Where shall we measure?
Results from the second phase of the Ice Thickness Models Intercomparison eXperiment (ITMIX2)

D. Farinotti, D.J. Brinkerhoff , J. J. Fürst, P. Gantayat, F. Gillet-Chaulet, M. Huss, P. W. Leclercq, H. Maurer, M. Morlighem, A. Pandit, A. Rabatel, R. Ramsankaran, T. J. Reerink, E. Robo, E. Rouges, E. Tamre, W. J. J. van Pelt, M. A. Werder, M. F. Azam, H. Li, and L. M. Andreassen

Laboratory of Hydraulics, Hydrology and Glaciology (VAW)
Swiss Federal Institute for Forest, Snow and Landscape Research WSL

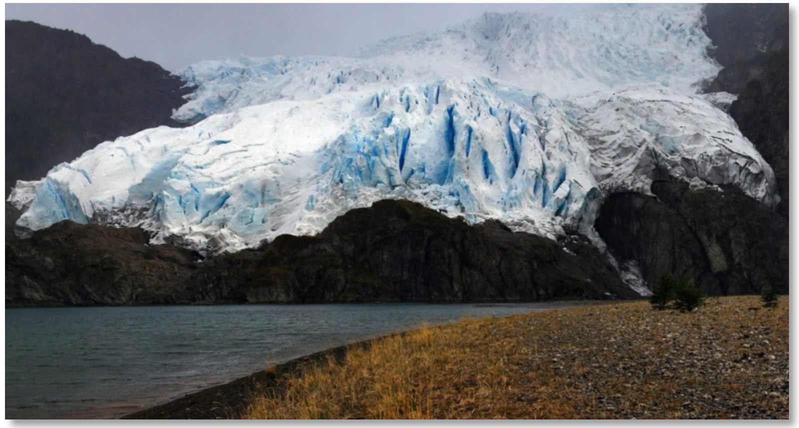
WE WANT TO KNOW GLACIER ICE THICKNESS!

B e c a u s e

Glaciers = sea level change

Glaciers = water resources

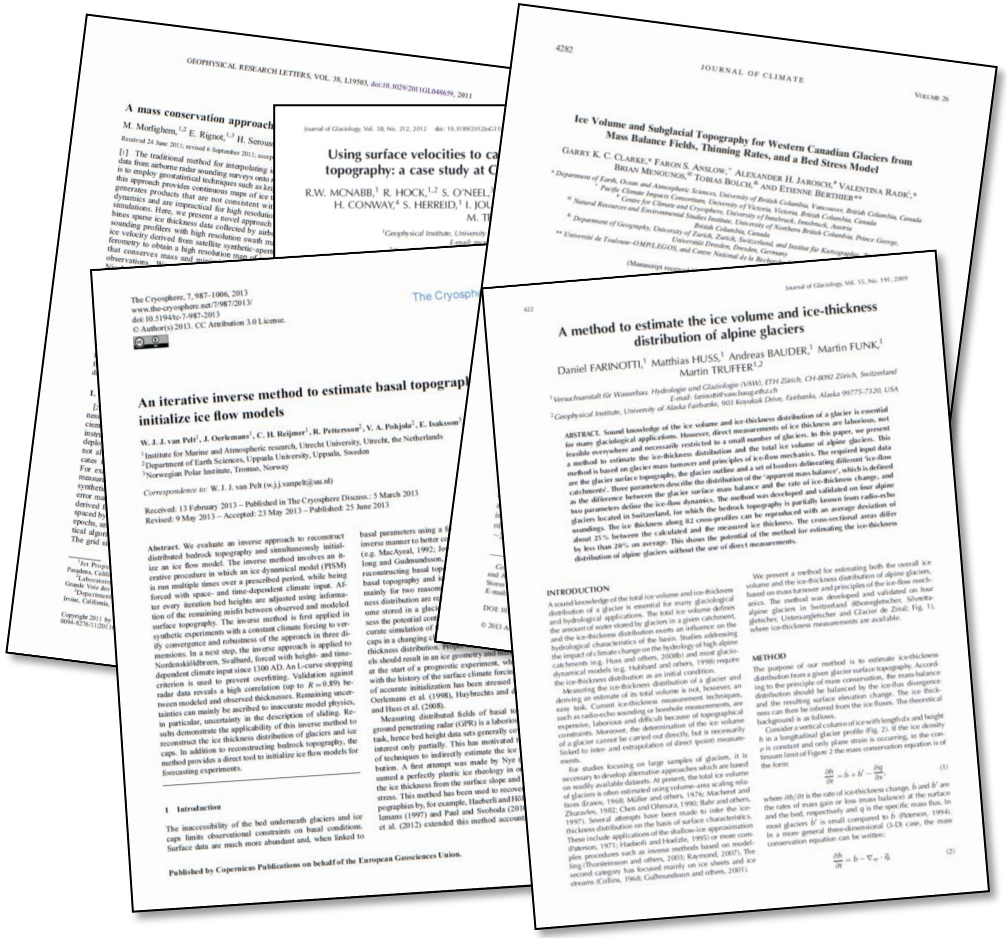
Glaciers = interesting



How to go about it?

Either M E A S U R E …

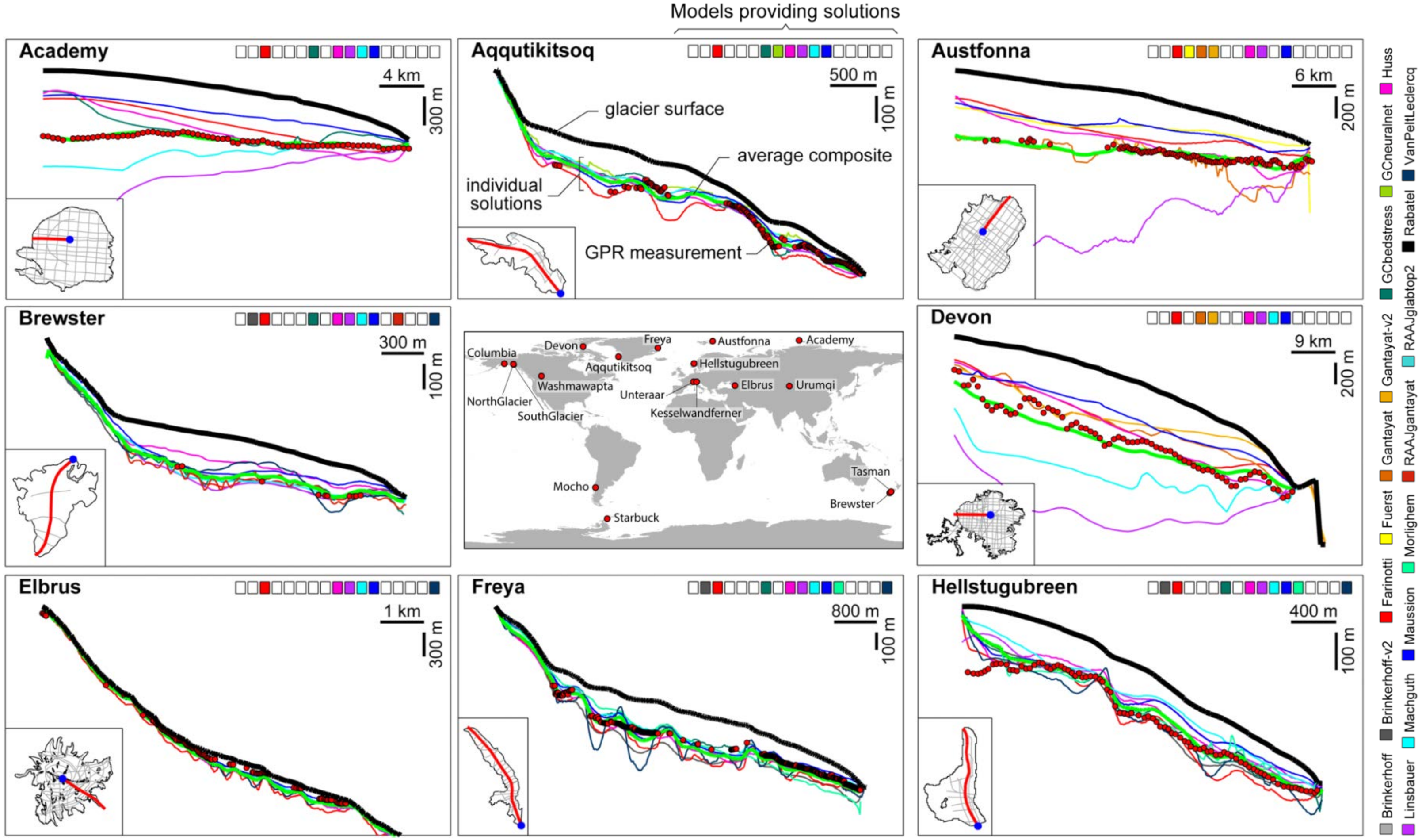
… or M O D E L !



But what model shall we chose?

Results from ITMIX1

A first answer to the question “which model shall we chose” was given by the
Ice Thickness Models Intercomparison eXperiment phase 1 …



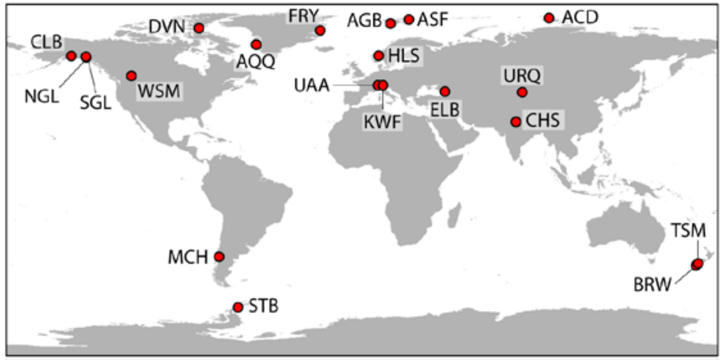
Farinotti et al., TC, 2017

… but the “game” of ITMIX1 was ***“given the surface, estimate the thickness”*** (i.e. no measured ice thickness was provided).

ITMIX2 thus addressed the case in which some sparse thickness data are available.

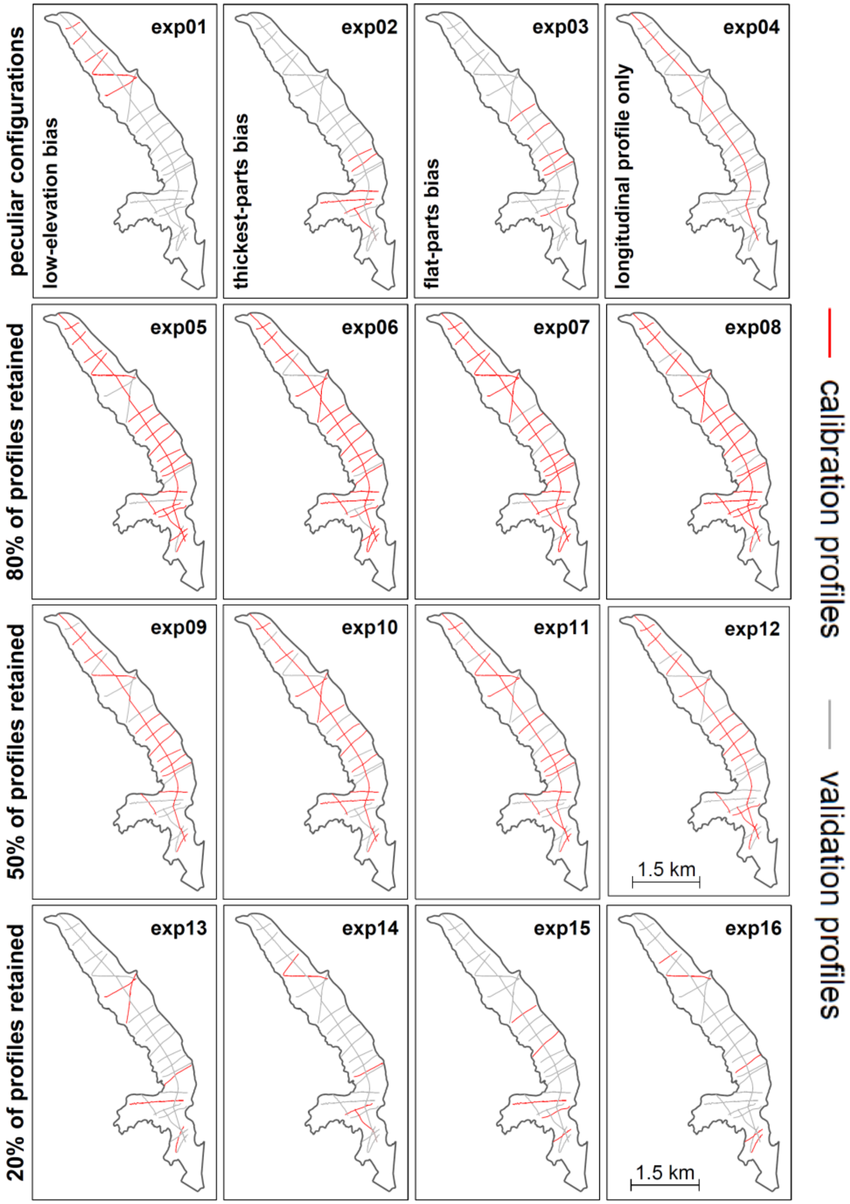
ITMIX2 setup

The same glaciers as in ITMIX1 (“test cases”) were considered
Provided info: glacier outline, surface DEM (and dh/dt, SMB, velocity when available)



Subsets of ice thickness measurements were provided in different combinations (“experiments”), to mimic different situations of data availability (plot to the right = example for Freya Glacier).

➡ 23 test cases with 16 experiments each

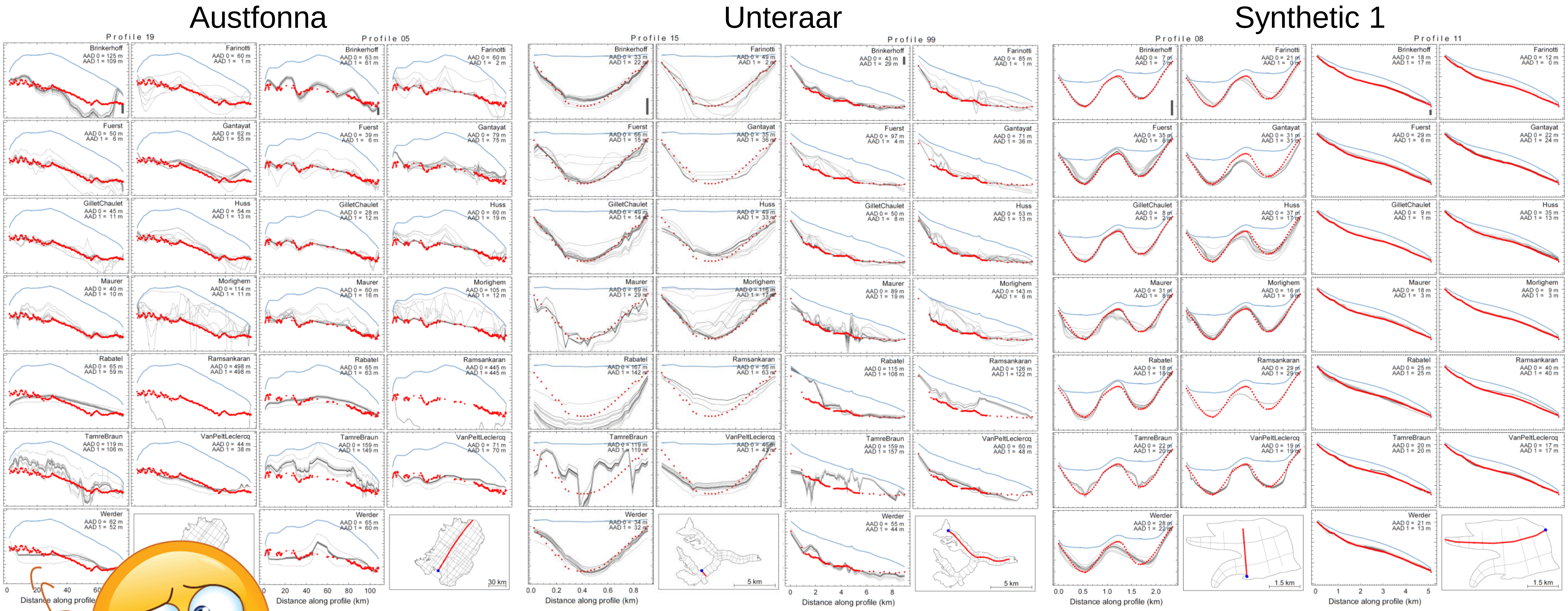


Rules of the game:

- Everyone can join (both published and unpublished models)
- Pick the test cases you like but at least the compulsory ones
- A test case is complete when providing results for all 16 exp.
- Tweak your model as you like but pretend the validation profiles not to be available when you do so

Results!

13 models, 159 individual solutions, ~500 profiles,



Maybe we need some summary statistics here?

Btw: For equivalent plots of all test cases, see the link to the Supplementary Materials on the last page.

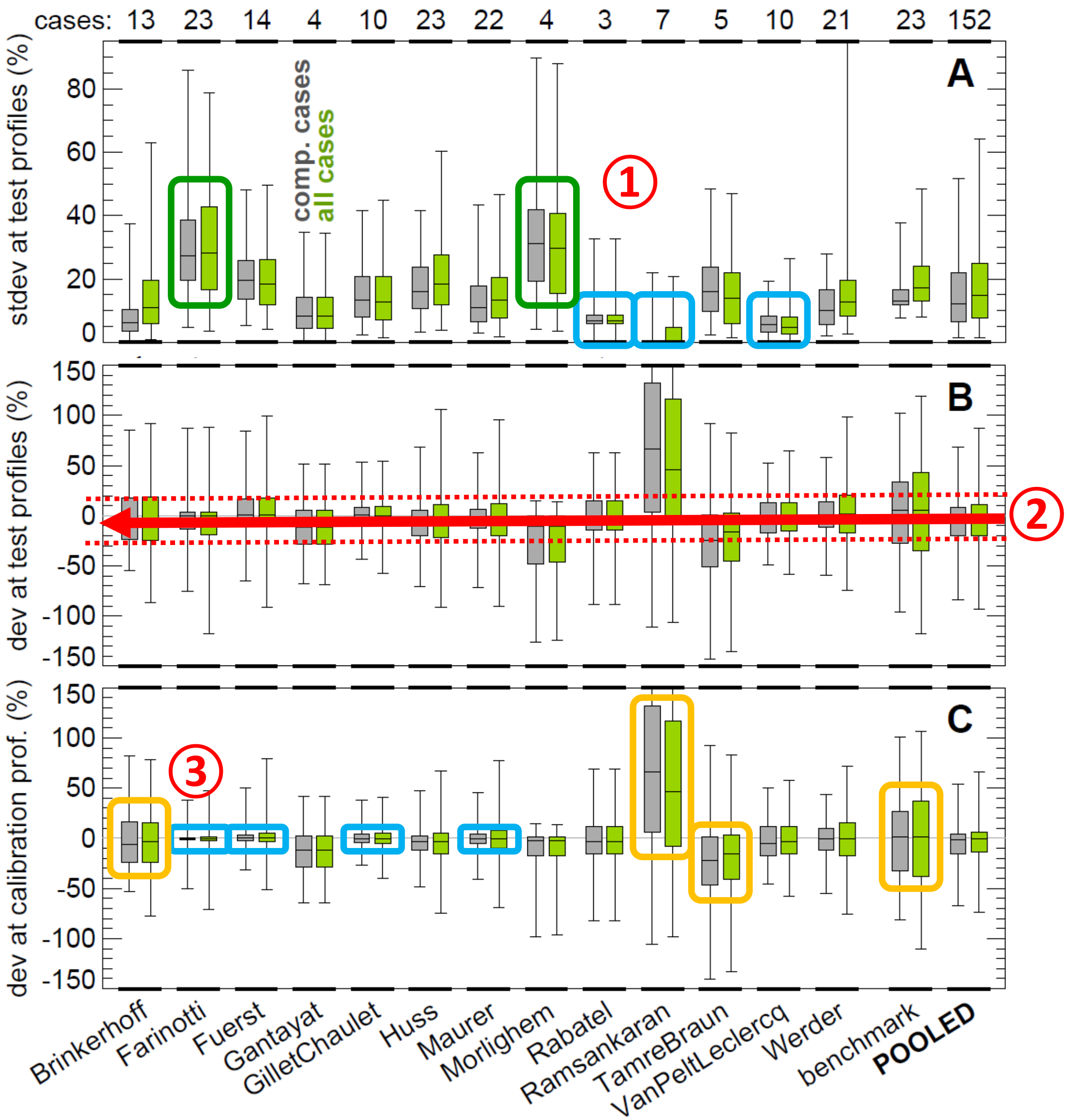
General characteristic

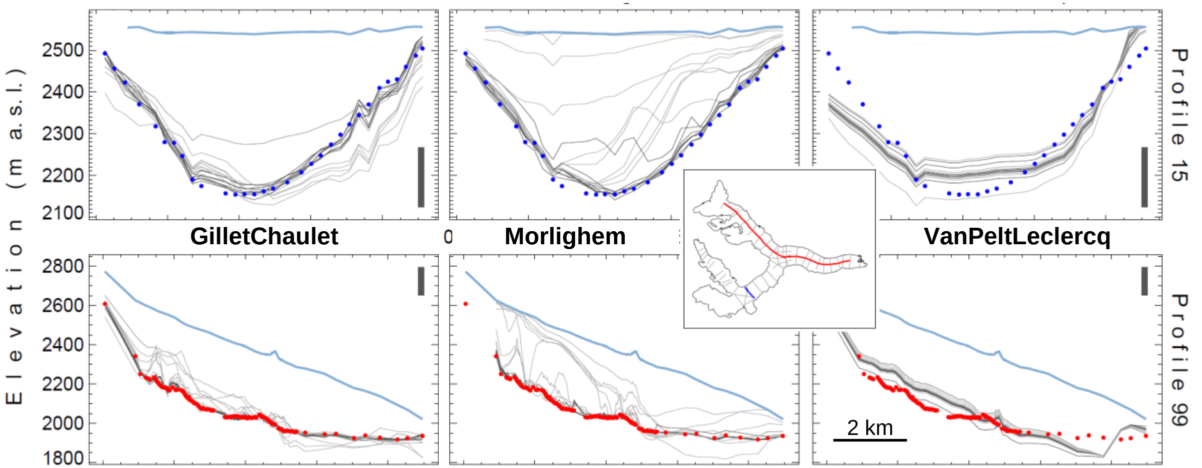
Method 1: Pool all test cases and experiments, stratify by model.
A: Model variability across experiments, for a given test case.
B: Deviations at profiles *not* seen during calibration.
C: Deviations at profiles *seen* during calibration

① Some models provide similar (σ <13%) results for all exp., others show large (σ >30%) deviations ("benchmark" has σ ≈17%).

② Models are unbiased (Δ≈–2%), with a typical point-deviation at locations not seen during calibration of ≈16%.

③ Some models aim at matching the measured thickness exactly (IQR<10%), other show some tolerance (IQR≈30-40%)





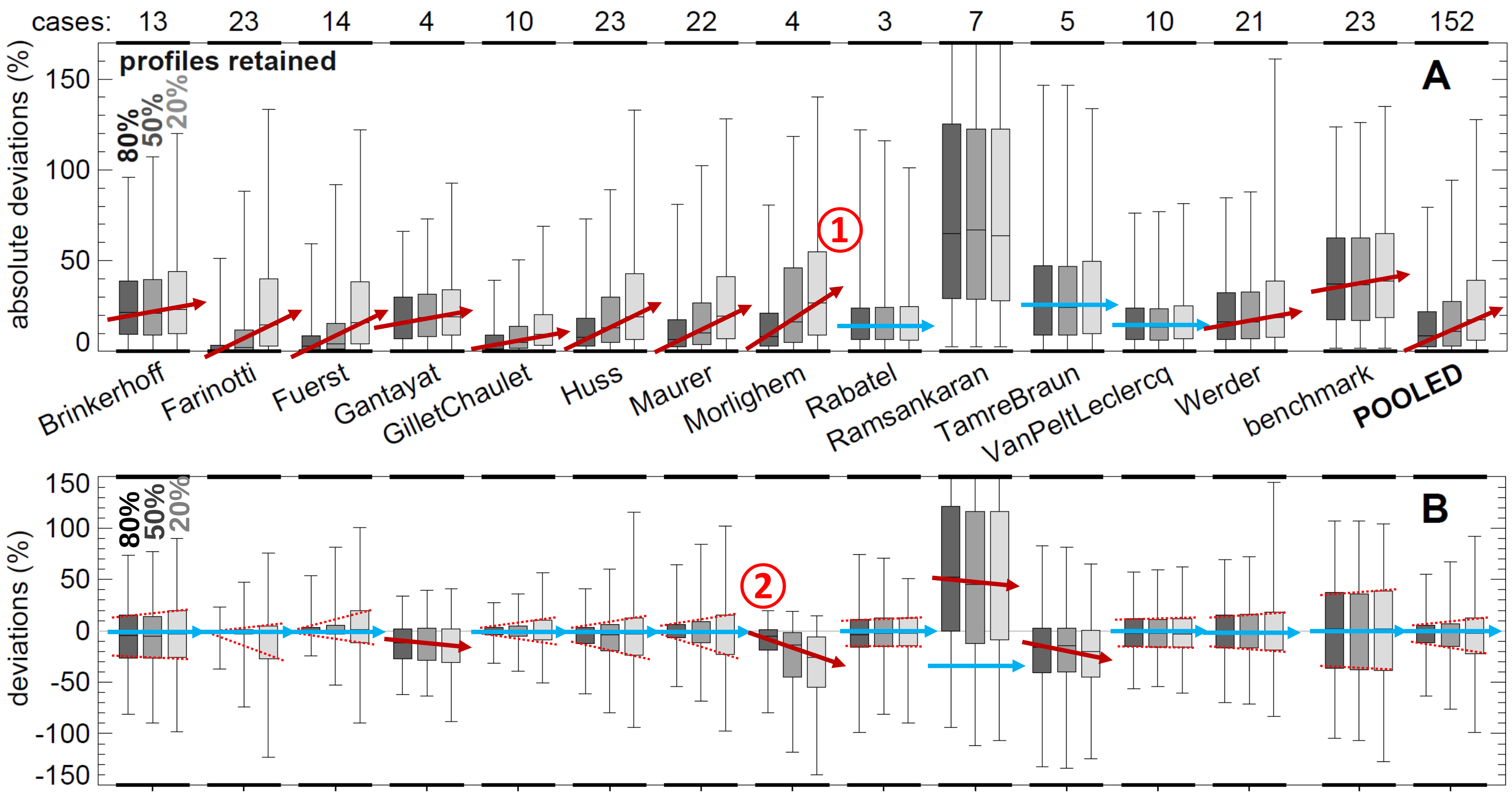
These plots are for visualizing what different σ imply.

ETH zürich

Influence of available number of profiles

Method 2: Pool all test cases, stratify by experiment and by model.
A: Absolute deviations for points *not* seen during calibration.
B: Actual deviations for points *not* seen during calibration.

1 Deviations increase if fewer measurements are available. (Deviations ≈ 8, 11, 18% for 80, 50, 20% of retained profiles, respectively)



Interpretation:
Additional information helps constraining the models, with some models being more sensitive than others.

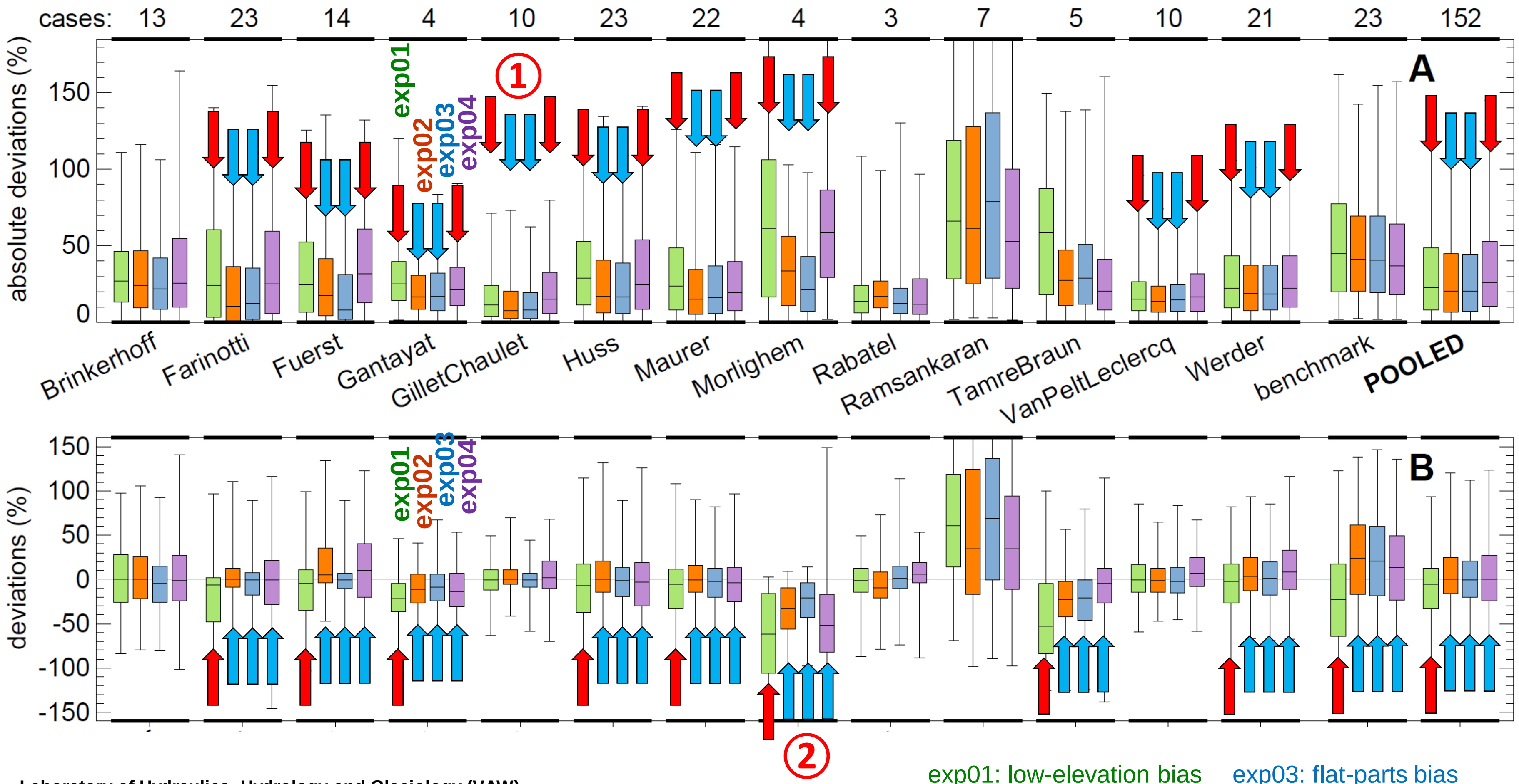
2 Most models are unbiased, even if relatively few measurements are available. Only some models drift.

Interpretation: A few observations are sufficient for estimating the total glacier volume.

ETH zürich

Influence of profile location

Method 3: Pool all test cases, stratify by experiment and by model.
A: Absolute deviations for points *not* seen during calibration.
B: Actual deviations for points *not* seen during calibration.



exp01: low-elevation bias
exp02: thickest-parts bias
exp03: flat-parts bias
exp04: longitudinal profile

① Thick- and flat-bias (conditions often coincide) resulting in ~6% lower deviations (5 models showing a difference >10%)

Interpretation: If we constrain the thickest parts, the deviations can't grow so large. (useful if we care about volume!)

② The low-elevation bias results in a ~5% thickness underestimation. (3 models showing Δ>20%)

Interpretation: The lowest elevations are the thinnest ("benchmark": Δ≈–22%). Some models also assume steady state (whilst most glaciers are actually retreating)

Influence of distance to the closest observation

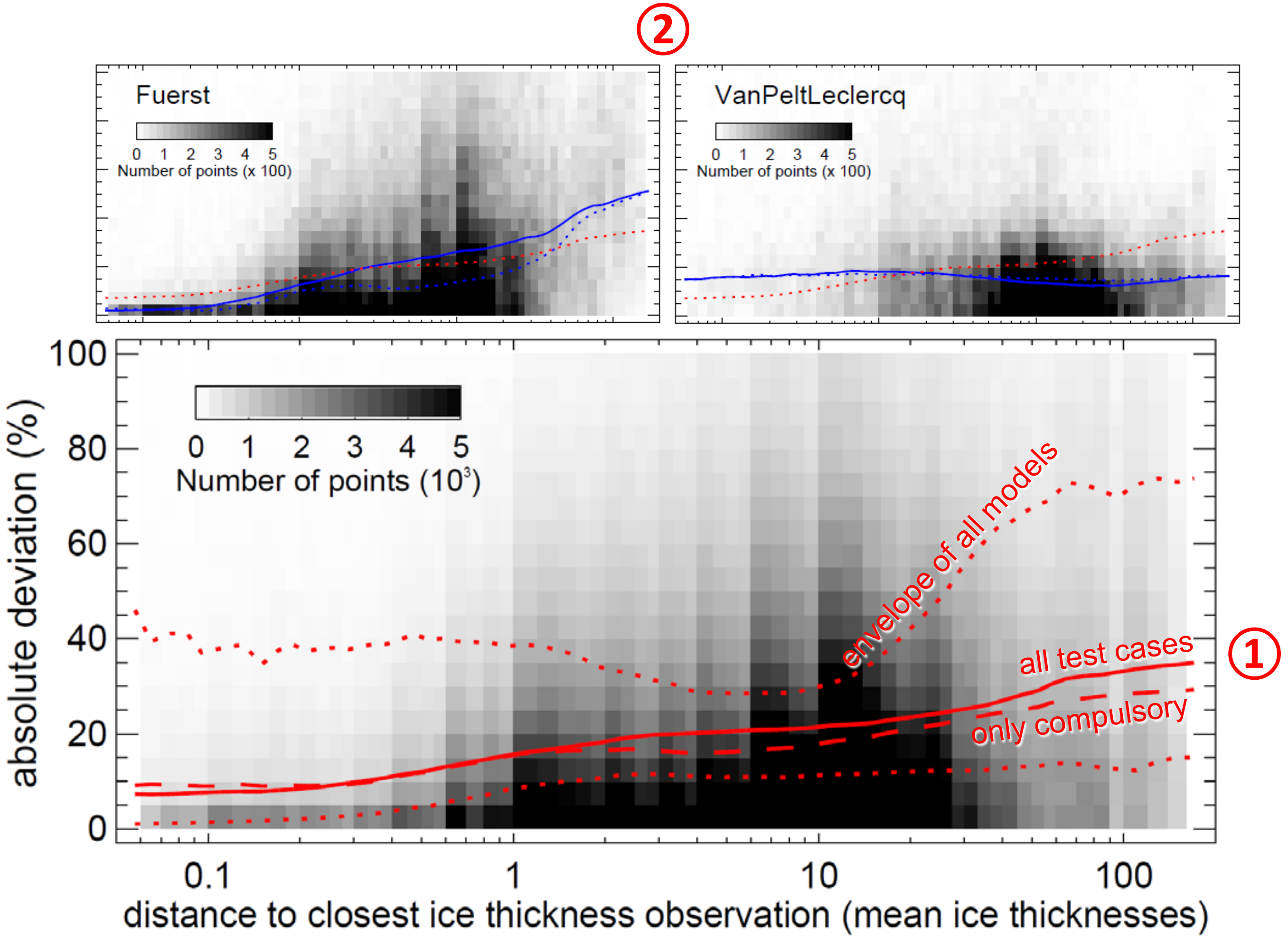
Method 4: Pool all test cases, all experiments, and all models.
Express deviations as a function of the closest point that was seen during calibration.

1 The deviations increase with the distance to the closest observation (as expected)

The increase is of roughly 8.5% for every ten-fold increase in distance (i.e. average deviations are of 8%, 16% and 21% for distances of 0.1, 1, and 10 mean thicknesses, respectively)

And if you dig deeper:

2 The function is specific to every model, i.e. the numbers given above hold true on average.



Is there a best model?

Method 5: Go back to all metrics discussed so far.

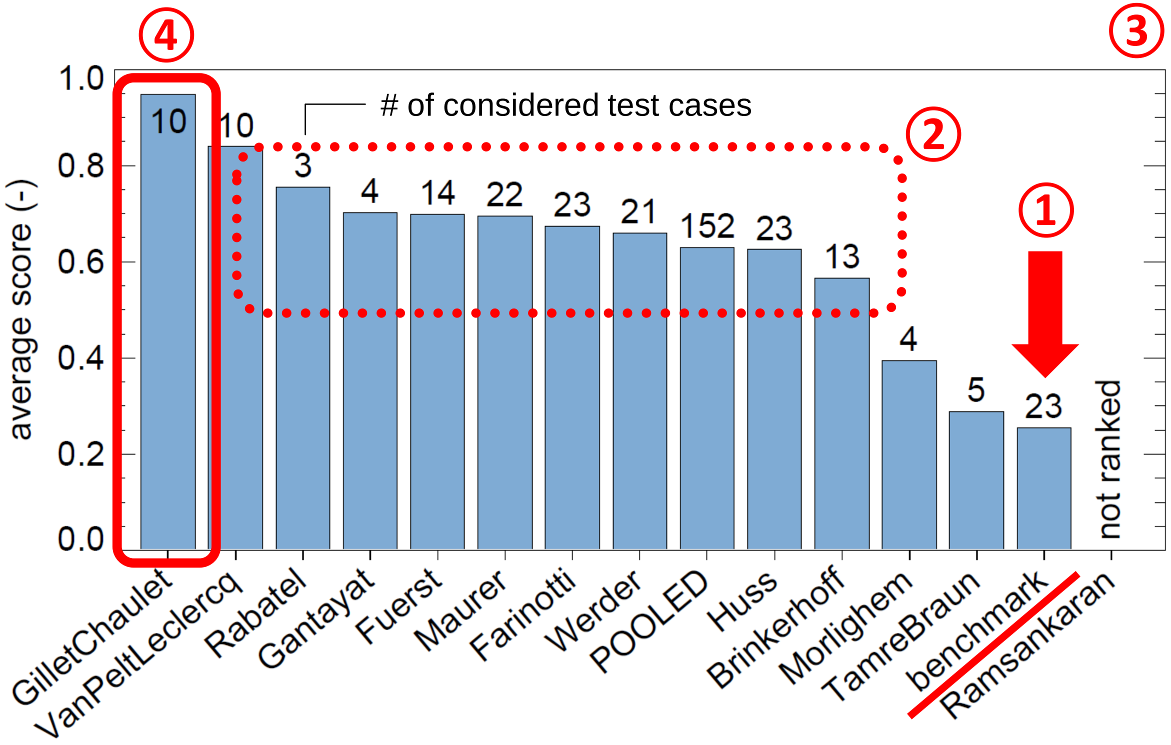
- For every metric, individually consider the (i) median, (ii) IQR, and (iii) 95% confidence interval (= *indicators*).
- For each metric and indicator, compute a score by normalizing the abs. value with the spread across models.
- For every model, compute the average score *S* (*S* = 1 (0): model is always ranked as the best (worst))

① All models outperform the "*benchmark*" model.

② Most models show a very similar performance.

③ Unfortunately, it is not possible to discern a cluster of "better" or of "less performant" models.

④ The only model that seem to consistently show a better performance is the ensemble-approach *"GilletChaulet"*.



➡ **The last result is consistent with the outcomes of ITMIX1.**

Conclusions

- ITMIX2 aimed at investigating the capabilities of ice thickness estimation models to make use of sparse thickness data.
- The experiment included 23 test cases with 16 experiments each and attracted the participation of 13 different models.
- Good news: - the models have skill!
  - the models are virtually unbiased (Δ≈–2%)
  - limited sets of thickness obs. are sufficient for constraining the volume
- For unmeasured locations, the typical deviation is of ~16% the mean ice thickness.
- The uncertainty is dependent on the location of the closest observation
  typically, an increase of 8.5% is to be expected per ten-fold increase in distance.
- It is advantageous to place the observations over the thickest parts of the glacier.
- Avoid configurations in which only the lowest elevations are surveyed.
- There is no single best model, and not even a set of best models.

  ➡ The best is to use a model ensemble!

FOLLOW US ON twitter @VAW_glaciology

Thank you for your interest!

For a **complete description** of the work, see **Farinotti et al. (2020)**, Results from the Ice Thickness Models Intercomparison eXperiment phase 2 (ITMIX2). *Frontiers in Earth Science*. DOI: https://doi.org/10.3389/feart.2020.571923

The complete **ITMIX2 dataset** as well as an extensive set of figures is **available at** https://doi.org/10.3929/ethz-b-000447116