More accurate models

Incorporate comprehensive, creative error estimation

Faster, more innovative parameterization development

Opportunities and ***pitfalls*** in automated calibration

Overfitting and compensating error

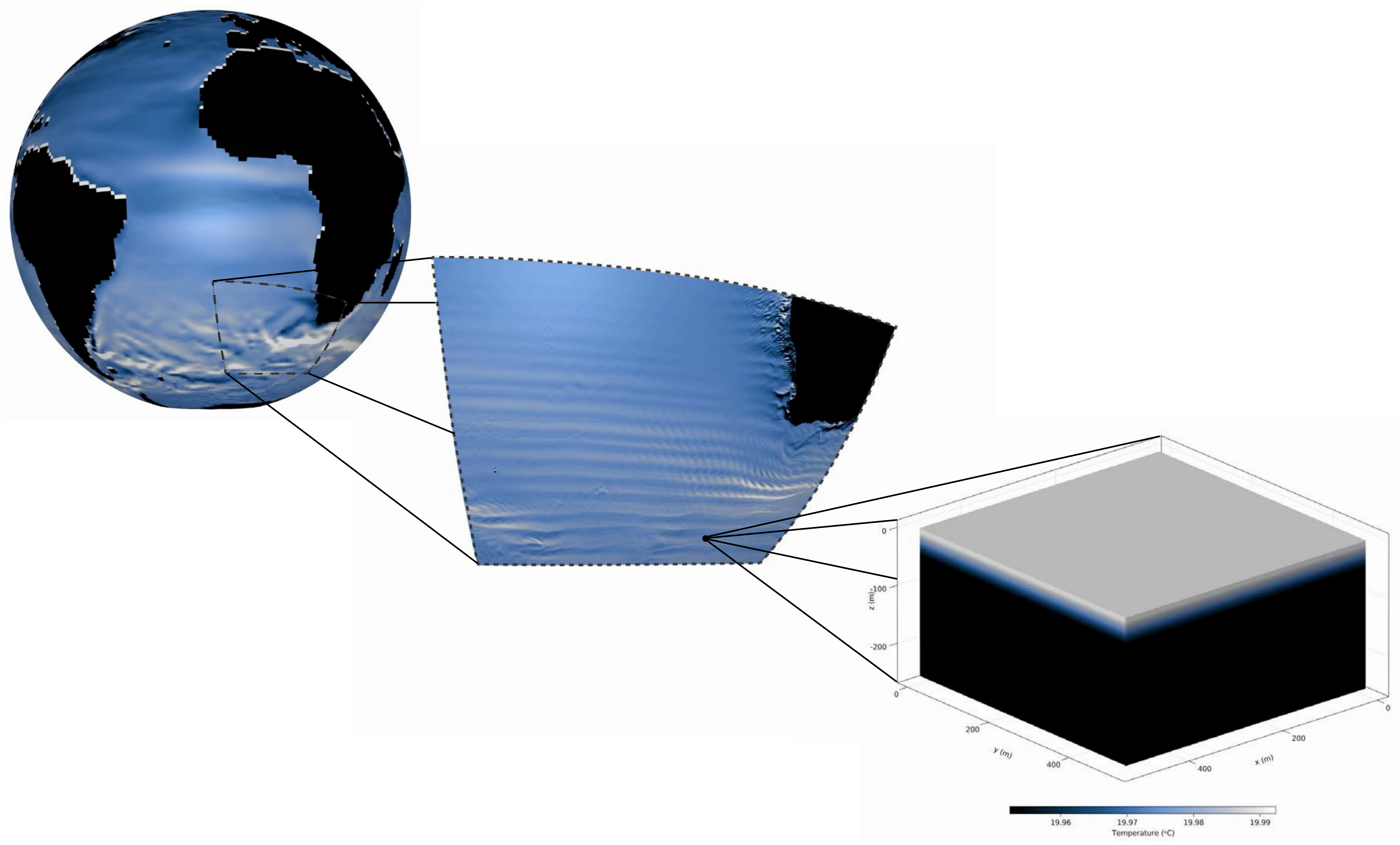
Ocean Model Development Workshop

Gregory Wagner (MIT) and the Clima-Ocean team

Automated calibration of parameterizations

→ **Finding parameters by minimizing error with a computational method**

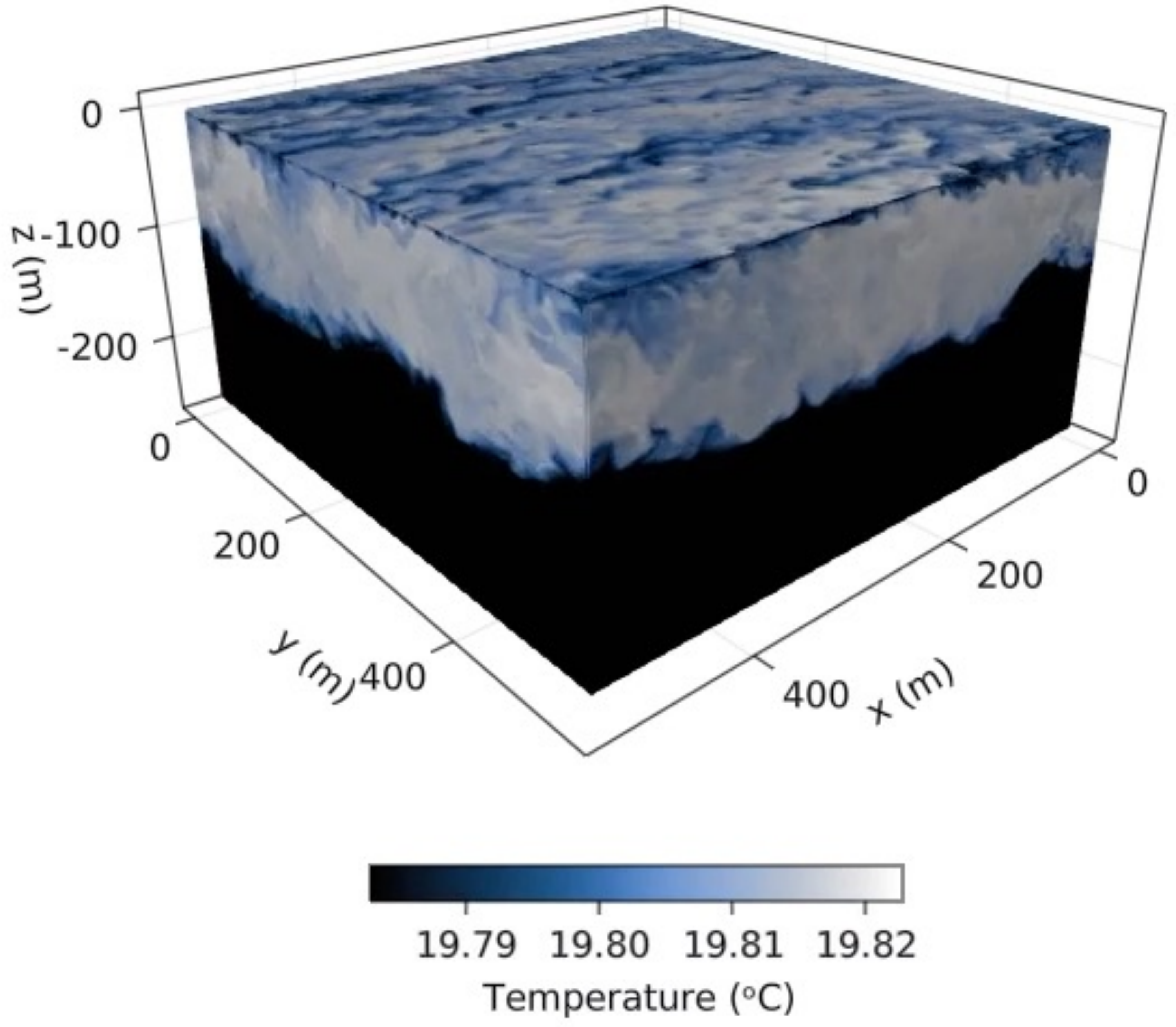
- Methods: stochastic gradient descent, Ensemble Kalman Inversion
- Repeatable and reproducible
- “Error” can be formulated flexibly
- Can include uncertainty quantification as an additional step

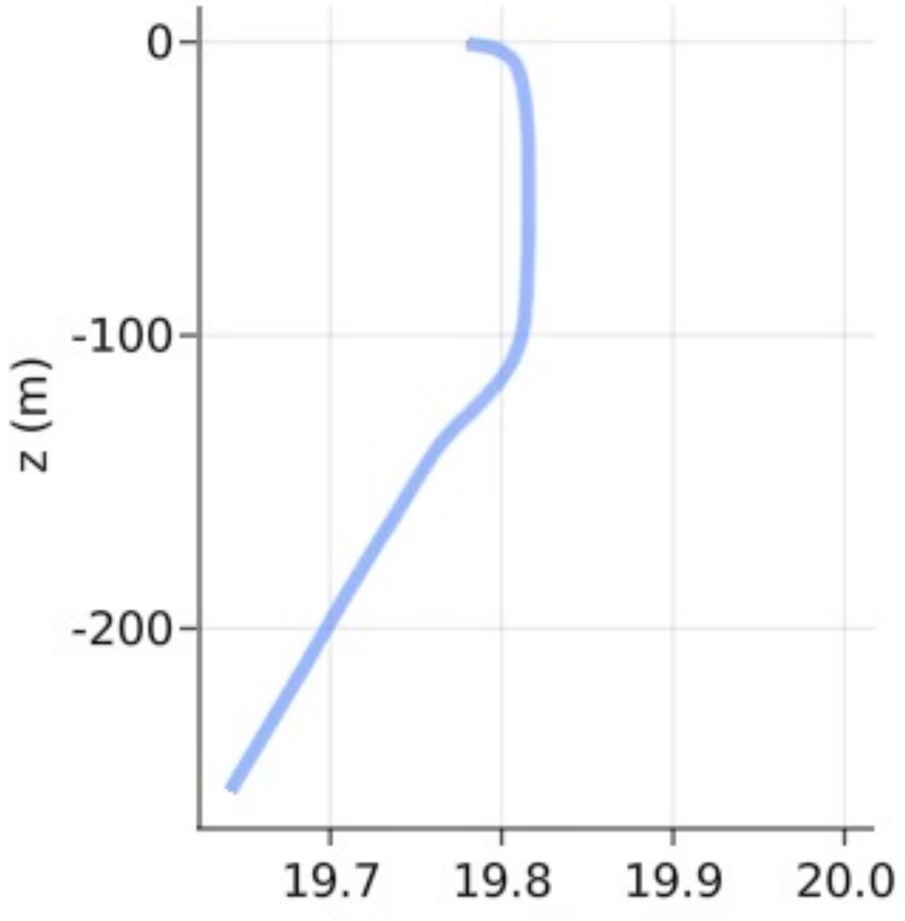


Automated calibration of mixing parameterizations

- High-fidelity LES = “truth”
- Parameterization embedded in a “single column model”

$$\partial_t \overline{T} = -\partial_z \overline{w'T'}(\mathbb{C})$$





Horizontally-averaged temperature (^{o}C)

- Error = difference between horizontally-averaged LES and model T, U, V

$$Error(\mathbb{C}) \sim \int_{-H}^{0} [\overline{T}_{\mathrm{LES}} - \overline{T}_{\mathrm{model}}]^2 dz$$

free parameters

horizontally-averaged

CATKE: a one-equation parameterization

Based on Convective Adjustment and Turbulent Kinetic Energy

Single-column temperature equation

$$\partial_t \overline{T} = \partial_z \left(\kappa_c \, \partial_z \overline{T} \right)$$

Single-column TKE equation

$$\partial_t e = \partial_z (\kappa_e \, \partial_z e) + \kappa_u | \, \partial_z \overline{\boldsymbol{u}} |^2 - \kappa_c \, \partial_z \overline{b} - \frac{e^{3/2}}{\ell_D}$$

Eddy diffusivity

$$\kappa_c = \ell_c \sqrt{e}$$

Mixing length

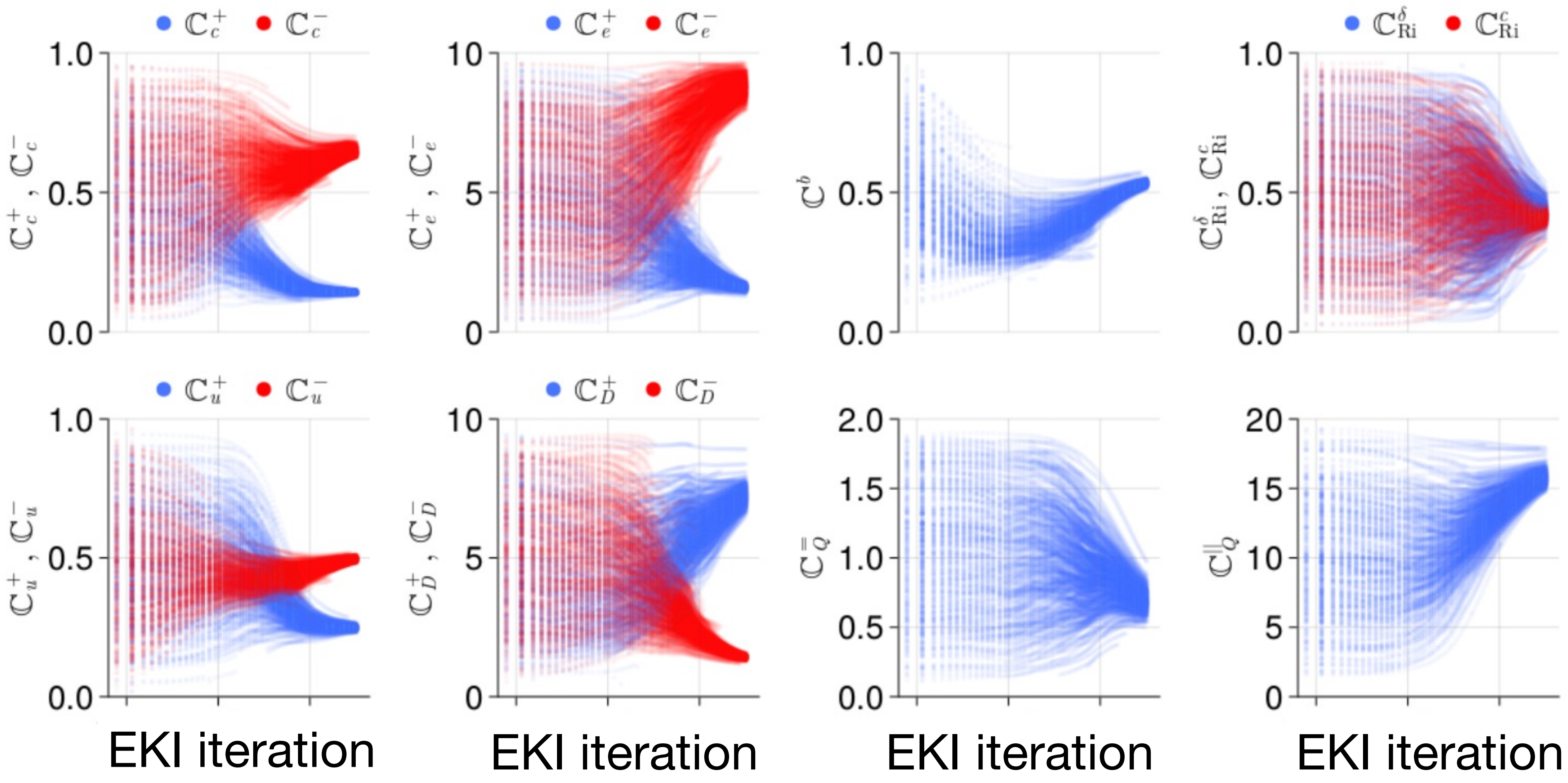
$$\ell_c = \ell_c^{\text{conv}} + \sigma_c(Ri) min \left(d, \mathbb{C}^b \frac{\sqrt{e}}{N} \right)$$

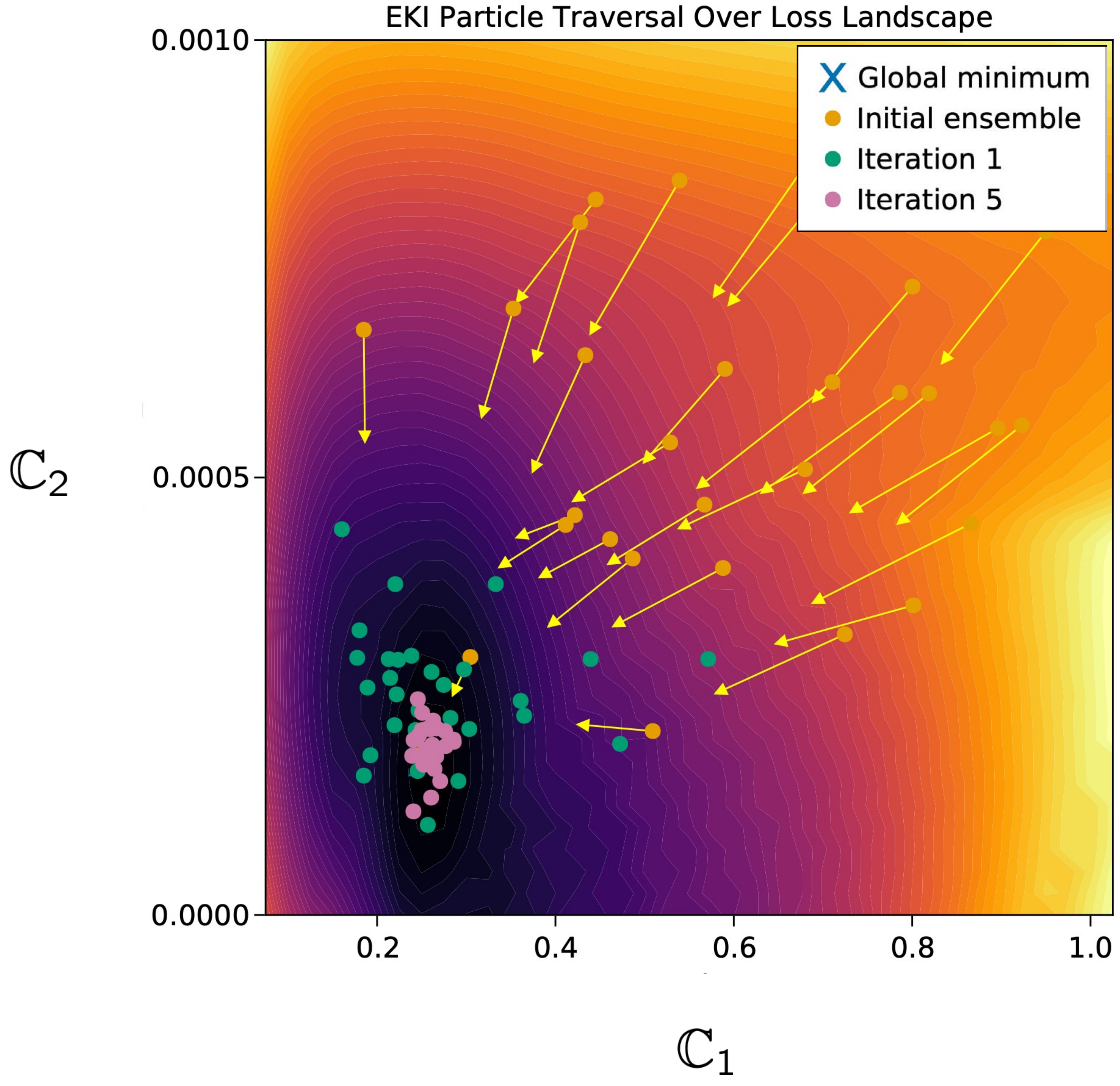
Convective mixing length

$$\ell_c^{\text{conv}} \sim \frac{e^{3/2}}{J^b} \, if N^2 > 0$$

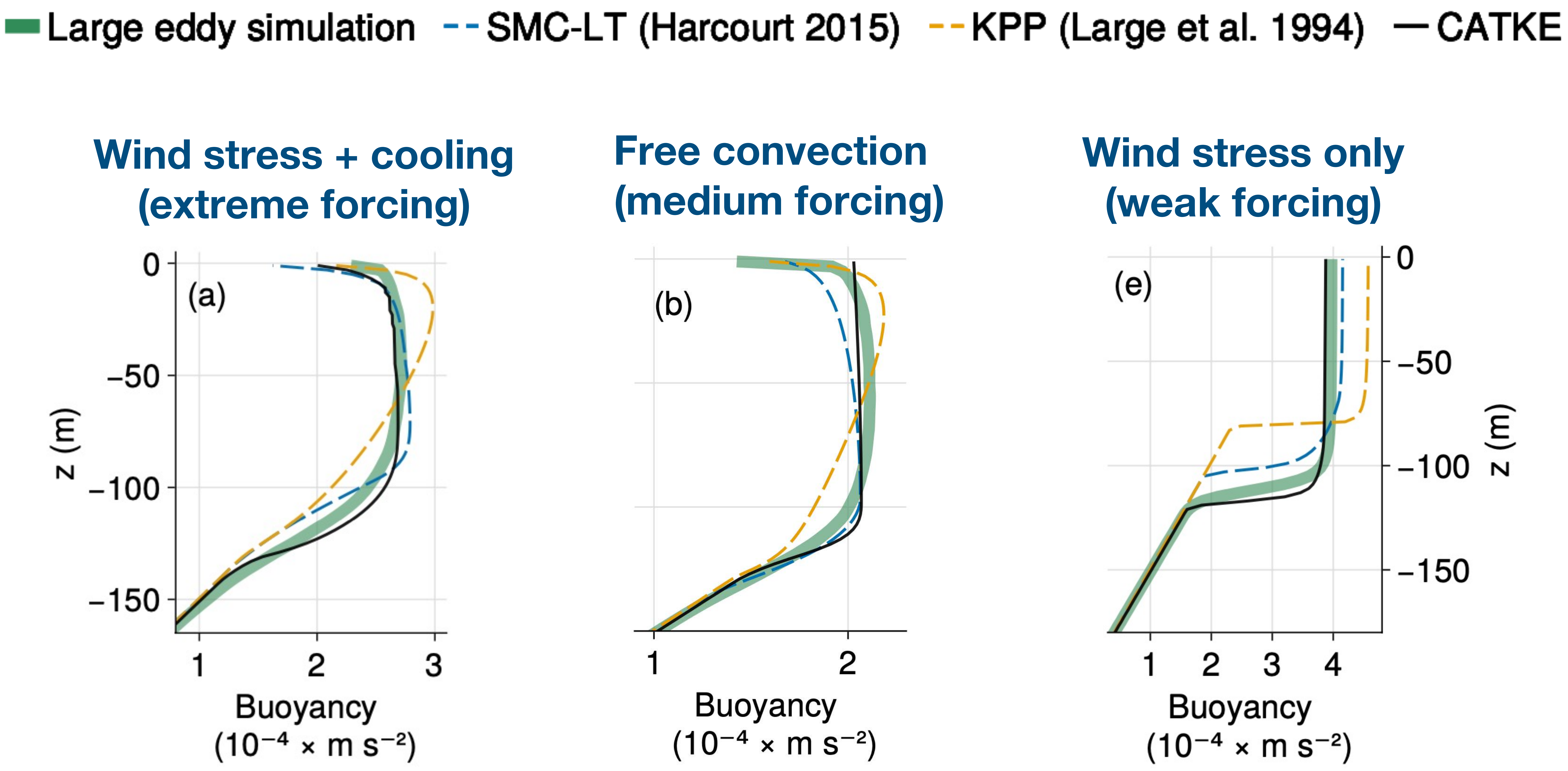
Calibration with Ensemble Kalman Inversion

- Iteratively improve an ensemble of models
- Data: 21 LES forced by constant surface fluxes (including both temperature and passive tracer)
- *Note:* incorporate error from simulations with 8m, 4m, and 2m vertical resolution

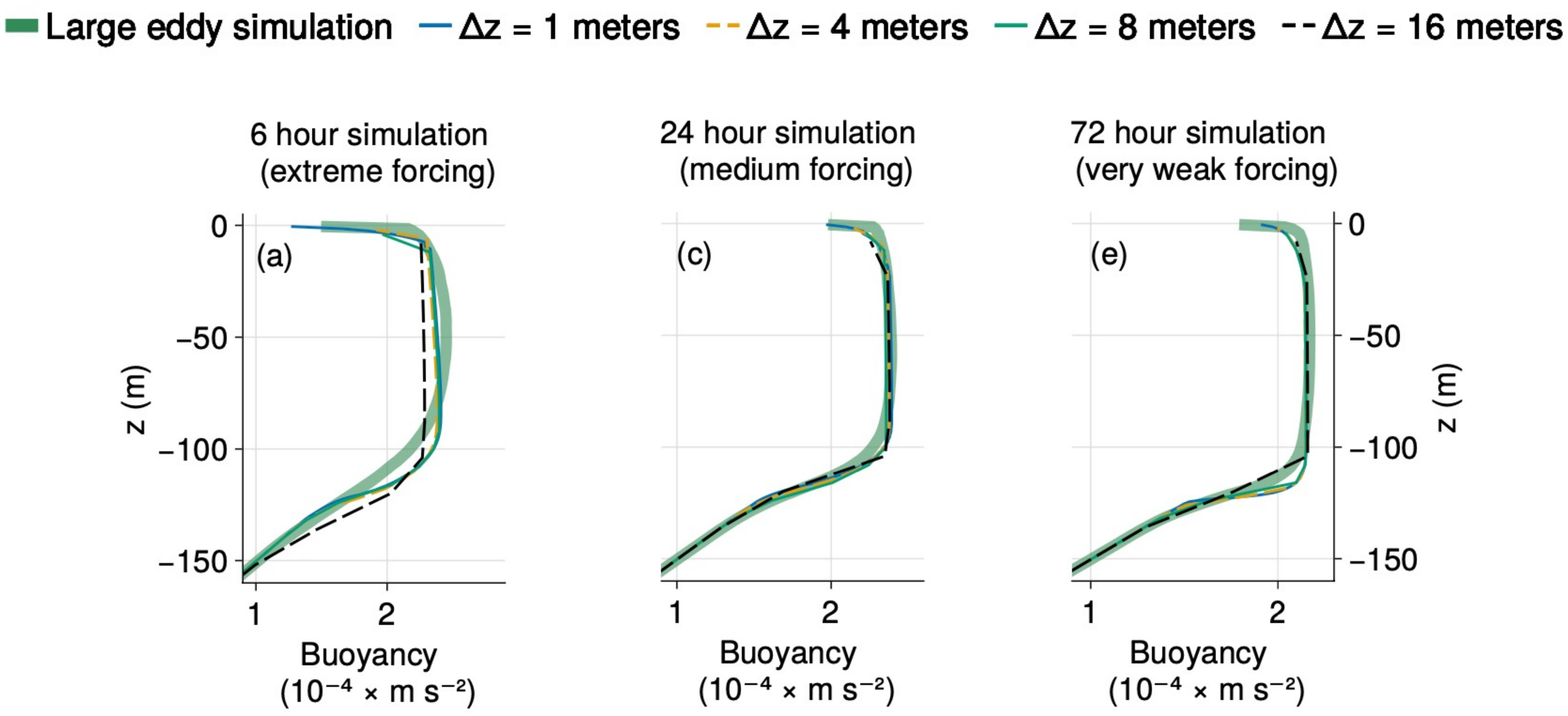




Realizing opportunity 1: more accurate models



Realizing opportunity 2: flexible error design



Realizing opportunity 3: accelerating model development

Evaluating three formulations of CATKE of increasing complexity

Minimalist

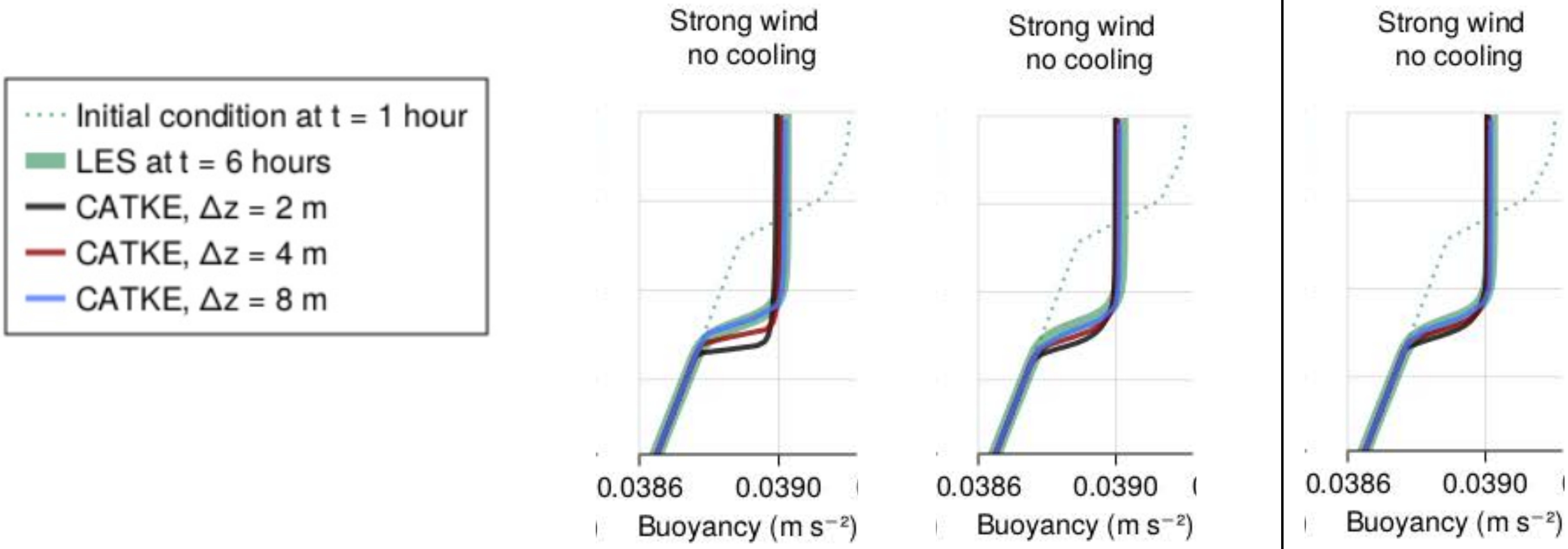
(7 parameters)

Variable Pr but no convective adjustment

(13 parameters)

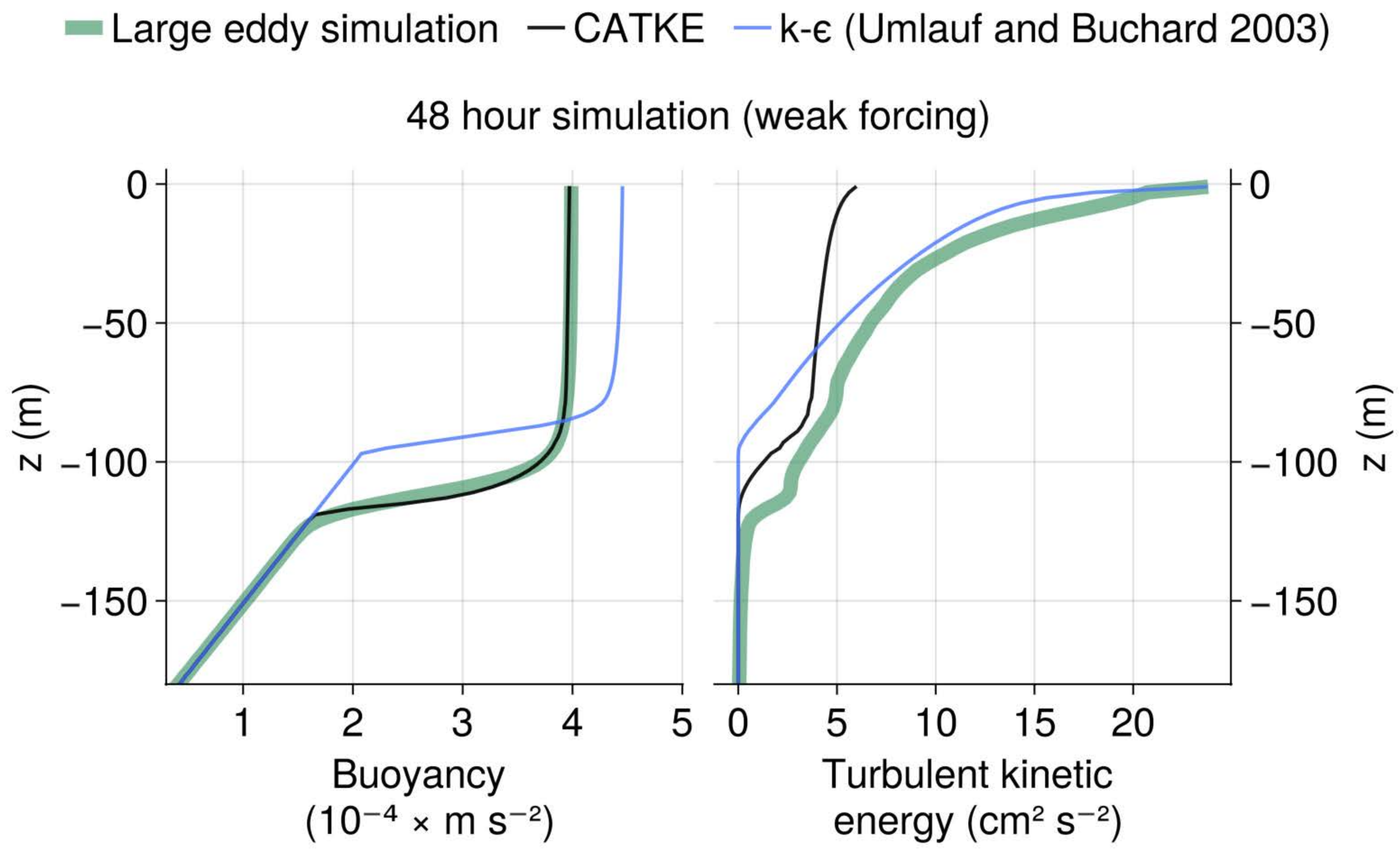
“Favorite”

(23 parameters)



Pitfall: compensating error

Do we get the right answer for the wrong reasons?



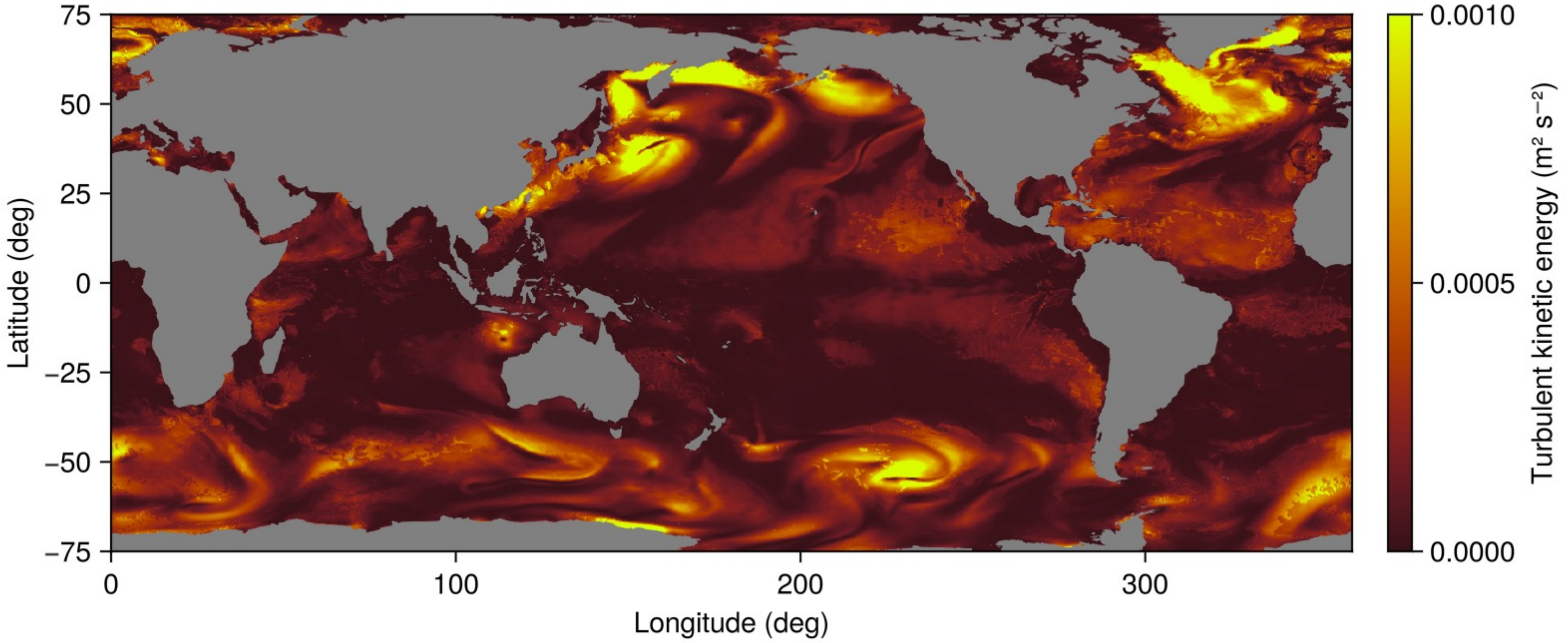
Solution

Reinterpret CATKE's "TKE" as a latent variable

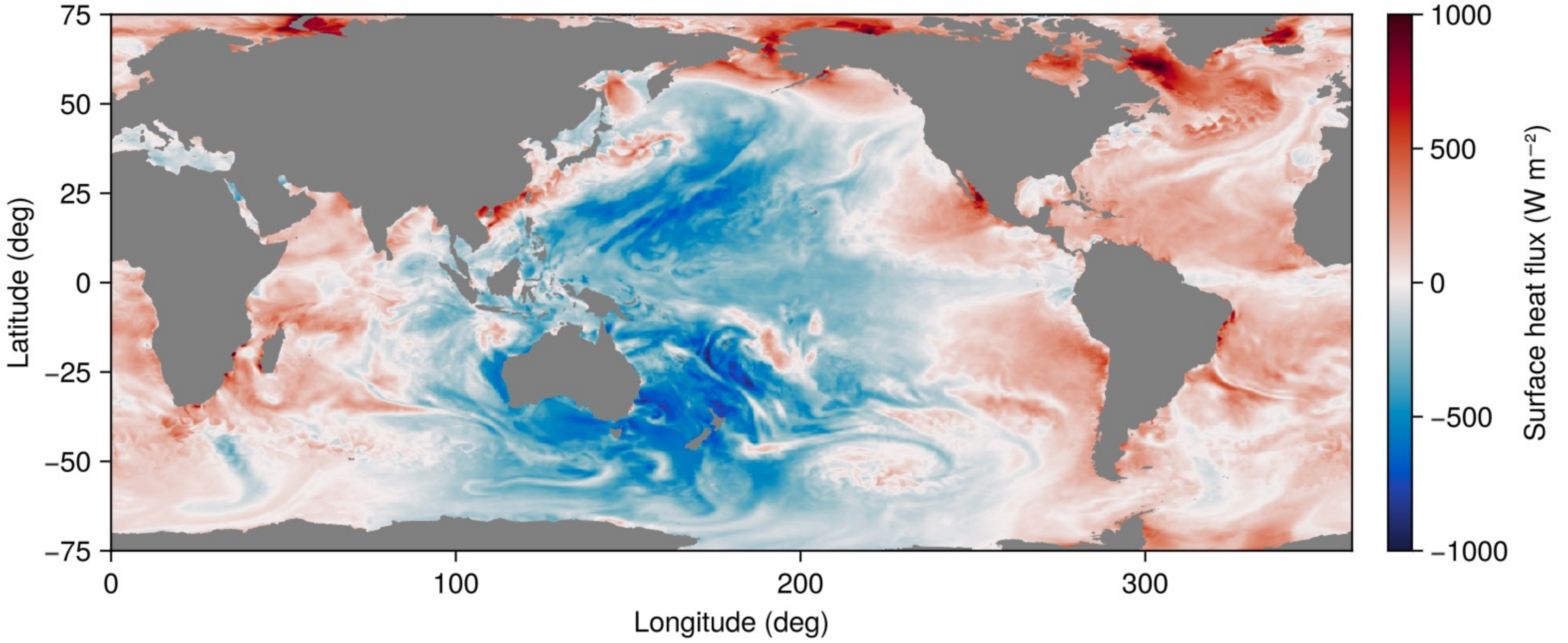
Pitfalls in calibrating against observations (1)

Re-calibrating CATKE together with surface flux parameterization

CATKE's turbulent kinetic energy



Surface heat flux



Solution

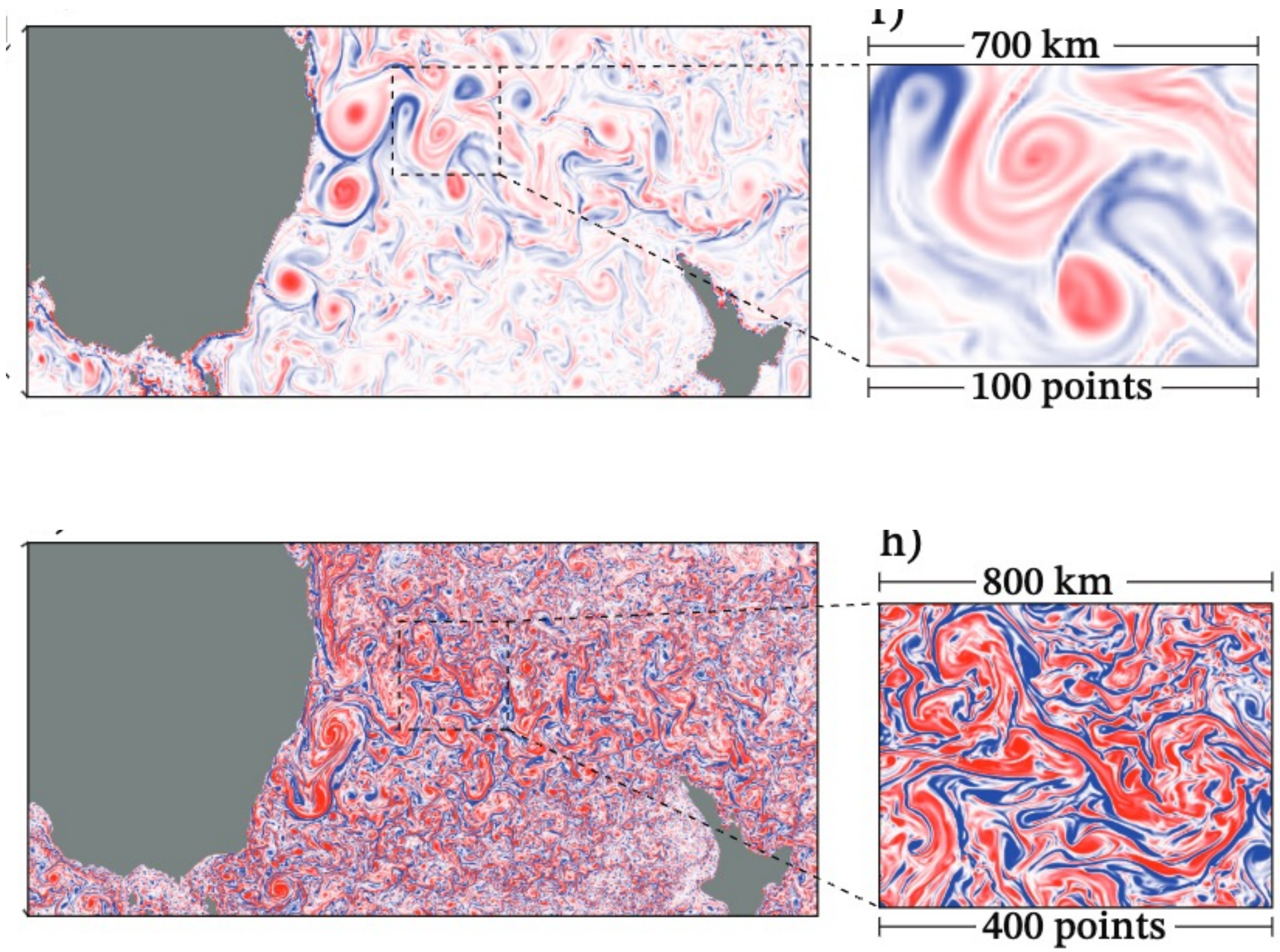
Need data on both mixed layer depth and SST

Pitfalls in calibrating against observations (2)

Re-calibrating CATKE without submesoscale restratification

~8 km resolution

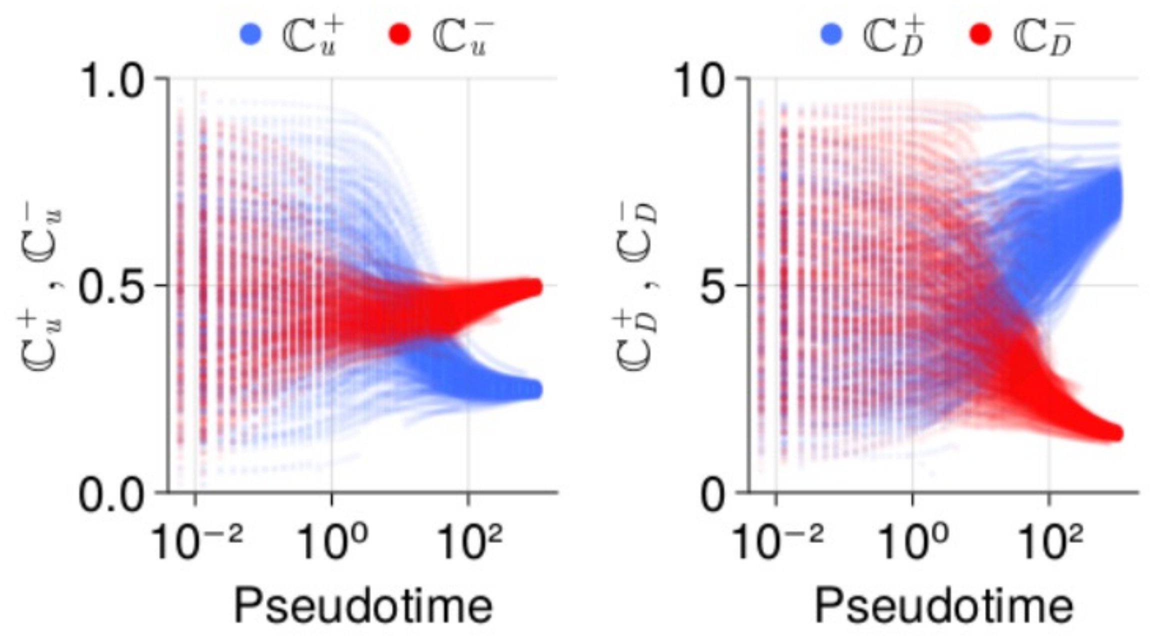
~2 km resolution



Fox-Kemper et al 2011 but see also Sinha and Callies 2023

Solution

Use uncertainty quantification to constrain CATKE parameters a priori



Summary (opportunities and pitfalls of automated calibration)

- **Opportunities when using automated calibration:**
  - More accurate models
  - Faster parameterization development
  - Flexible error design
- **Pitfall: compensating error**
- **But we have solutions:**
  - *A priori* constraints via automated calibration + uncertainty quantification
  - Careful definition of observable vs latent variables
  - Use more data