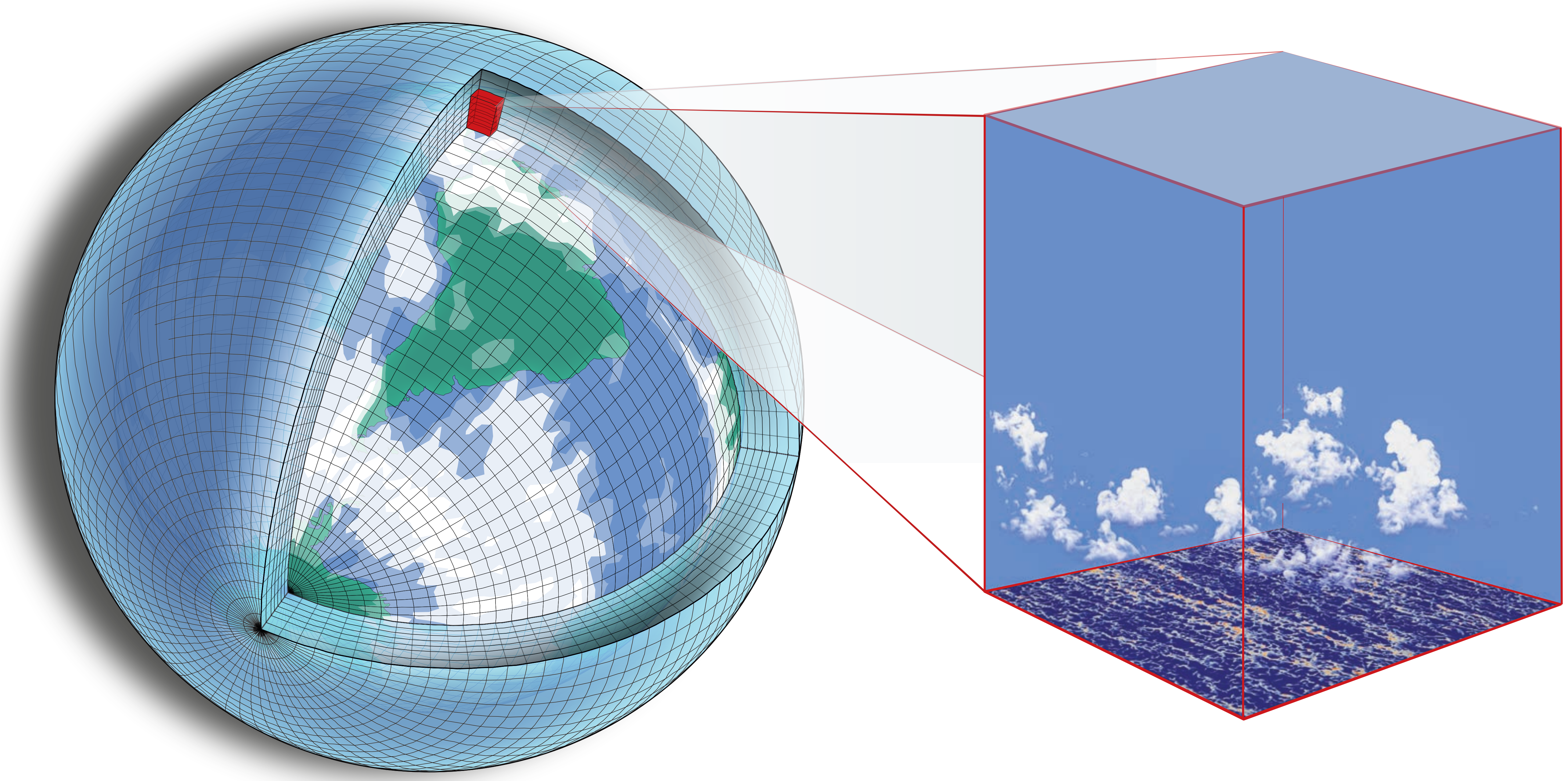
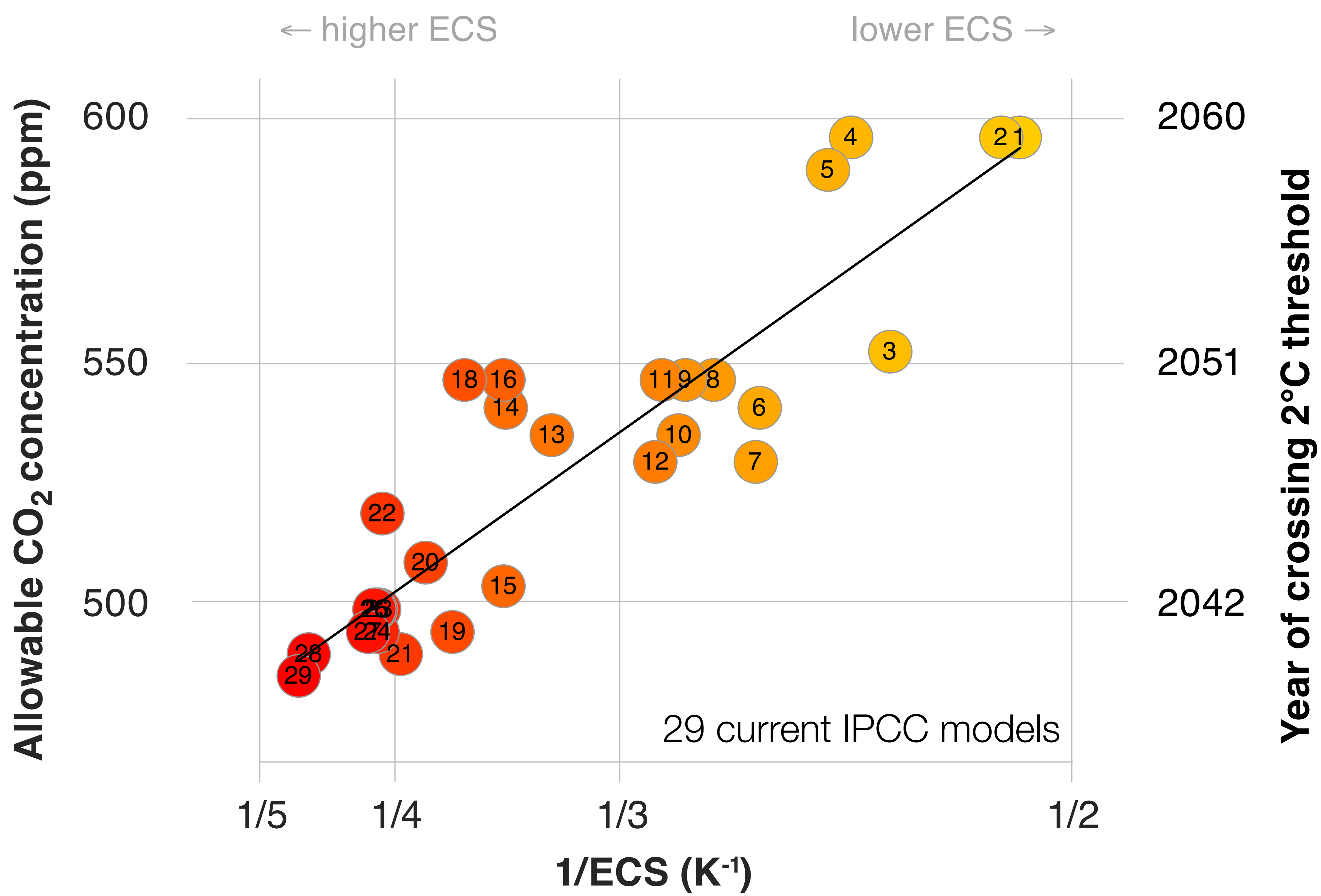


Unifying the modeling of boundary layers, convection, and clouds, and driving it with data

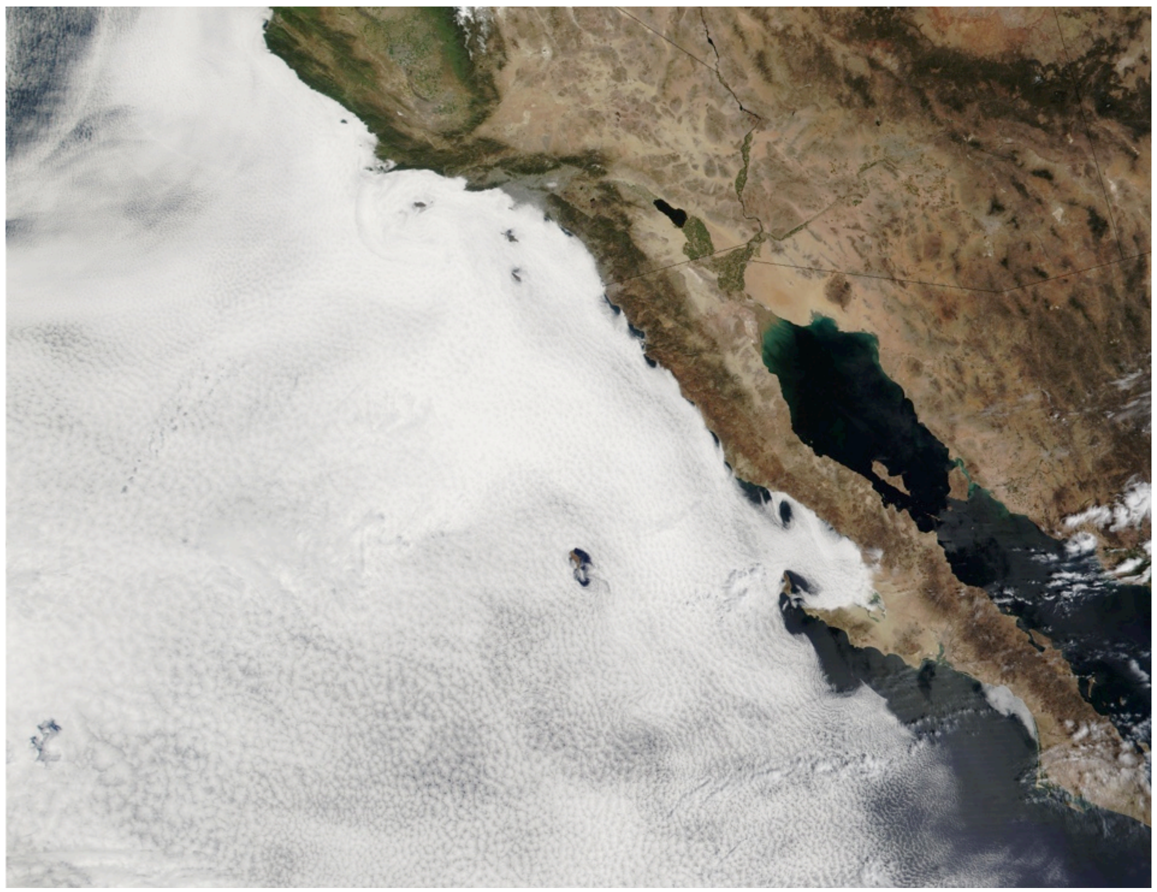
Tapio Schneider, Yair Cohen, Colleen Kaul, Shiwei Lan, Kyle Pressel, Andrew Stuart, Zhihong Tan, and Joao Teixeira

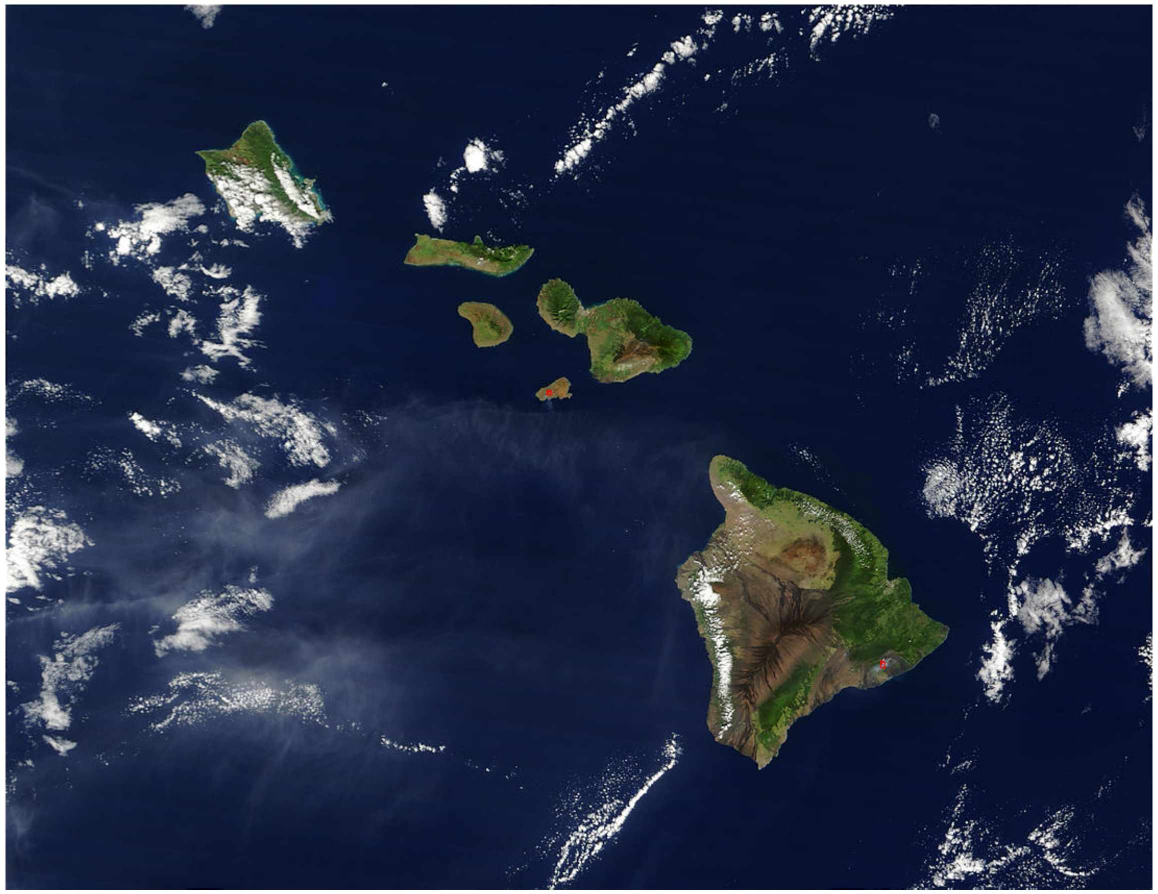
Climate predictions remain highly uncertain: E.g., allowable CO_2 concentration before crossing 2°C warming threshold



Schneider et al., *Nature Climate Change* 2017

Climate predictions are uncertain mostly because of low clouds



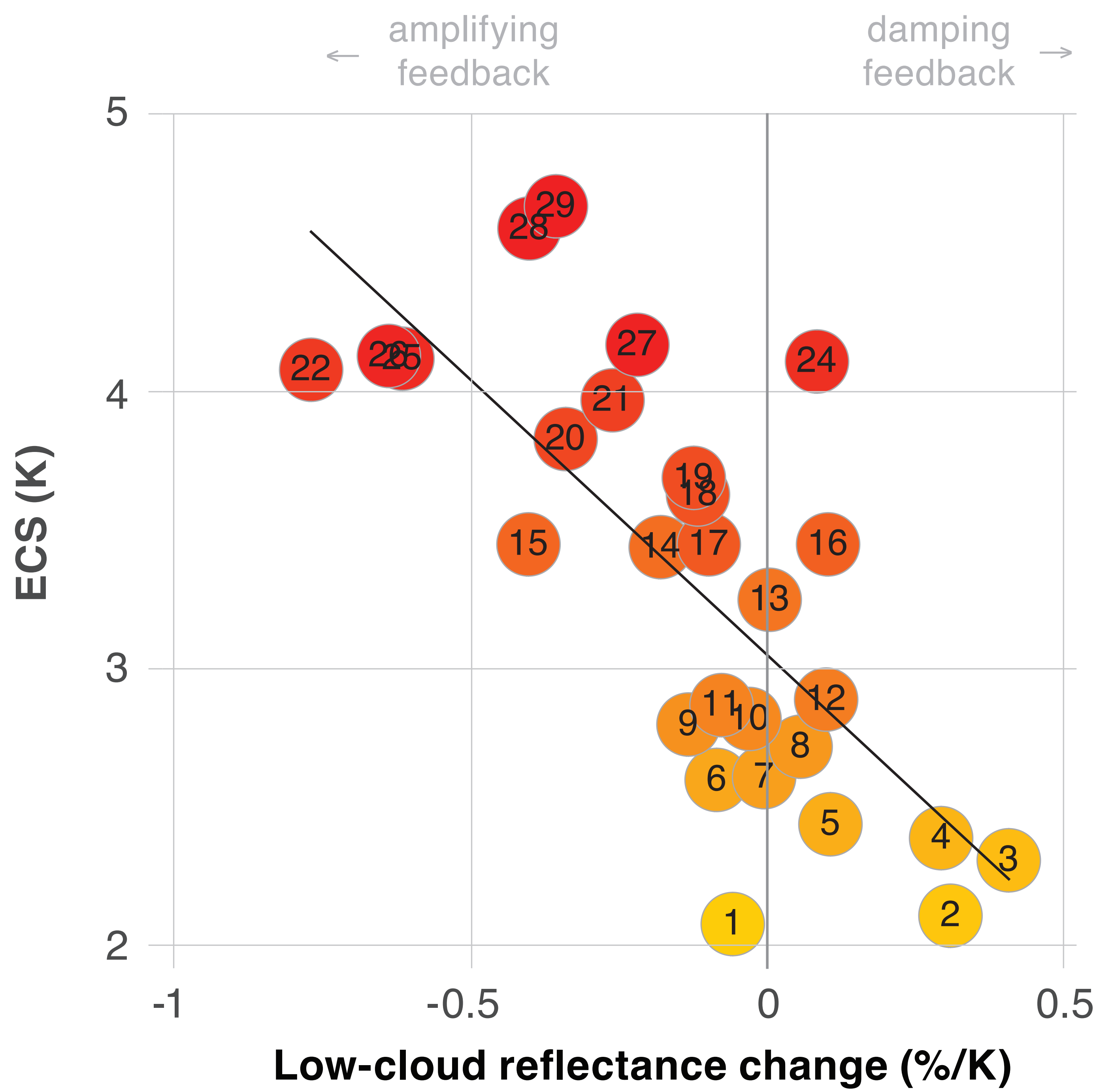


http://eoimages.gsfc.nasa.gov

Stratocumulus: colder

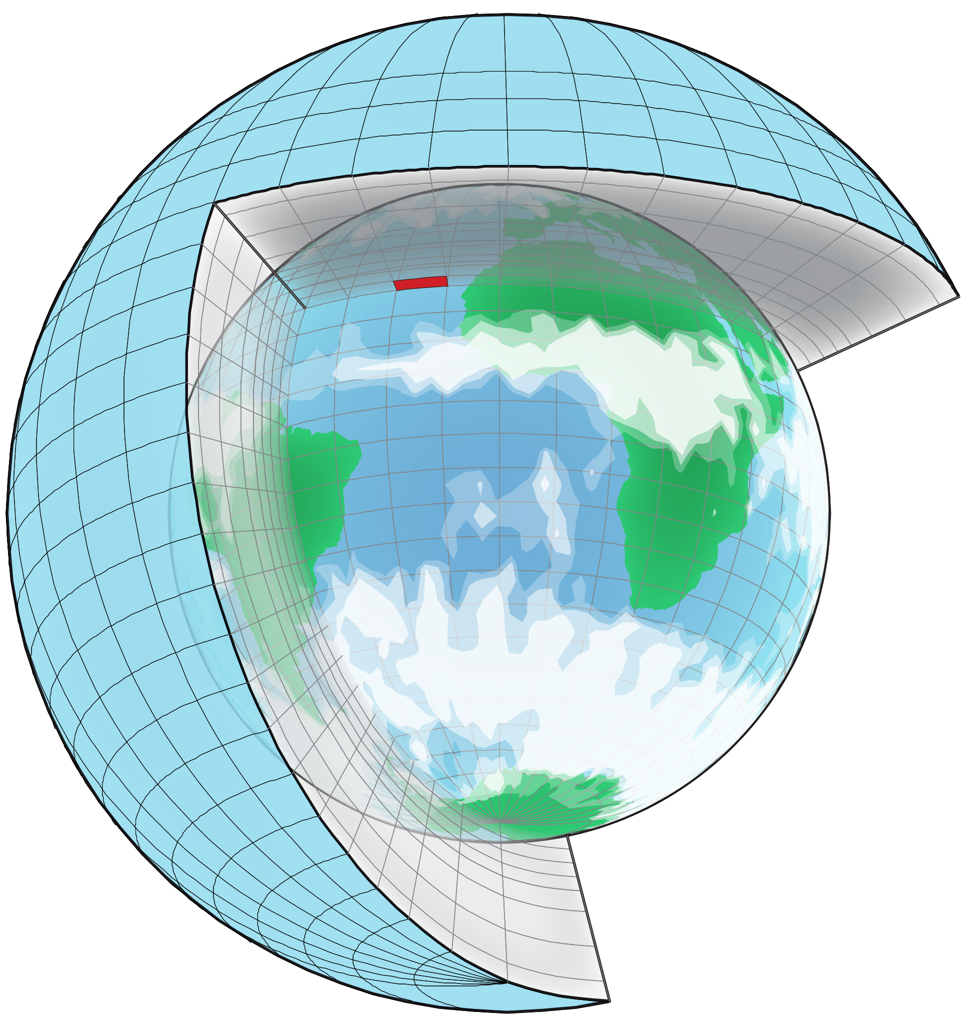
Cumulus: warmer

Low-cloud reflectance feedback accounts for majority of spread in climate response across models

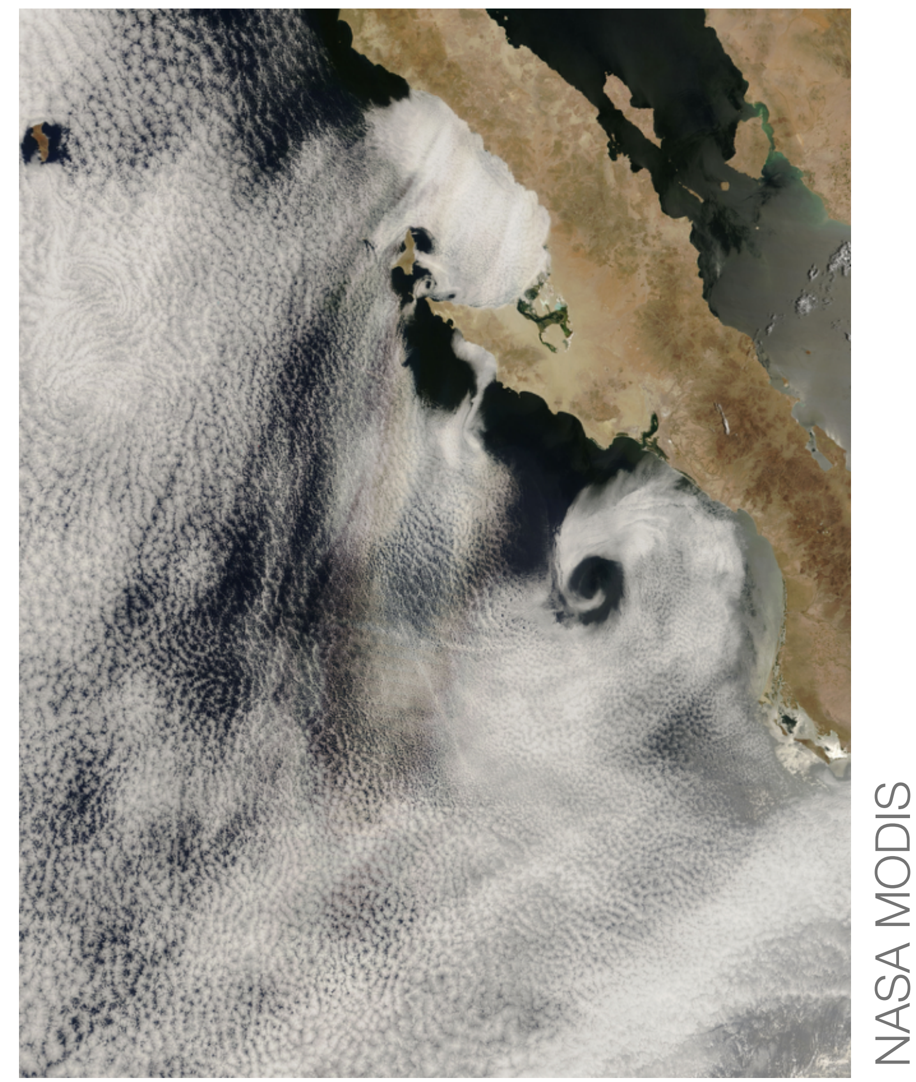


Schneider et al., Nature Climate Change 2017
(see also Cess et al. 1990, Bony & Dufresne 2005, and many others)

Climate models are too coarse to resolve clouds

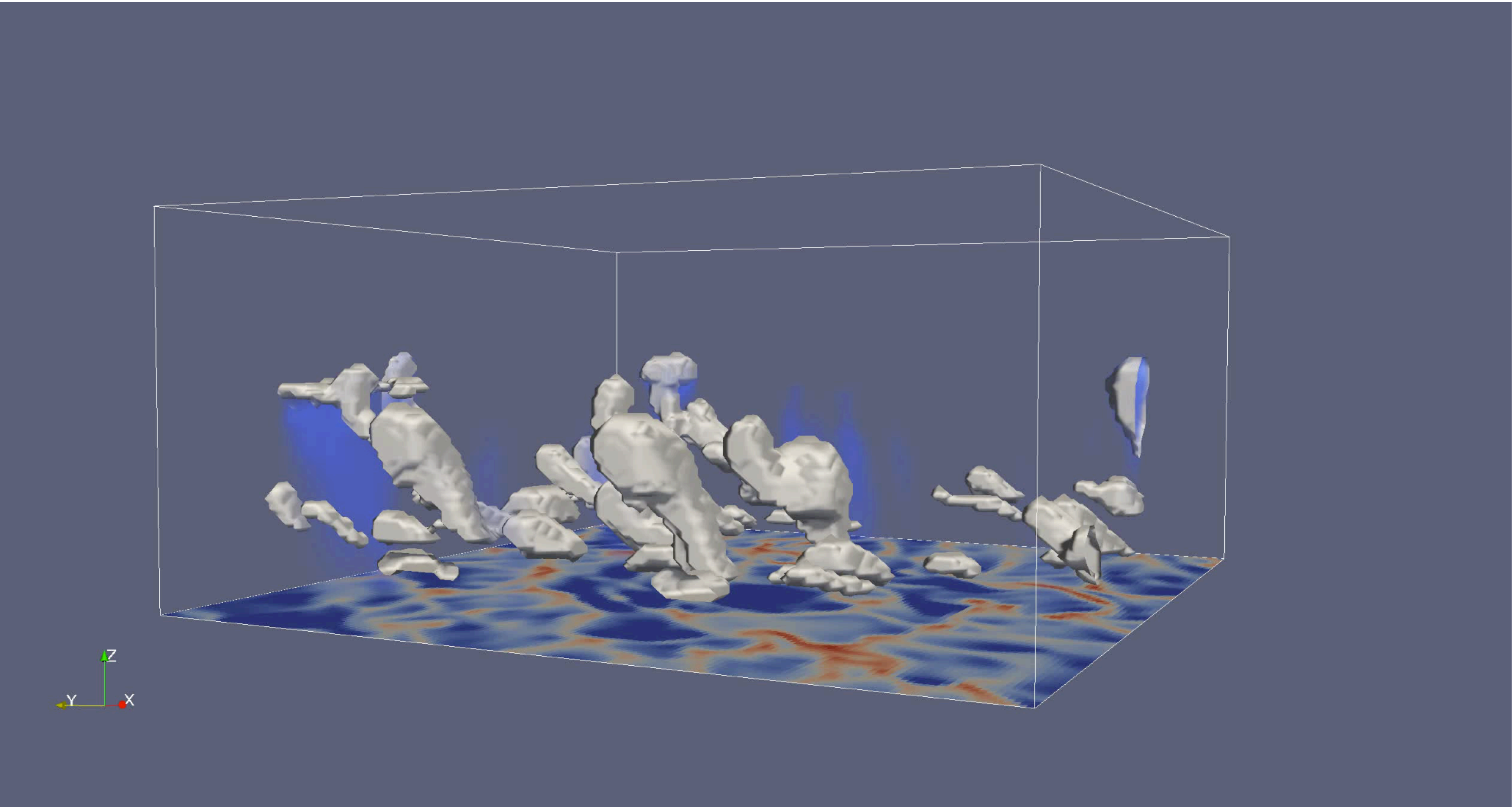


Global model:
~100 km resolution



Cloud scales: ~10-100 m

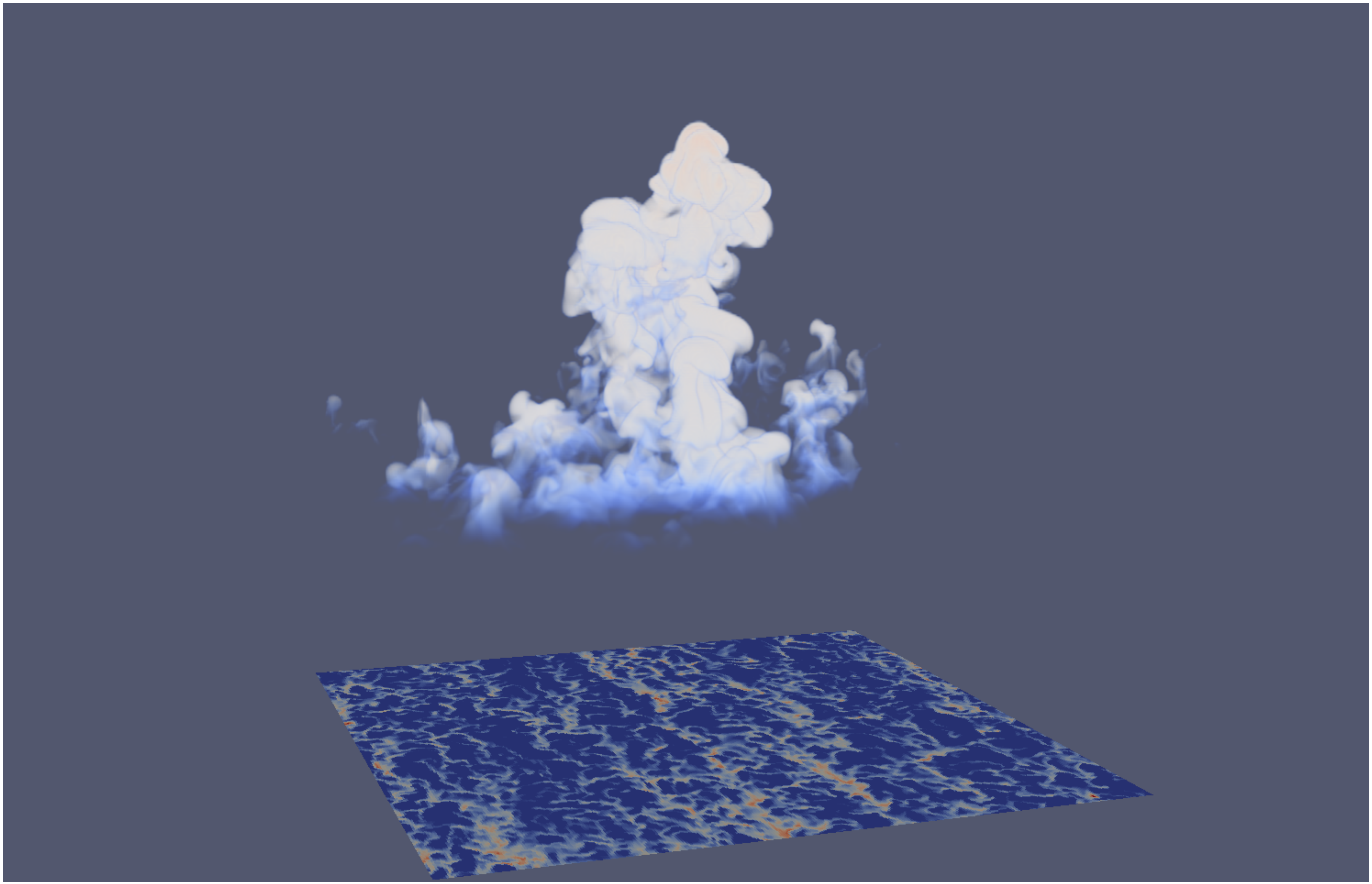
But we can simulate clouds in limited areas



Large-eddy simulation of tropical cumulus

Simulation with PyCLES (Pressel et al. 2015)

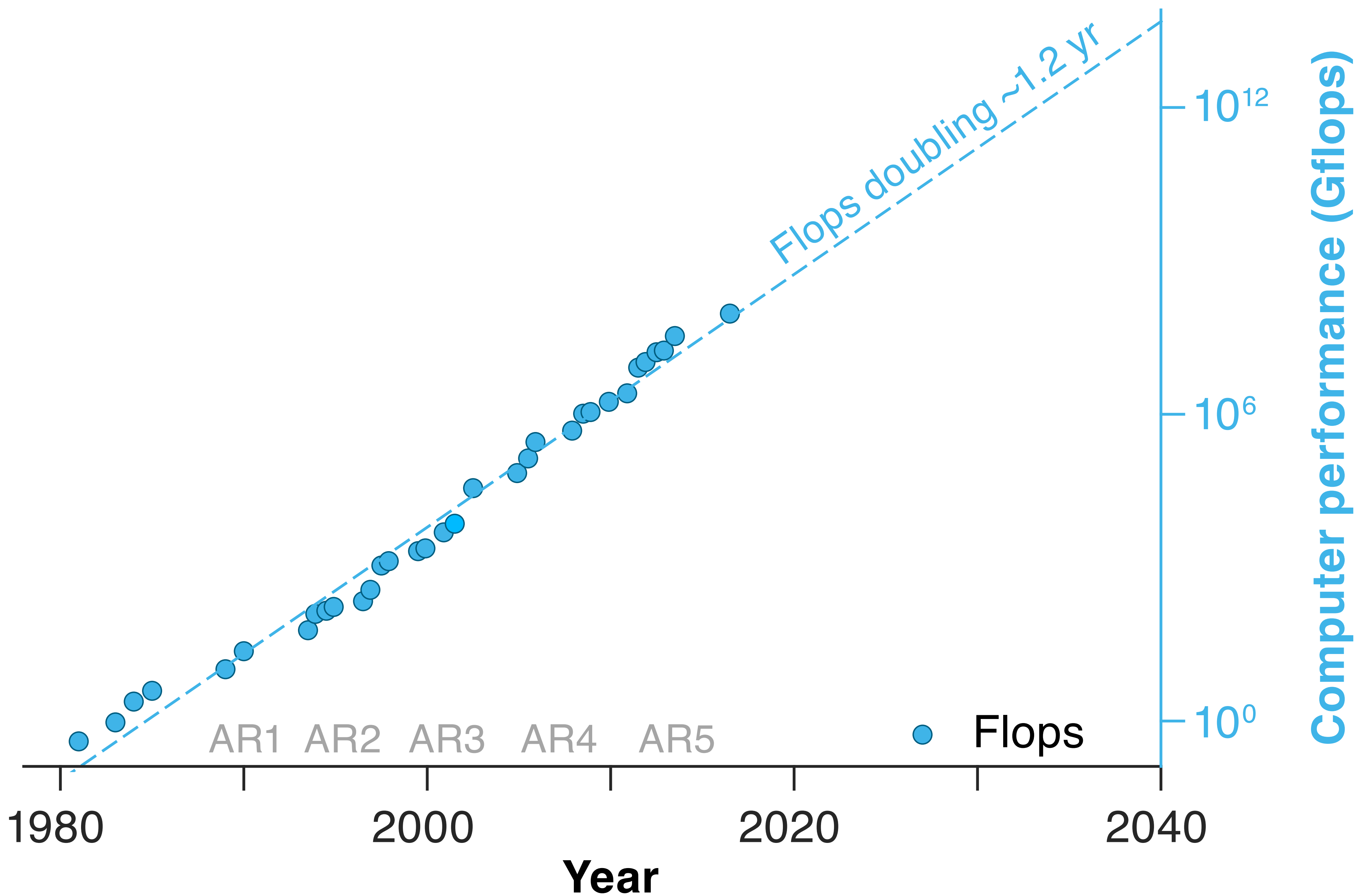
The simulations are so reliable, they are set in stone…



The simulations are so reliable, they are set in stone…

Karen LaMonte's *Cumulus* is at the Biennale now

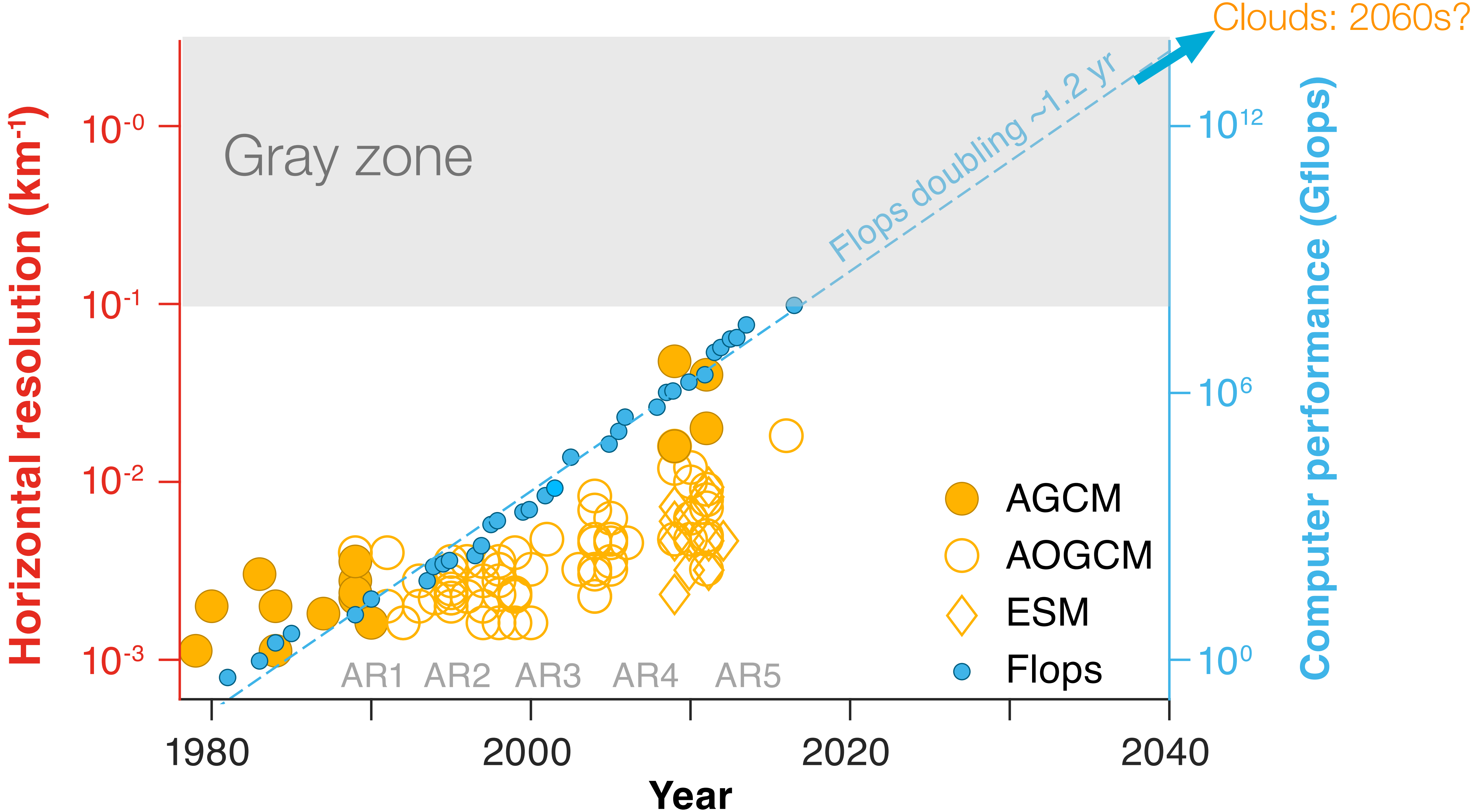
When will we be able to compute clouds globally?



Peak performance of fastest computer

Schneider et al. *Nature Climate Change*, 2017

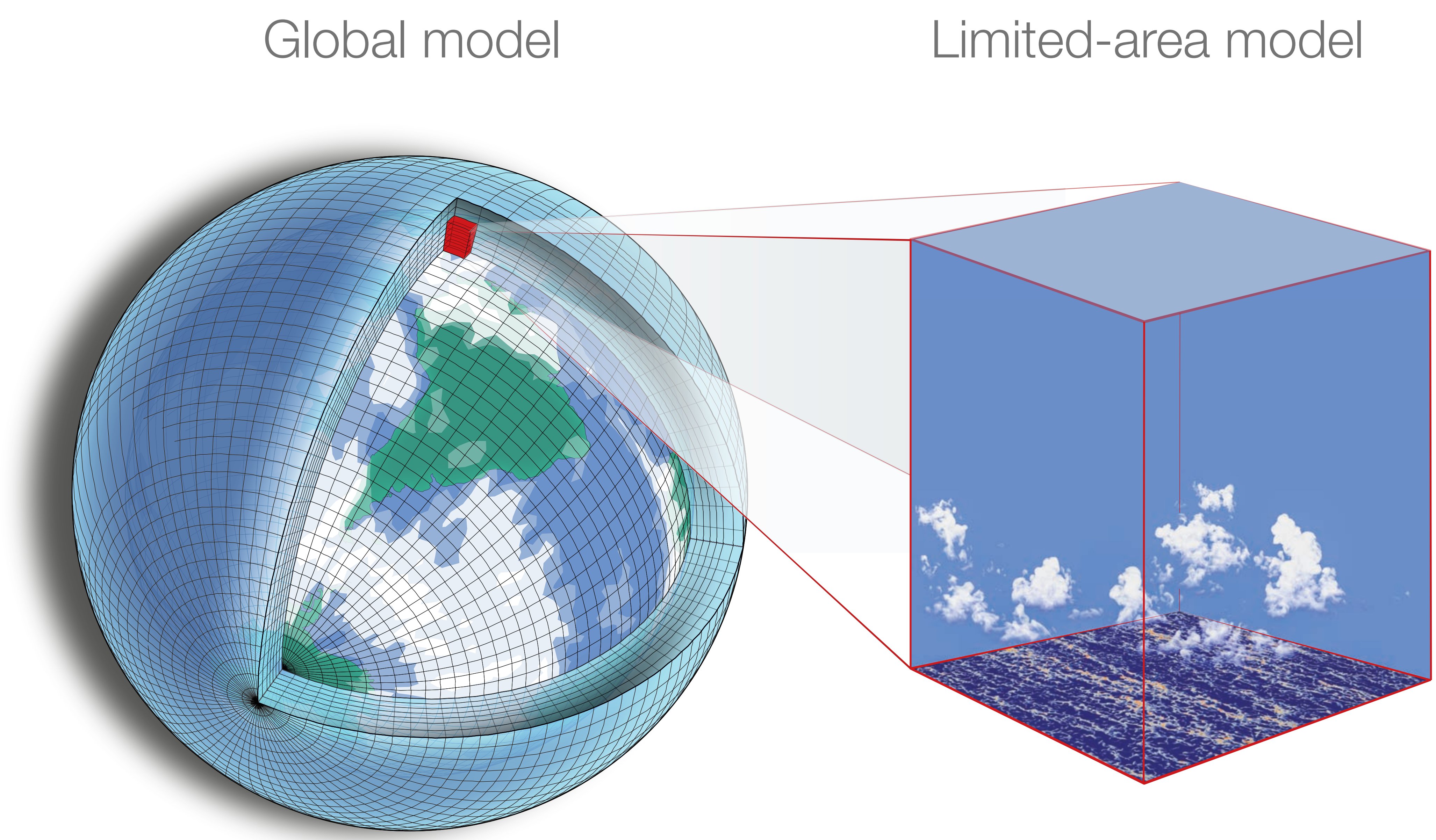
Global cloud resolving models not before 2060



Climate model resolution

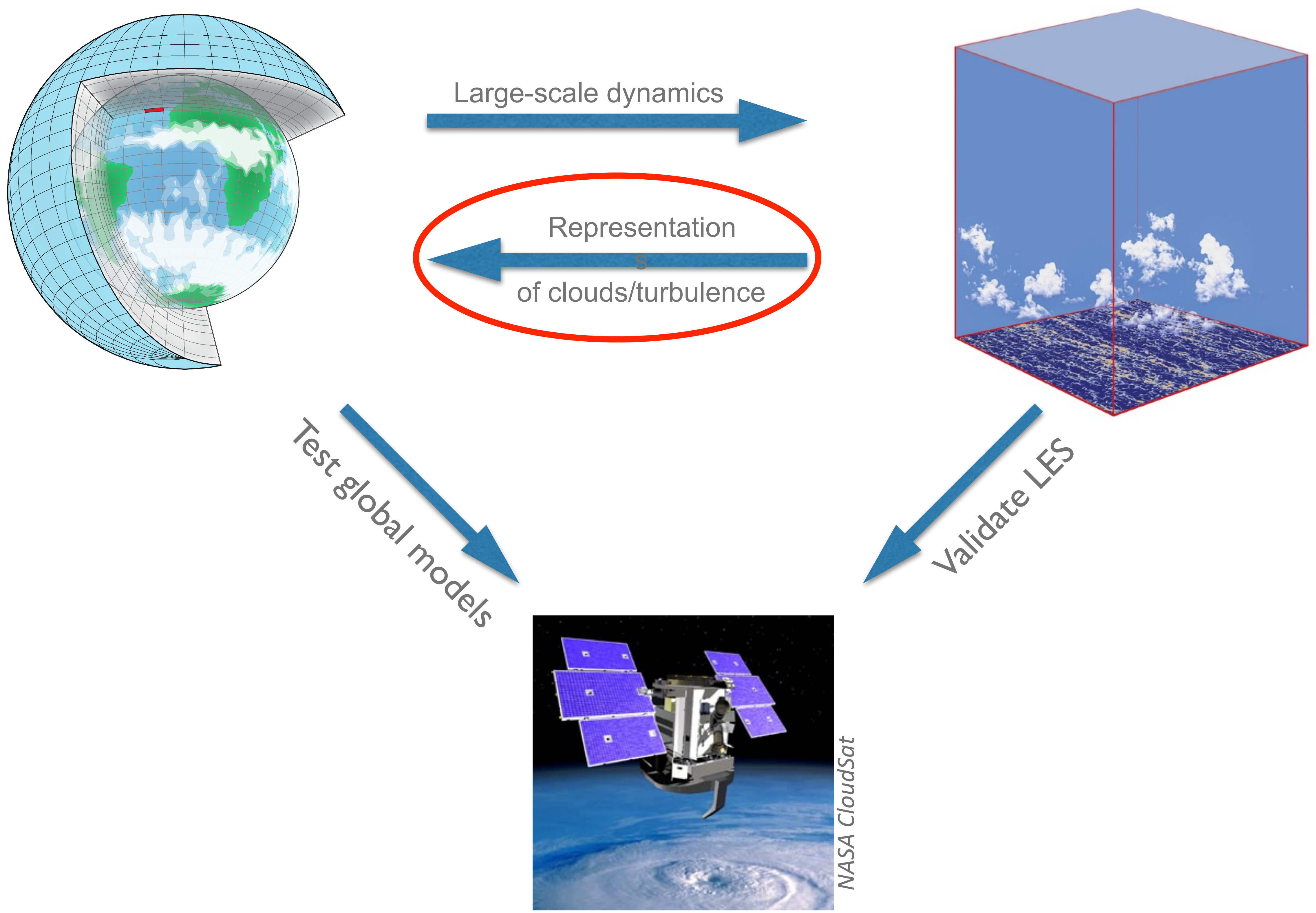
Schneider et al. *Nature Climate Change*, 2017

What we can do now



Use global and limited-area models in hierarchical framework to develop parameterizations

We can make progress exploiting model hierarchy and new data



Cloud/boundary layer turbulence schemes in current GCMs have unphysical discontinuities

- **Deep convection**: Often mass flux schemes (e.g., Arakawa & Schubert1974, Tiedtke 1989; Arakawa & Wu 2013)
- **Shallow convection**: Often also mass flux schemes, but with discontinuously different parameters (e.g., entrainment rates)
- **Boundary layer turbulence**: Often diffusive; difficult to match with cloud layer (e.g.,Troen & Mahrt 1986)

Parametric and structural discontinuities for processes with common (e.g., dry) limits

We use plumes/environment decomposition to develop unified representation of all SGS turbulence

Use adiabatically conserved variables $\phi = \{\theta_l, q_t\}$; partition fluxes into updraft, environment, and (possible) downdraft components (Siebesma & Cuijpers 1995, Soares et al. 2004):

$$\overline{w'\phi'} = a_u \overline{w'\phi'}_u + (1-a_u)\overline{w'\phi'}_e + a_u(1-a_u)(w_u - w_e)(\phi_u - \phi_e)$$

Traditionally plume variance is taken to be small:

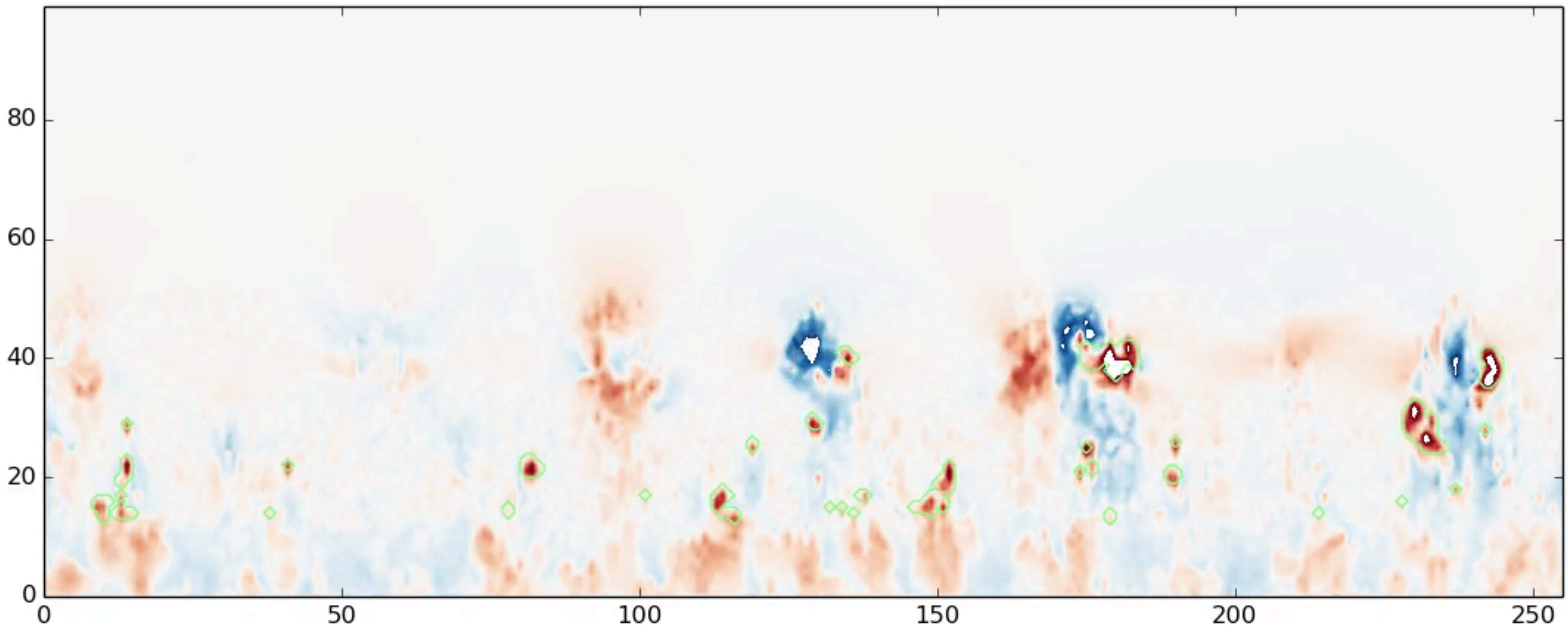
$$\overline{w'\phi'} = (1-a_u)\overline{w'\phi'}_e + a_u(1-a_u)(w_u - w_e)(\phi_u - \phi_e)$$

1st term focus in BL schemes, 2nd (mass flux) in convection. Keep both together!

Based on ideas by Siebesma, Teixeira, Soares et al.

Phenomenology of turbulence motivates plumes/ environment decomposition (BOMEX, shallow Cu)

Colors: vertical velocity (red up, blue down); green contours: cloud condensate



(Kyle Pressel, code available at climate-dynamics.org)

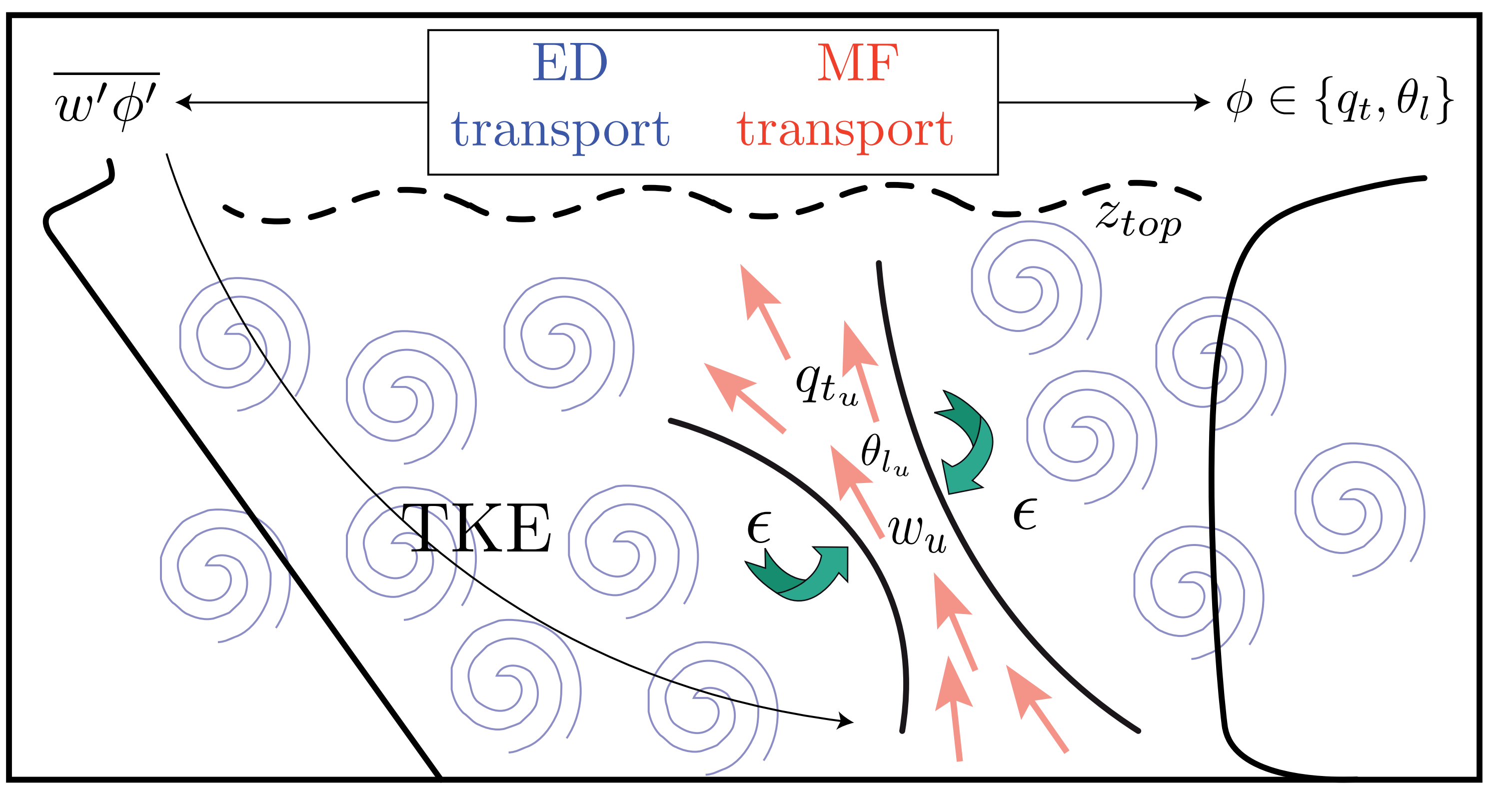
Eddy diffusion/mass flux scheme

Turbulent flux of conserved variables (Siebesma & Teixeira 2000)

$$\overline{w'\phi'} \approx -K(z)\frac{\partial\bar{\phi}}{\partial z} + M(z)(\phi_u - \bar{\phi})$$

ED term (environment)

MF term (updraft)



(Witek et al. 2011)

Structure of extended EDMF scheme

Decomposes domain into environment (i=0) and updrafts (i=1, …, N):

- Continuity:

$$\frac{\partial(\rho a_i)}{\partial t} + \frac{\partial(\rho a_i \overline{w}_i)}{\partial z} + \nabla_h \cdot (\rho a_i \langle \mathbf{u}_h \rangle) = \underbrace{\rho a_i \overline{w}_i \left(\sum_j \epsilon_{ij} - \delta_i \right)}_{\text{Mass entrainment/detrainment}}$$

- Scalar mean:

$$\frac{\partial(\rho a_i\overline{\phi}_i)}{\partial t} + \frac{\partial(\rho a_i \overline{w}_i\overline{\phi}_i)}{\partial z} + \nabla_h\cdot\left(\rho a_i\langle\mathbf{u}_h\rangle\overline{\phi}_i\right) = \underbrace{-\frac{\partial(\rho a_i\overline{w_i'\phi_i'})}{\partial z}}_{\text{Turbulent transport}} + \underbrace{\rho a_i\overline{w}_i\left(\sum_j\epsilon_{ij}\overline{\phi}_j-\delta_i\overline{\phi}_i\right)}_{\text{Entrainment/detrainment}} + \underbrace{\rho a_i \overline{S}_{\phi,i}}_{\text{Sources/sinks}}$$

- Scalar covariance

$$\frac{\partial(\rho a_i\overline{\phi_i'\psi_i'})}{\partial t} + \frac{\partial(\rho a_i \overline{w}_i\overline{\phi_i'\psi_i'})}{\partial z} + \nabla_h\cdot\left(\rho a_i\langle\mathbf{u}_h\rangle\overline{\phi_i'\psi_i'}\right) = \underbrace{-\rho a_i\overline{w_i'\psi_i'}\frac{\partial\overline{\phi}_i}{\partial z} - \rho a_i\overline{w_i'\phi_i'}\frac{\partial\overline{\psi}_i}{\partial z}}_{\text{Generation/destruction by cross}-\text{gradient flux}}$$

$$\underbrace{+\rho a_i\overline{w}_i\left[\sum_j\epsilon_{ij}\left(\overline{\phi_j'\psi_j'}+(\overline{\phi}_j-\overline{\phi}_i)(\overline{\psi}_j-\overline{\psi}_i)\right)-\delta_i\overline{\phi_i'\psi_i'}\right]}_{\text{Covariance entrainment/detrainment}} \underbrace{-\frac{\partial(\rho a_i\overline{w_i'\phi_i'\psi_i'})}{\partial z}}_{\text{Turbulent transport}} + \underbrace{\rho a_i\left(\overline{S_{\phi,i}'\psi_i'}+\overline{S_{\psi,i}'\phi_i'}\right)}_{\text{Sources/sinks}}.$$

(Tan et al., JAMES 2018)

EDMF assumptions

- Close environmental fluxes diffusively

$$\overline{w_0'\phi_0'} = -K \frac{\partial \overline{\phi}_0}{\partial z}$$

- Assume drafts have no variance (can represent variance through multiple drafts)

- Use TKE-mixing length (*l*) closure for diffusivity

$$K = c_K l \sqrt{\overline{e}_0},$$

with prognostic equation for environmental TKE

(Tan et al., JAMES 2018)

Key ingredients of new scheme

- Can represent all SGS turbulent motion in unified manner, from boundary-layer turbulence to deep convection
- Explicit time dependence and SGS memory: essential at short time steps/high resolution
- Closes second moments (e.g., TKE) consistently with drafts
- Has variable draft fraction, with prognostic equation for area fraction (needed for resolution adaptivity)
- Can be viewed as representing PDFs through Gaussian plus *N* delta functions (Lappen and Randall 2001)

(Tan et al., JAMES 2018)

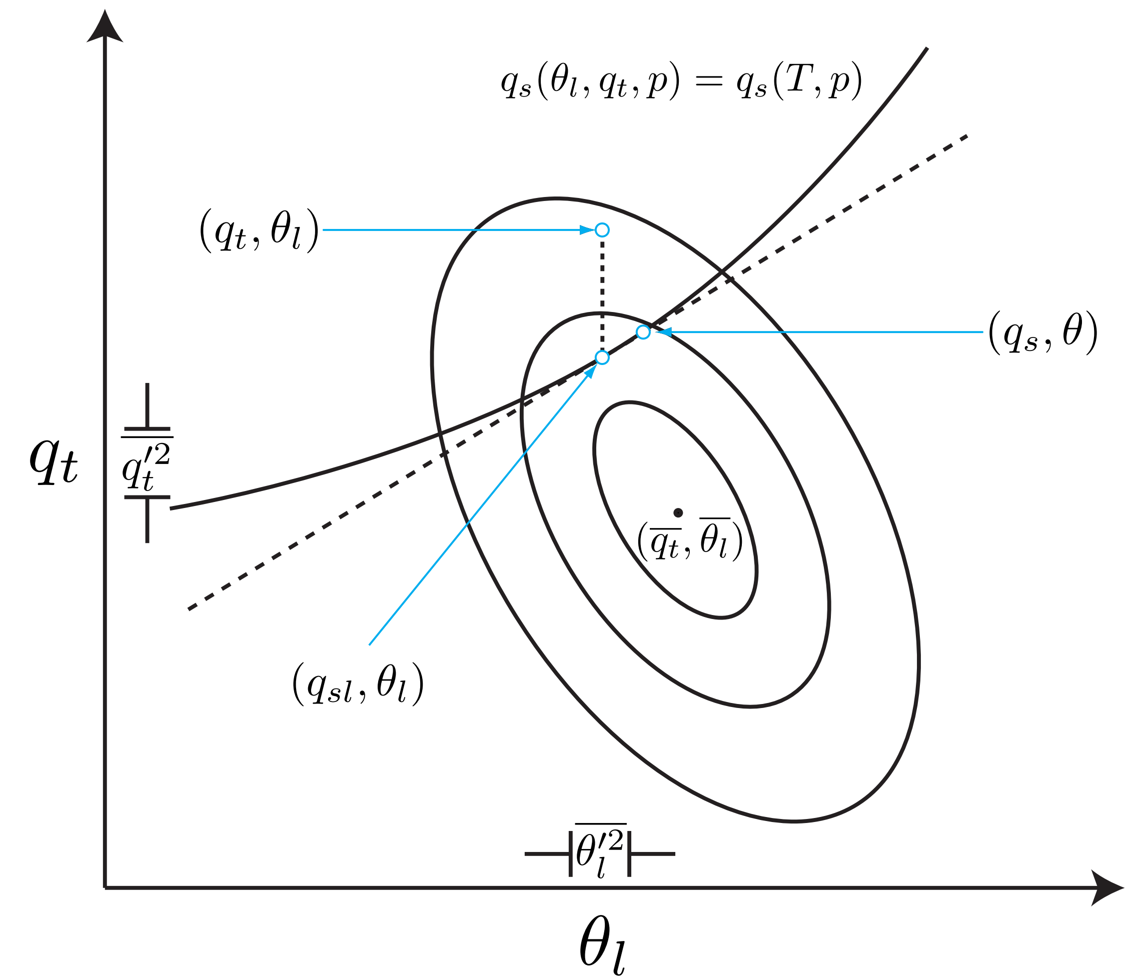
Couple EDMF scheme to probabilistic cloud scheme

Cloud fraction and liquid water assuming bivariate Gaussian *G* for SGS θ_l, q_t

(Sommeria & Deardorff 1977):

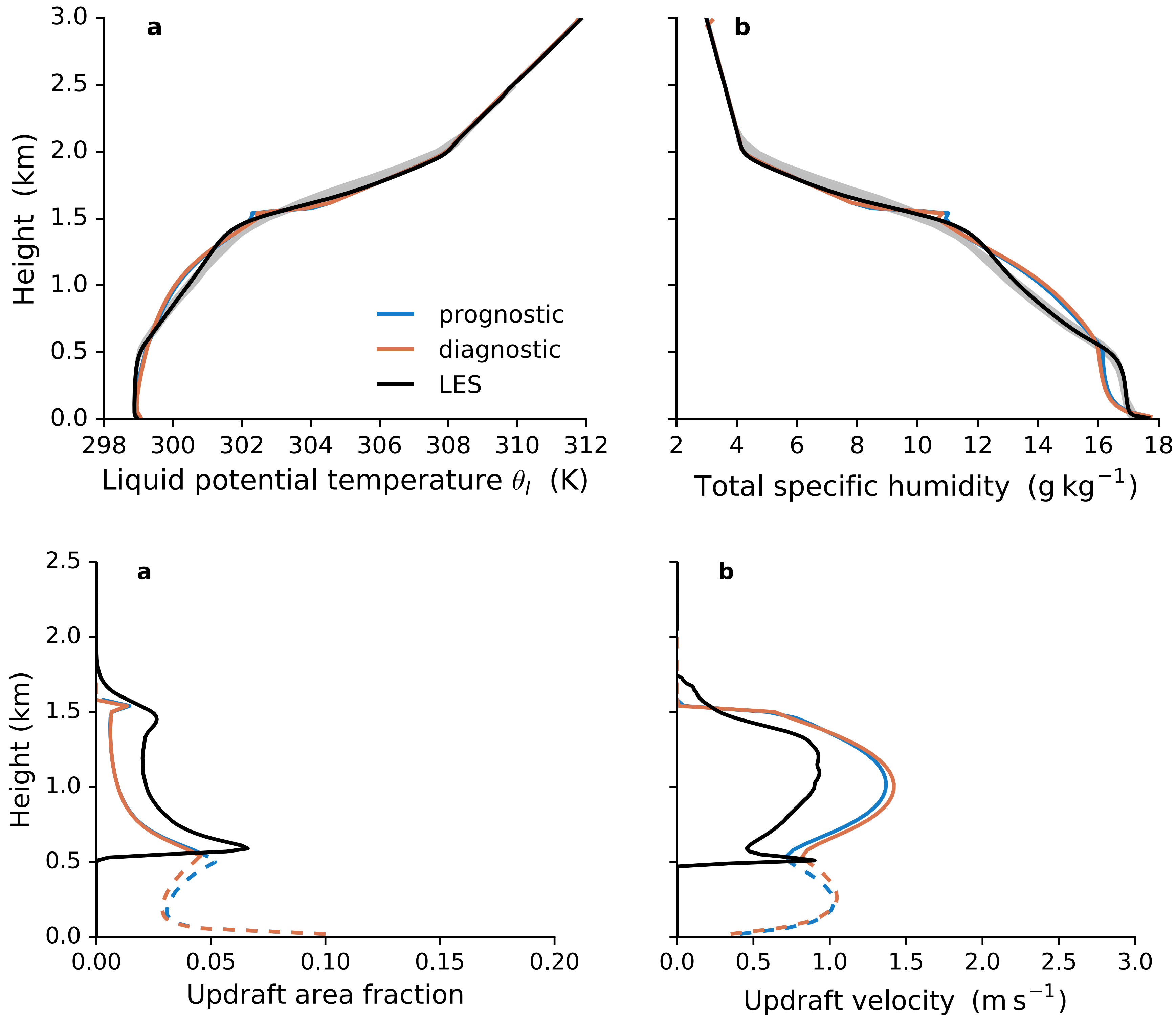
$$\mathrm{CF} = \int_{-\infty}^{\infty} \int_{q_{sl}}^{\infty} G \, dq_t \, d\theta_l$$

$$q_l = \int_{-\infty}^{\infty} \int_{q_s}^{\infty} (q_t - q_s) G \, dq_t \, d\theta_l$$



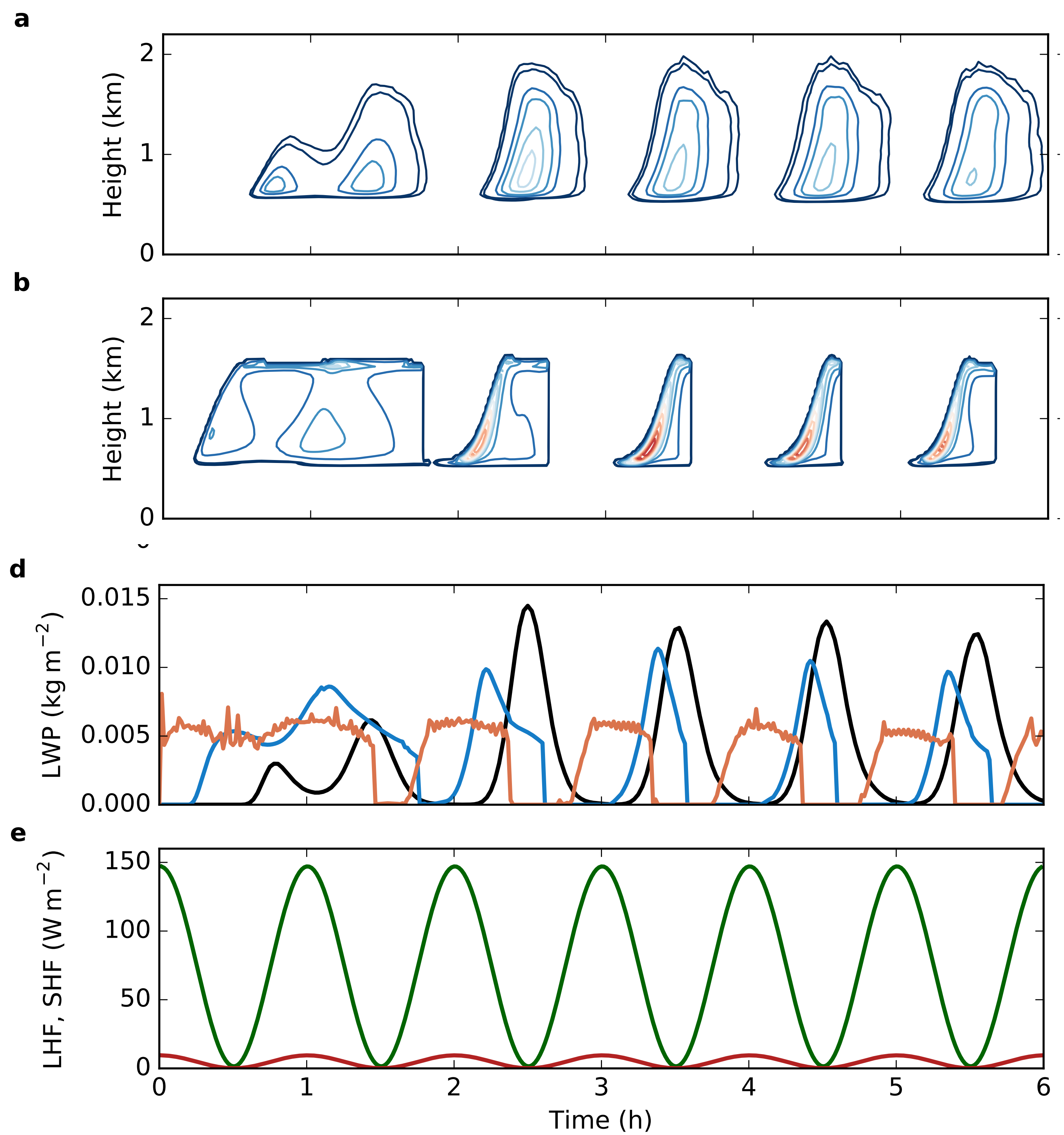
Approximately, cloud fraction and cloud water depend only on normalized saturation excess $\bar{s}/\sigma$, with $s = q_t - q_{sl}, \;\; \sigma^2 = \mathrm{var}(s)$

With simple closures, EDMF scheme works for Cu



Tan et al. JAMES, 2018

A more challenging test: convective lifecycles (oscillating BOMEX)



LES liquid water

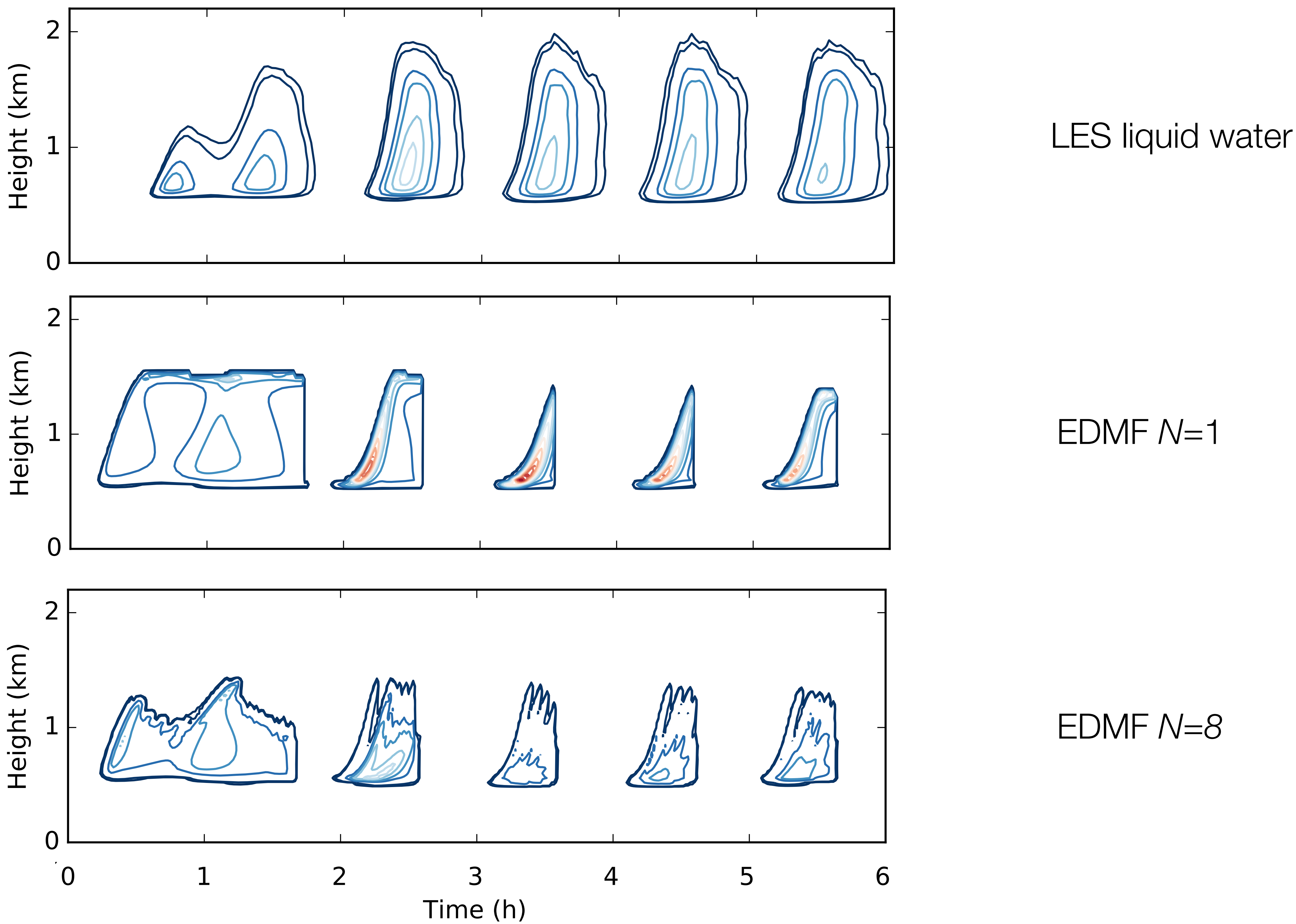
EDMF liquid water

Liquid water path

Surface fluxes

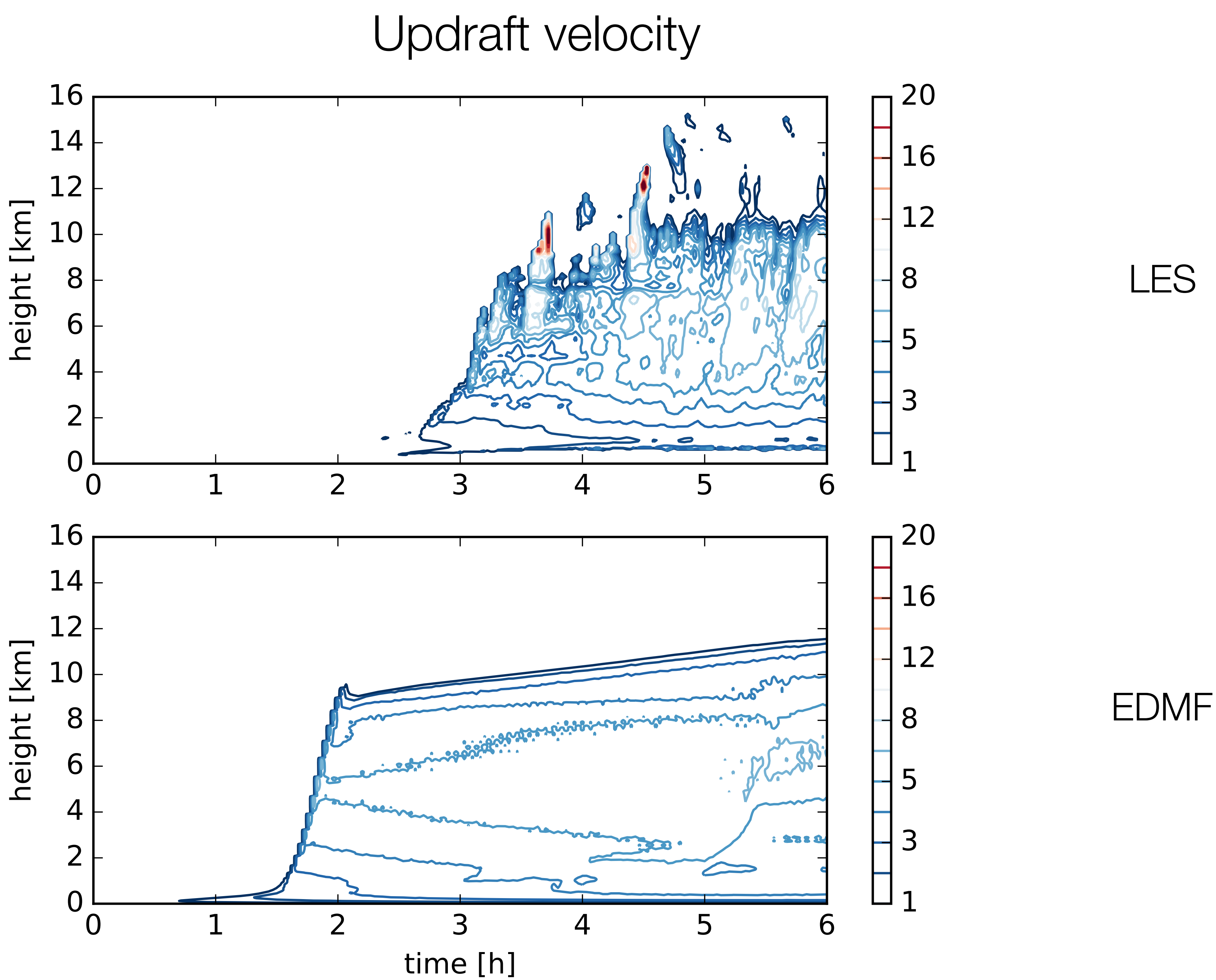
Tan et al. JAMES, 2018

Increasing number of updrafts increases accuracy



Yair Cohen

EDMF scheme can also simulate deep convection (TRMM LBA)



Yair Cohen

The new unified EDMF scheme…

- is prognostic (essential in gray zone)
- can simulate range of motion from boundary layer turbulence to deep convection
- reduces number of tunable parameters relative to the plethora of parameters in traditional schemes (for BL turbulence, shallow convection, deep convection)
- is scale adaptive (e.g., number of plumes can depend on grid size)

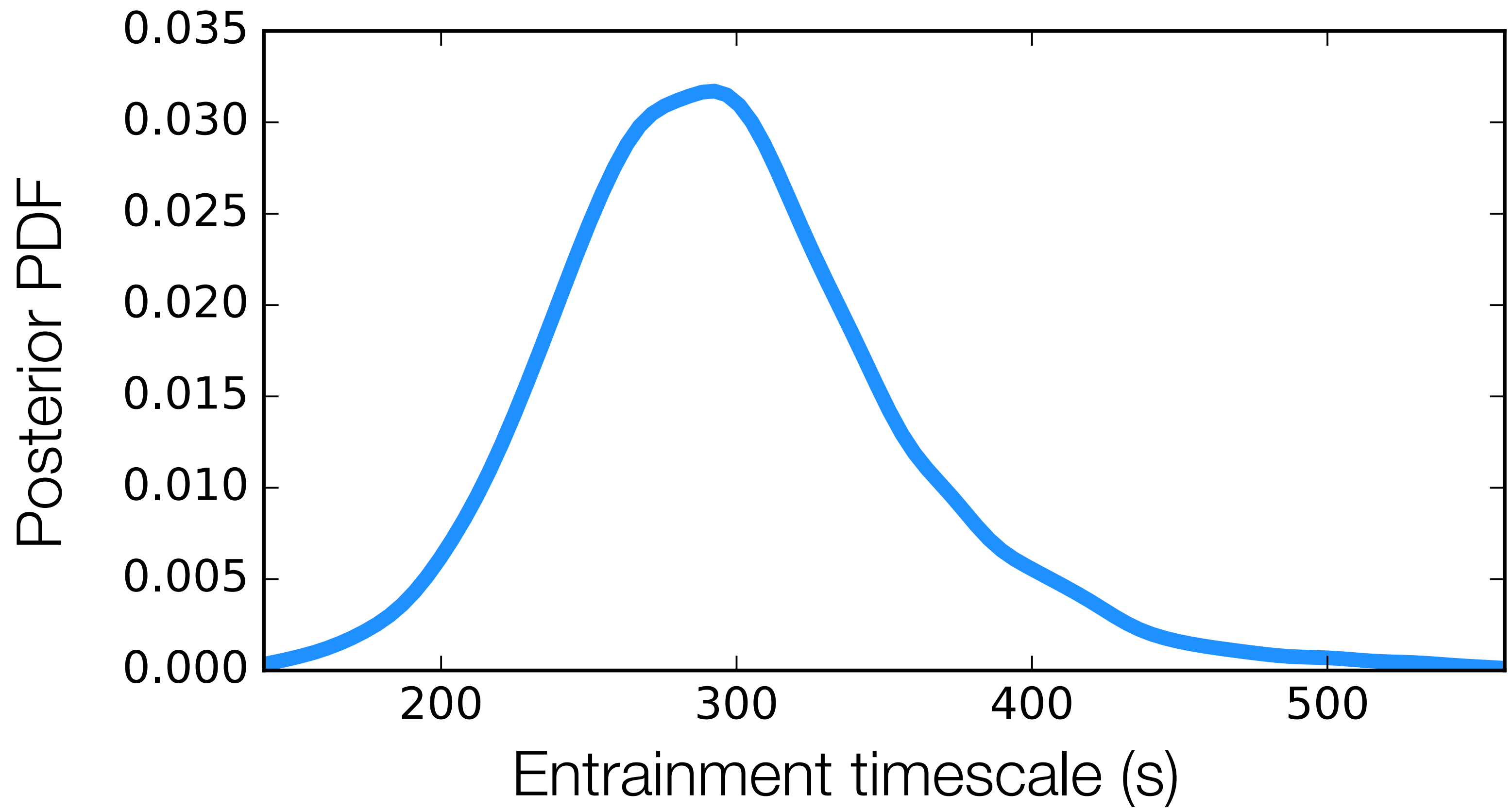
We are nearly done developing the structure of the scheme; next step is automated learning of closure parameters

- Learning by minimizing least-squares objective function

$$J_0(a) = \frac{1}{2}\left\| \left\langle f(lwp, cf, q_t, q_l) - f_{LES}(lwp, cf, q_t, q_l) \right\rangle \right\|_{\Sigma}^{2}$$

- Moment vector f=(LWP, CF, q_t, q_l)

- Posterior PDF with MCMC (e.g., entrainment timescale)



Ongoing work: physics-informed machine learning

- Model closure parameters as a flexible functions of all possible nondimensional groups

- E.g., entrainment rate

$$\epsilon_i = \frac{1}{a_i^{1/2}\Delta} g\left(\frac{\mathrm{TKE}_0}{\bar{w}_i^2}, \frac{\bar{b}_i}{\bar{w}_i^2}, \ldots\right)$$

- Estimate functions *g* as parametric (or even non-parametric) functions from data

- Data right now are a library of LES simulations driven by GCM output. Observational data later.

Broader vision: DA/machine learning can help improve unsatisfying closure parameter estimation

- Currently, climate models use global mean of TOA radiative balance and surface temperature in “tuning”
- More detailed space-based observations (CloudSat, CALIPSO etc.) are mostly used for model evaluation
- High-resolution simulations (LES) used only for specific regions/times, not broadly sampling globe
- Instead: use ***available observations*** and ***targeted high-resolution simulations*** in a climate model that learns automatically from data

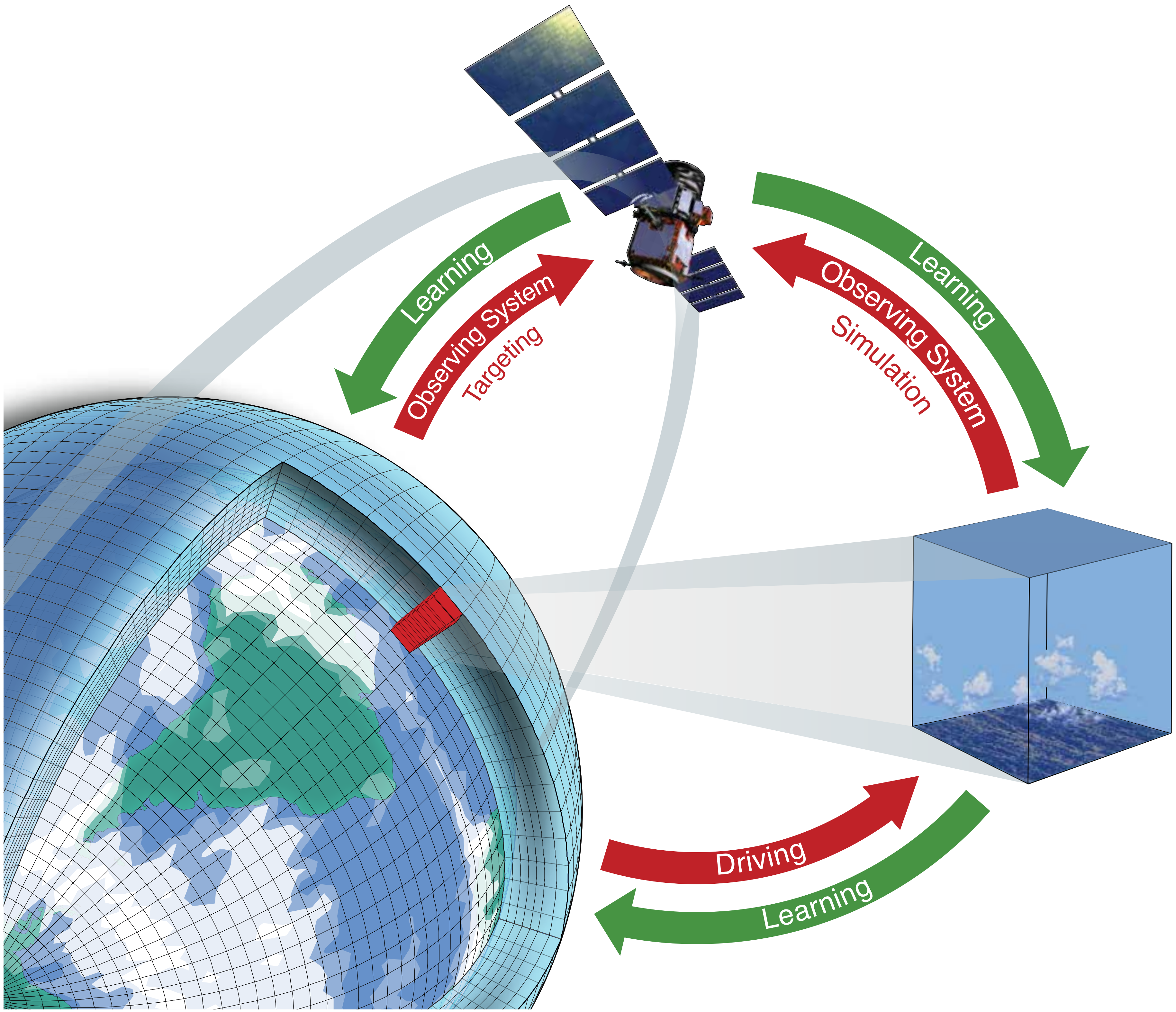
Earth System Modeling 2.0: Toward Learning Models

- Build models that learn adaptively from observations (especially space-based) and high-resolution simulations nested in “supercolumns”

- A model that is aware of its uncertainties can target supercolumn LES to reduce uncertainties where needed (similar to targeted observations in NWP)

- Integrate data and nested high-resolution simulations from the outset in a learning environment (ideally, for all relevant processes)

Goal is a hierarchical system that integrates data and models (and can also be used to design observing systems)



Climate goals and computing the future of clouds

Tapio Schneider, João Teixeira, Christopher S. Bretherton, Florent Brient, Kyle G. Pressel, Christoph Schär and A. Pier Siebesma

How clouds respond to warming remains the greatest source of uncertainty in climate projections. Improved computational and observational tools can reduce this uncertainty. Here we discuss the need for research focusing on high-resolution atmosphere models and the representation of clouds and turbulence within them.

Geophysical Research Letters

FRONTIER ARTICLE
10.1002/2017GL076101

Key Points:
- Earth system models (ESMs) and their parameterization schemes can be radically improved by data assimilation and machine learning
- ESMs can integrate and learn from global observations from space and from local high-resolution simulations
- Ensemble Kalman inversion and Markov chain Monte Carlo methods show promise as learning algorithms for ESMs

Earth System Modeling 2.0: A Blueprint for Models That Learn From Observations and Targeted High-Resolution Simulations

Tapio Schneider[1,2], Shiwei Lan[1], Andrew Stuart[1], and João Teixeira[2]

[1]California Institute of Technology, Pasadena, CA, USA, [2]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA

Abstract Climate projections continue to be marred by large uncertainties, which originate in processes that need to be parameterized, such as clouds, convection, and ecosystems. But rapid progress is now

...and computing ...ouds

...on, Florent Brient, Kyle G. Pressel, Christoph Schär

...ce of uncertainty in climate projections. Improved ...rtainty. Here we discuss the need for research ...esentation of clouds and turbulence within them.

CLIMATE MODELLING

A model revolution

Geophys. Res. Lett. http://doi.org/ch6g (2017)

While considerable progress has been made in our understanding of the climate system, projections of the future remain highly uncertain. Such relatively low confidence stems, in part, from uncertainties in the parameterization schemes of Earth system models (ESMs) — approximations of unresolved small-scale processes — including, for example, cloud dynamics. Tapio Schneider and colleagues from the California Institute of Technology, USA, envision a revolution in Earth system modelling using data as... ...on and machine learning to impro... ...ation schemes.

...search Letters

...A Blueprint for Models

Next-Generation Climate Models Could Learn, Improve on the Fly

Scientists propose development of new models that use machine learning techniques to reduce uncertainties in climate predictions.

Summary

- Unified parameterization based on EDMF framework holds promise
    - can be made scale-aware
    - can adapt to data through adjusting parameters and number of plumes
    - captures convective life cycle
- Parameterization can be improved through data assimilation and ML from LES and observations
- This approach can (and should) be extended to other process models, ideally in a automatically learning ESM