An Ice Sheet model validation framework for the Greenland ice sheet

S. Price[1], M. Hoffman[1], J. Bonin[6], I. Howat[4], T. Neumann[2], J. Saba[2],
I. Tezaur[5], J. Guerber[2], D. Chambers[6], K. Evans[3], J. Kennedy[3], J. Lanaerts[7],
W. Lipscomb[1], M. Perego[6], A. Salinger[5], R. Tuminaro, M. van den Broeke[7], S. Nowicki[2]

[1]Los Alamos National Laboratory
[2]NASA Goddard Space Flight Center
[3]Oak Ridge National Laboratory
[4]Ohio State University
[5]Sandia National Laboratories
[6]University of South Florida
[7]Utrecht University

Supported by DOE Office of Science ASCR & BER through SciDAC

U.S. DEPARTMENT OF ENERGY
Office of Science

NASA

SciDAC
Scientific Discovery through Advanced Computing

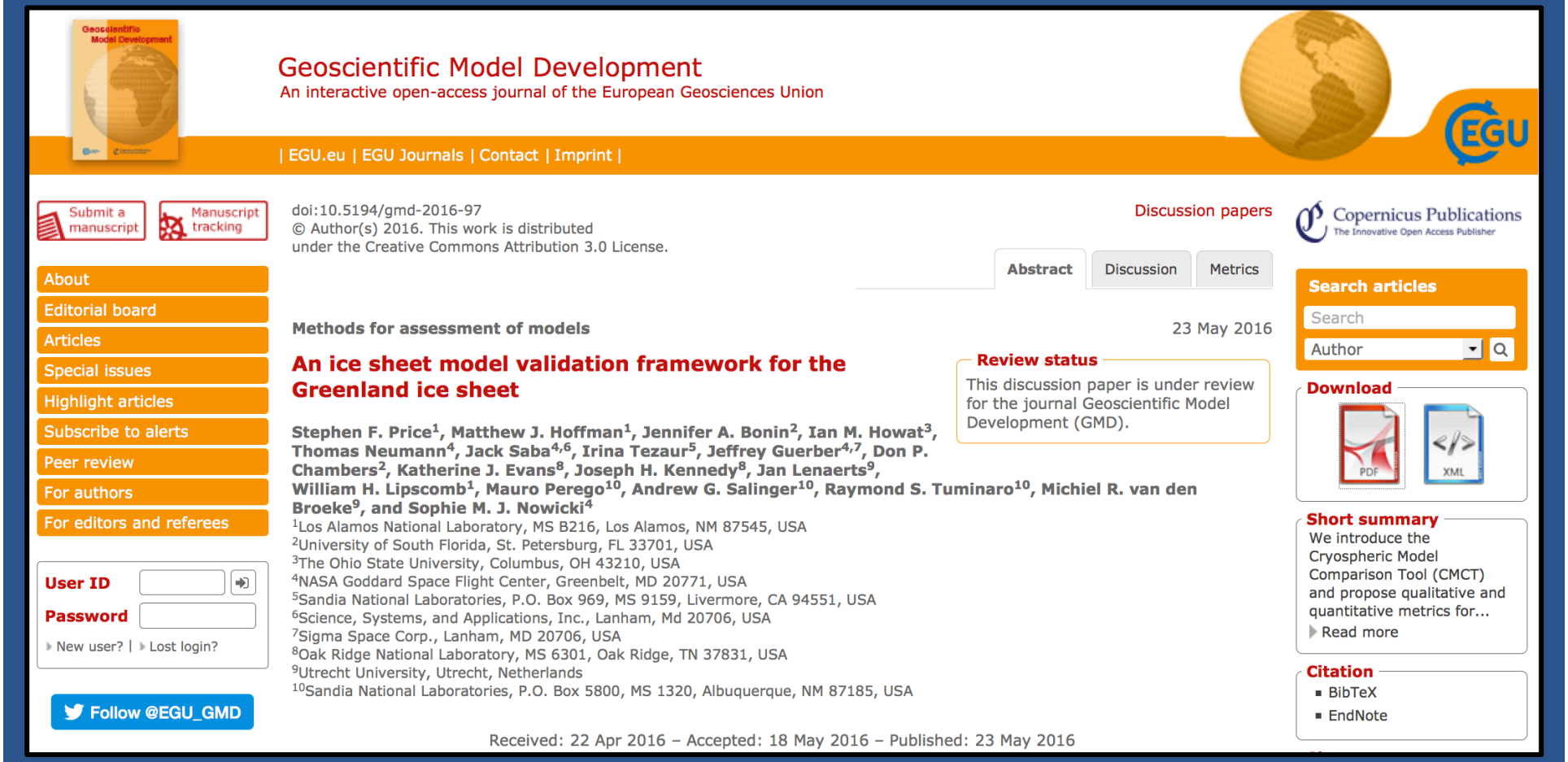

Motivation / Concept
Components
Results
Summary & Future Work

this

***again*?**

(yawn)

GMDD paper in review



Geoscientific Model Development
An interactive open-access journal of the European Geosciences Union

| EGU.eu | EGU Journals | Contact | Imprint |

Submit a manuscript | Manuscript tracking

About
Editorial board
Articles
Special issues
Highlight articles
Subscribe to alerts
Peer review
For authors
For editors and referees

User ID
Password
New user? | Lost login?

Follow @EGU_GMD

doi:10.5194/gmd-2016-97

Discussion papers

Abstract | Discussion | Metrics

Methods for assessment of models

23 May 2016

An ice sheet model validation framework for the Greenland ice sheet

Stephen F. Price[1], Matthew J. Hoffman[1], Jennifer A. Bonin[2], Ian M. Howat[3], Thomas Neumann[4], Jack Saba[4,6], Irina Tezaur[5], Jeffrey Guerber[4,7], Don P. Chambers[2], Katherine J. Evans[8], Joseph H. Kennedy[8], Jan Lenaerts[9], William H. Lipscomb[1], Mauro Perego[10], Andrew G. Salinger[10], Raymond S. Tuminaro[10], Michiel R. van den Broeke[9], and Sophie M. J. Nowicki[4]

[1]Los Alamos National Laboratory, MS B216, Los Alamos, NM 87545, USA
[2]University of South Florida, St. Petersburg, FL 33701, USA
[3]The Ohio State University, Columbus, OH 43210, USA
[4]NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA
[5]Sandia National Laboratories, P.O. Box 969, MS 9159, Livermore, CA 94551, USA
[6]Science, Systems, and Applications, Inc., Lanham, Md 20706, USA
[7]Sigma Space Corp., Lanham, MD 20706, USA
[8]Oak Ridge National Laboratory, MS 6301, Oak Ridge, TN 37831, USA
[9]Utrecht University, Utrecht, Netherlands
[10]Sandia National Laboratories, P.O. Box 5800, MS 1320, Albuquerque, NM 87185, USA

Received: 22 Apr 2016 – Accepted: 18 May 2016 – Published: 23 May 2016

Review status
This discussion paper is under review for the journal Geoscientific Model Development (GMD).

Copernicus Publications
The Innovative Open Access Publisher

Search articles
Search
Author

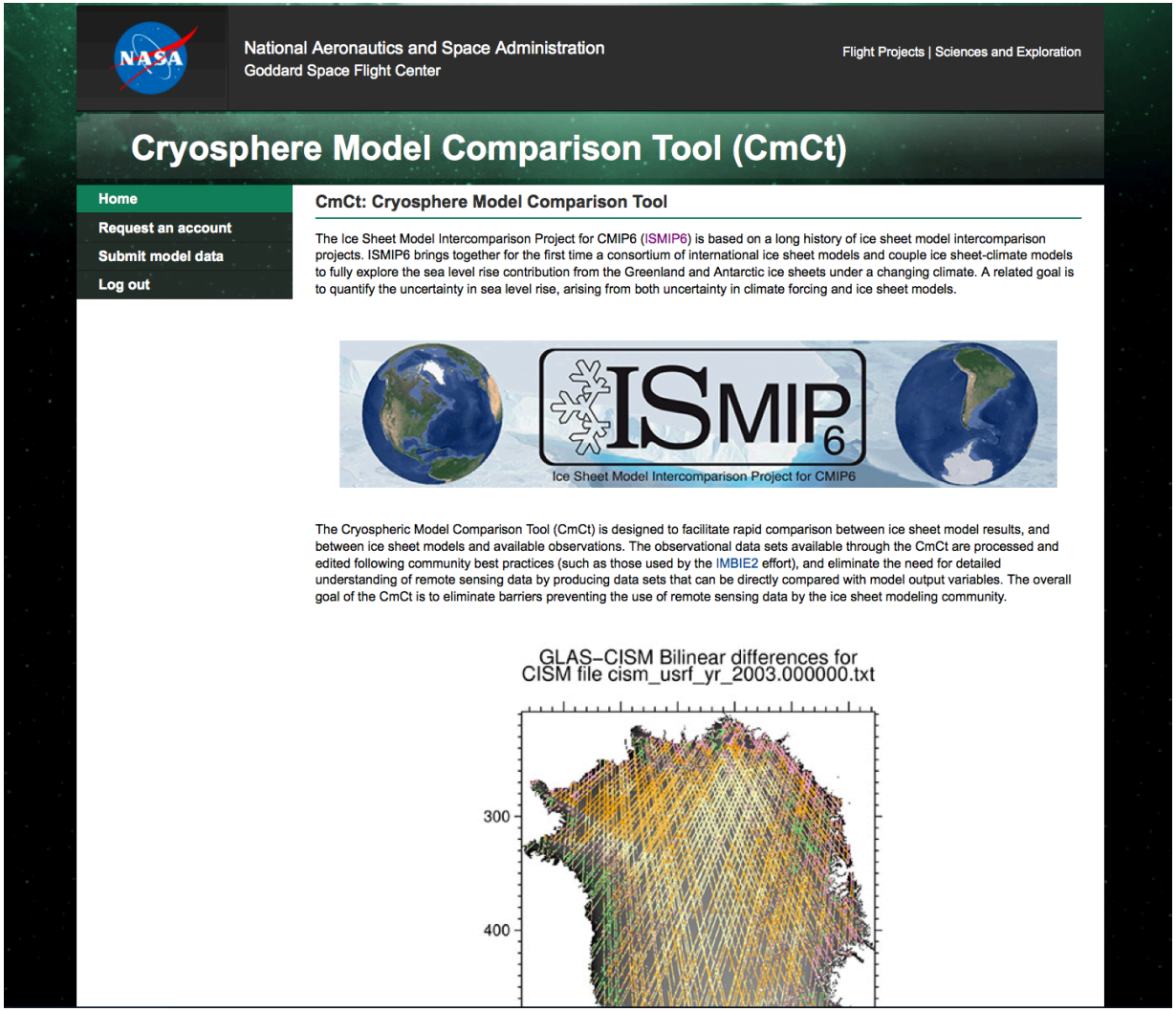
Download
PDF | XML

Short summary
We introduce the Cryospheric Model Comparison Tool (CMCT) and propose qualitative and quantitative metrics for...
Read more

Citation
- BibTeX
- EndNote

(two positive reviews so far)

Validation software is online and ready for testing by the community



https://ggsghpcc.sgt-inc.com/cmct/


Motivation / Concept
Components
Results
Summary & Future Work

Motivation

There are currently (up to) 2 decades of large-scale satellite observations of Greenland ice sheet geometry change:

ICESat1:	2003 – 2009
GRACE:	2002 – 201? (ongoing)

Future missions will extend these observational time series:

ICESat2:	2017 – 20??
GRACE “follow-on”:	2017 – 20??
GRACE2	2020’s - ?

These data can and should be used for ice sheet model *validation***, but no framework currently exists for doing so.

** *validation:* How well do our models represent the real ice sheet?

Concept

Run ice sheet model over some specified time period for which ICESat and / or GRACE observations exist

Process model output for comparison to these observations

Process observations for comparison to model output

Evaluate model performance relative to observations:

ICESat : ice sheet surface elevation [state comparison]

GRACE : rate of mass change [trend comparison]

Calculate *metrics* to quantify model performance (e.g., to gauge improvement as new dynamics, physics, boundary conditions, higher-resolution are added)

Concept

The Main goal in *this* work is to demonstrate whether or not we can use these observations to help validate models (as opposed to actually arguing that we have validated a particular model)

Questions:

(1) Can we develop standard *metrics* for comparing models with these observations?

(2) Can we demonstrate that these metrics are useful for distinguishing between relatively worse and relatively better model simulations?

Motivation / Concept

Components

Results

Summary & Future Work

Components: Observations

Validation

ICESat: 2003 – 2009

GRACE: 2003 – 2011 (CSR Release-05)

Model Forcing

RACMO2[1] (1960-present):

- monthly SMB (applied as anomalies)

Outlet Glacier Flux[2] (1990-present):

- mean-annual flux at grounding line

[1] van Angelen et al. (*Surv. Geophys.*, 2013)

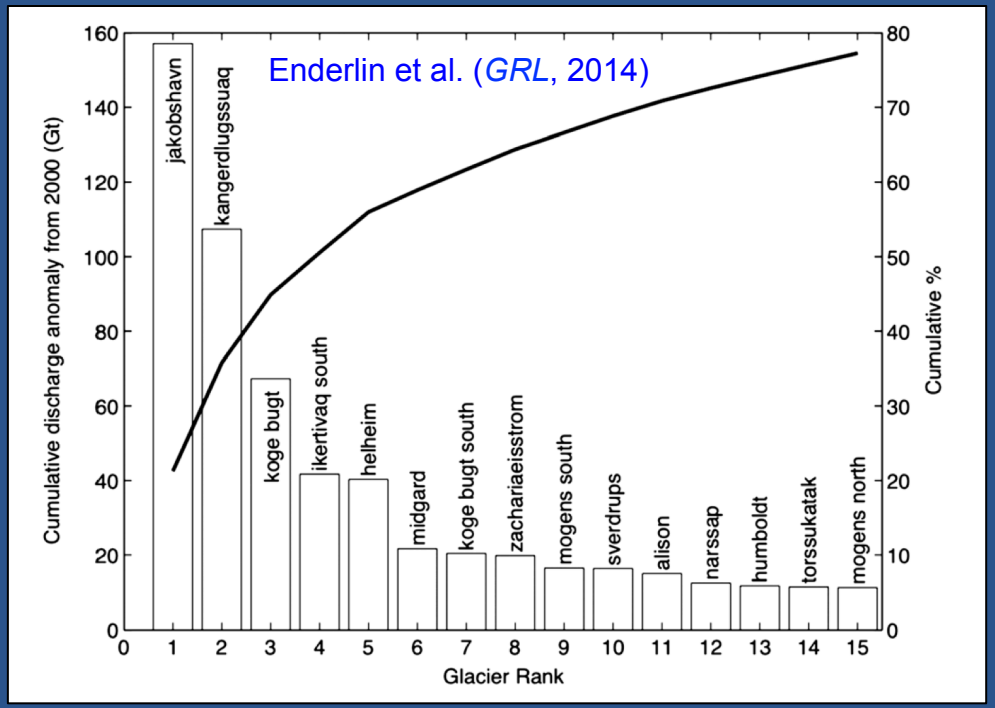
[2] Enderlin et al. (*GRL*, 2014)

Components: Model Forcing

RACMO[1] SMB over time period of interest is well validated over Greenland

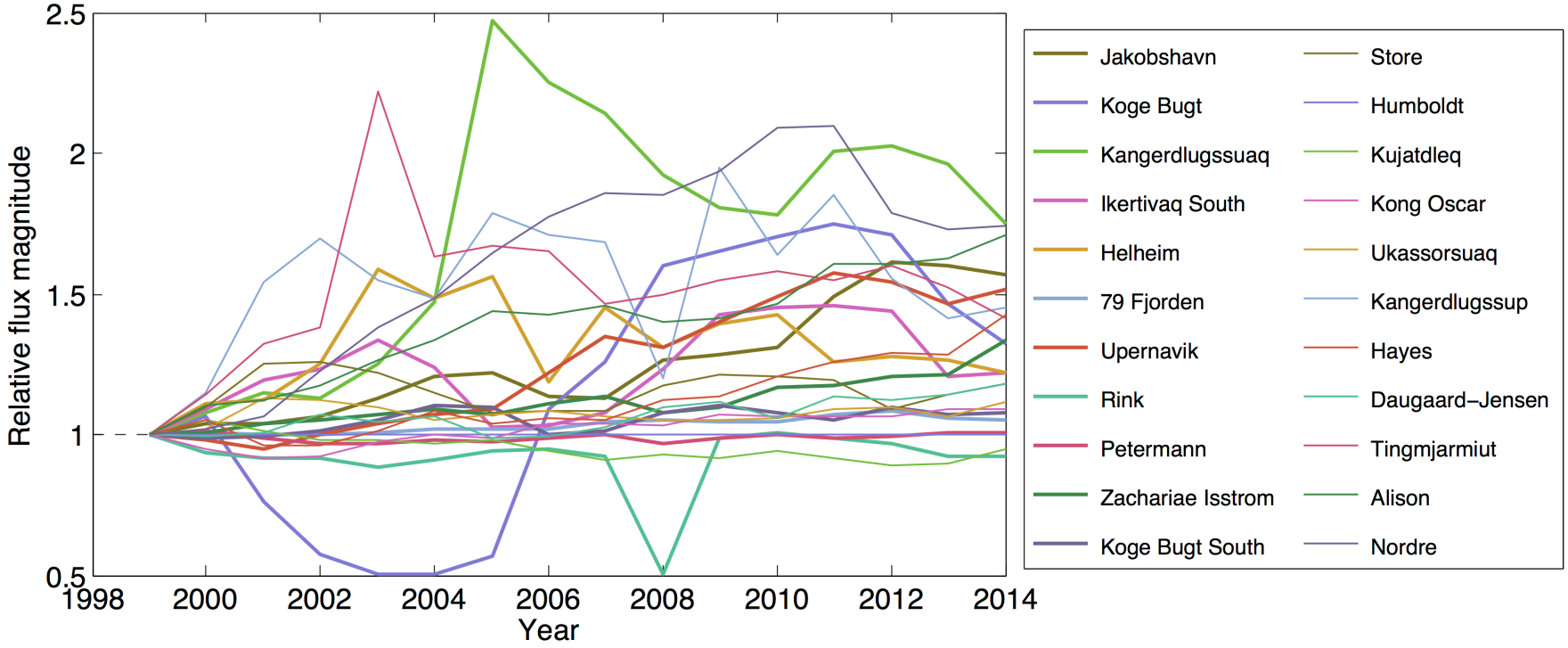
Dynamic thinning over same time (1990-present) is well captured by the flux time series from ~15 outlet glaciers[2]

We take these datasets as the "truth" (or at least our best guess) and use them as model forcing



[1] van Angelen et al. (*Surv. Geophys.*, 2014)
[2] Enderlin et al. (*GRL*, 2014)

Outlet Glacier Flux Forcing Time Series



Components: Models

Persistence Model:

- hold geometry constant in time
- useful for testing against altimetry observations *only*

SMB-only Model:

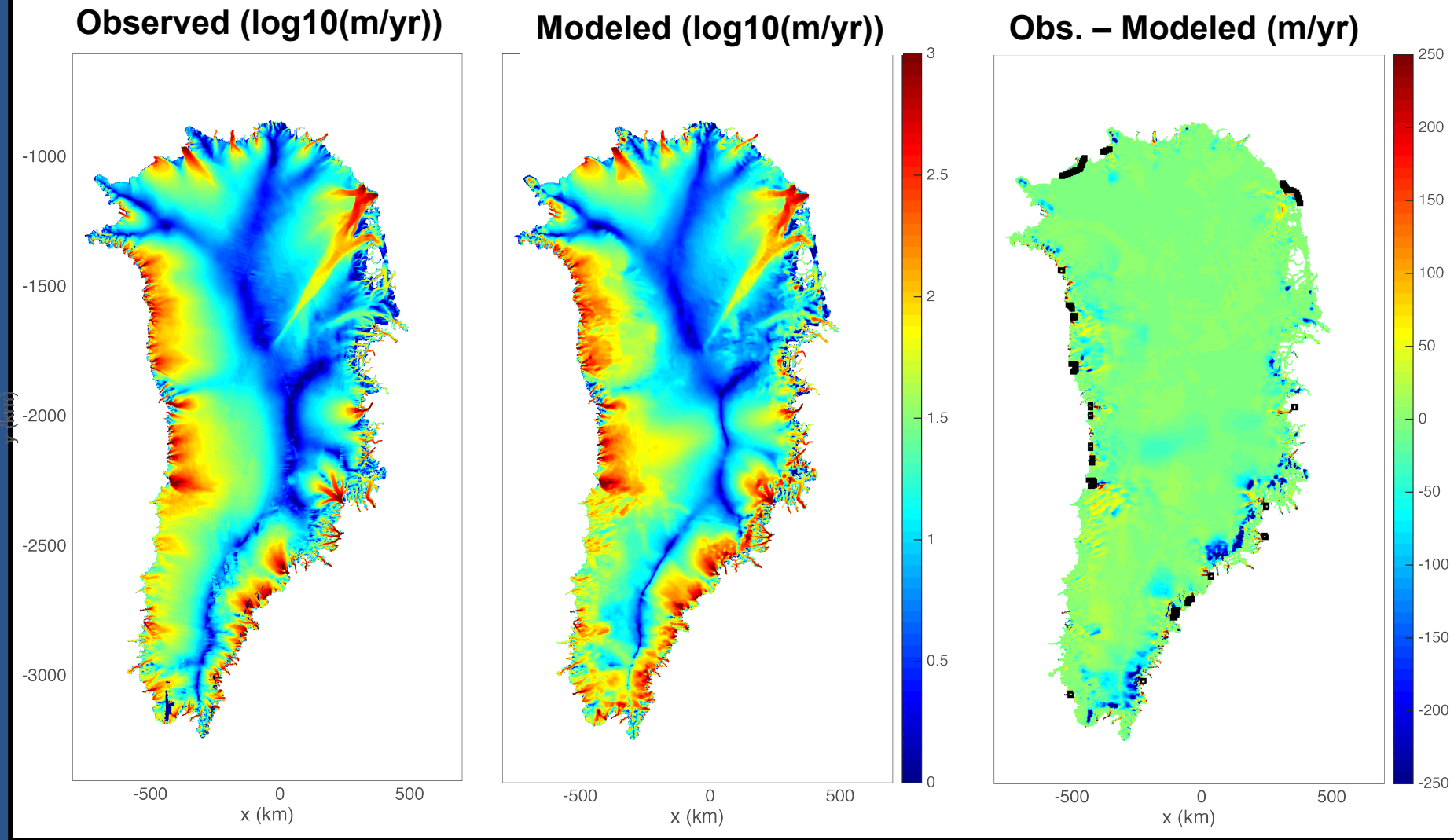
- Apply RACMO2 SMB time series to dH/dt
- include SS “discharge” using 1960-1990 mean SMB

Dynamic Model:

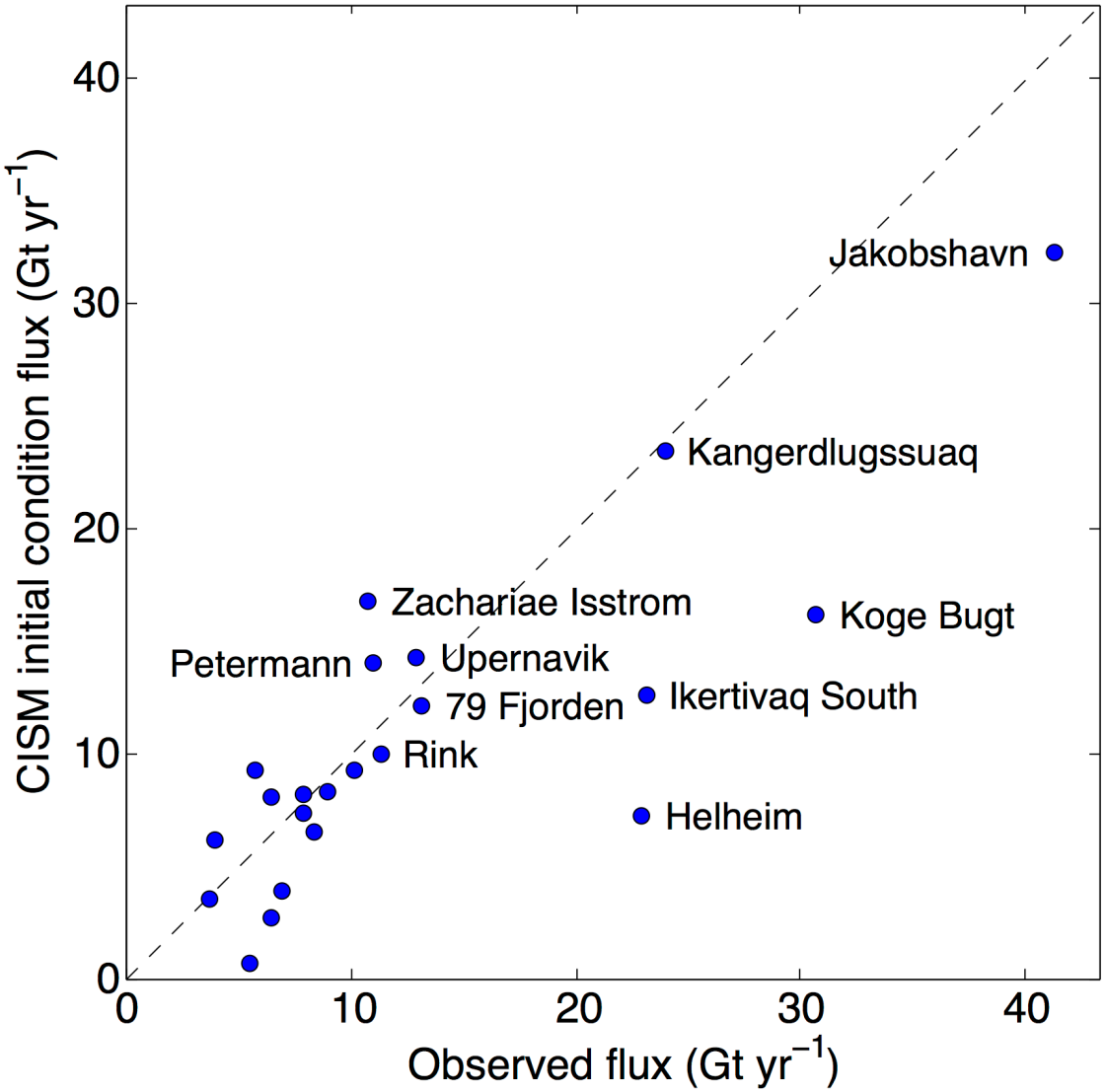
- CISM2.0[1] + FELIX-FO[2]
- parallel, 3d, first-order Stokes approximation
- coupled to CISM 2.0 as external dycore (hex. mesh)
- forced at boundaries using outlet glacier flux time series

[1] http://github.com/CISM/cism/ [2] Tezaur et al. (*GMDD*, 2014)

1 km res. initial condition: surface speed



1 km res. initial condition: ice flux



Data Processing

ICESAT / GRACE data processing

- stuff happens …

Model post-processing for ICESAT

- only minor steps needed

Model Post-processing for GRACE

- Model lat., lon. ice thickness binned at ½ x ½ degree
- Thickness in each bin converted to cm water equiv.
- Binned data transformed to 60x60 spherical harmonics
- Result is model “seen” at equiv. resolution to GRACE
- Harmonics mapped back to ½ x ½ degree bins for plotting

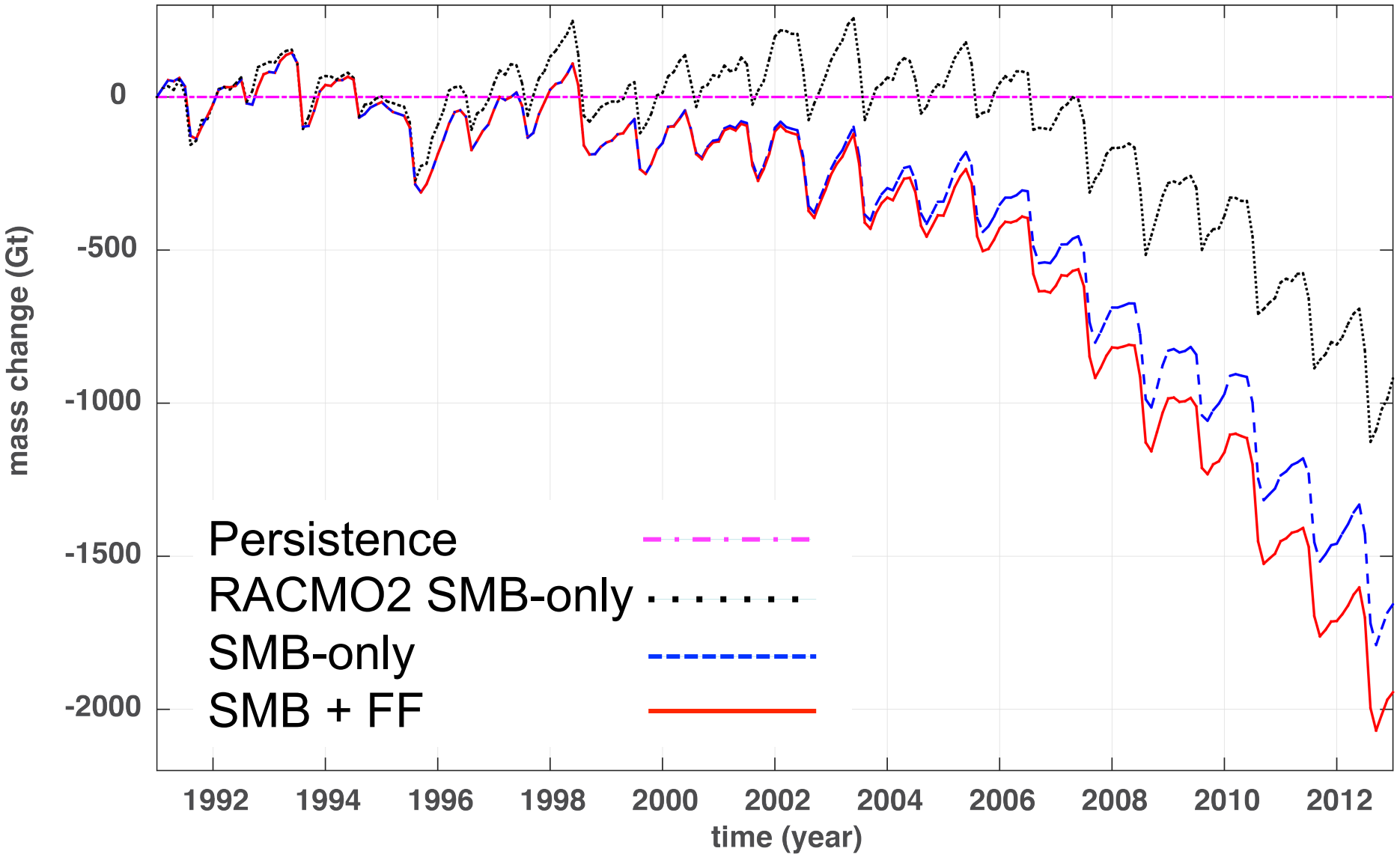
Motivation / Concept

Components

Results

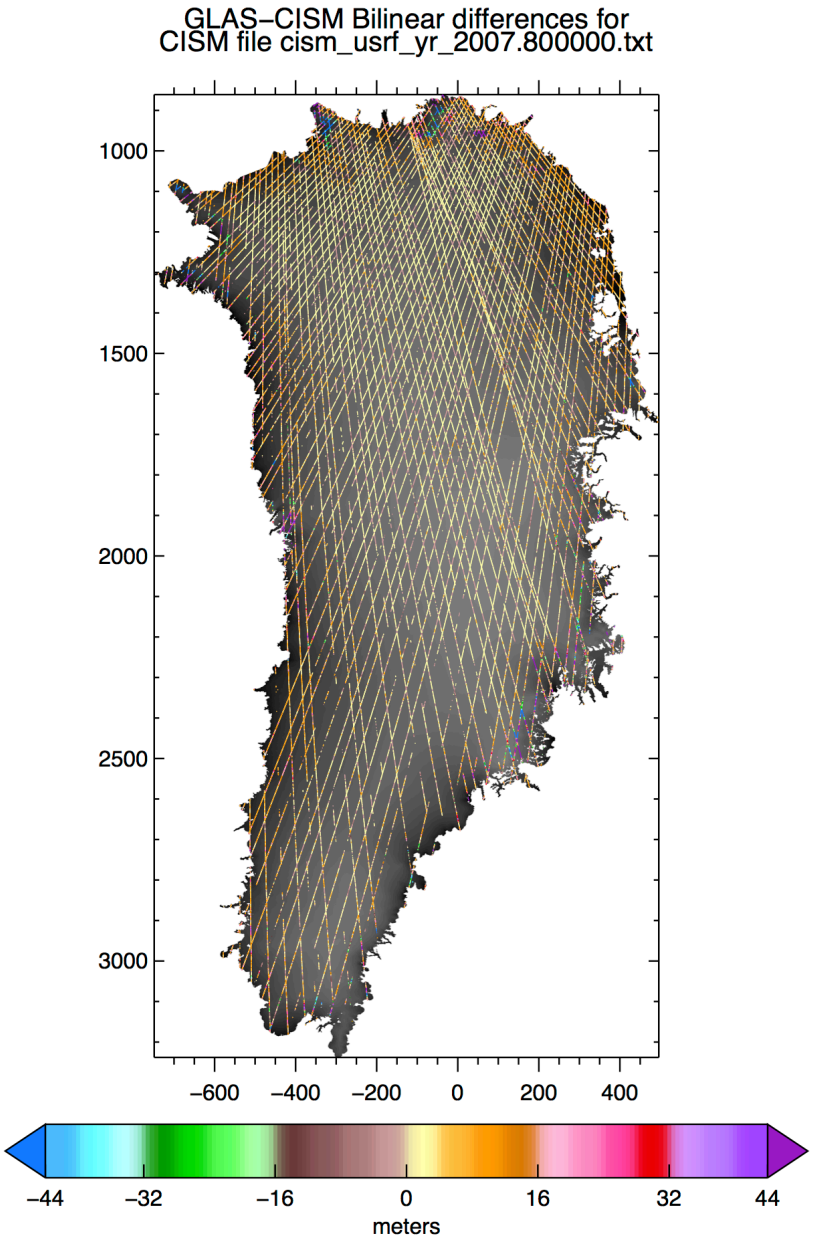
Summary & Future Work

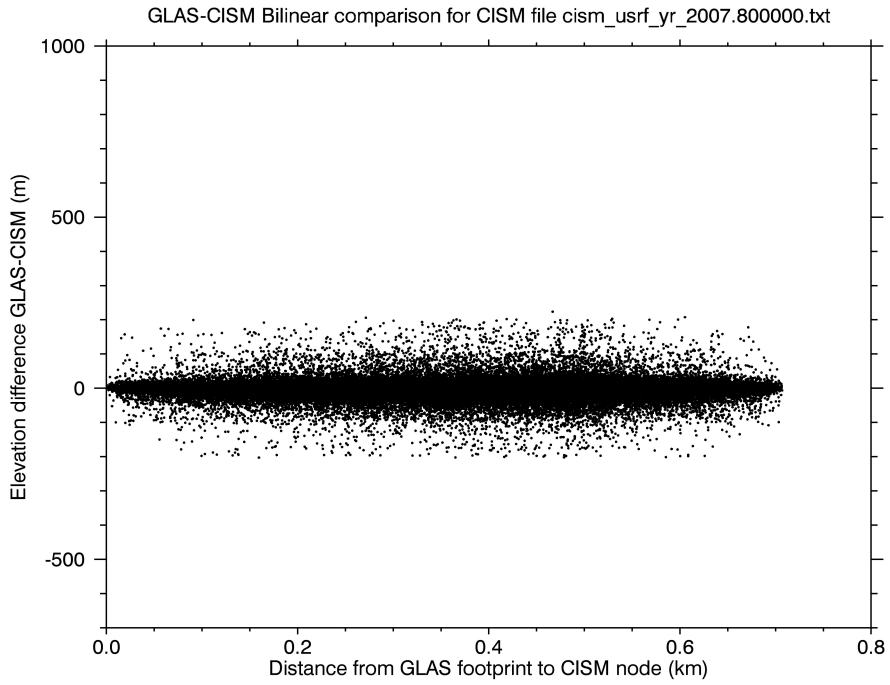
Results

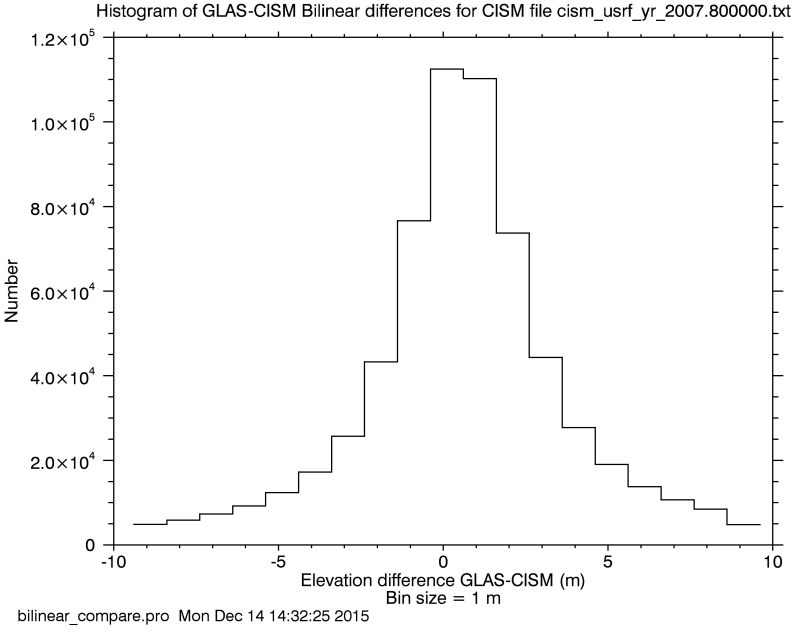

0
-500
-1000
-1500
-2000
mass change (Gt)
1992
1994
1996
1998
2000
2002
2004
2006
2008
2010
2012
time (year)
Persistence
RACMO2 SMB-only
SMB-only
SMB + FF

Results: ICESat

Oct. 2007: SMB + FF







Results: ICESat Metrics

Date (yr)	$\overline{\Delta z}$ (m)	$\sigma_{\Delta z}$ (m)	$\overline{\lvert\Delta z\rvert}$ (m)	$\sigma_{\lvert\Delta z\rvert}$ (m)
	CISM, RACMO2, Pers.	CISM, RACMO2, Pers.	CISM, RACMO2, Pers.	CISM, RACMO2, Pers.
2003.8	0.26, 0.20, 0.15	10.48, 10.43, 10.38	4.46, 4.40, 4.37	9.49, 9.45, 9.42
2004.8	0.11, 0.01, -0.05	10.87, 10.83, 10.78	4.61, 4.56, 4.52	9.84, 9.82, 9.79
2007.8	0.33, 0.19, 0.04	10.33, 10.28, 10.24	4.35, 4.29, 4.26	9.38, 9.34, 9.32

Whole-ice-sheet metrics for SMB-only simulation (top) and SMB+FF simulation (bottom)

Date (yr)	$\overline{\Delta z}$ (m)	$\sigma_{\Delta z}$ (m)	$\overline{\lvert\Delta z\rvert}$ (m)	$\sigma_{\lvert\Delta z\rvert}$ (m)
	CISM, RACMO2, Pers.	CISM, RACMO2, Pers.	CISM, RACMO2, Pers.	CISM, RACMO2, Pers.
2003.8	0.27, 0.20, 0.15	10.49, 10.43, 10.38	4.46, 4.40, 4.37	9.50, 9.45, 9.42
2004.8	0.12, 0.01, -0.05	10.89, 10.83, 10.78	4.62, 4.56, 4.52	9.86, 9.82, 9.79
2007.8	0.36, 0.19, 0.19	10.34, 10.28, 10.24	4.36, 4.29, 4.26	9.38, 9.34, 9.32

Results: ICESat Metrics

Date (yr)	$\overline{\Delta z}$ (m)	$\sigma_{\Delta z}$ (m)	$\overline{\lvert\Delta z\rvert}$ (m)	$\sigma_{\lvert\Delta z\rvert}$ (m)
	CISM, RACMO2, Pers.	CISM, RACMO2, Pers.	CISM, RACMO2, Pers.	CISM, RACMO2, Pers.
2003.8	0.26, 0.20, 0.15	10.48, 10.43, 10.38	4.46, 4.40, 4.37	9.49, 9.45, 9.42
2004.8	0.11, 0.01, -0.05	10.87, 10.83, 10.78	4.61, 4.56, 4.52	9.84, 9.82, 9.79
2007.8	0.33, 0.19, 0.04	10.33, 10.28, 10.24	4.35, 4.29, 4.26	9.38, 9.34, 9.32

Whole-ice-sheet metrics for SMB-only simulation (top) and SMB+FF simulation (bottom)

Date (yr)	$\overline{\Delta z}$ (m)	$\sigma_{\Delta z}$ (m)	$\overline{\lvert\Delta z\rvert}$ (m)	$\sigma_{\lvert\Delta z\rvert}$ (m)
	CISM, RACMO2, Pers.	CISM, RACMO2, Pers.	CISM, RACMO2, Pers.	CISM, RACMO2, Pers.
2003.8	0.27, 0.20, 0.15	10.49, 10.43, 10.38	4.46, 4.40, 4.37	9.50, 9.45, 9.42
2004.8	0.12, 0.01, -0.05	10.89, 10.83, 10.78	4.62, 4.56, 4.52	9.86, 9.82, 9.79
2007.8	0.36, 0.19, 0.19	10.34, 10.28, 10.24	4.36, 4.29, 4.26	9.38, 9.34, 9.32

Results: ICESat Metrics

Basin:

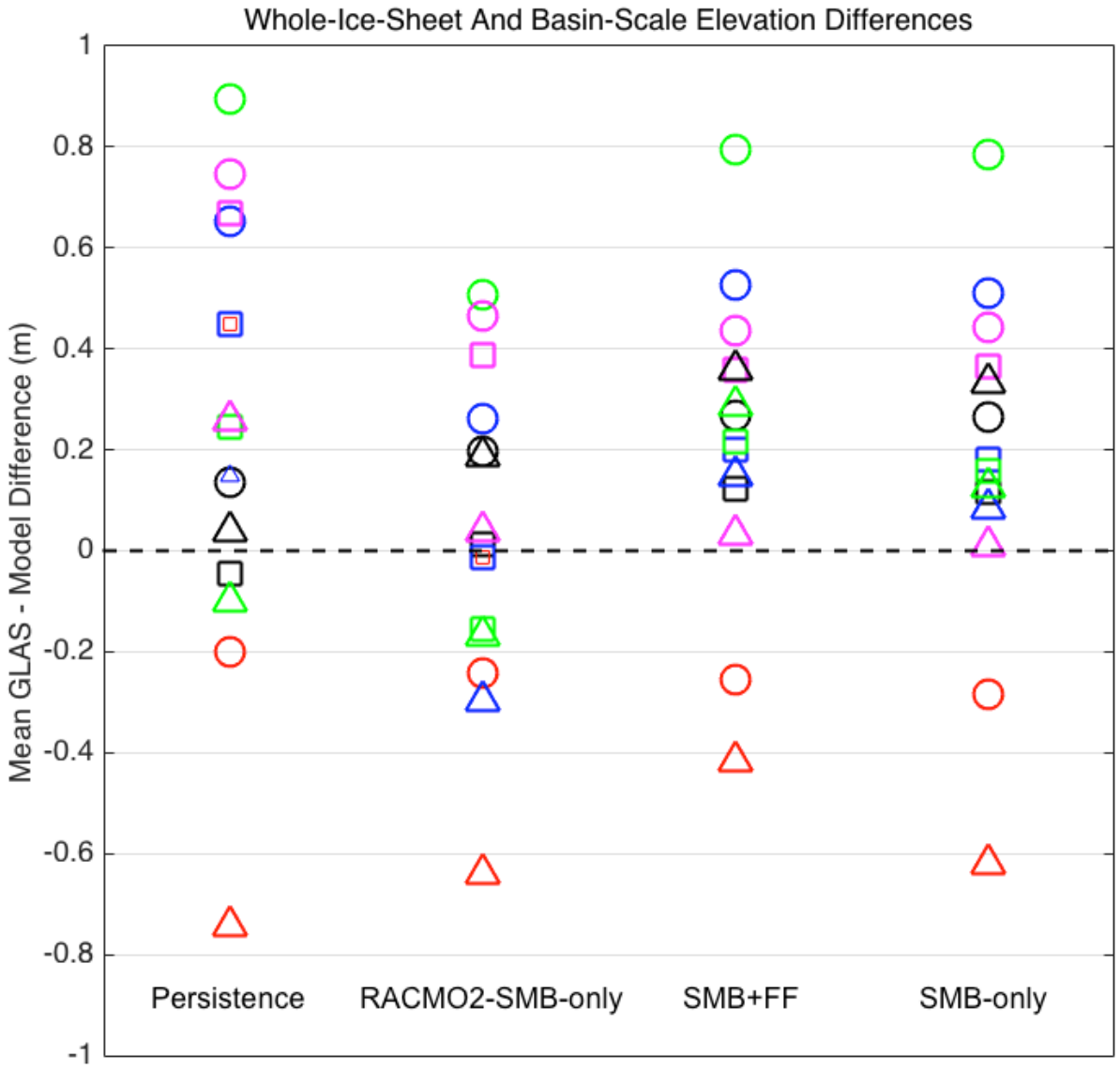
Whole Ice Sheet
Jacobshavn
Kang.
Helheim
Northwest

Date:

2003.8 ○

2004.8 □

2007.8 △



Results: ICESat

The model was initialized using the recent Bamber DEM[1], which uses ICESat data from the time span of our model run

Overall, the model / observation mismatch is small at the whole-ice-sheet and basin scale (mean differences <1 m)

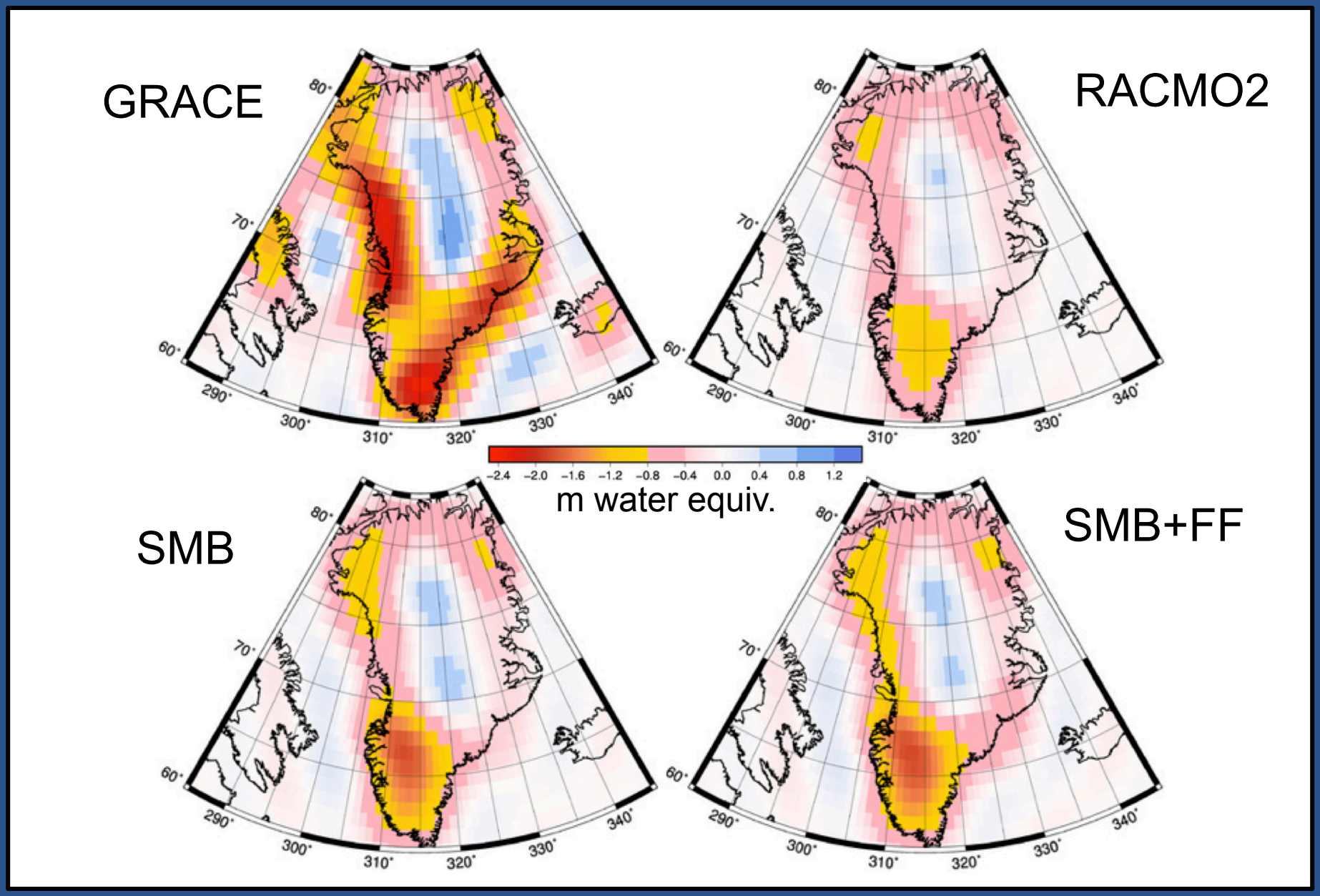
While unable to distinguish between simulations conducted here, ICESat observations confirm that:

- the model states are a good match to observations
- the models are clearly “doing no harm” w.r.t. respecting the observations / init. cond.

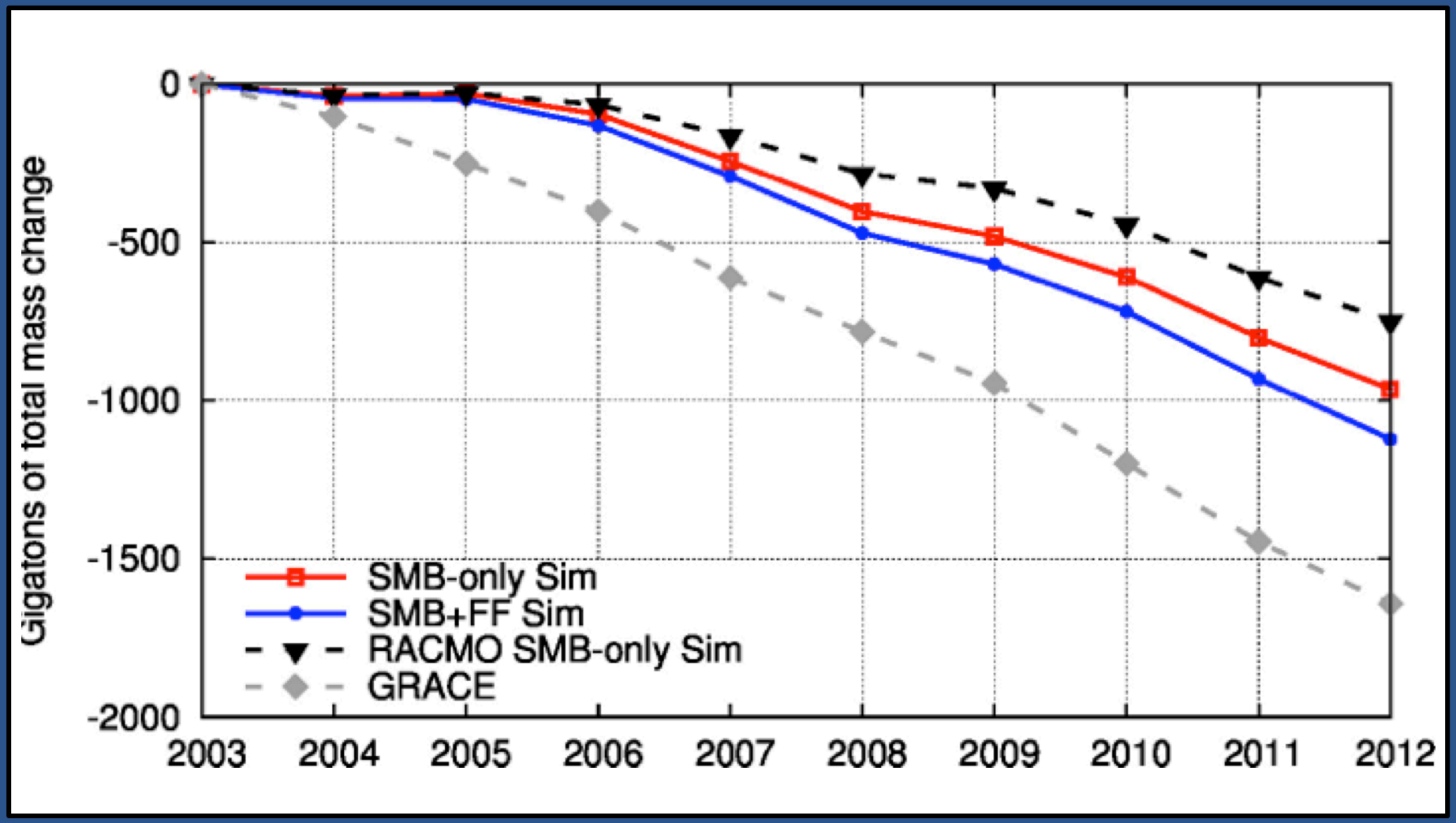
[1] Bamber et al. (*TC*, 2013)

Results: GRACE

Total mass change from 2003-2012



Whole ice-sheet mass trend



GRACE Metrics

Proposed metrics, M*, for use in evaluating model performance relative to GRACE observations:

$$PVE(x,y) = \frac{\sigma_{GRACE} - \sigma_{[\,GRACE\,-\,Model\,]}}{\sigma_{GRACE}} \times 100$$

M_{PVE} = ice sheet wide average of PVE

M_{Trend} = difference in spatial mean, linear trend

Simulation	Trend (Gt yr^{-1})	M_{Trend} (error)	M_{PVE} (%)
GRACE	-186.1	0 (0%)	100
RACMO2-SMB-only	-83.3	-102.8 (55%)	39.8
SMB-only	-100.4	-75.7 (41%)	46.5
SMB+FF	-121.0	-65.1 (35%)	49.7


Motivation / Concept
Components
Results
Summary & Future Work

Summary

Questions:

(1) Can we develop standard *metrics* for comparing models with these observations?

YES (other metrics possible too)

(2) Can we demonstrate that these metrics are useful for distinguishing between relatively worse and relatively better model simulations?

YES (with caveats …)

Summary

- Over this time span, and for this set of simulations, ICESat observations *ARE NOT* clearly able to distinguish between the different simulations

- GRACE observations *ARE* clearly able to distinguish between different simulations

- GRACE metrics confirm expectations:
    - model with SMB and dynamics forcing better than …
    - model with SMB forcing only, better than …
    - non-dynamic models

- Here, *trend* observations (GRACE) useful for distinguishing between simulations, *state* observations (ICESat) are not

Summary

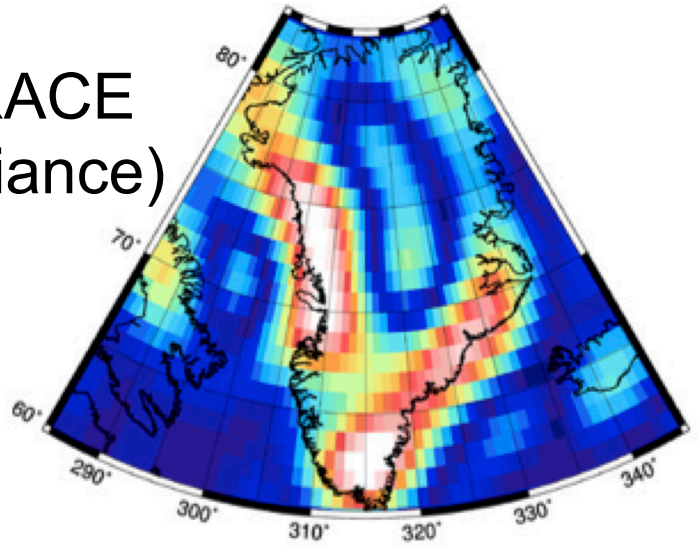
- Here, *trend* observations (GRACE) useful for distinguishing between simulations, *state* observations (ICESat) are not
- But state information still important – a good match to the observed trends would be much less compelling if match to state was poor
- Thus, validation against observations should account for both the *state* and *trend* (of the model and observations); different types of observations needed for validation
- Current generation ice sheet models, when appropriately forced, show skill at mimicking ice sheet observations

Future Work

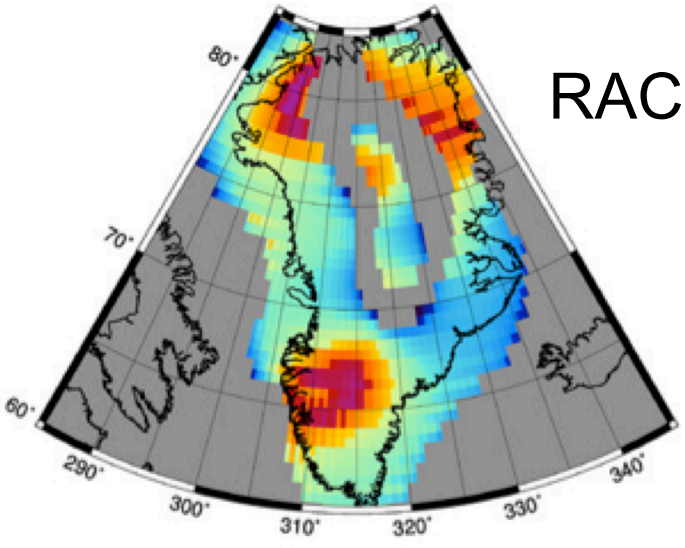
- Refine CMCT as beta version undergoes testing
- Include GRACE processing (currently offline from CMCT)
- Extend to Antarctic observations and model test cases
- Include other data from past, present, future:
    - 1990's NASA-P3, ERS, Envisat altimetry (past)
    - OIB, Cryosat altimetry (present)
    - ICESat2, GRACE2, Cryosat (future)
- Include other types of validation data (temperature, etc.)

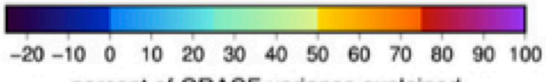
Results: GRACE PVE

GRACE (variance)

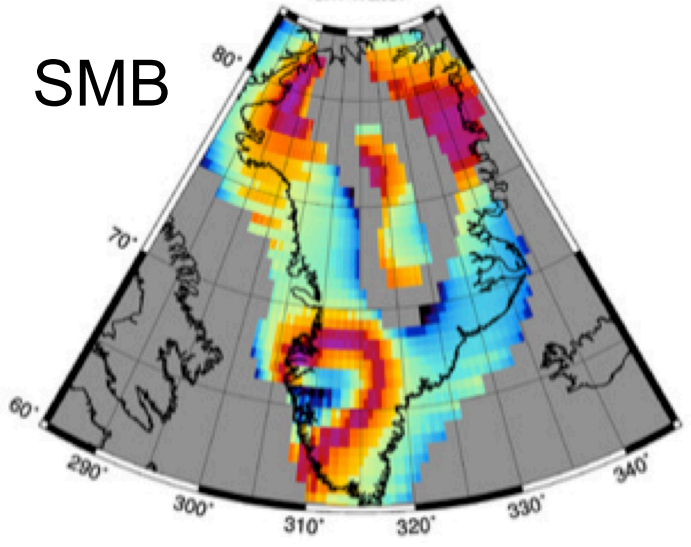


RACMO2

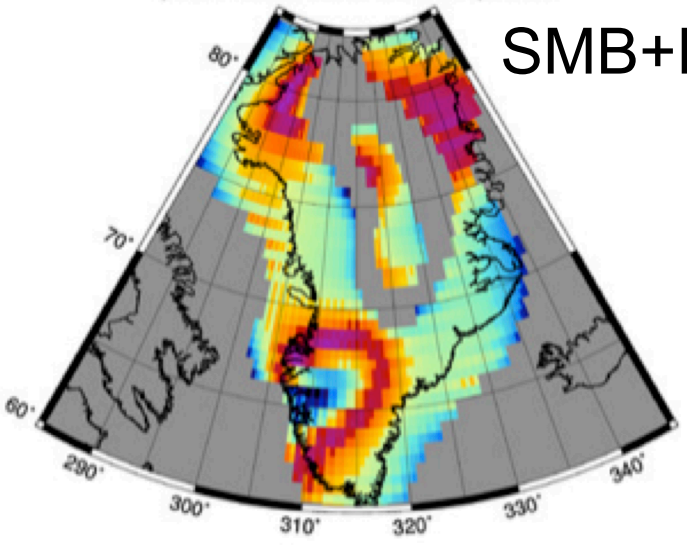




SMB



SMB+FF



ICESat Post Processing

- GIMP 90-m DEM mask used to filter GLAS rel. 64 data. GLAS points excluded …
  - if not within GIMP mask
  - if reflectivity < threshold value
  - if waveform stndev > threshold value
  - if | GIMP – GLAS | > 200 m (clouds, blowing snow)
- Annual model output compared to elevations from fall ICESat campaign of same year
- Model grid point elevations are interpolated to nearest GLAS footprint

GRACE Post Processing

- Ocean, atmos. trends filtered out
- C_{20} geocenter terms replaced with higher-accuracy SLR versions
- GIA model used to remove isostacy signal
- Monthly data averaged to annual
- No other smoothing applied to GRACE data