Global glacier ice thickness inversion with supervised learning

N. Maffezzoli[1,2]

... *and* E. Rignot, C. Barbante, T. Petersen, S. Vascon

[1]Ca' Foscari University of Venice
[2]University of California Irvine

HAMLET-PHYSICS 2025, Copenhagen, 21 August 2025

A glacier

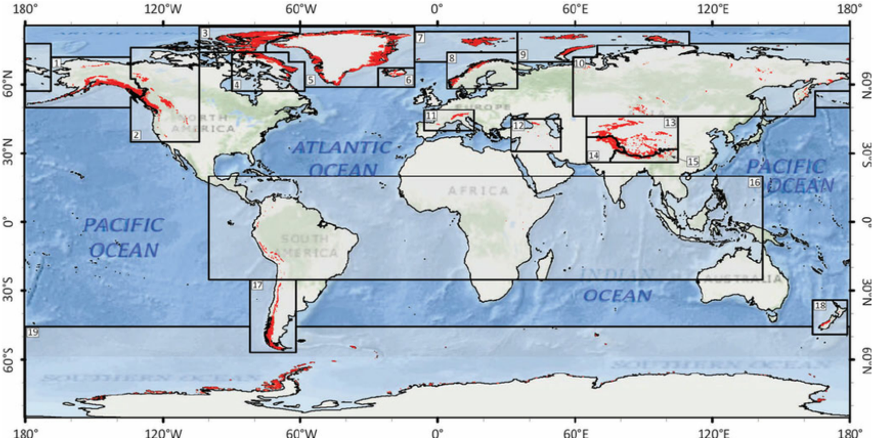
A glacier

A glacier

A glacier

Earth's glaciers: Randolph Glacier Inventory (RGI)

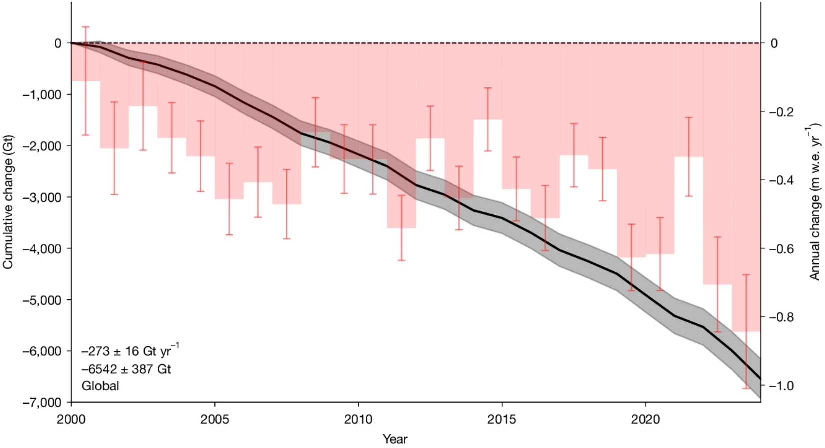
RGI Version v. 7: 274,531



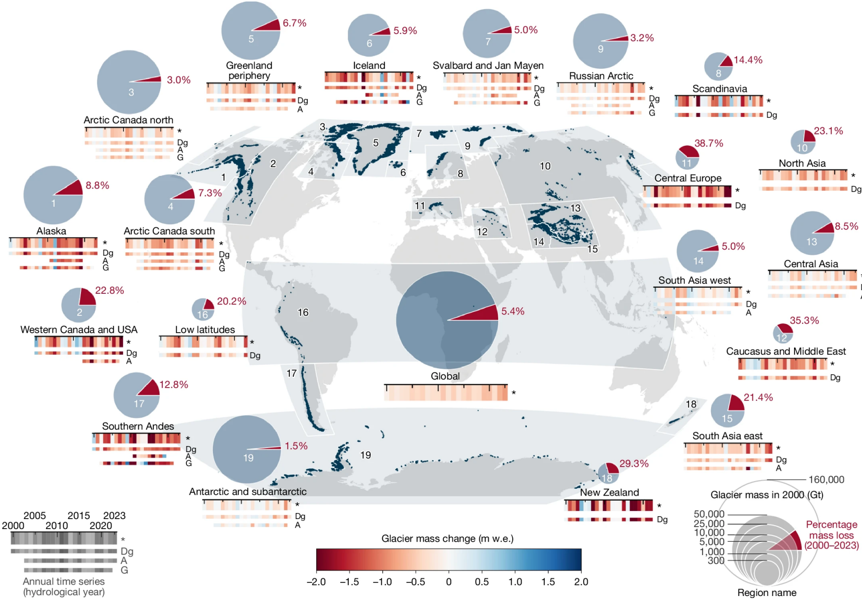
[1] Image from Randolph Glacier Inventory, version 6.

Motivations

- Glacier mass loss: 18% larger than the loss from the Greenland Ice Sheet and more than 2x Antarctic Ice Sheet.
- Global loss 273±16 Gt/yr from 2000 to 2023, with an increase of 36±10% over the last decade.



The GlaMBIE Team, Nature, 2025



The GlaMBIE Team, Nature, 2025

This is dM/dt (mass change)

But ...

How much ice mass is there ? M = ?

Current approaches for glacier ice thickness inversion

1. Area-volume scaling approaches, $V \simeq cA^{\gamma}$
2. Geometrical models of bedrock shape (e.g. u-shaped valleys)
3. Mass conservation: $\frac{dH}{dt} + \nabla \cdot (H\vec{v}) = M_s + M_b$
4. Approximations of full-Stokes model, e.g. Shallow Ice Approximation (SIA, 1984)

$$H = \left(\frac{v_s(1-\beta)(n+1)}{2A(\rho g)^n ||\vec{\nabla} z||^n} \right)^{\frac{1}{n+1}}$$

5. Machine (-deep) learning: few local attempts.

Global glacier models:

1. Farinotti et al. (2019), ensemble of (up to) 5 models
2. Millan et al. (2022), Shallow Ice Approximation

In-situ surveys: ice thickness measurements

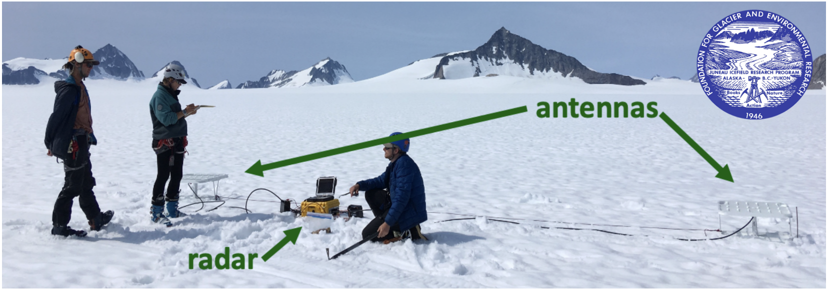
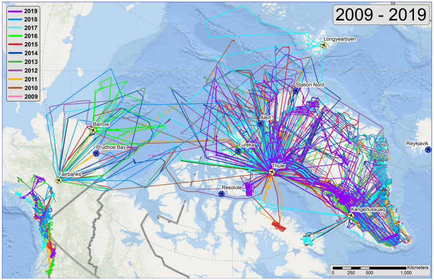


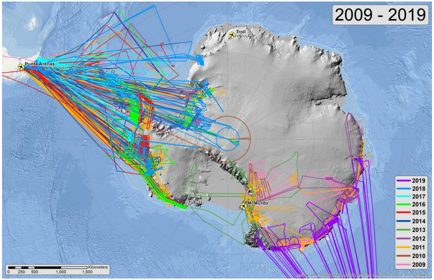
Photo credit: Elizabeth Case.

Ground-penetrating radar systems: emit radio waves and record the reflected signal.
Basal reflections originate from the strong contrast between the ice and the underlying sediment, bedrock, or water.

NASA IceBridge: a legacy of millions of data points

For eleven years from 2009 through 2019, the planes of NASA's Operation IceBridge flew above the Arctic, Antarctic and Alaska, gathering data on the height, depth, **thickness**, flow and change of sea ice, glaciers and ice sheets.



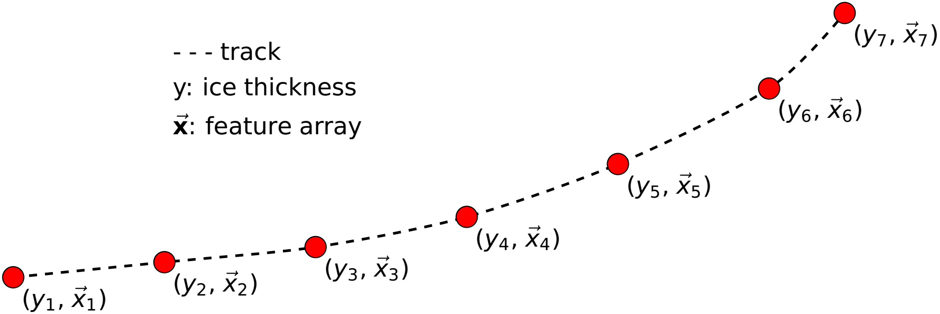


Goal

Build a ML model for ice thickness inversion of glaciers.

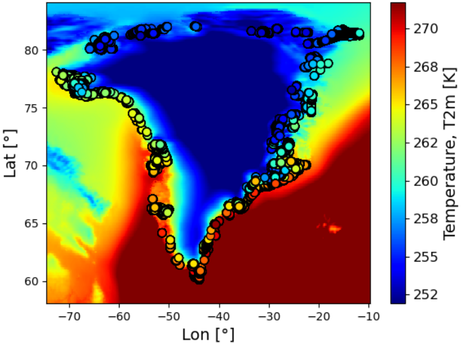
Methodology

- Supervised-learning approach
- **y**: target tabular data (ice thickness is measured discretely in space).
- $\vec{\mathbf{x}}$: array of model inputs. Some are raster, some are tabular.



Model inputs, $\vec{x}$

1. Distance to margin
2. Surface mass balance (+1)
3. Surface ice velocity (6x)
4. Elevation
5. Elevation normalized 0-1
6. Elevation from base
7. Slope (8x +1)
8. Curvature (6x + 1)
9. Glacier zmin, zmax, zmed
10. Glacier zmax-zmin
11. Glacier and Cluster Areas
12. Glacier Perimeter and Length
13. Glacier Aspect
14. Surface temperature
15. Distance to ocean



Ice thickness inversion via gradient-boosted trees (ICEBOOST)

Pipeline

- Training data: y (ice thickness) + 39 numerical features $\vec{x}$
- Gradient Boosted Decision Trees: XGBoost + CatBoost.

Training data

- **Target (y)**: 6.5 million thickness measurements
- Data from ca. 1400 glaciers (0.6% of all Earth's glaciers).

Hyperparameters

- Model is optimized globally (not stratified regionally)

[1] GlaThiDa Consortium (2020): Glacier Thickness Database 3.1.0. World Glacier Monitoring Service, Zurich, Switzerland.

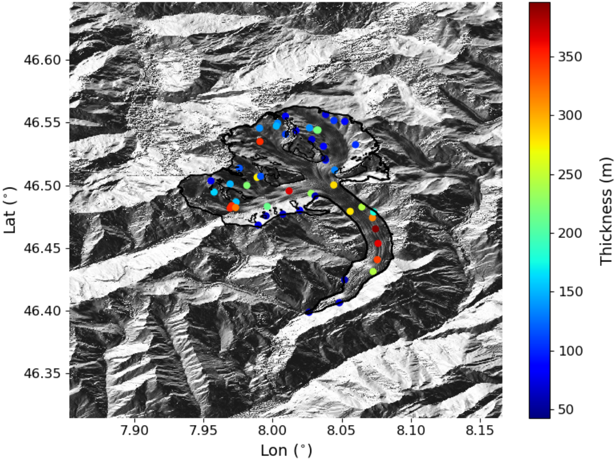
Model error: rmse, [m]

Region	IceBoost 2025	Millan 2022	Farinotti 2019
Alaska, US, West Canada	**116 (21)**	151	173
Arctic Canada North	**83 (7)**	131	129
Arctic Canada South	**58 (9)**	103	115
Greenland Periphery	**93 (23)**	112	112
Iceland	-	-	-
Svalbard	52 (7)	66	**51**
Scandinavia	**42 (6)**	60	53
Russian Arctic	-	-	-
North Asia	**15 (3)**	19	23
Central Europe	35 (5)	46	35
Caucasus and Middle East	56 (1)	65	56
Asia	**36 (12)**	62	37
Low Latitudes	-	-	-
Southern Andes	43 (8)	**35**	40
New Zealand	-	-	-
Antarctic and Islands	**109 (20)**	113	192

Maffezzoli et al., GMD, 2025

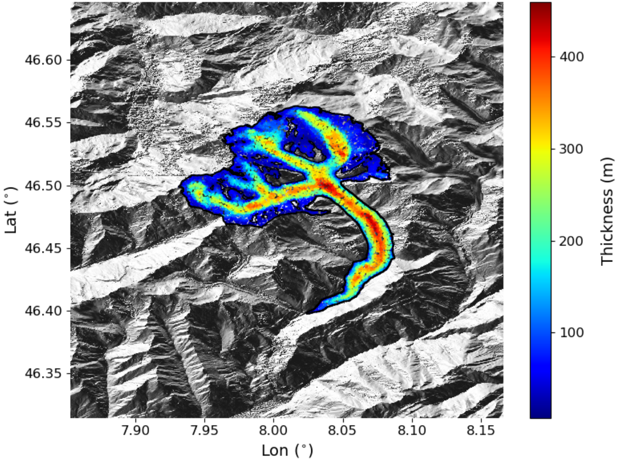
Model deploy

- Compute feature vector $\vec{x}$ at locations within the glacier.
- Query the model locally: $y = ICEBOOST(\vec{x})$.

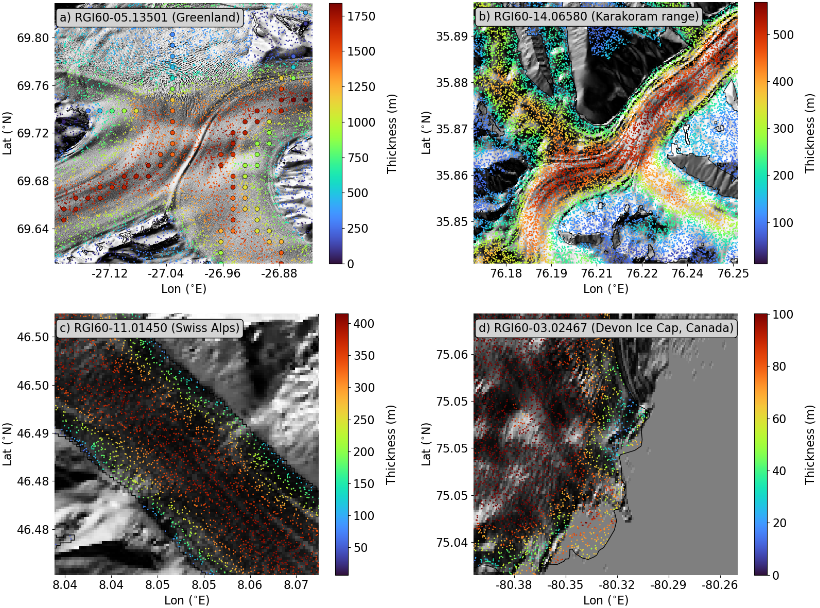


Model deploy

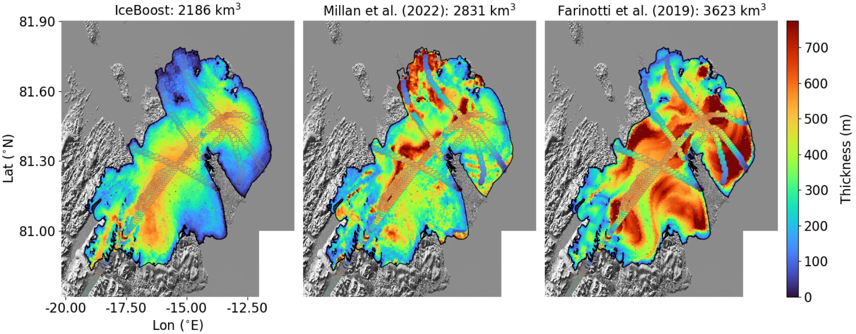
- Compute feature vector $\vec{x}$ at locations within the glacier.
- Query the model locally: $y = ICEBOOST(\vec{x})$.



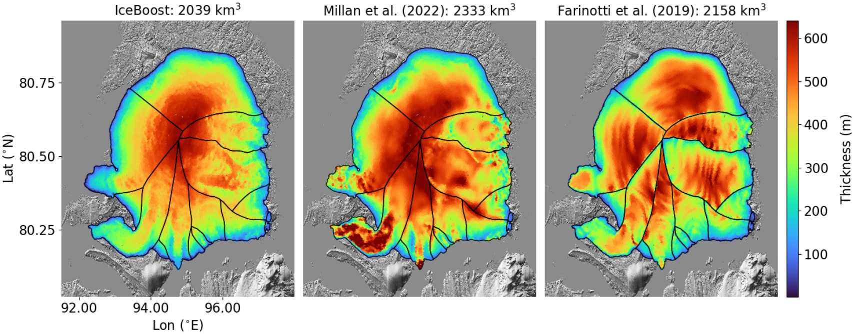
Model deploy



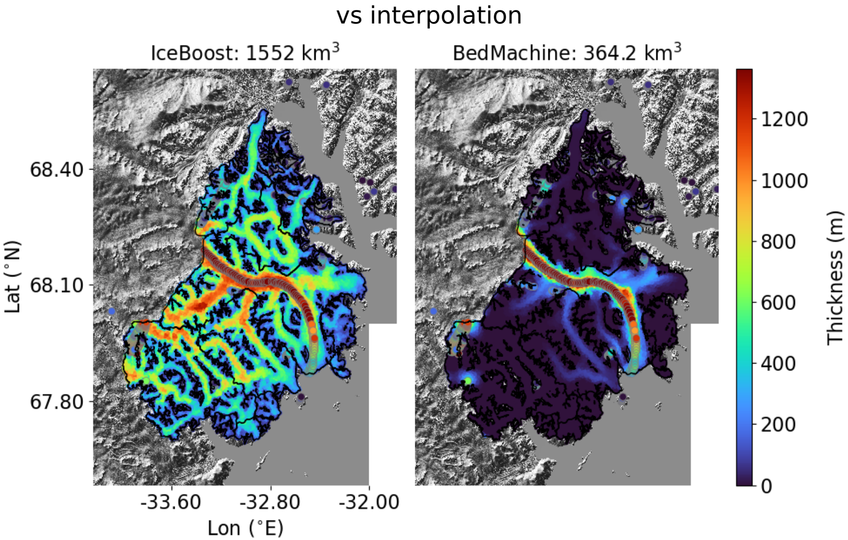
Model deploy



Model deploy

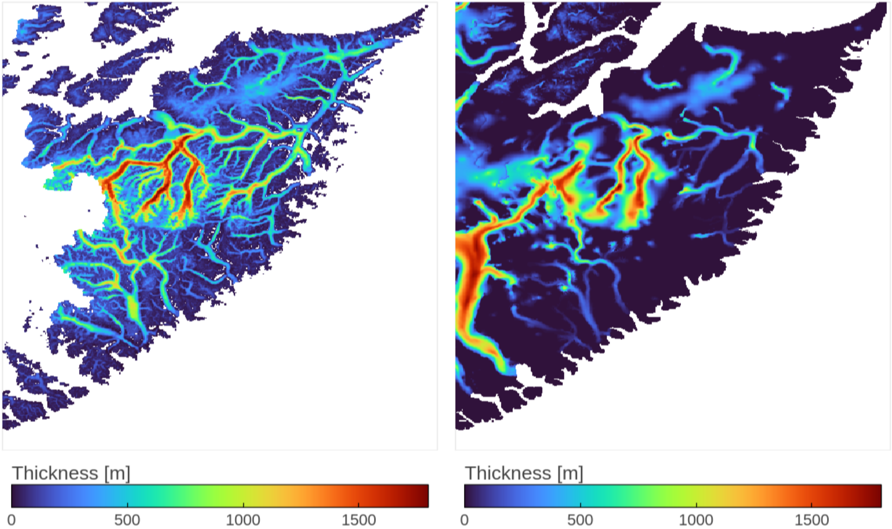


Model deploy

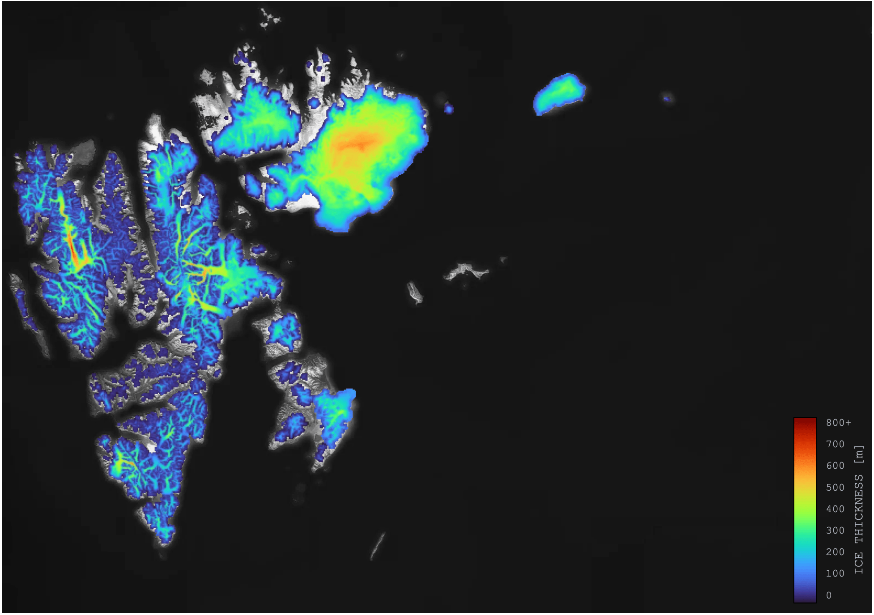


Geikie Plateau (Greenland)

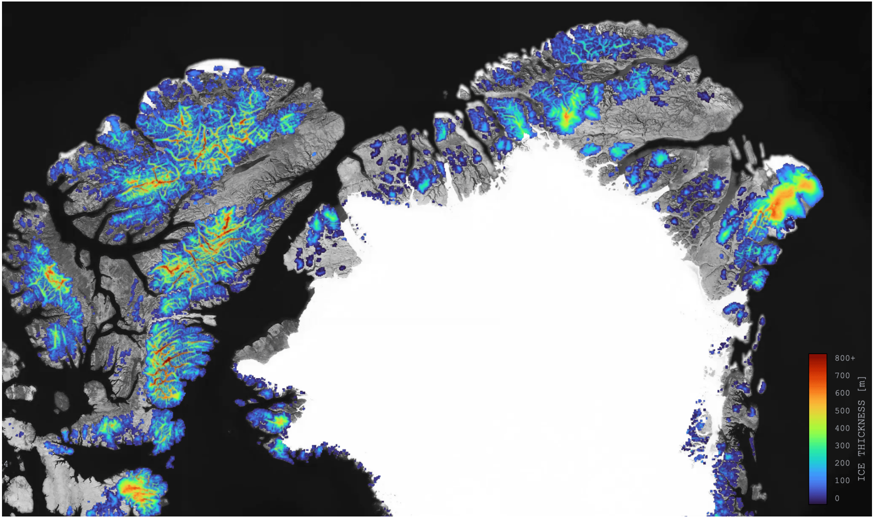
ICEBOOST (left) - BedMachine v5 (right)



Svalbard



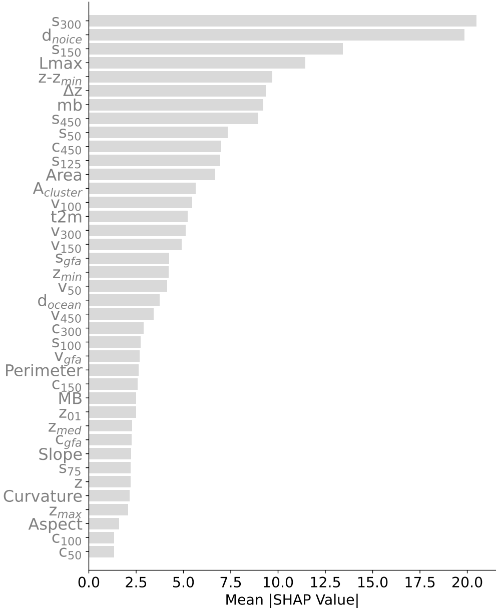
Arctic



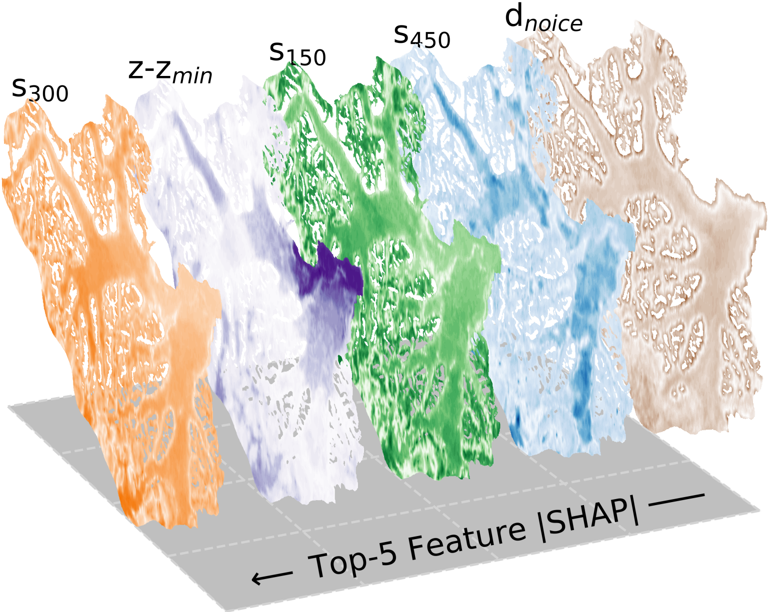
Global ice volumes on RGI v. 6

Region ($\cdot 10^3\ km^3$)	**IceBoost 2025**	**Millan 2022**	**Farinotti 2019**
Alaska, US, West Canada	18.74	19.2 ± 5.6	20.0 ± 5.0
Arctic Canada North	24.6	25.4 ± 7.2	28.3 ± 7.3
Arctic Canada South	7.1	7.0 ± 2.1	8.6 ± 2.2
Greenland Periphery	12.8	11.8 ± 3.7	15.7 ± 4.1
Iceland	3.98	3.7 ± 0.9	3.8 ± 1.0
Svalbard	6.90	7.0 ± 2.3	7.5 ± 1.9
Scandinavia	0.29	0.30 ± 0.10	0.30 ± 0.08
Russian Arctic	13.9	15.5 ± 3.9	14.6 ± 3.8
North Asia	0.16	0.11 ± 0.05	0.14 ± 0.04
Central Europe	0.11	0.12 ± 0.05	0.13 ± 0.03
Caucasus and Middle East	0.067	0.05 ± 0.03	0.06 ± 0.02
Asia	9.25	9.4 ± 3.7	7.02 ± 1.0
Low Latitudes	0.11	0.07 ± 0.04	0.10 ± 0.03
Southern Andes	5.97	5.9 ± 1.6	5.3 ± 1.4
New Zealand	0.077	0.07 ± 0.03	0.07 ± 0.02
Antarctic and Islands	44.1	35.1 ± 9.1	46 ± 12
Total ($\cdot 10^3\ km^3$)	148	141 ± 40	158 ± 41

SHAP analysis of model features



Feature ranking with SHAP



Computation

Hardware

- Hard Disk (Input products, .tif, .nc, .shp, .gpkg): 500 GB.
- Hard Disk (training dataset, .csv/.parquet): 3 GB.
- Hard Disk (Model): 10 MB.
- RAM: 32-64 GB.
- No GPU.
- Multi-core preferable, n_{jobs}=8.

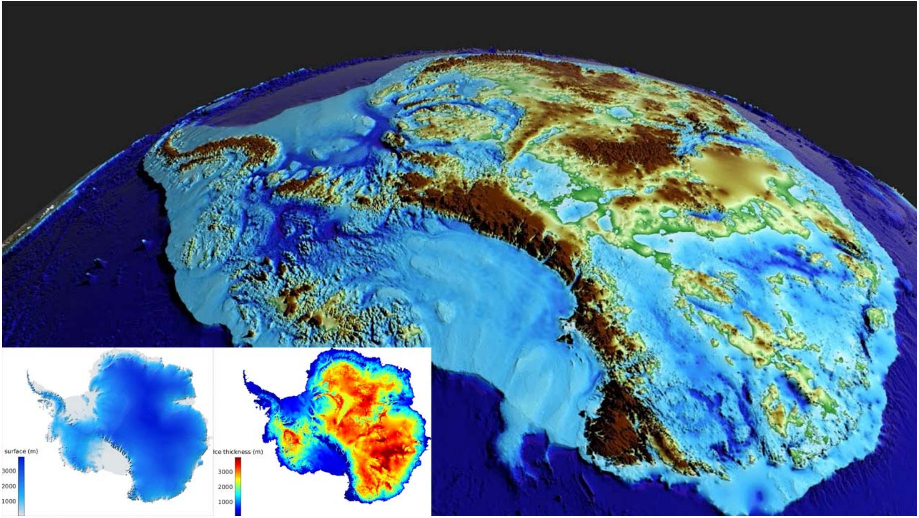
Software

- Python (+ optional cuda with NVIDIA Rapids)

Time

- Feature fetching on-the-fly: CPU. Time: Min 1 s/glacier. Max 60 s/glacier.
- Model run: 0.1 s/glacier on CPU.
- E.g. regional simulation: Svalbard (1615 glaciers): 6’ with n_{jobs}=8.

Can we model the ice sheets ?



Model: BedMachine v3 (credit: UCI-JPL-Dartmouth)

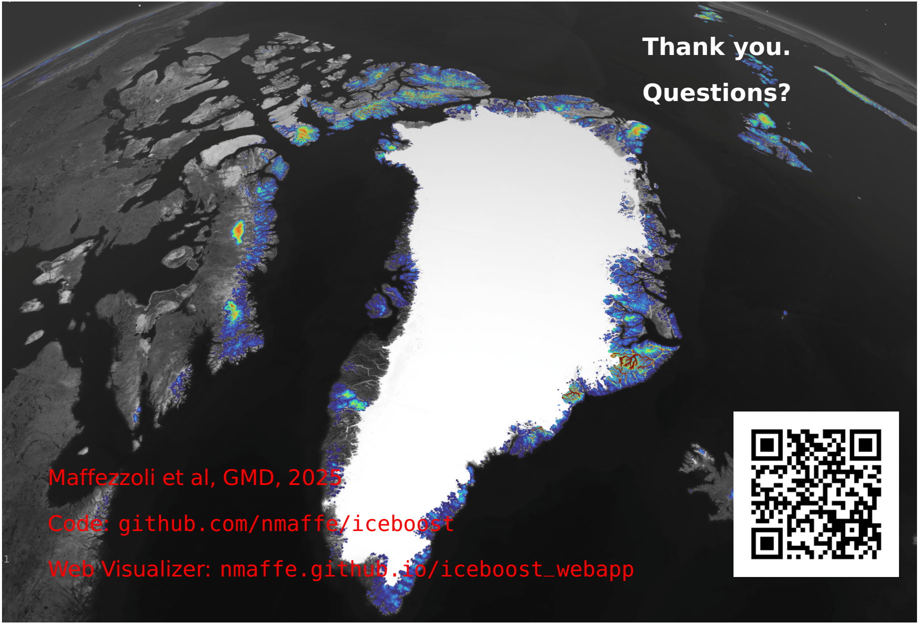
Some conclusions

Positives

- Lower error by up to 30-40 % at high latitudes compared to baseline models.
- Training dataset can be enlarged.
- Features can be added (and assessed and removed).
- Features can be improved (e.g. DEM, ice velocity, distributed mass-balance, glacier geometries).

Drawbacks

- Input features and target should well behave. Feature imputation is tricky.
- Decision boundaries of decision trees may appear.
- No physics.



Thank you.

Questions?

Maffezzoli et al, GMD, 2025

Code: `github.com/nmaffe/iceboost`

Web Visualizer: `nmaffe.github.io/iceboost_webapp`