Parameter Estimation and Uncertainty Quantification for Climate Modeling

- or -

too many simulations, too little time

Matthew Hecht, James Gattiker
and David Higdon
Los Alamos National Laboratory

too many simulations, too little time...

- when quantifying uncertainty in climate model results
  - ensembles are always small
    - even before consideration of input parameter variation
- when selecting optimal input parameters
  - particular when # parameters > one

Los Alamos NATIONAL LABORATORY EST.1943

U.S. DEPARTMENT OF ENERGY | Office of Science

UQ example

Los Alamos NATIONAL LABORATORY EST. 1943

U.S. DEPARTMENT OF ENERGY | Office of Science

Quantile Estimation in a Climate Model

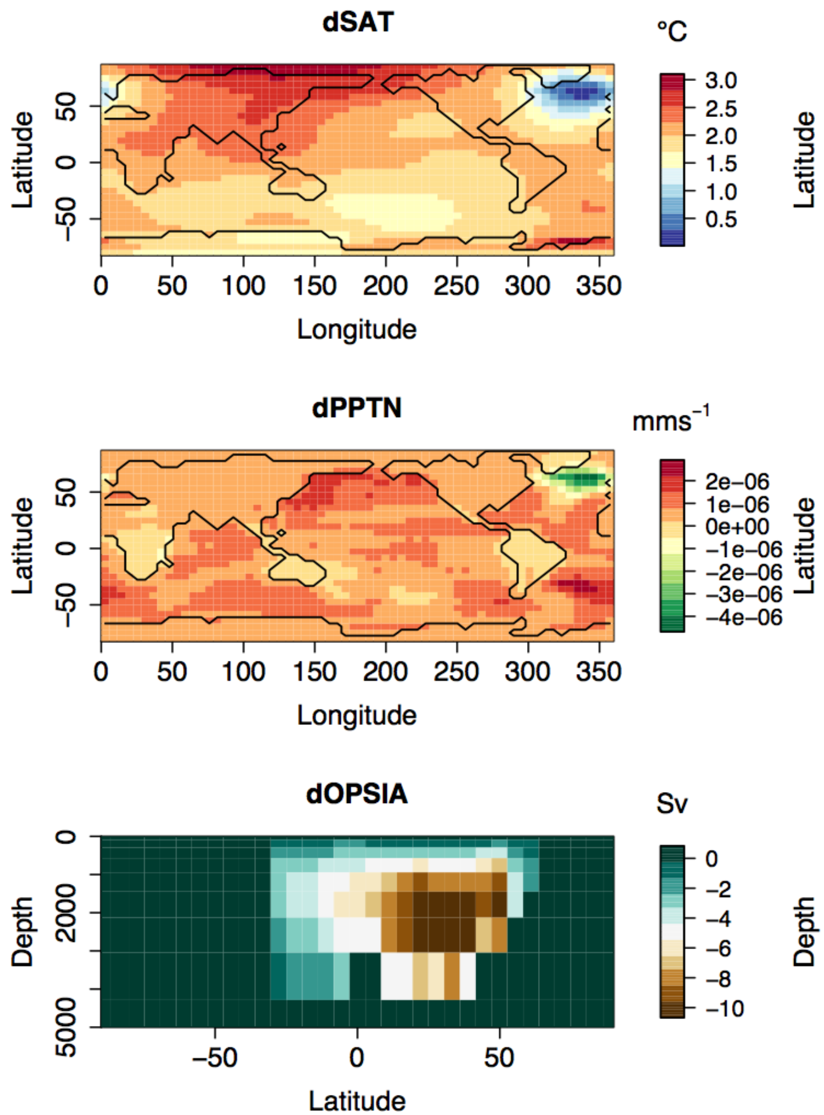
LA-UR 09-03674

James R. Gattiker, Scott Vander Wiel [1]
Doug McNeall[2], Peter Challenor [3]

[1]Statistical Sciences, Los Alamos Nat'l Lab
[2] UK Met Office
[3] UK National Oceanography Center

June 11, 2009

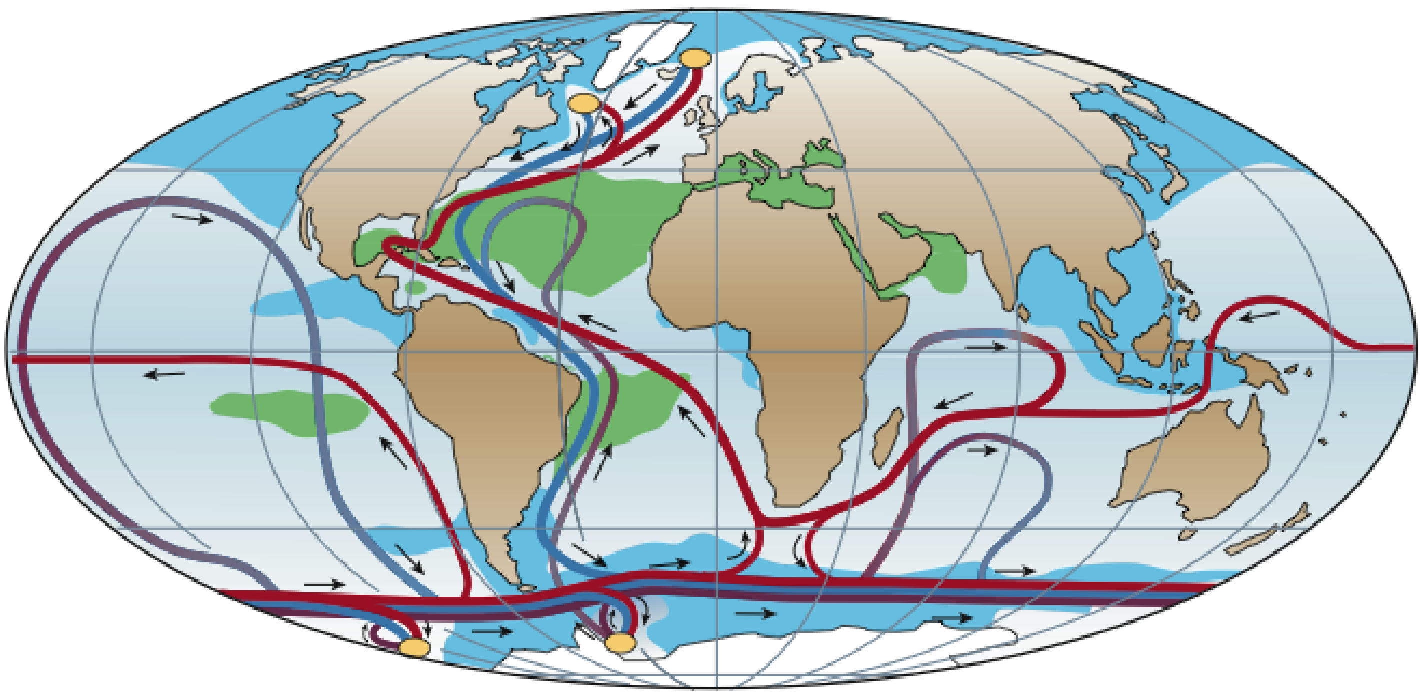
The Problem



We have the climate model GENIE-1:

- Intermediate complexity model.
- 64 long. $\times$ 30 lat. $\times$ 8 depth.
- Deterministic simulation.
- Derive a scalar measure of MOC strength.
- Limited runs available.

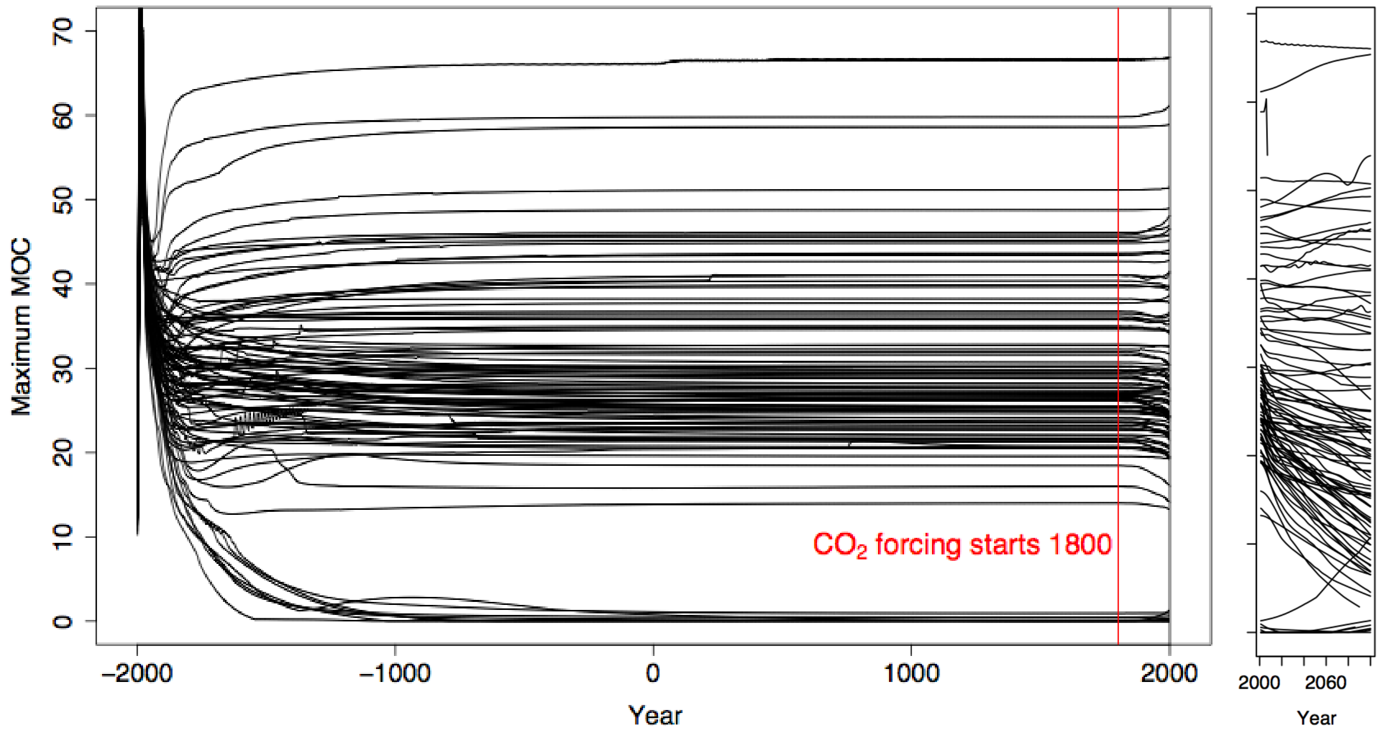
The Problem



From that model, we would like to estimate the likely decrease in MOC circulation by 2100.

The Approach

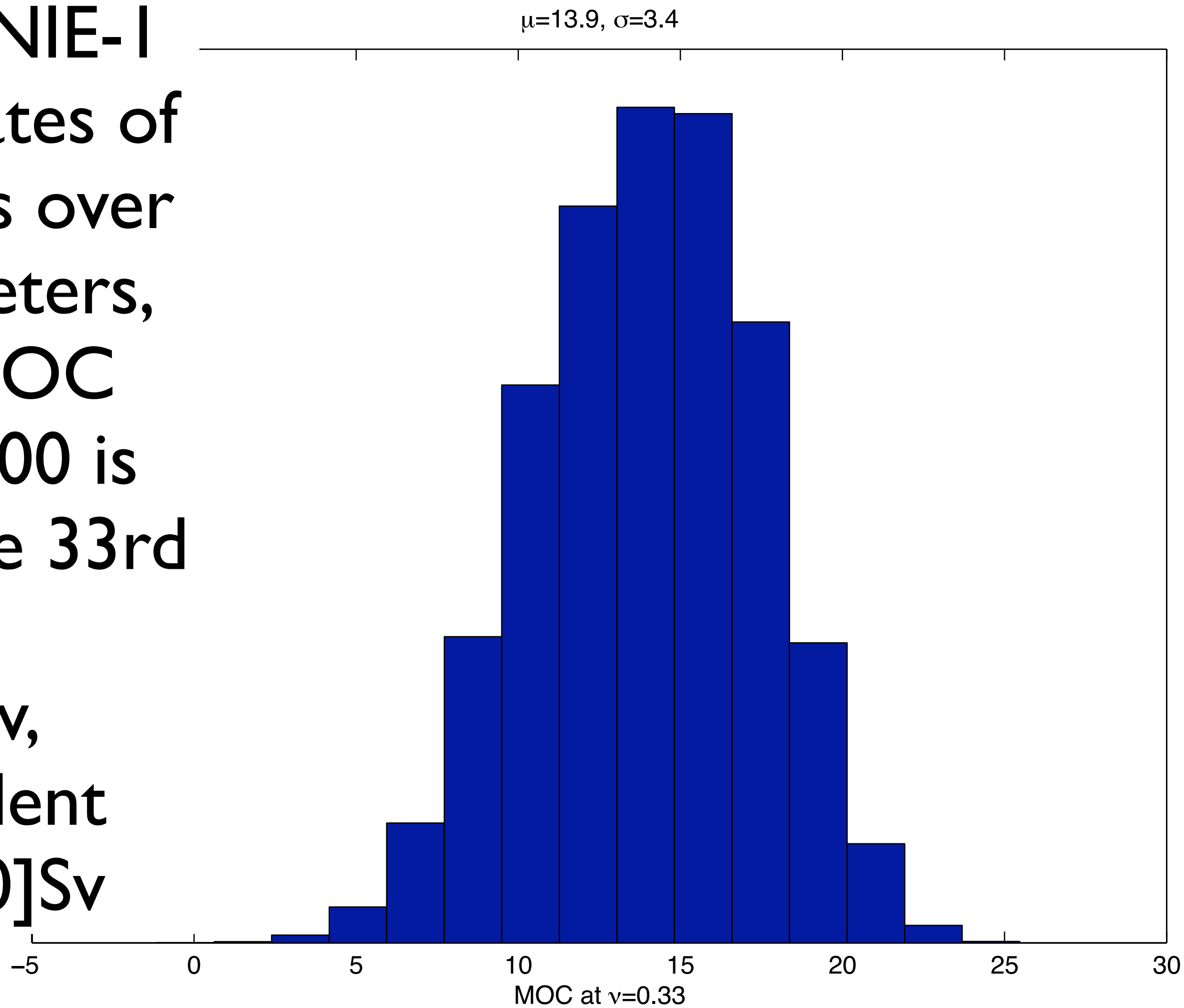
Design an ensemble of model runs, systematically varying uncertain inputs in a designed experiment (Latin Hypercube).



(forced by IPCC Scenario A1B)

Final MOC Distribution

According to GENIE-1 and expert estimates of input distributions over uncertain parameters, the change in MOC from 2000 to 2100 is very uncertain. The 33rd percentile is ~14 +/- 3.5 Sv, or a 90% confident interval of [8,20]Sv



Parameter Est. Example

- Three turbulence sub-gridscale parameterizations, one parameter each:
  - Gent-McWilliams isopycnal mixing/transport
  - Lagrangian-Averaged Navier Stokes alpha model (LANS-α)
  - eddy viscosity

Los Alamos NATIONAL LABORATORY EST. 1943

U.S. DEPARTMENT OF ENERGY Office of Science

aside on LANS-α

- Lagrangian-Averaged Navier Stokes α model:
  - involves 2 velocities:
    - smooth Eulerian-avg'd transport velocity
    - less smooth Lagrangian-avg'd velocity carried with the flow
  - preservation of Kelvin's Circulation Theorem
- onset of instability at ~2 or 3x coarser res

Los Alamos NATIONAL LABORATORY EST. 1943

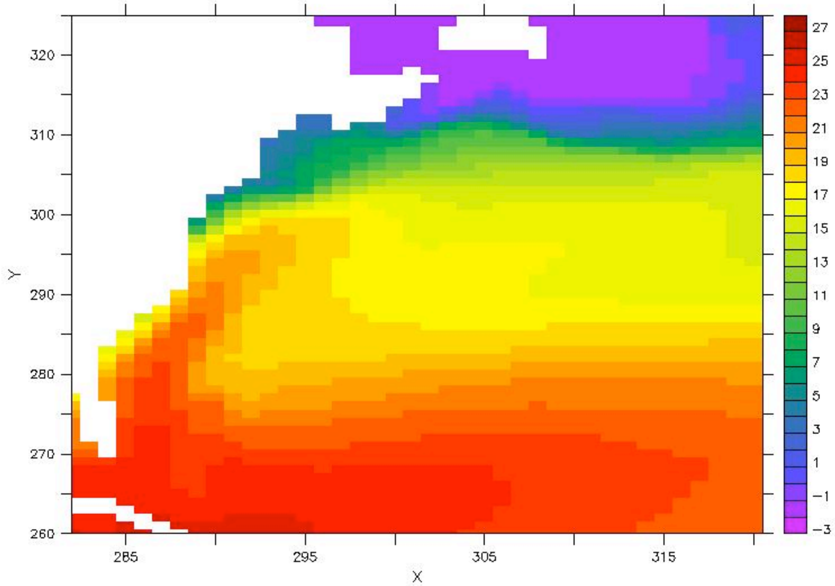
U.S. DEPARTMENT OF ENERGY Office of Science

candidate for param est:

- if eddy viscosity needed without LANS-α, then will be needed with LANS-α
  - probably with a larger value of eddy viscosity, as flow will be more energetic
- LANS-α not necessarily a replacement for isopycnal tracer mixing/transport (GM)
  - even if this parameter may be made smaller
- So, we want to pick good values for each of these three coefficients
  - and choice of one influences choice of others

a reminder: why we parameterize eddies

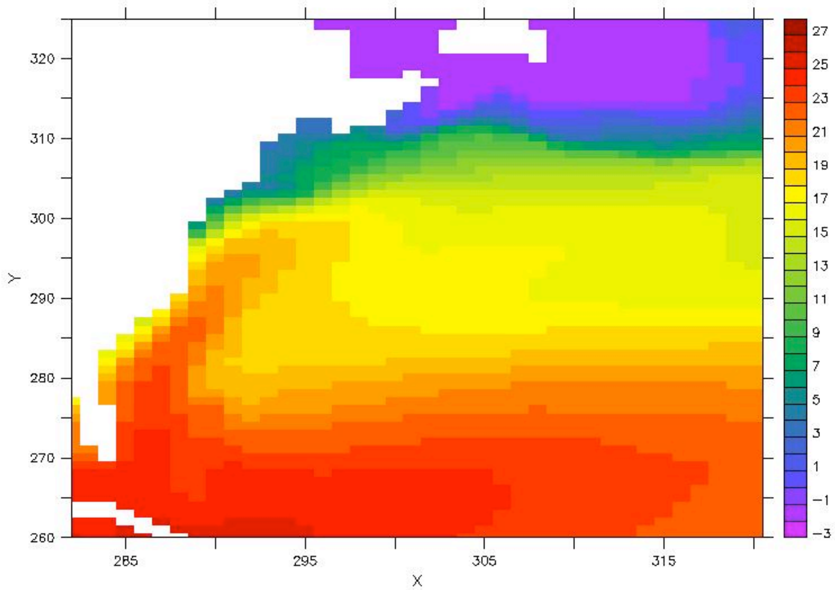
"IPCC-class" ocean model for Climate simulations



Surface temperature
1.0º x 1.0º grid

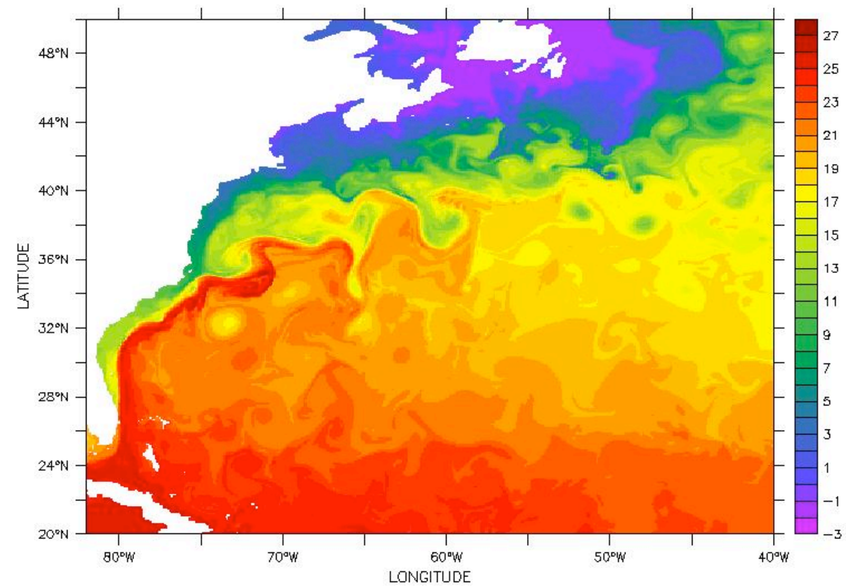
a reminder: why we parameterize eddies

"IPCC-class" ocean model for Climate simulations



Surface temperature
1.0º x 1.0º grid

Strongly eddying simulation



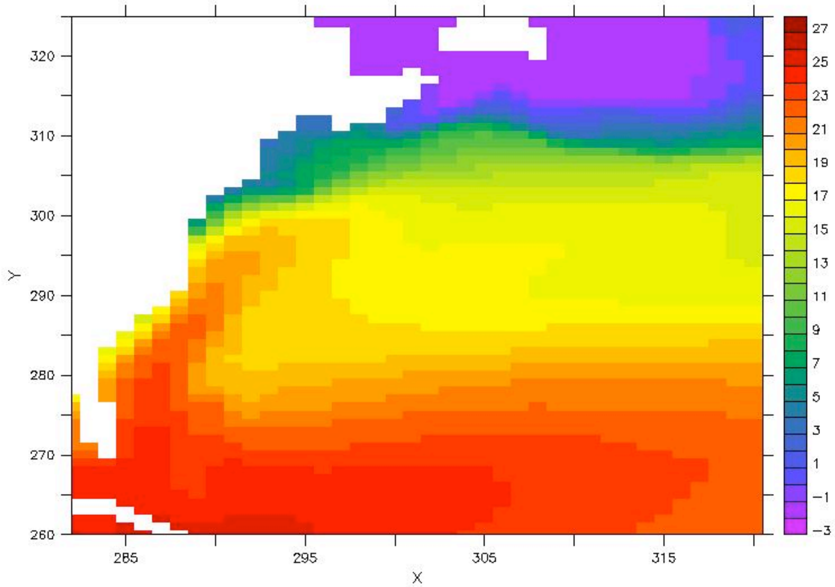
Surface temperature
0.1º x 0.1º grid

Los Alamos NATIONAL LABORATORY EST. 1943

U.S. DEPARTMENT OF ENERGY | Office of Science

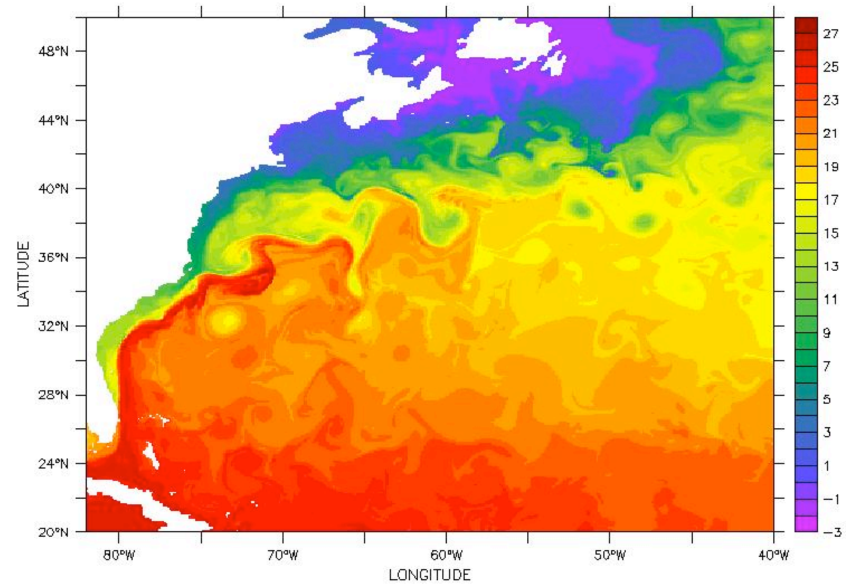
a reminder: why we parameterize eddies

"IPCC-class" ocean model for Climate simulations



Surface temperature
1.0º x 1.0º grid

Strongly eddying simulation



Surface temperature
0.1º x 0.1º grid

Factor of ~10 for each doubling of resolution

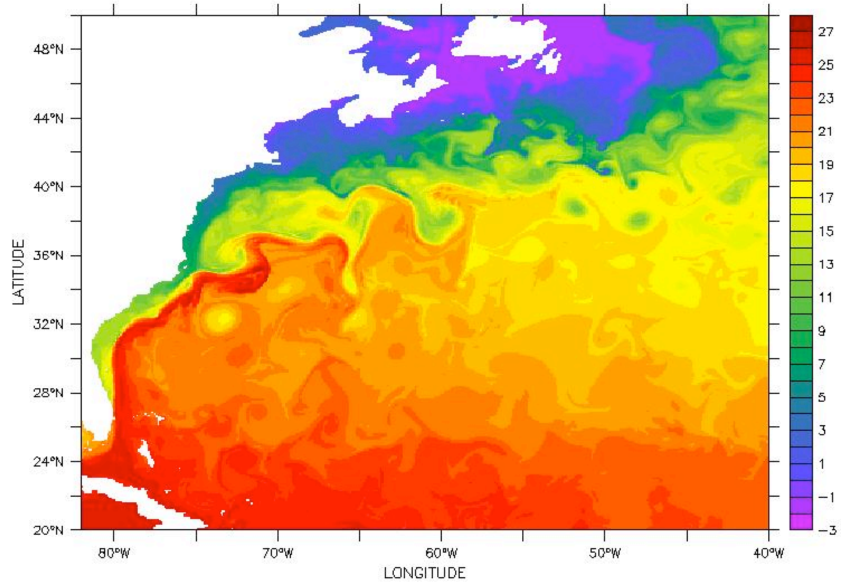
Los Alamos NATIONAL LABORATORY EST.1943

U.S. DEPARTMENT OF ENERGY Office of Science

isopycnal tracer mixing/transport schemes can parameterize much of the effect of eddies...

but some of the action of eddies has resisted parameterization

Strongly eddying simulation



Los Alamos NATIONAL LABORATORY EST. 1943

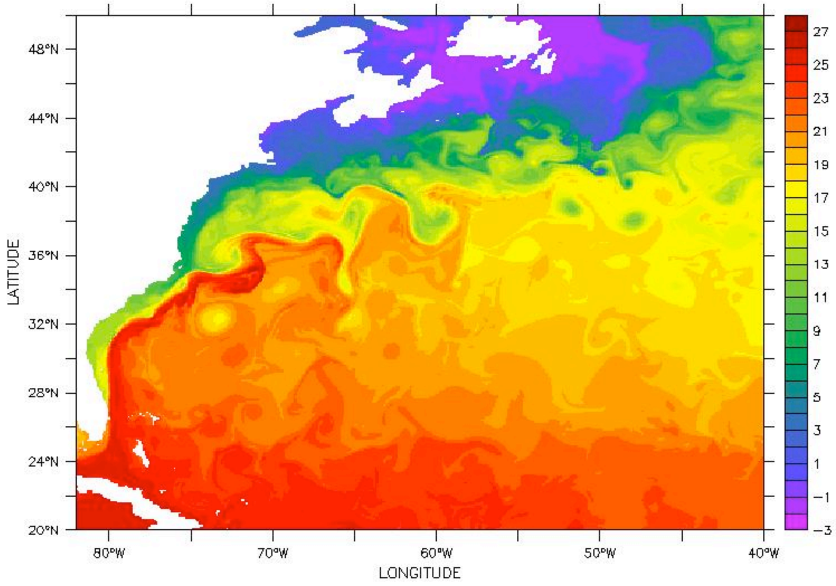
U.S. DEPARTMENT OF ENERGY | Office of Science

isopycnal tracer mixing schemes can parameterize much of the effect of eddies...

but some of the action of eddies has resisted parameterization

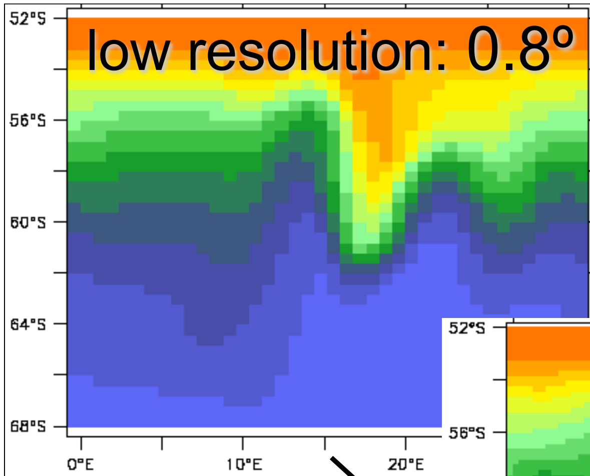
LANS-α turbulence model offers possibility of strongly eddying solution, but at ½ the resolution

Strongly eddying simulation

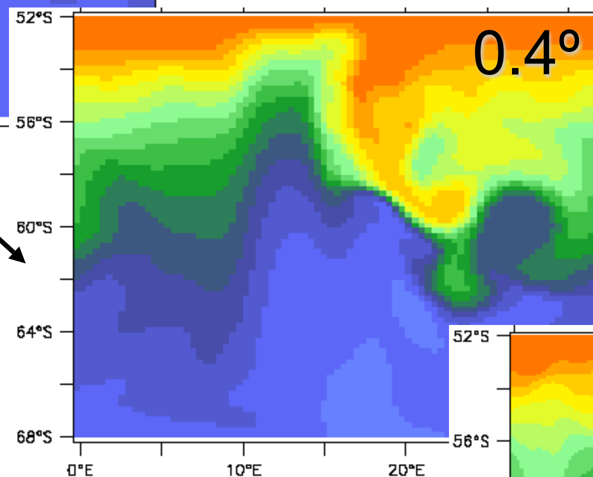


Los Alamos NATIONAL LABORATORY EST. 1943

U.S. DEPARTMENT OF ENERGY Office of Science

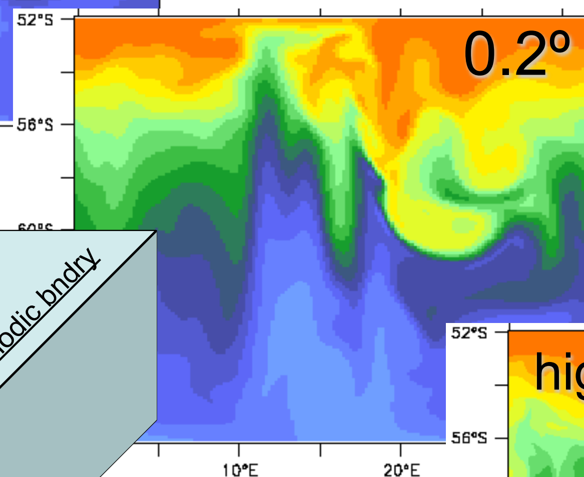


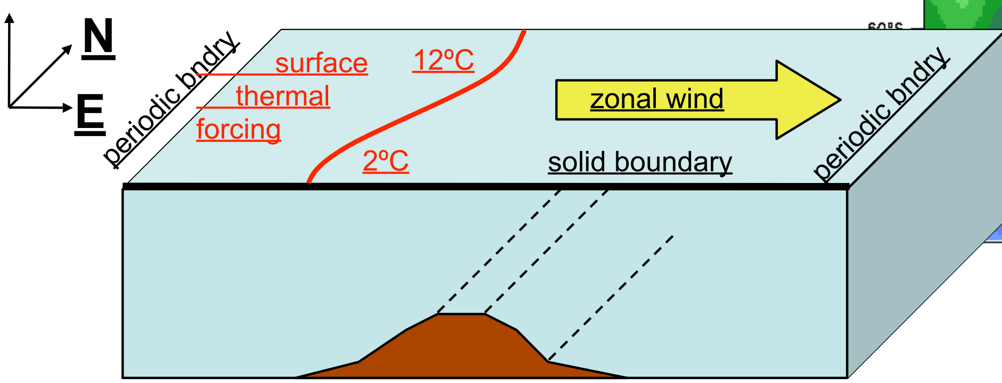
Here, we see increase in mesoscale eddy variability with resolution

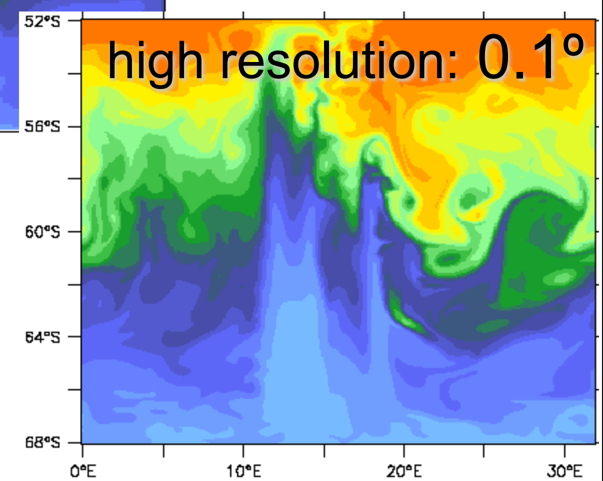


control simulations
(without use of LANS-α)

cost of doubling
horizontal grid
is factor of 10



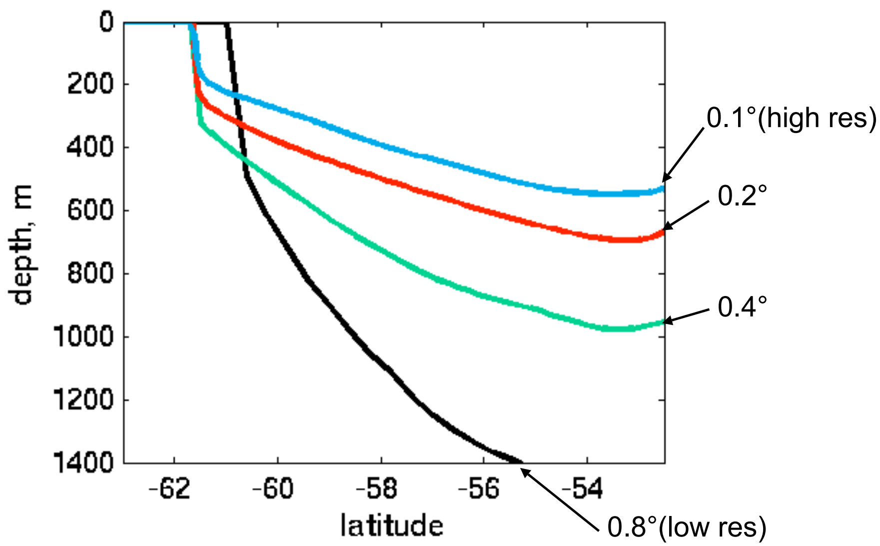




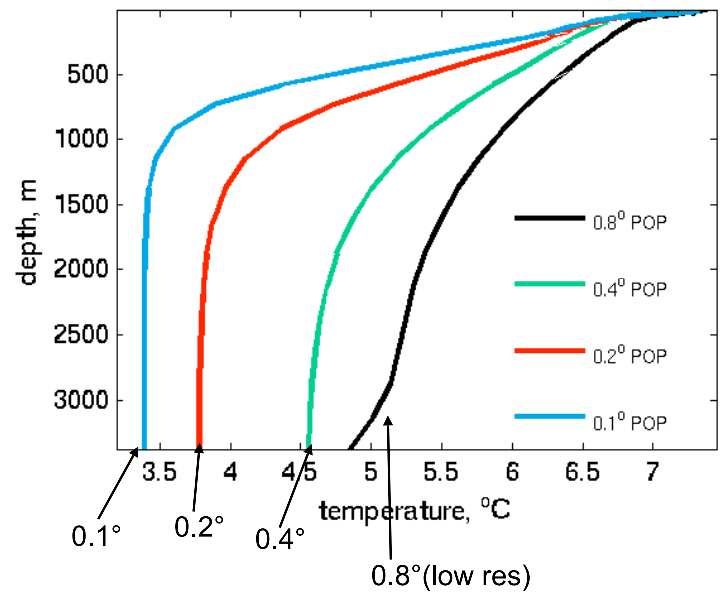
Los Alamos
NATIONAL LABORATORY
EST. 1943

Test Problem Results, POP and POP-α

6C isotherm



Vertical temperature profile

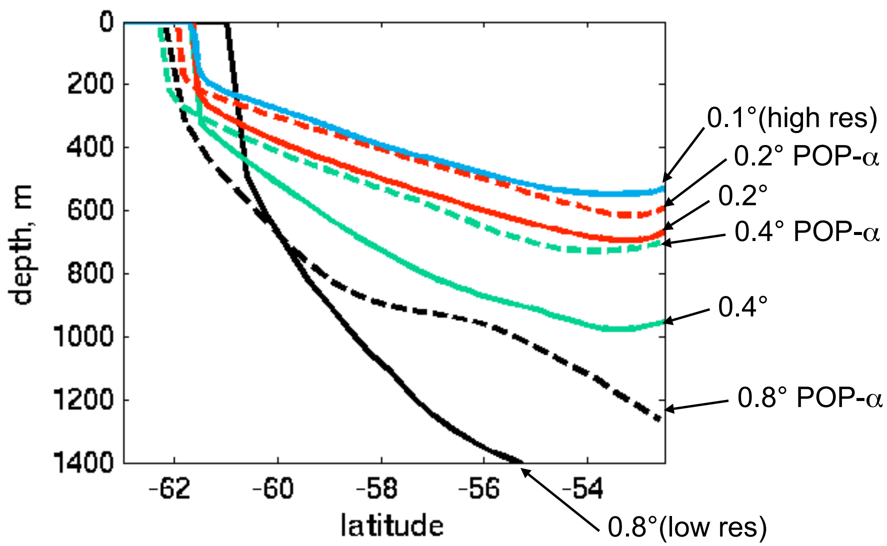


Los Alamos NATIONAL LABORATORY EST. 1943

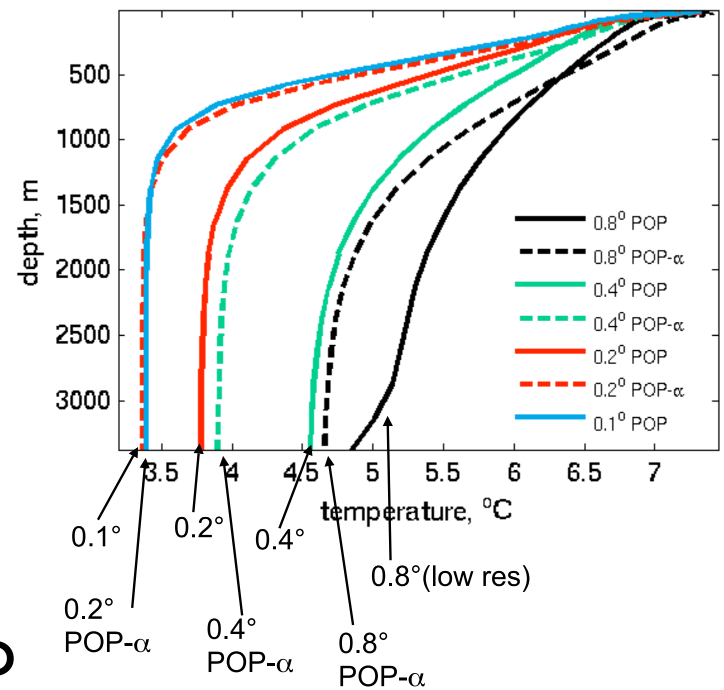
U.S. DEPARTMENT OF ENERGY | Office of Science

Test Problem Results, POP and POP-α

6C isotherm



Vertical temperature profile



Use of LANS-α comparable to a doubling of resolution, in these measures

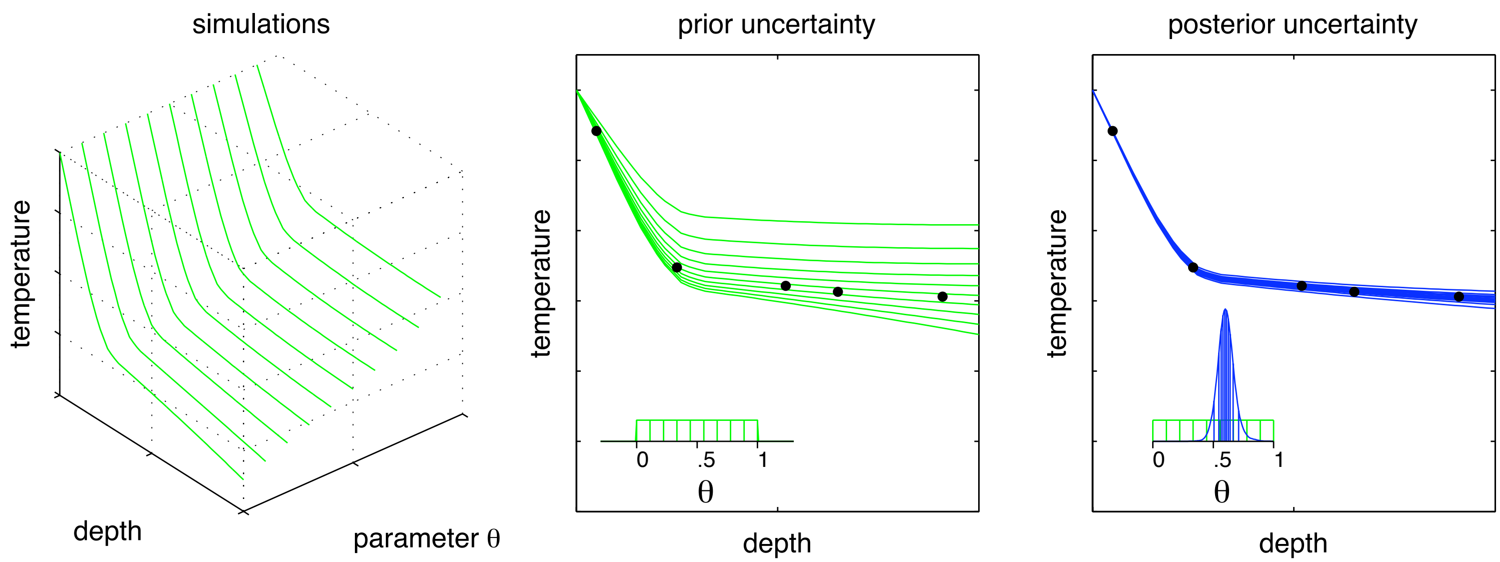
These results (from JCP 2008) produced with "standard" values of lateral viscosity, and without GM

Los Alamos NATIONAL LABORATORY EST. 1943

U.S. DEPARTMENT OF ENERGY Office of Science

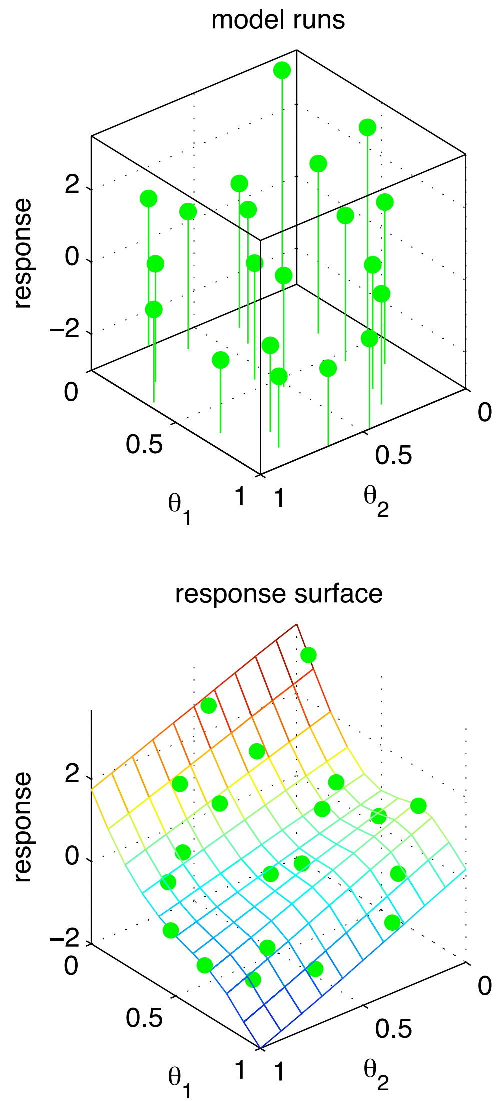
Parameter Estimation: in Concept

Use statistical approach to find input settings to match the target profile



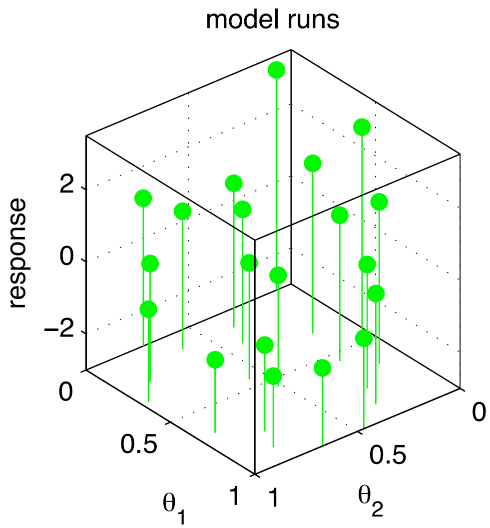
- Finds input parameter settings that best match the target temperature profile
- Requires that initial ranges for the parameters be specified
- Here we show a 1-d parameter space – actual application uses a 3-d parameter space.

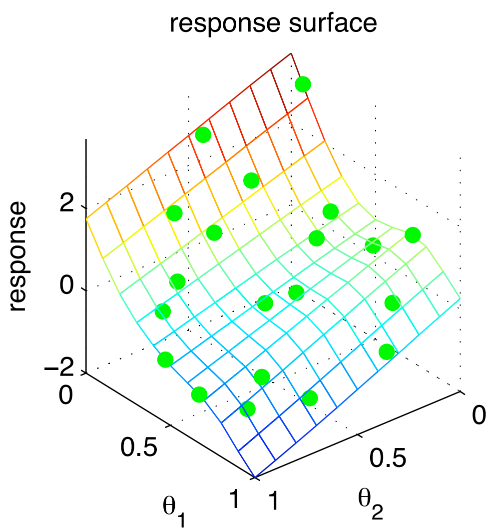
Construct a response surface of the simulation output to predict at untried settings



- Actual application requires a basis representation to predict temperature profiles
- Can use holdouts to assess accuracy of response surface
- Can carry out sensitivity analysis using response surface

Construct a response surface of the simulation output to predict at untried settings

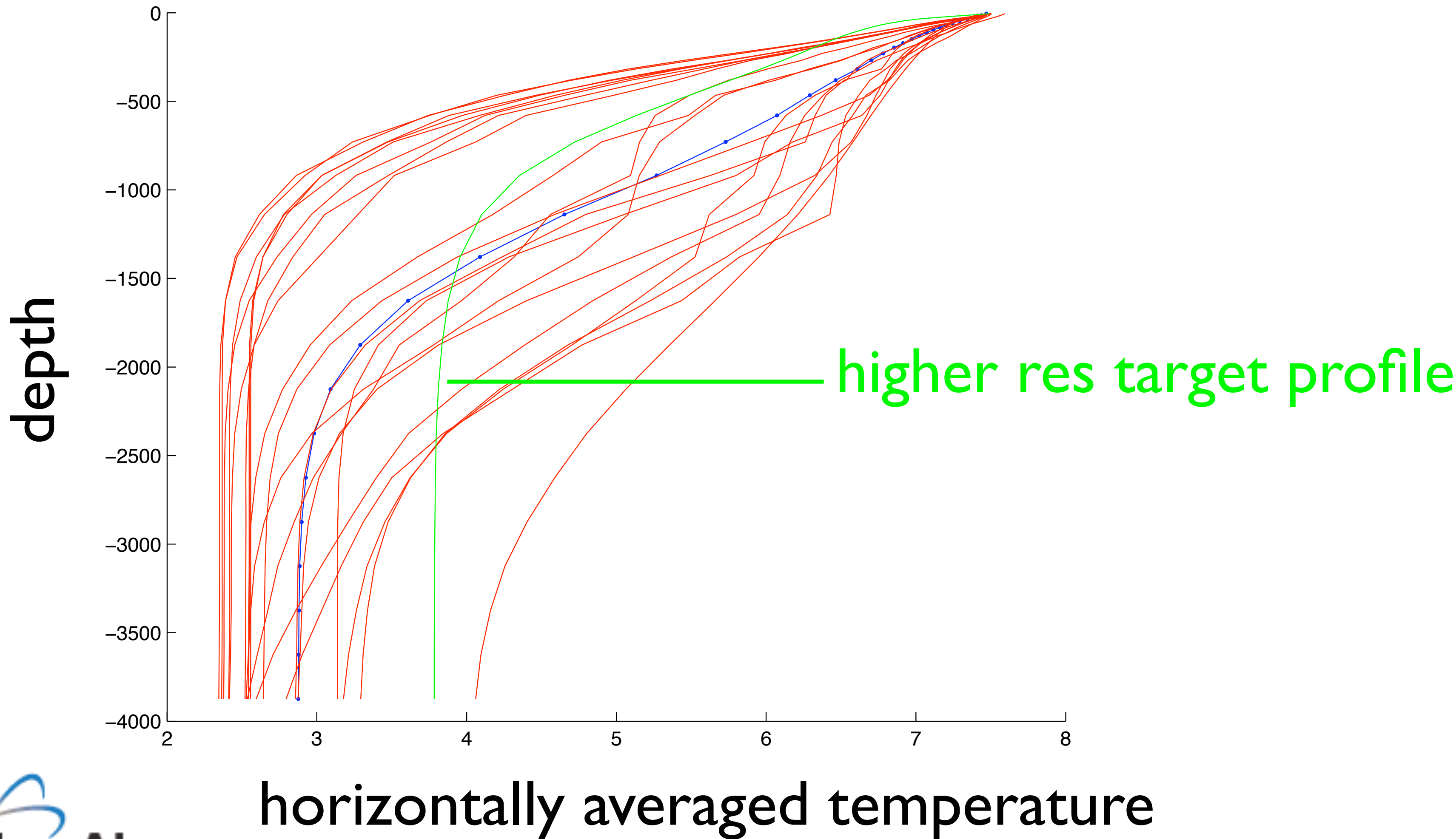




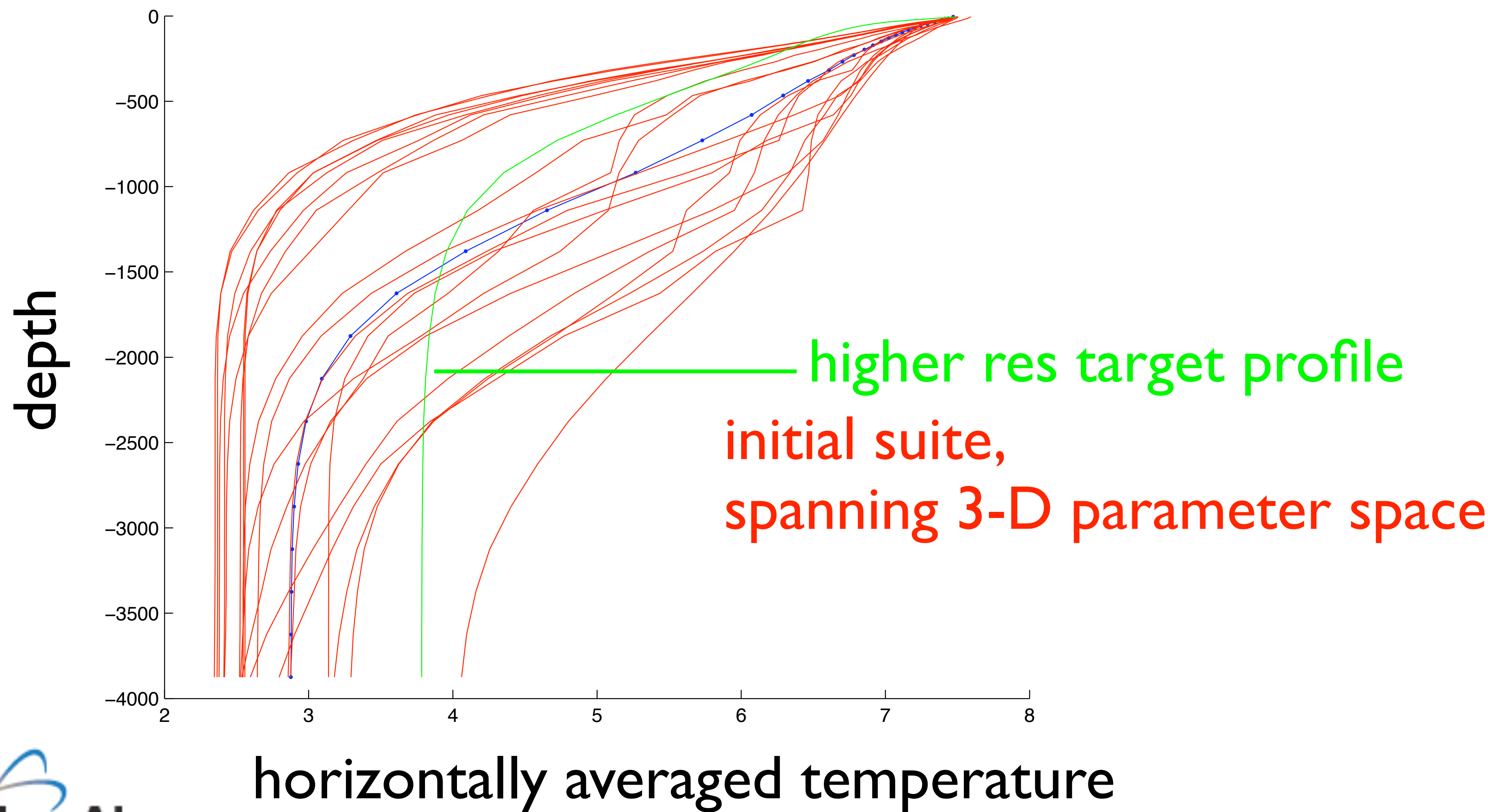
- Actual application requires a basis representation to predict temperature profiles
- Can use holdouts to assess accuracy of response surface
- Can carry out sensitivity analysis using response surface

Prediction at untried settings based on Gaussian process emulators (there's a literature on this)

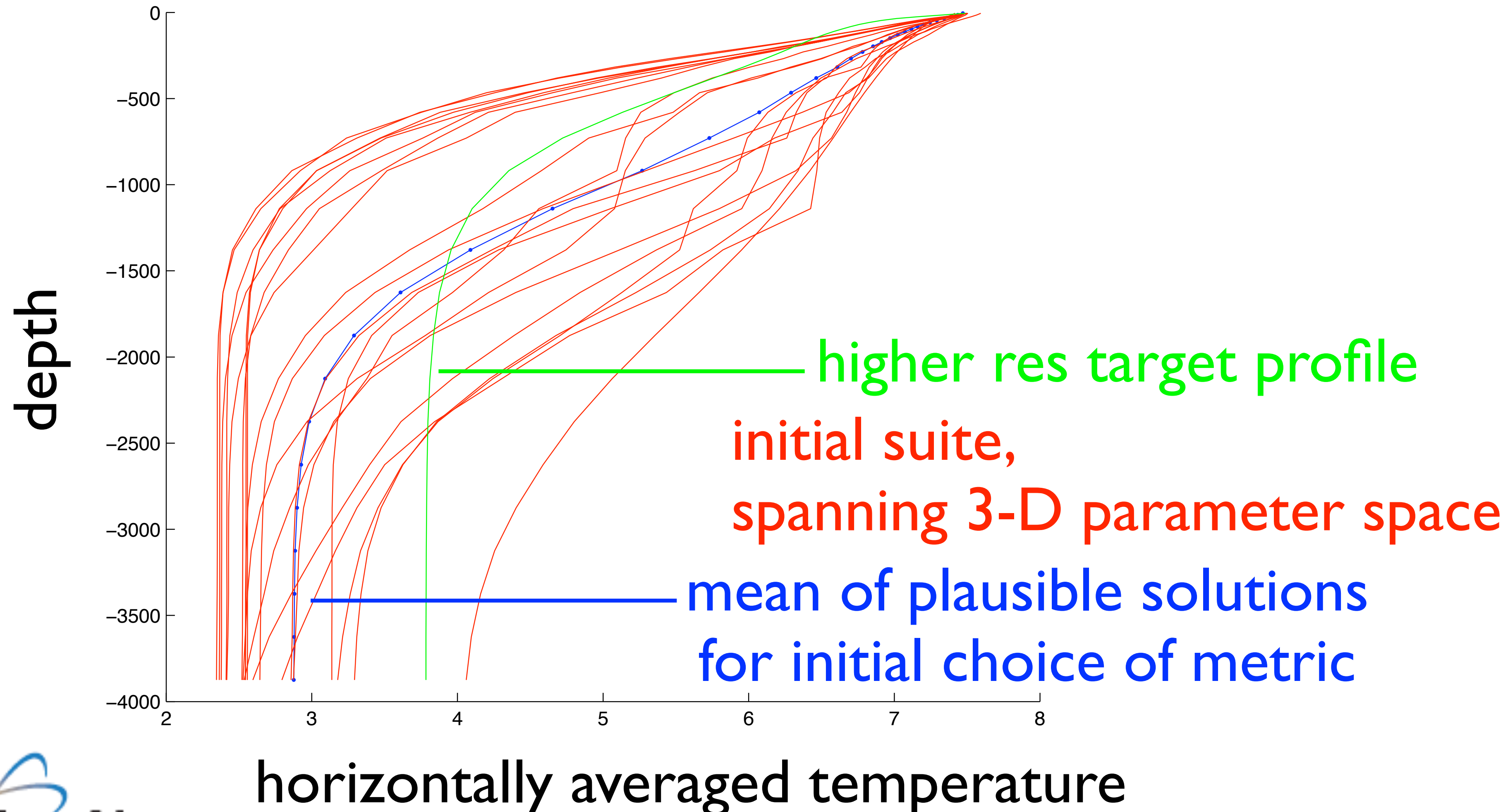
Parameter Estimation: in Practice



Parameter Estimation: in Practice



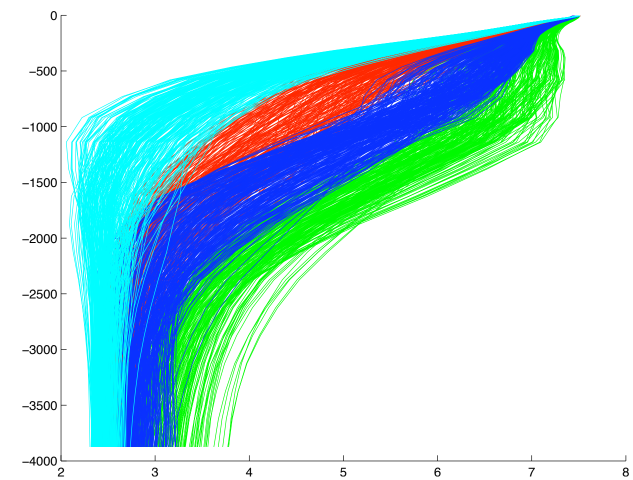
Parameter Estimation: in Practice



Los Alamos NATIONAL LABORATORY EST. 1943

U.S. DEPARTMENT OF ENERGY | Office of Science

hard to fill 3-D param space with ocean model runs -- but it's easy now with the emulator

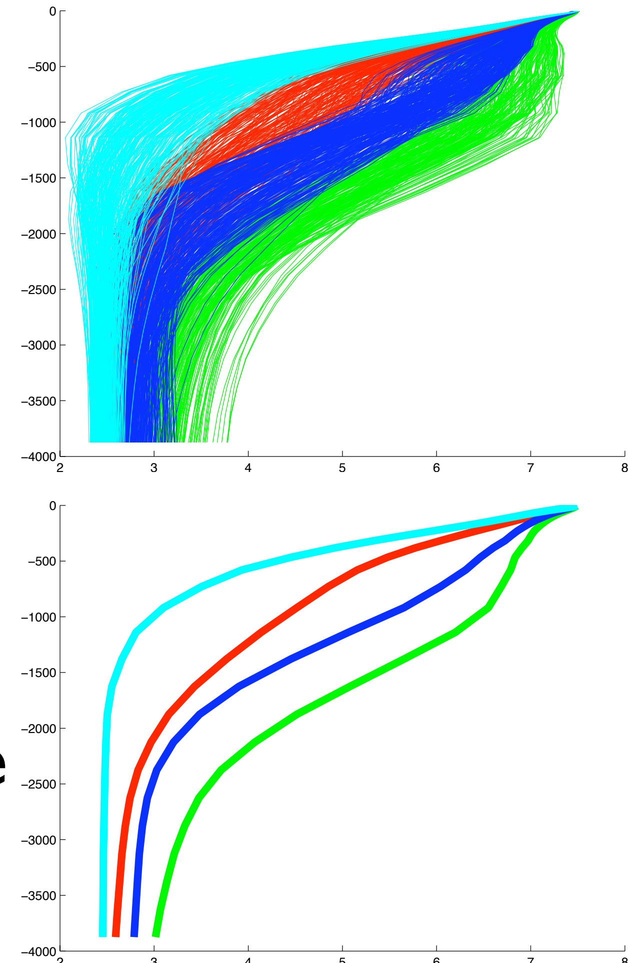


many profiles, produced by the emulator, grouped based on a clustering algorithm

Los Alamos NATIONAL LABORATORY EST. 1943

U.S. DEPARTMENT OF ENERGY Office of Science

hard to fill 3-D param space with ocean model runs --
but it's easy now with the emulator



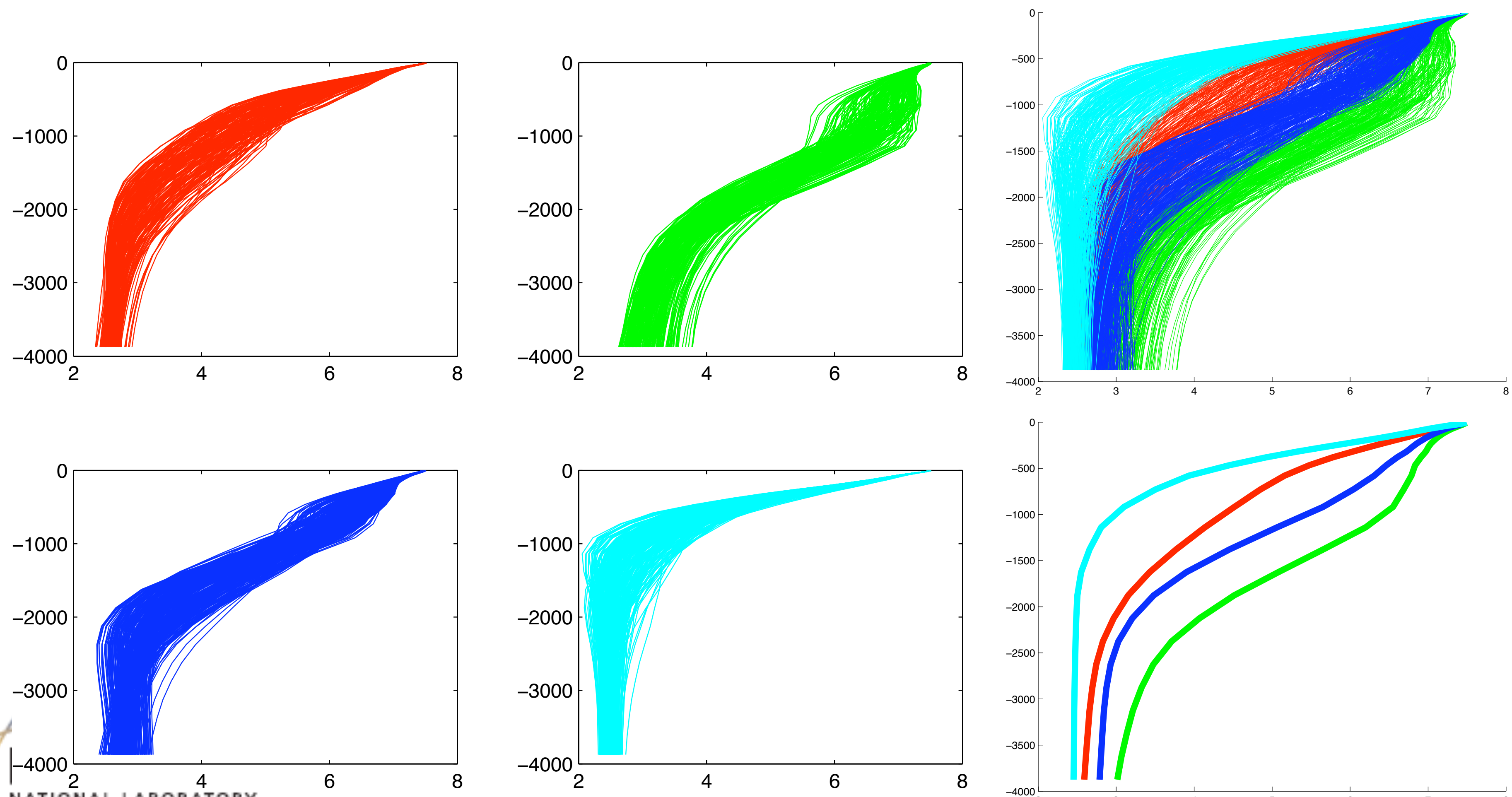
many profiles, produced by the emulator, grouped based on a clustering algorithm

then each associated with one representative profile

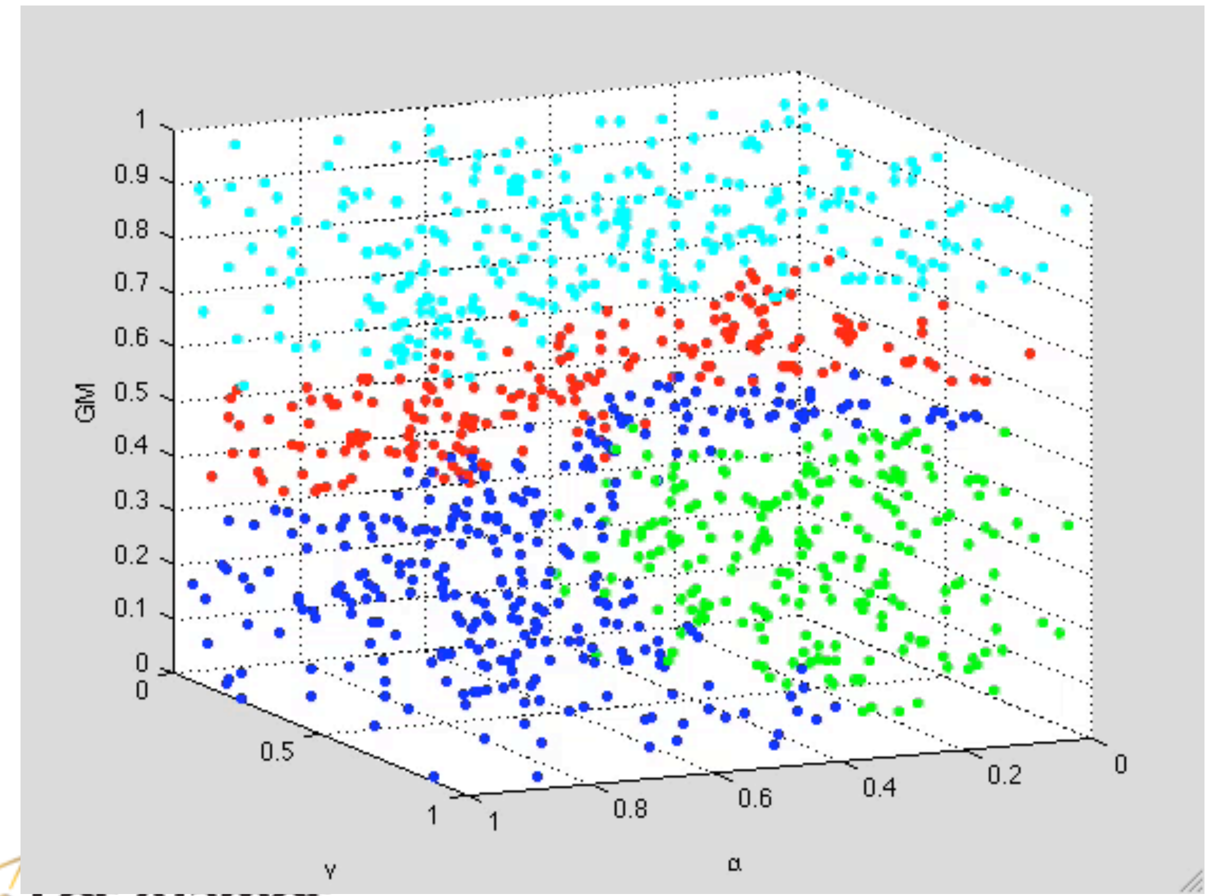
Los Alamos NATIONAL LABORATORY EST. 1943

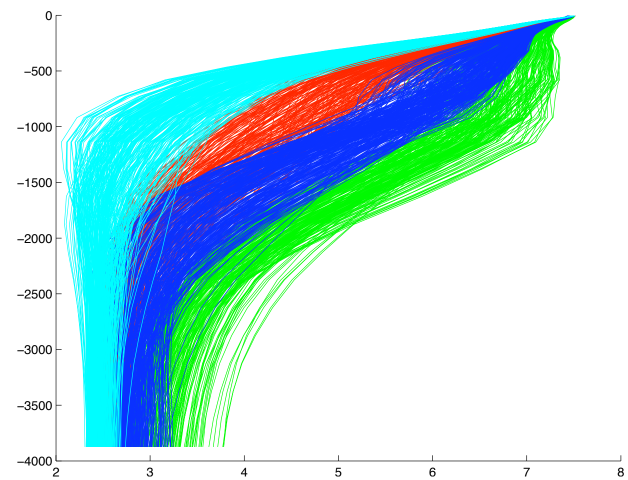
ENERGY | Science

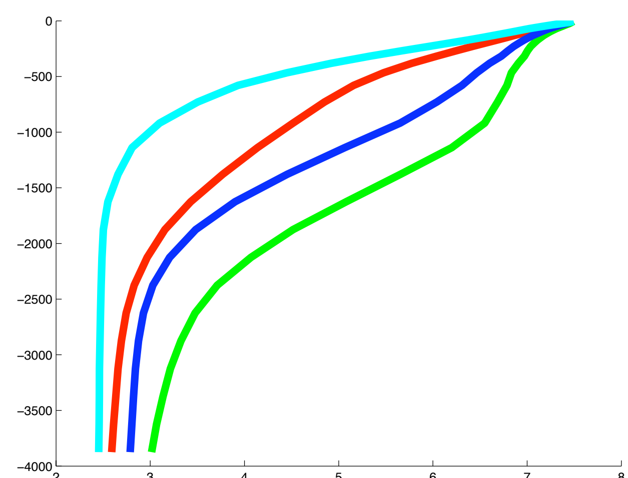
hard to fill 3-D param space with ocean model runs -- but it's easy now with the emulator



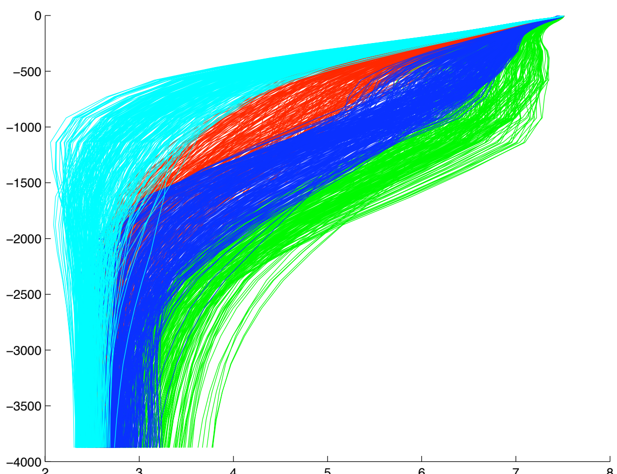
hard to fill 3-D param space with ocean model runs -- but it's easy now with the emulator

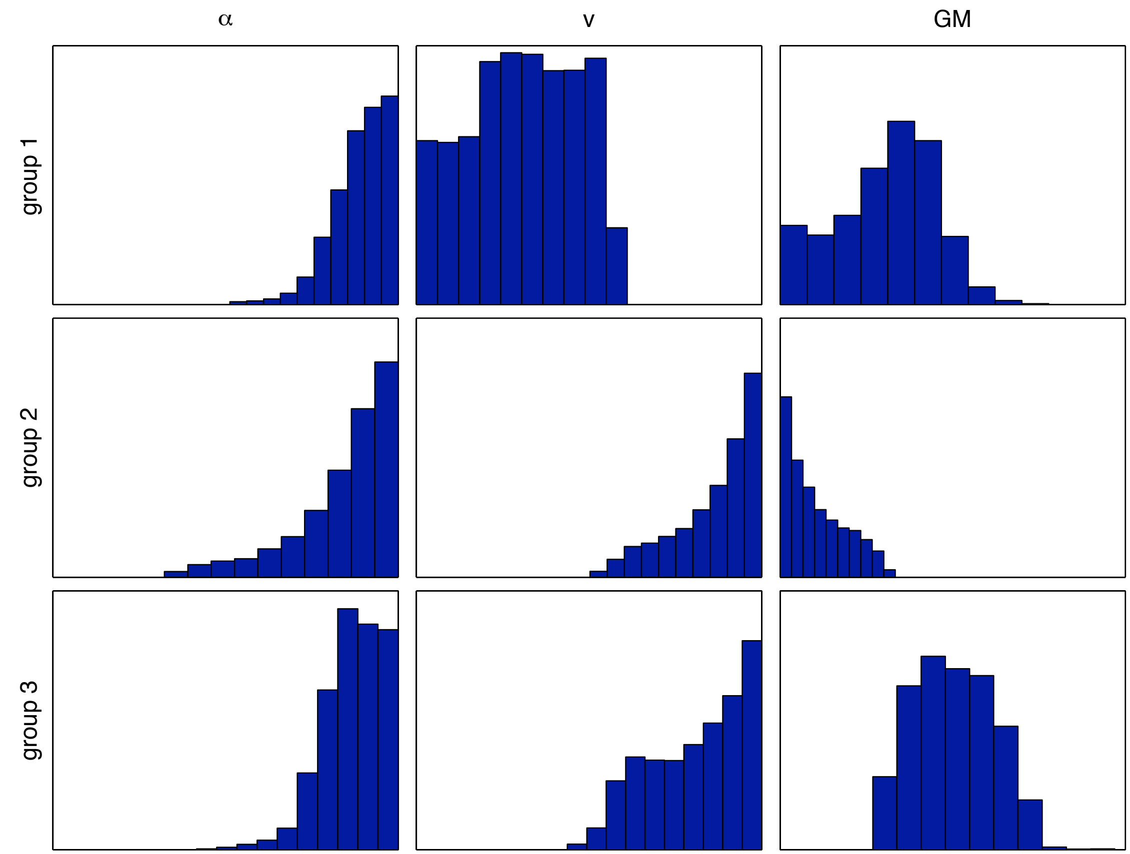






3 ways to produce more plausible solution



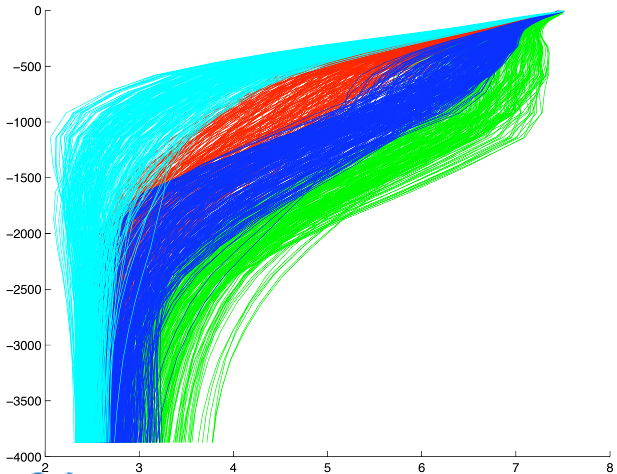


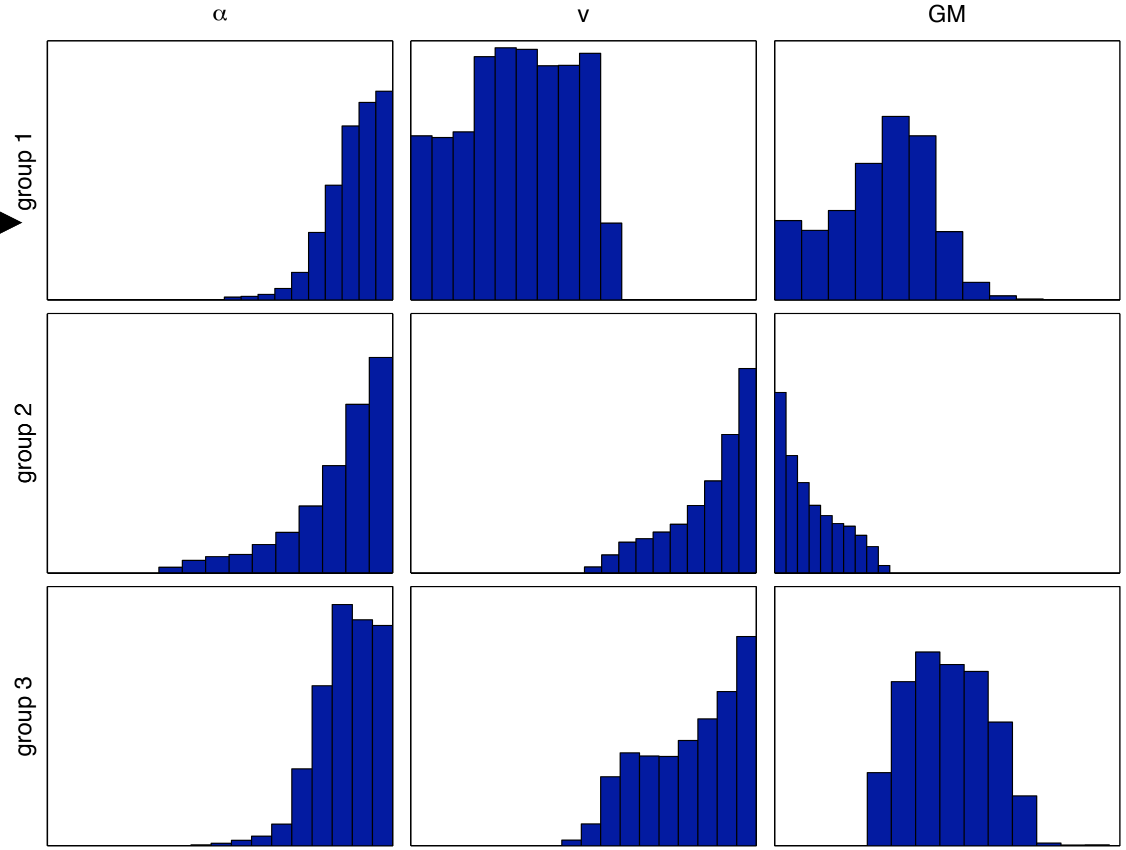
Los Alamos
NATIONAL LABORATORY
EST. 1943

U.S. DEPARTMENT OF ENERGY
Office of Science

3 ways to produce more plausible solution

this row indicates high α, mid-to-low viscosity, mid-to-low GM





My lessons learned

- For parameter estimation,
  - a sense of which output metrics will respond to the input parameters in question is essential
- and for quantification of uncertainty in climate model result:
  - should understand in advance which input parameters will cause strong variation in the output result

Smooth Eulerian-averaged **u** and rough Lagrangian-averaged **v**

Filtering of rough velocity v produces smooth velocity u:

$$\mathbf{u} = Filter(\mathbf{v})$$

$$Filter = (1 - \alpha^2 \nabla^2)^{-1}$$

Then apply Kelvin's Circulation Theorem around a closed loop within the fluid:

$$\frac{d}{dt} \oint_{\gamma(\mathbf{u})} \mathbf{v} \cdot dx = \oint_{\gamma(\mathbf{u})} \nu \nabla^2 \mathbf{v} + \mathbf{F}$$

after manipulation, modified eqns of motion:

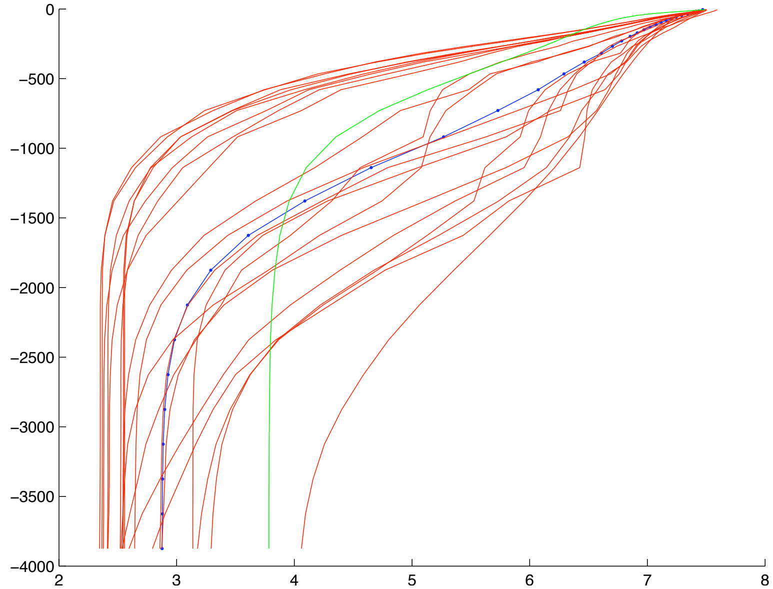
$$\frac{\partial \mathbf{v}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{v} + \nabla \mathbf{u}^{\mathrm{T}} \cdot \mathbf{v} + \nabla \pi = \nu \nabla^2 \mathbf{v} + F$$

extra nonlinear term

modified pressure

good choice of metric is essential:

for example, instead of minimizing distance from target, level-by-level:

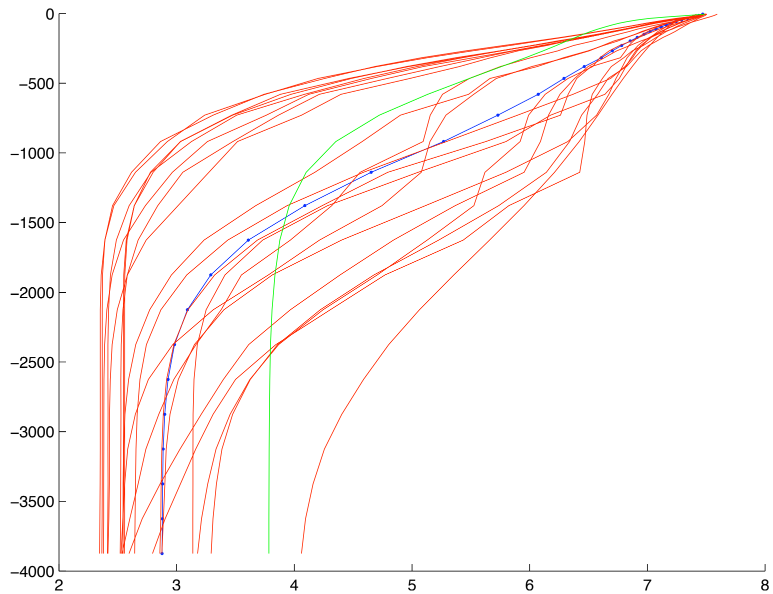


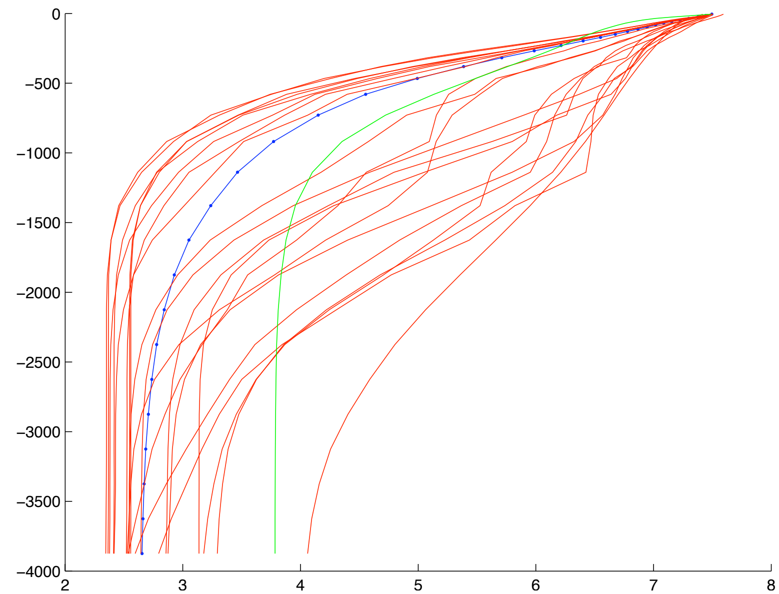
Los Alamos NATIONAL LABORATORY EST.1943

U.S. DEPARTMENT OF ENERGY | Office of Science

good choice of metric is essential:

for example, instead of minimizing distance from target, level-by-level:





can allow for depth-independent bias
(so, select based on shape)

Los Alamos NATIONAL LABORATORY EST. 1943

U.S. DEPARTMENT OF ENERGY | Office of Science