Uncertainty Quantification (UQ) in Global Ocean State Estimation

Alex Kalmikov and Patrick Heimbach

with Carl Wunsch

Department of Earth, Atmospheric and Planetary Sciences (EAPS)

Massachusetts Institute of Technology, Cambridge MA

2013 SIAM Annual Meeting (AN13)

San Diego, CA

July 8 2013

Estimating the Circulation and Climate of the Ocean

http://ecco-group.org

http://mitgcm.org

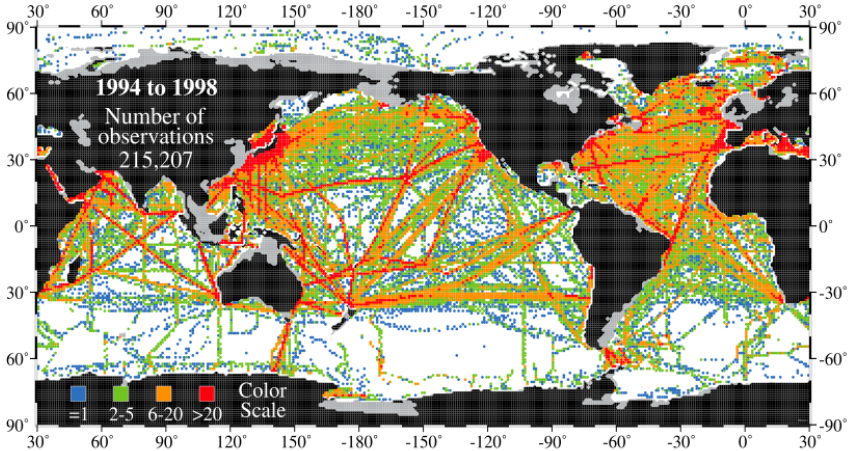
Outline

1. Global ocean state estimation for climate
   - motivation and objectives
2. Quantifying uncertainties of the state estimate and associated climate diagnostics
   - Algorithmic Differentiation (AD)
   - the Hessian method
3. Example: Volume transport through Drake Passage
   - Antarctic Circumpolar Current (ACC)
   - MITgcm in barotropic configuration
4. Results
   - inverse-to-forward UQ propagation
   - uncertainty propagation mechanisms
   - time-resolving UQ
   - posterior uncertainty reduction

Global Ocean State Estimation

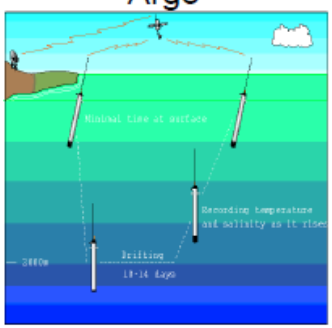
- Combination of ocean **models** with **observations**
- Dynamically and kinematically consistent reconstruction of ocean
  - fields, evolution, dynamics
- “Data assimilation”, “reanalysis”
- **Estimation** (interpolation) vs. **forecasting** (extrapolation) problem
- Sparse observations (*in space & time*), restricted (*mostly*) to upper ocean layers

In-situ observations



T/P, Jason

Argo



How to synthesize?
Estimation / optimal control problem:
Use a **model** (MITgcm) and its **adjoint**:

UQ framework for ocean state estimation

- Estimation problem: Data → Controls → Target

  $\mathbf{y} \rightarrow \mathbf{x} \rightarrow \mathbf{z}$

- Uncertainty problem: Data uncertainty → Controls uncertainty → Target uncertainty

  $\Delta\mathbf{y} \rightarrow \Delta\mathbf{x} \rightarrow \Delta\mathbf{z}$

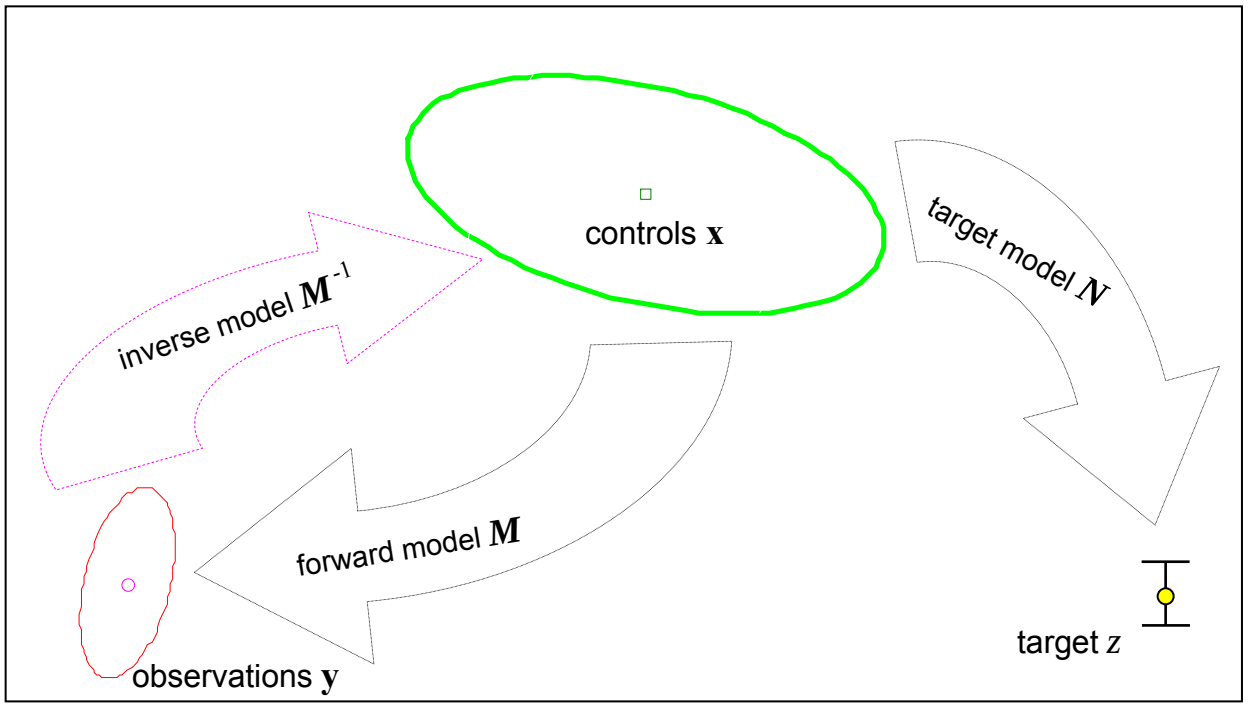
Examples: observed surface elevation

- initial surface elevation
- wind forcing

mass transport

$\mathbf{x}$ – unknowns (e.g. parameters, state, controls, model error)

- Observations (forward) model
  $\mathbf{y} = \boldsymbol{M}(\mathbf{x})$
- Inverse model
  $\mathbf{x} = \boldsymbol{M}^{-1}(\mathbf{y})$
- Target model
  $\mathbf{z} = \boldsymbol{N}(\mathbf{x}) = \boldsymbol{N}\left(\boldsymbol{M}^{-1}(\mathbf{y})\right)$



Adjoint method - estimation

- **Nonlinear Inverse Problem:** $\mathbf{y} = \boldsymbol{M}(\mathbf{x}) \rightarrow \mathbf{x} = \boldsymbol{M}^{-1}(\mathbf{y})$

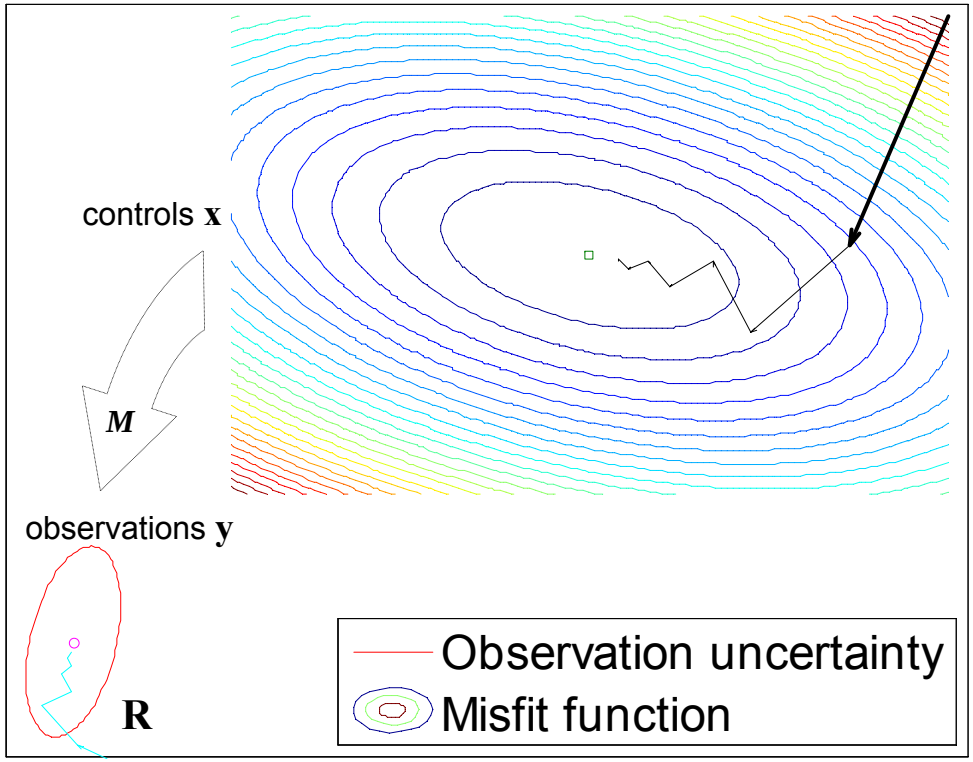
- **Method of Total Inversion** (Tarantola & Valette 1982)

Model–data misfit function: $J(\mathbf{x}) = \frac{1}{2}(\boldsymbol{M}(\mathbf{x}) - \mathbf{y})^T \mathbf{R}^{-1}(\boldsymbol{M}(\mathbf{x}) - \mathbf{y})$

– Weighted by data covariance: $\mathbf{R}$

Minimum of misfit function = Nonlinear least-squares solution $\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} J(\mathbf{x})$

Gradient based optimization =
Adjoint method
(Method of Lagrange multipliers)



Solution uncertainty?

ECCO so far stops here …

Hessian method - uncertainty quantification

- Uncertainty of adjoint method solution → <u>curvature</u> of misfit function!

$$J(\mathbf{x}) = \tfrac{1}{2}(\boldsymbol{M}(\mathbf{x}) - \mathbf{y})^T \mathbf{R}^{-1}(\boldsymbol{M}(\mathbf{x}) - \mathbf{y})$$

- Described by **Hessian matrix** of J

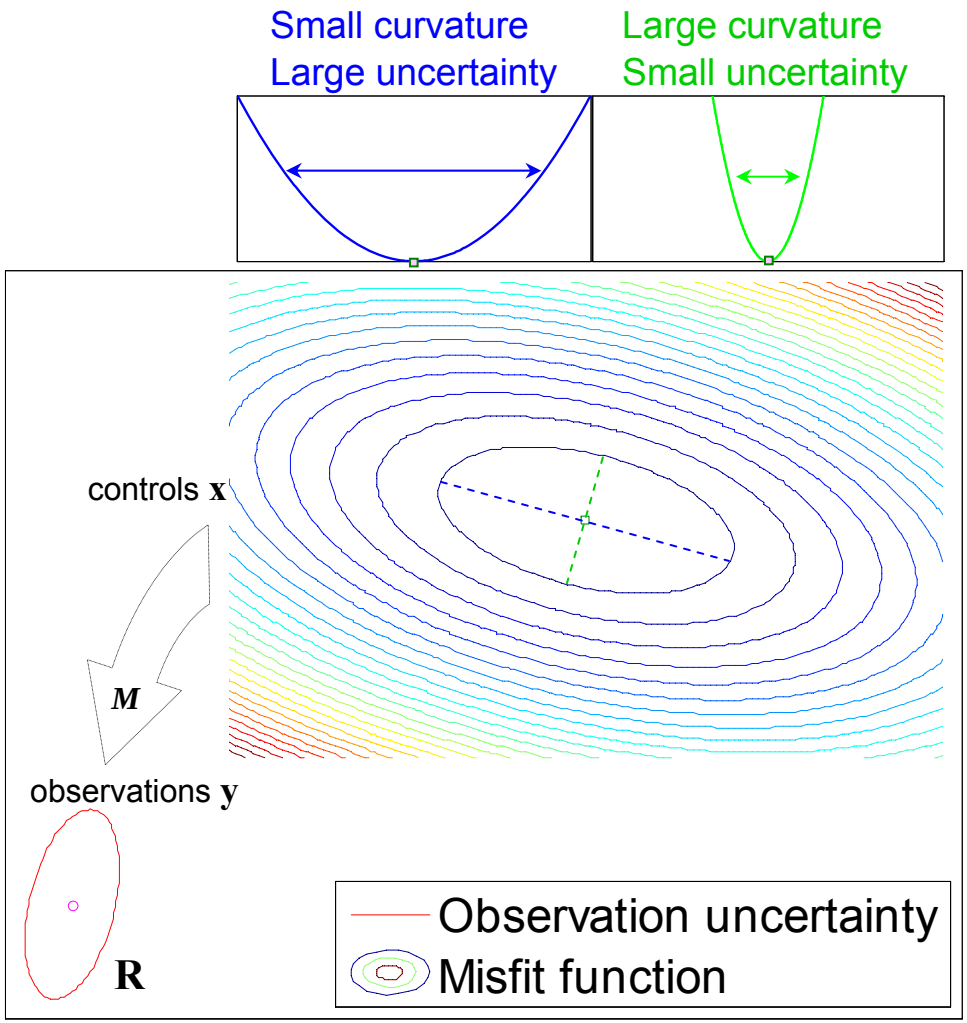
$$\mathbf{H} \equiv \frac{\partial^2 J}{\partial \mathbf{x}^2}(\hat{\mathbf{x}}) =$$

$$\underbrace{\left(\frac{\partial \boldsymbol{M}}{\partial \mathbf{x}^T}\right)^T \mathbf{R}^{-1}\left(\frac{\partial \boldsymbol{M}}{\partial \mathbf{x}^T}\right)}_{\text{Linear term}} + \underbrace{\left(\frac{\partial^2 \boldsymbol{M}}{\partial \mathbf{x}^2}\right)^T \mathbf{R}^{-1}\left(\boldsymbol{M}(\hat{\mathbf{x}}) - \mathbf{y}\right)}_{\text{Nonlinear term}}$$

For linear model or zero residual problem: **Hessian** is determined by product of model **Jacobians**

Posterior covariance of controls $\mathbf{x}$

≈ Inverse of Hessian matrix $\mathbf{P} \approx (\mathbf{H})^{-1}$



Hessian UQ methodology

- Adjoint method (first derivative)
  - State estimation

Data → Controls

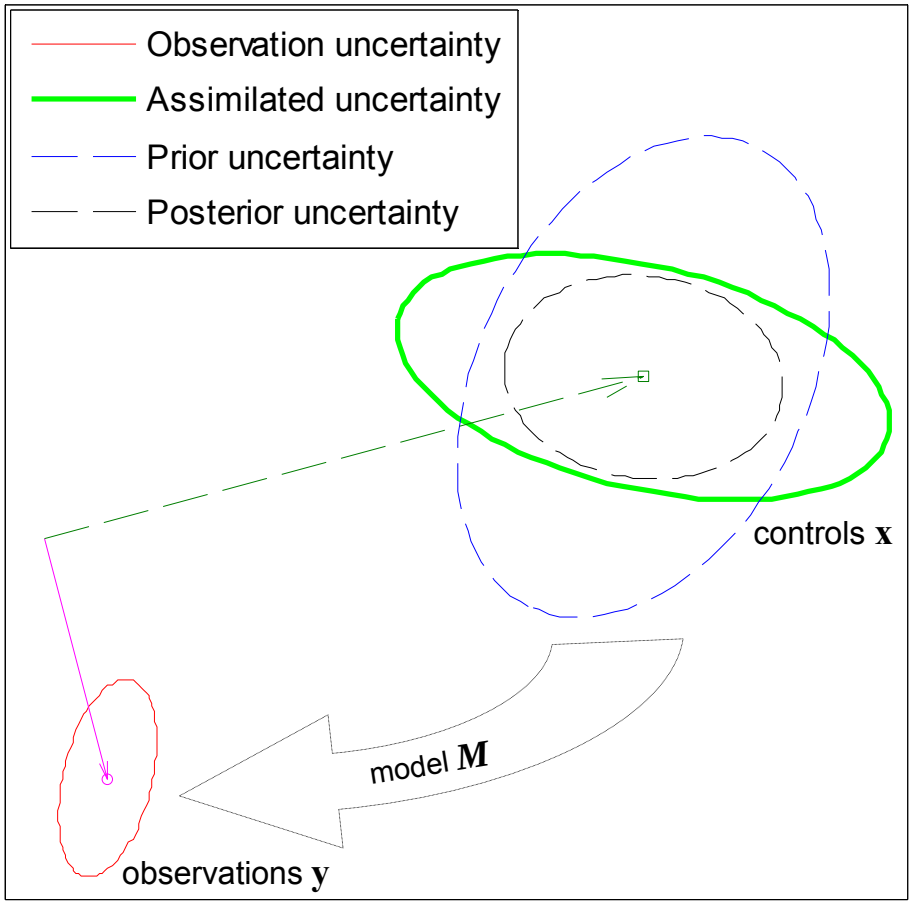
$\mathbf{y} \rightarrow \mathbf{x}$

- Hessian method (second derivative)
  - Uncertainty quantification

Data uncertainty → Controls uncertainty

$\mathbf{P}_{yy} \rightarrow \mathbf{P}_{xx}$

Invertible (well-posed) forward model



- Assimilation of observation uncertainty

$\mathbf{R} \rightarrow \mathbf{P}_{xx} = (\mathbf{H})^{-1}$

not exist for ill-posed problems

- Reduction of prior controls uncertainty

$\mathbf{P}_0$ and $\mathbf{P}_{xx} \rightarrow \mathbf{P} = \left(\mathbf{H} + \mathbf{P}_0^{-1}\right)^{-1}$

Uncertainty assimilation for ill-posed problem

❖ Forward model not invertible

- Assimilation of observation uncertainty

$\mathbf{R} \not\rightarrow \mathbf{P}_{\mathbf{xx}} = (\mathbf{H})^{-1}$ *singular matrix inversion, ill-conditioned inversion*

- Prior-to-posterior uncertainty reduction

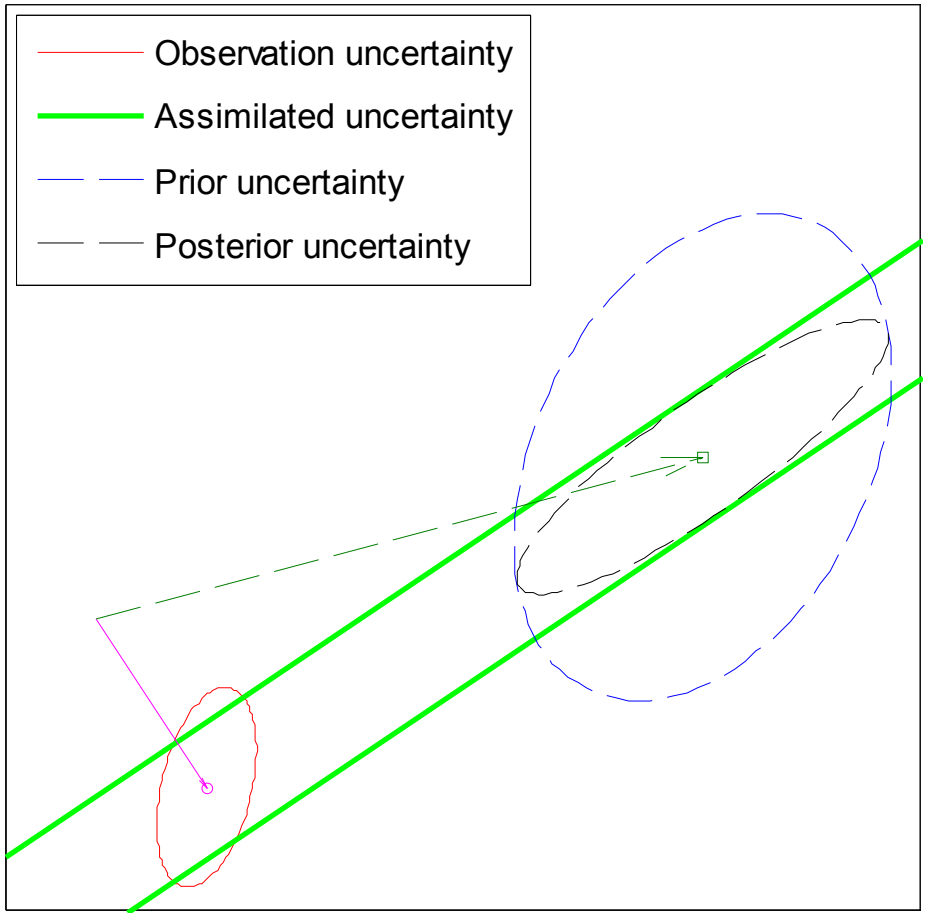
$\mathbf{P}_0$ and "$\mathbf{P}_{\mathbf{xx}}$" $\rightarrow$ $\mathbf{P} = \left(\mathbf{H} + \mathbf{P}_0^{-1}\right)^{-1}$

- Hessian spectral decomposition:

$$\mathbf{H} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T = \sum_i \lambda_i \mathbf{v}_i \mathbf{v}_i^T$$

- Hessian **Nullspace** $\{\mathbf{v}_i \mid \lambda_i = 0\}$
  - not constrained by data
  - infinite uncertainty
- Hessian **Range** $\{\mathbf{v}_i \mid \lambda_i \neq 0\}$
  - data-supported sub-domain
  - finite uncertainty constrained by data

➢ use nullspace to compress Hessian inversion



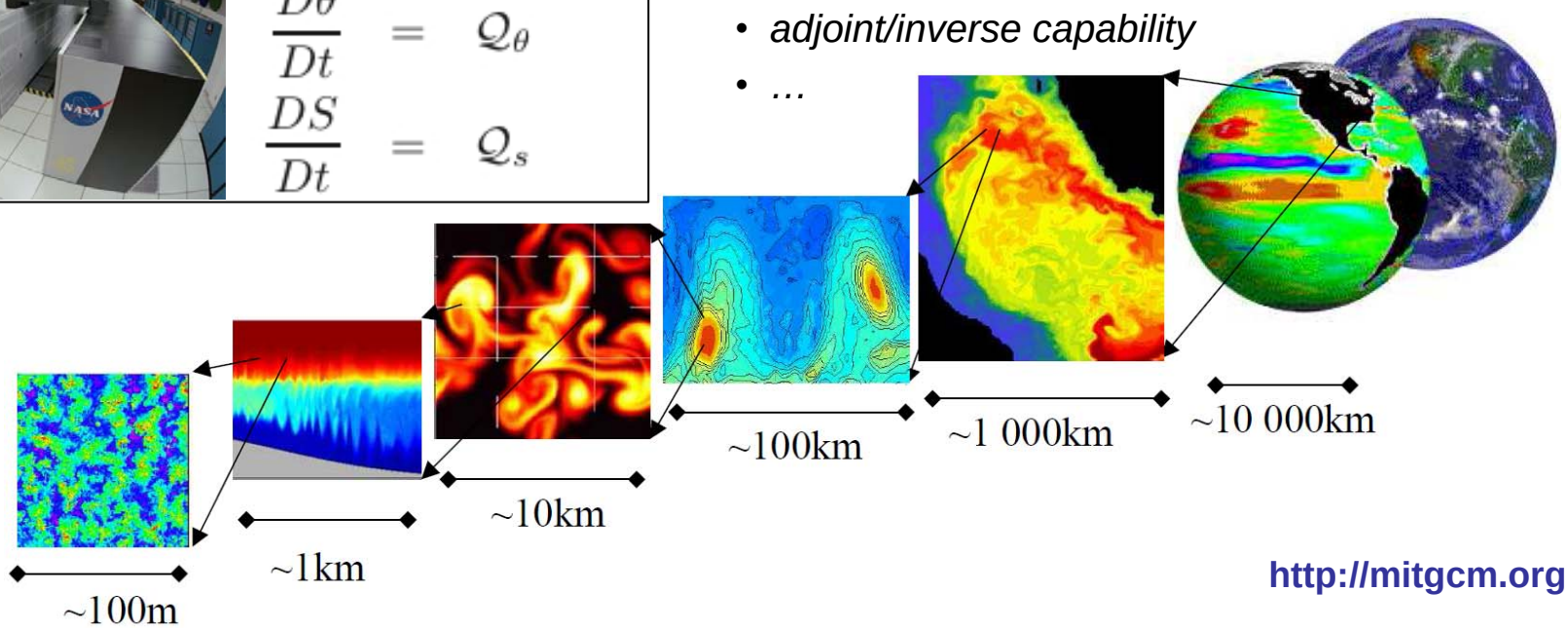
Implementation in MIT general circulation model (MITgcm)

- Key technical achievement
- Simplified configuration (2D, not all parameterizations)

$$
\begin{aligned}
\frac{D\vec{\mathbf{v}}_h}{Dt} + f\hat{\mathbf{k}} \times \vec{\mathbf{v}}_h + \frac{1}{\rho_c}\nabla_z p &= \vec{\mathcal{F}} \\
\epsilon_{nh}\frac{Dw}{Dt} + \frac{g\rho}{\rho_c} + \frac{1}{\rho_c}\frac{\partial p}{\partial z} &= \epsilon_{nh}\mathcal{F}_w \\
\nabla_z \cdot \vec{\mathbf{v}}_h + \frac{\partial w}{\partial z} &= 0 \\
\rho &= \rho(\theta, S) \\
\frac{D\theta}{Dt} &= \mathcal{Q}_\theta \\
\frac{DS}{Dt} &= \mathcal{Q}_s
\end{aligned}
$$

Approximated form of Navier-Stokes equations for an incompressible fluid, consisting of:

- *momentum equation (including Coriolis term)*
- *conservation of mass (continuity equation)*
- *conservation of tracers (heat, salt)*
- *equation of state*
- *subgrid-scale parameterizations*
- *scalable (domain decomposition)*
- *general curvilinear grid (incl. cubed-sphere)*
- *adjoint/inverse capability*
- *…*



http://mitgcm.org

Hessian computation and inversion

- **A**lgorithmic **D**ifferentiation (**AD**) of Fortran code
  - "forward-over-reverse" differentiation of adjoint code
- **Matrix-free** Hessian evaluation, Hessian-vector product code $\mathbf{H}\cdot\begin{bmatrix}0\\1\\0\\\vdots\\0\end{bmatrix}=col_2(\mathbf{H})$
- Lanczos matrix-free eigendecomposition
- Large matrix – $O(10^{10})$ elements, O(60GB) of storage

Matrix inversion dimensionality reduction:

Control space >> independent observations with nonzero information

$\dim(\mathbf{H}) >> \mathrm{rank}(\mathbf{H})$

$O(10^5)$ >> $O(100)$

→ *Details in Kalmikov and Heimbach (submitted to SISC, 2013)*

- Direct differentiation with AD!
  - **Hessian code**

→ Required for nonlinear and large residual problems

- Linearized Hessian
  - **Jacobian code**

$$\mathbf{H}_L=\left(\frac{\partial \boldsymbol{M}}{\partial \mathbf{x}^T}\right)^T \mathbf{R}^{-1}\left(\frac{\partial \boldsymbol{M}}{\partial \mathbf{x}^T}\right)$$

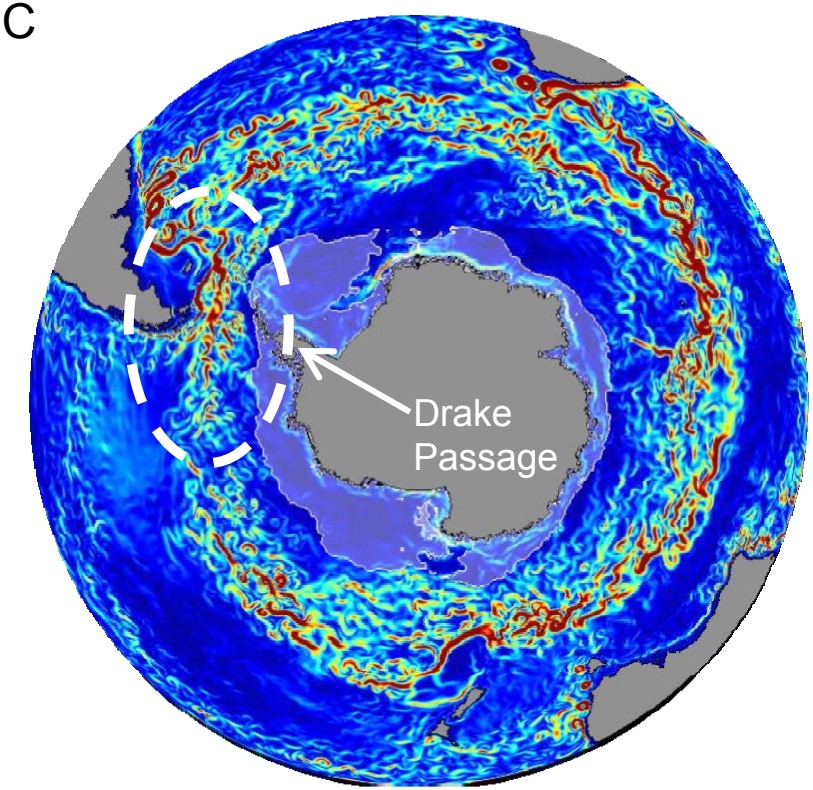
- Low rank Hessian approximation
  - **Adjoint code**

Antarctic Circumpolar Current (ACC) Transport

- Drake Passage – the narrowest constriction of ACC

- Mass transport in Southern Ocean
  - Important oceanographic index
  - Common modeling metric

- Simplified modeling in barotropic framework

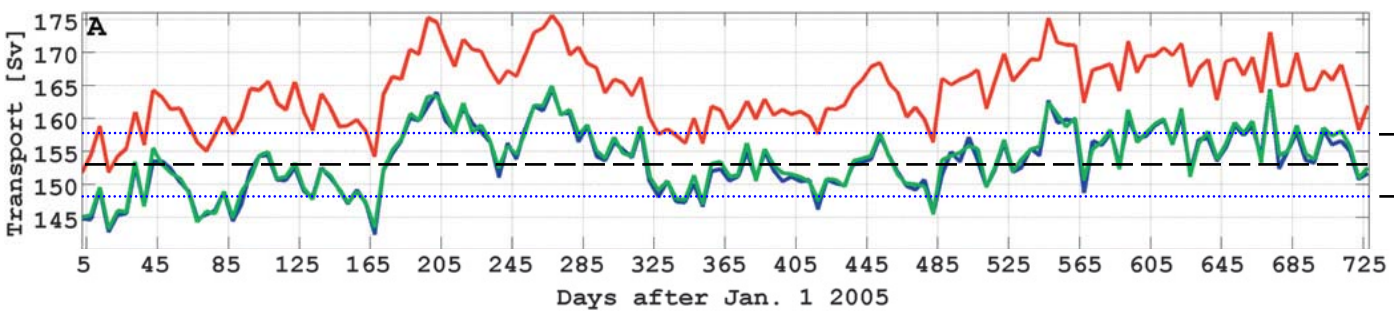


Southern Ocean State Estimation (SOSE)
Mazloff, Heimbach and Wunsch (2010)

- Ad-hoc uncertainty estimates:
- ➢ **Standard deviations of temporal variability**

→ *Need:* **Formal** uncertainty measure based on uncertainties of observations and modeling

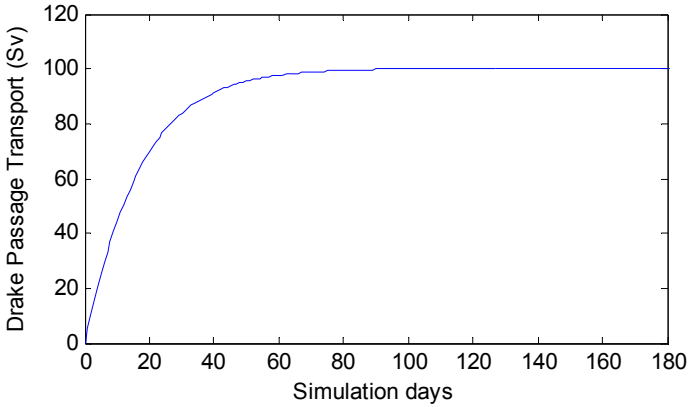
- SOSE Drake Passage transport estimate: 153 ± 5 Sv (Sv = 10^6 m^3/s)

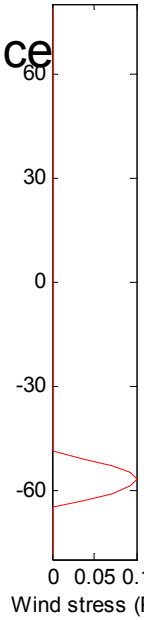


$$\sigma_{\text{variability}}(T) = \sqrt{\left\langle \left(T - \langle T \rangle\right)^2 \right\rangle}$$

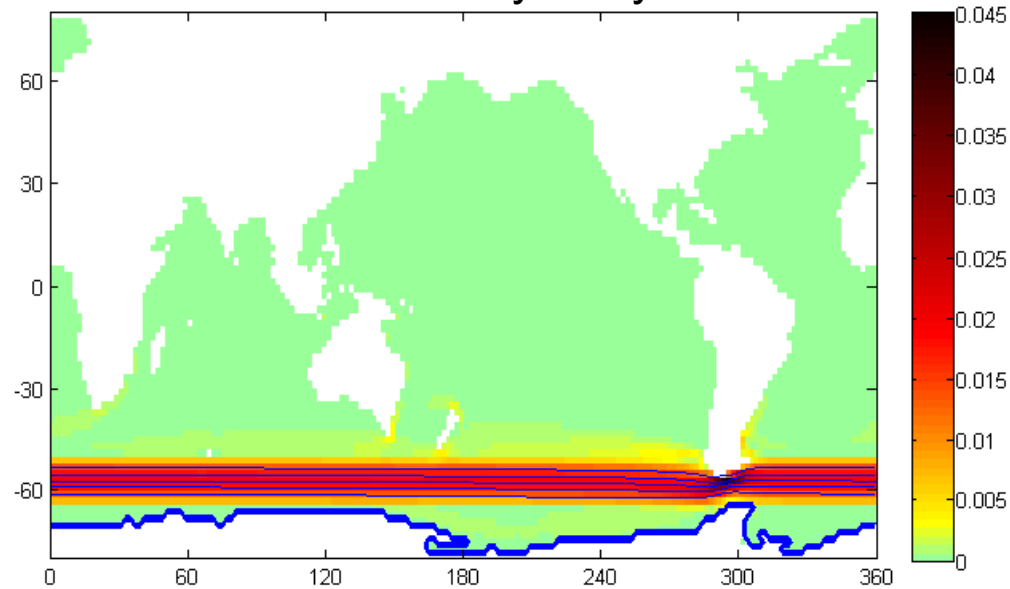
MITgcm modeling of ACC

- Barotropic configuration
- Wind forcing – Bottom drag balance
- Drake Passage transport

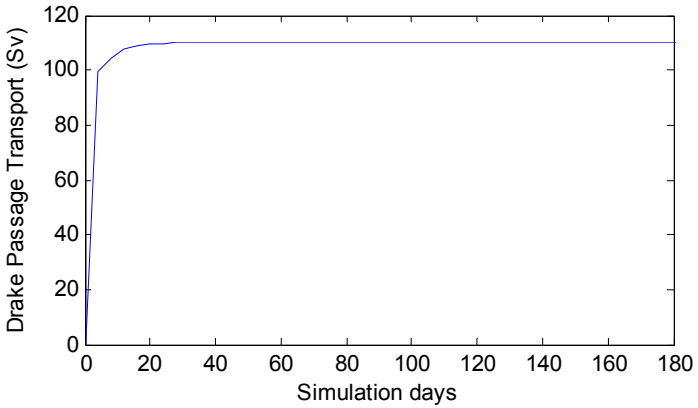


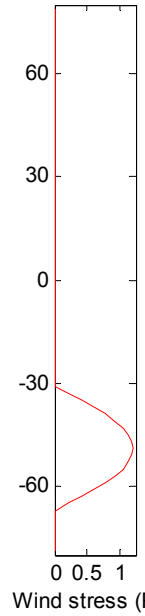


Flat bathymetry

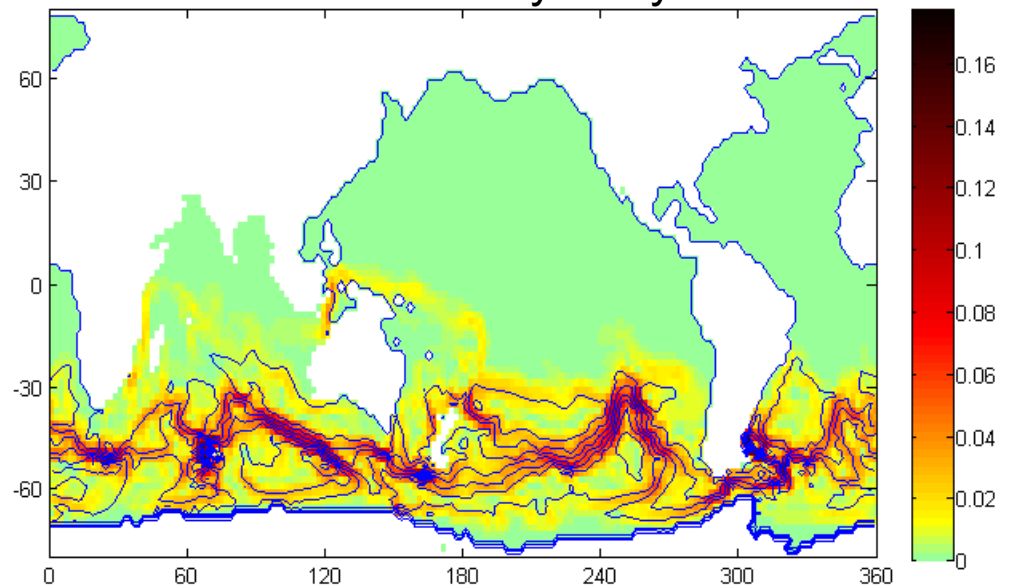


- Constrained by bottom topography
- Flow follows geostrophic contours
- Faster adjustment to steady state





Real bathymetry



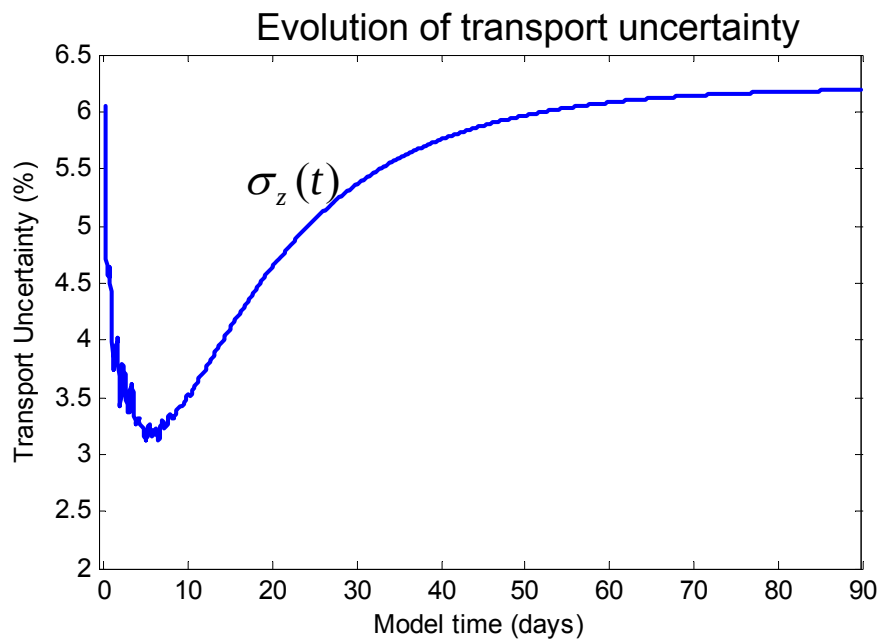
Forward uncertainty propagation

Controls uncertainty → Target uncertainty
$\Delta\mathbf{x}$ → $\Delta\mathbf{z}$

Covariance propagation: $$\sigma_z^2(t) = \left(\frac{\partial N}{\partial \mathbf{x}^T}\right)\mathbf{P}_{\mathbf{xx}}\left(\frac{\partial N}{\partial \mathbf{x}^T}\right)^T$$

- Nonlinear forward model: $z(t) = N(t, \mathbf{x})$
- Model controls: $\mathbf{x} = \left[\underbrace{\tau_x, \tau_y, r}_{\text{BC}}, \underbrace{u_0, v_0, \eta_0}_{\text{IC}}\right]^T$

  (initial & boundary conditions)



- Time dependent model
- Time dependent model derivatives
- Time dependent model uncertainty

A priori control uncertainty:

- correlations unknown
- assume independent **prior uncertainties**

Diagonal prior covariance:

$$\mathbf{P}_{\mathbf{xx}} = \mathbf{P}_0 = \begin{bmatrix} \sigma_{x_1}^2 & 0 & 0 \\ 0 & \sigma_{x_2}^2 & 0 \\ 0 & 0 & \sigma_{x_3}^2 \end{bmatrix}$$

Quadrature rule: $$\sigma_z^2 = \sum_i \left(\frac{\partial N}{\partial x_i}\right)^2 \sigma_{x_i}^2 + \cancel{\sum_i \sum_{j \neq i} \frac{\partial N}{\partial x_i}\frac{\partial N}{\partial x_j}\sigma_{x_i}\sigma_{x_j}\rho_{ij}}$$

- addition of independent uncertainties

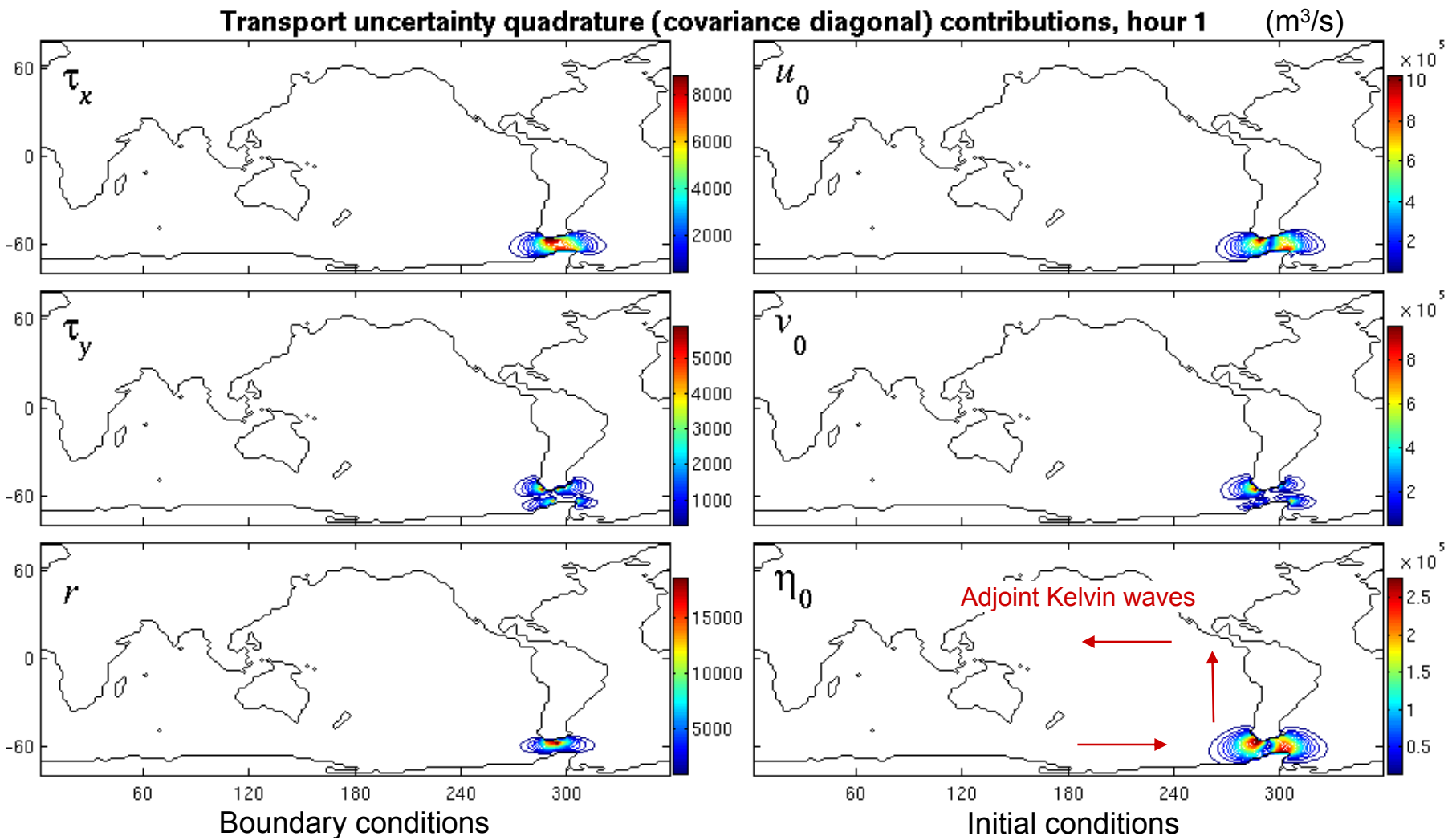
Uncertainty propagation mechanisms

Quadrature sum: $\sigma_z^2 = \sum_i \left(\frac{\partial N}{\partial x_i} \sigma_{x_i} \right)^2$

Control fields map

Wind Forcing X	Current Velocity U
Wind Forcing Y	Current Velocity V
Bottom Drag	Surface Elevation

- **Uncertainty waves**
- Reverse (adjoint) dynamics
- Adjoint Kelvin (gravity) waves



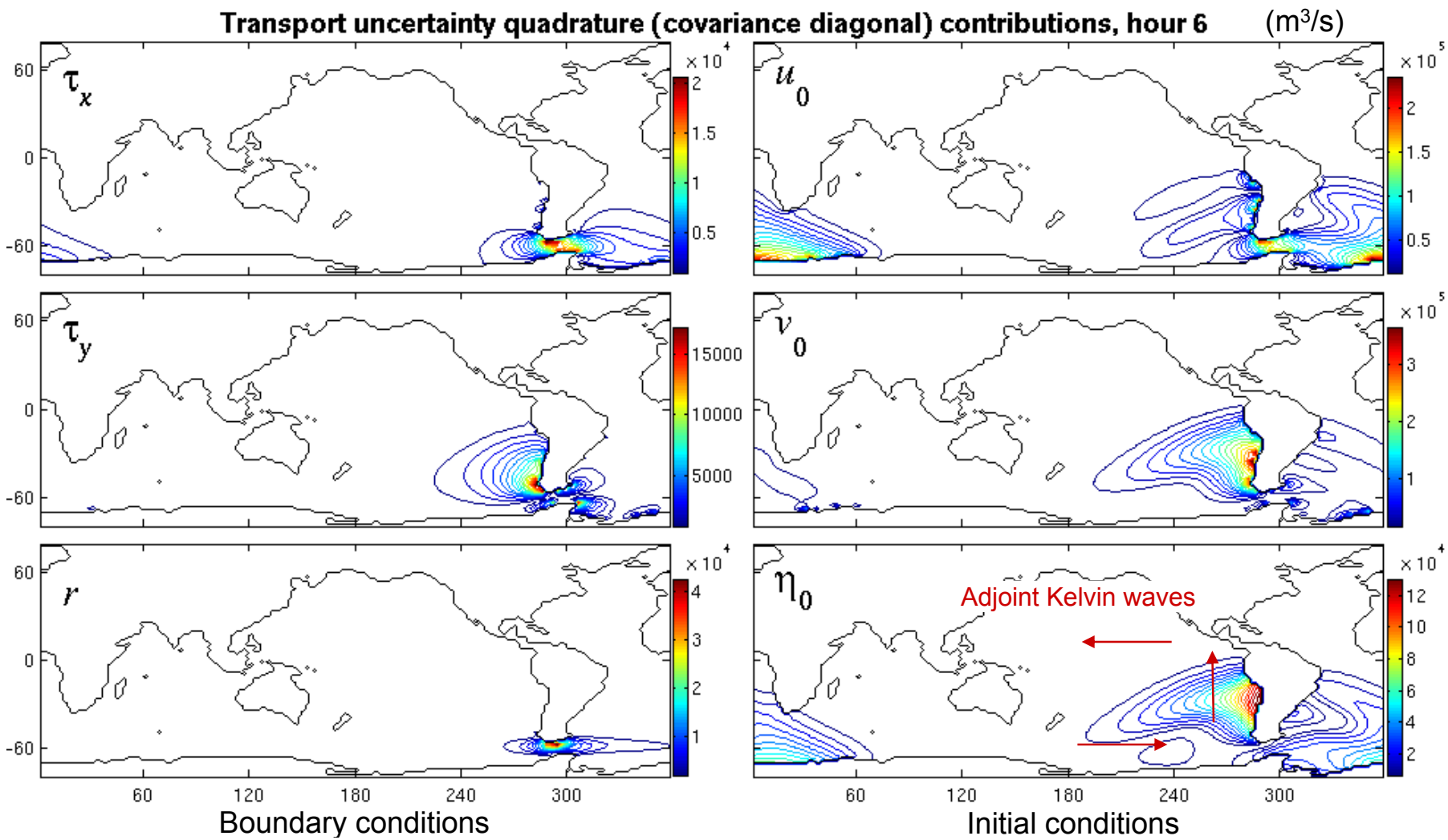
Uncertainty propagation mechanisms

Quadrature sum: $\sigma_z^2 = \sum_i \left(\frac{\partial N}{\partial x_i} \sigma_{x_i} \right)^2$

Control fields map

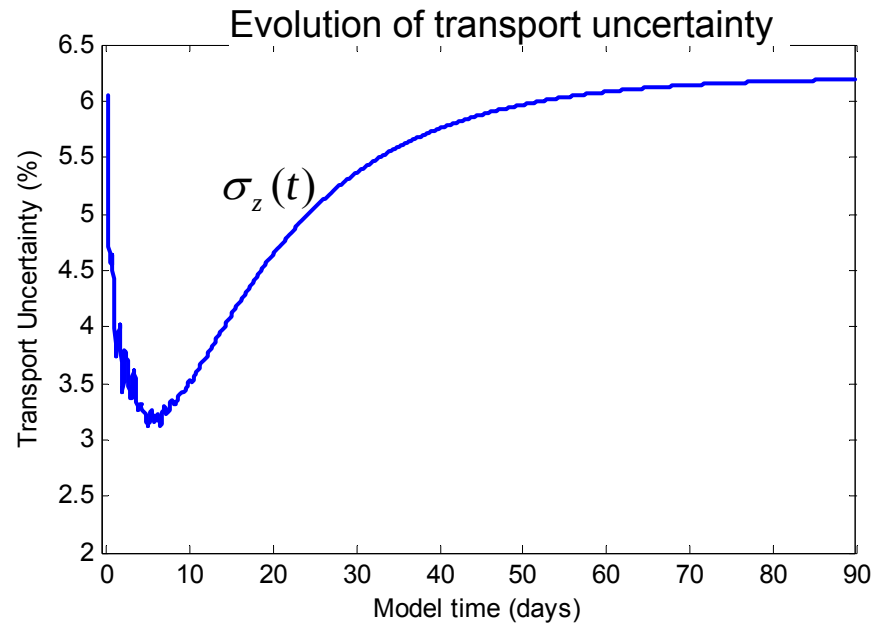
Wind Forcing X	Current Velocity U
Wind Forcing Y	Current Velocity V
Bottom Drag	Surface Elevation

- **Uncertainty waves**
- Reverse (adjoint) dynamics
- Adjoint Kelvin (gravity) waves



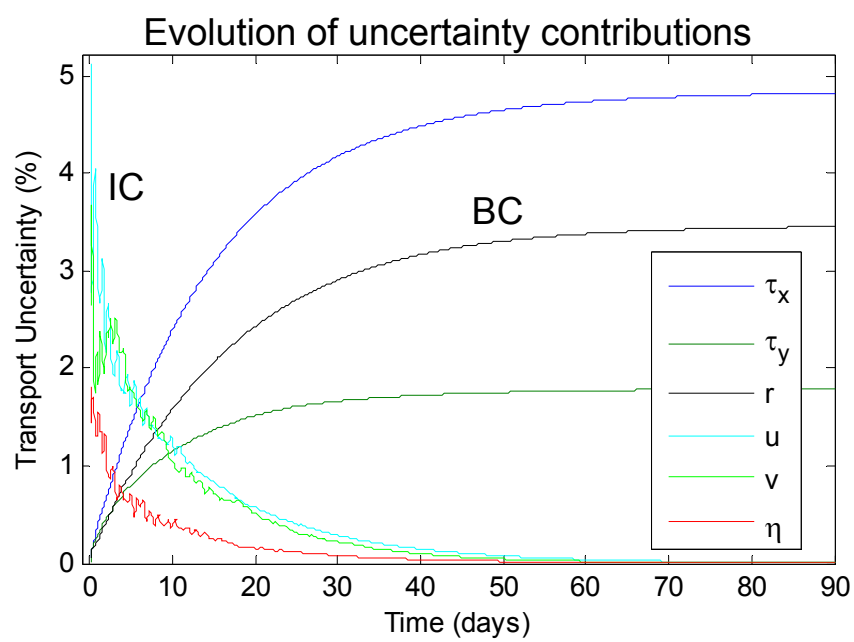
Forward uncertainty propagation dynamics

Covariance propagation: $\sigma_z^2(t) = \left(\frac{\partial N}{\partial \mathbf{x}^T}\right)\mathbf{P}_{\mathbf{xx}}\left(\frac{\partial N}{\partial \mathbf{x}^T}\right)^T$



Understanding the dynamics of uncertainty evolution

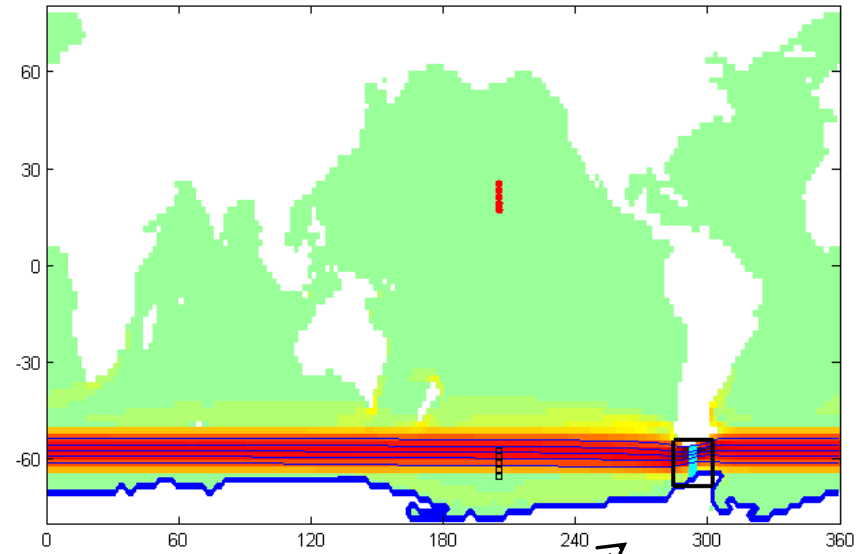
- Transport uncertainty due to **Initial Conditions** decays
  - Transient effects
- Transport uncertainty due to **Boundary Conditions** grows
  - Stationary effects



Quadrature sum: $\sigma_z^2 = \sum_i \left(\frac{\partial N}{\partial x_i}\sigma_{x_i}\right)^2$

Uncertainty assimilation experiments

- *Different observation systems*
- *Locations and geographic coverage*
- *Assimilation times*
- *Number of data points*
- *Type of variables*
- *Different control priors*



- How sea surface elevation observations constrain the Drake Passage transport?
- Effects of the uncertainty of sea surface elevation measurements on the uncertainty of Drake Passage transport?

UQ algorithm schemes and time dependence

Data uncertainty → Controls uncertainty → Target uncertainty

$\mathbf{P}_{\mathbf{yy}}$ → $\mathbf{P}_{\mathbf{xx}}$ → σ_z^2

Pure Prior-independent UQ

$$\mathbf{R} \longrightarrow \boxed{\mathbf{P}_{\mathbf{xx}} = \left(\mathbf{H}\right)^{-1}} \longrightarrow \sigma_z^2 = \mathbf{g}^T \mathbf{P}_{\mathbf{xx}} \mathbf{g}$$

Prior-to-Posterior uncertainty reduction

$$\mathbf{R}(t_a) \longrightarrow \boxed{\begin{array}{l} \mathbf{P}_0 \\ \downarrow \\ \mathbf{P}(t_a) = \left(\mathbf{H} + \mathbf{P}_0^{-1}\right)^{-1} \end{array}} \begin{array}{l} \longrightarrow \sigma_z^2(t) = \mathbf{g}^T \mathbf{P}_0 \mathbf{g} \\ \longrightarrow \sigma_z^2(t_a, t) = \mathbf{g}^T \mathbf{P} \mathbf{g} \end{array}$$

Observation / Assimilation time t_a

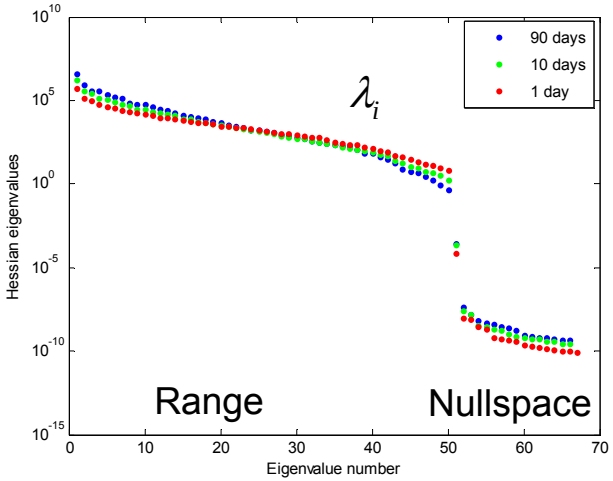
Forward propagation time t

Inverse uncertainty propagation (pure)

Hessian eigenvalue spectra

$$\mathbf{H} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T = \sum_i \lambda_i \mathbf{v}_i \mathbf{v}_i^T$$

- Spectral cut-off
- Estimation nullspace



$$\mathbf{R} \longrightarrow \boxed{\mathbf{P}_{xx} = (\mathbf{H})^{-1}} \longrightarrow \sigma_z^2 = \mathbf{g}^T \mathbf{P}_{xx} \mathbf{g}$$

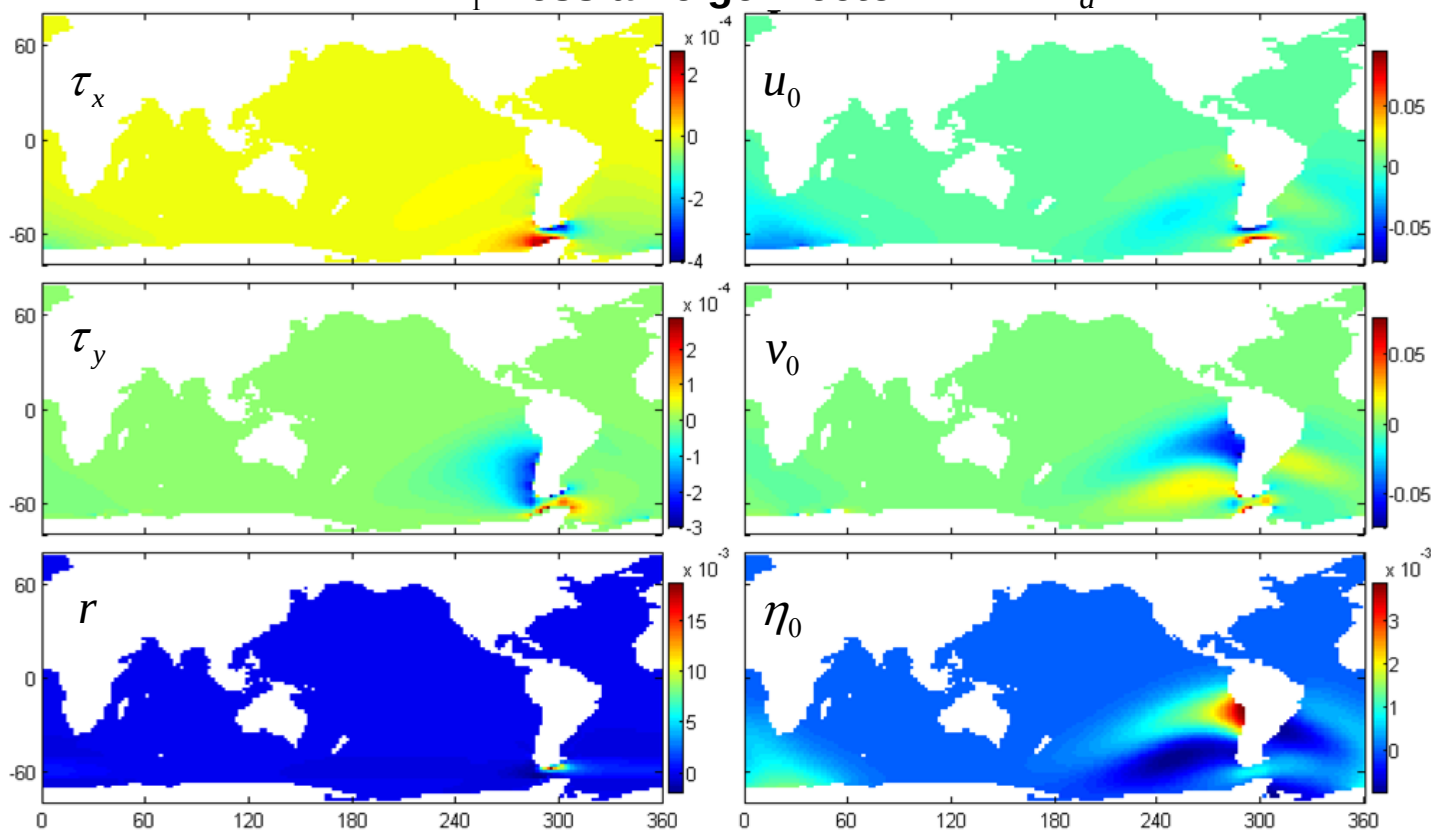
$$\mathbf{P}_0 \longrightarrow \sigma_z^2 = \mathbf{g}^T \mathbf{P}_0 \mathbf{g}$$

$$\downarrow$$

$$\mathbf{R} \longrightarrow \mathbf{P} = \left(\mathbf{H} + \mathbf{P}_0^{-1}\right)^{-1} \longrightarrow \sigma_z^2 = \mathbf{g}^T \mathbf{P} \mathbf{g}$$

- Principal uncertainty patterns
- Wavy propagating structures
  - Hessian waves
  - Uncertainty waves!

$\mathbf{v}_1$ Hessian eigenvector 1 t_a= 6 hours



Model uncertainty reduction

- Constraining independent prior uncertainties

Relative Prior-to-Posterior Uncertainty Reduction

$$\frac{\sigma_{prior} - \sigma_{post}}{\sigma_{prior}} \cdot 100$$

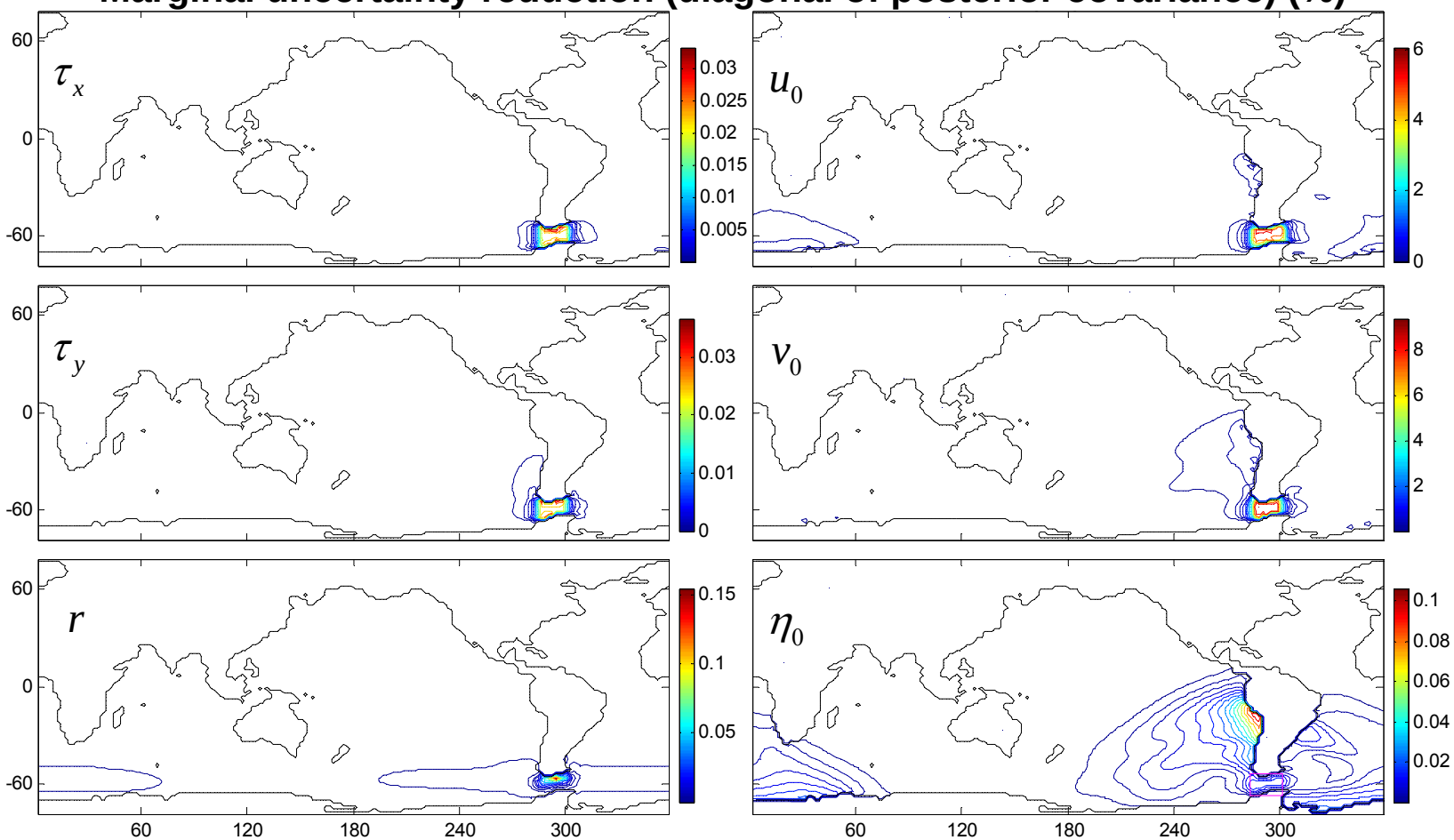
$$\mathbf{R} \longrightarrow \mathbf{P}_{xx} = (\mathbf{H})^{-1} \longrightarrow \sigma_z^2 = \mathbf{g}^T \mathbf{P}_{xx} \mathbf{g}$$

$$\mathbf{P}_0 \longrightarrow \sigma_z^2 = \mathbf{g}^T \mathbf{P}_0 \mathbf{g}$$

$$\mathbf{R} \longrightarrow \mathbf{P} = \left(\mathbf{H} + \mathbf{P}_0^{-1}\right)^{-1} \longrightarrow \sigma_z^2 = \mathbf{g}^T \mathbf{P} \mathbf{g}$$

Marginal uncertainty reduction (diagonal of posterior covariance) (%)

t_a= 6 hours



Boundary conditions

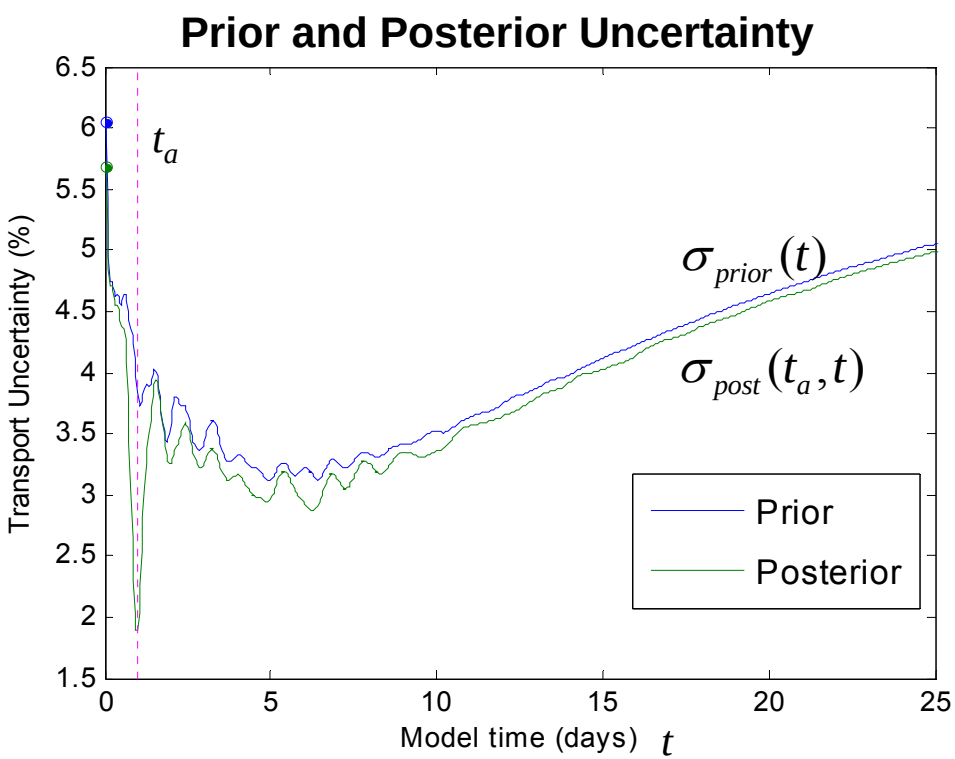
Initial conditions

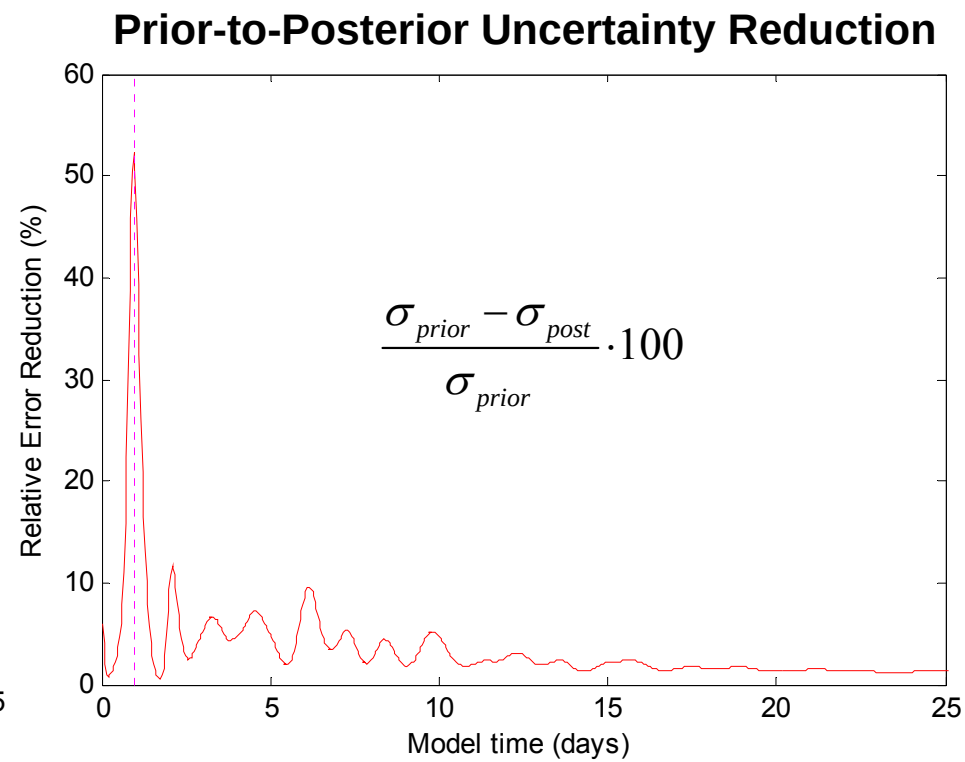
Forward uncertainty propagation & reduction

$$\begin{array}{ccccc} \mathbf{R} & \longrightarrow & \mathbf{P}_{xx} = (\mathbf{H})^{-1} & \longrightarrow & \sigma_z^2 = \mathbf{g}^T \mathbf{P}_{xx} \mathbf{g} \\ & & \mathbf{P}_0 & \longrightarrow & \boxed{\sigma_z^2 = \mathbf{g}^T \mathbf{P}_0 \mathbf{g}} \\ & & \downarrow & & \\ \mathbf{R} & \longrightarrow & \mathbf{P} = \left(\mathbf{H} + \mathbf{P}_0^{-1}\right)^{-1} & \longrightarrow & \boxed{\sigma_z^2 = \mathbf{g}^T \mathbf{P} \mathbf{g}} \end{array}$$

Uncertainty evolution of Drake Passage Transport

- **Time-resolving algorithm**
- Uncertainty assimilation on day 1 (t_a)
- **Transient** uncertainty reduction





Forward uncertainty propagation & reduction

$$\mathbf{R} \longrightarrow \mathbf{P}_{xx} = (\mathbf{H})^{-1} \longrightarrow \sigma_z^2 = \mathbf{g}^T \mathbf{P}_{xx} \mathbf{g}$$

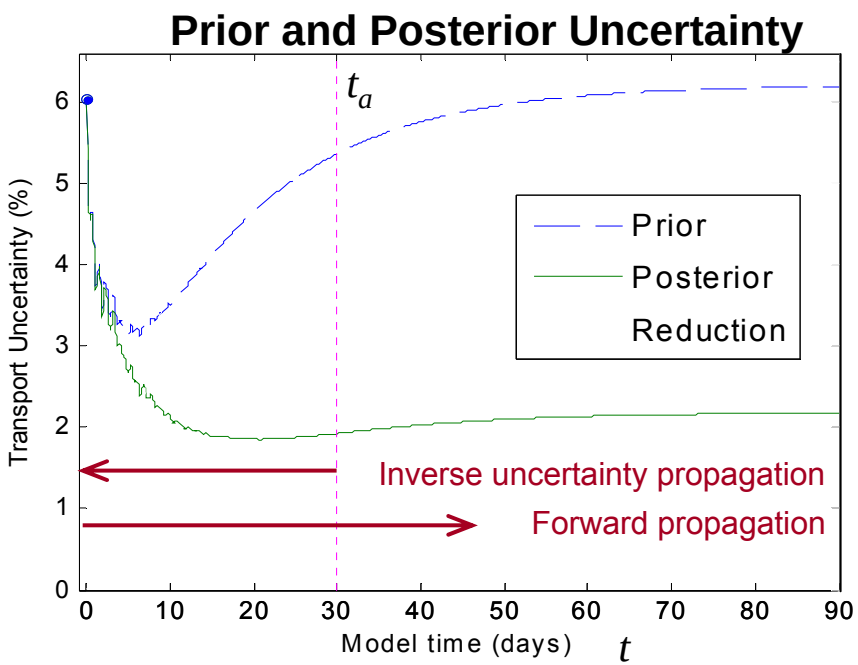
$$\mathbf{P}_0 \longrightarrow \sigma_z^2 = \mathbf{g}^T \mathbf{P}_0 \mathbf{g}$$

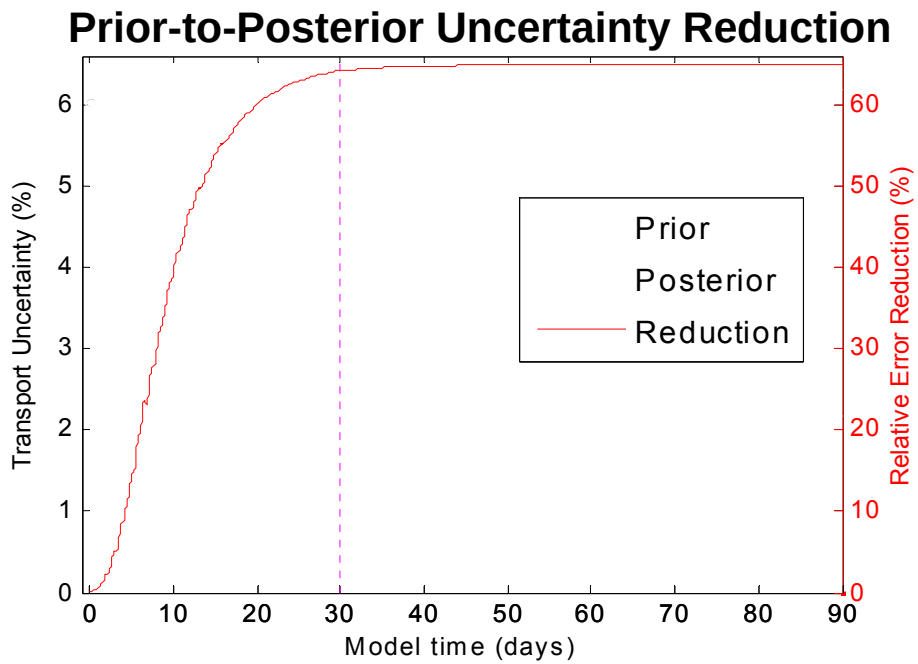
$$\downarrow$$

$$\mathbf{R} \longrightarrow \mathbf{P} = \left(\mathbf{H} + \mathbf{P}_0^{-1}\right)^{-1} \longrightarrow \sigma_z^2 = \mathbf{g}^T \mathbf{P} \mathbf{g}$$

Uncertainty evolution of Drake Passage Transport

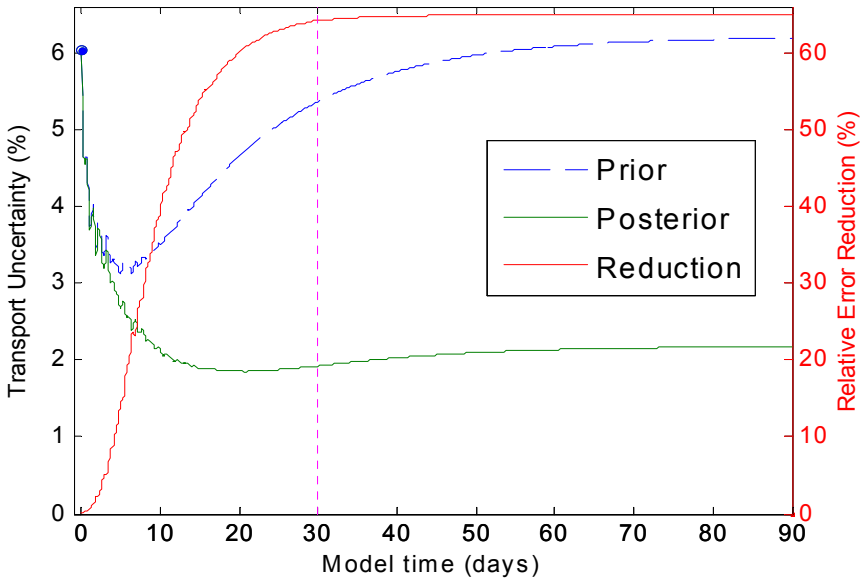
- **Time-resolving algorithm**
- Uncertainty assimilation on day 30 (t_a)
- **Stationary** uncertainty reduction

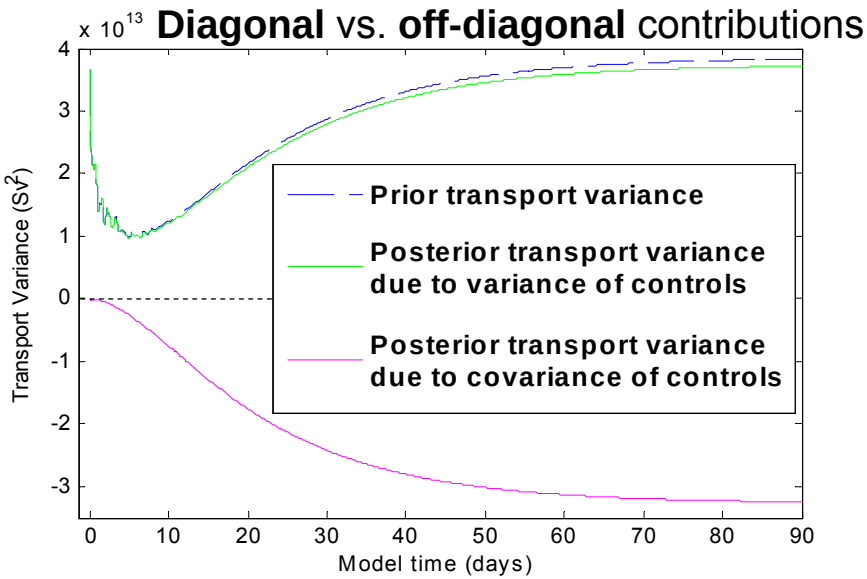




Posterior covariance coupling

- Uncertainty assimilation on day 30 (t_a)
- **Stationary** uncertainty reduction





$$\mathbf{R} \longrightarrow \mathbf{P}_{xx} = (\mathbf{H})^{-1} \longrightarrow \sigma_z^2 = \mathbf{g}^T \mathbf{P}_{xx} \mathbf{g}$$

$$\mathbf{P}_0 \longrightarrow \sigma_z^2 = \mathbf{g}^T \mathbf{P}_0 \mathbf{g}$$

$$\mathbf{P}_0 \downarrow$$

$$\mathbf{R} \longrightarrow \mathbf{P} = \left(\mathbf{H} + \mathbf{P}_0^{-1}\right)^{-1} \longrightarrow \sigma_z^2 = \mathbf{g}^T \mathbf{P} \mathbf{g}$$

- **Diagonal** terms of posterior covariance
  vs.
- **Off-diagonal** terms of posterior covariance

$$\sigma_z^2 = \sum_i \left(\frac{\partial N}{\partial x_i}\right)^2 \sigma_{x_i}^2 + \sum_i \sum_{j \neq i} \frac{\partial N}{\partial x_i} \frac{\partial N}{\partial x_j} \sigma_{x_i} \sigma_{x_j} \rho_{ij}$$

- Observation uncertainty assimilation:
  - Reduces uncertainty of model controls (*small*)
  - Couples model controls (*large*)
- Correlations dominate uncertainty reduction of transport

Summary and conclusions

A significant step in development of **formal uncertainty quantification methodology** for large-scale ill-posed inverse problems:

- implemented a Hessian-based UQ framework within a global ocean state estimation system (**ECCO**)
- demonstrated for barotropic transport through Drake Passage within **MITgcm** and synthetic observations

Scalable numerical machinery

- **Inverse** and **Forward** uncertainty propagation
- Prior-independent UQ (pure) & prior-to-posterior uncertainty reduction (regularized UQ)
- **Range-nullspace** decomposition of estimation uncertainty
- Nonlinear AD **Hessian code**
- **Linearized** uncertainty description
- Depends on elimination of uncertainty **nullspace** when possible

Oceanographic insight

- Resolving **time evolution** of uncertainty - t**ransient** and s**tationary** uncertainty regimes
- Identifying the physical **mechanisms** of uncertainty dynamics (waves)
- Uncertainty **coupling** dominates uncertainty **reduction**

• **Next:** baroclinic high-resolution MITgcm with realistic forcing

Thank you!