# Mass Conservation on Rails

# Rethinking Physics-Informed Learning of Ice Flow Vector Fields


Kim Bente[1] Roman Marchant[2] Fabio Ramos[1,3]

[1] School of Computer Science, The University of Sydney [2] Human Technology Institute, University of Technology Sydney [3] NVIDIA

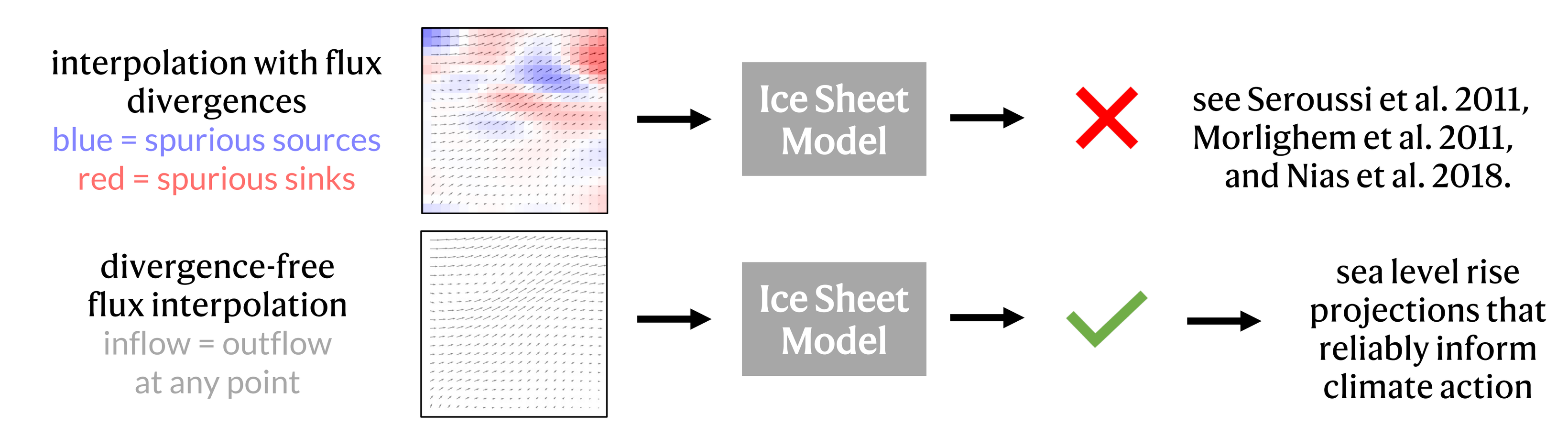


## Summary

- To reliably project **sea level rise**, ice sheet models require physically plausible and skilful interpolations of ice flow from sparse & noisy measurements.
- Mass Conservation can be achieved by imposing a **divergence-free constraint** on ice flux vector fields ( $\nabla \cdot \boldsymbol{v} = 0$ ), where flux is the ice mass transport rate.
- Existing **Physics-Informed ML** approaches only use soft-constrained PINNs, which can't guarantee physics and often suffer from spectral bias and convergence issues.
- We propose **hard-constrained dfNNs** (divergence-free Neural Networks), which guarantee Mass Conservation, as well as a **directional guidance** learning strategy that leverages dense satellite data.
- Reconstruction experiments on **real Antarctic ice flux data** highlight the **physics & accuracy benefits** of using hard-constrained dfNNs over soft-constrained PINNs or unconstrained NNs.


## Problem

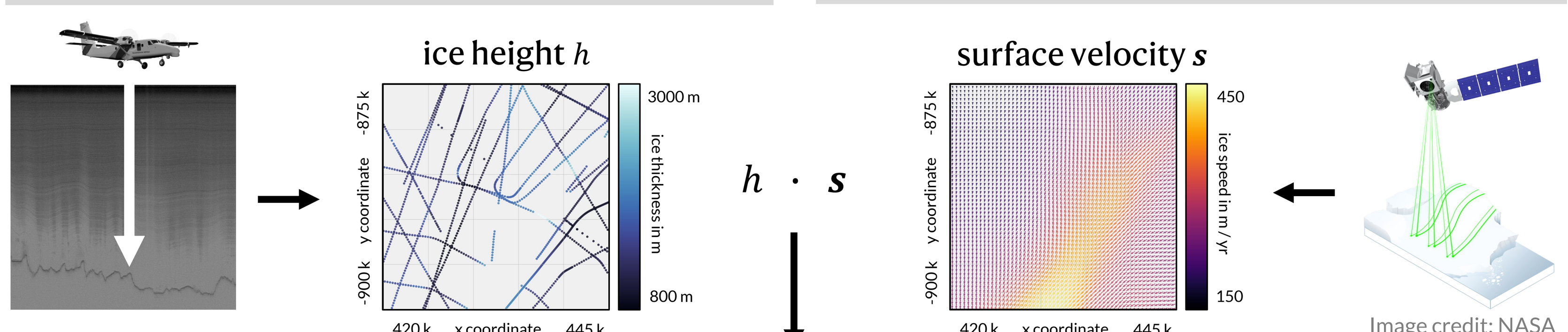
Climate projection models rely on inputs that adhere to fundamental physical principles:




→ Numerical ice sheet models require ice flux initialisations to respect **Mass Conservation (MC).**

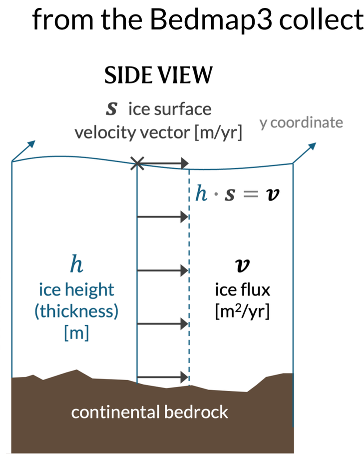
## Real Antarctic Data

| sparse ice **height** point data $h$ from airborne | dense **surface velocity** vectors $s$ from satellites |
|---|---|



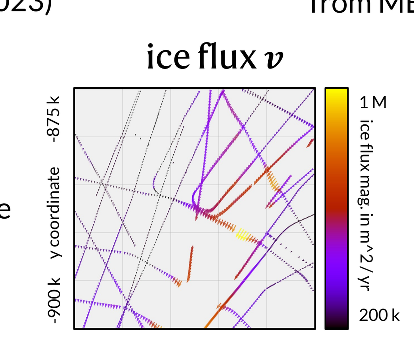

from the Bedmap3 collection (Fremand et al. 2023)

from MEaSUREs InSAR velocity map (Rignot et al. 2017)




$\boldsymbol{v} = h \cdot \boldsymbol{s}$

**Ice flux** describes the transport of ice mass/volume per unit time for any point in 2D space.




After removing negligible terms from the mass continuity equation,

$$\frac{\partial h}{\partial t} + \nabla \cdot (h\, \boldsymbol{s}) = \text{smb} - \text{bmb}$$

it simplifies to a divergence-free constraint.

$$\nabla \cdot \boldsymbol{v} = 0$$

→ Local Mass Conservation can be enforced via a **divergence-free constraint.**

## Physics-Informed Machine Learning (PIML)

Three PIML strategies after Karniadakis et al. 2021

| **Observational bias**<br>unconstrained models | **Learning bias**<br>soft-constrained models | **Inductive bias**<br>hard-constrained models |
|---|---|---|
| learn from observations that embody physics | additional physics violation loss term promotes physics | model or kernel architecture enforces physics exactly |
| Implicit strategy, but observations of ice flux are limited and noisy. | Teisberg et al. 2021 & Steidl et al. 2025 propose PINNs for ice flux. Generally, PINNs lack guarantees, and can incur convergence issues and spectral bias. | No glaciology applications yet. Generally, hard-constrained models guarantee physics, and may converge and generalise better especially when data is sparse & noisy. |
| e.g. NNs | e.g. PINNs | → e.g. dfNNs (proposed) |

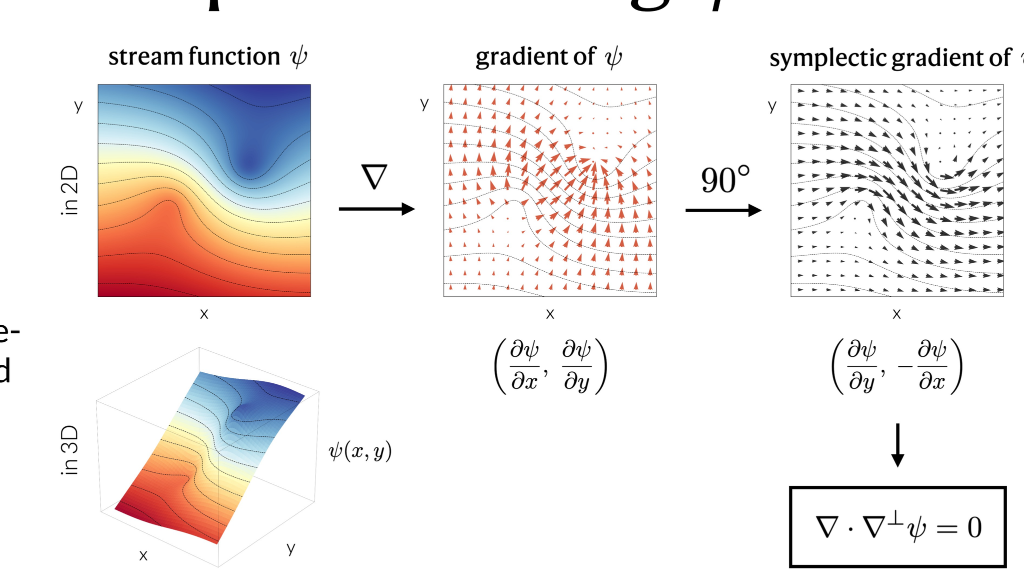
## A fluid dynamics trick – parametrising $\psi$

→ Let $\psi(x, y)$ (i.e. "psi") be a sufficiently smooth scalar field defined on a 2D spatial domain.

→ Rotating the gradient $\nabla\psi$ of this surface by 90° yields the symplectic gradient, denoted $\nabla^{\perp}\psi$.

→ The resulting vector field is divergence-free by construction ($\nabla \cdot \nabla^{\perp}\psi = 0$) and thus enforces mass conservation.

→ In 2D the symplectic vector field is equivalent to the Hamiltonian vector field.




→ By parametrising a stream function $\psi$ we can model **guaranteed divergence-free vector fields.**

## The divergence-free Neural Network (dfNN)

Based on hard-constrained methods introduced in


- Model Inclusive Learning [Kuroe et al. 1998] – early work proposed for learning vector field decompositions.
- Hamiltonian Neural Networks (HNNs) [Greydanus et al. 2019] – initially developed for energy conservation.
- Neural Conservation Laws [Richter-Powell et al. 2022] – generalisation for solving the continuity equation.


→ We propose the first application of dfNNs to glaciology.

→ Our implementation leverages PyTorch's automatic differentiation in the forward pass.

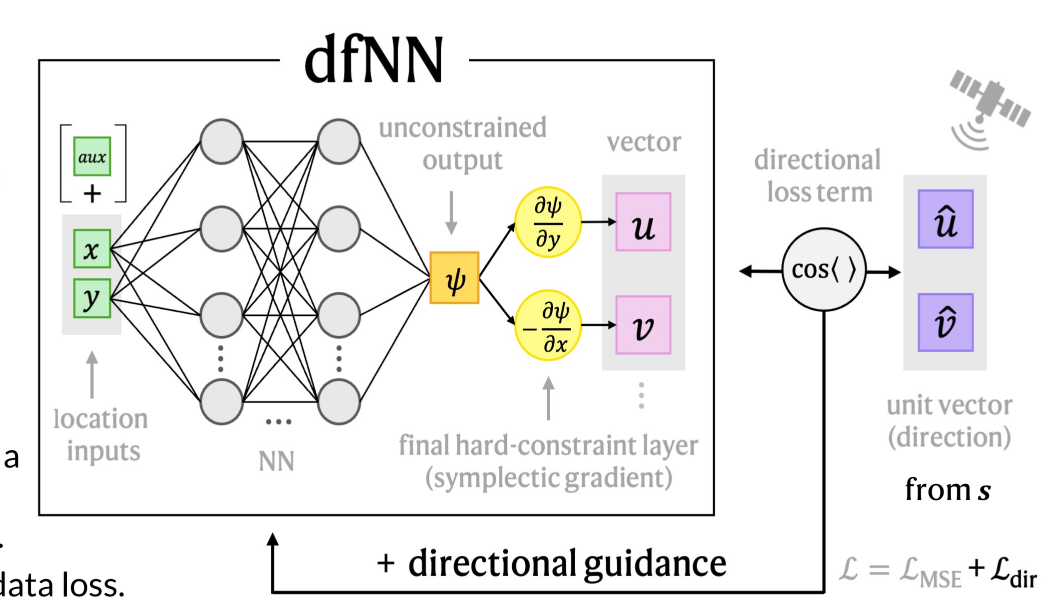
→ Using a SiLU activation function benefitted model convergence.

**Directional guidance (proposed)**

→ We also propose directional guidance, a learning strategy that leverages dense directional information from satellites.

→ A PINN-style loss $\mathcal{L}_{\text{dir}}$ is added to the data loss.

→ Can be applied to any NN (including dfNNs).

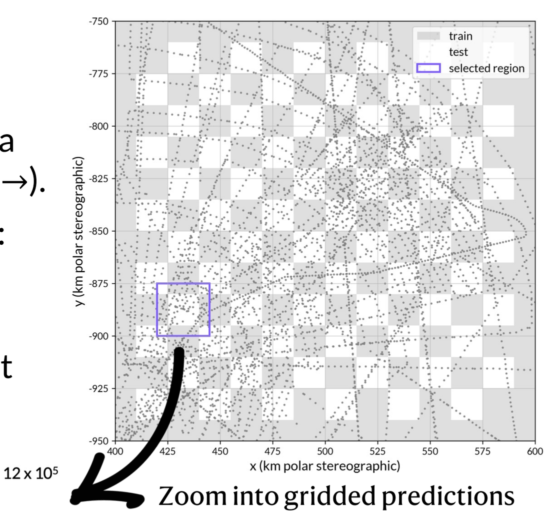


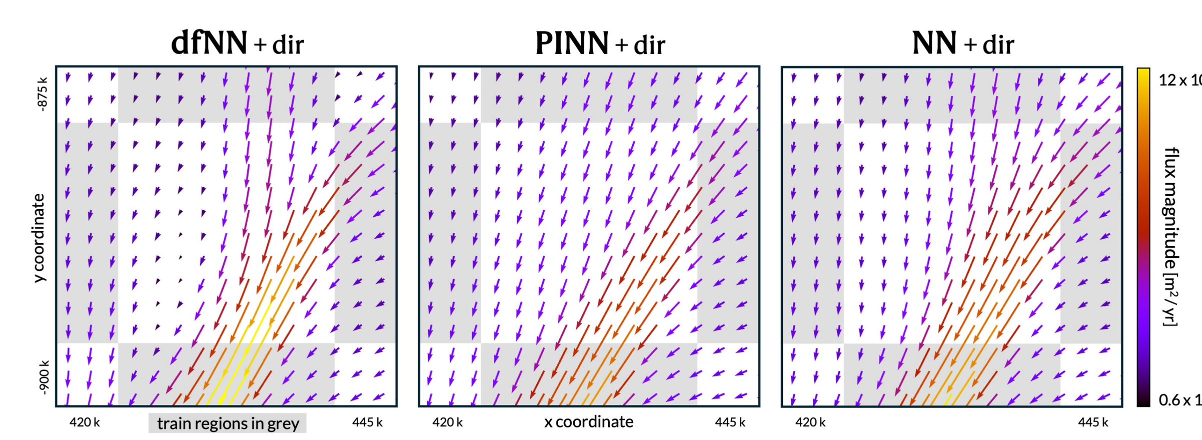

## Experiments over Byrd Glacier

**Train & Test** – We divide a 200×200 km region over Byrd Glacier, Antarctica into train (27k) and test regions (22k) following a chequerboard pattern (see →).

**Models** – **dfNN**, PINN, and NN (ceteris paribus) +2 additional variants each:
 **+ dir** – directional guidance, using dense surface velocity observations.
 **+ aux** – auxiliary inputs; here we use surface elevation from satellites.

**Metrics** – RMSE & MAE ↓ and MAD ↓ (Mean Absolute Divergence) over test



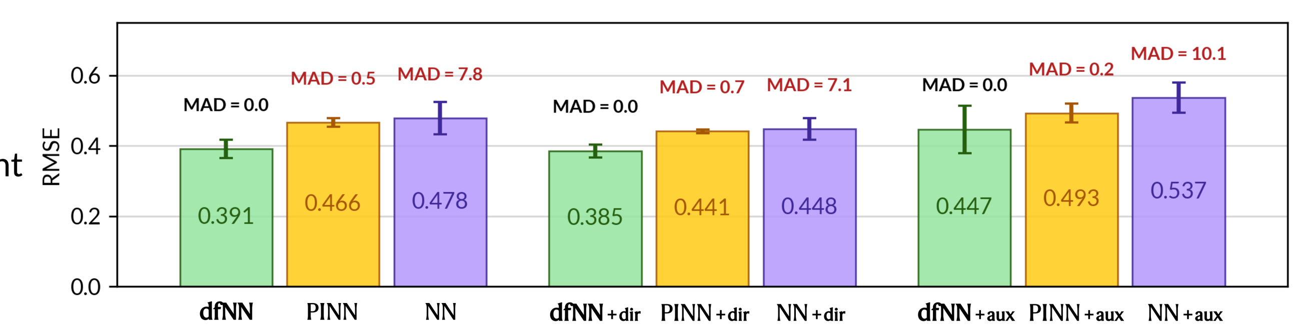




→ **dfNN** + dir (best model) naturally captures curved flow routing.

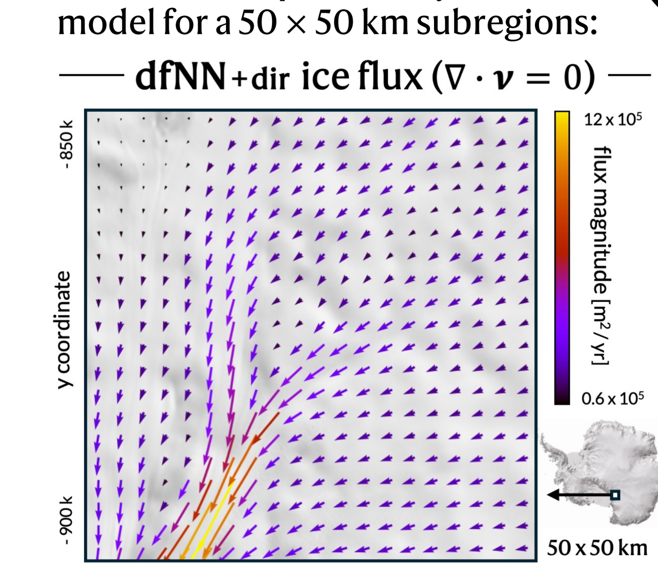
→ PINN + dir and NN + dir fit local variations less well (spectral bias) and produce overly linear flows.

## Results

**dfNN** + dir which enforces the diverge-free constraint by design, is the overall best model.




Gridded interpolation by our best model for a 50 × 50 km subregions:




**dfNNs** models perform best, followed by PINNs and NNs: hard-constraining > soft-constraining > not constraining

- Soft-constraining significantly reduces divergence, but only hard-constrained models produce fully divergence-free vector fields (MAD = 0.0).
- Surface elevation as an auxiliary predictor (+aux), introduced more noise than signal and did not improve predictions. Presumably, flow controls over Byrd glacier are not expressed as variations in surface elevation.

→ **hard-constrained dfNNs are the more accurate and physically reliable than PINNs or NNs.**

→ **directional guidance (+dir) boosts all models.**

## Ongoing and future work

✓ Preprint now available: **Probabilistic model variant (dfNGP)** incorporating the dfNN into a divergence-free Gaussian Process to quantify spatially varying uncertainty.

❑ Applications to **other environmental flows such as of ground water, wind, ocean currents, pollutants, or the atmosphere.**

❑ Extension to **higher dimensions** and spatio-temporal solutions to the continuity equation (building on Richter-Powell et al. 2022)

Join the ML4Cryo community

For the paper, the dfNN implementation in PyTorch, fully reproducible experiments, and full references scan


This research was supported by the Australian Government through the Australian Research Council's Industrial Transformation Training Centre in Data Analytics for Resources and Environments (DARE) (project IC190100031) as well as through the Australian Government Research Training Program (RTP) Scholarship.