Prof. Sebastian Schemm, IAC

Prof. Siddhartha Mishra, SAM

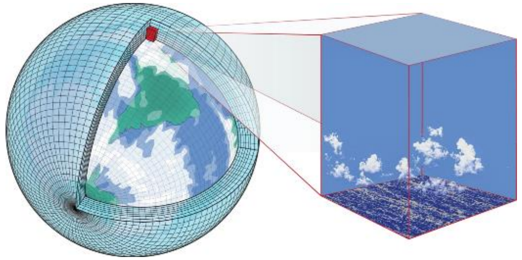
Prof. Ulrike Lohmann, IAC

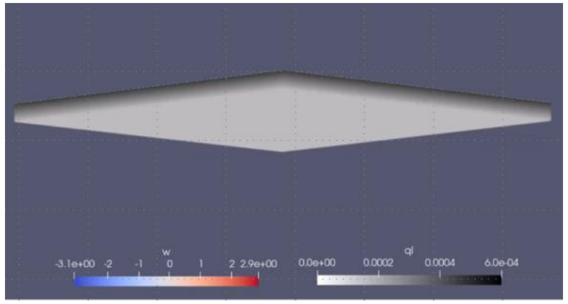
Dana Grund
IAC and SAM

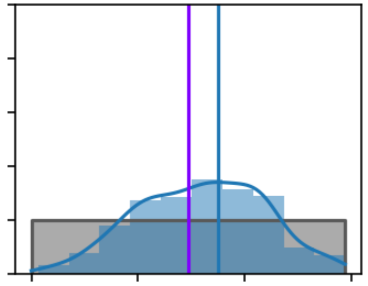
SAM = Seminar for Applied Mathematics
IAC = Institute for Atmospheric and Climate Science

Calibration of High-Resolution Atmospheric Models

Calibration of High-Resolution Atmospheric Models







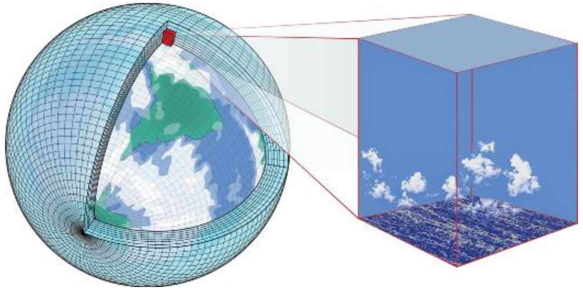
1. Models

2. Clouds & Measurements

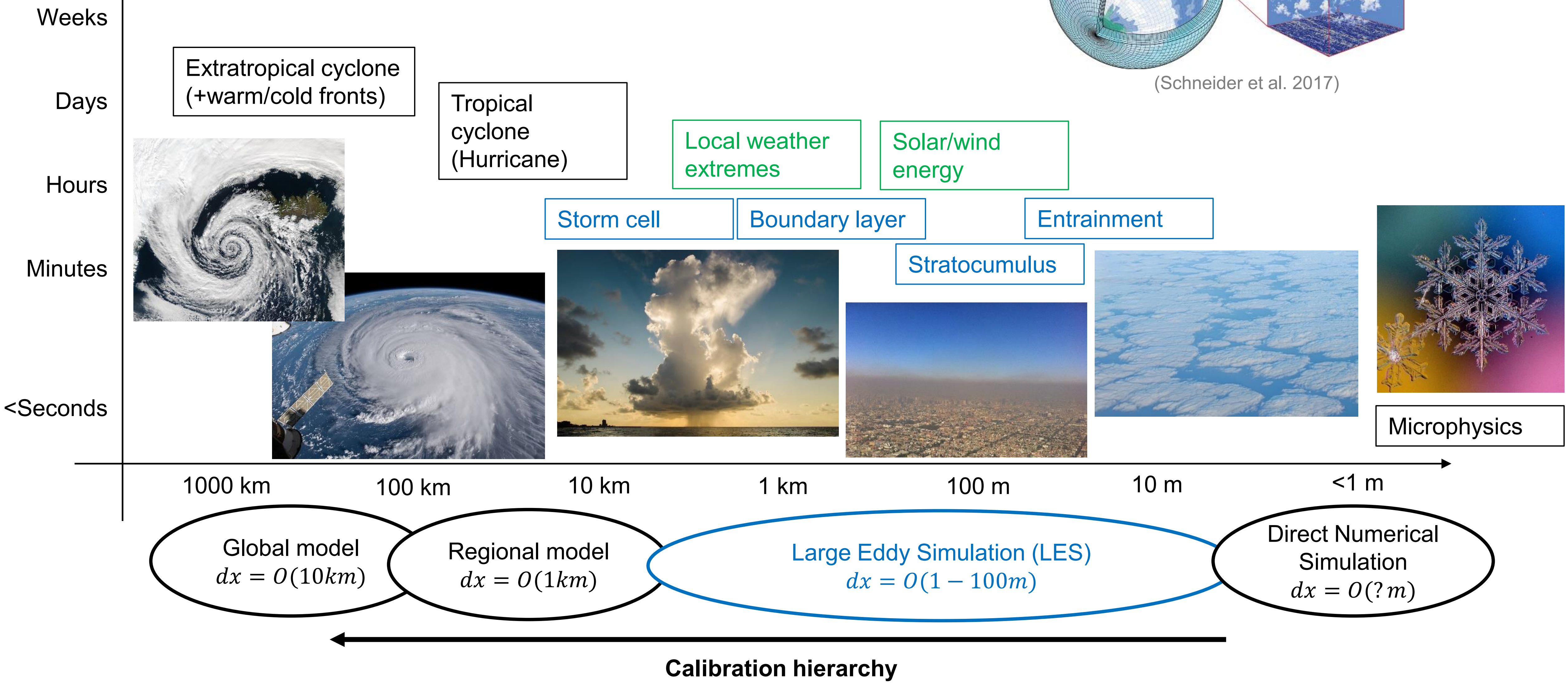
3. Simulation

4. Calibration

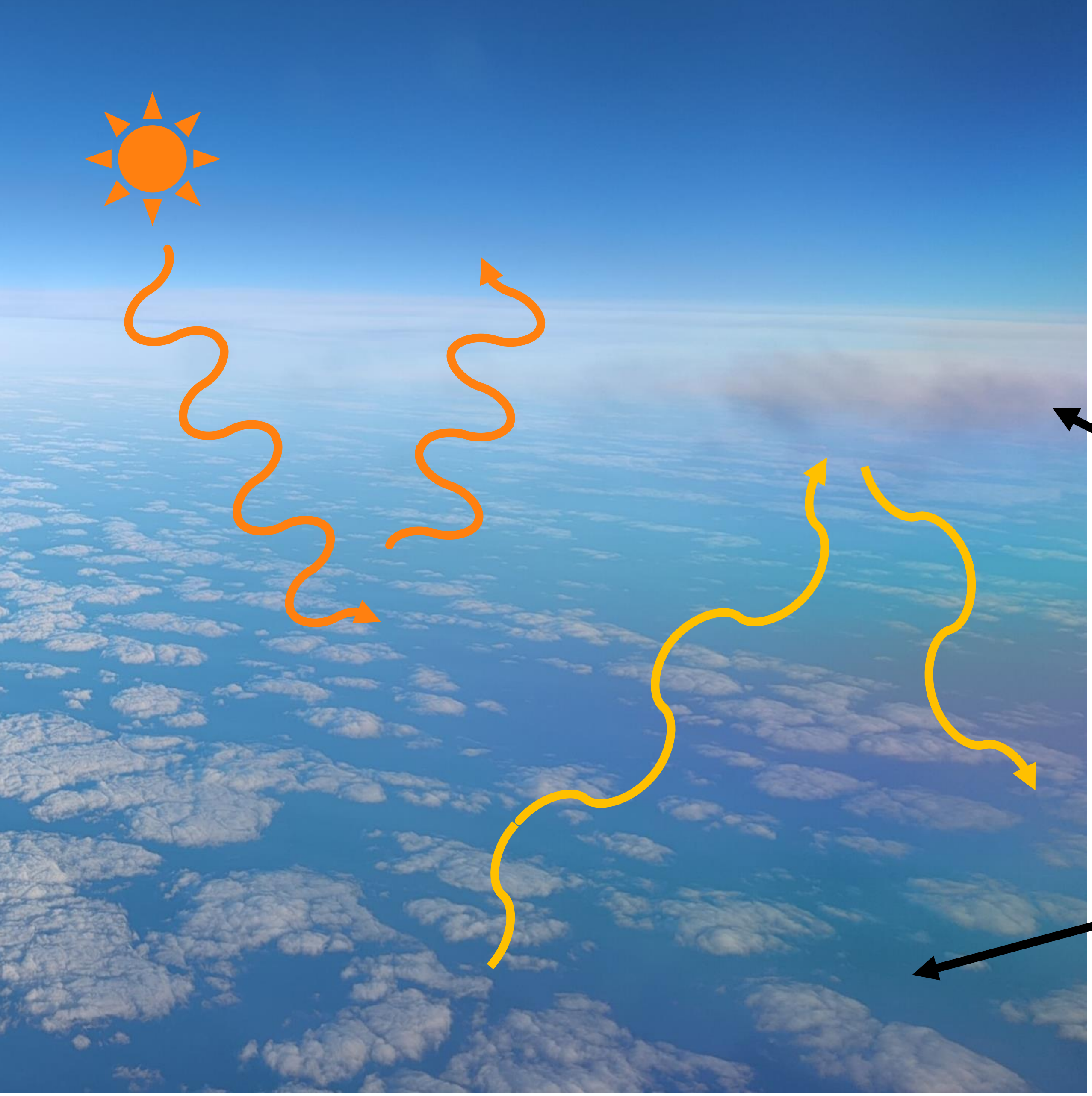
Models of the atmosphere



(Schneider et al. 2017)



Images: Wikipedia



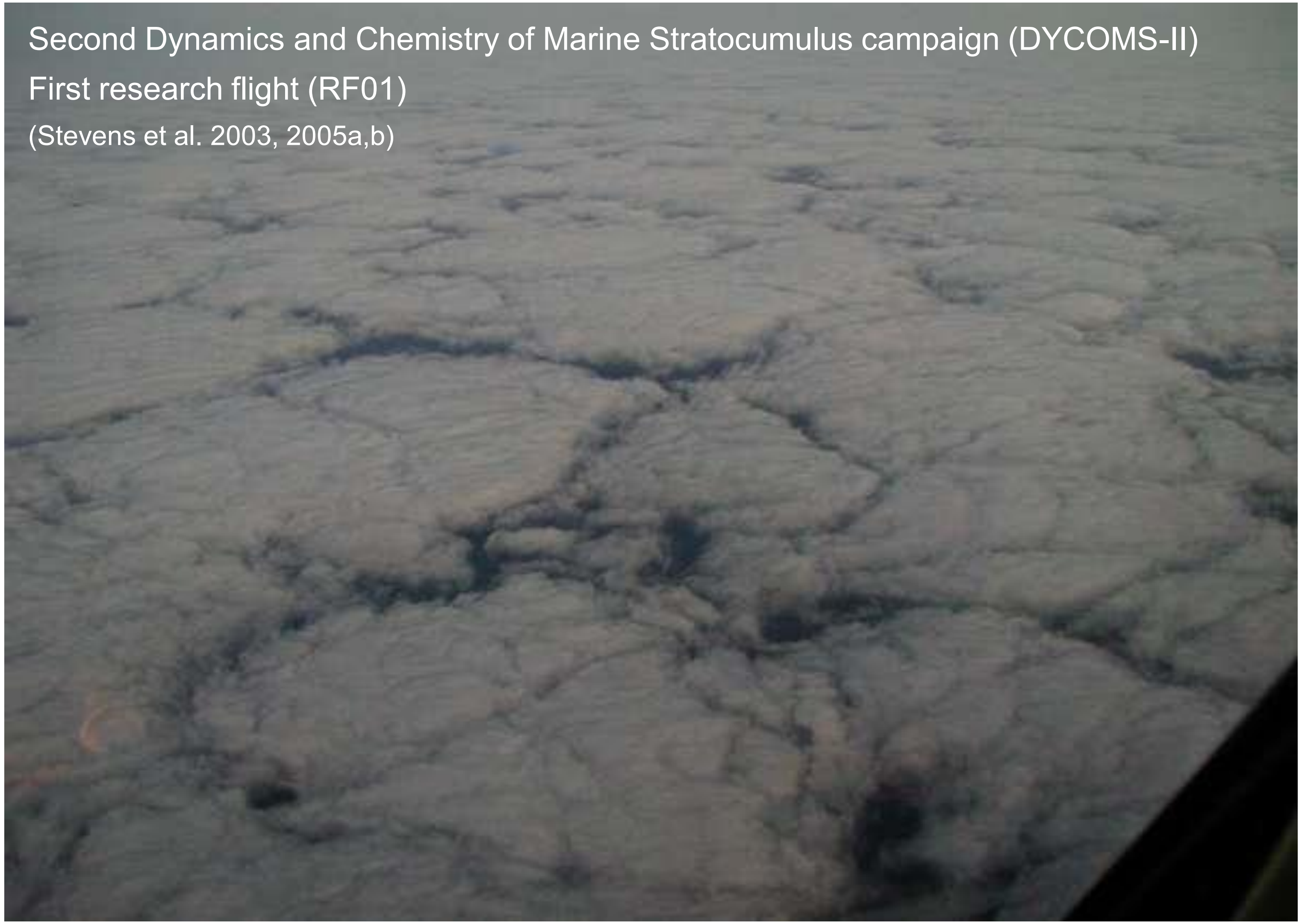
Low clouds are „good“ *

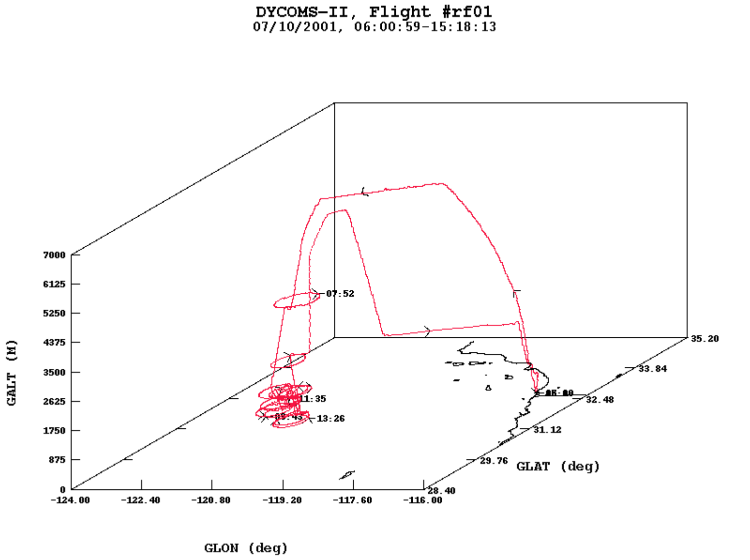
Contrails (high)

Marine stratocumulus clouds (low)

*of course, there are more details to this story.

Marine stratocumulus clouds

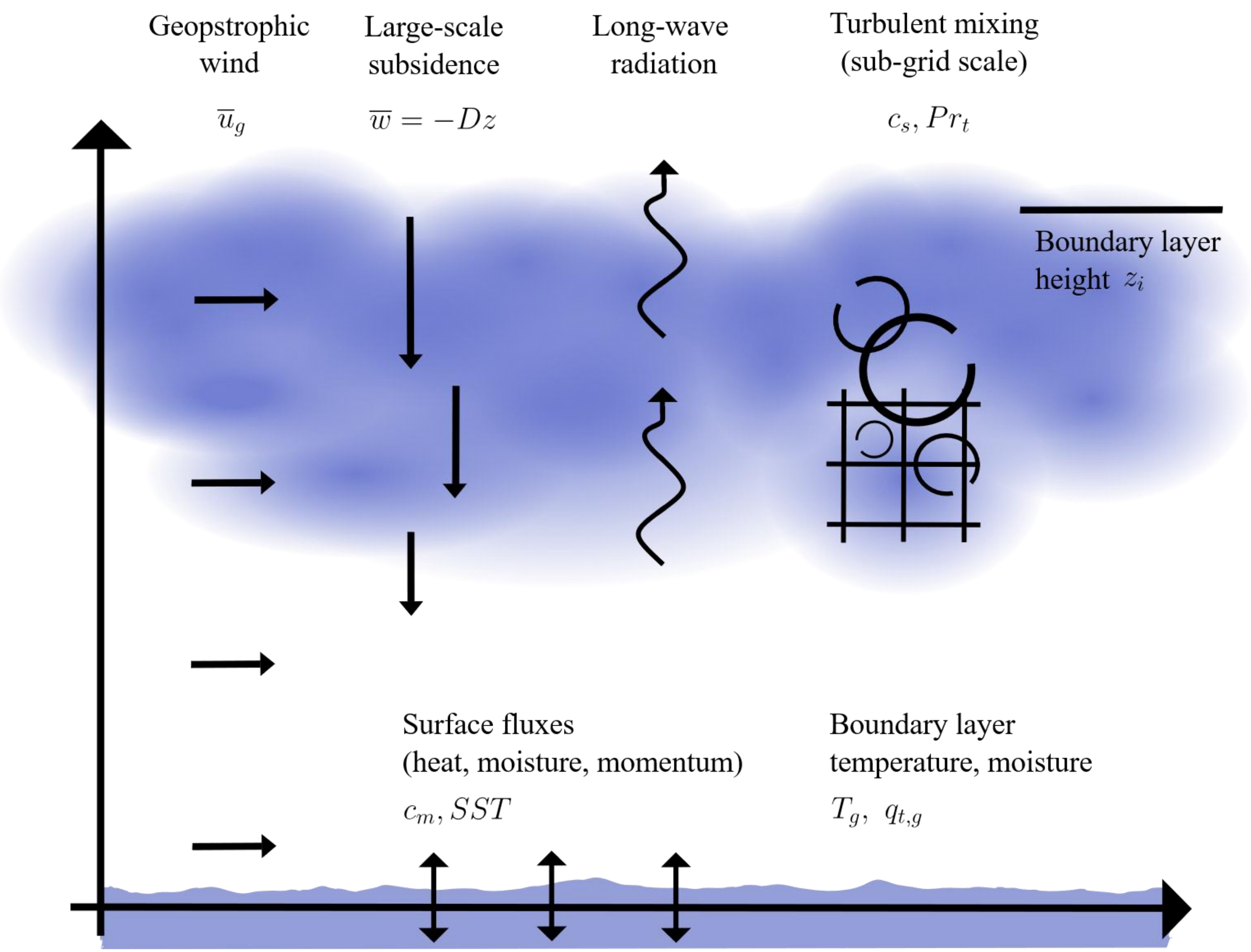




Guiding questions

- Cloud properties?
- Stability over time?
- Sensitivities?

Marine stratocumulus clouds



Uncertain parameters θ

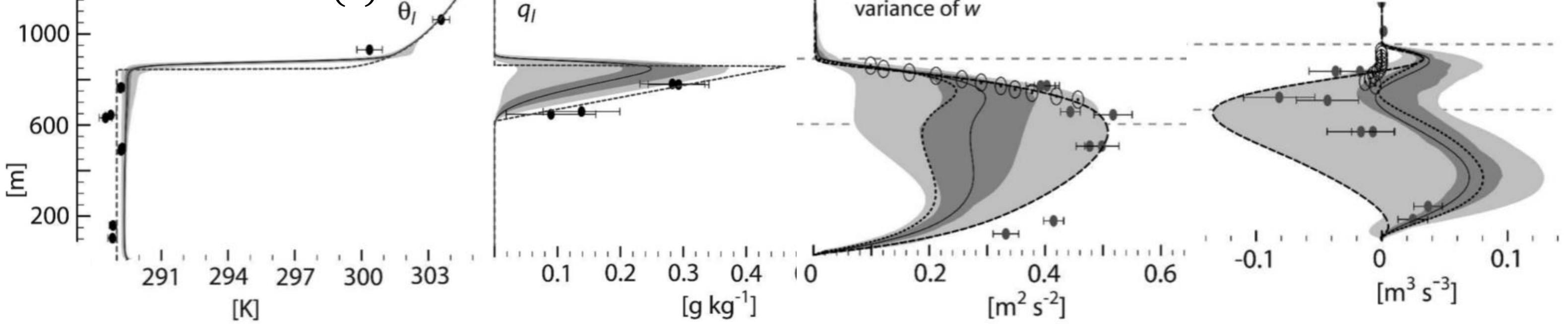
T_g	Initial BL temperature
$q_{t,g}$	Initial BL total humidity
z_i	Initial BL height
u_g	Forcing: Horizontal wind (geostrophic)
D	Forcing: Vertical wind (diverence ~ subsidence)
SST	Forcing: Sea surface temperature
c_m	Aerodyn. bulk coefficient (surface exchange)
c_s	Smagorinsky constant (turbulent viscosity)
Pr_t	Turbulent Prandtl number (turbulent diffusion)

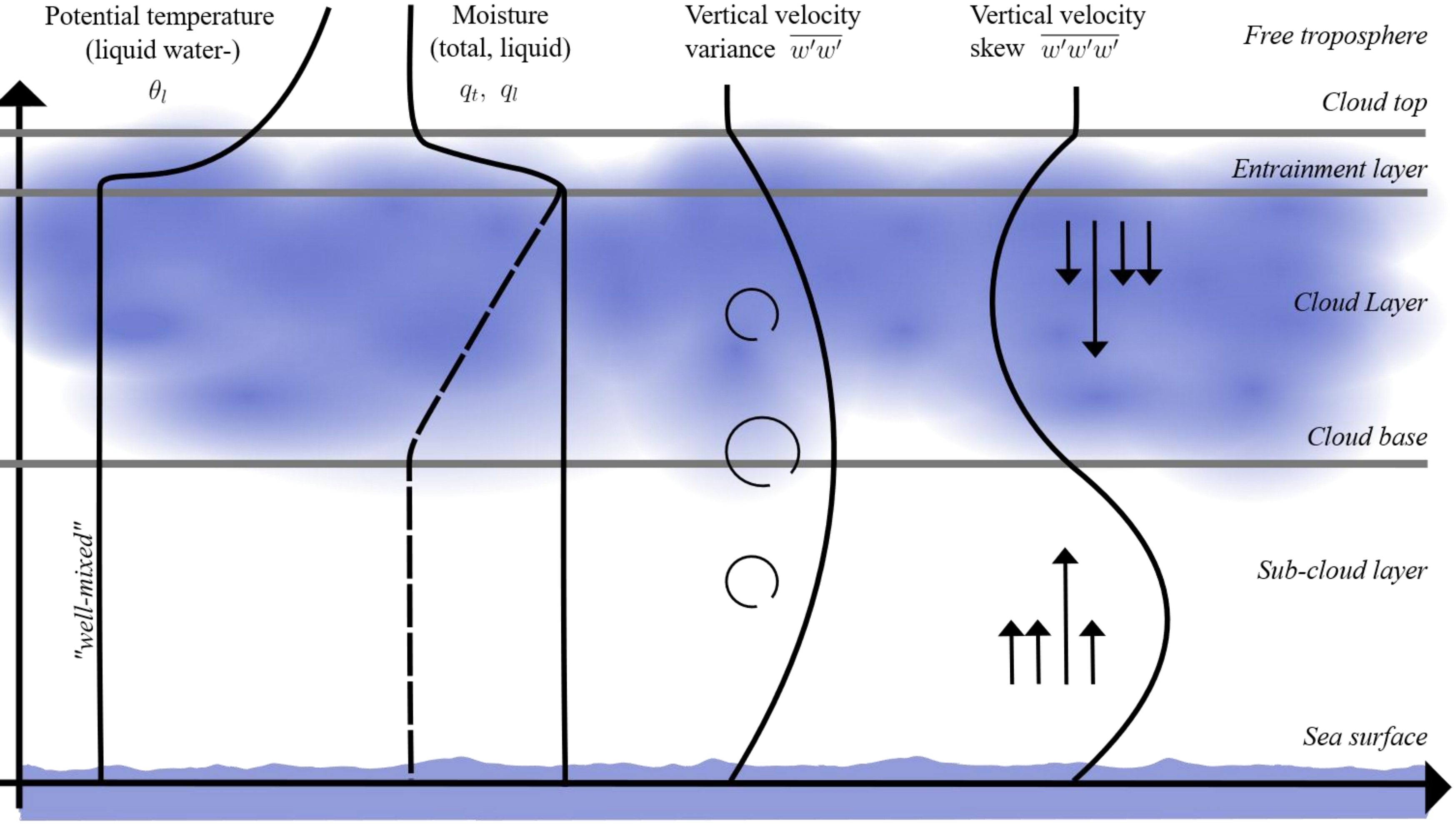
Guiding questions

- Cloud properties?
- Stability over time?
- Sensitivities?

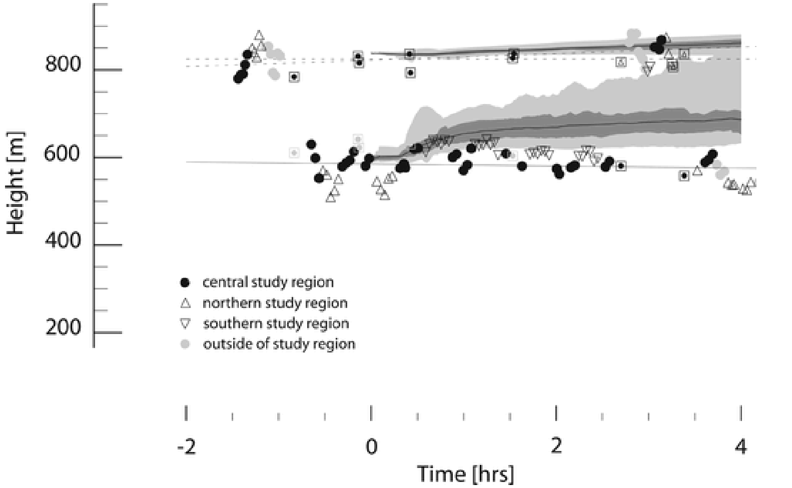
Meas. Data

Profiles $d(z)$





Timeseries $d(t)$
Cloud top, base heights

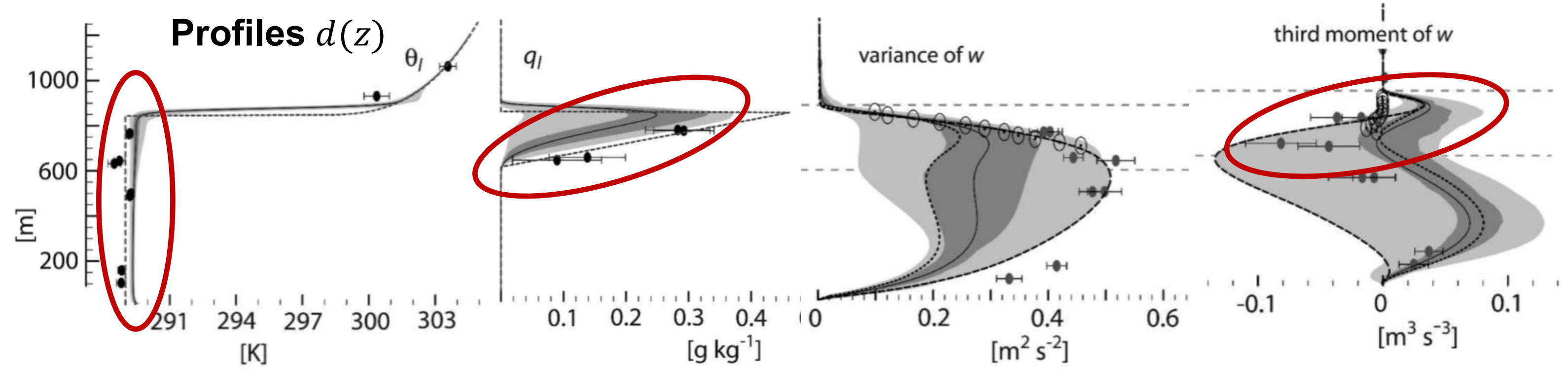


Data d

$d(z)$ Profiles: Horizontal and time averages at measurement heights

$d(t)$ Timeseries: Horizontal averages (Linearly approximated)

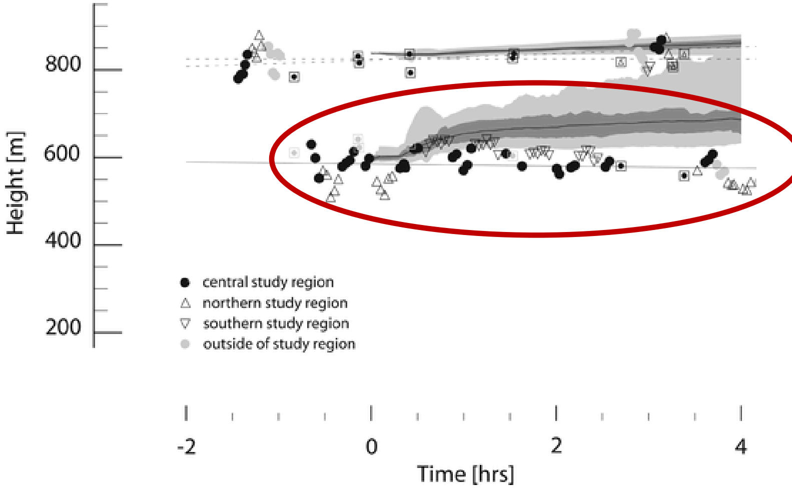
Profiles $d(z)$

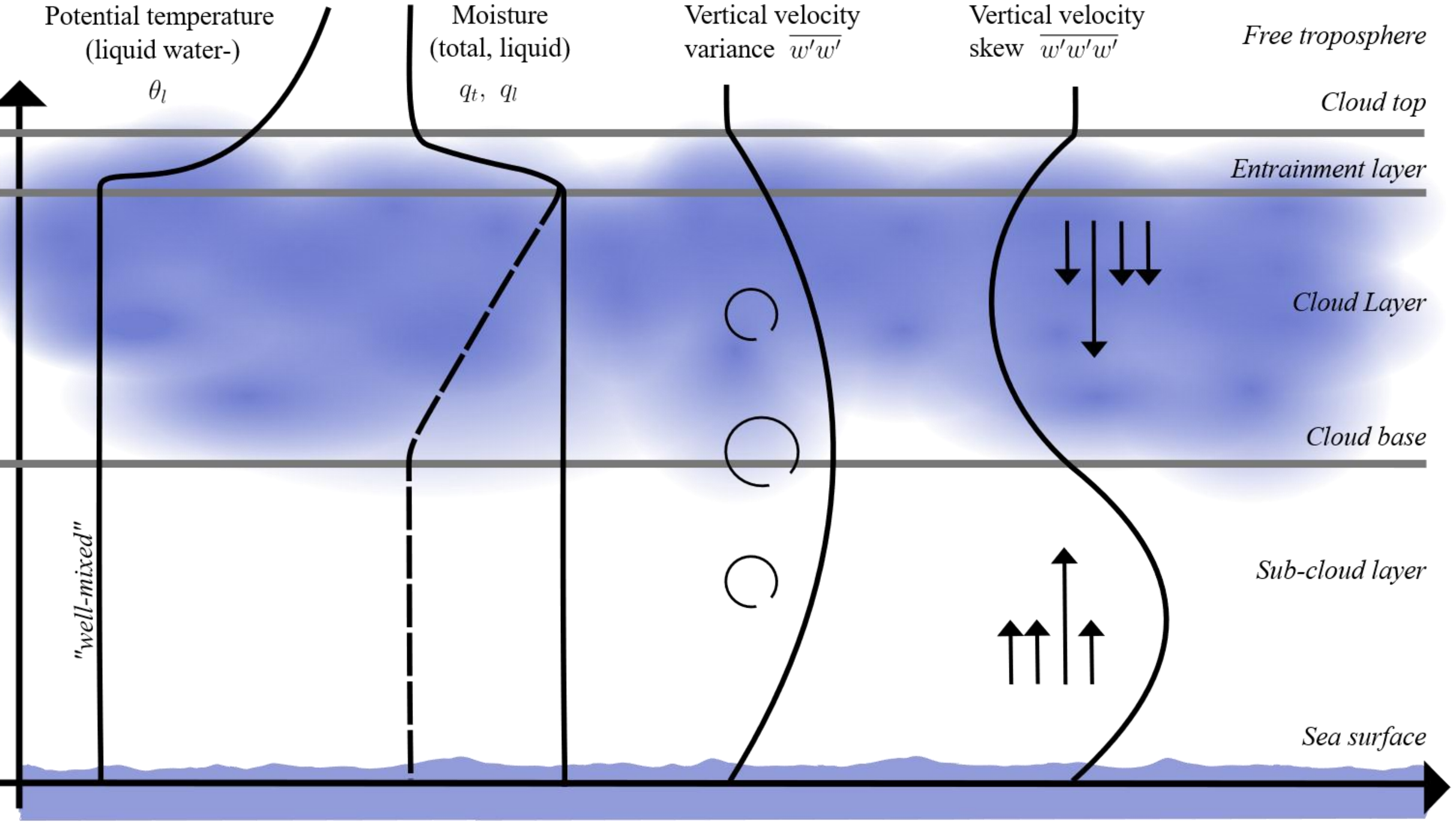


Meas. Data

Timeseries $d(t)$

Cloud top, base heights





LES intercomparison

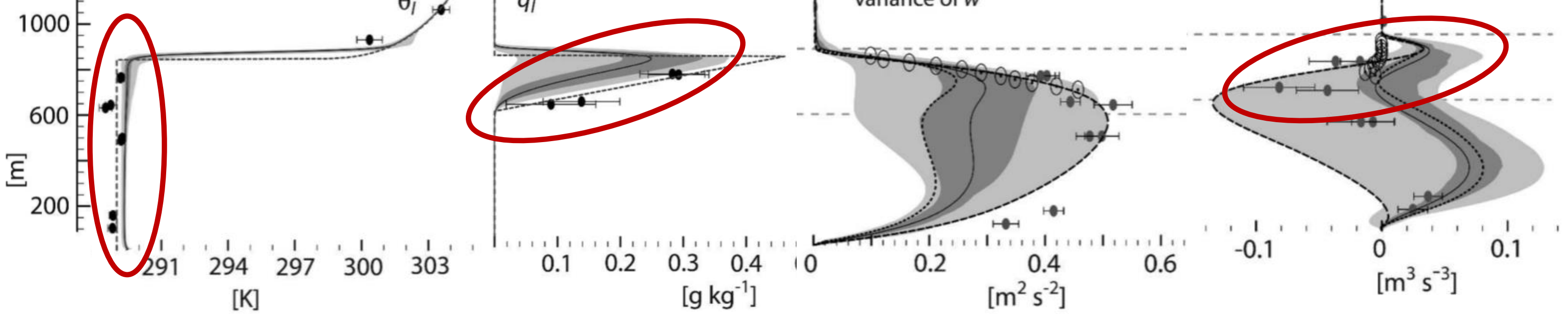
(Stevens et al. 2005)

Biases: θ_l ↑, q_l ↓

Missed: Cloud base sinking

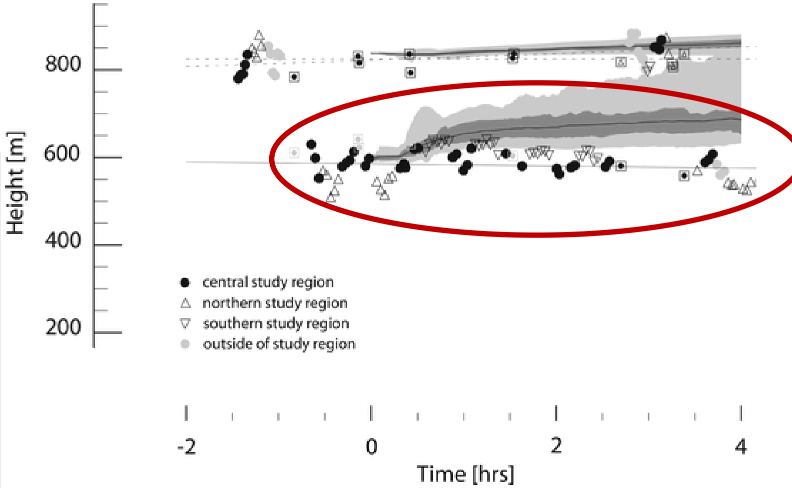
Missed: negative skew

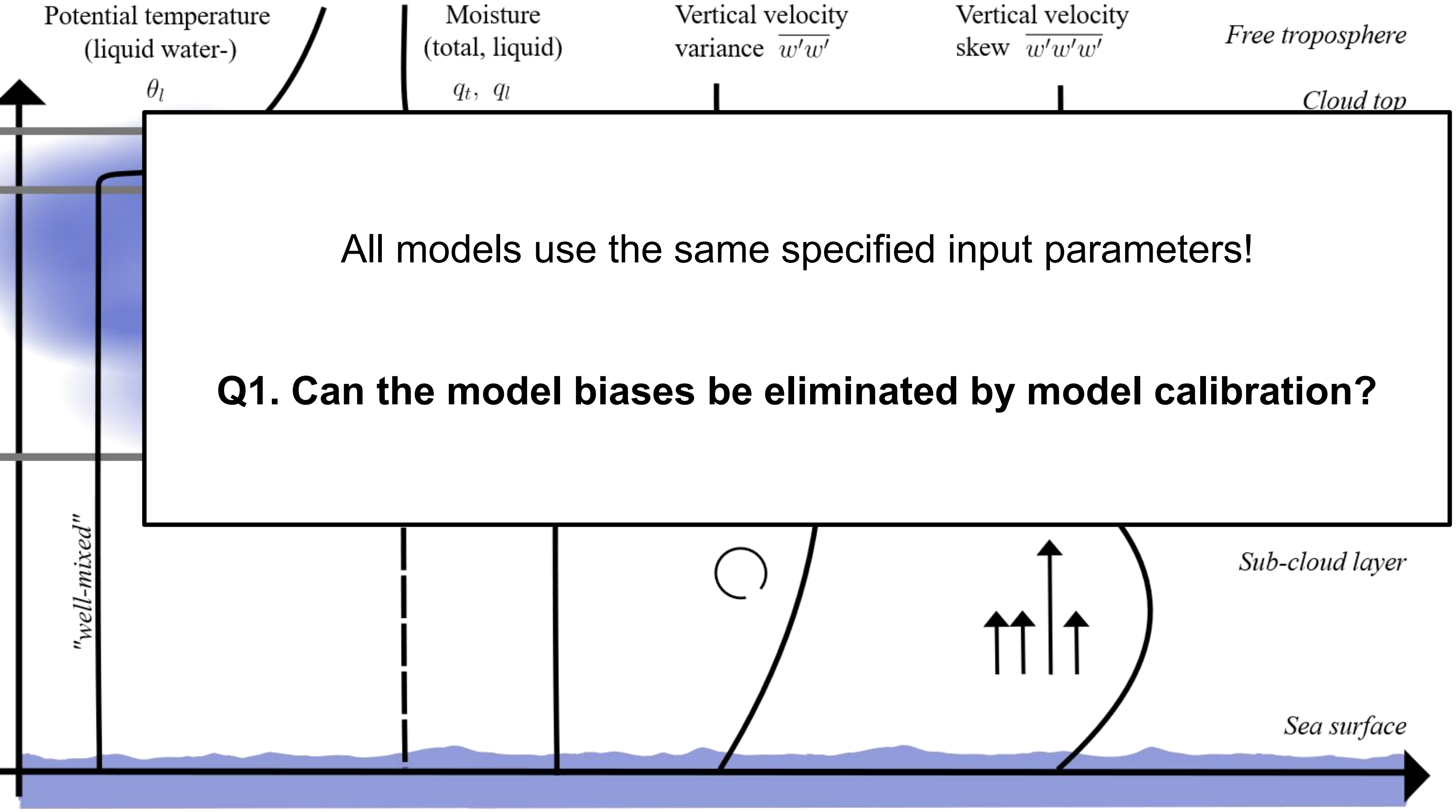
Profiles $d(z)$



Meas. Data

Timeseries $d(t)$
Cloud top, base heights





All models use the same specified input parameters!

Q1. Can the model biases be eliminated by model calibration?

LES intercomparison
(Stevens et al. 2005)

Biases: $\theta_l \uparrow$, $q_l \downarrow$

Missed: Cloud base sinking

Missed: negative skew

Simulation of marine stratocumulus clouds

Large-eddy simulation in an anelastic framework with closed water and entropy balances (Pressel et al. 2015) „PyCLES“ = Python Cloud LES

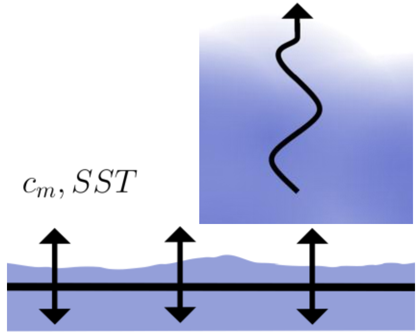
Kyle G. Pressel[1], Colleen M. Kaul[1], Tapio Schneider[1,2], Zhihong Tan[1,2], and Siddhartha Mishra[3]

[1]Department of Earth Sciences, ETH Zürich, Zürich, Switzerland, [2]California Institute of Technology, Pasadena, California, USA, [3]Seminar for Applied Mathematics, ETH Zürich, Zürich, Switzerland

Numerics: Grid-scale

- Common in the field: central differences
  - *Dispersive, needs numerical diffusion to stabilize*
- PyCLES: Finite volumes 5th order WENO
  - *Dissipative, does NOT need extra diffusion*
- Radiation: RRTM
- Surface fluxes: constant (in space and time)

c_m, SST



Numerics: Sub-grid scale (SGS)

- Smagorinsky scheme (simplest!)
  - SGS stresses

$$\tau_{ij} = -2\nu_t S_{ij} \qquad S_{ij} = \frac{1}{2}\left(\frac{\partial u_j}{\partial x_i} + \frac{\partial u_i}{\partial x_j}\right)$$

  - SGS viscosity, diffusivity

$$\nu_t = (c_S \Delta)^2 f_B |S| \qquad D_t = \nu_t / \mathrm{Pr}_t$$

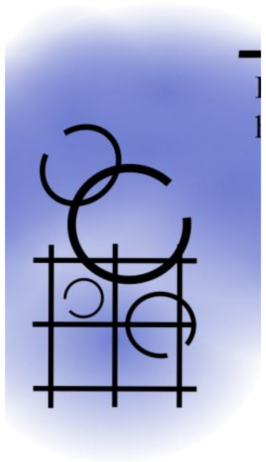
  - Flow-dependence through moist buoyancy frequency N

$$f_B = \begin{cases} 1 & \text{for} \quad N^2 \leq 0, \\ \max\left[0,\ 1 - N^2/(\mathrm{Pr}_t |S|^2)\right]^{1/2} & \text{for} \quad N^2 > 0. \end{cases}$$

→ SGS schemes add numerical diffusion.

Turbulent mixing
(sub-grid scale)

c_s, Pr_t



Simulation of marine stratocumulus clouds

Large-eddy simulation in an anelastic framework with closed water and entropy balances (Pressel et al. 2015) „PyCLES“ = Python Cloud LES

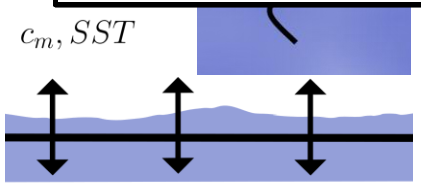
Kyle G. Pressel[1], Colleen M. Kaul[1], Tapio Schneider[1,2], Zhihong Tan[1,2], and Siddhartha Mishra[3]

[1]Department of Earth Sciences, ETH Zürich, Zürich, Switzerland, [2]California Institute of Technology, Pasadena, California, USA, [3]Seminar for Applied Mathematics, ETH Zürich, Zürich, Switzerland

Numerics: Grid-scale

- Common in the field: central differe
  - *Dispersive, needs numerical diffusio*
- PyCLES: Finite volumes 5th order
  - *Dissipative, does NOT need extra di*
- Radiation: RRTM
- Surface fluxes: constant (in space and time)

c_m, SST



Numerics: Sub-grid scale (SGS)

- Smagorinsky scheme (simplest!)

PyCLES works best without SGS scheme,
i.e. $c_s = 0$ (“implicit LES”).
(Pressel et al. 2017)

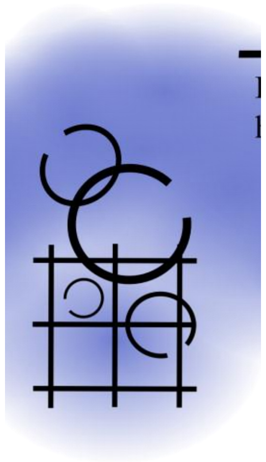
Q2. How do different numerics affect the calibration?

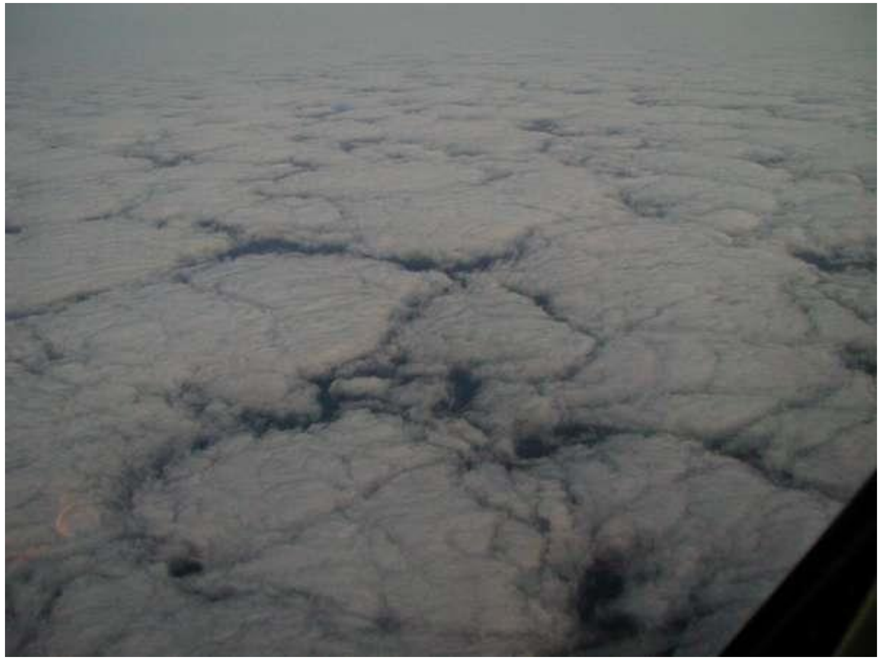
$\max\left[0,\ 1-N^2/(\mathrm{Pr}_t|S|^2)\right]$ for $N^2 > 0$.

→ SGS schemes add numerical diffusion.

Turbulent mixing (sub-grid scale)

c_s, Pr_t

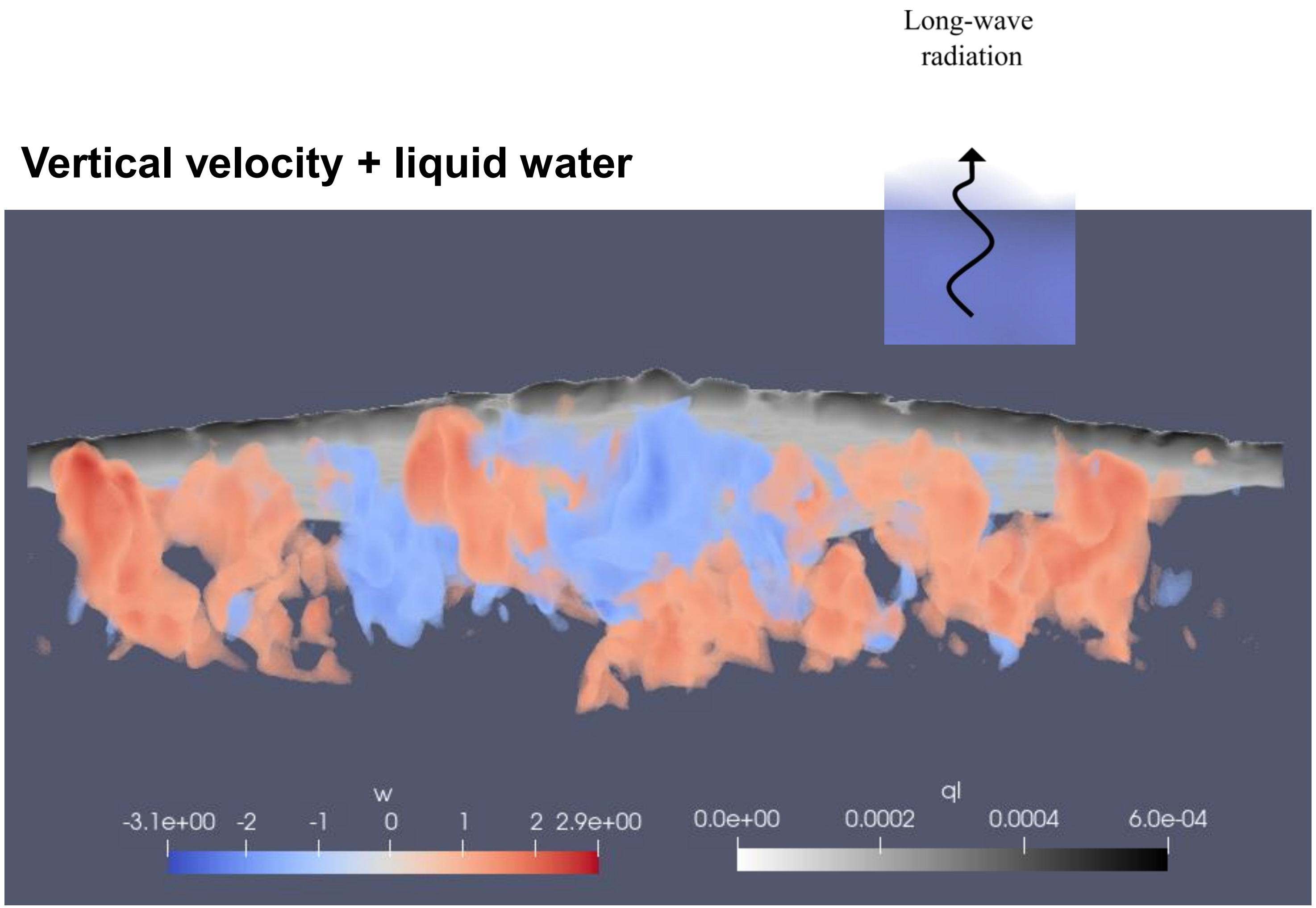




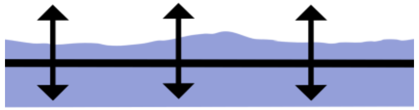
Resolution
- $dx = 35\,m,\ dz = 5\,m$
- $nx^2 \times nz = 96^2 \times 300 = 2.8\text{M}$ pixels
- $T = 4h,\ dt \approx 1 - 10s$

Ressources
- Memory: 166M per time step
- Compute: 4:15h on 16 cores = 68 CPUh

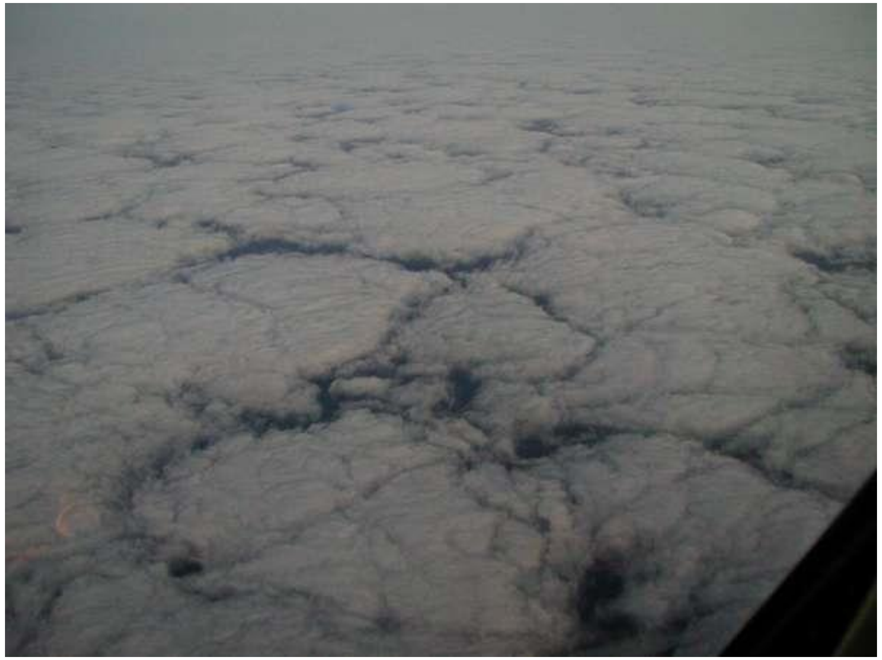
Vertical velocity + liquid water

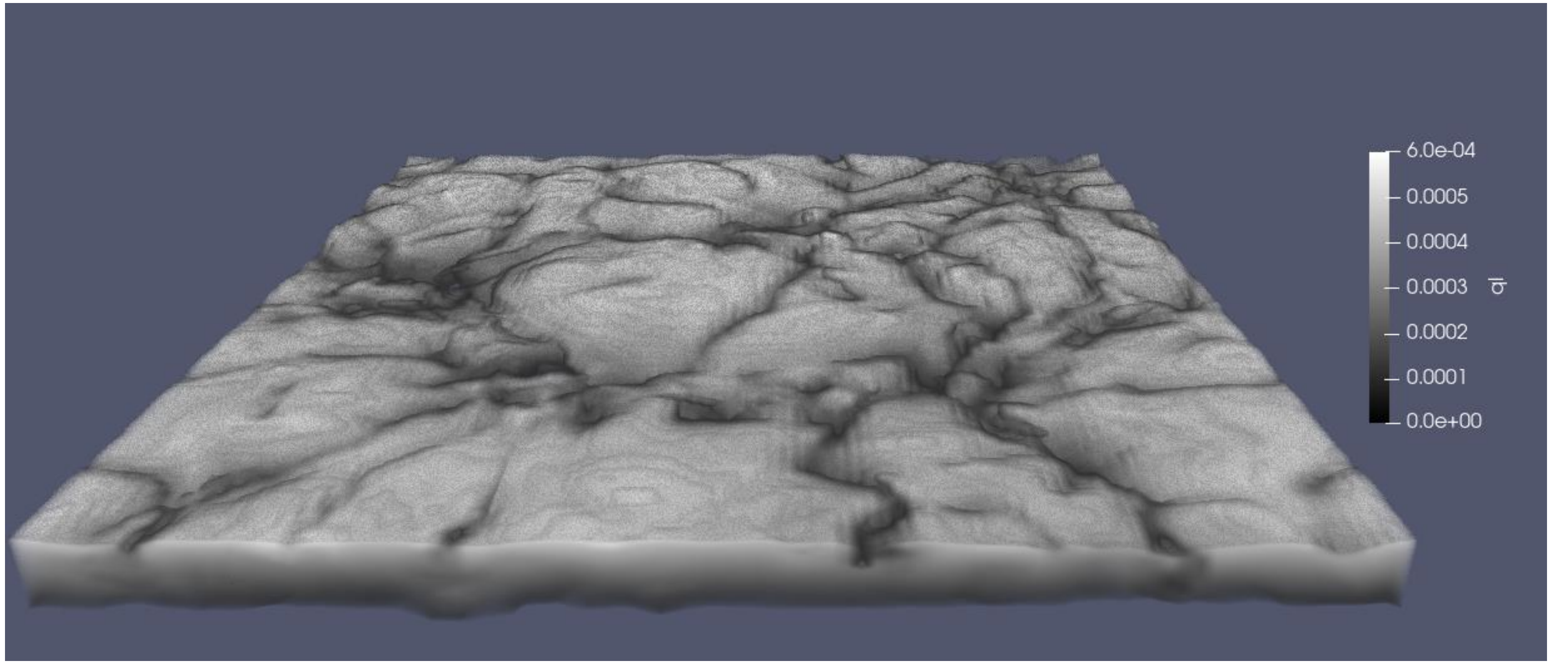


Surface fluxes
(heat, moisture, momentum)
c_m, SST



Liquid water





Resolution

- $\mathrm{d}x = 35\, m,\ dz = 5\, m$
- $nx^2 \times nz = 96^2 \times 300 = 2.8\mathrm{M}$ pixels
- $T = 4h,\ dt \approx 1 - 10s$

Ressources

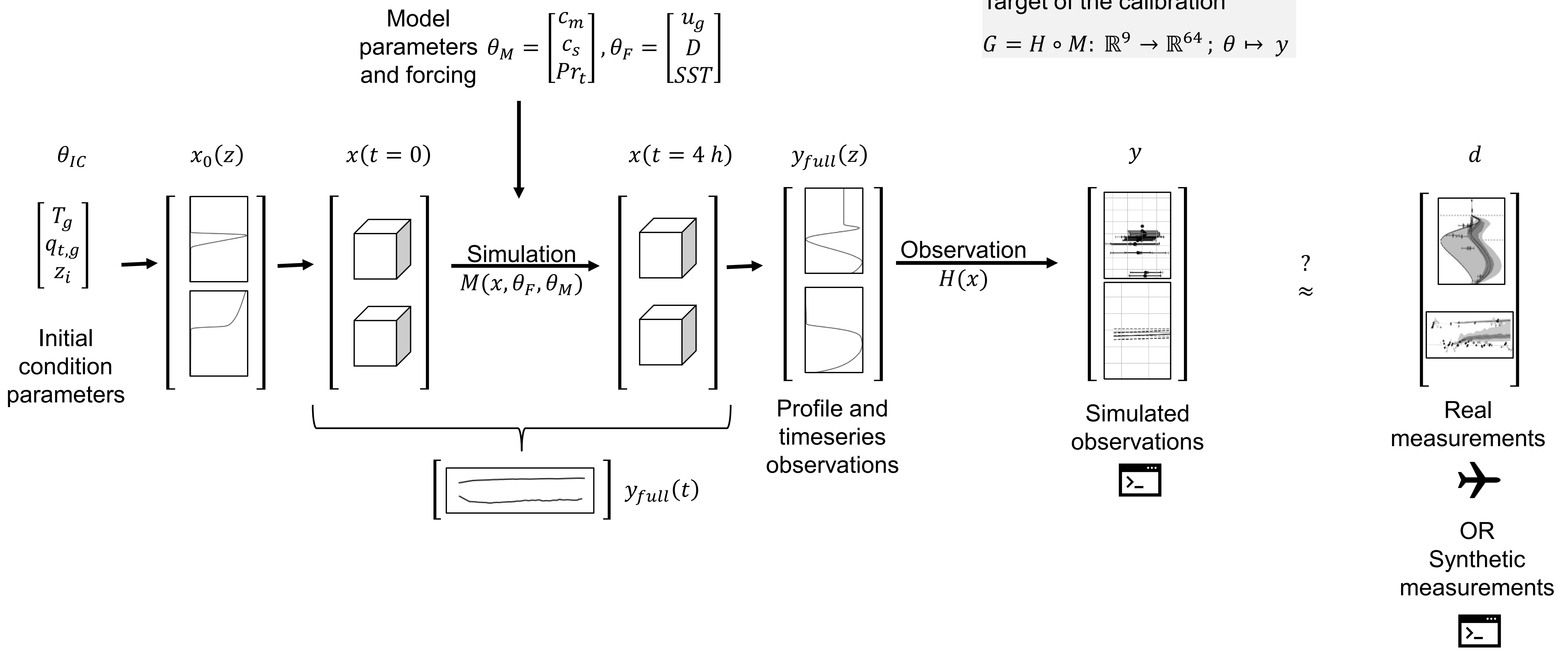
- Memory: 166M per time step
- Compute: 4:15h on 16 cores = 68 CPUh

Data assimilation setup

Parameter-to-data map

Target of the calibration

$G = H \circ M\colon \mathbb{R}^9 \to \mathbb{R}^{64};\ \theta \mapsto y$



Data assimilation setup

Target: parameter-to-data map $\mathcal{G}: \theta \mapsto y \approx d$

Bayes law: $p(\theta|d) \propto p(d|\theta)\, p(\theta)$

posterior $\propto$ *likelihood * prior*

analysis $\propto$ *observations * forecast*

Assume $\mathcal{N}$: $p(\theta) = \mathcal{N}(\mu_\theta, C_\theta),\ p(d|\theta) = \mathcal{N}(\mu_j, C_\epsilon)$

DA Method: EnKF (perturbed observations)

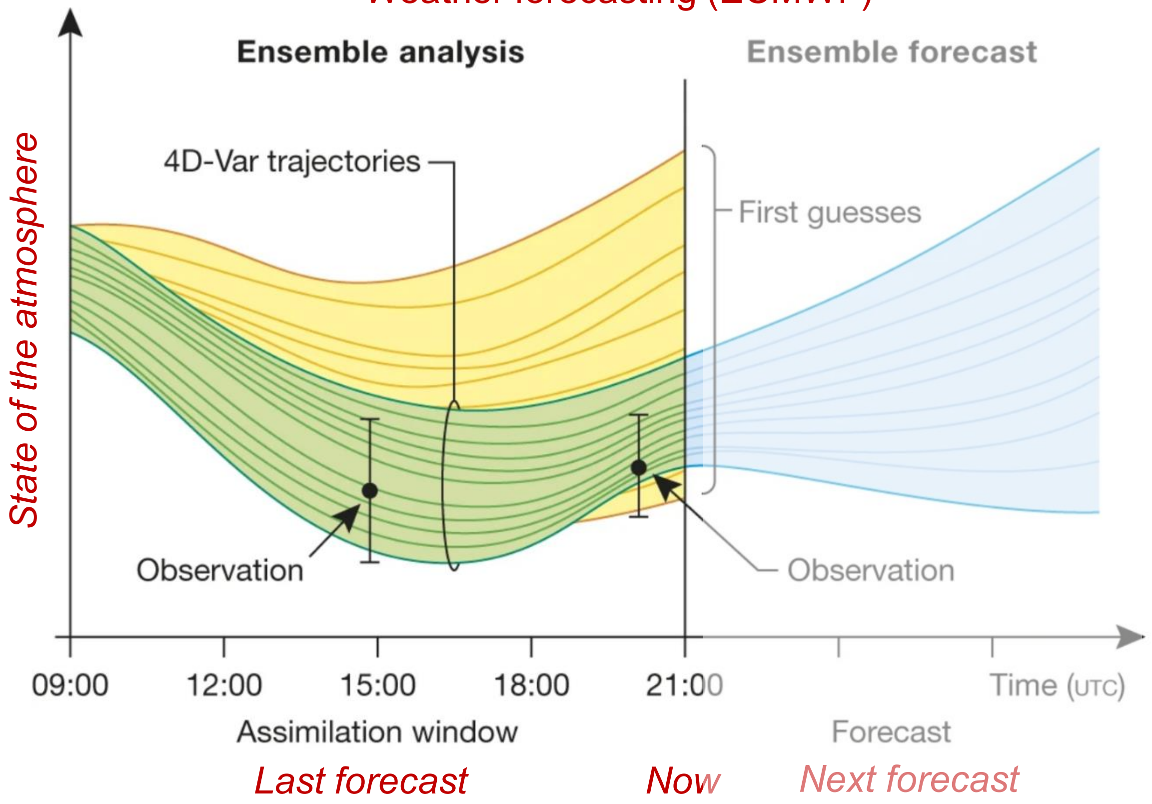
[ETKF (sqrt) gave similar results]

DAPPER package (credit to P. Raanes!)

Single-event calibration ≠ Weather forecasting

- θ = parameters (not state)
- No time cycling („smoothing“, not „filtering“)

Weather forecasting (ECMWF)



Bayesian Setup

Bayes law: $p(\theta|d) \propto p(d|\theta)\, p(\theta)$

*posterior ∝ likelihood * prior*

*analysis ∝ observations * forecast*

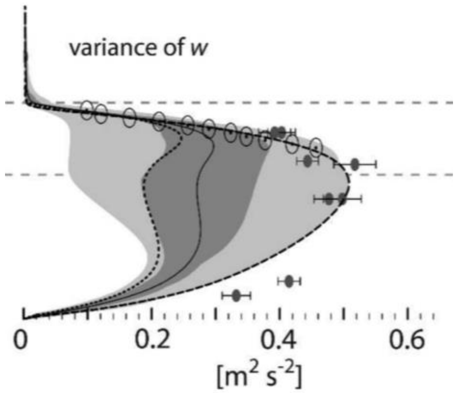
Assume $\mathcal{N}$: $p(\theta) = \mathcal{N}(\mu_\theta, C_\theta),\ p(d|\theta) = \mathcal{N}(\mu_j, C_\epsilon)$

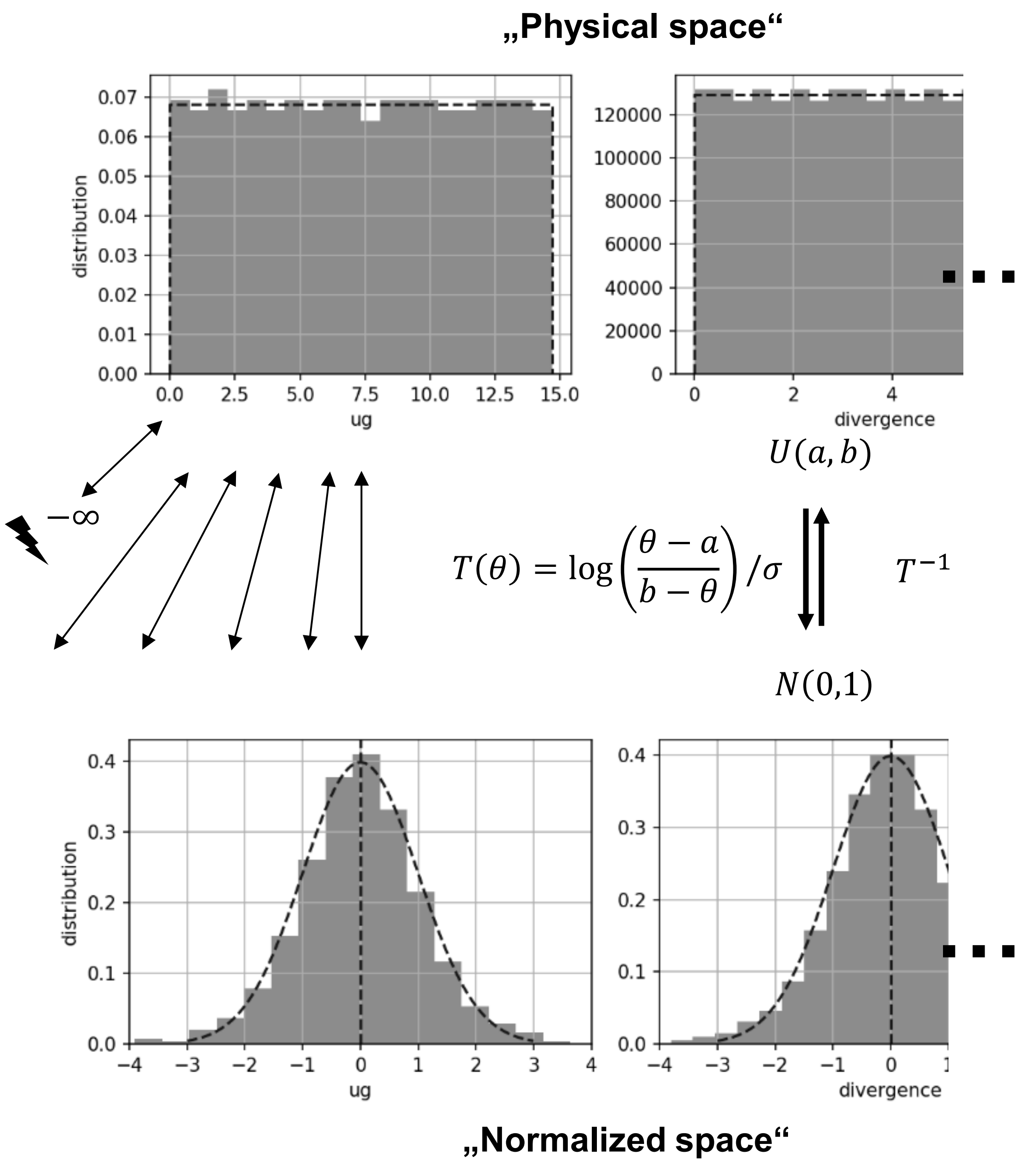
→ Data: σ_j^2 =meas. Error

$C_\epsilon = diag(\sigma_j^2)$: no model error, no cross-correlations

→ Parameters: Transform uniform bounds $U(a,b)$ to $\mathcal{N}(0,1)$: $C_\theta = I$

Bounds around nature parameters from the intercomparison

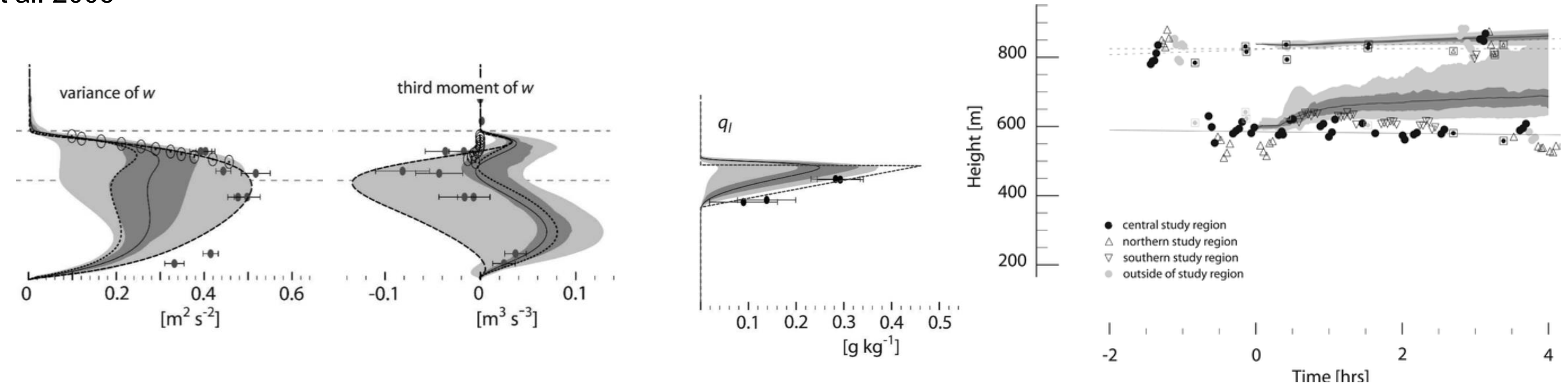




Prior ensemble

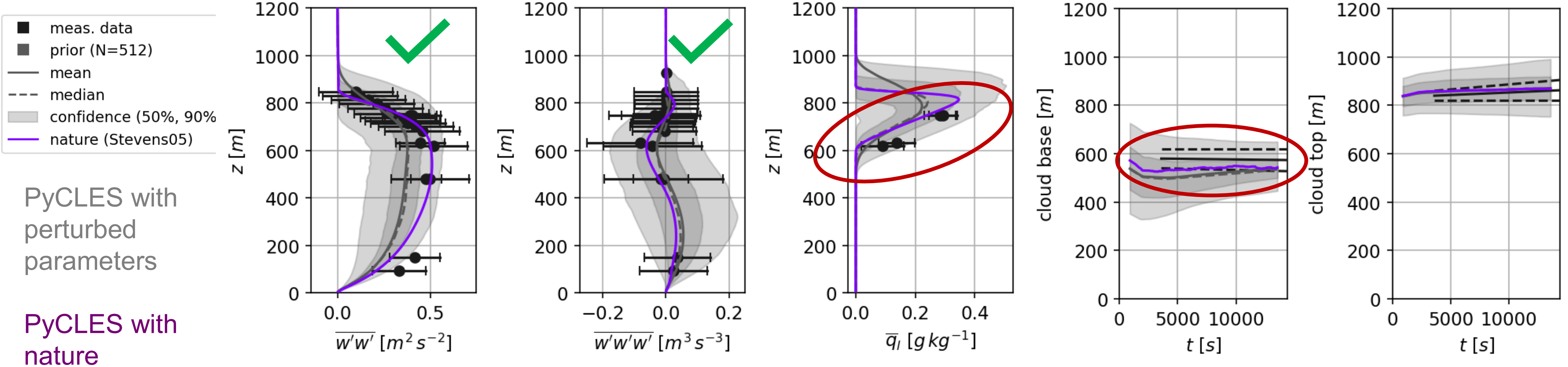
PPE = Perturbed Physics Ensemble, N=512
Compared to Stevens et al. 2005

Other LES with nature parameters



PyCLES with perturbed parameters

PyCLES with nature parameters



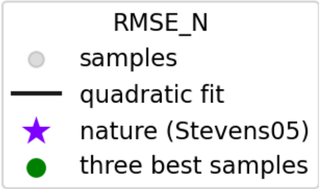
w profiles: good performance

Moisture and cloud base: similar bias to Stevens05

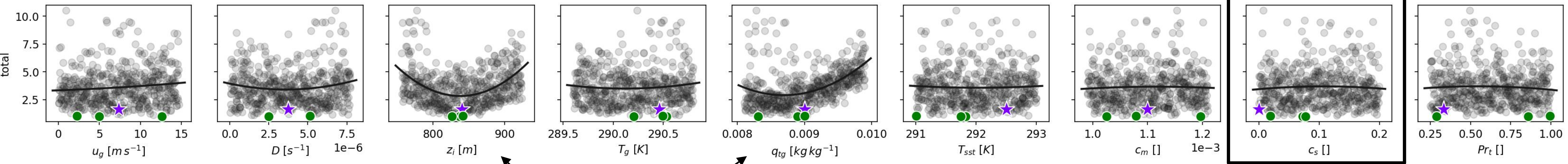
Prior ensemble

Gaussian RMSE, N=512

$$RMSE_{\mathcal{N}}(\theta) = \left(\frac{1}{M}\sum_{j=1}^{M}\left(d_j - y_j(\theta)\right)^2/\sigma_j^2\right)^{1/2}$$



Gaussian RMSE over all observations

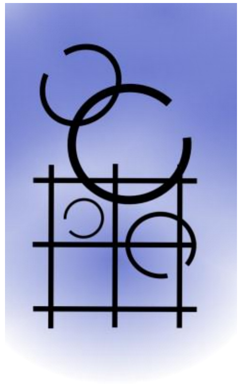


Clear minimum only for initial moisture and boundary layer height

No sign of correlation for the Smagorinsky constant.

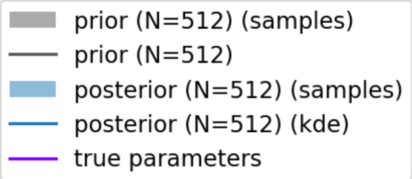
Uncertain parameters θ

T_g	Initial BL temperature
$q_{t,g}$	Initial BL total humidity
z_i	Initial BL height
u_g	Forcing: Horizontal wind (geostrophic)
D	Forcing: Vertical wind (diverence ~ subsidence)
SST	Forcing: Sea surface temperature
c_m	Aerodyn. bulk coefficient (surface exchange)
c_s	Smagorinsky constant (turbulent viscosity)
Pr_t	Turbulent Prandtl number (turbulent diffusion)



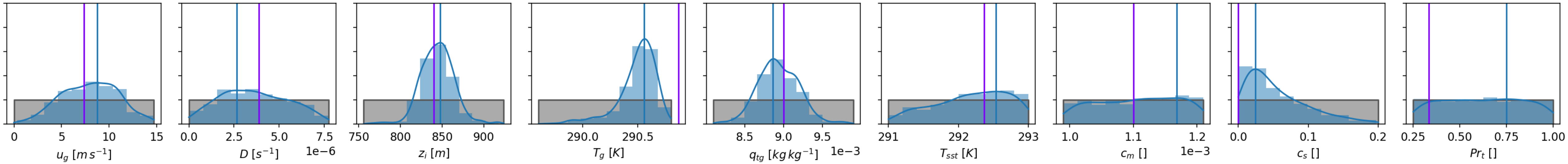
Assimilating synthetic observations

Ensemble Smoother (EnKF with perturbed observations), N=512



Synthetic data: nature sample

≈



Skewed (due to prior transform T)

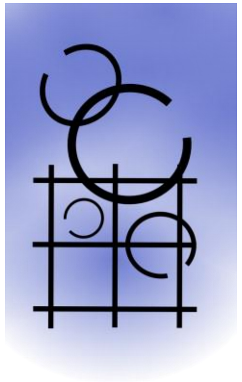
Not identifiable

Identifiable ☺

Not identifiable

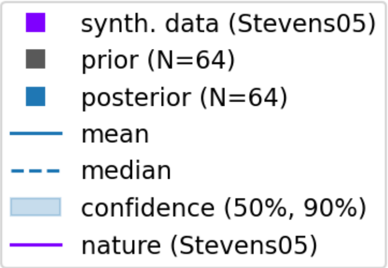
Uncertain parameters θ

T_g	Initial BL temperature
$q_{t,g}$	Initial BL total humidity
z_i	Initial BL height
u_g	Forcing: Horizontal wind (geostrophic)
D	Forcing: Vertical wind (diverence ~ subsidence)
SST	Forcing: Sea surface temperature
c_m	Aerodyn. bulk coefficient (surface exchange)
c_s	Smagorinsky constant (turbulent viscosity)
Pr_t	Turbulent Prandtl number (turbulent diffusion)



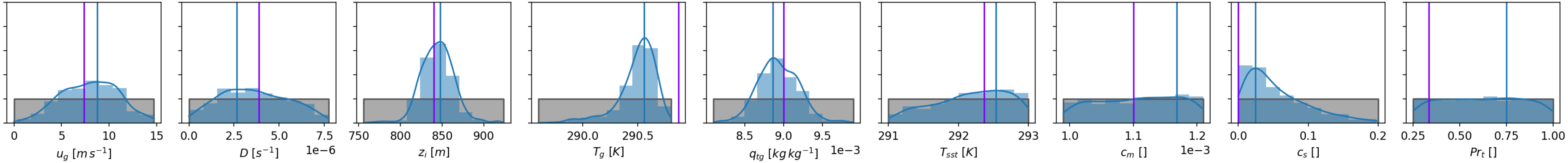
Assimilating synthetic observations

Ensemble Smoother (EnKF with perturbed observations), N=512



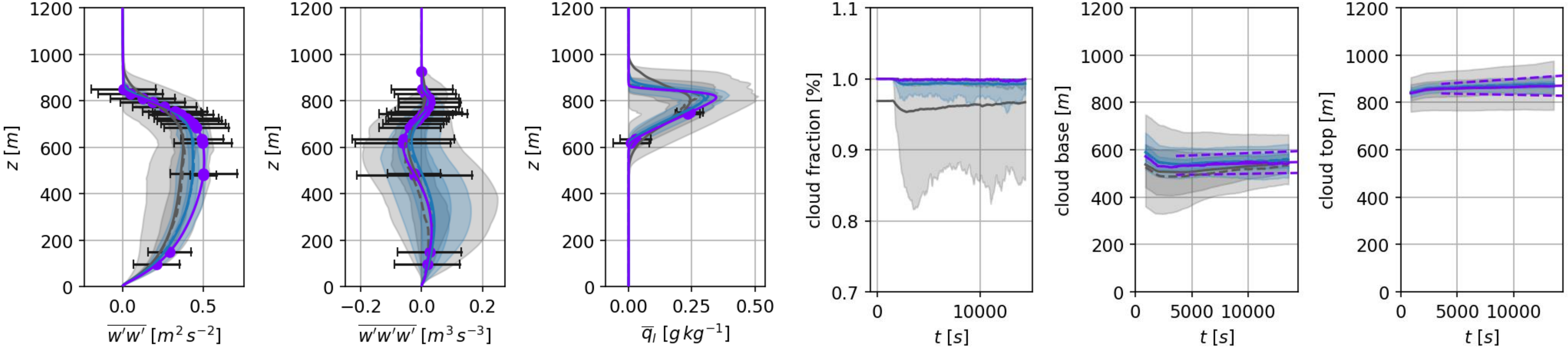
Synthetic data: nature sample

≈



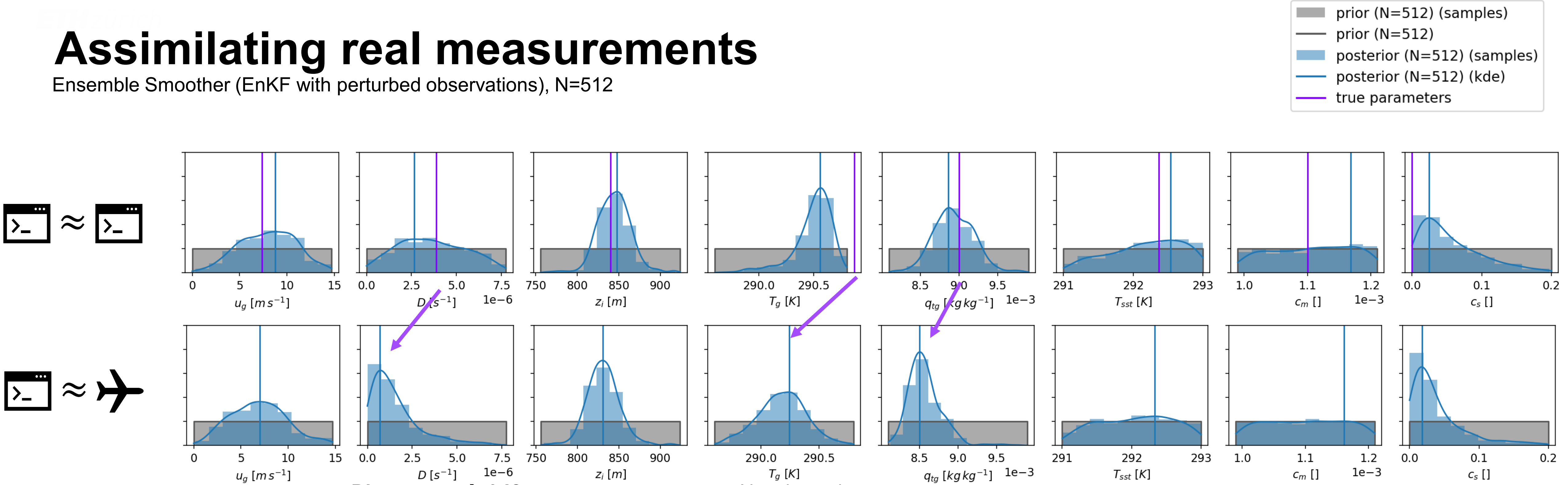
Simulate posterior ensemble (N=64) …

Good fits to the synthetic data ☺



Assimilating real measurements

Ensemble Smoother (EnKF with perturbed observations), N=512



Divergence → 0 !?

Not skewed.

Not identifiable

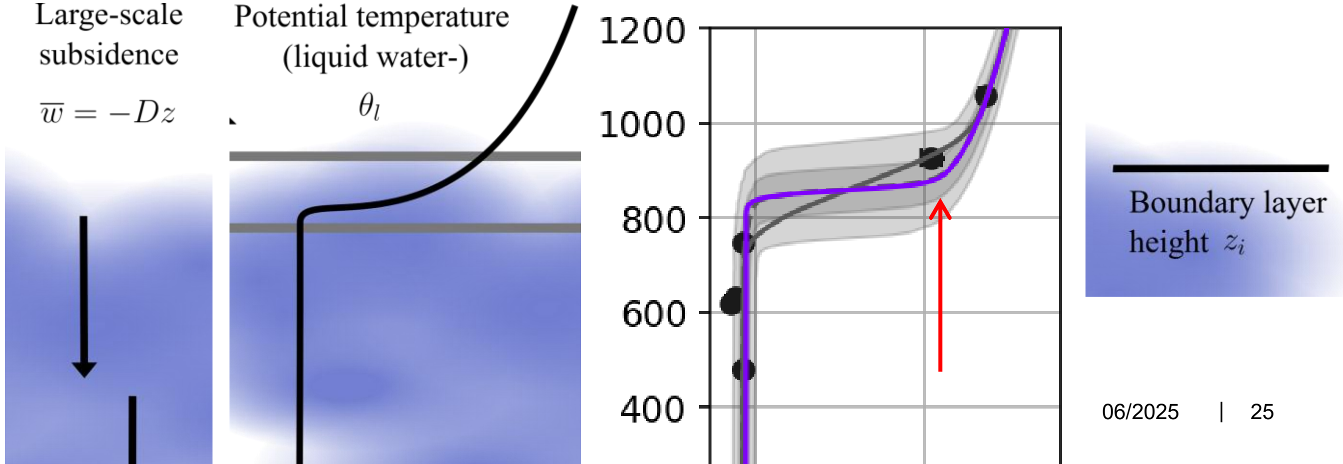
Identifiable ☺

Why divergence → 0 ?

Less „squeezing“ helps to reduce bias just above cloud top*

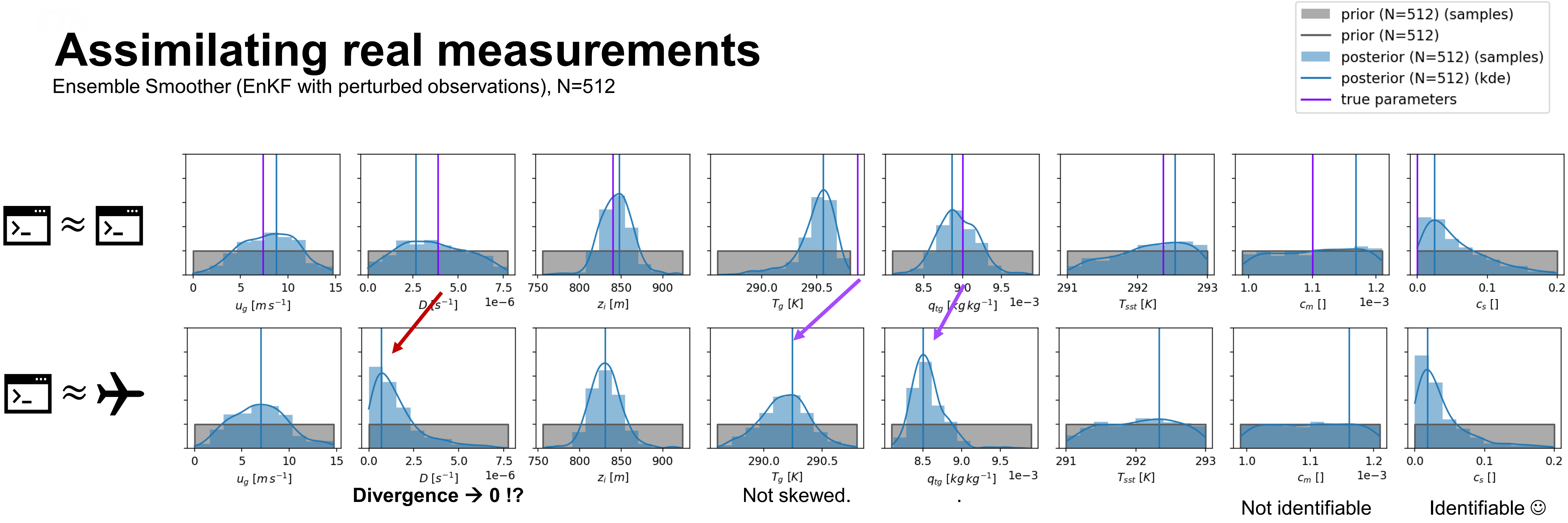
Why not increase the initial layer height z_i?

z_i is strongly constrained by cloud top height and w profiles*



Assimilating real measurements

Ensemble Smoother (EnKF with perturbed observations), N=512



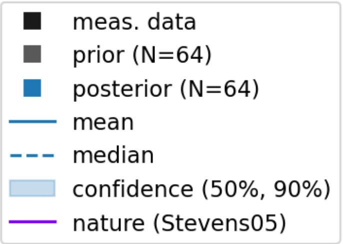
Q1. Can the model biases be eliminated by model calibration?

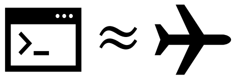
For PyCLES, adapt the nature parameter as follows:

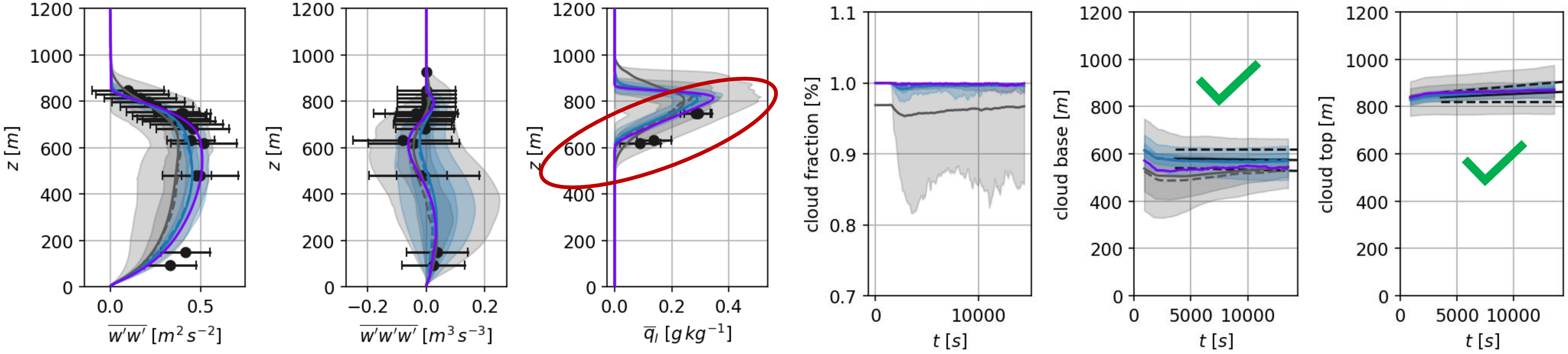
- Remove input biases $T_g \downarrow,\ q_{t,g} \downarrow$
- Correct misspecification (?) $D \rightarrow 0$
- Choose model parameter $c_s \rightarrow 0$

Assimilating real measurements

Ensemble Smoother (EnKF with perturbed observations), N=64

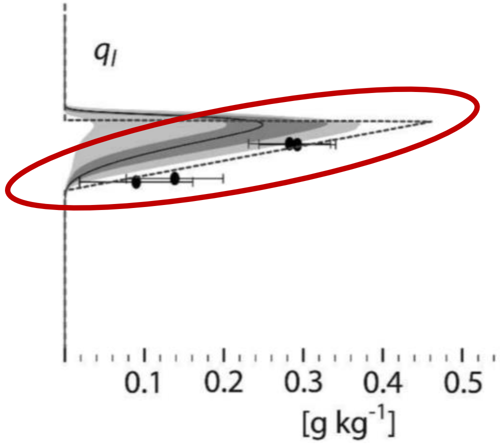


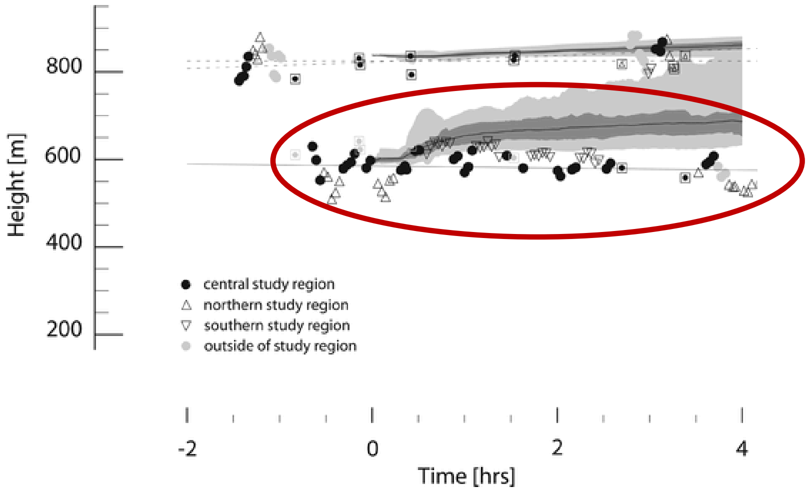


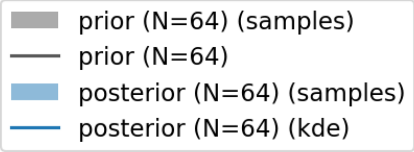


Lack of liquid water remains.

Better timeseries fit.
Growing cloud layer ☺







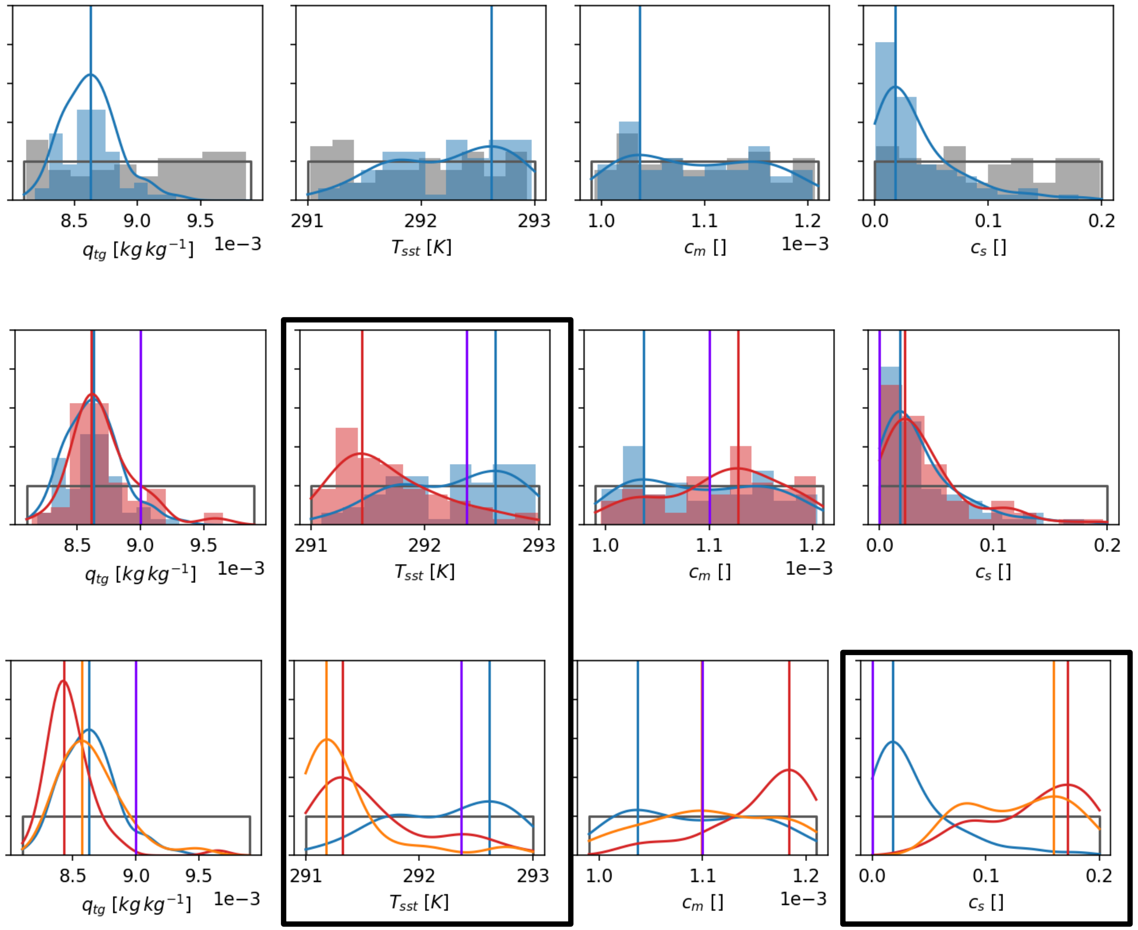
Calibrating different numerics

Ensemble Smoother (EnKF with perturbed observations), N=64

Default
dx=35, dz=5
WENO-WENO

Low resolution
dx=50, dz=15

Other numerics
central-WENO
central-central



Q2. How do different numerics affect the calibration?

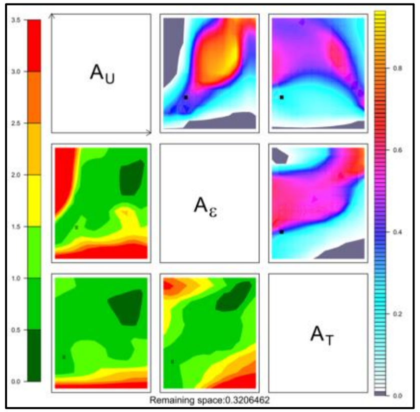
Both numerics and resolution affect the posterior estimates.

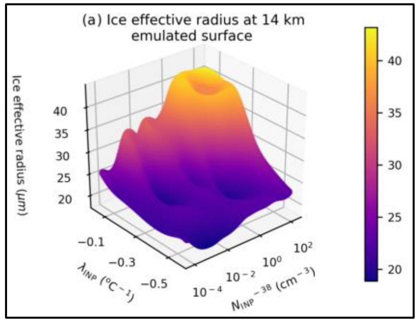
For implicit LES numerics, the Smagorinsky constant $c_s \rightarrow 0$.

Forcing bias depends on resolution and numerics.

Model parameter c_s depends on numerics, but not on resolution.

Next Steps





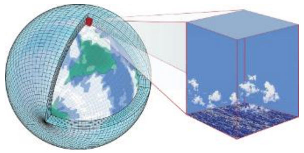
1. Compare to other calibration methods

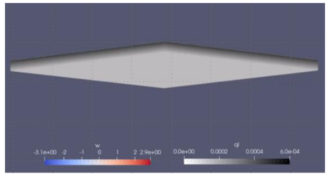
- Data assimilation: ES-MDA
- Atmospheric community: Typically O(1-10) parameters
  - Latin Hypercube sampling + performance score
    (Hawker et al. 2021, Liu et al. 2022)
  - History matching (ruling out parameter space regions)
    (Covreux et. al 2021)
  - Gaussian Process + Markov-Chain Monte—Carlo
    (Cleary et. al 2021)

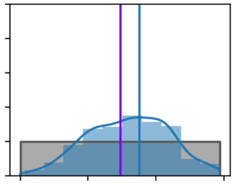
2. Increase problem complexity: $\boldsymbol{G}\colon \mathbb{R}^{\mathbf{9}} \to \mathbb{R}^{\mathbf{64}}\,;\ \boldsymbol{\theta} \mapsto \boldsymbol{y}$

- Infer full profiles → $\theta \in \mathbb{R}^{O(50)}$
- Assimilate full timeseries $y \in \mathbb{R}^{O(100)}$
- More complex dynamics → More nonlinear G
  - Differnt atmospheric conditions (test case)
  - Include microphysics parameters (precipitation)

Calibration of High-Resolution Atmospheric Models







1. Models

2. Clouds & Measurements

3. Simulation

4. Calibration

1. Large Eddy Simulations serve to calibrate lower-resolution models.

2. The DYCOMS-II RF01 marine stratocumulus measurements are a well-known intercomparison case for LES, but the models share similar biases.

3. PyCLES is an implicit LES which performs best without sub-grid scale scheme.

4. EnKF can be used to tune PyCLES to DYCOMS-II RF01.

   Q1: We obtain improved input biases and model parameters.

   Q2: The posterior estimates are sensitive to resolution and numerics.

References

DYCOMS-RF01 test case

Stevens, B., D. H. Lenschow, I. Faloona, et al. (2003). “On Entrainment Rates in Nocturnal Marine Stratocumulus”. In: Quarterly Journal of the Royal Meteorological Society 129.595, pp. 3469–3493. doi: 10.1256/qj.02.202.

Stevens, B., D. H. Lenschow, G. Vali, et al. (2003). “Dynamics and Chemistry of Marine Stratocumulus—DYCOMS-II”. In: Bulletin of the American Meteorological Society 84.5, pp. 579–594. doi: 10.1175/BAMS-84-5-579.

Stevens, B. et al. (2005). “Evaluation of Large-Eddy Simulations via Observations of Nocturnal Marine Stratocumulus”. In: Monthly Weather Review 133.6, pp. 1443–1462. doi: 10 . 1175 /MWR2930.1.

PyCLES model

Pressel, K. G., C. M. Kaul, T. Schneider, Z. Tan, and S. Mishra (2015). “Large-Eddy Simulation in an Anelastic Framework with Closed Water and Entropy Balances”. In: Journal of Advances in Modeling Earth Systems 7.3, pp. 1425–1456. doi: 10.1002/2015MS000496.

Pressel, K. G., S. Mishra, T. Schneider, C. M. Kaul, and Z. Tan (2017). “Numerics and Subgrid-Scale Modeling in Large Eddy Simulations of Stratocumulus Clouds”. In: Journal of Advances in Modeling Earth Systems 9.2, pp. 1342–1365. doi: 10.1002/2016MS000778.

Model calibration

Cleary, E., A. Garbuno-Inigo, S. Lan, T. Schneider, and A. M. Stuart (2021). “Calibrate, Emulate, Sample”. In: Journal of Computational Physics 424, p. 109716. doi: 10.1016/j.jcp.2020.109716.

Couvreux, F. et al. (2021). “Process-Based Climate Model Development Harnessing Machine Learning: I. A Calibration Tool for Parameterization Improvement”. In: Journal of Advances in Modeling Earth Systems 13.3, e2020MS002217. doi: 10.1029/2020MS002217.

Hawker, R. E., A. K. Miltenberger, J. S. Johnson, J. M. Wilkinson, A. A. Hill, B. J. Shipway, P. R. Field, B. J. Murray, and K. S. Carslaw (2021). “Model Emulation to Understand the Joint Effects of Ice-Nucleating Particles and Secondary Ice Production on Deep Convective Anvil Cirrus”. In: Atmospheric Chemistry and Physics 21.23, pp. 17315–17343. doi: 10.5194/acp-21-17315-2021.

Liu, S., C. Zeman, S. L. Sørland, and C. Schär (2022). “Systematic Calibration of a Convection-Resolving Model: Application Over Tropical Atlantic”. In: Journal of Geophysical Research: Atmospheres 127.23, e2022JD037303. doi: 10.1029/2022JD037303.

Schneider, T., J. Teixeira, C. S. Bretherton, F. Brient, K. G. Pressel, C. Schär, and A. P. Siebesma (2017). “Climate Goals and Computing the Future of Clouds”. In: Nature Climate Change 7.1, pp. 3–5. doi: 10.1038/nclimate3190.