



RESEARCH ARTICLE

10.1029/2024JH000169

Key Points:

- We present the application of physics-informed neural networks, a unified framework solving forward and inverse, to ice sheet modeling
- We are able to infer parameters that traditional numerical methods cannot invert for using momentum conservation, such as ice thickness
- We show the capability of the framework: simultaneous inversion of basal friction while interpolating sparse ice thickness observations

Correspondence to:

G. Cheng,
gong.cheng@dartmouth.edu

Citation:

Cheng, G., Morlighem, M., & Francis, S. (2024). Forward and inverse modeling of ice sheet flow using physics-informed neural networks: Application to Helheim Glacier, Greenland. *Journal of Geophysical Research: Machine Learning and Computation*, 1, e2024JH000169. <https://doi.org/10.1029/2024JH000169>

Received 9 FEB 2024
Accepted 24 JUL 2024

Author Contributions:

Conceptualization: Gong Cheng, Mathieu Morlighem
Data curation: Gong Cheng, Sade Francis
Formal analysis: Gong Cheng
Funding acquisition: Mathieu Morlighem
Investigation: Gong Cheng, Sade Francis
Methodology: Gong Cheng, Mathieu Morlighem
Project administration: Gong Cheng, Mathieu Morlighem
Resources: Gong Cheng
Software: Gong Cheng
Supervision: Gong Cheng, Mathieu Morlighem
Validation: Gong Cheng
Visualization: Gong Cheng

© 2024 The Author(s). Journal of Geophysical Research: Machine Learning and Computation published by Wiley Periodicals LLC on behalf of American Geophysical Union.
This is an open access article under the terms of the [Creative Commons Attribution License](#), which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Forward and Inverse Modeling of Ice Sheet Flow Using Physics-Informed Neural Networks: Application to Helheim Glacier, Greenland

Gong Cheng¹ , Mathieu Morlighem¹ , and Sade Francis¹
¹Department of Earth Sciences, Dartmouth College, Hanover, NH, USA

Abstract Predicting the future contribution of the ice sheets to sea level rise over the next decades presents several challenges due to a poor understanding of critical boundary conditions, such as basal sliding. Traditional numerical models often rely on data assimilation methods to infer spatially variable friction coefficients by solving an inverse problem, given an empirical friction law. However, these approaches are not versatile, as they sometimes demand extensive code development efforts when integrating new physics into the model. Furthermore, this approach makes it difficult to handle sparse data effectively. To tackle these challenges, we use the Physics-Informed Neural Networks (PINNs) to seamlessly integrate observational data and governing equations of ice flow into a unified loss function, facilitating the solution of both forward and inverse problems within the same framework. We illustrate the versatility of this approach by applying the framework to two-dimensional problems on the Helheim Glacier in southeast Greenland. By systematically concealing one variable (e.g., ice speed, ice thickness, etc.), we demonstrate the ability of PINNs to accurately reconstruct hidden information. Furthermore, we extend this application to address a challenging mixed inversion problem. We show how PINNs are capable of inferring the basal friction coefficient while simultaneously filling gaps in the sparsely observed ice thickness. This unified framework offers a promising avenue to enhance the predictive capabilities of ice sheet models, reducing uncertainties, and advancing our understanding of poorly constrained physical processes.

Plain Language Summary Our ability to predict the future contribution of the ice sheets to future sea-level rise is limited due to the lack of observations, especially at the base of the ice sheets. Traditional computer models infer basal sliding from observations at the surface based on ice flow physics, a process that becomes complex and inflexible when incorporating new information or a more sophisticated description of ice flow. Our solution involves Physics-Informed Neural Networks that seamlessly integrate data and physical laws in a unified framework. We demonstrate the versatility of Physics-Informed Neural Networks (PINNs) on Helheim Glacier in Southeast Greenland, showcasing their ability to handle missing or incomplete data. Additionally, we extend PINNs to address a challenging problem, which consists of inferring basal sliding while filling gaps in sparsely observed ice thickness at the same time. This unified approach holds promise for improving ice sheet predictions and advancing our understanding of complex ice dynamics.

1. Introduction

The Greenland and Antarctic ice sheets are collectively responsible for more than half of the observed global sea-level rise in recent decades (Frederikse et al., 2020). This mass loss is primarily driven by the dynamic behavior of marine-terminating glaciers, which discharge icebergs into the ocean (Mouginot et al., 2019; Rignot et al., 2019). Current projections of sea level rise rely on transient simulations (Eyring et al., 2016; Nowicki et al., 2016) that capture the response of these outlet glaciers to climate forcings. These numerical models solve forward problems to compute the evolution of ice velocity and geometry (Bueler & Brown, 2009; Colinge & Blatter, 1998; Durand et al., 2009; Mangeney & Califano, 1998), which, in turn, affect the discharge of ice into the ocean.

Among all the physical processes influencing current and future ice discharge, basal friction is a major control on the dynamics of these outlet glaciers. Basal friction characterizes a relationship between the ice basal velocity and the stress exerted by the underlying bed (Budd et al., 1979; Gagliardini et al., 2007; Weertman, 1957). This relationship has a significant influence on how the ice sheets respond to external forcings, particularly in the context of climate change (Barnes et al., 2021; Brondex et al., 2019; Yu et al., 2018). Therefore, improving our

Writing – original draft: Gong Cheng,
Mathieu Morlighem
Writing – review & editing: Gong Cheng,
Mathieu Morlighem, Sade Francis

understanding of basal conditions and accurately representing them in ice sheet numerical models is critical (Åkesson et al., 2021; Brondex et al., 2017).

Empirical friction laws typically incorporate a friction “coefficient,” which is spatially and sometimes temporally variable, encapsulating less-understood physical processes. This friction coefficient cannot be directly measured and is generally inferred from surface observations using data assimilation methods (D. MacAyeal, 2002; Morlighem et al., 2013; Petra et al., 2012; Tarantola, 2005). These methods involve solving inverse problems aimed at minimizing an objective or cost function, which is constructed based on the misfit between observed and modeled variables, such as surface velocity or surface elevation (Goldberg & Sergienko, 2011; Morlighem & Goldberg, 2023). These observational data are often associated with noise, and the uncertainties, as well as model error, propagate through the inverse model and end up lumped into the friction coefficient (Cheng & Lotstedt, 2020; Karniadakis et al., 2021), which contributes significantly to the overall uncertainty in ice sheet modeling. Furthermore, solving inverse problems often requires specially designed numerical methods, demanding different formulations and sometimes complex computer codes (Griewank et al., 1996; Vogel, 2002). These challenges underscore the pressing need for novel approaches to potentially alleviate these issues and enhance our ability to model and understand the dynamics of ice sheets.

Recent advances in machine learning techniques, coupled with the wealth of new remote sensing data, offer opportunities to develop new approaches to better constrain numerical models with data. Among all the emerging methods in recent years, neural networks have shown to be a promising tool to complement traditional ice flow models. These traditional models often rely on specifically designed numerical methods to solve forward and inverse problems, for example, Durand et al. (2009); Mangeney and Califano (1998); Bueler and Brown (2009); Colinge and Blatter (1998); Tarantola (2005); D. MacAyeal (2002); Petra et al. (2012); Morlighem et al. (2013), etc. In the remainder of this paper, these numerical methods will be referred to as “traditional methods.” Machine learning-based methods offer significant benefits in three key application areas: data assimilation (Bolibar et al., 2023; Iwasaki & Lai, 2023; Juvet & Cordonnier, 2023; Riel & Minchew, 2023; Riel et al., 2021), GPU-based forward ice flow emulator (Brinkerhoff, 2022; He et al., 2023; Juvet, 2022; Rahnemoonfar & Koo, 2024), and hybrid approaches that combine these advantages to create powerful new methods for ice flow modeling (Juvet & Cordonnier, 2023). Particularly, Physics-Informed Neural Networks have been designed to tackle complex problems associated with intricate mathematical operations (Karniadakis et al., 2021; Lu, Meng, et al., 2021; Raissi et al., 2019). PINNs seamlessly integrate observational data with underlying physical laws and have been used in a wide range of applications (e.g., Karniadakis et al., 2021; Lu, Meng, et al., 2021; Riel et al., 2021; Teisberg et al., 2021; Wang et al., 2022). The PINN loss function represents a weighted sum of data misfits and the residuals of the governing equations, typically in the form of partial differential equations (PDEs). This contrasts with standard deep neural networks, which solely learn from available data. By enforcing physical laws as soft constraints, PINNs make it possible to infer quantities without direct observations.

Here, we explore the use of PINNs in ice sheet modeling by solving two-dimensional forward and inverse problems on a real Greenland glacier. By training PINNs with a carefully designed loss function, while knowing all variables except one, we try to recover the unknown. Depending on the choice of training data, we use this framework as a forward solver to get the ice velocity, or as an inverse solver to infer the basal friction coefficient. To evaluate whether PINNs can effectively and accurately reconstruct the unknown field, we compare its solution with the true solution obtained from a conventional numerical model. Furthermore, we try to infer other parameters within the governing equations using this framework, even those challenging to invert using traditional numerical methods. In a second set of experiments, we extend the application of PINNs to address a challenging mixed-inversion problem: we test the ability of the framework to infer the basal friction coefficient, while simultaneously filling in gaps in sparsely observed ice thickness. This type of mixed inversion problems has been previously explored in fluid mechanics applications (Raissi et al., 2020), mountain glaciers (Juvet & Cordonnier, 2023), initial ice thickness (Perego et al., 2014), but remains challenging for traditional numerical methods.

2. Method

2.1. Physical Model—Shelfy Stream Approximation

An ice sheet or glacier usually occupies a three-dimensional shallow volume with a small aspect ratio, where the width of the domain is several orders of magnitudes larger than the height of the domain. Therefore, the variation

in the vertical direction can generally be neglected, by assuming a plug flow in fast flowing areas, which leads to the so-called Shelfy Stream Approximation (MacAyeal, 1989, SSA).

Consider a two-dimensional domain $\mathbf{x} = (x, y) \in \Omega$, where the ice velocity is denoted as $\mathbf{u} = (u, v)^T$. The governing ice dynamics are described by the SSA (MacAyeal, 1989), expressed as a system of PDEs:

$$\nabla \cdot \boldsymbol{\sigma} + \boldsymbol{\tau}_b = \rho_i g H \nabla s \quad (1)$$

where $\boldsymbol{\tau}_b = (\tau_{bx}, \tau_{by})^T$ represents the basal shear stress, ρ_i is the ice density, g is the gravitational acceleration, s is the surface elevation, and H is the ice thickness. The stress tensor $\boldsymbol{\sigma}$ of the SSA model is defined as

$$\boldsymbol{\sigma} = \mu H \begin{bmatrix} 4 \frac{\partial u}{\partial x} + 2 \frac{\partial v}{\partial y} & \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \\ \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} & 2 \frac{\partial u}{\partial x} + 4 \frac{\partial v}{\partial y} \end{bmatrix}. \quad (2)$$

The ice viscosity, μ , is determined by Glen's flow-law (Glen, 1958), which in two dimensions reads:

$$\mu = \frac{B}{2} \left(\left(\frac{\partial u}{\partial x} \right)^2 + \left(\frac{\partial v}{\partial y} \right)^2 + \frac{1}{4} \left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \right)^2 + \frac{\partial u}{\partial x} \frac{\partial v}{\partial y} \right)^{\frac{1-n}{2n}}, \quad (3)$$

where $n = 3$ is the flow-law exponent, and B is the pre-factor dependent on ice temperature (Cuffey & Paterson, 2010), among other factors.

We assume here that the basal shear stress $\boldsymbol{\tau}_b$ is related to the ice velocity $\mathbf{u}$ by Weertman's friction law (Weertman, 1957)

$$\boldsymbol{\tau}_b = -C^2 |\mathbf{u}|^{m-1} \mathbf{u}, \quad (4)$$

where C is a spatially varying friction coefficient and $m = 1/3$. While various empirical friction laws exist (e.g., Budd et al., 1979; Gagliardini et al., 2007; Weertman, 1957), for simplicity, we focus on Weertman's law in this work. Nevertheless, the methodology presented can be generalized to accommodate other friction laws.

Many of the Greenlandic glaciers are marine-terminating glaciers, which have calving fronts at the ice-ocean interface. The boundary conditions on the calving front Γ are defined as

$$\boldsymbol{\sigma} \mathbf{n} = \frac{1}{2} g (\rho_i H^2 - \rho_w b^2) \mathbf{n} \quad (5)$$

where $\mathbf{n} = (n_x, n_y)^T$ is the outward normal unit vector on Γ , ρ_w is the density of sea water, and $b = s - H$ is the bed elevation. For the remaining outer boundaries of the domain, we apply Dirichlet boundary conditions.

2.2. Data

To effectively train the PINNs, we rely on both direct measurements and reanalysis models. For simplicity, we will refer to these training data sets as “data” throughout this manuscript. These data sets encompass a range of variables, including ice velocity, surface elevation, ice front positions, ice thickness, and inferred basal friction coefficients. In this work, we focus on Helheim Glacier in Southeast Greenland, as shown in Figure 1.

The ice surface velocity mosaic is from NASA's MEaSUREs products (Joughin et al., 2018). We denote the velocity observations at N_u distinct locations $\{\mathbf{x}_j^u\}_{j=1}^{N_u}$ as $\hat{\mathbf{u}}_j = (\hat{u}_j, \hat{v}_j)^T$. Surface elevation data $\hat{s}_j$ at the location $\{\mathbf{x}_j^s\}_{j=1}^{N_s}$ are from the Greenland Ice Mapping Project (Howat et al., 2014), while ice thickness $\hat{H}_j$ at $\{\mathbf{x}_j^H\}_{j=1}^{N_H}$ is from

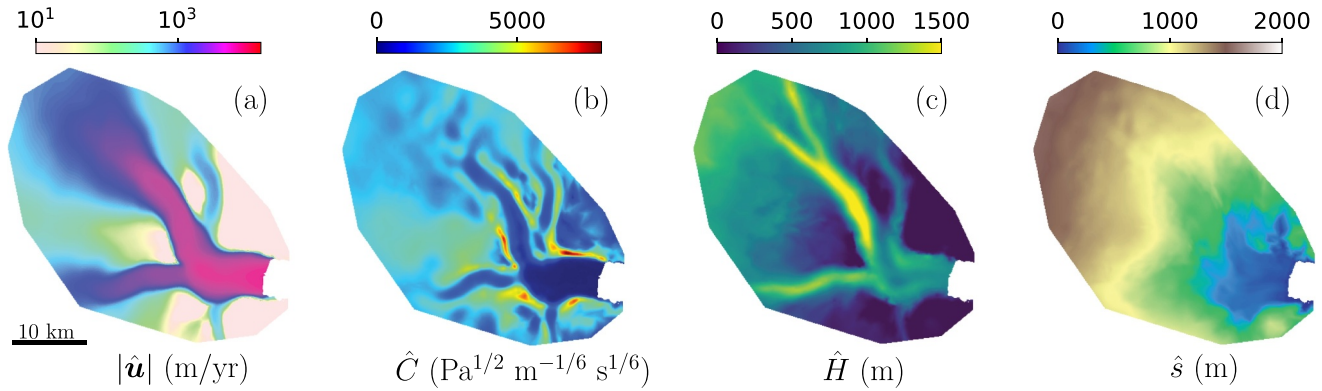


Figure 1. Reference data from Helheim Glacier. (a) Velocity magnitude (b) basal friction coefficient (c) ice thickness (d) surface elevation. The color map in (a) is on a logarithmic scale, and the rest of subplots are on linear scales.

BedMachine Greenland v6.1 (Morlighem et al., 2017). The ice front position is derived using the ice mask from Howat et al. (2014).

To complete the data used in training the PINNs and assessing their performance as forward and inverse models, we require the friction coefficient $\hat{C}_j$ at specific locations $\{\mathbf{x}_j^C\}_{j=1}^{N_c}$. Since there is no direct observation of the friction coefficient, we perform an inversion using the Ice-sheet and Sea-level System Model (ISSM, Larour et al., 2012), using the aforementioned data sets to infer the friction coefficient following the method described in Morlighem et al. (2013).

2.3. Physics-Informed Neural Networks

The fundamental difference between this work and many other physics-informed neural networks in ice sheet modeling, such as those discussed in Jouvet and Cordonnier (2023); Bolibar et al. (2023); Brinkerhoff (2022); He et al. (2023); Rahnmounfar and Koo (2024); Riel et al. (2021), lies in what and how the neural network learns. Specifically, the neural networks in He et al. (2023); Bolibar et al. (2023); Rahnmounfar and Koo (2024) act as emulators, learning the relationships among the variables from the numerical solutions of the PDE. In Jouvet and Cordonnier (2023), the neural networks also function similarly as emulators but learn the relationship directly from the PDE by integrating the PDE residual into the loss function during training. Meanwhile, in Riel et al. (2021), the neural networks learn simultaneously from the data and from physics-based properties, such as the smoothness and sign of the basal drag.

Our approach, following Raissi et al. (2019); Iwasaki and Lai (2023); Wang et al. (2022), differs by having the neural network learn all the variables involved in the PDE as functions of the coordinates (x, y) in the domain. We implement a PINN to assimilate both the physical model detailed in Section 2.1 and the data presented in Section 2.2. The architecture of the PINN is schematically illustrated in Figure 2. In this configuration, fully connected parallel neural networks are employed, with the inputs designated as the coordinates x and y in the two-dimensional plane (or x in the one-dimensional cases). The PINN generates predictions for five different outputs. Specifically, three parallel neural networks, each comprising 6 layers with 20 nodes per layer, are utilized to generate predictions for ice velocity, ice geometry, and friction coefficient, respectively. We employ the hyperbolic tangent activation function for all nodes in the PINN. We normalize the input variables to the range $[-1, 1]$ across the computational domain before feeding them into the network. Additionally, the network outputs are denormalized from the range $[-1, 1]$ to their actual values as used in the governing PDEs.

2.4. The Loss Function

We formulate a loss function $\mathcal{L}(\theta)$ to optimize the parameters θ of the neural networks. This loss function takes the form

$$\mathcal{L} = \mathcal{L}_u + \mathcal{L}_g + \mathcal{L}_C + \mathcal{L}_\varphi \quad (6)$$

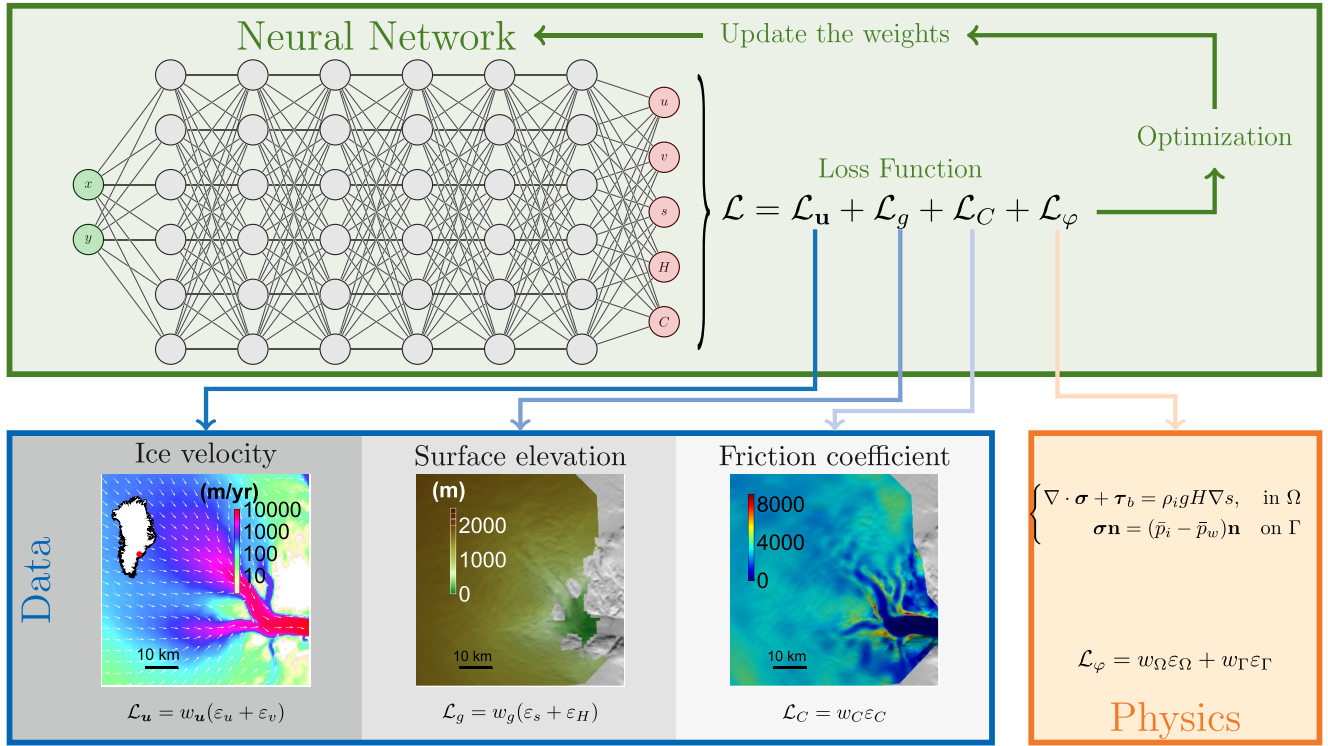


Figure 2. Illustration of the Physics-Informed Neural Network. Note that the neural network architecture depicted here is for illustrative purposes only, and does not represent the actual configuration used in this study described in the text.

where each term represents the contributions to the loss function associated with subscripts u (velocity), g (ice geometry), C (friction coefficients), and φ (PDE residuals). Specifically, our loss function comprises a weighted sum of data misfits and residuals from the governing PDEs. We further denote these misfits and residuals by ε , and the weights by w , with the same subscripts as the ones described above.

The data misfits are calculated by the mean-square errors (MSE) of the PINN's predictions at locations where the corresponding observational data is available, and they contribute to the loss function as

$$\mathcal{L}_u = w_u(\varepsilon_u + \varepsilon_v) = \frac{w_u}{N_u} \sum_{j=1}^{N_u} \left((u(\mathbf{x}_j^u) - \hat{u}_j)^2 + (v(\mathbf{x}_j^u) - \hat{v}_j)^2 \right), \quad (7)$$

$$\mathcal{L}_g = w_g(\varepsilon_s + \varepsilon_H) = \frac{w_g}{N_s} \sum_{j=1}^{N_s} (s(\mathbf{x}_j^s) - \hat{s}_j)^2 + \frac{w_g}{N_H} \sum_{j=1}^{N_H} (H(\mathbf{x}_j^H) - \hat{H}_j)^2, \quad (8)$$

$$\mathcal{L}_C = w_C \varepsilon_C = \frac{w_C}{N_C} \sum_{j=1}^{N_C} (C(\mathbf{x}_j^C) - \hat{C}_j)^2, \quad (9)$$

where $u(\mathbf{x})$, $v(\mathbf{x})$, $s(\mathbf{x})$, $H(\mathbf{x})$, and $C(\mathbf{x})$ are the output of the PINN as shown in Figure 2. The weights, w_u , w_g , and w_C , are scaled according to the typical values of the ice velocity, ice thickness, and friction coefficient in the International System of Units (SI), to balance their contributions to the total loss function. For Helheim Glacier, the typical values of these variables and the corresponding weights are shown in Table 1.

The PDEs in Equation 1 are evaluated using the output of the PINN at a distinct set of collocation points $\{\mathbf{x}_j^\Omega\}_{j=1}^{N_\Omega}$, which are different from the observational data sets in Section 2.2. The residual of the PDEs is expressed as the mean-square errors among these collocation points together with the boundary conditions. Specifically, for the

Table 1
Typical Values of Variables in Helheim Glacier

Variable	Typical value	Weights	Value
$ u $	10^4 m yr^{-1}	w_u	$10^{-8} \times (315,360,00^2) \text{ m}^{-2} \text{ s}^2$
s, H	10^3 m	w_g	10^{-6} m^{-2}
C	$10^4 \text{ Pa}^{1/2} \text{ m}^{-1/6} \text{ s}^{1/6}$	w_C	$10^{-8} \text{ Pa}^{-1} \text{ m}^{1/3} \text{ s}^{-1/3}$
τ_b	10^5 Pa	w_Ω	10^{-10} Pa^{-2}
$\rho_i g H^2$	10^9 Pa m	w_Γ	$10^{-18} \text{ Pa}^{-2} \text{ m}^{-2}$

Helheim Glacier, we include the calving front boundary condition along the points at $\{\mathbf{x}_j^\Gamma\}_{j=1}^{N_\Gamma}$ in evaluating the residual of the PDEs as

$$\begin{aligned} \mathcal{L}_\varphi &= w_\Omega \varepsilon_\Omega + w_\Gamma \varepsilon_\Gamma \\ &= \frac{w_\Omega}{N_\Omega} \sum_{i=1}^{N_\Omega} \left\| \nabla \cdot \boldsymbol{\sigma} + \boldsymbol{\tau}_b - \rho_i g H \nabla s \right\|^2 \\ &\quad + \frac{w_\Gamma}{N_\Gamma} \sum_{i=1}^{N_\Gamma} \left\| \boldsymbol{\sigma} \mathbf{n} - \frac{1}{2} g (\rho_i H^2 - \rho_w b^2) \mathbf{n} \right\|^2, \end{aligned} \quad (10)$$

where the weights w_Ω and w_Γ are chosen by scaling the basal shear stress and boundary force as in Table 1.

Effectively constructing a robust loss function for the PINN poses a significant challenge due to the diverse nature of its components, each representing distinct physical quantities with inherently different orders of magnitudes. Importantly, these components serve as soft constraints within the loss function, collectively influencing the learning process of the neural network (Iwasaki & Lai, 2023; Lu, Meng, et al., 2021). The success of PINNs in ice sheet modeling critically relies on the careful selection of weights assigned to each component in the loss function. To optimize the predictive capabilities of the PINN before applying it to real problems at Helheim Glacier, we perform an extensive grid search to find the best weights outlined in Table 1. We fix w_u and vary the other four weights by several orders of magnitude around the typical values. This search aims to validate that the values in Table 1 are indeed optimal for the diverse components in the loss function. We systematically vary w_g and w_C by ± 2 orders of magnitude, while adjusting w_Ω by ± 5 orders of magnitude. For w_Γ , the ratio w_Ω/w_Γ is fixed at 10^{-8} . To ensure robustness, each experiment is repeated at least 25 times using a Glorot normal initializer with distinct random seeds. To speed up the grid search procedure, we perform these validation experiments exclusively on an inverse problem along a flowline of Helheim Glacier, treating it as a one-dimensional problem with fewer data points. The main results are shown in Section 3.1.

2.5. Numerical Experiments 1–3

After validating the choice of the weights, we apply the PINN to the two-dimensional data sets obtained from Helheim Glacier, as detailed in Section 2.2. Subsequently, we conduct three distinct sets of numerical experiments by intentionally concealing one component within the training data sets.

In the first experiment, we test the ability of the PINN to solve the forward problem. This involves utilizing the friction coefficient, $\hat{C}$, and the ice geometry, $\hat{s}$ and $\hat{H}$, to solve for the ice velocity, $\mathbf{u}$. The second experiment tests the ability of the PINN to solve an inverse problem and infer basal friction based on observed velocities. Here, we train the network with ice velocity, $\hat{\mathbf{u}}$, and ice geometry, $\hat{s}$ and $\hat{H}$, and infer the basal friction coefficient, C . In the third experiment, we test the ability of the PINN to infer the ice thickness, H . This is accomplished using the ice velocity, $\hat{\mathbf{u}}$, surface elevation, $\hat{s}$, and the basal friction coefficient, $\hat{C}$, as training data sets. For all these experiments, we calculate the misfits as the PINN predictions minus the “true” solutions.

2.6. Experiment 4: Dual Inversion

In many real-world scenarios, the basal friction coefficient is unknown and needs to be inferred from surface observations, and the ice thickness is only known along flight lines, leaving large gaps that are generally filled using interpolation methods. Figure 3 illustrates all available flight tracks around Helheim Glacier, with dots representing resampled points at 200 m intervals along the tracks. These flight track data are notably sparse, even along the main branch of Helheim Glacier, where only one flight track is present in the center of the ice stream. Various numerical methods have been developed to leverage flight track data along with other observations to fill gaps in regions lacking direct measurements. Some examples include the BedMachine Greenland and Antarctica models (Morlighem et al., 2017, 2020), which use mass conservation principles to constrain ice thickness.

Given the flexibility of the PINN, we perform one more test here to assess its ability to address a *dual* inversion problem. Here we would like to test the ability of the PINN to infer the basal friction coefficient, C , while simultaneously filling gaps in sparsely observed ice thickness, H . Following the same procedure as the ones

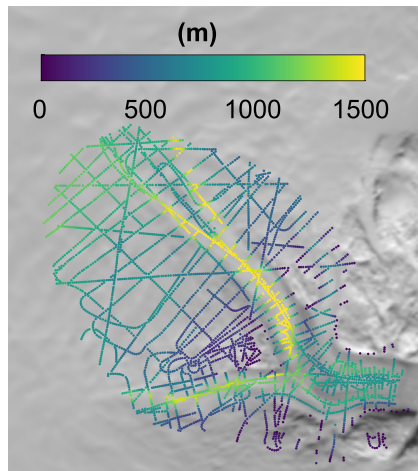


Figure 3. Available ice thickness data in the region of interest. The dots are resampled at 200 m intervals, overlaid with an image map from MEaSUREs MODIS Mosaic of Greenland (Haran et al., 2018).

described above, we expose the model to ice velocity, $\hat{\mathbf{u}}$, surface elevation, $\hat{s}$, and ice thickness only along flight tracks, $\hat{H}$, as shown in Figure 3.

For all these four experiments, we use 4,000 data points for the training sets, and 9,000 collocation points for evaluating the PDE residual. The calving front is described by 121 data points with spacing of 200 m, and the Dirichlet boundary around the domain has 541 data points. All the experiments are trained with Adam optimization up to 1,000,000 epochs on the NVIDIA Tesla V100 SXM2 32GB GPU. We use TensorFlow 2.4.1 implementation of the Adam optimizer, with a learning rate at 0.001, and hyper-parameters $\beta_1 = 0.99$, $\beta_2 = 0.999$, along with a tolerance of $\epsilon = 0.1$. Further details regarding the definition and algorithm can be found in Kingma and Ba (2014).

3. Results

3.1. Choice of Weights

In total, we conducted over 15,000 experiments, systematically exploring various combinations of weights. Perhaps unsurprisingly, balancing the different data misfits is relatively straightforward by scaling to the same

magnitude within the SI unit system. However, the balance between the data misfit and PDE residual is particularly delicate in order to obtain satisfactory results.

In Figure 4, we vary w_Ω while keeping all other weights fixed, as specified in Table 1. Each point in the figure represents an experiment, and due to randomization in the training process, we repeated these experiments multiple times to derive an average behavior of the PINN. Figures 4a–4c exhibit distinct “L-curve” patterns. As w_Ω decreases from 10^{-7} to 10^{-15} Pa⁻², the data misfit diminishes, reflecting the increased focus of the loss function on the data, while placing less emphasis on the PDE constraints. The minimum data misfit is achieved at $w_\Omega = 10^{-11}$ Pa⁻². However, errors, particularly the MSE of the velocity, start to diverge, which is indicative of overfitting (Iwasaki & Lai, 2023). This trend becomes more pronounced for $w_\Omega < 10^{-11}$ Pa⁻². Figure 4d further illustrates this by presenting the mean and standard deviation of test errors for each w_Ω . Test error is computed by comparing the PINN's prediction of C with the reference data $\hat{C}$. From this figure, we conclude that the optimal choice for w_Ω is 10^{-10} Pa⁻². This value corresponds to the square of $1/(100 \text{ kPa})$, a typical driving stress in ice sheet modeling, serving as the right-hand side of the PDE. This choice of weight aligns well with physical expectations.

3.2. Forward Problem

Now that we know how to best choose the weights of the cost functions, we can perform our first set of experiments. We train the PINN with the data set of friction coefficient, $\hat{C}$, ice thickness, $\hat{H}$, and surface elevation, $\hat{s}$, described in Section 2.2. In this case, the PINN is treated as a solver for the forward problem, aiming to determine the ice velocity $\mathbf{u}$. Since we are not exposing the ice velocity to the PINN, the ice velocity is solely inferred by the PDE constraint in the loss function. The PINN's predictions and the corresponding misfits with respect to the reference data are shown in Figure 5. For comparative analysis, we evaluate the PINN's predictions on a 400×400 regular grid, resulting in a mesh resolution of approximately 100 m, which is finer than the density of the collocation or data points. In order to calculate the misfit, the reference data is projected onto the same grid using a piecewise cubic, continuously differentiable interpolation. All predictions agree well with the reference data in Figure 1, particularly where training data is available, that is, C , H , and s , as shown in Figures 5f–5h. The root mean squared error (RMSE) of these misfits is presented in Table 2.

The PINN's prediction of the flow velocity closely matches the true solution over most of the area, capturing all branches of Helheim Glacier. Although there are some large misfits locally (e.g., ± 700 m/yr) in the fast flow region in Figure 5e, the RMSE of the misfit is 193.75 m/yr. This represents approximately less than 10% of the average flow velocity over the entire domain (2,028.69 m/yr) and about 2.7% of the highest velocity (7,152.93 m/yr).

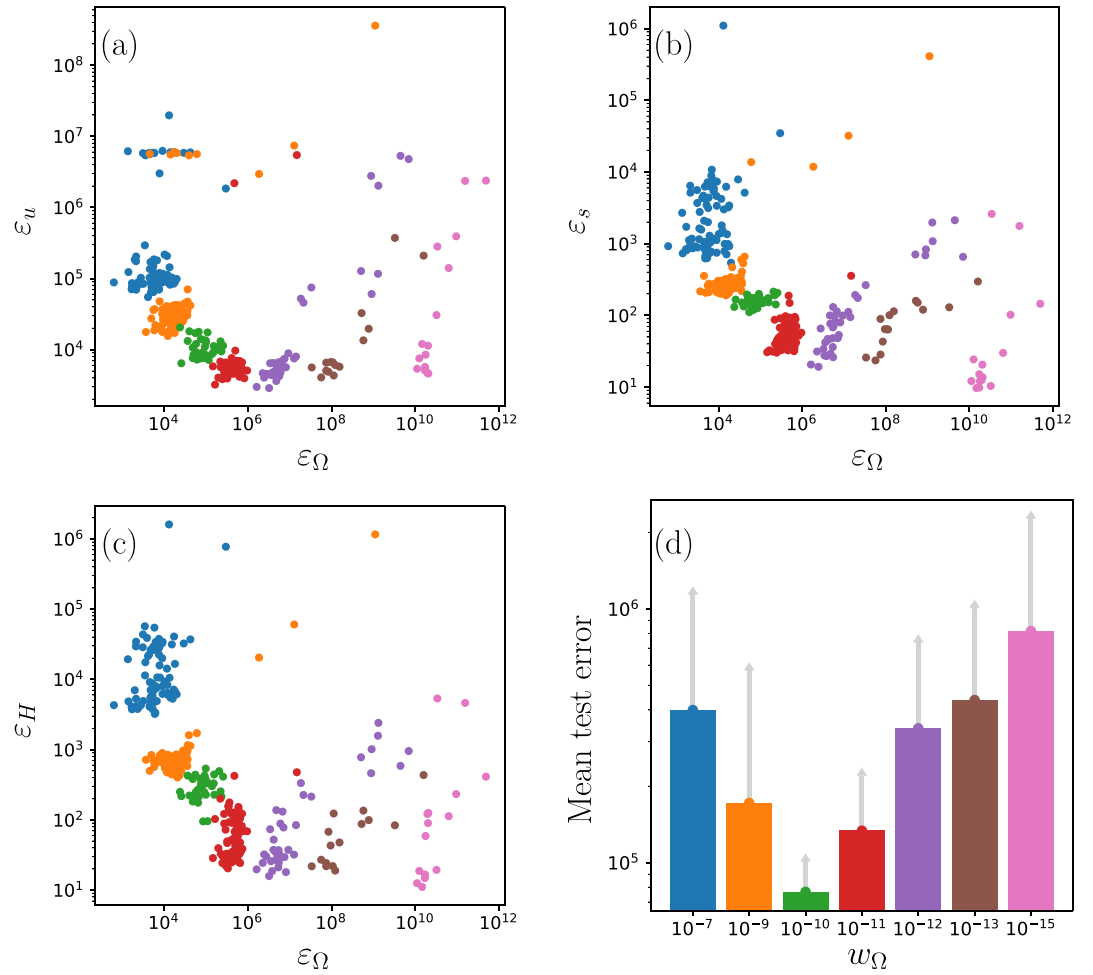


Figure 4. MSE of the (a) velocity, (b) surface elevation, and (c) ice thickness versus the PDE residual ε_Ω . (d) The mean test error of the PINNs predictions using different weights w_Ω .

3.3. Inverse Problem

We change the training data set to use ice velocity $\hat{u}$, ice thickness $\hat{H}$, and surface elevation $\hat{s}$. In this configuration, the PINN serves as an inverse solver to infer the basal friction coefficient C . Again, because we don't expose the PINN to the “true” friction coefficient from the ISSM model inversion, the PINN is inferring C solely based on the PDE constraint that is linking the friction coefficient to the other variables that the PINN is exposed to. The predictions and misfits are presented in Figure 6, and the RMSE of the misfit is provided in Table 2. Similar to the forward problem in Section 3.2, the predictions of PINN align well with the “true” solution. Particularly for those learning from the reference data, the relative errors are all below 3% (the average ice thickness is 716.61 m, and the average surface elevation is 987.66 m).

The RMSE of the misfit in C is $589.61 \text{ Pa}^{1/2} \text{ m}^{-1/6} \text{ s}^{1/6}$, with a relative error 6.69%. However, as shown in Figure 6f, the pattern of large errors is located primarily in the slow-moving region (velocity $< 10 \text{ m/yr}$), where there is minimal ice coverage (ice thickness $< 30 \text{ m}$). In contrast, the predicted friction coefficient is highly accurate in the fast-flow region, capturing all features, including branches and shear margins. The RMSE of misfit in C over the region with $|u| > 10$ is $392.54 \text{ Pa}^{1/2} \text{ m}^{-1/6} \text{ s}^{1/6}$, with a relative error 4.45%.

3.4. Inferring Ice Thickness

The flexibility of the PINN framework provides the possibility to infer the ice thickness H using ice velocity $\hat{u}$, surface elevation $\hat{s}$, and friction coefficient $\hat{C}$. The PINN predictions and their corresponding misfits are shown in

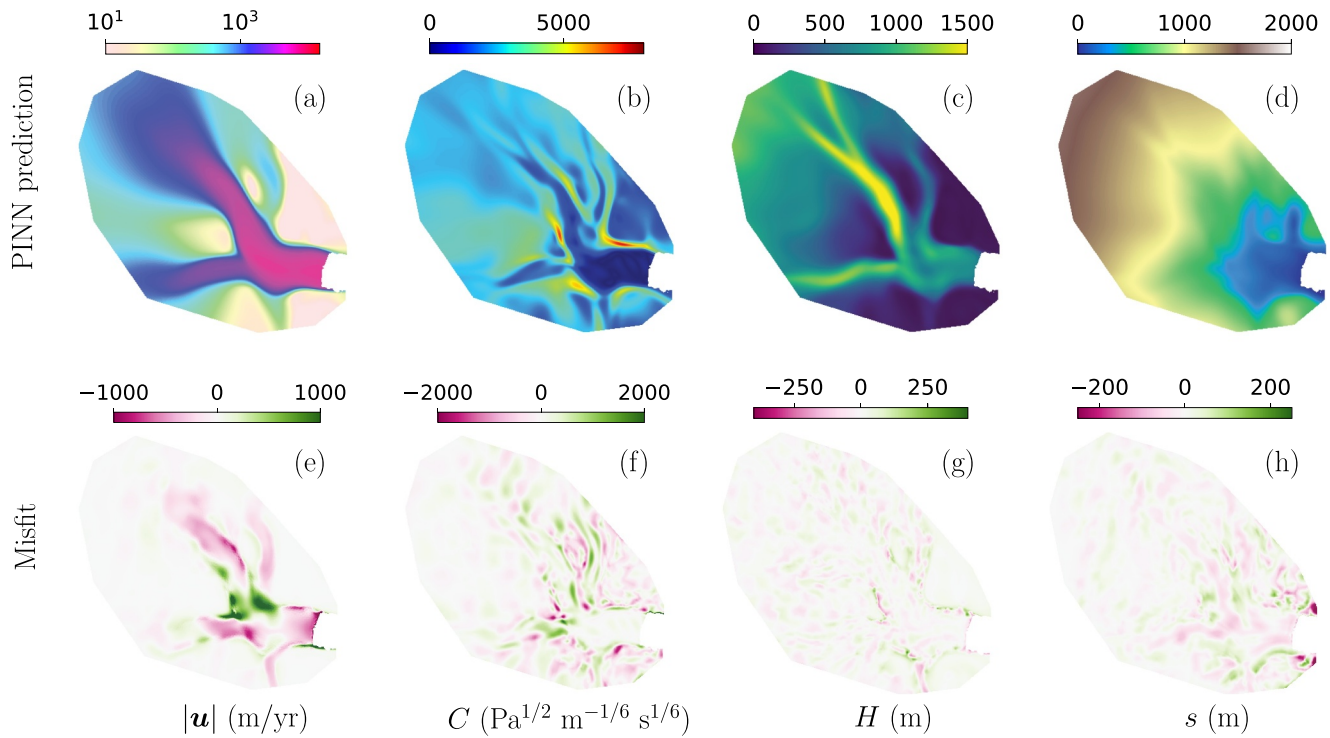


Figure 5. (a)–(d) Predictions of the PINN solving a forward problem without exposure to ice velocity during training. (e)–(h) Corresponding misfits between the predictions and the “true” solutions in Figure 1. The color map in (a) is on a logarithmic scale, and the rest of subplots are on linear scales.

Figure 7. Similar to Sections 3.2 and 3.3, the predictions of u , s , and C align well with the training data, with relative errors of 3.5%, 4.3%, and 1.3%, respectively. The PINN’s prediction of the ice thickness in Figure 7c is a smoothed version of the true solution shown in Figure 1c. The misfit in Figure 7g is distributed fairly evenly in the entire domain with an RMSE of 69.47 m with a relative error of 9.69%, primarily located along the sharp transition of the H , such as the ice front and a part of the northern branch of Helheim Glacier.

3.5. Dual Inversion

The predictions from the PINN and their corresponding misfits are presented in Figure 8. Notably, the PINN predictions for ice velocity and surface elevation align well with the true solutions (shown in Figure 1), and the RMSE of the misfits are 126.83 m/yr for the velocity and 22.08 m for the surface elevation. Both are below those obtained in the forward problem (193.75 m/yr and 26.99 m). The predicted ice thickness closely reproduces the shape and magnitude observed in the true solution as well. While the predicted friction coefficient shows a high misfit in slow-moving regions, as expected given the limitations of SSA in slow-moving regions discussed above, it aligns well with the true solution in fast-flow regions. The RMSE values for both C and H are comparable to those obtained in the individual inversions presented in Sections 3.3 and 3.4 (see Table 2).

4. Discussion

4.1. A Unified Framework

The results presented above show how a single PINN architecture can solve both forward and inverse problems within the same framework. It is worth noting that there are no inherent distinctions between solving different types of problems within this framework. In other words, solving for ice velocity or inferring basal traction or ice thickness is implemented by switching on and off terms in the loss function during the training process. Depending on the training data provided, the PINN learns variables from both observational

Table 2
RMSE for All Experiments

Experiments	$ u $ (m/yr)	C ($\text{Pa}^{1/2} \text{ m}^{-1/6} \text{ s}^{1/6}$)	H (m)	s (m)
Forward problem	193.75	269.54	30.33	26.99
Inverse problem	56.01	589.61	18.96	14.11
Invert for H	70.64	114.55	69.47	12.99
Dual inversion	126.83	899.53	88.39	22.08

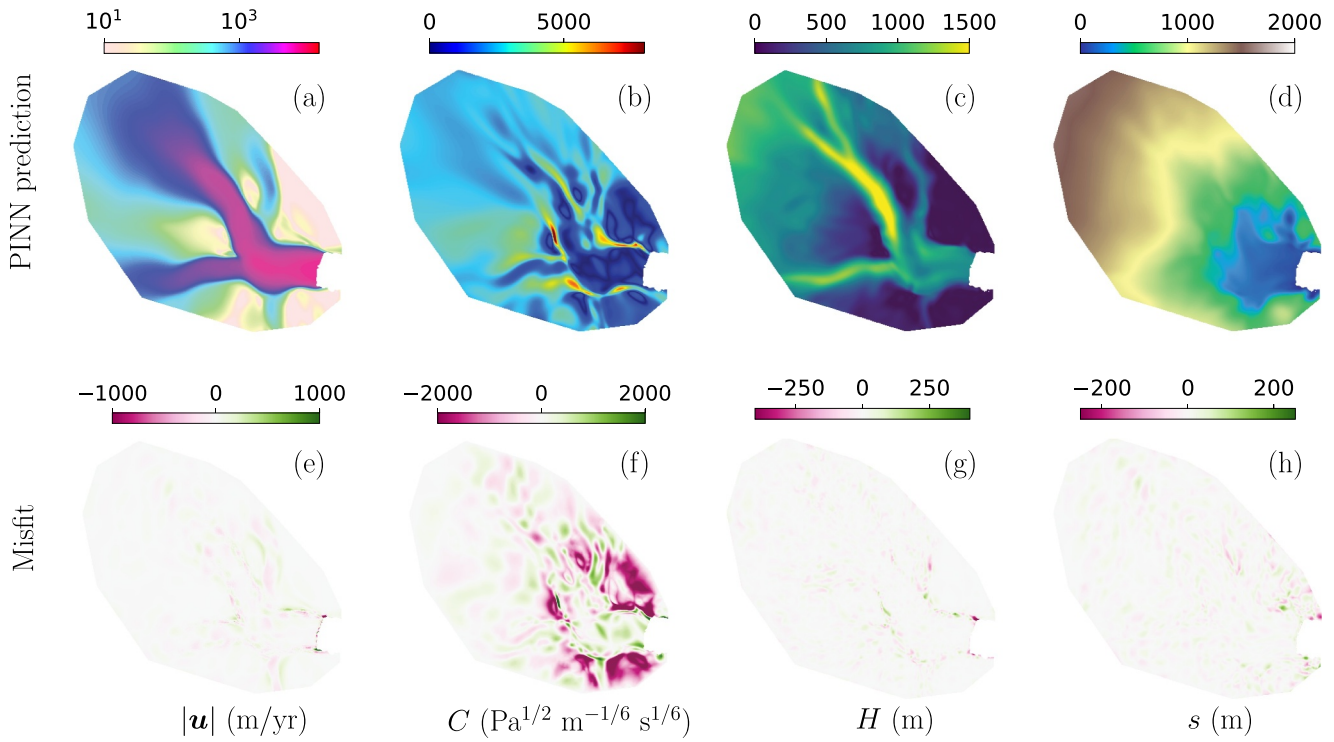


Figure 6. (a)–(d) Predictions of the PINN solving an inverse problem without exposure to the friction coefficient during training. (e)–(h) Corresponding misfits between the predictions and the “true” solutions in Figure 1. The color map in (a) is on a logarithmic scale, and the other panels are on linear scales.

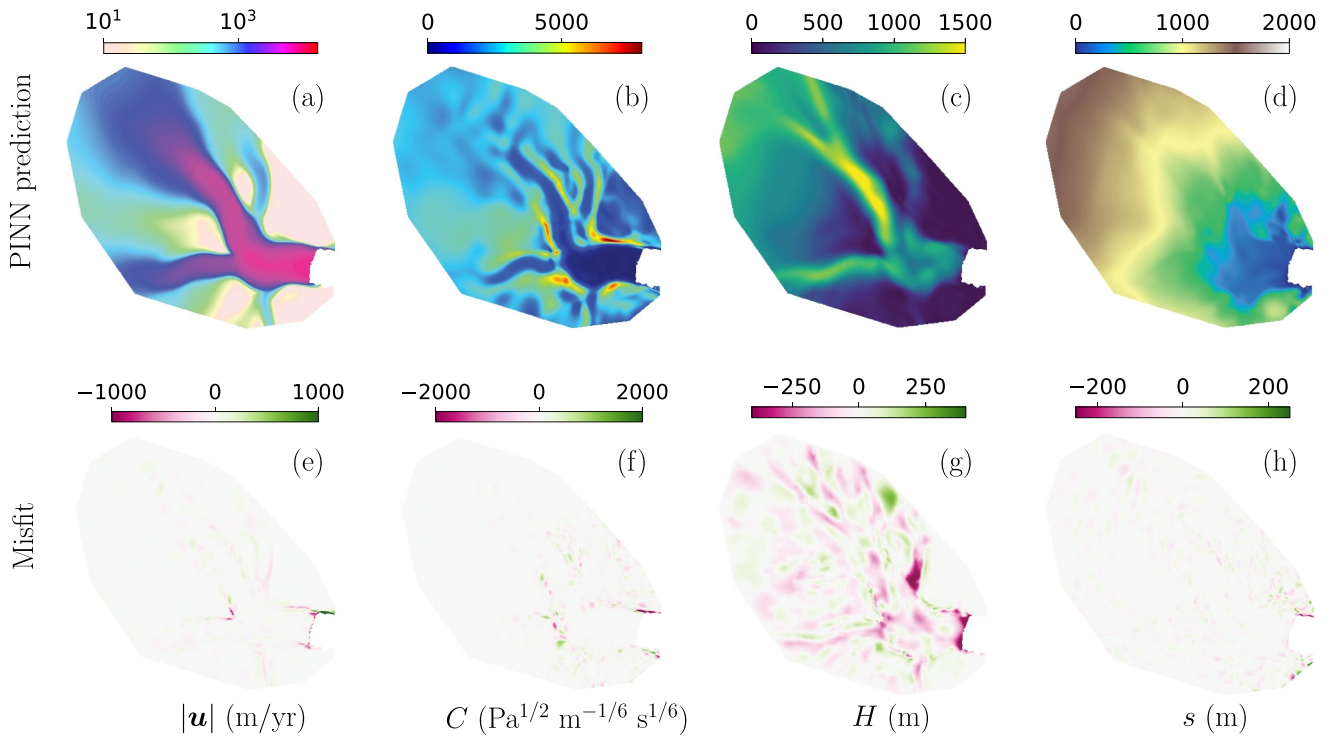


Figure 7. (a)–(d) Predictions of the PINN inferring ice thickness using ice velocity $\hat{\mathbf{u}}$, surface elevation $\hat{s}$, and friction coefficient $\hat{C}$ in the training procedure. (e)–(h) Corresponding misfits between the predictions and their corresponding reference data in Figure 1. The color map in (a) is on a logarithmic scale, and the other panels are on linear scales.

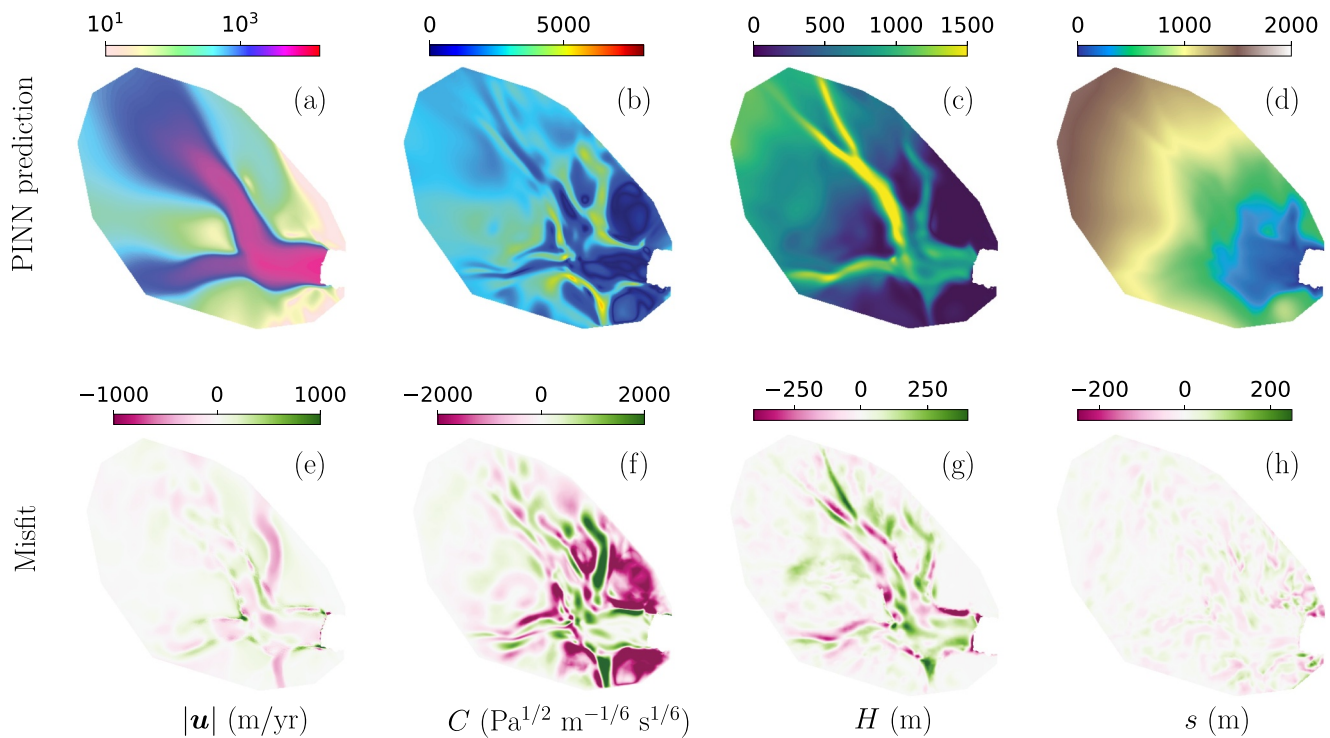


Figure 8. (a)–(d) Predictions of the PINN inferring ice thickness and basal friction coefficient using ice velocity $\hat{\mathbf{u}}$, surface elevation $\hat{s}$, and flight track data $\hat{H}$ (as in Figure 3) in the training procedure. (e)–(h) Corresponding misfits between the predictions and their corresponding reference data in Figure 1. The color map in (a) is on a logarithmic scale, and the rest of subplots are on linear scales.

data and the underlying physics. This approach considerably simplifies the code development process compared to conventional numerical methods.

The training data serves as a foundational basis for the PINN to learn from, while the governing PDEs act as soft constraints, guiding the estimation of unknown parameters in the system. The PDE constraints provide additional relationships between all the variables involved, and act as a regularizing operator, particularly for the terms involving gradients (Karniadakis et al., 2021). For variables with available data during training, the PINN acts as a smooth interpolator. On the contrary, for variables without data, the PINN behaves as a solver, effectively satisfying the soft constraints imposed by the governing PDEs.

This unified framework, illustrated in Figure 2, offers flexibility for incorporating new variables and physics into the system. The addition of new variables involves introducing additional outputs from the neural network, while the inclusion of new physics requires adding an extra term to the loss function to assess the residual of the new governing equation. Furthermore, this framework can be readily extended to time-dependent problems by introducing time as an input variable to the neural network and adding time-dependent PDEs, such as the conservation of mass, in the loss function (Raissi et al., 2020; Riel et al., 2021).

4.2. Model Performance

The results presented in Section 3 show that the PINN is able to produce solutions for all four experiments with reasonable accuracy. The relative errors in all cases are below 5%. For the variables learned directly from the data (e.g., surface topography), the errors are evenly distributed throughout the entire domain, with some larger errors associated with steep changes in the data. On the other hand, when inferring solutions based on the PDE constraints, the errors depend on the location. In the forward run, the error in the predicted velocity ($\mathbf{u}$) is larger in the high-velocity region, especially close to the shear margin and ice front. Similarly, when inferring the ice thickness, larger errors are observed in the fast-flowing region and at sharp transition zones in the ice thickness.

However, in the basal friction inverse experiment, the errors in the inferred parameter C show a different pattern. Larger errors are observed in the slower-moving regions with thinner ice. This difference is likely due to the MSE of the velocity misfit that tends to prioritize the fast-flowing regions over the slower-moving ones, and the errors in the slow moving regions have limited impact on the overall accuracy of the solutions. Considering the extensive range of ice velocities observed in Helheim Glacier, spanning five orders of magnitude, the solution for C in the inverse problem detailed in Figure 6 exhibits more substantial errors in regions characterized by low ice velocities (less than 10 m/yr) compared to those in the fast-moving regions. This situation is not unique to this approach, and traditional numerical methods also face this problem when inverting for the basal friction coefficient. The solution to tackle this problem in conventional inverse methods is to complement the L^2 norm for the velocity misfit with a logarithmic norm of the velocity that will place more weight on the slow-moving regions (Morlighem et al., 2010). Additionally, the SSA equations (Equation 1) may not be a good approximation of the ice dynamics in slow-moving regions where vertical shear cannot be neglected. Since the flow velocity does not strongly depend on the friction coefficient in these slow-moving regions, the error is expected to be high.

To assess the quality of the PINN solutions, we conduct two additional forward problem experiments using friction coefficient C inverted from ISSM (Figure 1b) and PINN (Figure 6b), respectively. These experiments are solved by the Finite Element method using ISSM on a triangular mesh at a resolution of 200 m (23,484 degrees of freedom). The velocity magnitude solutions are shown in Figures 9a and 9b and the misfits between these solutions and the “true” solution (as in Figure 1a) are shown in Figures 9c and 9d, where the RMSE of the misfit are 106.58 m/yr and 474.14 m/yr, respectively. We find similar patterns of the misfit at fast-flowing regions and along the shear margins, as observed in the forward solution from PINN in Figure 5e. While the ISSM forward solution (Figure 9a) exhibits slightly better accuracy, the difference is not substantial. Notably, in the experiment using C from PINN in Figure 9b, unrealistically large velocities (over 5,000 m/yr) are observed in the thin ice region north of the ice front. This area is sparsely covered by ice in reality, as depicted in the flight track data in Figure 3. As previously discussed, incorporating additional terms in the loss function or imposing additional physical constraints may be necessary to improve the representation of the slow velocity region.

The experiment in Section 3.4 is particularly noteworthy as it demonstrates the capability of PINNs to infer ice thickness using momentum conservation (SSA). To our knowledge, this has not yet been achieved using conventional numerical methods. Together with the other experiments presented in Section 3, these results underscore the inherent flexibility of the PINN framework. Ideally, this approach can be applied to infer any variable within the governing equation in a similar manner.

It is also important to note that in the dual inversion problem (Section 3.5), only the ice velocity, surface elevation, and ice thickness along flight lines are incorporated into the training procedure and exposed to the PINN. The governing equation in the PINN is based on momentum conservation rather than mass conservation, which is the principle employed by BedMachine for inferring ice thickness. Consequently, discrepancies between the PINN predictions and the reference ice thickness from BedMachine are expected, constituting the likely primary reason for the observed misfit in Figure 8g. Furthermore, considering that the reference friction coefficient is inferred from ISSM using the ice thickness from BedMachine, differences are expected, particularly in regions where the two ice thickness data sets diverge.

4.3. Regularization

Previous studies (e.g., Brinkerhoff, 2022; He et al., 2023; Jouvét, 2022; Jouvét & Cordonnier, 2023; Riel & Minchew, 2023; Riel et al., 2021) enforce spatial smoothness by adding a term to the loss function that penalizes high gradients in basal drag calculated from the PDEs (Riel & Minchew, 2023; Riel et al., 2021), or through other regularization terms (generally involving the gradient of the friction coefficients). This regularization helps the inversion converge faster and makes the loss function more convex (Brinkerhoff, 2022; He et al., 2023; Jouvét, 2022; Jouvét & Cordonnier, 2023). Our approach directly computes the residual of the entire PDEs without the need to augment the loss function with extra regularization terms.

The inherent smoothness in the solutions generated by the PINNs can be attributed to several factors within the framework (Chen et al., 2021; Seo, 2024). These factors include the neural network architecture, the choice of activation functions, the formulation of the loss function, and the optimization method, typically stochastic gradient descent (SGD) (Poggio et al., 2017). Specifically, the relatively shallow architecture of the neural

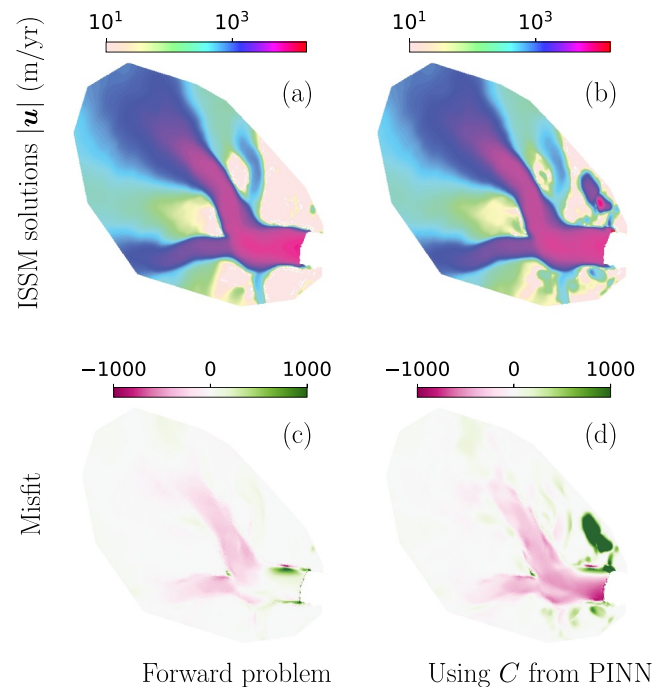


Figure 9. (a) The numerical solution of the ice velocity u using ISSM to solve the same forward problem as in Figure 5. (b) The same forward problem solved by ISSM as in (a), except the friction coefficient C is inverted from PINN as in Figure 6. (c)–(d) Corresponding misfits between the ISSM solutions and the reference velocity data in Figure 1a. The color maps in (a)–(b) are on a logarithmic scale, and the rest of subplots are on linear scales.

network, coupled with the differentiable hyperbolic tangent activation function, provides a smooth representation of the PINN output. Moreover, the loss function of the PINN consists of both the PDE residual and the data misfit, acting as a soft constraint and implicit regularization mechanism (Bolibar et al., 2023). Importantly, the prior knowledge from the PDEs introduces essential structure that effectively regularizes the minimization procedure during neural network training (Raissi et al., 2020). Finally, SGD tends to converge to smooth solutions by selecting global minima and mitigating overfitting risks (Jin et al., 2020; Lu, Pestourie, et al., 2021; Poggio et al., 2017).

Unlike previous studies (e.g., Brinkerhoff, 2022; He et al., 2023; Jouvét, 2022; Jouvét & Cordonnier, 2023; Riel & Minchew, 2023; Riel et al., 2021), our approach directly utilizes the residual of the PDEs in the loss function, ensuring that the loss function converges to zero if the output of the PINN converges to the true solution that fully satisfies the PDEs.

4.4. Limitations

While our study highlights the capabilities of PINNs in ice sheet modeling, certain limitations should be acknowledged. For the forward model, which is mathematically well-posed, traditional grid-based solvers clearly outperform PINNs (Karniadakis et al., 2021). For instance, while training the PINN for a forward problem (Section 3.2) requires approximately 10 hr on one GPU, the same problem can be solved within minutes using established solvers like ISSM with 40 CPUs for a mesh of approximately 20,000 elements. Another challenge is that the governing equations are imposed as soft constraints in the loss function and compete with the data misfit during the optimization, causing occasional non-convergence. Furthermore, it is well known that SSA serves as a reliable approximation for ice dynamics in fast-flowing regions but its assumptions break down in the interior of the ice sheet. Generalizing this approach to the entire Greenland Ice Sheet may necessitate the use of alternative physics or a combination of different physics to infer ice thickness, for example,

Future research directions will need to address the identified limitations and further enhance the application of PINNs in ice sheet modeling. To enhance its efficiency, the training process could be optimized and potentially integrate parallel computing strategies for faster execution. The handling of PDEs as soft constraints in the PINN

framework could be revised in order to mitigate convergence issues. Finally, improving the accuracy of the ice sheet interior will involve alternative physics or hybrid approaches that better capture the complexities of ice dynamics in slow-moving regions. These steps will collectively contribute to advancing the robustness, accuracy, and computational efficiency of PINNs for comprehensive ice sheet modeling.

5. Conclusion

This study explores several applications of PINNs in typical problems of ice sheet modeling. In contrast to traditional numerical methods, we utilize PINNs to construct a unified framework for both forward and inverse modeling. The inherent adaptability of PINNs is particularly easy to use and expand, enabling the inclusion of new physical parameters into the numerical model. This approach offers a promising avenue for enhancing the flexibility of ice sheet models and data assimilation, beyond the traditional categories of forward or inverse problems.

The dual inversion case presented in this study further demonstrates the ability of PINNs to simultaneously infer the basal friction coefficient and fill in gaps in partially sparse ice thickness observations. Serving as complements rather than replacements to traditional numerical methods, PINNs offer simplicity in implementation, rapid development, and easy testing of novel concepts, thereby introducing new perspectives to data-driven ice sheet modeling. This study suggests the potential of PINNs in improving our understanding of ice dynamics and eventually enabling more accurate predictions of future sea-level rise in glaciology and climate science.

Data Availability Statement

The data and the code of the simulations are available at Cheng (2024). ISSM is open source and available at ISSM Team (2023).

Acknowledgments

This work was supported by the Heising Simons Foundation Grants 2019-1161 and 2021-3059, and the National Science Foundation 2118285 and 2147601.

References

- Åkesson, H., Morlighem, M., O'Regan, M., & Jakobsson, M. (2021). Future projections of Petermann glacier under ocean warming depend strongly on friction law. *Journal of Geophysical Research: Earth Surface*, 126(6), e2020JF005921. <https://doi.org/10.1029/2020JF005921>
- Barnes, J. M., dos Santos, T. D., Goldberg, D., Gudmundsson, G. H., Morlighem, M., & De Rydt, J. (2021). The transferability of adjoint inversion products between different ice flow models. *The Cryosphere*, 15(4), 1975–2000. <https://doi.org/10.5194/tc-15-1975-2021>
- Bolibar, J., Sapienza, F., Maussion, F., Lguensat, R., Wouters, B., & Pérez, F. (2023). Universal differential equations for glacier ice flow modelling. *Geoscientific Model Development*, 16(22), 6671–6687. <https://doi.org/10.5194/gmd-16-6671-2023>
- Brinkerhoff, D. J. (2022). Variational inference at glacier scale. *Journal of Computational Physics*, 459, 111095. <https://doi.org/10.1016/j.jcp.2022.111095>
- Bronxex, J., Gagliardini, O., Gillet-Chaulet, F., & Durand, G. (2017). Sensitivity of grounding line dynamics to the choice of the friction law. *Journal of Glaciology*, 63(241), 854–866. <https://doi.org/10.1017/jog.2017.51>
- Bronxex, J., Gillet-Chaulet, F., & Gagliardini, O. (2019). Sensitivity of centennial mass loss projections of the Amundsen basin to the friction law. *The Cryosphere*, 13(1), 177–195. <https://doi.org/10.5194/tc-13-177-2019>
- Budd, W. F., Keage, P. L., & Blundy, N. A. (1979). Empirical studies of ice sliding. *Journal of Glaciology*, 23(89), 157–170. <https://doi.org/10.3189/s0022143000029804>
- Bueler, E., & Brown, J. (2009). Shallow shelf approximation as a “sliding law” in a thermomechanically coupled ice sheet model. *Journal of Geophysical Research*, 114(F3), 1–21. <https://doi.org/10.1029/2008JF001179>
- Chen, Z., Liu, Y., & Sun, H. (2021). Physics-informed learning of governing equations from scarce data. *Nature Communications*, 12(1), 6136. <https://doi.org/10.1038/s41467-021-26434-1>
- Cheng, G. (2024). Forward and Inverse modeling of ice sheet flow using physics-informed neural network [Software]. *Zenodo*. <https://doi.org/10.5281/zenodo.10627691>
- Cheng, G., & Lotstedt, P. (2020). Parameter sensitivity analysis of dynamic ice sheet models - Numerical computations. *The Cryosphere*, 14(2), 673–691. <https://doi.org/10.5194/tc-14-673-2020>
- Colinge, J., & Blatter, H. (1998). Stress and velocity fields in glaciers: Part I. Finite-Difference schemes for higher-order glacier models. *Journal of Glaciology*, 44(148), 448–456. <https://doi.org/10.1017/s0022143000001969>
- Cuffey, K. M., & Paterson, W. S. B. (2010). In B. Heinemann (Ed.), *The physics of glaciers* (4th ed.). Elsevier.
- Durand, G., Gagliardini, O., Zwinger, T., Le Meur, E., & Hindmarsh, R. C. A. (2009). Full Stokes modeling of marine ice sheets: Influence of the grid size. *Annals of Glaciology*, 50(52), 109–114. <https://doi.org/10.3189/172756409789624283>
- Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the Coupled Model Inter-comparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, 9(5), 1937–1958. <https://doi.org/10.5194/gmd-9-1937-2016>
- Frederikse, T., Landerer, F., Caron, L., Adhikari, S., Parkes, D., Humphrey, V. W., et al. (2020). The causes of sea-level rise since 1900. *Nature*, 584(7821), 393–397. <https://doi.org/10.1038/s41586-020-2591-3>
- Gagliardini, O., Cohen, D., Raback, P., & Zwinger, T. (2007). Finite-element modeling of subglacial cavities and related friction law. *Journal of Geophysical Research*, 112(F2), 1–11. <https://doi.org/10.1029/2006JF000576>
- Glen, J. W. (1958). The flow law of ice: A discussion of the assumptions made in glacier theory, their experimental foundations and consequences. *IASH Publ*, 47, 171–183.

- Goldberg, D. N., & Sergienko, O. V. (2011). Data assimilation using a hybrid ice flow model. *The Cryosphere*, 5(2), 315–327. <https://doi.org/10.5194/tc-5-315-2011>
- Griewank, A., Juedes, D., & Utke, J. (1996). ADOL-C: A package for the automatic differentiation of algorithms written in C/C++. *ACM Transactions on Mathematical Software*, 22(2), 131–167. <https://doi.org/10.1145/229473.229474>
- Haran, T., Bohlander, J., Scambos, T., Painter, T., & Fahnestock, M. (2018). *MEaSUREs MODIS mosaic of Greenland (MOG) 2005, 2010, and 2015 image maps, version 2*. NASA National Snow and Ice Data Center Distributed Active Archive Center. <https://doi.org/10.5067/9ZO79PHOTYE5>
- He, Q., Perego, M., Howard, A. A., Karniadakis, G. E., & Stinis, P. (2023). A hybrid deep neural operator/finite element method for ice-sheet modeling. *Journal of Computational Physics*, 492, 112428. <https://doi.org/10.1016/j.jcp.2023.112428>
- Howat, I. M., Negrete, A., & Smith, B. E. (2014). The Greenland Ice Mapping Project (GIMP) land classification and surface elevation datasets. *The Cryosphere*, 8(4), 1509–1518. <https://doi.org/10.5194/tc-8-1509-2014>
- ISSM Team. (2023). Ice-sheet and sea-level system model source code, v4.23 r27696 [Software]. *Zenodo*. <https://doi.org/10.5281/zenodo.7850841>
- Iwasaki, Y., & Lai, C.-Y. (2023). One-dimensional ice shelf hardness inversion: Clustering behavior and collocation resampling in physics-informed neural networks. *Journal of Computational Physics*, 492, 112435. <https://doi.org/10.1016/j.jcp.2023.112435>
- Jin, P., Lu, L., Tang, Y., & Karniadakis, G. E. (2020). Quantifying the generalization error in deep learning in terms of data distribution and neural network smoothness. *Neural Networks*, 130, 85–99. <https://doi.org/10.1016/j.neunet.2020.06.024>
- Joughin, I., Smith, B. E., & Howat, I. M. (2018). A complete map of Greenland ice velocity derived from satellite data collected over 20 years. *Journal of Glaciology*, 64(243), 1–11. <https://doi.org/10.1017/jog.2017.73>
- Jouvet, G. (2022). Inversion of a Stokes glacier flow model emulated by deep learning. *Journal of Glaciology*, 69(273), 1–14. <https://doi.org/10.1017/jog.2022.41>
- Jouvet, G., & Cordonnier, G. (2023). Ice-flow model emulator based on physics-informed deep learning. *Journal of Glaciology*, 1–15. <https://doi.org/10.1017/jog.2023.73>
- Karniadakis, G. E., Kevrekidis, I. G., Lu, L., Perdikaris, P., Wang, S., & Yang, L. (2021). Physics-informed machine learning. *Nature Reviews Physics*, 3(6), 422–440. <https://doi.org/10.1038/s42254-021-00314-5>
- Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.
- Larour, E., Seroussi, H., Morlighem, M., & Rignot, E. (2012). Continental scale, high order, high spatial resolution, ice sheet modeling using the Ice Sheet System Model (ISSM). *Journal of Geophysical Research*, 117(F01022), 1–20. <https://doi.org/10.1029/2011JF002140>
- Lu, L., Meng, X., Mao, Z., & Karniadakis, G. E. (2021). DeepXDE: A deep learning library for solving differential equations. *SIAM Review*, 63(1), 208–228. <https://doi.org/10.1137/19M1274067>
- Lu, L., Pestourie, R., Yao, W., Wang, Z., Verdugo, F., & Johnson, S. G. (2021). Physics-informed neural networks with hard constraints for inverse design. *SIAM Journal on Scientific Computing*, 43(6), B1105–B1132. <https://doi.org/10.1137/21M1397908>
- MacAyeal, D. (1989). Large-scale ice flow over a viscous basal sediment: Theory and application to Ice Stream B, Antarctica. *Journal of Geophysical Research*, 94(B4), 4071–4087. <https://doi.org/10.1029/jb094ib04p04071>
- MacAyeal, D. (2002). Notes on ice stream inverse code finite element solution for ice stream equations. (copy of a document printed in 2002).
- Mangeney, A., & Califano, F. (1998). The shallow ice approximation for anisotropic ice: Formulation and limits. *Journal of Geophysical Research*, 103(B1), 691–705. <https://doi.org/10.1029/97jb02539>
- Morlighem, M., & Goldberg, D. (2023). Data assimilation in glaciology. In *Applications of data assimilation and inverse problems in the Earth sciences*. Cambridge University Press.
- Morlighem, M., Rignot, E., Binder, T., Blankenship, D., Drews, R., Eagles, G., et al. (2020). Deep glacial troughs and stabilizing ridges unveiled beneath the margins of the Antarctic ice sheet. *Nature Geoscience*, 13(2), 132–137. <https://doi.org/10.1038/s41561-019-0510-8>
- Morlighem, M., Rignot, E., Seroussi, H., Larour, E., Ben Dhia, H., & Aubry, D. (2010). Spatial patterns of basal drag inferred using control methods from a full-Stokes and simpler models for Pine Island Glacier, West Antarctica. *Geophysical Research Letters*, 37(L14502), 1–6. <https://doi.org/10.1029/2010GL043853>
- Morlighem, M., Seroussi, H., Larour, E., & Rignot, E. (2013). Inversion of basal friction in Antarctica using exact and incomplete adjoints of a higher-order model. *Journal of Geophysical Research*, 118(3), 1746–1753. <https://doi.org/10.1002/jgrf.20125>
- Morlighem, M., Williams, C. N., Rignot, E., An, L., Arndt, J. E., Bamber, J. L., et al. (2017). BedMachine v3: Complete bed topography and ocean bathymetry mapping of Greenland from multi-beam echo sounding combined with mass conservation. *Geophysical Research Letters*, 44(21), 11051–11061. <https://doi.org/10.1002/2017GL074954>
- Mouginot, J., Rignot, E., Björk, A. A., van den Broeke, M., Millan, R., Morlighem, M., et al. (2019). Forty-six years of Greenland ice sheet mass balance from 1972 to 2018. *Proceedings of the National Academy of Sciences*, 116(19), 9239–9244. <https://doi.org/10.1073/pnas.1904242116>
- Nowicki, S. M. J., Payne, A. J., Larour, E., Seroussi, H., Goelzer, H., Lipscomb, W. H., et al. (2016). Ice Sheet Model Intercomparison Project (ISMIP6) contribution to CMIP6. *Geoscientific Model Development*, 9(12), 4521–4545. <https://doi.org/10.5194/gmd-9-4521-2016>
- Perego, M., Price, S., & Stadler, G. (2014). Optimal initial conditions for coupling ice sheet models to Earth system models. *Journal of Geophysical Research: Earth Surface*, 119(9), 1–24. <https://doi.org/10.1002/2014JF003181>. Received
- Petra, N., Zhu, H., Stadler, G., Hughes, T. J. R., & Ghattas, O. (2012). An inexact Gauss-Newton method for inversion of basal sliding and rheology parameters in a nonlinear Stokes ice sheet model. *Journal of Glaciology*, 58(211), 889–903. <https://doi.org/10.3189/2012JoG11J182>
- Poggio, T., Kawaguchi, K., Liao, Q., Miranda, B., Rosasco, L., Boix, X., et al. (2017). Theory of deep learning III: Explaining the non-overfitting puzzle. arXiv. <https://doi.org/10.48550/arxiv.1801.00173>
- Rahnemoonfar, M., & Koo, Y. (2024). Graph neural networks as fast and high-fidelity emulators for finite-element ice sheet modeling. arXiv. <https://doi.org/10.48550/arxiv.2402.05291>
- Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378, 686–707. <https://doi.org/10.1016/j.jcp.2018.10.045>
- Raissi, M., Yazdani, A., & Karniadakis, G. E. (2020). Hidden fluid mechanics: Learning velocity and pressure fields from flow visualizations. *Science*, 367(6481), 1026–1030. <https://doi.org/10.1126/science.aaw4741>
- Riel, B., & Minchew, B. (2023). Variational inference of ice shelf rheology with physics-informed machine learning. *Journal of Glaciology*, 69(277), 1167–1186. <https://doi.org/10.1017/jog.2023.8>
- Riel, B., Minchew, B., & Bischoff, T. (2021). Data-driven inference of the Sechanics of slip Slong glacier beds using physics-informed neural networks: Case study on Rutford Ice Stream, Antarctica. *Journal of Advances in Modeling Earth Systems*, 13(11), e2021MS002621. <https://doi.org/10.1029/2021MS002621>

- Rignot, E., Mouginot, J., Scheuchl, B., van den Broeke, M., van Wessem, M. J., & Morlighem, M. (2019). Four decades of Antarctic Ice Sheet mass balance from 1979–2017. *Proceedings of the National Academy of Sciences*, 116(4), 1095–1103. <https://doi.org/10.1073/pnas.1812883116>
- Seo, J. (2024). Solving real-world optimization tasks using physics-informed neural computing. *Scientific Reports*, 14(1), 202. <https://doi.org/10.1038/s41598-023-49977-3>
- Tarantola, A. (2005). *Inverse problem theory*. SIAM.
- Teisberg, T. O., Schroeder, D. M., & MacKie, E. J. (2021). A machine learning approach to mass-conserving ice thickness interpolation. In *2021 IEEE international geoscience and remote sensing symposium IGARSS* (pp. 8664–8667).
- Vogel, C. R. (2002). *Computational methods for inverse problems*. Society for Industrial and Applied Mathematics.
- Wang, Y., Lai, C.-Y., & Cowen-Breen, C. (2022). Discovering the rheology of Antarctic ice shelves via physics-informed deep learning. <https://doi.org/10.21203/rs.3.rs-2135795/v1>
- Weertman, J. (1957). On the sliding of glaciers. *Journal of Glaciology*, 3(21), 33–38. <https://doi.org/10.3189/s0022143000024709>
- Yu, H., Rignot, E., Seroussi, H., & Morlighem, M. (2018). Retreat of Thwaites Glacier, West Antarctica, over the next 100 years using various ice flow models, ice shelf melt scenarios and basal friction laws. *The Cryosphere*, 12(12), 3861–3876. <https://doi.org/10.5194/tc-12-3861-2018>