The need for hindcasting in ice sheet modeling

or why we can get the right answer for the wrong reason

Andy Aschwanden & Constantine Khroulev

Geophysical Institute
University of Alaska Fairbanks, USA

Northwest Glaciologists Meeting, October 2012

What's the weather tomorrow?

What's the weather tomorrow?

What's the weather tomorrow?

Weather forecasting 100

B. Taylor says

$$weather(tomorrow) = \underbrace{weather(today)}_{0th \text{ order}} + \underbrace{weather'(today)\Delta t}_{1th \text{ order}}$$

Bottom line

- if you don't know the weather today, you're unlikely to get tomorrow's weather right...
- you also need to monitor changes in weather

Ice sheet “weather” forecasting 100

Because ice sheets change more slowly than the atmosphere, predicting their behavior over the coming century has more in common with short-term weather prediction:
small errors in the initial state could systematically affect a forecast throughout the 21st century.

(Arthern & Gudmundsson, 2010, J. Glaciol)

Hindcast

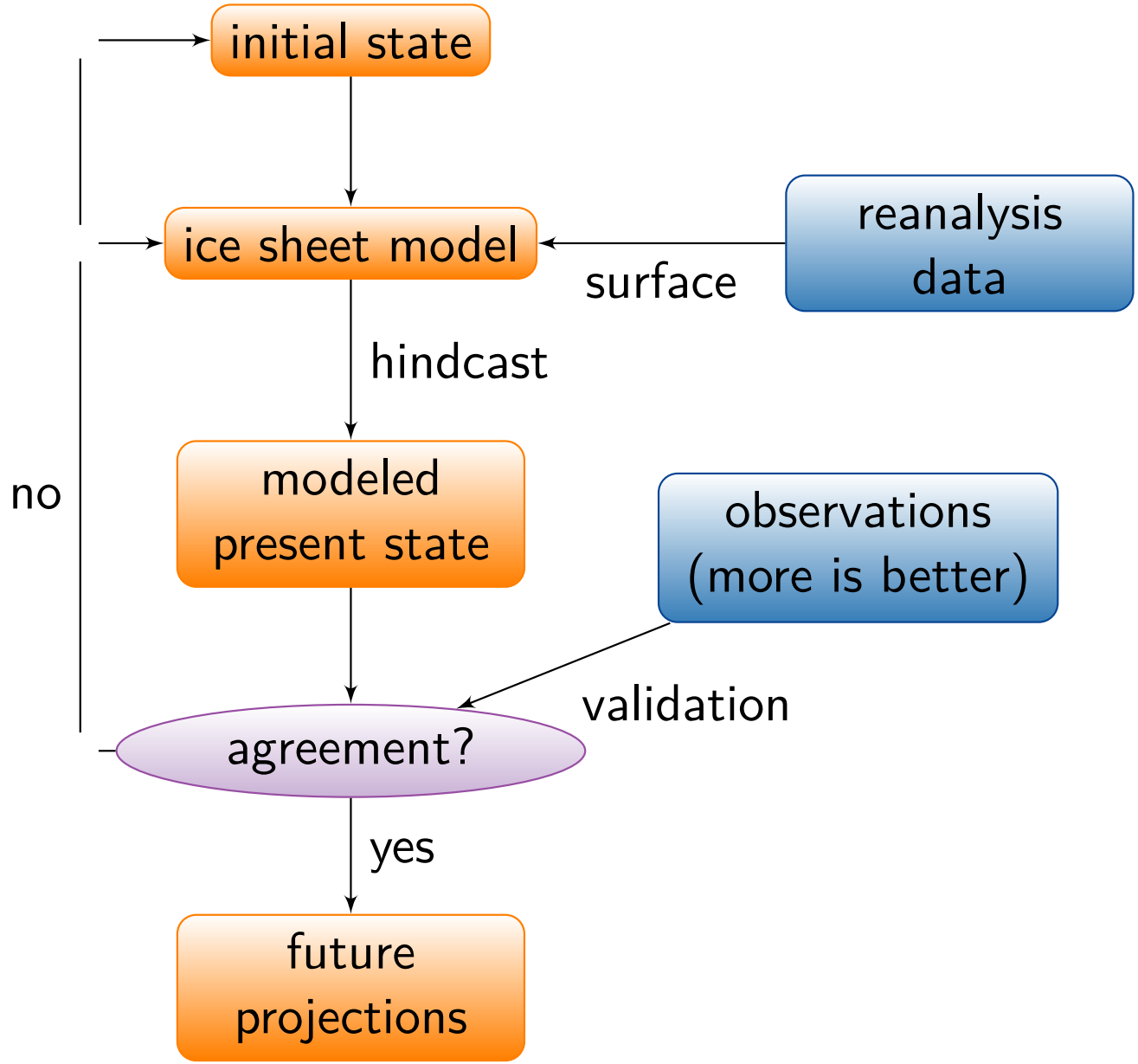
From wikipedia 

A hindcast is a way of testing a mathematical model. Known or closely estimated inputs for past events are entered into the model to see how well the output matches the known results.

Example

Entering climate forcings (events that force change) into a climate model. If the hindcast accurately showed weather events that are known to have occurred, the model would be considered successful.

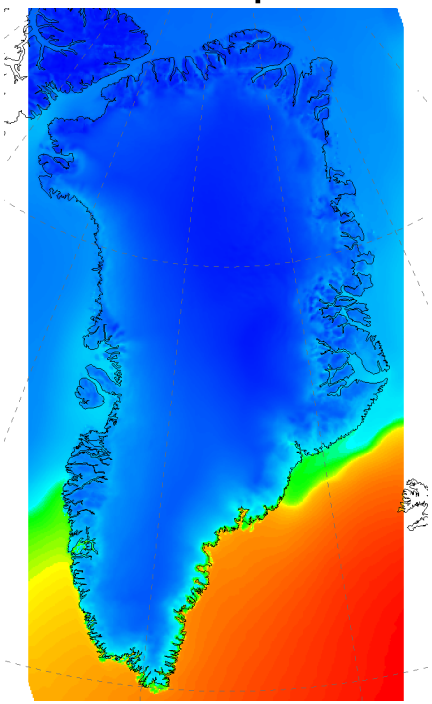
How to test your ice sheet initial states



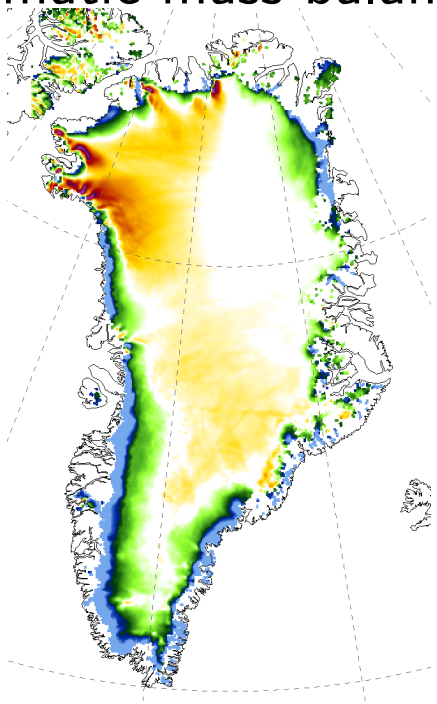
Hindcasting, the ice sheet way

- reanalysis from a regional climate model HIRHAM5 (5km) as climate forcing (yes, you can use RACMO2 too)
- time-series from 1989-2011 with monthly values of

2m air temperature



climatic mass balance



Initial states

As an example we test 3 initial states (ways to initialize a model)

- constant-climate steady-state using present-day climate
- paleo-climate has a memory of past Ice Ages
- flux-corrected paleo-climate combines paleo-climate with information about present-day ice thickness

Details about initializations are irrelevant in this talk

Validation: ice volume and ice thickness

Probably the most common validation metric is ice volume

- **X** ice thickness was used to obtain "flux-corrected" initial state, thus not available for validation

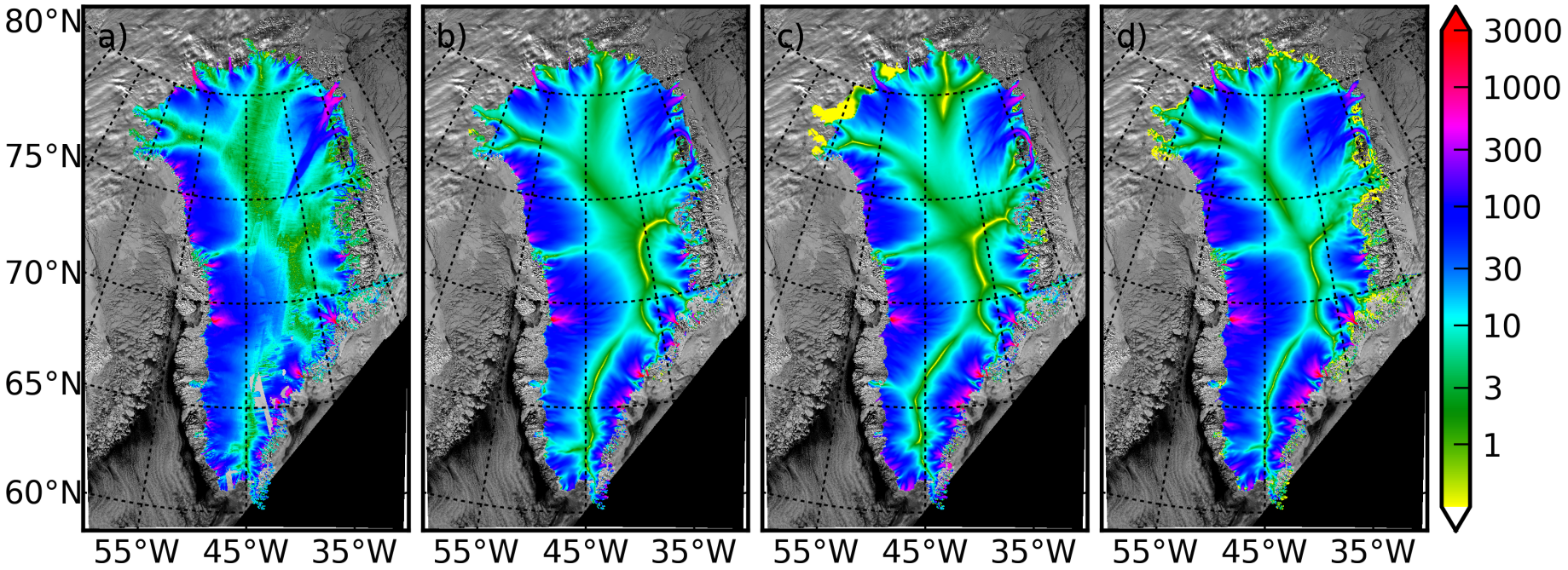
	observed	constant-climate	paleo-climate	flux-corrected
ice volume				
initial volume $[10^6\,\mathrm{km}^3]$	2.93	3.18	3.37	X
ice thickness				
avg difference [m]		99	121	X
rms difference [m]		199	244	X

observed ice thickness is from Griggs & Bamber (unpublished)

- how well do we know ice thickness?

$\Rightarrow$ we will see that ice volume is a very weak metric

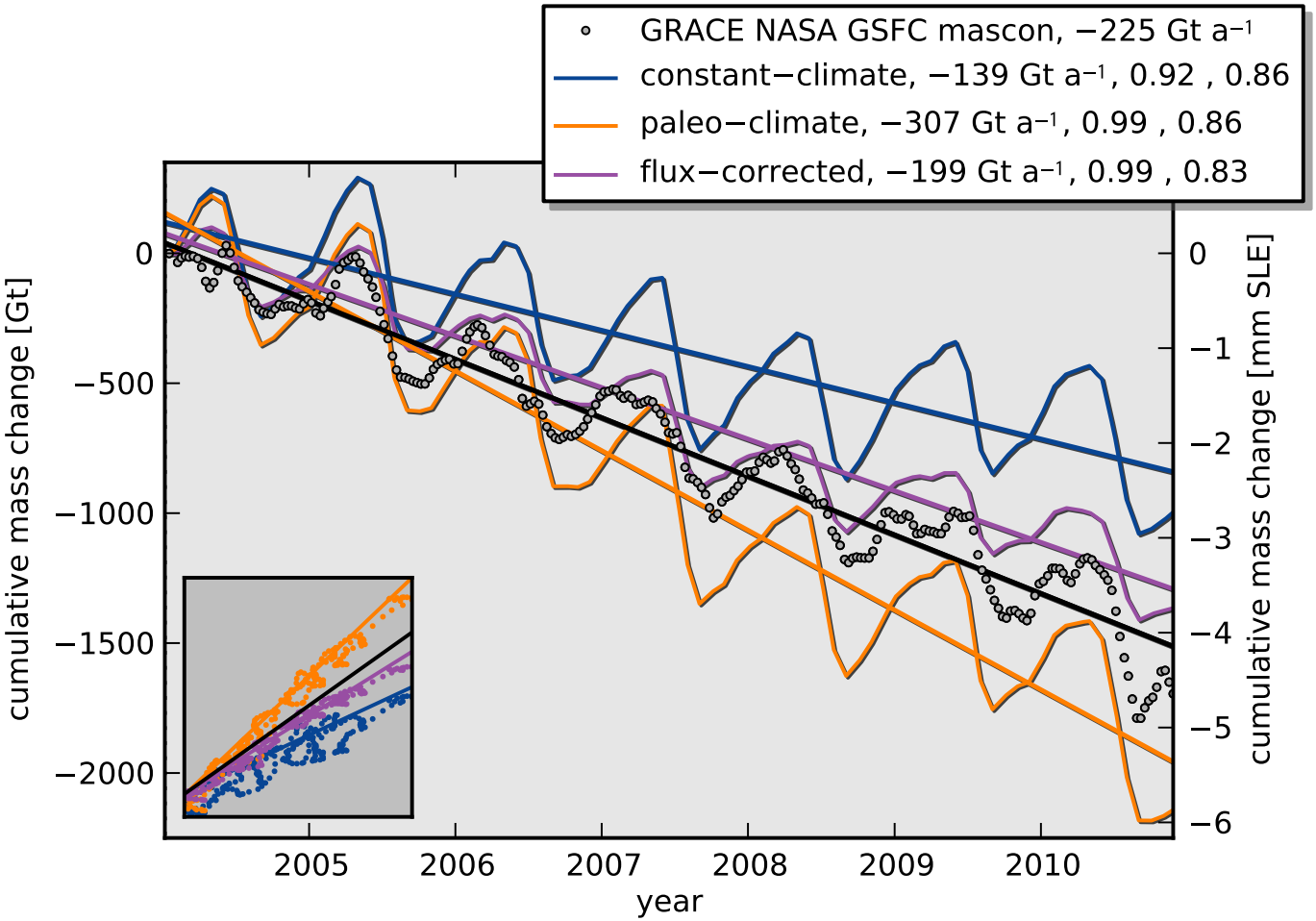
Validation: surface speeds



a) SAR (Joughin et al., 2010); b) constant-climate; c) paleo-climate; d) flux-corrected. Values in m/a.

surface speeds were not used for initialization

Validation: cumulative total mass changes



mass changes are well reproduced

Validation: mass fluxes from 1990 to 2010

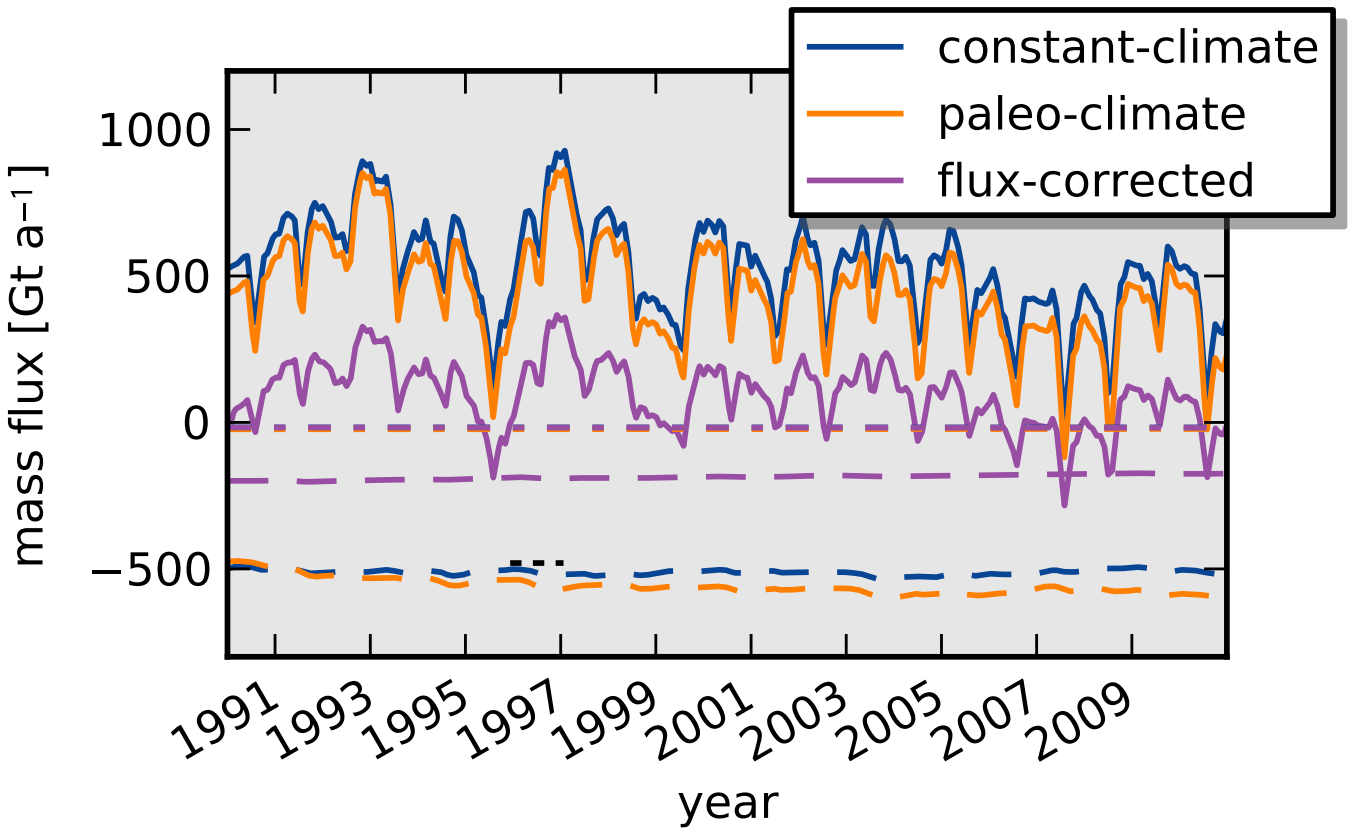
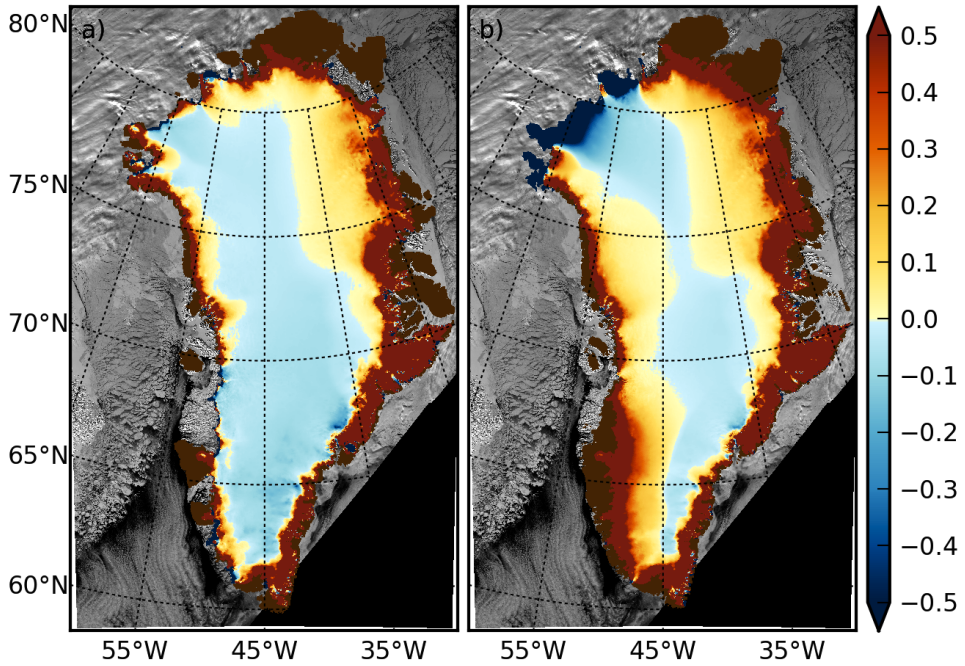


Figure: Climatic mass balance (solid line) and ice discharge (dashed line). For comparison the ice discharge estimate for 1996 *(Van den Broeke et al, 2009)* is shown (black dotted line).

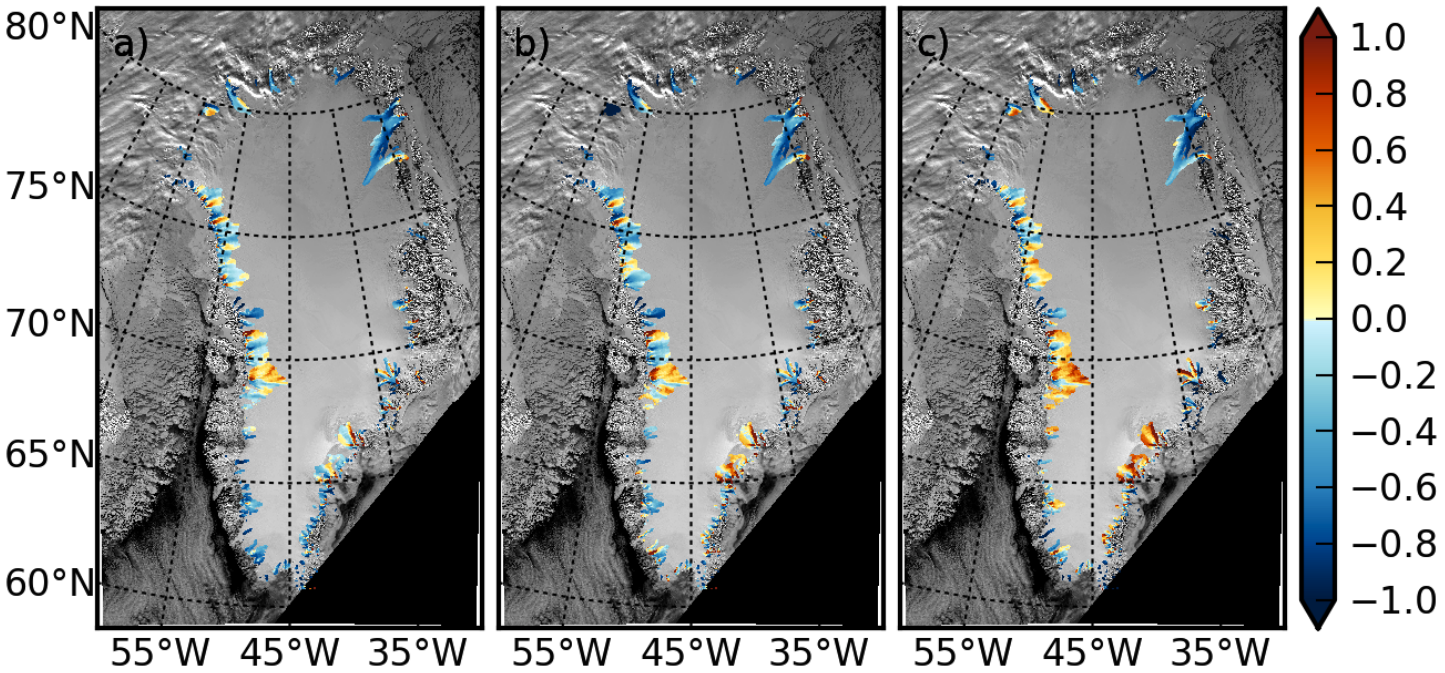
Validation: relative ice thickness differences



a) constant-climate; b) paleo-climate. Observed ice thickness is from Griggs & Bamber (unpublished)

- ice thickness was used to obtain "flux-corrected" initial state, thus not available for validation

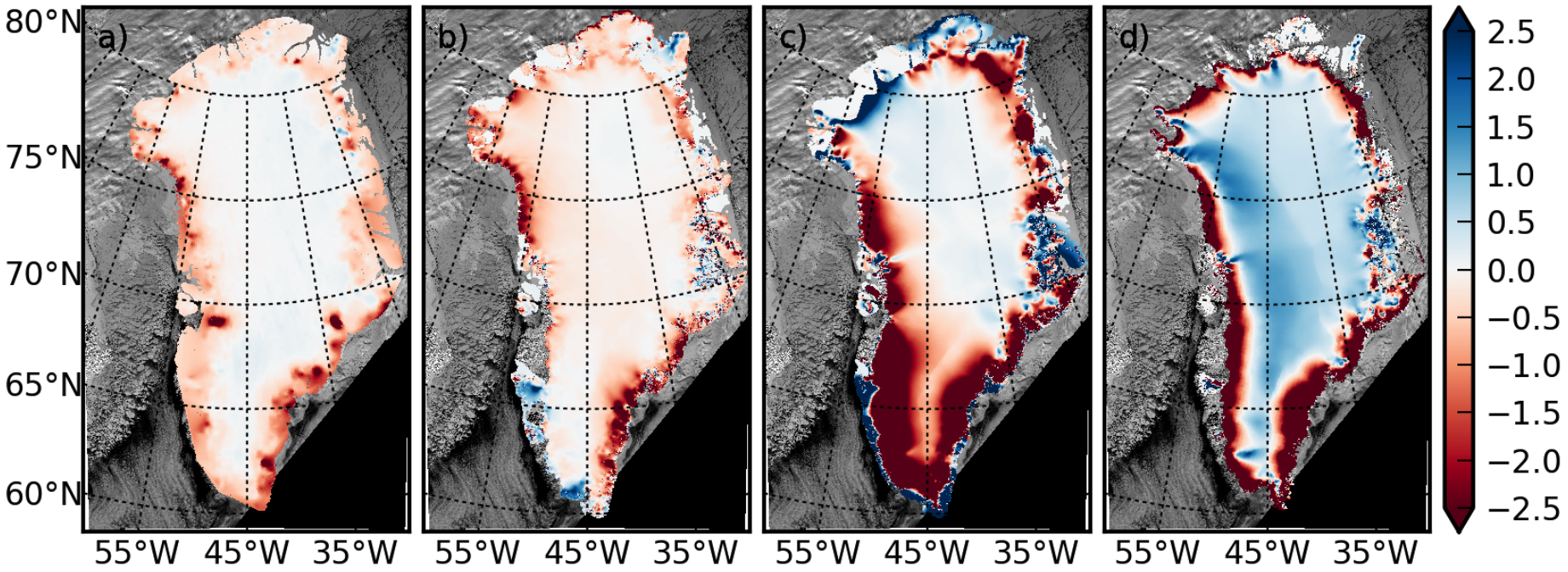
Validation: relative surface speeds differences



a) constant-climate; b) paleo-climate; c) flux-corrected.

flow speeds are under-estimated in most outlet glaciers

Validation: surface elevation change 2003–2009



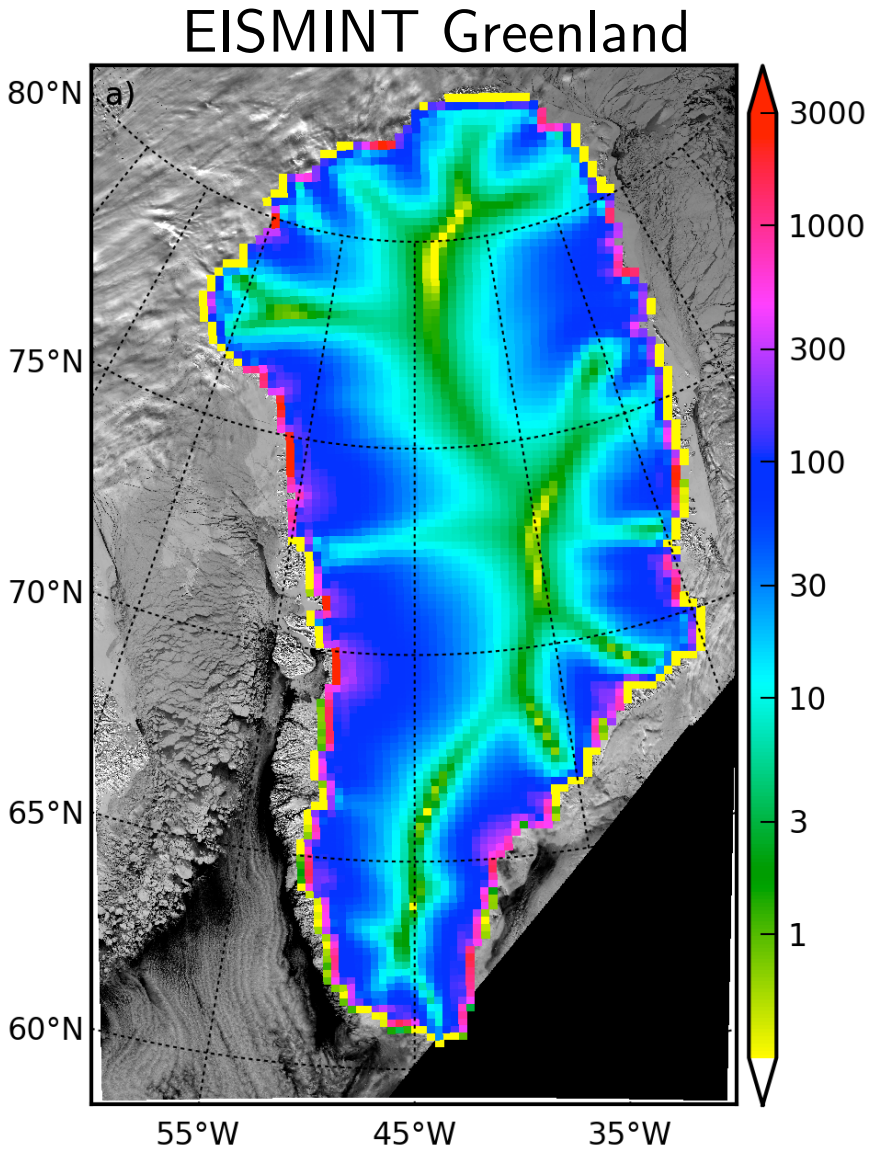
a) ICESat (Sørensen, 2011); b) constant-climate; c) paleo-climate; d) flux-corrected. Values in m.

- not very well reproduced

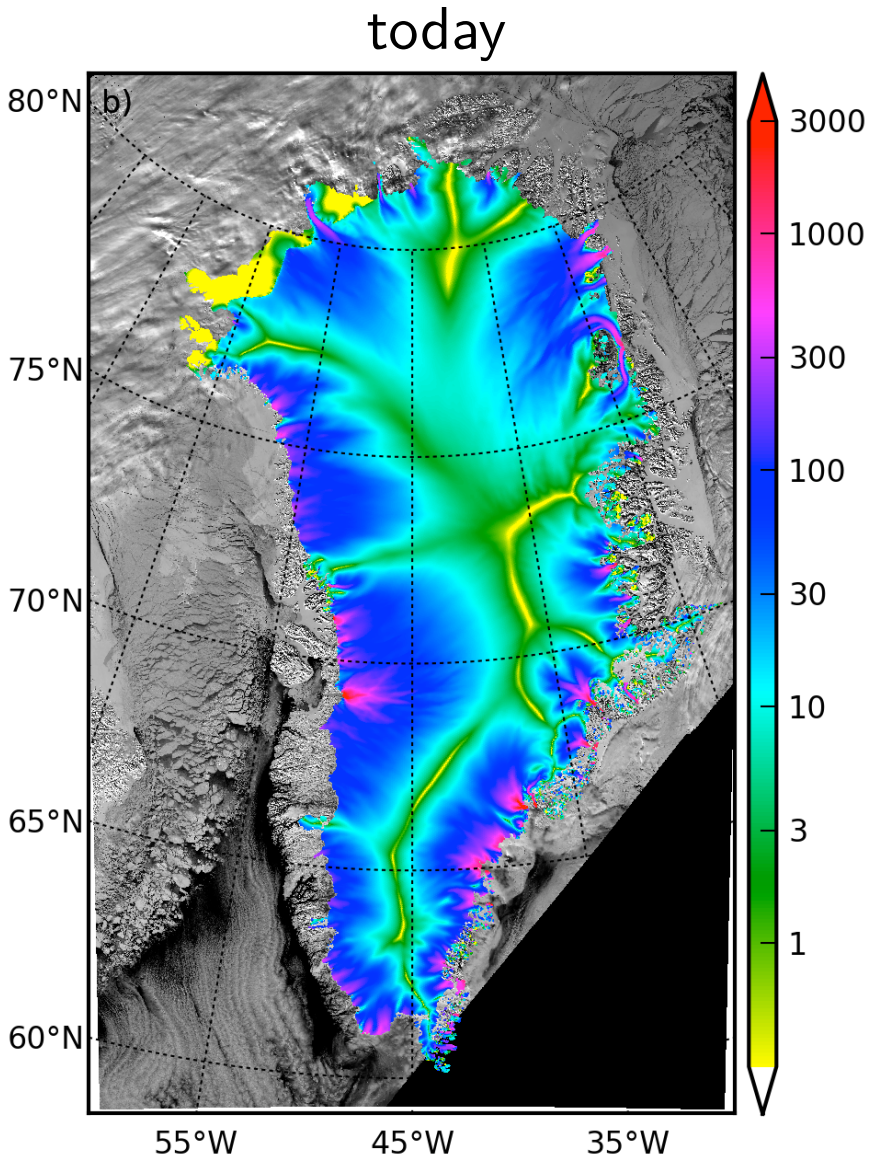
Summary

- using inappropriate metrics, it's possible to get the right result for the wrong reason
- hindcasting is an excellent strategy to expose model limitations and unrealistic environmental forcings
- do you really trust prognostic simulations obtained with models performing poorly in hindcasts?

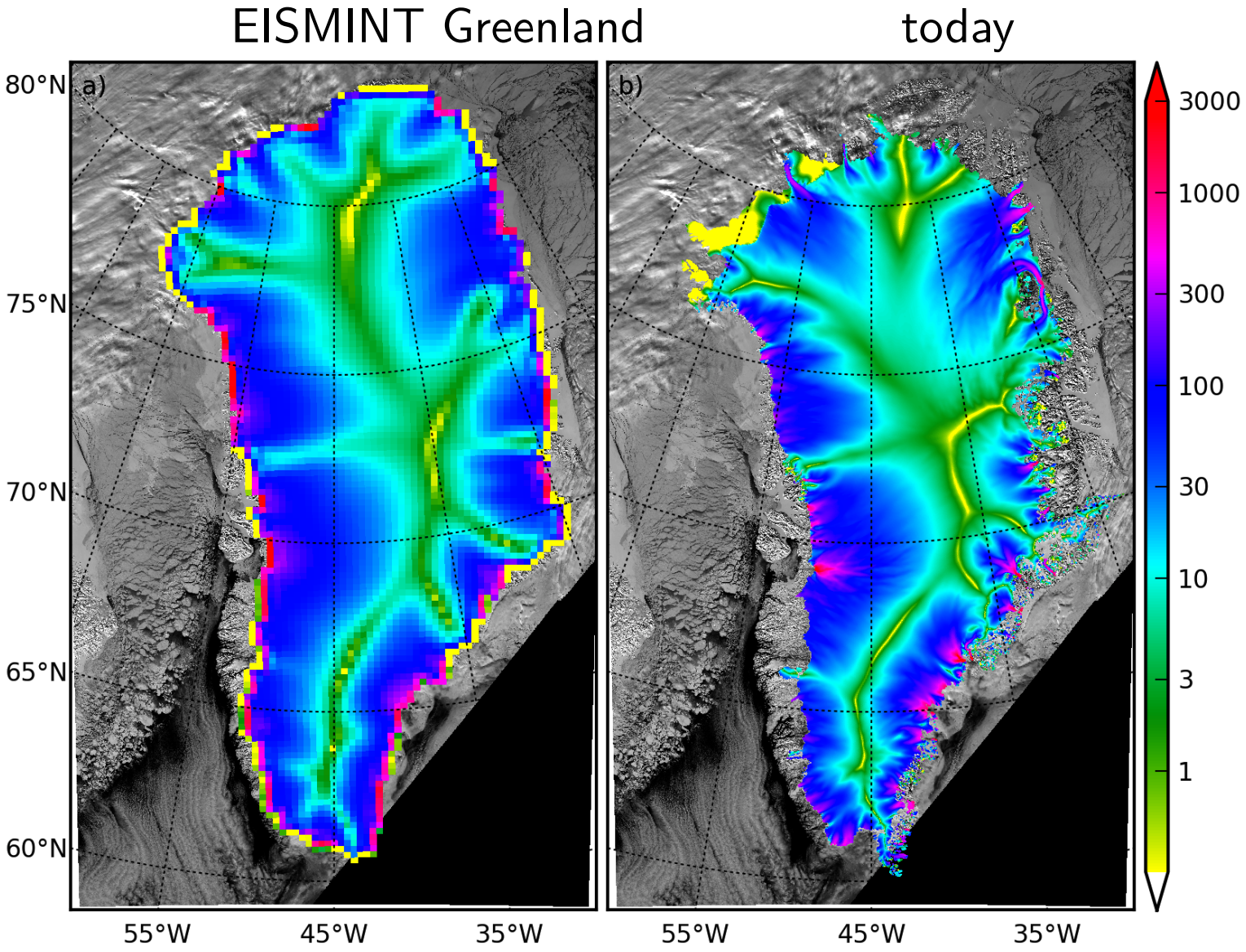
We've come a long ways...



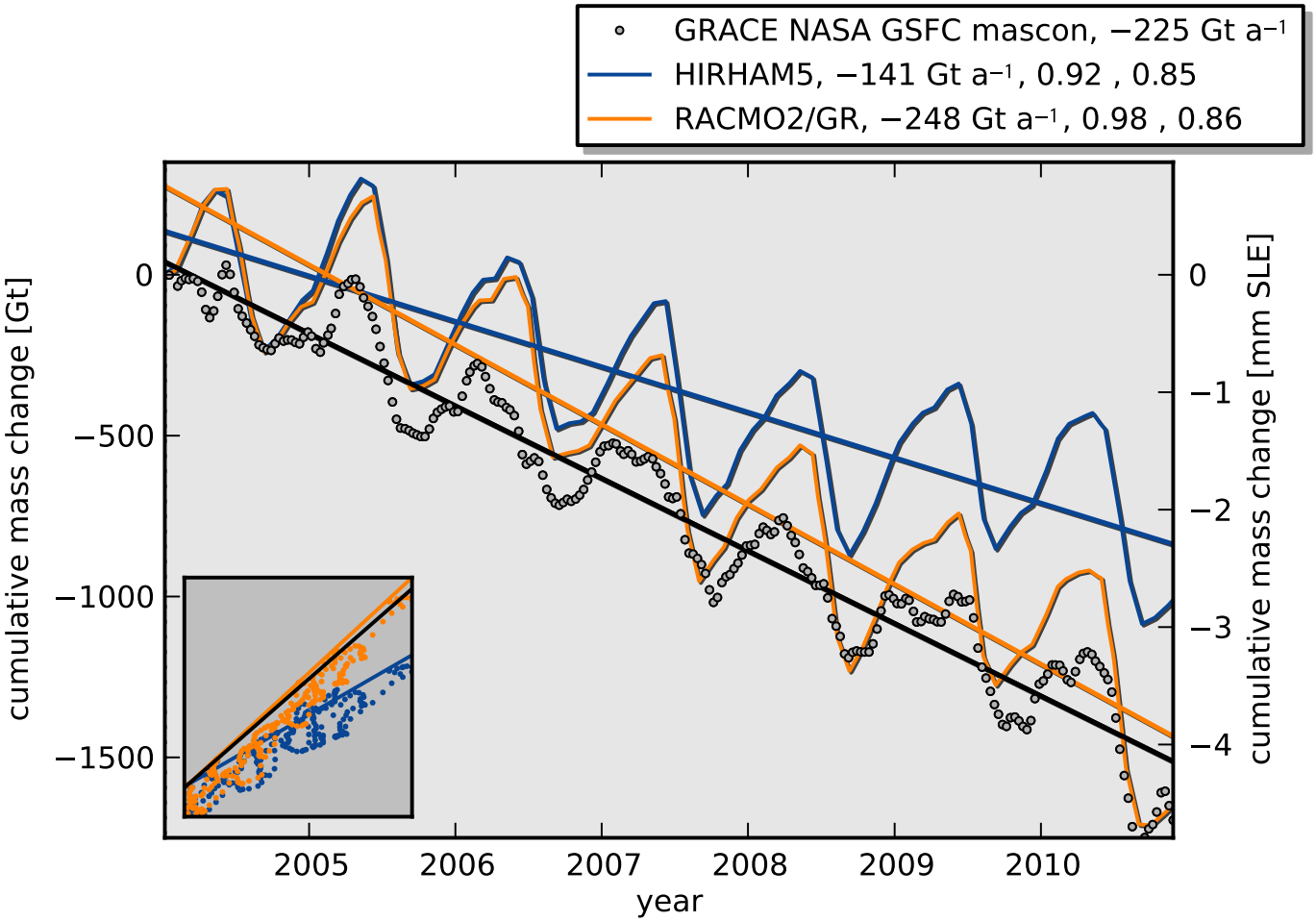
We've come a long ways...



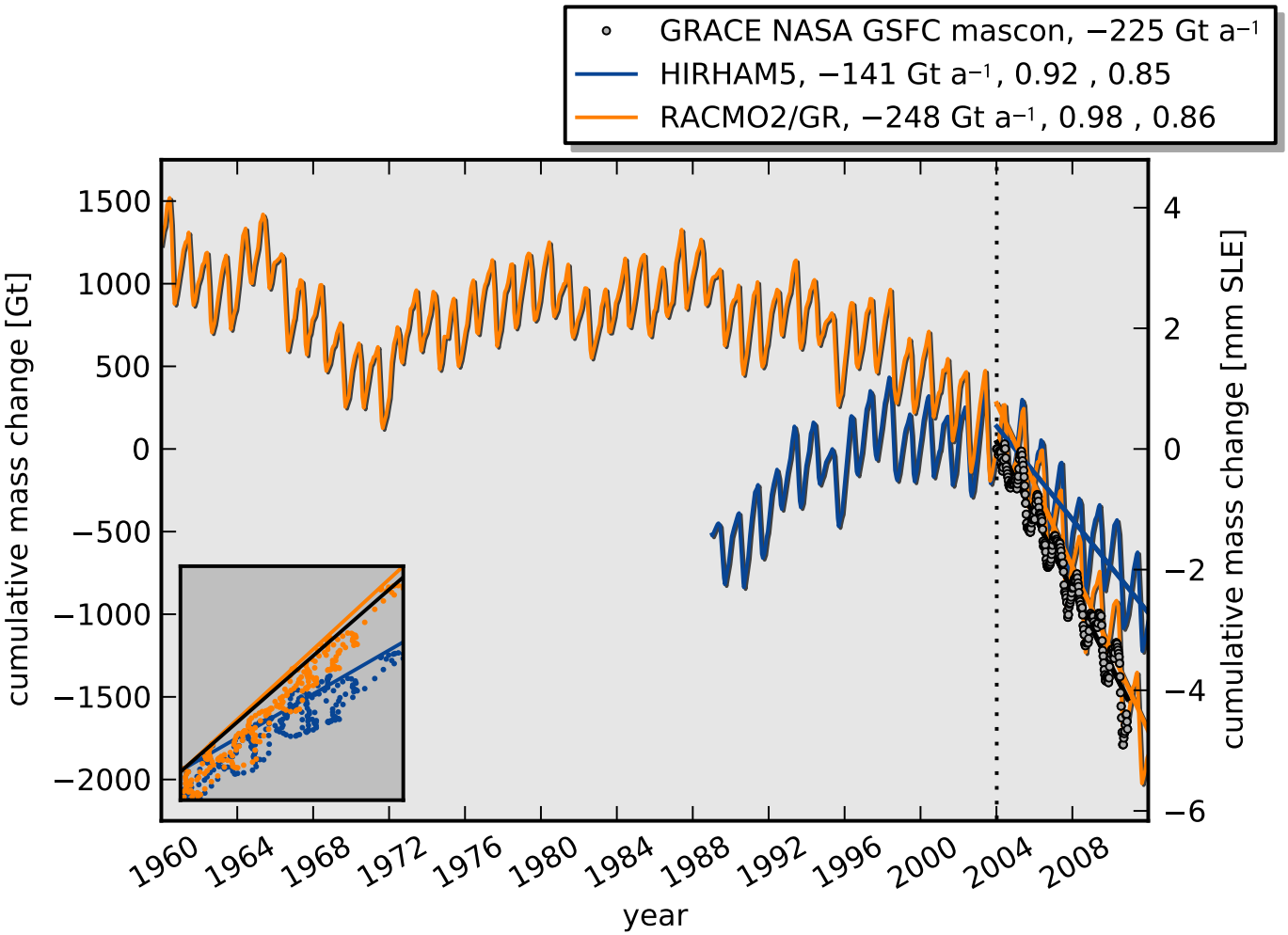
We've come a long ways. . .



HIRHAM5 vs RACMO2/GR



HIRHAM5 vs RACMO2/GR



Thoughts on metrics

To quantify the predictive skill, we define an objective function J,

$$J = \sum_i \alpha_i \left\| \Psi_i^{\mathrm{obs}} - \Psi_i^{\mathrm{m}} \right\|_p , \tag{1}$$

- "i" is an observable (e.g. surface elevation, mass change, surface speed)
- "obs" (observations), "m" (modeled state)

we seek to penalize the use of small numbers of observables