2024GL111902. <https://doi.org/10.1029/2024GL111902>
- Gleason, C. J., Yang, K., Feng, D., Smith, L. C., Liu, K., Pitcher, L. H., Chu, V. W., Cooper, M. G., Overstreet, B. T., Rennermalm, A. K., & Ryan, J. C. (2021). Hourly surface meltwater routing for a Greenlandic

- supraglacial catchment across hillslopes and through a dense topological channel network. *The Cryosphere*, 15(5), 2315–2331. <https://doi.org/10.5194/tc-15-2315-2021>
- Goldstein, S. N., Ryan, J. C., How, P. R., Esenther, S. E., Pitcher, L. H., LeWinter, A. L., Overstreet, B. T., Kyzivat, E. D., Fayne, J. V., & Smith, L. C. (2023). Proglacial river stage derived from georectified time-lapse camera images, Inglefield Land, Northwest Greenland. *Frontiers in Earth Science*, 11. <https://doi.org/10.3389/feart.2023.960363>
- Hall, D. K., Box, J. E., Casey, K. A., Hook, S. J., Shuman, C. A., & Steffen, K. (2008). Comparison of satellite-derived and in-situ observations of ice and snow surface temperatures over Greenland. *Remote Sensing of Environment*, 112(10), 3739–3749. <https://doi.org/10.1016/j.rse.2008.05.007>
- Han, X., Zhang, Z., Ding, N., Gu, Y., Liu, X., Huo, Y., Qiu, J., Yao, Y., Zhang, A., Zhang, L., Han, W., Huang, M., Jin, Q., Lan, Y., Liu, Y., Liu, Z., Lu, Z., Qiu, X., Song, R., ... Zhu, J. (2021). Pre-trained models: Past, present and future. *AI Open*, 2, 225–250. <https://doi.org/10.1016/j.aiopen.2021.08.002>
- Harper, J., Humphrey, N., Pfeffer, W. T., Brown, J., & Fettweis, X. (2012). Greenland ice-sheet contribution to sea-level rise buffered by meltwater storage in firn. *Nature*, 491(7423), 240–243. <https://doi.org/10.1038/nature11566>
- Hermann, M., Box, J. E., Fausto, R. S., Colgan, W. T., Langen, P. L., Mottram, R., Wuite, J., Noël, B., van den Broeke, M. R., & van As, D. (2018). Application of PROMICE Q-Transect in Situ Accumulation and Ablation Measurements (2000–2017) to Constrain Mass Balance at the Southern Tip of the Greenland Ice Sheet. *Journal of Geophysical Research: Earth Surface*, 123(6), 1235–1256. <https://doi.org/10.1029/2017JF004408>
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., ... Thépaut, J.-N. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730), 1999–2049. <https://doi.org/10.1002/qj.3803>
- Hill, T., & Dow, C. F. (2023). The impact of surface melt rate and catchment characteristics on Greenland Ice Sheet moulin inputs. *The Cryosphere*, 17(7), 2607–2624. <https://doi.org/10.5194/tc-17-2607-2023>
- Höge, M., Kauzlaric, M., Siber, R., Schönenberger, U., Horton, P., Schwanbeck, J., Floriancic, M. G., Viviroli, D., Wilhelm, S., Sikorska-Senoner, A. E., Addor, N., Brunner, M., Pool, S., Zappa, M., & Fenicia, F. (2023). CAMELS-CH: Hydro-meteorological time series and landscape attributes for 331 catchments in hydrologic Switzerland. *Earth System Science Data*, 15(12), 5755–5784. <https://doi.org/10.5194/essd-15-5755-2023>
- Hosna, A., Merry, E., Gyalmo, J., Alom, Z., Aung, Z., & Azim, M. A. (2022). Transfer learning: A friendly introduction. *Journal of Big Data*, 9(1), 102. <https://doi.org/10.1186/s40537-022-00652-w>
- Igneczi, A., & Bamber, J. L. (2025). A high-resolution pan-Arctic meltwater discharge dataset from 1950 to 2021. *Earth System Science Data*, 17(7), 3203–3218. <https://doi.org/10.5194/essd-17-3203-2025>
- Karagali, I., Barfod Suhr, M., Mottram, R., Nielsen-Englyst, P., Dybkjær, G., Ghent, D., & Høyer, J. L. (2022). A new Level 4 multi-sensor ice surface temperature product for the Greenland Ice Sheet. *The Cryosphere*, 16(9), 3703–3721. <https://doi.org/10.5194/tc-16-3703-2022>
- Karlsson, N. B., Solgaard, A. M., Mankoff, K. D., Gillet-Chaulet, F., MacGregor, J. A., Box, J. E., Citterio, M., Colgan, W. T., Larsen, S. H., Kjeldsen, K. K., Korsgaard, N. J., Benn, D. I., Hewitt, I. J., & Fausto, R. S. (2021). A first constraint on basal melt-water production of the Greenland ice sheet. *Nature Communications*, 12(1), 3461. <https://doi.org/10.1038/s41467-021-23739-z>
- Kingma, D. P., & Ba, J. (2017). *Adam: A Method for Stochastic Optimization* (arXiv:1412.6980). arXiv. <https://doi.org/10.48550/arXiv.1412.6980>
- Kondo, K. (2020). *Discharge measurement at the outlet stream of Qaanaaq Glacier in the summer 2017–2019* (Version 2) [Dataset]. GEUS Dataverse. https://doi.org/10.22008/hokkaido/data/meltwater_discharge/qaanaaq
- Lefebvre, F., Fettweis, X., Gallée, H., Van Ypersele, J.-P., Marbaix, P., Greuell, W., & Calanca, P. (2005). Evaluation of a high-resolution regional climate simulation over Greenland. *Climate Dynamics*, 25(1), 99–116. <https://doi.org/10.1007/s00382-005-0005-8>
- Lim, Y.-K., Schubert, S. D., Nowicki, S. M. J., Lee, J. N., Molod, A. M., Cullather, R. I., Zhao, B., & Velicogna, I. (2016). Atmospheric summer teleconnections and Greenland Ice Sheet surface mass variations: Insights from MERRA-2. *Environmental Research Letters*, 11(2), 024002. <https://doi.org/10.1088/1748-9326/11/2/024002>
- Loritz, R., Dolich, A., Acuña Espinoza, E., Ebeling, P., Guse, B., Götte, J., Hassler, S. K., Hauffe, C., Heidbüchel, I., Kiesel, J., Mälicke, M., Müller-Thomy, H., Stölzle, M., & Tarasova, L. (2024). CAMELS-DE: Hydro-meteorological time series and attributes for 1582 catchments in Germany. *Earth System Science Data*, 16(12), 5625–5642. <https://doi.org/10.5194/essd-16-5625-2024>

- Ma, K., Feng, D., Lawson, K., Tsai, W.-P., Liang, C., Huang, X., Sharma, A., & Shen, C. (2021). Transferring Hydrologic Data Across Continents – Leveraging Data-Rich Regions to Improve Hydrologic Prediction in Data-Sparse Regions. *Water Resources Research*, 57(5), e2020WR028600. <https://doi.org/10.1029/2020WR028600>
- Machguth, H., MacFerrin, M., van As, D., Box, J. E., Charalampidis, C., Colgan, W., Fausto, R. S., Meijer, H. A. J., Mosley-Thompson, E., & van de Wal, R. S. W. (2016). Greenland meltwater storage in firn limited by near-surface ice formation. *Nature Climate Change*, 6(4), 390–393. <https://doi.org/10.1038/nclimate2899>
- Mankoff, K. (2025). *Streams, Outlets, Basins, and Discharge [k=1.0]* (Version 8) [Dataset]. GEUS Dataverse. <https://doi.org/10.22008/FK2/XKQVL7>
- Mankoff, K. D., Noël, B., Fettweis, X., Ahlström, A. P., Colgan, W., Kondo, K., Langley, K., Sugiyama, S., van As, D., & Fausto, R. S. (2020). Greenland liquid water discharge from 1958 through 2019. *Earth System Science Data*, 12(4), 2811–2841. <https://doi.org/10.5194/essd-12-2811-2020>
- Mankoff, K. D., & Tulaczyk, S. M. (2017). The past, present, and future viscous heat dissipation available for Greenland subglacial conduit formation. *The Cryosphere*, 11(1), 303–317. <https://doi.org/10.5194/tc-11-303-2017>
- Martinez-Ramon, M., Ajith, M., & Kurup, A. R. (2024). The Multilayer Perceptron. In *Deep Learning: A Practical Introduction* (pp. 1–39). Wiley. <https://doi.org/10.1002/9781119861898.ch1>
- Neath, A. A., & Cavanaugh, J. E. (2012). The Bayesian information criterion: Background, derivation, and applications. *WIREs Computational Statistics*, 4(2), 199–203. <https://doi.org/10.1002/wics.199>
- Newman, A. J., Sampson, K., Clark, M., Bock, A., Viger, R., Blodgett, D., Addor, N., & Mizukami, M. (2022). *CAMELS: Catchment Attributes and Meteorology for Large-sample Studies* [Dataset]. Zenodo. <https://doi.org/10.5065/D6MW2F4D>
- Nghiem, S. V., Hall, D. K., Mote, T. L., Tedesco, M., Albert, M. R., Keegan, K., Shuman, C. A., DiGirolamo, N. E., & Neumann, G. (2012). The extreme melt across the Greenland ice sheet in 2012. *Geophysical Research Letters*, 39(20). <https://doi.org/10.1029/2012GL053611>
- Noël, B., van Kampenhout, L., Lenaerts, J. T. M., van de Berg, W. J., & van den Broeke, M. R. (2021). A 21st Century Warming Threshold for Sustained Greenland Ice Sheet Mass Loss. *Geophysical Research Letters*, 48(5), e2020GL090471. <https://doi.org/10.1029/2020GL090471>
- Otosaka, I. N., Horwath, M., Mottram, R., & Nowicki, S. (2023). Mass Balances of the Antarctic and Greenland Ice Sheets Monitored from Space. *Surveys in Geophysics*, 44(5), 1615–1652. <https://doi.org/10.1007/s10712-023-09795-8>
- Ougahi, J. H., & Rowan, J. S. (2026). Investigating Deep Learning Knowledge Transfer in Streamflow Prediction From Global to Local Catchment. *Water Resources Research*, 62(2), e2025WR041194. <https://doi.org/10.1029/2025WR041194>
- Rae, J. G. L., Aðalgeirsdóttir, G., Edwards, T. L., Fettweis, X., Gregory, J. M., Hewitt, H. T., Lowe, J. A., Lucas-Picher, P., Mottram, R. H., Payne, A. J., Ridley, J. K., Shannon, S. R., van de Berg, W. J., van de Wal, R. S. W., & van den Broeke, M. R. (2012). Greenland ice sheet surface mass balance: Evaluating simulations and making projections with regional climate models. *The Cryosphere*, 6(6), 1275–1294. <https://doi.org/10.5194/tc-6-1275-2012>
- Rahmstorf, S., Box, J. E., Feulner, G., Mann, M. E., Robinson, A., Rutherford, S., & Schaffernicht, E. J. (2015). Exceptional twentieth-century slowdown in Atlantic Ocean overturning circulation. *Nature Climate Change*, 5(5), 475–480. <https://doi.org/10.1038/nclimate2554>
- Rennermalm, A. K., Smith, L. C., Chu, V. W., Forster, R. R., Box, J. E., & Hagedorn, B. (2012). Proglacial river stage, discharge, and temperature datasets from the Akuliarusiarsuup Kuua River northern tributary, Southwest Greenland, 2008–2011. *Earth System Science Data*, 4(1), 1–12. <https://doi.org/10.5194/essd-4-1-2012>
- Rennermalm, A. K., Smith, L. C., Hammann, A. C., Leidman, S. Z., Cooper, M. G., Cooley, S. W., & Pitcher, L. (2017). River discharge at station AK-004-001, 2008–2016, version 3.0 [Dataset]. In *The State University of New Jersey*. PANGAEA. <https://doi.org/10.1594/PANGAEA.876357>
- Ridder, K. D., & Gallée, H. (1998). *Land Surface-Induced Regional Climate Change in Southern Israel*. https://journals.ametsoc.org/view/journals/apme/37/11/1520-0450_1998_037_1470_lsirc_2.0.co_2.xml
- Rosenblatt, F. (1962). *Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms*. Spartan Books.
- Rosier, S. H. R., Bull, C. Y. S., Woo, W. L., & Gudmundsson, G. H. (2023). Predicting ocean-induced ice-shelf melt rates using deep learning. *The Cryosphere*, 17(2), 499–518. <https://doi.org/10.5194/tc-17-499-2023>
- Ryan, J. C. (2024). Contribution of surface and cloud radiative feedbacks to Greenland Ice Sheet meltwater production during 2002–2023. *Communications Earth & Environment*, 5(1), 538. <https://doi.org/10.1038/s43247-024-01714-y>

- Ryan, J. C., Cooley, S. W., & Webb, E. E. (2026). Revisiting Trends in Greenland Ice Sheet Albedo Using the Combined MODIS and VIIRS Record. *Earth and Space Science*, 13(3), e2025EA004731. <https://doi.org/10.1029/2025EA004731>
- Ryan, J. C., Smith, L. C., van As, D., Cooley, S. W., Cooper, M. G., Pitcher, L. H., & Hubbard, A. (2019). Greenland Ice Sheet surface melt amplified by snowline migration and bare ice exposure. *Science Advances*, 5(3), eaav3738. <https://doi.org/10.1126/sciadv.aav3738>
- Schlager, E., Scher, S., Mottram, R. H., & Langen, P. L. (2026). Learning to melt: Emulating Greenland surface melt from a polar RCM with machine learning. *The Cryosphere*, 20(6), 3313–3343. <https://doi.org/10.5194/tc-20-3313-2026>
- Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, 6(2), 461–464. <https://doi.org/10.1214/aos/1176344136>
- Slater, D. A., & Straneo, F. (2022). Submarine melting of glaciers in Greenland amplified by atmospheric warming. *Nature Geoscience*, 15(10), 794–799. <https://doi.org/10.1038/s41561-022-01035-9>
- Slater, T., Shepherd, A., McMillan, M., Leeson, A., Gilbert, L., Muir, A., Munneke, P. K., Noël, B., Fettweis, X., van den Broeke, M., & Briggs, K. (2021). Increased variability in Greenland Ice Sheet runoff from satellite observations. *Nature Communications*, 12(1), 6069. <https://doi.org/10.1038/s41467-021-26229-4>
- Smith, L. C., Andrews, L. C., Pitcher, L. H., Overstreet, B. T., Rennermalm, Å. K., Cooper, M. G., Cooley, S. W., Ryan, J. C., Miège, C., Kershner, C., & Simpson, C. E. (2021). Supraglacial River Forcing of Subglacial Water Storage and Diurnal Ice Sheet Motion. *Geophysical Research Letters*, 48(7). <https://doi.org/10.1029/2020GL091418>
- Smith, L. C., Yang, K., Pitcher, L. H., Overstreet, B. T., Chu, V. W., Rennermalm, Å. K., Ryan, J. C., Cooper, M. G., Gleason, C. J., Tedesco, M., Jeyaratnam, J., van As, D., van den Broeke, M. R., van de Berg, W. J., Noël, B., Langen, P. L., Cullather, R. I., Zhao, B., Willis, M. J., ... Behar, A. E. (2017). Direct measurements of meltwater runoff on the Greenland ice sheet surface. *Proceedings of the National Academy of Sciences*, 114(50), E10622–E10631. (world). <https://doi.org/10.1073/pnas.1707743114>
- Stone, M. (1974). Cross-Validatory Choice and Assessment of Statistical Predictions. *Journal of the Royal Statistical Society. Series B (Methodological)*, 36(2), 111–147.
- Swingedouw, D., Houssais, M.-N., Herbaut, C., Blaizot, A.-C., Devilliers, M., & Deshayes, J. (2022). AMOC Recent and Future Trends: A Crucial Role for Oceanic Resolution and Greenland Melting? *Frontiers in Climate*, 4. <https://doi.org/10.3389/fclim.2022.838310>
- Tedesco, M., Matai, R., & Fettweis, X. (2026). The Modèle Atmosphérique Régional – Intelligence Artificielle (MAR-IA): Surface meltwater over Greenland. *The Cryosphere*, 20(7), 4235–4253. <https://doi.org/10.5194/tc-20-4235-2026>
- Tedstone, A. (2017). *Proglacial discharge measurements, Leverett Glacier, south-west Greenland (2009-2012)* (p. 4.5 MB, 12 files) [Text/csv]. Polar Data Centre; British Antarctic Survey, Natural Environment Research Council; Cambridge, CB3 0ET, UK. <https://doi.org/10.5285/17C400F1-ED6D-4D5A-A51F-AAD9EE61CE3D>
- Timmermans, G., Noël, B., Kittel, C., Dethinne, T., Ghilain, N., & Fettweis, X. (2026). Evaluation of MARv3.14 over the Greenland Ice Sheet. *EGU sphere*, 1–30. <https://doi.org/10.5194/egusphere-2026-2341>
- Xu, Y., Rignot, E., Menemenlis, D., & Koppes, M. (2012). Numerical experiments on subaqueous melting of Greenland tidewater glaciers in response to ocean warming and enhanced subglacial discharge. *Annals of Glaciology*, 53(60), 229–234. <https://doi.org/10.3189/2012AoG60A139>
- Yang, K., Smith, L. C., Fettweis, X., Gleason, C. J., Lu, Y., & Li, M. (2019). Surface meltwater runoff on the Greenland ice sheet estimated from remotely sensed supraglacial lake infilling rate. *Remote Sensing of Environment*, 234, 111459. <https://doi.org/10.1016/j.rse.2019.111459>
- Yang, K., Sommers, A., Andrews, L. C., Smith, L. C., Lu, X., Fettweis, X., & Li, M. (2020). Intercomparison of surface meltwater routing models for the Greenland ice sheet and influence on subglacial effective pressures. *The Cryosphere*, 14(10), 3349–3365. <https://doi.org/10.5194/tc-14-3349-2020>
- Young, T. J., Christoffersen, P., Bougamont, M., Tulaczyk, S. M., Hubbard, B., Mankoff, K. D., Nicholls, K. W., & Stewart, C. L. (2022). Rapid basal melting of the Greenland Ice Sheet from surface meltwater drainage. *Proceedings of the National Academy of Sciences*, 119(10), e2116036119. <https://doi.org/10.1073/pnas.2116036119>
- Zhang, Y., Geng, S., & Wang, Y. (2026). Satellite-Derived Summer Albedo Variations on the Greenland Ice Sheet from 1979 to 2024 Linked with Climatic Indices. *Remote Sensing*, 18(2), 295. <https://doi.org/10.3390/rs18020295>