

# DIFFICE-jax: Differentiable neural-network solver for data assimilation of ice shelves in JAX

**Yongji Wang**  <sup>1,2</sup>¶ and **Ching-Yao Lai**  <sup>1</sup>

<sup>1</sup> Department of Geophysics, Stanford University, USA <sup>2</sup> Department of Mathematics, New York University, USA ¶ Corresponding author

DOI: [10.21105/joss.07254](https://doi.org/10.21105/joss.07254)

## Software

- [Review](#) ↗
- [Repository](#) ↗
- [Archive](#) ↗

---

Editor: [Anjali Sandip](#) 

Reviewers:

- [@daniel-cheng](#)
- [@RahulSundar](#)

Submitted: 09 July 2024

Published: 01 May 2025

License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC BY 4.0](#)).

## Summary

The flow of Antarctic ice shelves is controlled by their viscosity structure, which cannot be directly measured at the continental scale. Misrepresenting viscosity in ice-dynamics simulations can lead to imprecise forecasts of ice sheet mass loss into the oceans and its consequential impact on global sea-level rise. With the continent-wide remote-sensing data available over the past decades, the viscosity of the ice shelves can be inferred by solving an inverse problem. We present DIFFICE\_jax: a DIFFerentiable solver using physics-informed neural networks (PINNs) ([Raissi et al., 2019](#)) for data assimilation and inverse modeling of ICE shelves written in JAX. This Python package converts discretized remote-sensing data into meshless and differentiable functions, and infers the viscosity profile by directly solving the Shallow Shelf Approximation (SSA) equations for ice shelves. The inversion algorithm is implemented in JAX ([Bradbury et al., 2018](#)). The DIFFICE\_jax package includes several advanced features beyond vanilla PINNs algorithms, including collocation points resampling, non-dimensionalization of data and equations, extended-PINNs (XPINNs) ([Jagtap & Karniadakis, 2020](#)), and viscosity exponential scaling function, which are essential for accurate inversion. The package is designed to be user-friendly and accessible for beginners. The GitHub repository also provides tutorial examples with Colab notebooks for users at different levels to reproduce the results and modify the code for their specific problems of interest.

## Statement of need

One of the unsolved questions in the cryosphere, or the broader field of geophysics, is how to uncover the hidden physical properties of various geophysical flows on Earth, such as ice-shelf flows ([Millstein et al., 2022](#); [Ranganathan & Minchew, 2024](#)). Ice shelves are the floating extensions of grounded ice sheet that play a critical role in slowing ice discharging into the ocean, mitigating the global sea level rise. A critical physical properties required to predict ice dynamics is the effective ice viscosity. However, continent-wide in-situ measurement of the viscosity is challenging if not impossible. Instead, with high-resolution remote-sensing data available for both ice-shelf velocity ( $u, v$ ) and thickness  $h$  across Antarctica, effective viscosity  $\mu$  can be inferred via solving an inverse problem constrained by the SSA equations ([Douglas R. MacAyeal, 1989](#)).

Conventional method for solving inverse problems in glaciology includes the control method, or called the adjoint method ([Barnes et al., 2021](#); [Fürst et al., 2015](#); [Goldberg & Sergienko, 2011](#); [Douglas R. MacAyeal, 1993](#); [Mathieu Morlighem et al., 2010](#); [M. Morlighem et al., 2013](#); [Perego et al., 2014](#); [Vieli & Payne, 2003](#)), which is a PDE-constraint optimization method that requires deriving extra adjoint equations ([Douglas R. MacAyeal, 1993](#); [M. Morlighem et al., 2013](#)). Data used in inverse problems often contains noise. Thus, the adjoint method requires additional regularization techniques to prevent error propagation. In contrast, neural networks

(NN) can inherently de-noise the data while solving the inverse problem (Iwasaki & Lai, 2023; Riel & Minchew, 2023; Wang et al., 2022) without regularization terms in the cost function. That said, users should be careful about the NN's tendency to miss high-frequency signals due to their spectral biases (Rahaman et al., 2019; Xu et al., 2020) which can be mitigated by XPINNs (Jagtap & Karniadakis, 2020). Moreover, the solver itself is differentiable; the gradient of the loss function with respect to the NN parameters are calculated via automatic differentiation (AD). This avoids the tedious efforts of writing the adjoint and is particularly advantageous when exploring new PDEs like the anisotropic equations. Benefits of AD for glaciological inverse problems are also shown in other deep-learning based emulators (Jouvet et al., 2022; Jouvet, 2023; Jouvet & Cordonnier, 2023).

The DIFFICE\_jax package supports the direct use of thickness data from radar profiles available only at flight lines as PINN predicts the complete thickness profile, in addition to the viscosity profile, via physics-informed interpolation. In addition, the outputs of PINNs (velocity, thickness, and viscosity fields) are continuous functions parameterized by a fixed number of weights and biases, requiring less memory than the output of classical methods with discretized grid points when higher resolutions and precision are demanded (Wang & Lai, 2024). Finally, the DIFFICE\_jax package, along with other deep-learning-based solver for ice-dynamics (Bolibar et al., 2023; Brinkerhoff et al., 2021; Cheng et al., 2024; He et al., 2023; Iwasaki & Lai, 2023; Jouvet et al., 2022; Jouvet, 2023; Jouvet & Cordonnier, 2023; Riel et al., 2021; Riel & Minchew, 2023; Wang et al., 2022), leverages GPUs for computational speedup.

Note that off-the-shelf PINNs code are not adequate for accurately inferring effective viscosity  $\mu$  from the governing equations and large-scale real-world data. Advanced algorithm features summarized below are all essential for the success of viscosity inversion via PINNs. The package provides comprehensive details of the key algorithms involved, including comments and examples, enabling users to bypass substantial hyper-parameter tuning.

## Algorithm features

Key features of DIFFICE\_jax that go beyond off-the-shelf PINNs, and the necessity of these features to ensure the success and robustness of large-scale viscosity inference are explained below:

**(1) Data and equation normalization/non-dimensionalization:** Proper training of NNs requires both input and output of the NN to be normalized. However, the values of observational data of ice velocity and thickness differ by several order of magnitude in their original units. The DIFFICE\_jax package provides the algorithm that can automatically normalize the observational data and derive normalized SSA equations for the training.

**(2) Optimal setting of equation weight:** The cost function of PINNs involves two terms: the data loss  $\mathcal{L}_d$  and the equation loss  $\mathcal{L}_e$ . The weighting pre-factors of the data and equation loss are optimally set in the package to ensure effective convergence of training loss.

**(3) Design of NNs to enforce positive-definiteness:** Considering that the effective viscosity  $\mu$  is positive with large spatial variation within the ice shelf, we introduce the viscosity expression as

$$\mu = \exp(\text{NN}),$$

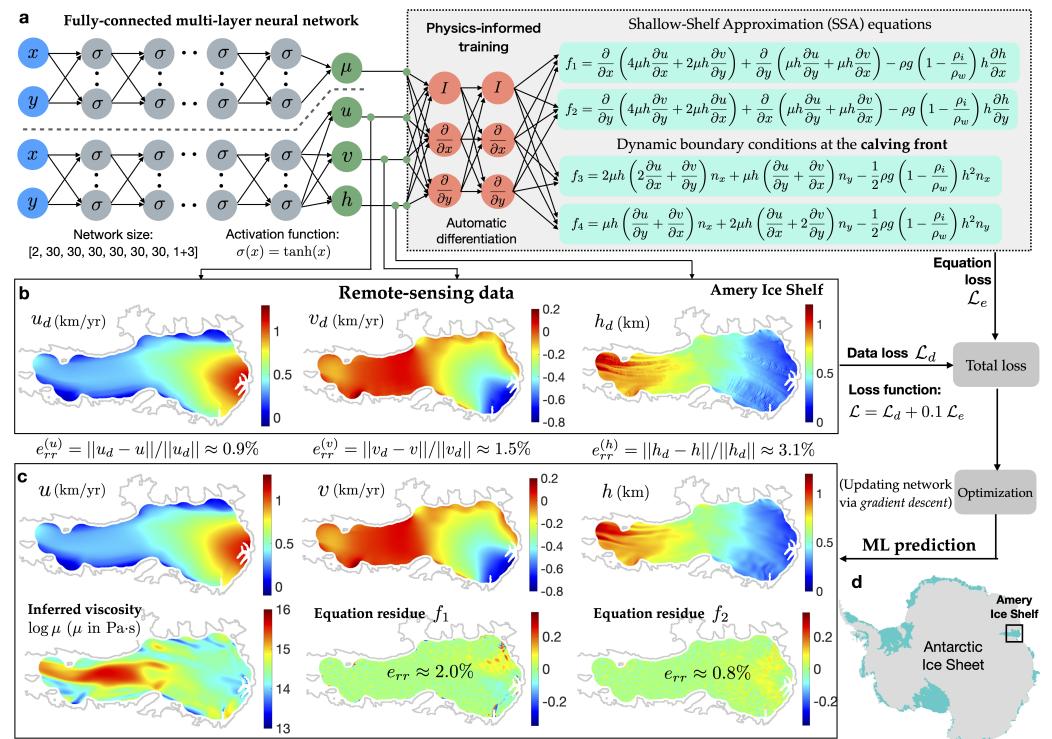
where  $\text{NN}_\mu$  is a fully-connected network for  $\mu$ . This setting ensures the positive-definiteness of the inferred viscosity and enhance the training to capture the large-varying viscosity profile with high accuracy.

**(4) Residual-based re-sampling of collocation points during training:** Due to errors and noise, PINN training with observational data often “cheats”, where the networks overfit the training data (Cowen-Breen et al., 2024; Wang et al., 2022). To prevent the issue, the DIFFICE\_jax

package uses residual-based re-sampling scheme during training (Daw et al., 2022; Lu et al., 2021), where more training data and collocation points are sampled in the area with large training residue.

**(5) Extended-PINNs (XPINNs) for large ice shelf:** Regular PINN training with a single network cannot capture the rich spatial variation of large ice shelves, such as Ross. The DIFFICE\_jax package adopts the approach of extended PINNs (XPINNs) (Jagtap & Karniadakis, 2020) for studying large ice shelves. This method divides the training domains into several sub-regions, with different networks assigned to each. Detailed description of XPINNs are provided in the doc folder of the GitHub repository.

**(6) Inversion of anisotropic viscosity:** Prior studies have shown that the viscosity of Antarctica Ice Shelves could be anisotropic (Wang et al., 2025). The DIFFICE\_jax package involves the first algorithm designed to infer anisotropic viscosity. The governing equations, associated boundary conditions, and the cost function for inferring anisotropic viscosity are described in the docs folder.



**Figure 1: PINN setup.** (a) The structure and workflow of physics-informed neural networks (PINNs) for inferring ice viscosity  $\mu$  from (b) the remote-sensing data of ice velocity  $u, v$  (Mouginot et al., 2019) and thickness  $h$  (M. Morlighem, 2020), and governing equations ( $f_1, f_2, f_3$ , and  $f_4 = 0$ ). The loss function  $\mathcal{L}$  contains two terms, the data loss  $\mathcal{L}_d$  and the equation loss  $\mathcal{L}_e$ . (c) Prediction of trained neural network for velocity ( $u, v$ ) and thickness  $h$ , which shows high agreement with the remote-sensing data with relatively error around 1-3%. The inferred viscosity well satisfies the equation with small residue values, indicating the accuracy of the inversion.

## Acknowledgements

We thank Charlie Cowen-Breen for the discussion on PINNs optimization, and Ming-Ruey Chou and Robert Clapper for Python Programming and Github preparation. We acknowledge the Stanford Doerr Discovery Grant and the Office of the Dean for Research at Princeton University for partial funding support via the Dean for Research Fund for New Ideas in the

Natural Sciences. C.-Y.L acknowledge the National Science Foundation for funding via Grant No. DMS-2245228.

## References

- Barnes, J. M., Dias dos Santos, T., 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 Discussions*, 2023, 1–26. <https://doi.org/10.5194/gmd-16-6671-2023>
- Bradbury, J., Frostig, R., Hawkins, P., Johnson, M. J., Leary, C., Maclaurin, D., Necula, G., Paszke, A., VanderPlas, J., Wanderman-Milne, S., & Zhang, Q. (2018). JAX: Composable transformations of Python+NumPy programs (Version 0.3.13). <http://github.com/google/jax>
- Brinkerhoff, D., Aschwanden, A., & Fahnestock, M. (2021). Constraining subglacial processes from surface velocity observations using surrogate-based Bayesian inference. *Journal of Glaciology*, 67(263), 385–403. <https://doi.org/10.1017/jog.2020.112>
- 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(3), e2024JH000169. <https://doi.org/10.1029/2024JH000169>
- Cowen-Breen, C., Wang, Y., Bates, S., & Lai, C.-Y. (2024). Euler operators for mis-specified physics-informed neural networks. *ICML 2024 AI for Science Workshop*. <https://openreview.net/forum?id=kkGR5fNq2J>
- Daw, A., Bu, J., Wang, S., Perdikaris, P., & Karpatne, A. (2022). Mitigating propagation failures in physics-informed neural networks using retain-resample-release (r3) sampling. *arXiv Preprint arXiv:2207.02338*. <https://doi.org/10.48550/arXiv.2207.02338>
- Fürst, J., Durand, G., Gillet-Chaulet, F., Merino, N., Tavard, L., Mouginot, J., Gourmelen, N., & Gagliardini, O. (2015). Assimilation of Antarctic velocity observations provides evidence for uncharted pinning points. *The Cryosphere*, 9, 1427–1443. <https://doi.org/10.5194/tc-9-1427-2015>
- 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>
- 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>
- 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>
- Jagtap, A. D., & Karniadakis, G. E. (2020). Extended physics-informed neural networks (XPINNs): A generalized space-time domain decomposition based deep learning framework for nonlinear partial differential equations. *Communications in Computational Physics*, 28(5). <https://doi.org/10.4208/cicp.oa-2020-0164>
- Jouvet, G. (2023). Inversion of a Stokes glacier flow model emulated by deep learning. *Journal of Glaciology*, 69(273), 13–26. <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>

- Jouvet, G., Cordonnier, G., Kim, B., Lüthi, M., Vieli, A., & Aschwanden, A. (2022). Deep learning speeds up ice flow modelling by several orders of magnitude. *Journal of Glaciology*, 68(270), 651–664. <https://doi.org/10.1017/jog.2021.120>
- 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>
- MacAyeal, Douglas R. (1989). Large-scale ice flow over a viscous basal sediment: Theory and application to ice stream B, Antarctica. *Journal of Geophysical Research: Solid Earth*, 94, 4071–4087. <https://doi.org/10.1029/JB094iB04p04071>
- MacAyeal, Douglas R. (1993). A tutorial on the use of control methods in ice-sheet modeling. *Journal of Glaciology*, 39, 91–98. <https://doi.org/10.3189/S0022143000015744>
- Millstein, J. D., Minchew, B. M., & Pegler, S. S. (2022). Ice viscosity is more sensitive to stress than commonly assumed. *Communications Earth & Environment*, 3, 1–7. <https://doi.org/10.1038/s43247-022-00385-x>
- Morlighem, M. (2020). *MEaSUREs BedMachine Antarctica, Version 2*. NASA National Snow; Ice Data Center Distributed Active Archive Center. <https://nsidc.org/data/NSIDC-0756/> versions/2
- Morlighem, Mathieu, 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(14). <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: Earth Surface*, 118(3), 1746–1753. <https://doi.org/10.1002/jgrf.20125>
- Mouginot, J., Rignot, E., & Scheuchl, B. (2019). *MEaSUREs Phase-Based Antarctica Ice Velocity Map, Version 1*. NASA National Snow; Ice Data Center Distributed Active Archive Center. <https://nsidc.org/data/NSIDC-0754/> versions/1
- 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), 1894–1917. <https://doi.org/10.1002/2014JF003181>
- Rahaman, N., Baratin, A., Arpit, D., Draxler, F., Lin, M., Hamprecht, F. A., Bengio, Y., & Courville, A. (2019). On the spectral bias of neural networks. *arXiv Preprint arXiv:1806.08734*. <https://doi.org/10.48550/arXiv.1806.08734>
- Raiassi, 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>
- Ranganathan, M., & Minchew, B. (2024). A modified viscous flow law for natural glacier ice: Scaling from laboratories to ice sheets. *Proceedings of the National Academy of Sciences*, 121(23), e2309788121. <https://doi.org/10.1073/pnas.2309788121>
- 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 mechanics of slip along glacier beds using physics-informed neural networks: Case study on Rutford Ice Stream, Antarctica. *Journal of Advances in Modeling Earth Systems*, 13, e2021MS002621. <https://doi.org/10.1029/2021MS002621>

- Vieli, A., & Payne, A. J. (2003). Application of control methods for modelling the flow of Pine Island Glacier, West Antarctica. *Annals of Glaciology*, 36, 197–204. <https://doi.org/10.3189/172756403781816338>
- Wang, Y., & Lai, C.-Y. (2024). Multi-stage neural networks: Function approximator of machine precision. *Journal of Computational Physics*, 112865. <https://doi.org/10.1016/j.jcp.2024.112865>
- Wang, Y., Lai, C.-Y., & Cowen-Breen, C. (2022). Discovering the rheology of Antarctic ice shelves via physics-informed deep learning. *SSRN Preprint*. <https://doi.org/10.21203/rs.3.rs-2135795/v1>
- Wang, Y., Lai, C.-Y., David J. Prior, and, & Cowen-Breen, C. (2025). Deep learning the flow law of Antarctic ice shelves. *Science*. <https://doi.org/10.1126/science.adp3300>
- Xu, Z.-Q. J., Zhang, Y., Luo, T., Xiao, Y., & Ma, Z. (2020). Frequency principle: Fourier analysis sheds light on deep neural networks. *Communications in Computational Physics*, 28(5), 1746–1767. <https://doi.org/10.4208/cicp.OA-2020-0085>