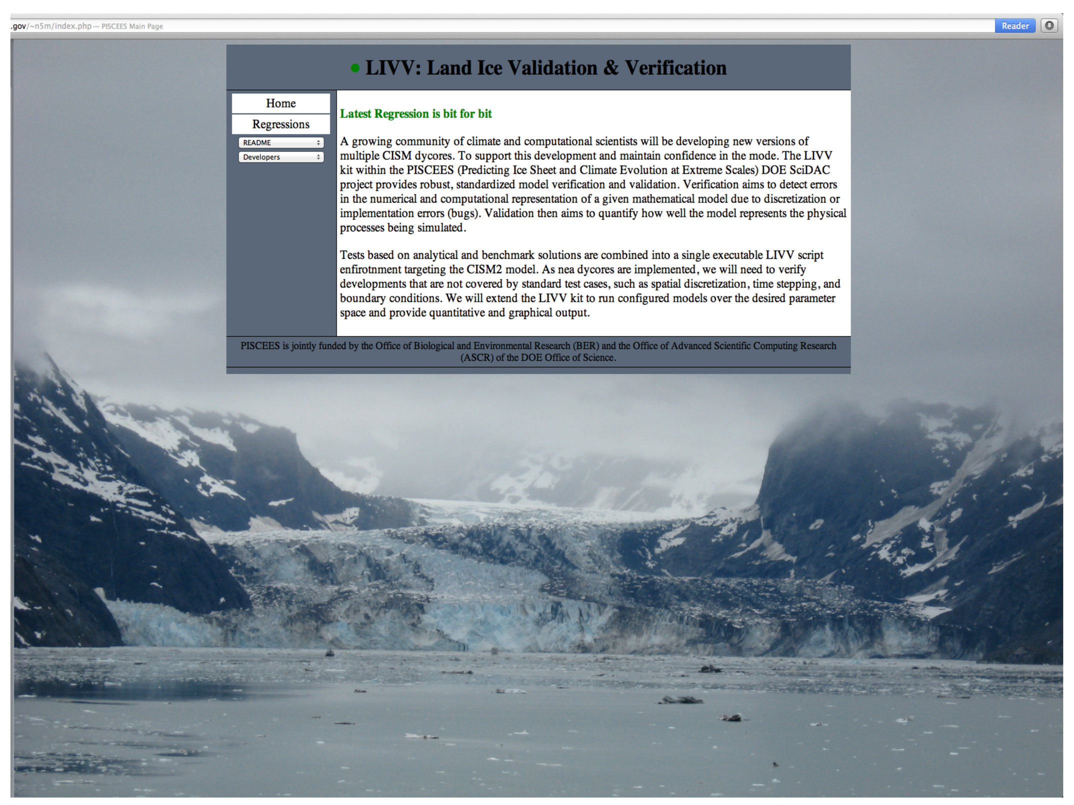
Validation and verification software for a robust evaluation of continental-scale ice sheet models

Katherine J. Evans*, Patrick H. Worley*, Adrianna Boghozian*, Matthew Norman*, Stephen Price+, Jeremy Fyke,+ Miren Vizcaino†
*Oak Ridge National Laboratory, +Los Alamos National Laboratory, †CN Delft, Netherlands

Verification and Auto-Testing

With the development of multiple dycores and multiple new feautures and parameter settings, a robust testing and verification software and automated regression testing system has been developed.

The existing suite of test cases are being combined into a collection of cases to be run automatically using a python -based extensble software package that comes with the CESM code. It is designed for easy use and extension to aid in the maintenance the original glide dycore and development of the new FELIX and BISICLES dycores.



v/~4ue/livv/test_suite.html — Test Suite Diagnostics

Test Suite Diagnostics

Test Suite Descriptions

Diagnostic Dome 30 Test: Bit-for-Bit

Velocity Solver Details
Case and Parameter Settings Details
Plots
Time of Last Simulation: 01/27/2014 08:34 AM

Evolving Dome 30 Test: Bit-for-Bit

Velocity Solver Details
Case and Parameter Settings Details
Plots
Time of Last Simulation: 01/27/2014 08:34 AM

Circular Shelf Test: Bit-for-Bit

Velocity Solver Details
Case and Parameter Settings Details
Plots
Time of Last Simulation: 01/27/2014 08:34 AM

Confined Shelf Test: Bit-for-Bit

Velocity Solver Details
Case and Parameter Settings Details
Plots
Time of Last Simulation: 01/27/2014 08:34 AM

ISMIP HOM A 80KM Test: Bit-for-Bit

Velocity Solver Details
Case and Parameter Settings Details
Plots
Time of Last Simulation: 01/27/2014 08:34 AM

ISMIP HOM A 20KM Test: Bit-for-Bit

Velocity Solver Details

Link to LIVV example

Case Details:

Configure File Settings

Output available from test run: thk uvel vvel velnorm
Grid Size (vert by ew by ns): 5x43x44
Grid Spacing (ew by ns): 5000x5000
Start/End Time: 0.,0., Number of time steps = 0.0

Parameters
----
flow_factor = 1
default_flwa = 5.7e-18

Options
----
dycore = 2 different than benchmark value: 1
flow_law = 0
evolution = 3
temperature = 0
marine_margin = 0

HO Options
----
which_ho_babc = 5
which_ho_efvs = 2
which_ho_sparse = 3 different than benchmark value: 4
which_ho_nonlinear = 0 different than benchmark value: 1

Velocity Solver Settings:

XML File Settings

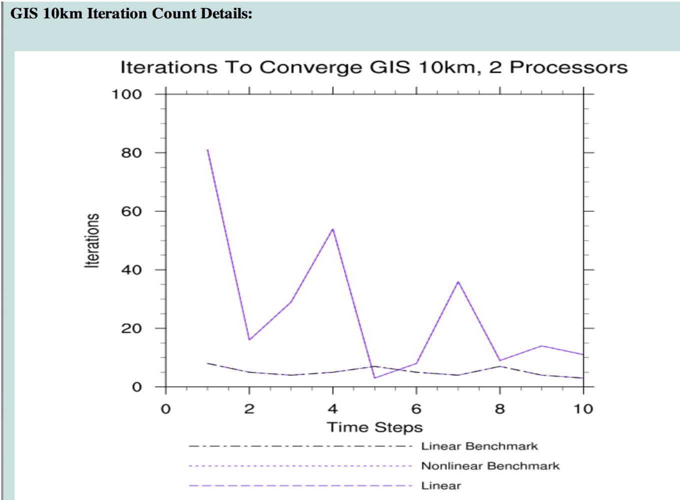
Preconditioner: Picard

Block GMRES: Convergence Tolerance = 1e-08
Block GMRES: Maximum Iterations = 300
Preconditioner Type = Ifpack
Prec Type = ILU
Overlap = 1
Fact: Level-of-Fill = 4

Solver: NK

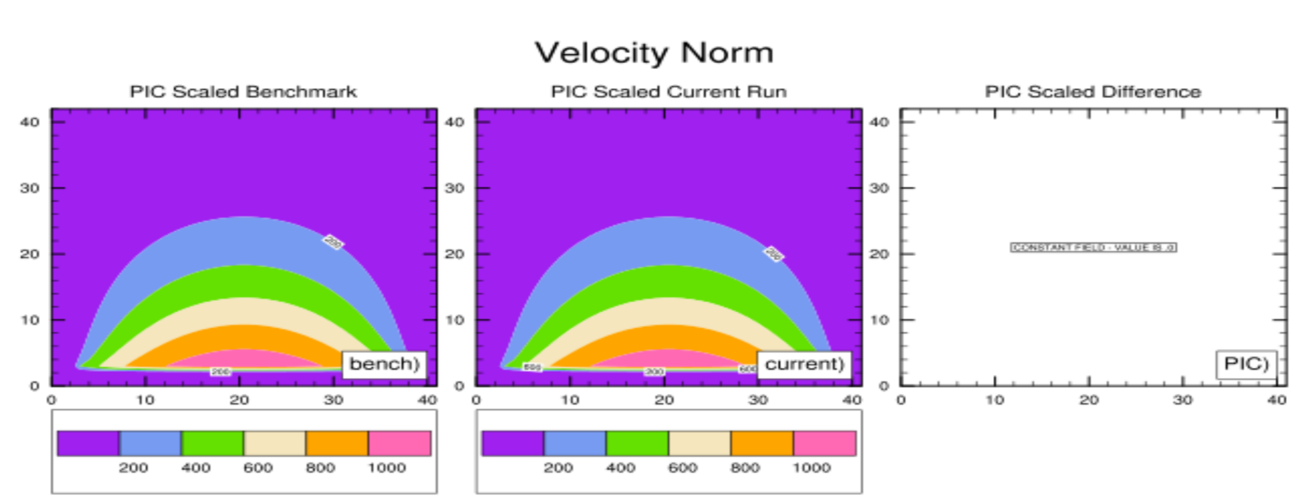
When the LIVV script is activated, it calls a set of python modules to run a suite of cases and creates a website (above left) with a link to a list of all the cases and 3 links for each (above right) to show the configuration information (left), velocity solver details (right), and plots of the results against a benchmark solution (below).

GIS 10km Iteration Count Details:



New Run: gis_10km.JFNK.trilinos.10.gnu.out

Number of Processors = 128
Number of Nonlinear Iterations = 81, 16, 29, 54, 3, 8, 36, 9, 14, 11, 15
Average Number of Linear Iterations per Time-Step = 8.127, 5.923, 4.821, 5.115, 7.333, 5.857, 4.333, 7.25, 4.25, 3.9, 2.714



Vizcaino, M., W. H. Lipscomb, W. J. Sacks, J. H. Van Angelen, B. Wouters, and M. R. VanDen Broeke (2013). "Greenland surface mass balance as simulated by the Community Earth System Model. Part 1: Model Evaluation and 1850-2005 Results." *J. Climate*, **26**: 7793-7812.

Acknowledgements

A special thanks to the U.S. Department of Energy Office of Science BER SciDAC program, who is funding this as part of the Predicting Ice Sheet and Climate Evolution at Extreme Scales (PISCEES) project. We performed the computer simulations and postprocessing using a ALCC allocation on DOE's Office of Science OLCF systems. We acknowledge the contributions of all the members of CESM's Land Ice working group in furthering development and analysis of CISM.

Performance V&V

We have created a capability in LIVV to provide information to model developers as to whether new model features (1) introduce undesirable convergence or performance issues that affect model solution time and robustness (verification of performance) and/or (2) provide acceptable simulation value for the additional expense (pLIVV). Performance benchmarks that account for machine variability are provided and deviations from the min/max values are highlighted to uncover issues with simulation at scale.

The main page lists a suite of runs for the diagnostic dome case and the GIS for a range of problem sizes to exercise the scaling and efficiency of the CISM model. One can select which cases to run within the input script based on the system being used.

First time runs: create a benchmark by running the job 10 times and create run stats. Then new runs are tested against the benchmark and highlighted if the run is outside typical or max/min behavior of the original runs (left). If outside the defined range, then links to data about the solver and timing files can be investigated.

Performance and Analysis Test Suite

Test Suite Descriptions

Diagnostic Dome 60 Test: Test Faster Than Expected Performance Range

Velocity Solver Details
Case and Parameter Settings Details
Timing Details
Time of Last Simulation: 05/05/2014 06:53 PM

Diagnostic Dome 120 Test: Test Within Expected Performance Range

Velocity Solver Details
Case and Parameter Settings Details
Timing Details
Time of Last Simulation: 05/05/2014 06:54 PM

Home

Timing Data

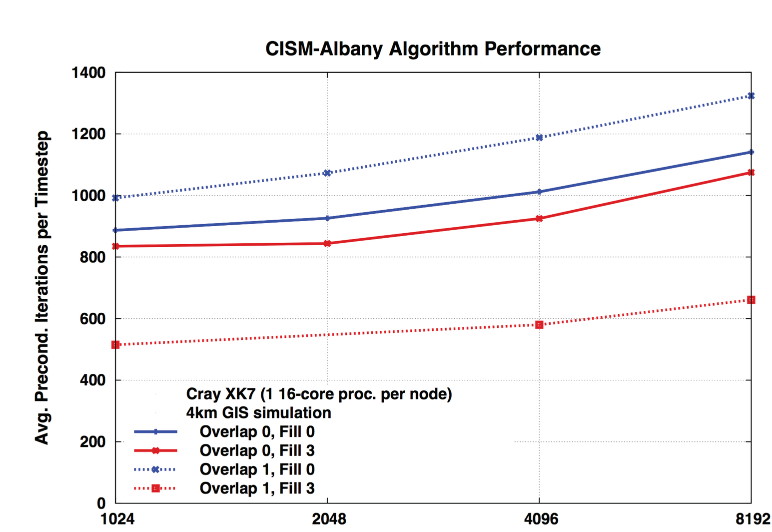
	glam			glissade
	Avg	Max	Min	Current Run
Simple Glide	28.6825246914	34.5996666667	24.9682222222	8.01977111111
Initial Diag Var Solver	28.6454925926	34.5602	24.9331888889	7.97007
Velo Driver	28.6430691358	34.5579222222	24.9307444444	7.96742333333
Glide IO Writeall	0.0105509412346	0.0106680811111	0.0104526111111	0.0107393055556

Less than the Min

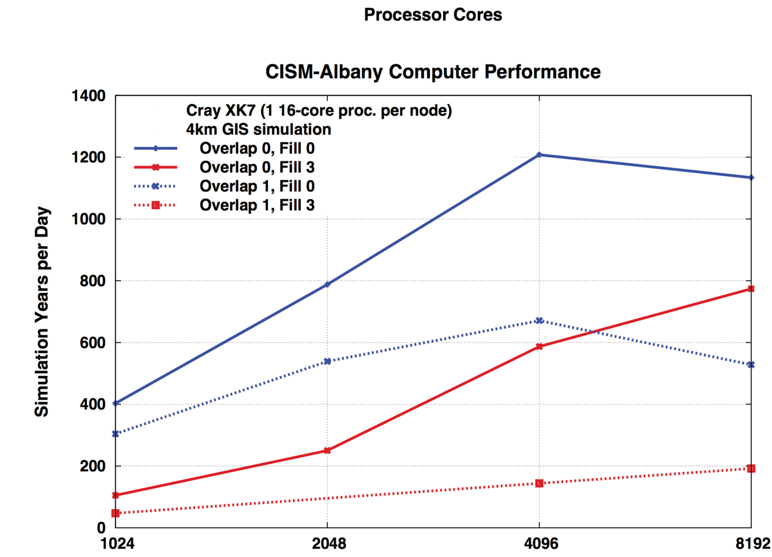
Greater than the Max

The above file is generated when pLIVV is run for a dome and/or GIS case for a given processor size. The timing for each portion of the run is listed for each row, and the columns show the average, max, and min for each one.

In summary, the above software infrastructure is designed for performance verification, which will determine whether observed performance is within the range of expected performance for a given model, test case, and computer platform. The same infrastructure can be used to support performance evaluation during model development, intermodel performance comparisons, and performance tuning.



Left: performance is evaluated for a range of algorithm parameters, indicating that performance is quite sensitive to these parameters (and so tuning is important). It also indicates that minimizing the iteration count in the linear and nonlinear solver is not the only metric of importance - the serial and parallel costs per iteration are also critical.

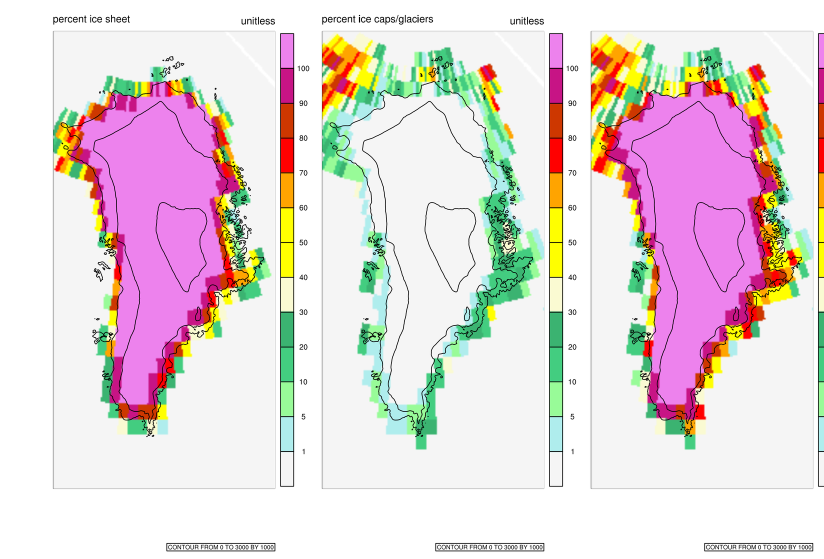


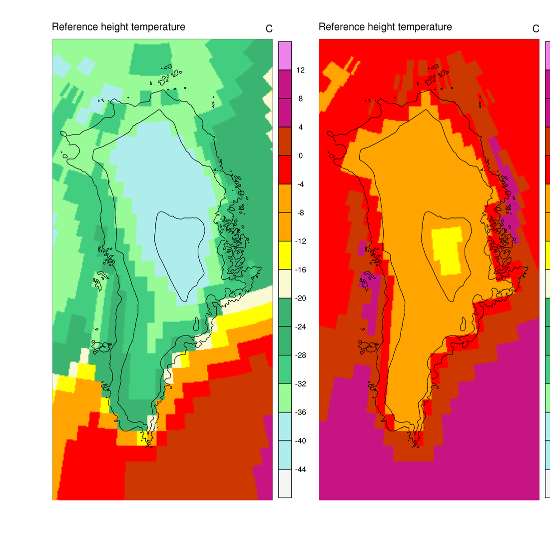
Left: Performance LIVV has a hierarchical structure, and additional detail can be displayed, for example, to determine why performance differs from the baseline. In the figure, the iteration count was used to illuminate some of the performance issues. Improved user support and automation of these more detailed studies is under development for future releases of the pLIVV component of LIVV.

Validation of CISM within CESM

The novelty of continental scale ice sheet modeling with fully coupled climate models along with the paucity of remote and in-situ datasets of ice sheets has prevented comprehensive model validation to date. LIVV is being extended to include a suite of plots and tables of key parameters from CISM runs, with some CESM parameters as well to monitor the boundary forcing data. The initial plots for variables such as ice sheet coverage, reference height T, solar radiation (incoming and absorbed), longwave radiation, sensible heat, and ground evaporation are being incorporated for automated display, with comparison to benchmark CISM model runs, tuned regional model runs, and observational data coming soon.

The plots to the left and below are snapshots from one month of a fully coupled CESM with active CISM 1.0 simulation to show the capability of a validation software package to automate model evaluation. The model run is a transient 1850-2005 simulation with 1 degree resolution atmosphere and ocean with a 5km resolution Greenland ice sheet. The details of the run are outlined in Vizcaino et al. (2013).

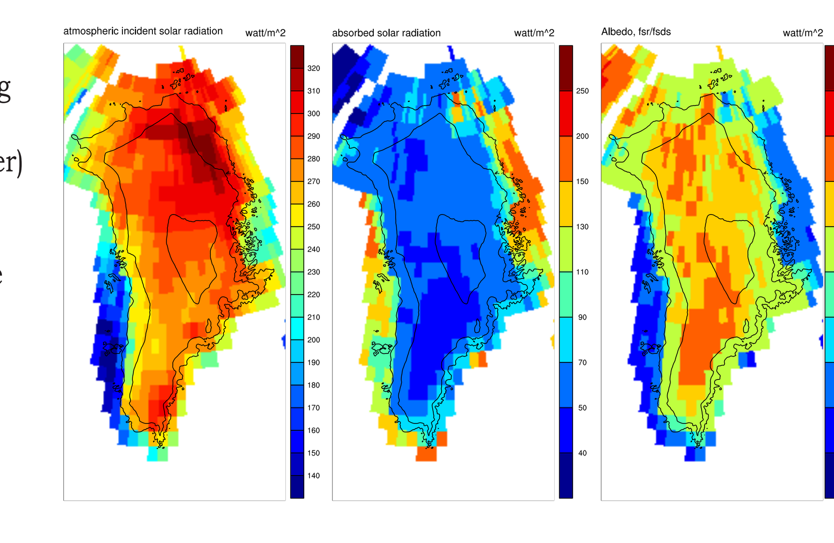


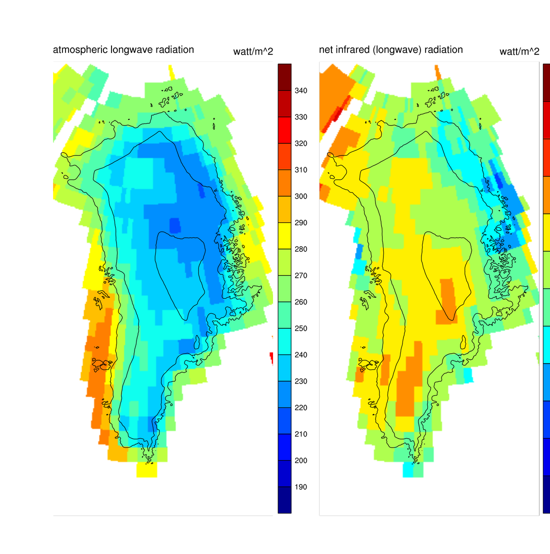


(Above) Ice sheet fraction (left), fraction of grid cell covered by glaciers and ice caps (center) and total ice fraction in the land component (CLM) of the CESM as defined in the surface datasets (right).
(Left) Reference temperature (°C) from the atmosphere component (CAM) of the CESM for January (left) and July (right) of 1960.

Elevation contours for all plots are in black for 0, 1000, 2000, and 3000m.

(Right) Mean shortwave downwelling solar flux (FSDS) from CAM (left), mean absorbed solar flux FSA (center) and surface albedo (FSR) (right) for July 1960.
(Below) Mean downwelling longwave radiation (FLDS) from CAM (left), mean net IR radiation (FIRA) from CLM (right) for July 1960.





(Below) Net surface radiation (FSA-FIRA) (left), sensible heat flux (FSH) (center), and the ground evaporation (QSOIL) (right) from CLM for July 1960. Elevation contours are plotted in black for 0, 1000, 2000, and 3000m.

These are a sampling of the breadth of plots and information to be created and presented on the LIVV website for a given CESM simulation. Relevant plots from all components are presented to ID and diagnose sensitivities and issues with ice sheet behavior within the coupled system.