A MACHINE LEARNING APPROACH TO MASS-CONSERVING ICE THICKNESS INTERPOLATION

Thomas O. Teisberg[1], *Dustin M. Schroeder*[1,2], *Emma J. MacKie*[2]

1) Department of Electrical Engineering, Stanford University,
2) Department of Geophysics, Stanford University

ABSTRACT

The subglacial topography of the Earth's ice sheets is a critical input to models of the evolution of ice sheets and sea level rise. Direct measurements of ice thickness, however, are sparse, necessitating techniques for interpolating these measurements. One class of interpolation methods enforces physical constraints to transform the problem into an inversion. A challenge with these approaches is that multiple unknown parameters must be solved for simultaneously. We introduce a new numerical approach to solving for mass conservation-constrained ice thickness maps. This technique, based on a physics-informed neural network, allows for the flexible incorporation of a range of soft constraints. In the future, this could enable simultaneous estimation of ice velocity, bed topography, and sliding parameters.

Index Terms— interpolation, ice-penetrating radar, physics-informed neural networks

1. INTRODUCTION

One of the fundamental challenges of glaciology is that it is difficult to observe processes and conditions occurring kilometers below the ice surface, but understanding such subsurface processes is critical to projecting the evolution of glaciers and ice sheets. Detailed subsurface measurements can be made at specific locations through labor-intensive field work to help understand the physical processes at play, but modelling entire glaciers almost always relies on inversions to estimate the subsurface conditions from surface measurements, sometimes with spatially sparse subsurface measurements.

In practice, conditions under the ice surface are described by multiple inter-related phenomena and characteristics. Inversions for a particular parameter (for instance topography, basal friction, or velocity) must make assumptions about other parameters to avoid making the problem hopelessly under-determined.

As ice sheet models become more sophisticated and our understanding of ice dynamics grows, it is increasingly important to be able to simultaneously invert for multiple parameters, while placing physically-based constraints on different dimensions of the problem.

As an initial step towards this approach, we propose a machine learning-based method for mass conservation-constrained subglacial topography inversion that allows for flexibly combining physical constraints and regularizations.

In Section 2, we review mass conservation-constrained ice thickness interpolation methods. We then demonstrate our proposed machine learning approach to this problem in Section 3. Section 4 discusses the extension of this method to the two-dimensional case and the introduction of simultaneous constraints on ice thickness and velocity. Finally, we discuss the potential we see in this approach in Section 5.

2. MASS CONSERVATION ICE THICKNESS INTERPOLATION

Maps of ice thickness over Earth's ice sheets are key inputs into models used to predict the contribution of these ice sheets to sea level rise. While accurate surface elevation measurements are widely available from satellite date, measurements of the topography below ice sheets are typically derived from airborne radar sounders. These instruments use the two-way travel time of radio waves through the ice to measure a profile of ice thickness along the flight path. As a result of the cost and logistical complexity of this process, ice thickness measurements are sparse in comparison to surface measurements. Maps of ice thickness must be made by interpolating between these sparse lines of measurements.

Kriging interpolation has been widely used to produce these maps [1], however Kriging interpolations produce maps that do not satisfy the conservation of mass when accounting for reasonable bounds on the ice flow velocity [2]. More recently, mass conservation interpolation methods have been introduced for fast-flowing regions of ice sheets [3][4]. This method is based on enforcing the condition that mass is conserved at every point given an estimated flow velocity. This constraint can be expressed in a two-dimensional form by using a depth-averaged ice velocity vector $\bar{v}$:

$$\nabla \cdot (h\bar{v}) = \dot{a} \tag{1}$$

Here h is the ice thickness and $\dot{a}$ is the apparent mass balance due to accumulation, ablation, and melting from the surface and bed. We make the approximation throughout this paper that ice has a constant density. Although depth-averaged ice velocity cannot easily be measured, surface velocity is available from satellite-born InSAR measurements [5] and can be used as a proxy for depth-averaged velocity.

Combining Eq. 1 with constraints on h where radar sounder ice thickness data exists has been effectively used to interpolate ice thickness in fast-flowing (e.g. $>$50 m/year) areas of ice sheets [4].

3. METHOD OVERVIEW

We apply a physics-informed neural network (PINN) [6] to predicting ice thickness and depth-averaged velocity as a function of one- or two-dimensional spatial coordinates. The feed-forward network's only input is a spatial coordinate and the only outputs are the predicted ice thickness and depth-averaged velocity vector at that point. Using automatic differentiation, we evaluate a physics-based loss function that penalizes deviations from Eq. 1 (see Section 4.2). We add additional loss terms that penalize differences between the predicted values and available measurements.

Among other advantages, our approach is entirely mesh-free and does not rely upon finite differences to compute the derivatives. As such, the network may be trained on data without making any decisions about mesh resolution and without the need to grid or otherwise interpolate the input data. Once trained, the network may be queried at any point of interest within the training domain. As a result, there is no need to select an output grid resolution.

As an illustrative example, we begin with a synthetic 1D (streamline) model. We assume that the apparent mass balance is zero throughout the domain. With this simplification, conservation of mass in one dimension simplifies to a constant product of depth-averaged velocity and ice thickness:

$$\frac{d}{dx}(h\bar{v}) = 0 \tag{2}$$

In practice, measurements are only available at specific locations. Points on the domain where velocity or thickness measurements are provided to the network are marked with dots in Figure 1. Unless a measurement exists at a location, the value is not available to the network.

Our interpolation network is trained using the sum of three categories of loss functions: radar data fit, velocity data fit, and mass conservation. The radar and velocity data fit terms rewards the predicted values for being consistent with measurements where they are available. The mass conservation term minimizes non-physical implied creation or destruction of mass that can occur in other interpolation schemes. Each term can be weighted by a factor. For the 1D case, each term is given equal weight.

Figure 1 shows the results of the network's interpolation with and without the mass conservation loss term. Without

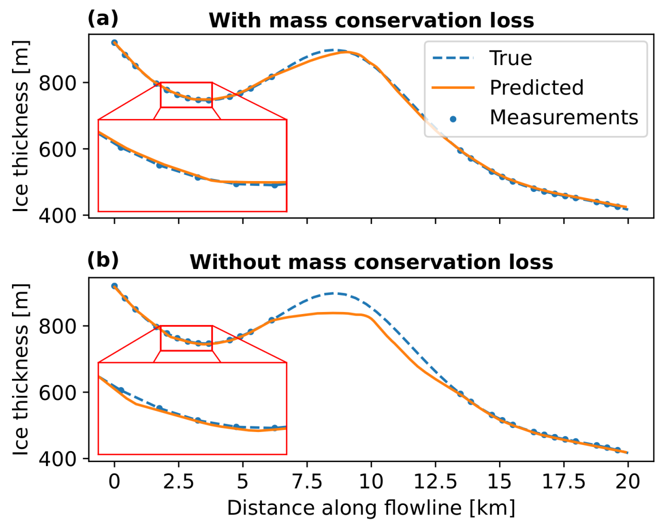

Fig. 1. (a) PINN-based interpolation results incorporating a conservation of mass constraint (b) Neural network interpolation without physics-based constraints

the mass conservation loss term, the network produces somewhat arbitrary ice thickness predictions in the gap between dense measurements. Even in the areas with dense measurements available, there are some sharp corners and other artifacts visible between data points (see red zoom rectangles in the figure). Introducing the mass conservation loss produces an interpolation result that is smoother in the dense regions and tracks the true thickness relatively well in the area without ice thickness data.

4. APPLICATION TO BYRD GLACIER

While illustrative, the one-dimensional case without noise is a trivial problem. For a more interesting two-dimensional example, we apply our method to a section of Byrd Glacier. We use radar data from surveys conducted in 2011 and 2017 by the University of Kansas [7] and between 2009 and 2012 by the University of Texas at Austin [8][9]. We use the MEaSUREs dataset for surface velocity measurements [5].

4.1. Network architecture and training

The interpolation network is a feed-forward fully connected neural network with 5 hidden layers of width 1000 and hyperbolic tangent activation functions. As in the one-dimensional case, the input is a spatial coordinate, now in two dimensions. For the purposes of the velocity loss functions described in Section 4.3, the network outputs the predicted ice thickness, the components of the predicted depth-averaged velocity vector, and the interpolated components of the surface velocity. This last output is necessary to provide a smooth interpolation of the surface velocity measurements. The network is trained

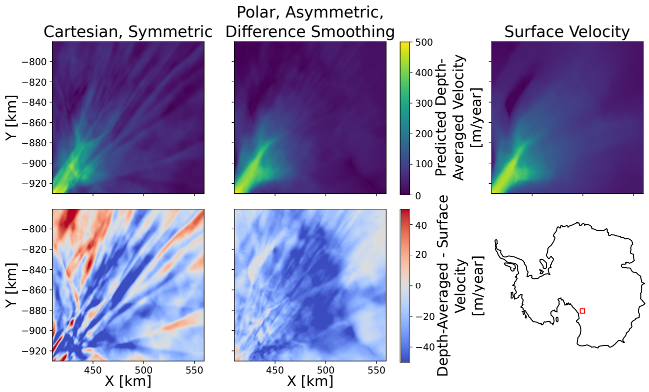

Fig. 2. Estimated depth-averaged velocities (top row) and differences between the depth-averaged and surface velocity (bottom row) for the two velocity loss functions considered

for 100 epochs using the Adam optimizer with a learning rate of 0.0002, $\beta_1 = 0.9$, and $\beta_2 = 0.99$.

The training set includes every point where a surface velocity or radar ice thickness measurement is available. In addition, the training set includes 500,000 random points sampled uniformly from the entire domain. While radar and velocity data misfit loss functions can only be evaluated at the points where those measurements are available, the mass conservation loss term and the velocity difference smoothing loss term (see Section 4.3) can be evaluated at any point. The inclusion of these randomly sampled points helps to ensure that these non-data terms are applied across the entire domain, not just where measurements exist.

Radar data constraints are enforced by a mean squared error loss between predicted thickness and measurements. Other components of the loss function are described below.

4.2. Mass conservation loss

The physics-informed loss function leverages automatic differentiation to compute the spatial derivatives of the predicted depth-averaged velocity and ice thickness. The loss term is the mean squared residual of Eq. 1.

For the purposes of the results shown in this paper, we assume the apparent mass balance $\dot{a}$ is zero throughout the domain but note that incorporating non-zero apparent mass balance into the proposed method is straight-forward.

4.3. Velocity data fit and regularization loss functions

Velocity data constraints are complicated by the difference between measured surface velocities and the depth-averaged velocities in Eq. 1. The surface velocity represents the sum of velocity due to the ice sliding along the bed and internal deformation of the ice, which generally produces a velocity profile which decreases with depth [10].

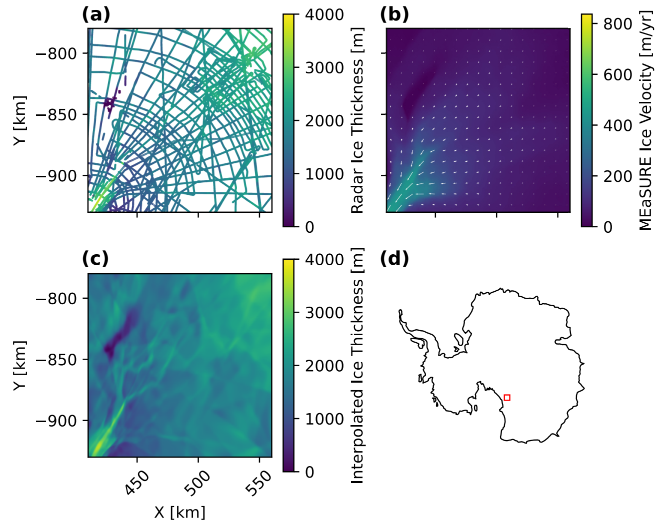

Fig. 3. Input radar data (a) and surface velocity (b) for a section of Byrd glacier (red square in (d)) used to produce the interpolated ice thickness map in (c)

We compare two options for the velocity loss functions. A straight-forward option is to allow fixed, symmetric error bounds around the surface velocity measurement. This is implemented in our method by penalizing the square error outside of the fixed bounds on $\bar{v}_x$ and $\bar{v}_y$.

This approach may produce physically-implausible results for a few reasons. First, in general, the depth-averaged velocity cannot exceed the surface velocity. Second, allowing a wide range of velocities at each point can lead to unrealistic spatial patterns. Surface and depth-averaged velocities are strongly correlated, so it would be unlikely to see a sharp change in the difference between surface and depth-averaged velocity. Finally, while the magnitude of the velocity may vary significantly with depth, the direction of the velocity vector would not be expected to change much.

An alternative formulation of this loss function is to separate out loss functions on the direction and magnitude of the depth-averaged velocity vectors. We also introduce asymmetric error bounds, allowing the magnitude of the depth-averaged velocity to be up to 50 m/year less than the surface velocity but allowing no additional error above the surface velocity magnitude. Our loss formulation penalizes the square error outside of this range of magnitudes.

Finally, we introduce a smoothing term that penalizes the mean square of the derivatives of the difference between the surface and depth-averaged velocity. This regularization term allows for sharp edges in depth-averaged velocity only where that edge is reflected in the surface velocity.

Figure 2 compares the predicted depth-averaged velocity using simple box constraints with the result using the latter set of loss functions described here.

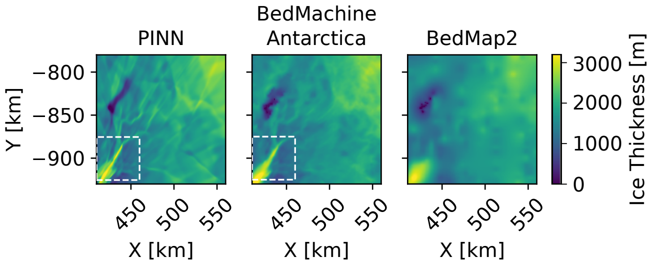

Fig. 4. Comparison between this paper's method, BedMachine Antarctica, and BedMap2 interpolated ice thickness maps. White outline shows the detail region in Figure 5.

4.4. Results

Figure 3 shows the selected domain, input surface velocity data, input radar data, and the resulting interpolated ice thickness map. Figure 4 compares our results to ice thickness maps from BedMachine Antarctica [4] and BedMap2 [11]. Our results show strong similarity to BedMachine Antarctica, as expected due to the closely-related methods.

This interpolation problem is highly under-constrained, so the choice of regularization terms is important. Figure 5 shows a comparison between our results and BedMachine Antarctica in a small part of the fast-flowing trough with the error between the interpolated result and the radar data overlaid. In this region, our results have a mean absolute error from the radar data of 70 meters, versus 131 meters for BedMachine Antarctica. The extent of this difference motivates our interest in flexibly incorporating geophysically-realistic regularization terms such as velocity difference smoothing.

5. DISCUSSION

We have demonstrated the application of an alternative numerical method for solving PDE-constrained interpolation problems to ice thickness interpolation. Our purpose is not to propose an alternative ice thickness map for Byrd Glacier but rather to introduce this method and its potential advantages.

Our proposed method offers flexibility in incorporating multiple types of physical constraints and a variety of types of data collected at the boundaries or throughout the domain. We have shown here that this flexibility allows for adding regularization terms that produce more realistic estimated depth-average velocity maps. In future work, we hope to demonstrate that this flexibility can be leveraged to incorporate additional physical terms into the interpolation, producing results which more accurately capture the multiple physical processes simultaneous occurring below the ice surface. We further believe that this approach can serve as a stepping stone for new methods in uncertainty quantification that take into account the coupling of multiple physical processes.

Source code to reproduce our methods is available at `github.com/thomasteisberg/igarss2021`

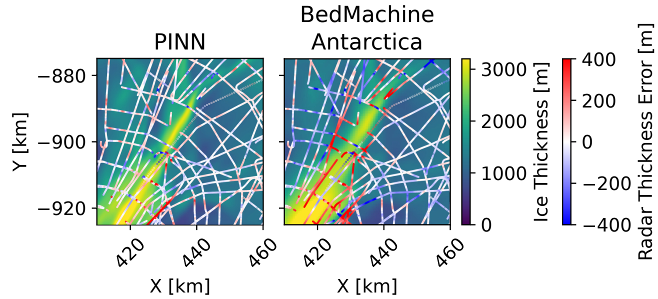

Fig. 5. In this enlarged comparison, the overlaid lines indicate locations of radar data used in this work (not necessarily the same as in [4]) with red indicating the predicted thickness exceeds the radar thickness.

6. REFERENCES

[1] J. L. Bamber, R. L. Layberry, and S. P. Gogineni, "A new ice thickness and bed data set for the Greenland ice sheet: 1. Measurement, data reduction, and errors," *Journal of Geophysical Research: Atmospheres*, vol. 106, no. D24, pp. 33 773–33 780, Dec. 2001. [Online]. Available: http://doi.wiley.com/10.1029/2001JD900054

[2] H. Seroussi *et al.*, "Ice flux divergence anomalies on 79north Glacier, Greenland," *Geophysical Research Letters*, vol. 38, no. 9, pp. n/a–n/a, May 2011. [Online]. Available: http://doi.wiley.com/10.1029/2011GL047338

[3] M. Morlighem *et al.*, "A mass conservation approach for mapping glacier ice thickness," *Geophysical Research Letters*, vol. 38, no. 19, pp. n/a–n/a, Oct. 2011. [Online]. Available: http://doi.wiley.com/10.1029/2011GL048659

[4] M. Morlighem *et al.*, "Deep glacial troughs and stabilizing ridges unveiled beneath the margins of the Antarctic ice sheet," *Nature Geoscience*, Dec. 2019. [Online]. Available: http://www.nature.com/articles/s41561-019-0510-8

[5] J. Mouginot, E. Rignot, and B. Scheuchl, "Continent-Wide, Interferometric SAR Phase, Mapping of Antarctic Ice Velocity," *Geophysical Research Letters*, p. 9, 2019.

[6] M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Physics Informed Deep Learning (Part I): Data-driven Solutions of Nonlinear Partial Differential Equations," *arXiv:1711.10561*, Nov. 2017, arXiv: 1711.10561. [Online]. Available: http://arxiv.org/abs/1711.10561

[7] CReSIS, "Radar Depth Sounder Data," 2020, lawrence, Kansas, USA. [Online]. Available: http://data.cresis.ku.edu/

[8] D. D. Blankenship *et al.*, "IceBridge HiCARS 1 L2 geolocated ice thickness, version 1," Boulder, Colorado, USA. [Online]. Available: https://doi.org/10.5067/F5FGUT9F5089

[9] D. D. Blankenship *et al.*, "IceBridge HiCARS 2 L2 geolocated ice thickness, version 1," Boulder, Colorado, USA. [Online]. Available: https://doi.org/10.5067/9EBR2T0VXUDG

[10] K. Cuffey and W. S. B. Paterson, "The Flow of Ice Masses," in *The physics of glaciers*, 4th ed. Burlington, MA: Butterworth-Heinemann/Elsevier, 2010, pp. 285–398, oCLC: ocn488732494.

[11] P. Fretwell *et al.*, "Bedmap2: improved ice bed, surface and thickness datasets for Antarctica," *The Cryosphere*, vol. 7, no. 1, pp. 375–393, Feb. 2013. [Online]. Available: https://www.the-cryosphere.net/7/375/2013/

We acknowledge the use of data and/or data products from CReSIS generated with support from the University of Kansas, NASA Operation IceBridge grant NNX16AH54G, NSF grants ACI-1443054, OPP-1739003, and IIS-1838230, Lilly Endowment Incorporated, and Indiana METACyt Initiative.