Introduction to Computational and Variational Methods for Inverse Problems
CSE 397/GEO 391/ME 397/ORI 397
Fall 2015

Omar Ghattas

Departments of Geological Sciences and Mechanical Engineering
Institute for Computational Engineering and Sciences
The University of Texas at Austin

slides downloadable from:
http://users.ices.utexas.edu/~omar/inverse_problems/intro.pdf

The inverse problem: Determining cause from effect

- Integrating data and models to infer uncertain parameters is an inverse problem (also called data assimilation, history-matching, parameter estimation, model calibration, system identification, etc.)

The inverse problem: Determining cause from effect

- Integrating data and models to infer uncertain parameters is an inverse problem (also called data assimilation, history-matching, parameter estimation, model calibration, system identification, etc.)
- Applications across a broad spectrum of natural and engineered systems (seismic inversion, medical imaging, weather forecasting, non-destructive testing, inverse scattering, reservoir parameter estimation, remote sensing, source localization, ...)

The inverse problem: Determining cause from effect

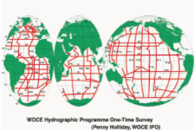
- Integrating data and models to infer uncertain parameters is an inverse problem (also called data assimilation, history-matching, parameter estimation, model calibration, system identification, etc.)
- Applications across a broad spectrum of natural and engineered systems (seismic inversion, medical imaging, weather forecasting, non-destructive testing, inverse scattering, reservoir parameter estimation, remote sensing, source localization, ...)
- Some key milestones in regularization-based inversion
  - 1920s: Hadamard lays down theoretical foundations for ill-posed problems
  - 1950s-60s: Marchuk introduces idea of adjoint method
  - 1960s-70s: Tikhonov's work on regularization provides a systematic means for computing inverse solutions
  - 2000s: Advent of terascale supercomputers and advanced algorithms permits solution of some large-scale regularized inverse problems

The inverse problem: Determining cause from effect

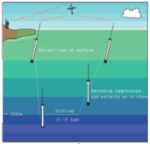
- Integrating data and models to infer uncertain parameters is an inverse problem (also called data assimilation, history-matching, parameter estimation, model calibration, system identification, etc.)
- Applications across a broad spectrum of natural and engineered systems (seismic inversion, medical imaging, weather forecasting, non-destructive testing, inverse scattering, reservoir parameter estimation, remote sensing, source localization, ...)
- Some key milestones in Bayesian inversion
  - 1763: Bayes' theorem
  - Metropolis (1953) and Hastings (1970) lay groundwork for Markov chain Monte Carlo Method, providing a method for computing solutions to Bayesian inverse problems
  - 2010s: Advent of petascale supercomputers and advanced algorithms permits solution of some large-scale Bayesian inverse problems

Ocean dynamics example: Observational data

WOCE

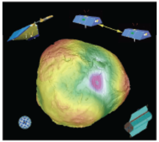


Argo

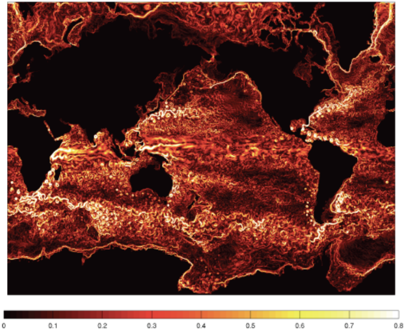


T/P, Jason

GRACE



How to integrate observational data into predictive model?



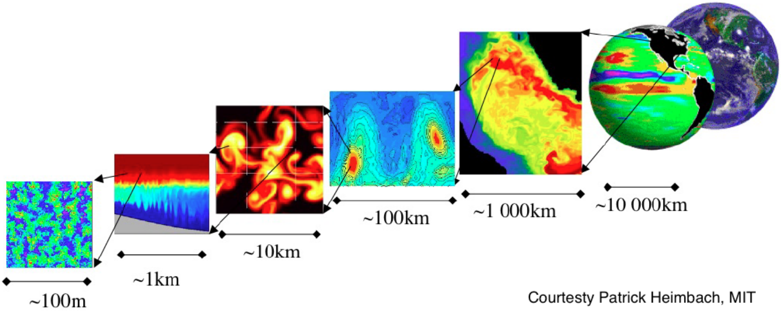
Courtesy Patrick Heimbach, MIT

Ocean dynamics example: Model

Navier-Stokes equations

- Conservation of mass
- Conservation of momentum
- Conservation of energy
- Conservation of salinity
- Equation of state
- Subgrid parameterizations

TACC *Stampede* supercomputer

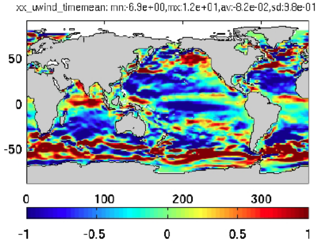


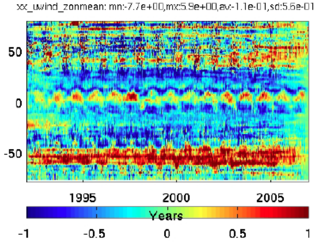
Courtesty Patrick Heimbach, MIT

Ocean dynamics example: Uncertain parameters

2D and 3D parameter fields

- 3D initial temperature and salinity fields
- 2D time-varying atmospheric state at ocean–atmosphere interface
    - surface air temperature
    - specific humidity
    - downwelling shortwave radiation
    - zonal and meridional wind speed
- 3D subgrid model parameter fields
    - vertical mixing coefficient
    - GM coefficient (geostrophic eddy mixing)
    - Redi coefficient (along-isypycnal mixing)





Courtesy Patrick Heimbach, ICES/GEO

Anatomy of an inverse problem

Input parameters, computational model, and output observables

The forward problem

- Given model parameters m, solve forward model $\mathcal{F}$ to yield output observables d

$$\mathcal{F}(m) \longrightarrow d$$

- Well-posed: solution exists, is unique, and is stable to perturbations in inputs
- Causal: later-time solutions depend only on earlier time solutions
- Local: the forward operator includes derivatives that couple nearby solutions in space and time

Anatomy of an inverse problem

Input parameters, computational model, and output observables

The inverse problem

- Given output observations d_{obs} and forward model $\mathcal{F}$, infer model parameters m

$$m \longleftarrow \mathcal{F}^{-1} d_{\text{obs}}$$

- Ill-posed: observations are usually sparse; many different parameter values may be consistent with the data
- Non-causal: the inverse operator couples earlier time solutions with later time ones
- Global: the inverse operator couples solution values across all of space and time

Anatomy of an inverse problem

Input parameters, computational model, and output observables

Occam's approach to ill-posedness

Employ regularization to penalize unwanted solution features, guarantee unique solution:

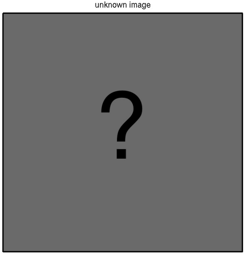
$$\min_{m} \ \frac{1}{2} \| \mathcal{F}(m) - d_{\text{obs}} \|_W^2 + \frac{\alpha}{2} \| m - m_{\text{ref}} \|_R^2$$

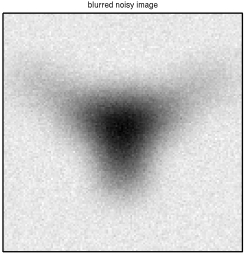
Bayesian approach to ill-posedness

Describe probability of all models that are consistent with the data and any prior knowledge of the parameters:

$$\pi(m|d_{\text{obs}}) \propto \exp\left(-\frac{1}{2} \| \mathcal{F}(m) - d_{\text{obs}} \|^2_{C_{\text{d}}^{-1}} - \frac{1}{2} \| m - m_{\text{pr}} \|^2_{C_{\text{m}}^{-1}} \right)$$

Ill-posedness demonstration: Image deblurring/denoising

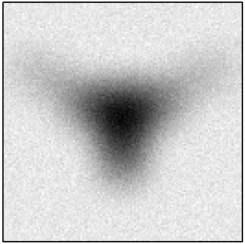




$$\xrightarrow{\mathcal{F}}$$

- forward model $\mathcal{F}$: blurring (convolution with a Gaussian)
- model parameters m: unknown left image
- observations d_{obs}: blurred and noisy right image

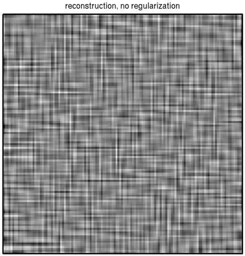
Ill-posedness demonstration: Image deblurring/denoising

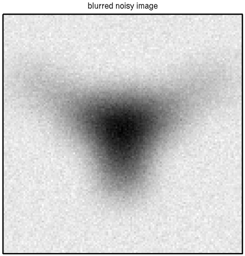


$$\xleftarrow{\mathcal{F}^{-1}}$$

- forward model $\mathcal{F}$: blurring (convolution with a Gaussian)
- model parameters m: unknown left image
- observations d_{obs}: blurred and noisy right image

Ill-posedness demonstration: Image deblurring/denoising

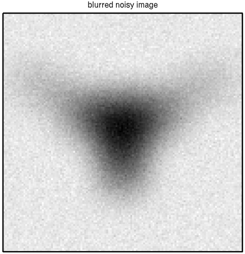




$$\xleftarrow{\mathcal{F}^{-1}}$$

- forward model $\mathcal{F}$: blurring (convolution with a Gaussian)
- model parameters m: unknown left image
- observations d_{obs}: blurred and noisy right image

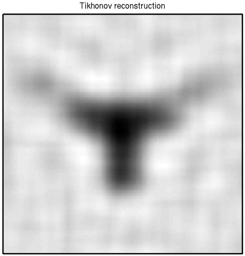
Ill-posedness demonstration: Image deblurring/denoising

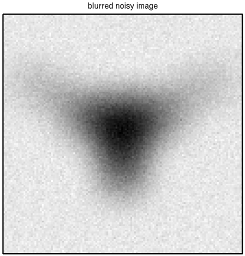


$$\xleftarrow{\min \ \frac{1}{2}\|\mathcal{F}m - d_{\text{obs}}\|^2 + \frac{\alpha}{2}\|m\|^2}$$

- forward model $\mathcal{F}$: blurring (convolution with a Gaussian)
- model parameters m: unknown left image
- observations d_{obs}: blurred and noisy right image
- regularization: Tikhonov

Ill-posedness demonstration: Image deblurring/denoising

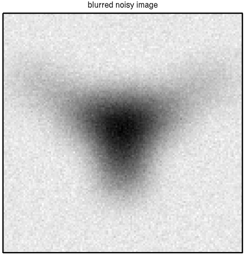




$$\min \; \frac{1}{2}\|\mathcal{F}m - d_{\text{obs}}\|^2 + \frac{\alpha}{2}\|m\|^2$$

- forward model $\mathcal{F}$: blurring (convolution with a Gaussian)
- model parameters m: unknown left image
- observations d_{obs}: blurred and noisy right image
- regularization: Tikhonov

Ill-posedness demonstration: Image deblurring/denoising

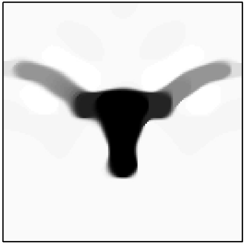


$$\min \frac{1}{2}\|\mathcal{F}m - d_{\text{obs}}\|^2 + \frac{\alpha}{2}\int (\nabla m \cdot \nabla m)^{0.5}$$

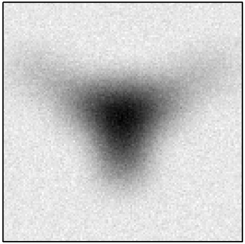
- forward model $\mathcal{F}$: blurring (convolution with a Gaussian)
- model parameters m: unknown left image
- observations d_{obs}: blurred and noisy right image
- regularization: total variation

Ill-posedness demonstration: Image deblurring/denoising

TV reconstruction



blurred noisy image



$$\min \frac{1}{2}\|\mathcal{F}m - d_{\mathsf{obs}}\|^2 + \frac{\alpha}{2}\int(\nabla m \cdot \nabla m)^{0.5}$$

$$\longleftarrow$$

- forward model $\mathcal{F}$: blurring (convolution with a Gaussian)
- model parameters m: unknown left image
- observations d_{obs}: blurred and noisy right image
- regularization: total variation

Example: Inverse problem for Antarctic ice sheet

- Ice flows from interior of polar ice sheets to ocean is primary contributor to sea level rise (200 billion tons/yr currently)
- Flow rates of outlet glaciers in Antarctica have been increasing over past several decades
- Thinning of ice shelves due to increased mixing in ocean driven by intensification of polar winds, bringing warmer water to surface;
- Ice shelf thinning leads to loss of buttressing effect and retreat of ice sheet
- 0.5 m sea level rise by 2070 estimated to jeopardize 136 largest port cities, with 150M inhabitants and $35 trillion in assets
- We need to develop predictive models, with quantified uncertainties, to better anticipate future sea level rise

Motivation, cont.

- Recent evidence suggests that sea level rose abruptly at the end of the last interglacial (∼118 kyr ago) by ∼5-6m; the likely cause is catastrophic collapse of polar ice sheets
  - *Ice sheet collapse following a prolonged period of stable sea level during the last interglacial*, MJ O'Leary, PJ Hearty, WG Thompson, ME Raymo, JX Mitrovica, JM Webster, **Nature Geoscience**, 6, 796800, 2013.
- Recent work indicates that retreat of the Amundsen Sea Embayment (a portion of the West Antarctic ice sheet) is accelerating with no major bed obstacles to prevent draw down of the entire basin.
  - *Widespread, rapid grounding line retreat of Pine Island, Thwaites, Smith and Kohler glaciers, West Antarctica from 1992 to 2011*, E. Rignot, J. Mouginot, M. Morlighem, H. Seroussi, and B. Scheuchl, **Geophysical Research Letters**, 41(10):3502–3509, 2014.

Dynamics of the Antarctic ice sheet and sea level rise

Glaciers flow thousands of miles from the continent's deep interior to its coast

Credit: NASA Goddard Space Flight Center/JPL-Caltech

"Full Stokes" ice sheet model

Balance of linear momentum, mass, and energy

$$-\boldsymbol{\nabla}\cdot[\eta(\theta,\boldsymbol{u})\,\dot{\boldsymbol{\varepsilon}}-\boldsymbol{I}p]=\rho\boldsymbol{g} \qquad [\dot{\boldsymbol{\varepsilon}}=\tfrac{1}{2}(\boldsymbol{\nabla}\boldsymbol{u}+\boldsymbol{\nabla}\boldsymbol{u}^T)]$$

$$\boldsymbol{\nabla}\cdot\boldsymbol{u}=0$$

$$\rho c\left(\frac{\partial\theta}{\partial t}+\boldsymbol{u}\cdot\boldsymbol{\nabla}\theta\right)-\boldsymbol{\nabla}\cdot(K\boldsymbol{\nabla}\theta)=2\,\eta\,\mathrm{tr}(\dot{\boldsymbol{\varepsilon}}^2)$$

Constitutive relation

$$\eta(\theta,\boldsymbol{u})=\left\{A_0\exp\left(-\frac{Q}{R\theta}\right)\right\}^{-\frac{1}{n}}\dot{\boldsymbol{\varepsilon}}_{\mathrm{II}}^{\frac{1-n}{2n}} \qquad [\dot{\boldsymbol{\varepsilon}}_{\mathrm{II}}=\tfrac{1}{2}\mathrm{tr}(\dot{\boldsymbol{\varepsilon}}^2)]$$

Boundary conditions

$$\theta|_{\Gamma_{FS}}=\theta_{FS} \qquad \frac{Dz}{Dt}|_{\Gamma_{FS}}=a \qquad \boldsymbol{\sigma}\boldsymbol{n}|_{\Gamma_{FS}}=\boldsymbol{0}$$

$$K\boldsymbol{\nabla}\theta\cdot\boldsymbol{n}|_{\Gamma_B}=q_B \qquad \boldsymbol{u}\cdot\boldsymbol{n}|_{\Gamma_B}=0 \qquad (\boldsymbol{I}-\boldsymbol{n}\otimes\boldsymbol{n})\,(\boldsymbol{\sigma}\boldsymbol{n}+\beta\boldsymbol{u})\,|_{\Gamma_B}=\boldsymbol{0}$$

"Full Stokes" ice sheet model

Balance of linear momentum, mass, and energy

$$-\boldsymbol{\nabla} \cdot [\eta(\theta, \boldsymbol{u})\, \dot{\boldsymbol{\varepsilon}} - \boldsymbol{I} p] = \rho \boldsymbol{g} \qquad [\dot{\boldsymbol{\varepsilon}} = \tfrac{1}{2}(\boldsymbol{\nabla} \boldsymbol{u} + \boldsymbol{\nabla} \boldsymbol{u}^T)]$$

$$\boldsymbol{\nabla} \cdot \boldsymbol{u} = 0$$

$$\rho c \left(\frac{\partial \theta}{\partial t} + \boldsymbol{u} \cdot \boldsymbol{\nabla} \theta \right) - \boldsymbol{\nabla} \cdot (K \boldsymbol{\nabla} \theta) = 2\, \eta\, \mathrm{tr}(\dot{\boldsymbol{\varepsilon}}^2)$$

Constitutive relation

Arrhenius-type thinning near melting point

$$\eta(\theta, \boldsymbol{u}) = \left\{ A_0 \exp\left(-\frac{Q}{R\theta} \right) \right\}^{-\frac{1}{n}} \dot{\boldsymbol{\varepsilon}}_{\mathrm{II}}^{\frac{1-n}{2n}} \qquad [\dot{\boldsymbol{\varepsilon}}_{\mathrm{II}} = \tfrac{1}{2}\mathrm{tr}(\dot{\boldsymbol{\varepsilon}}^2)]$$

Boundary conditions

$$\theta|_{\Gamma_{FS}} = \theta_{FS} \qquad \frac{Dz}{Dt}|_{\Gamma_{FS}} = a \qquad \boldsymbol{\sigma n}|_{\Gamma_{FS}} = \boldsymbol{0}$$

$$K \boldsymbol{\nabla} \theta \cdot \boldsymbol{n}|_{\Gamma_B} = q_B \qquad \boldsymbol{u} \cdot \boldsymbol{n}|_{\Gamma_B} = 0 \qquad (\boldsymbol{I} - \boldsymbol{n} \otimes \boldsymbol{n})\, (\boldsymbol{\sigma n} + \beta \boldsymbol{u})\, |_{\Gamma_B} = \boldsymbol{0}$$

"Full Stokes" ice sheet model

Balance of linear momentum, mass, and energy

$$-\boldsymbol{\nabla}\cdot[\eta(\theta,\boldsymbol{u})\,\dot{\boldsymbol{\varepsilon}} - \boldsymbol{I}p] = \rho\boldsymbol{g} \qquad [\dot{\boldsymbol{\varepsilon}} = \tfrac{1}{2}(\boldsymbol{\nabla}\boldsymbol{u} + \boldsymbol{\nabla}\boldsymbol{u}^T)]$$

$$\boldsymbol{\nabla}\cdot\boldsymbol{u} = 0$$

$$\rho c\left(\frac{\partial\theta}{\partial t} + \boldsymbol{u}\cdot\boldsymbol{\nabla}\theta\right) - \boldsymbol{\nabla}\cdot(K\boldsymbol{\nabla}\theta) = 2\,\eta\,\mathrm{tr}(\dot{\boldsymbol{\varepsilon}}^2)$$

Constitutive relation   shear thinning with 2nd strain rate invariant

$$\eta(\theta,\boldsymbol{u}) = \left\{A_0\exp\left(-\frac{Q}{R\theta}\right)\right\}^{-\frac{1}{n}} \dot{\boldsymbol{\varepsilon}}_{\mathrm{II}}^{\frac{1-n}{2n}} \qquad [\dot{\boldsymbol{\varepsilon}}_{\mathrm{II}} = \tfrac{1}{2}\mathrm{tr}(\dot{\boldsymbol{\varepsilon}}^2)]$$

Boundary conditions

$$\theta|_{\Gamma_{FS}} = \theta_{FS} \qquad \frac{Dz}{Dt}|_{\Gamma_{FS}} = a \qquad \boldsymbol{\sigma}\boldsymbol{n}|_{\Gamma_{FS}} = \boldsymbol{0}$$

$$K\boldsymbol{\nabla}\theta\cdot\boldsymbol{n}|_{\Gamma_B} = q_B \qquad \boldsymbol{u}\cdot\boldsymbol{n}|_{\Gamma_B} = 0 \qquad (\boldsymbol{I} - \boldsymbol{n}\otimes\boldsymbol{n})\,(\boldsymbol{\sigma}\boldsymbol{n} + \beta\boldsymbol{u})\,|_{\Gamma_B} = \boldsymbol{0}$$

"Full Stokes" ice sheet model

Balance of linear momentum, mass, and energy

$$-\boldsymbol{\nabla} \cdot [\eta(\theta, \boldsymbol{u})\, \dot{\boldsymbol{\varepsilon}} - \boldsymbol{I} p] = \rho \boldsymbol{g} \qquad [\dot{\boldsymbol{\varepsilon}} = \tfrac{1}{2}(\boldsymbol{\nabla u} + \boldsymbol{\nabla u}^T)]$$

$$\boldsymbol{\nabla} \cdot \boldsymbol{u} = 0$$

$$\rho c \left(\frac{\partial \theta}{\partial t} + \boldsymbol{u} \cdot \boldsymbol{\nabla} \theta \right) - \boldsymbol{\nabla} \cdot (K \boldsymbol{\nabla} \theta) = 2\, \eta\, \mathrm{tr}(\dot{\boldsymbol{\varepsilon}}^2)$$

Constitutive relation

$$\eta(\theta, \boldsymbol{u}) = \left\{ A_0 \exp\left(-\frac{Q}{R\theta} \right) \right\}^{-\frac{1}{n}} \dot{\boldsymbol{\varepsilon}}_{\mathrm{II}}^{\frac{1-n}{2n}} \qquad [\dot{\boldsymbol{\varepsilon}}_{\mathrm{II}} = \tfrac{1}{2} \mathrm{tr}(\dot{\boldsymbol{\varepsilon}}^2)]$$

Boundary conditions

dynamic free surface

$$\theta|_{\Gamma_{FS}} = \theta_{FS} \qquad \frac{Dz}{Dt}|_{\Gamma_{FS}} = a \qquad \boldsymbol{\sigma n}|_{\Gamma_{FS}} = \boldsymbol{0}$$

$$K \boldsymbol{\nabla} \theta \cdot \boldsymbol{n}|_{\Gamma_B} = q_B \qquad \boldsymbol{u} \cdot \boldsymbol{n}|_{\Gamma_B} = 0 \qquad (\boldsymbol{I} - \boldsymbol{n} \otimes \boldsymbol{n})\, (\boldsymbol{\sigma n} + \beta \boldsymbol{u})\, |_{\Gamma_B} = \boldsymbol{0}$$

"Full Stokes" ice sheet model

Balance of linear momentum, mass, and energy

$$-\boldsymbol{\nabla} \cdot [\eta(\theta, \boldsymbol{u})\, \dot{\boldsymbol{\varepsilon}} - \boldsymbol{I} p] = \rho \boldsymbol{g} \qquad [\dot{\boldsymbol{\varepsilon}} = \tfrac{1}{2}(\boldsymbol{\nabla u} + \boldsymbol{\nabla u}^T)]$$

$$\boldsymbol{\nabla} \cdot \boldsymbol{u} = 0$$

$$\rho c \left(\frac{\partial \theta}{\partial t} + \boldsymbol{u} \cdot \boldsymbol{\nabla} \theta \right) - \boldsymbol{\nabla} \cdot (K \boldsymbol{\nabla} \theta) = 2\, \eta\, \mathrm{tr}(\dot{\boldsymbol{\varepsilon}}^2)$$

Constitutive relation

$$\eta(\theta, \boldsymbol{u}) = \left\{ A_0 \exp\left(-\frac{Q}{R\theta} \right) \right\}^{-\frac{1}{n}} \dot{\boldsymbol{\varepsilon}}_{\mathrm{II}}^{\frac{1-n}{2n}} \qquad [\dot{\boldsymbol{\varepsilon}}_{\mathrm{II}} = \tfrac{1}{2}\mathrm{tr}(\dot{\boldsymbol{\varepsilon}}^2)]$$

Boundary conditions

basal sliding coefficient

$$\theta|_{\Gamma_{FS}} = \theta_{FS} \qquad \frac{Dz}{Dt}|_{\Gamma_{FS}} = a \qquad \boldsymbol{\sigma n}|_{\Gamma_{FS}} = \boldsymbol{0}$$

$$K \boldsymbol{\nabla} \theta \cdot \boldsymbol{n}|_{\Gamma_B} = q_B \qquad \boldsymbol{u} \cdot \boldsymbol{n}|_{\Gamma_B} = 0 \qquad (\boldsymbol{I} - \boldsymbol{n} \otimes \boldsymbol{n})\, (\boldsymbol{\sigma n} + \beta \boldsymbol{u})\, |_{\Gamma_B} = \boldsymbol{0}$$

Regularization approach to ice sheet inverse problem

Minimize regularized data misfit

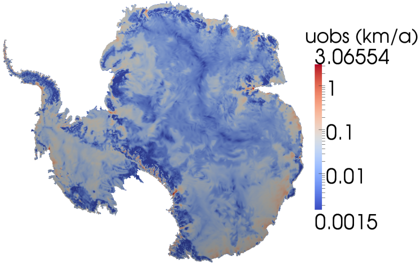
$$\min_{\beta} \mathcal{J}(\beta) := \frac{1}{2}\int_{\Gamma_t} (\mathcal{B}\boldsymbol{u}(\beta) - \boldsymbol{d}_{\text{obs}})^2 \, d\boldsymbol{s} + \frac{\alpha}{2}\int_{\Gamma_b} \nabla_\Gamma \beta \cdot \nabla_\Gamma \beta \, d\boldsymbol{s}$$

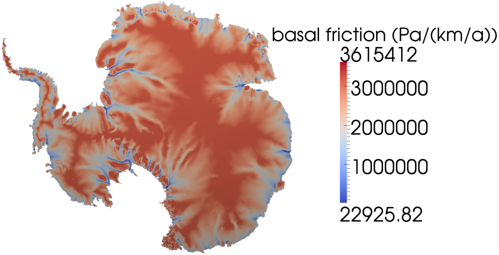
where u and p satisfy the *forward (nonlinear) Stokes equations*

$$\begin{aligned}
\boldsymbol{\nabla} \cdot \boldsymbol{u} &= 0 && \text{in } \Omega \\
-\boldsymbol{\nabla} \cdot [\eta(\boldsymbol{u})(\boldsymbol{\nabla}\boldsymbol{u} + \boldsymbol{\nabla}\boldsymbol{u}^T) - \boldsymbol{I}p] &= \rho \boldsymbol{g} && \text{in } \Omega \\
\boldsymbol{\sigma}_{\boldsymbol{u}}\boldsymbol{n} &= \boldsymbol{0} && \text{on } \Gamma_t \\
\boldsymbol{u} \cdot \boldsymbol{n} = 0, \quad (\boldsymbol{\sigma}_{\boldsymbol{u}}\boldsymbol{n})_\Gamma + \exp(\beta)\boldsymbol{u}_\Gamma &= \boldsymbol{0} && \text{on } \Gamma_b
\end{aligned}$$

- $\boldsymbol{u}$: velocity, β: log basal sliding coefficient field
- $\boldsymbol{d}_{\text{obs}}$: observed surface velocity, α: regularization parameter
- $\mathcal{B}$: observation operator,
- Γ: indicates surface tangential component or surface operator

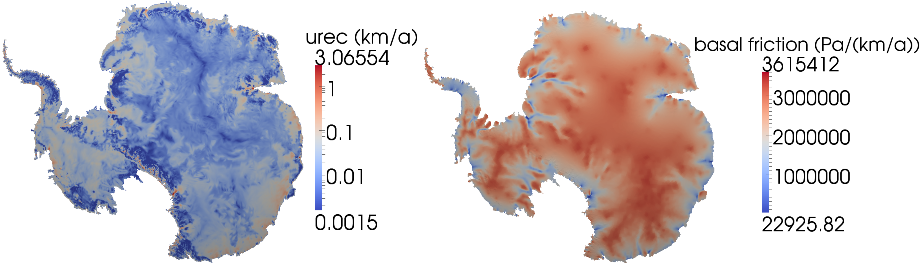
Antarctic ice sheet inversion for basal sliding field: Synthetic data





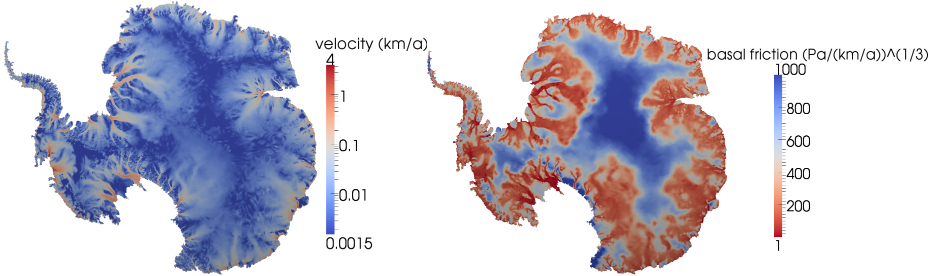
Left: Synthetic surface velocity observations
Right: "Truth" basal sliding field

Antarctic ice sheet inversion for basal sliding field: Synthetic data



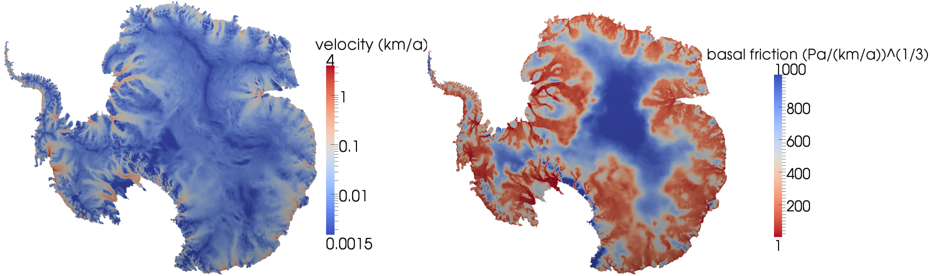
Left: Reconstructed surface velocity field
Right: Inferred basal sliding field

Antarctic ice sheet inversion for basal sliding field: InSAR data



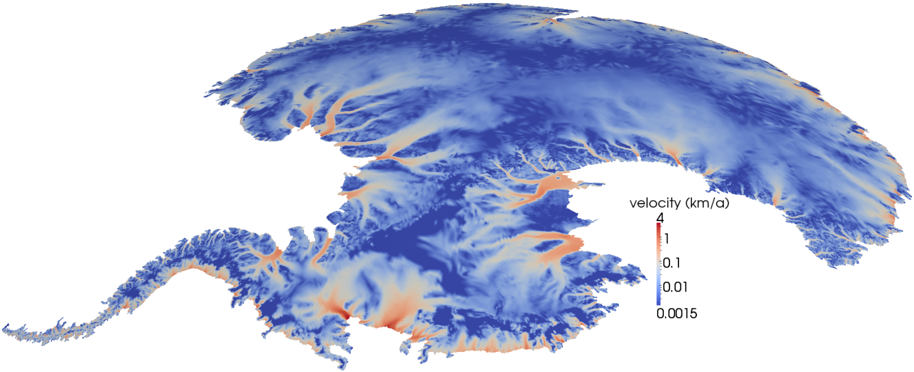
Left: InSAR-based Antarctica ice surface velocity observations
Right: Inferred basal sliding field

Antarctic ice sheet inversion for basal sliding field: InSAR data



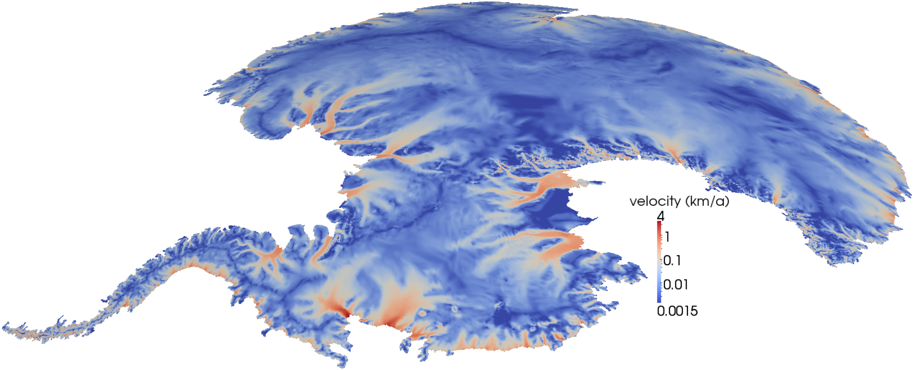
Left: Reconstructed ice surface velocity field
Right: Inferred basal sliding field

Antarctic ice sheet inversion for basal sliding field: InSAR data



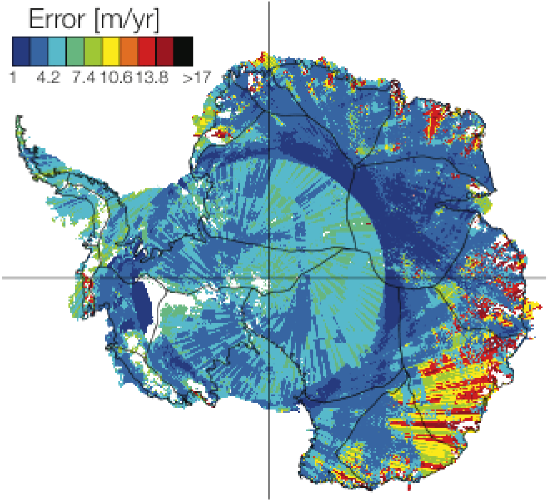
InSAR-based Antarctica ice surface velocity observations

Antarctic ice sheet inversion for basal sliding field: InSAR data



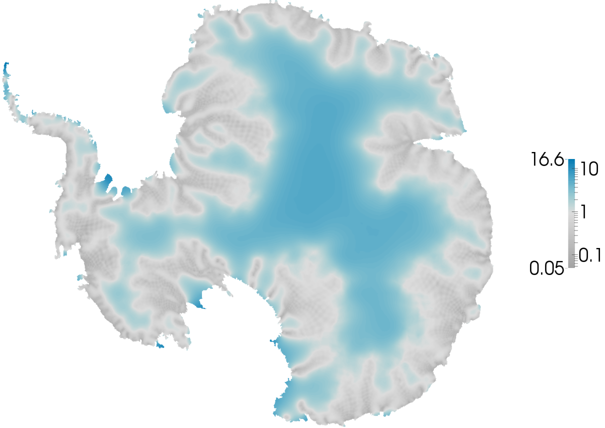
Reconstructed ice surface velocity field

Antarctic ice sheet inversion for basal sliding field: InSAR data



Error in velocity observations

Bayesian approach to inverse problem: Standard deviation of basal friction



Posterior standard deviation