ODEN INSTITUTE
FOR COMPUTATIONAL ENGINEERING & SCIENCES

ECCO & RELATED IN TEXAS …

Patrick Heimbach, Ivana Escobar, Shreyas Gaikwad, Matthew Goldberg, An T. Nguyen, Helen R. Pillar, Kirstin Schulz, Max Trostel

University of Texas at Austin

Oden Institute | Jackson School of Geosciences | Institute for Geophysics

https://crios-ut.github.io

Towards uncertainty quantification in ECCO

What is Uncertainty Quantification in the context of Data Assimilation?

- *Parameter estimation (calibration)*
  - **uncertainties in the parameters**
- *State estimation / reconstruction / synthesis (interpolation)*
  - **uncertainties in reconstructed state or derived quantities of interest (QoIs)**
- *Forecast initialization (extrapolation)*
  - **uncertainties in the forecast**

Origins of uncertainty in the context of DA

- **Observations:**
  - measurement error & irregular sampling (in space and time)
- **Assimilation scheme:**
  - DA algorithm (and approximations)
  - how observations are ingested in DA
- **Model:**
  - Parametric uncertainties
  - Structural uncertainties (discretization, model inadequacy)
- **All initial & boundary conditions:**
  - Initial state, external forcing, bathymetry, lateral boundaries
- **Use of prior knowledge:**
  - error covariances, representation error

Origins of uncertainty in the context of DA

- **Observations:**
  - measurement error & irregular sampling (in space and time)
- **Assimilation scheme:**
  - DA algorithm (and approximations)
  - how observations are ingested in DA
- ***Model:***
  - ***Parametric uncertainties***
  - Structural uncertainties (discretization, model inadequacy)
- ***All initial & boundary conditions:***
  - ***Initial state, external forcing, bathymetry, lateral boundaries***
- **Use of prior knowledge:**
  - error covariances, representation error

Proposed approaches

Proposed approaches

1. **Impact of surface atmospheric forcing uncertainty on state estimate**
   – via ensemble approaches – perturbation of atmospheric state
   – seek to maintain covariances among atmospheric perturbations
   – derive maximum covariance patterns of atmospheric forcing anomalies in deriving ensemble members
   – perturbations should not use leading modes of variability since we assume to properly represent phasing of these modes

Proposed approaches

2. **Impact of model parameter uncertainty**
   – via *Stochastically Perturbed Parametrization Tendencies Scheme*
   – adding a stochastic component in these subgrid-scale parameters yields improved representation of oceanic variability
   – *E.g., [Palmer et al, ECMWF 2009]; [Juricke, Palmer, Zanna, J. Clim. 2017]; [Subramanian, Juricke, Dueben, Palmer, BAMS 2019]*

3. **Impact of initial condition uncertainty**
   – via online SVD / optimal initial condition perturbation
   – exploring the use of the leading singular vectors to generate ensemble of perturbed initial conditions
   – *E.g., Zanna, Heimbach, Moore, Tziperman, [JPO 2010] / [J. Clim. 2011] / [QJRMS 2012]*

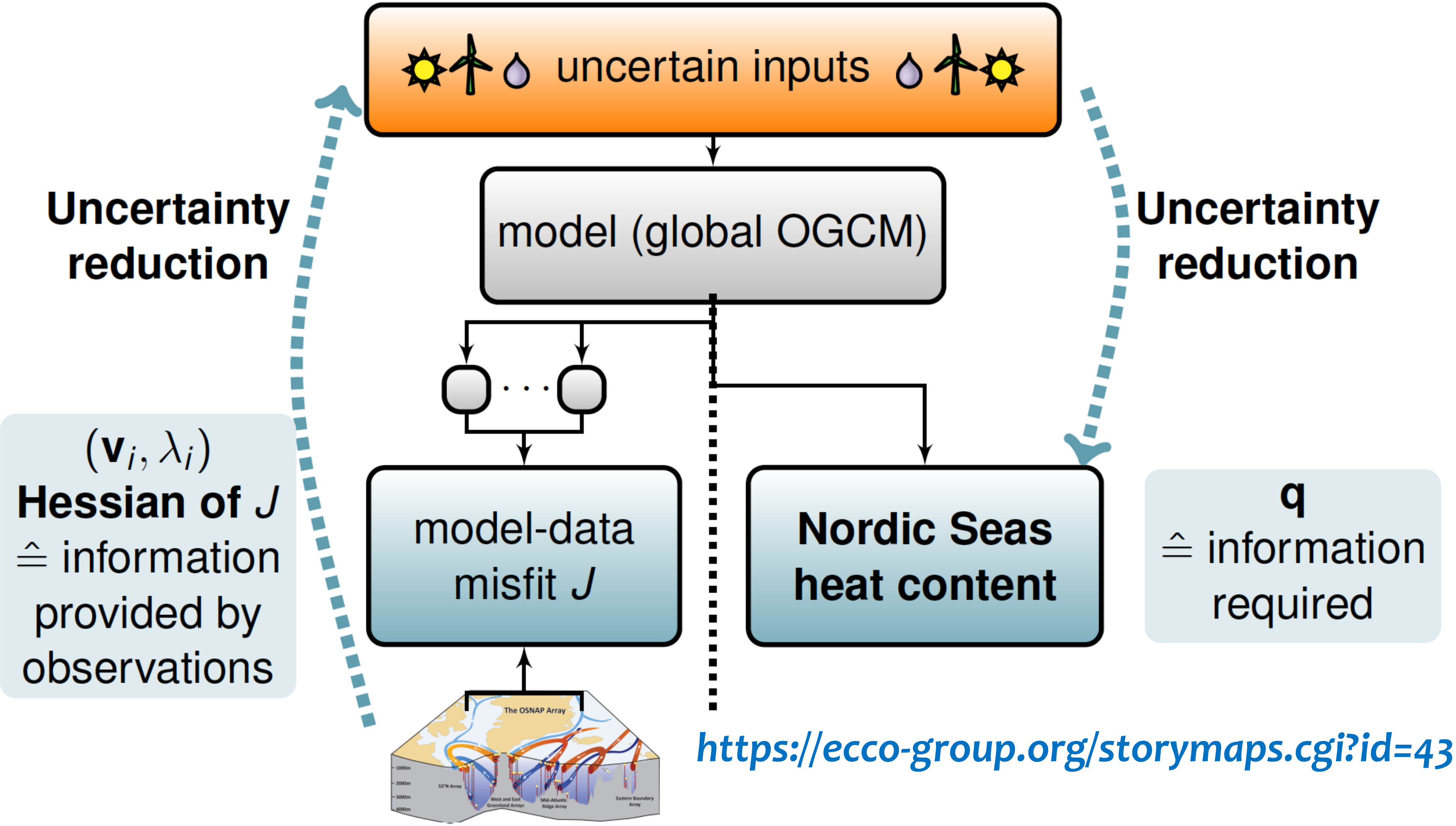
4. Optimal observing system design via Hessian uncertainty quantification

Courtesy Nora Loose, (CU Boulder)

Loose et al. JGR (2020)

Loose et al. J. Adv. Model. Earth Syst. (2021)

Pillar et al., in prep.



https://ecco-group.org/storymaps.cgi?id=43

OSNAP constraints on Nordic Seas heat content
= **uncertainty reduction** along

$$= \sum_i \frac{\lambda_i}{\lambda_i + 1} (\mathbf{v}_i \bullet \mathbf{q})^2$$

ODEN INSTITUTE FOR COMPUTATIONAL ENGINEERING & SCIENCE

Automatic differentiation & adjoints for the masses

Shreyas Gaikwad[1], Laurent Hascoët[2], Krishna Narayanan[3], P.H.[1]

[1] UT Austin, TX

[2] Inria, Sophia-Antipolis, France

[3] Argonne National Lab, Chicago, IL

ODEN INSTITUTE
FOR COMPUTATIONAL ENGINEERING & SCIENCES

Introduction to Tapenade

- TAPENADE is an **Open-Source** Automatic Differentiation (**AD**) Engine developed at Inria
- Developers – Laurent Hascoët and Valérie Pascual
- It takes in as input a source code and list of independent and dependent variables
- Outputs a differentiated code that evaluates the gradients of the dependent variables with respect to the independent variables
- In ML context - this is just like autograd, PyTorch, Tensorflow, etc.

Current capabilities with Tapenade

COMPATIBLE VERIFICATION EXPERIMENTS -

- tutorial_tracer_adjsens
- tutorial_global_oce_biogeo
- isomip
- global_ocean.cs32x15
- OpenAD
- global_with_exf
- halfpipe_streamice
- lab_sea (no ecco)

COMPATIBLE PACKAGES -

- exch2
- gfd
- gmredi
- exf
- cal
- seaice
- thsice
- diagnostics
- cd_code
- streamice
- monitor
- shelfice
- cost
- dic
- gchem
- ggl90
- kpp
- ptracers
- tapenade
- thsice
- ctrl
- grdchk

COMING SOON -
ecco, profiles, mnc

Ease of use with Tapenade

- Setup time **< 1 hour**
- Code in PR #685 - https://github.com/MITgcm/MITgcm/pull/685
- Documentation – See readthedocs associated with PR #685
- The only changes users typically make in their workflow are in red.

```
$ make CLEAN

$ ../../../tools/genmake2 -tap -of ../../../tools/build_options/linux_amd64_ifort -mods ../code_tap

$ make depend
$ make -j 8 tap_adj

$ cd ../run
$ rm -r *
$ ln -s ../input_tap/* .
$ ../input_tap/prepare_run

$ ln -s ../build/mitgcmuv_tap_adj .
$ ./mitgcmuv_tap_adj > output_tap_adj.txt 2>&1
```

Timing analysis – TAF vs Tapenade

- Two experiments
  - globals_with_exf : 100 time-steps of size 0.5 day, 50 days of simulation time
  - global_ocean.cs32x15 : 360 time-steps of size 1 day, 360 days of simulation time
- Experiments run on Sverdrup compute nodes

Time in seconds	TAF	Tapenade	Ratio (Tapenade/TAF)
globals_with_exf	18.60	38.66	2.078
global_ocean.cs32x15	362	1025	2.831

https://crios-ut.github.io

ODEN INSTITUTE
FOR COMPUTATIONAL ENGINEERING & SCIENCES