Use and implementation of adjoint methods in ice sheet models

Jesse Johnson, Jean Utke

30 June, 2010
Land Ice Working Group Session

Objective function

Using $g(\mathbf{x}, \mathbf{p})$

Given model output $\mathbf{x}$, it is desirable to compute a scalar valued function

$$g(\mathbf{x}, \mathbf{p})$$

that also depends on model parameters $\mathbf{p}$.

Examples of $g(\mathbf{x}, \mathbf{p})$ include;

- flow rate,
- streses,
- energy, or
- model agreement with data.

Refer to $g(\mathbf{x}, \mathbf{p})$ as an *objective function*.

Utility of derivatives

Why $\frac{dg}{dt}$ is needed, why its hard to get.

Differentiating the objective function, $\frac{dg}{dp}$ provides;

- the *sensitivity* of the objective function with respect to the parameters, and
- the *search direction* to be used in conjunction with conjugate gradients to determine the minumum of $g(\mathbf{x}, \mathbf{p})$.

The chain rule gives:

$$\frac{dg}{dp} = g_{\mathbf{x}}\mathbf{x}_{\mathbf{p}} + g_{\mathbf{p}}$$

Underscoring the problem, $\mathbf{x}_{\mathbf{p}}$ is tough to evaluate!

Assuming $\mathbf{x}$ can be written $A\mathbf{x} = \mathbf{b}$, and its derivative is $A_{p_i}\mathbf{x} + A\mathbf{x}_{p_i} = \mathbf{b}_{p_i}$, each $\mathbf{x}_{p_i}$ is solved with $\mathbf{x}_{p_i} = A^{-1}(\mathbf{b}_{p_i} - A_{p_i}\mathbf{x})$
...one linear system for each parameter!

Avoiding the M linear solves

Add zero in a cunning way

- Rewrite the objective function

$$\tilde{g} = g - \lambda^T \mathbf{f}$$

  where $f = A\mathbf{x} - b$, which is zero, making λ arbitrary.
- Strategy is to choose λ such that $\mathbf{x_p}$ is eliminated

$$\frac{dg}{d\mathbf{p}}\Big|_{\mathbf{f}=0} = \frac{d\tilde{g}}{dp}\Big|_{\mathbf{f}=0} = g_{\mathbf{p}} - \lambda^T \mathbf{f_p} + (g_{\mathbf{x}} - \lambda^T \mathbf{f_x})\mathbf{x_p}$$

- $\mathbf{x_p}$ is eliminated if

$${\mathbf{f_x}}^T \lambda = g_{\mathbf{x}}^T$$

- $A = \mathbf{f_x}$, so what we really require is that λ satisfies the *adjoint equation*

$$A^T \lambda = g_{\mathbf{x}}^T.$$

  Hence $\frac{dg}{dp}$ comes from the evaluation of a *single* linear system!

The conclusion

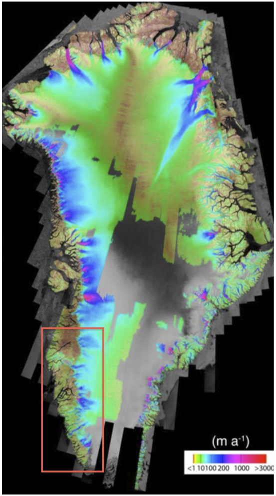
Having solved the adjoint system for λ, the gradient is written

$$\frac{dg}{d\mathbf{p}} = g_{\mathbf{p}} - \lambda^T (A_{\mathbf{p}}\mathbf{x} - \mathbf{b}_{\mathbf{p}})$$

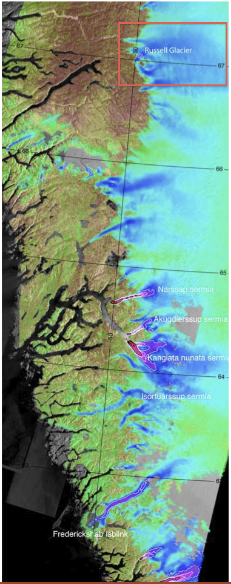
noting that;

- $\mathbf{x}$ is the result of solving the forward model,
- Computing $A_{\mathbf{p}}$ and $\mathbf{b}_{\mathbf{p}}$ are assumed to be analytic expressions, and can be treated *"automatically"*.
- Automatic differentiation (AD) is done with openAD (Utke).

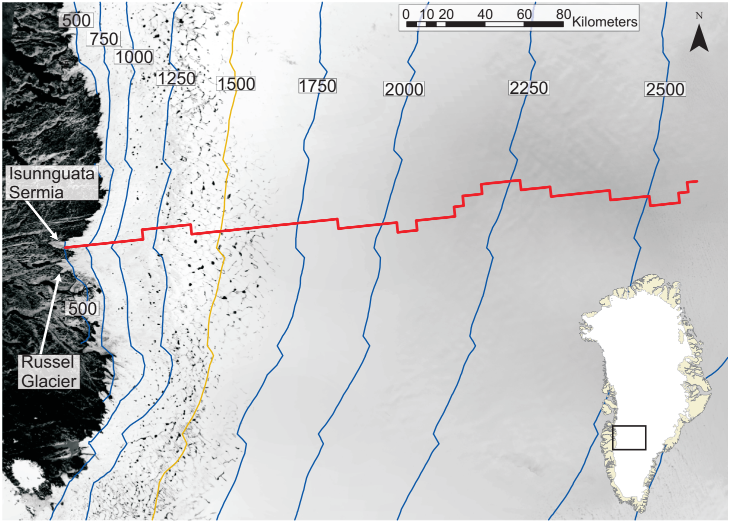
Greenland ice sheet
Velocities from Joughin 2010



Greenland ice sheet
Velocities from Joughin 2010



Greenland ice sheet
Profile



Conservation of energy

$$\frac{1}{\rho c_p} \nabla \cdot k_i \nabla \theta - \mathbf{u} \cdot \nabla \theta + 2\eta \dot{\epsilon}_{\Pi}^2 = 0$$

Conservation of momentum

$$\nabla \cdot 2\eta \dot{\epsilon} - \nabla p = \rho \mathbf{g}$$

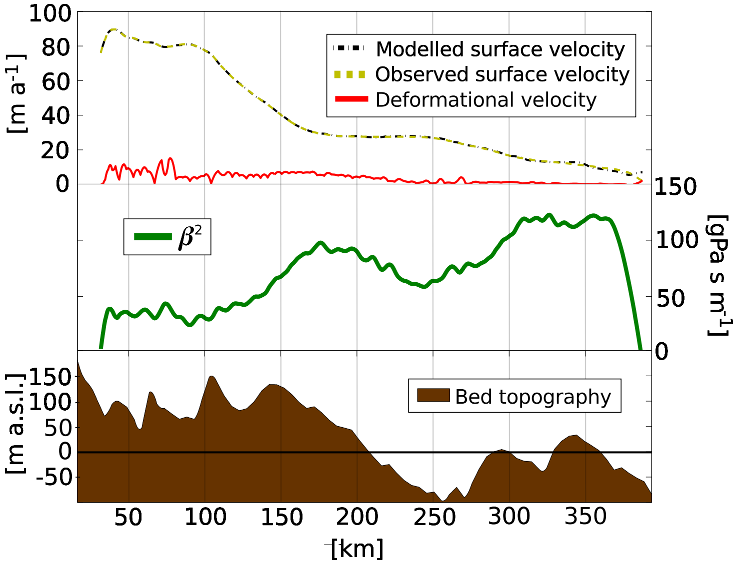
Boundary conditions

$$[-p\mathbf{I} + 2\eta \dot{\epsilon}]\hat{\mathbf{n}} = 0 \quad \text{(Free surface)},$$

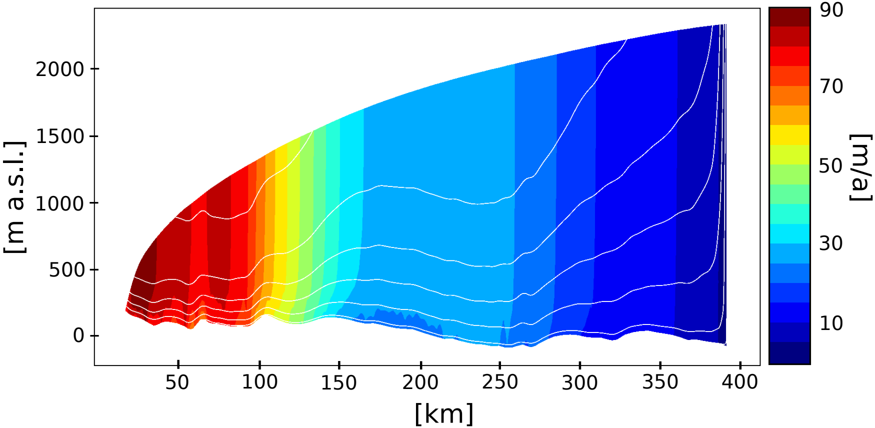
$$\tau_b = \beta^2 \cdot \mathbf{u} \quad \text{(Basal traction)},$$

$$-\hat{\mathbf{n}} k_i \nabla \theta = Q \quad \text{(Basal heat flow)}.$$

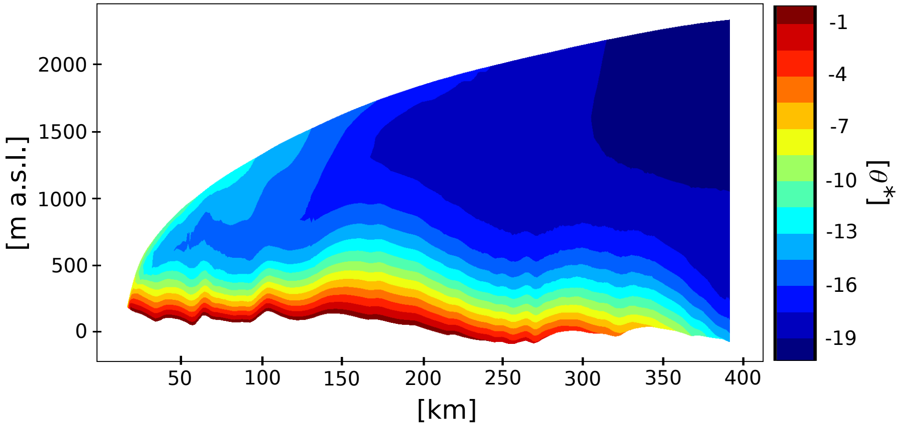
Objective function minimization
Using quasi-Newton method



Resulting fields
Temperature and velocity

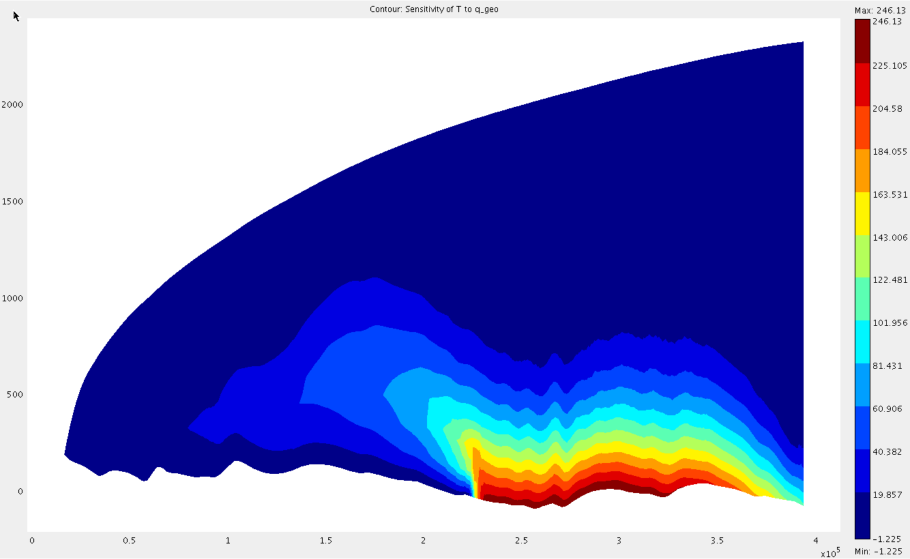


Resulting fields
Temperature and velocity



Sensitivity

Sensitivity of temperature to heat flow



Sensitivity
Sensitivity of velocity to heat flow