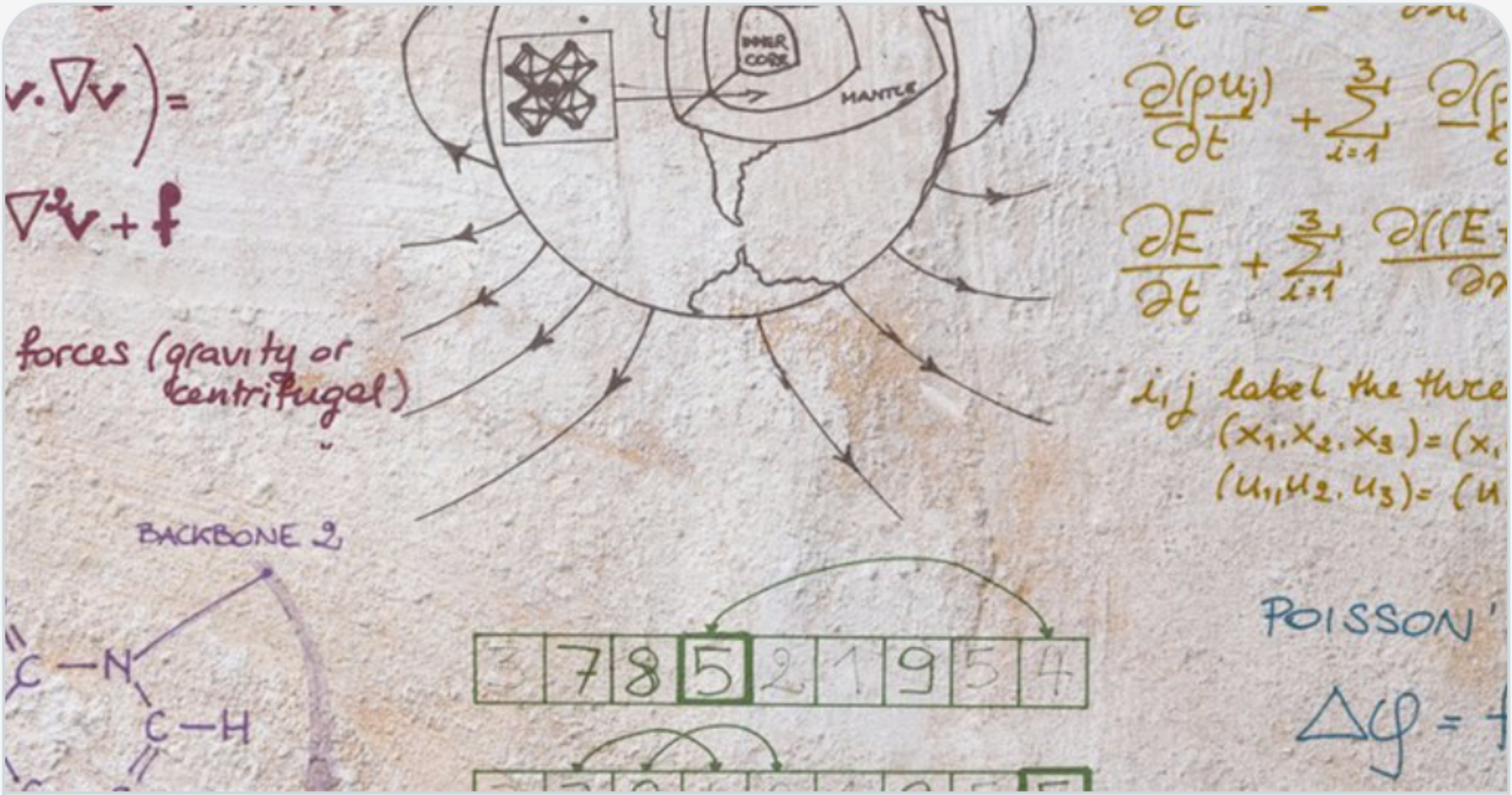
COLUMBIA CLIMATE SCHOOL
LAMONT-DOHERTY EARTH OBSERVATORY

Atmospheric radiation: using machine learning for the unknowable and uncomputable

Robert Pincus

Mike Pritchard @Pritchard_UCI · Oct 29

Interdiciplinary computer science conference seeks mini-symposia proposals: your chance to give four of your favorite speakers at the computing/earth science interface half an hour each to simulate and provoke. Deadline Nov 13.



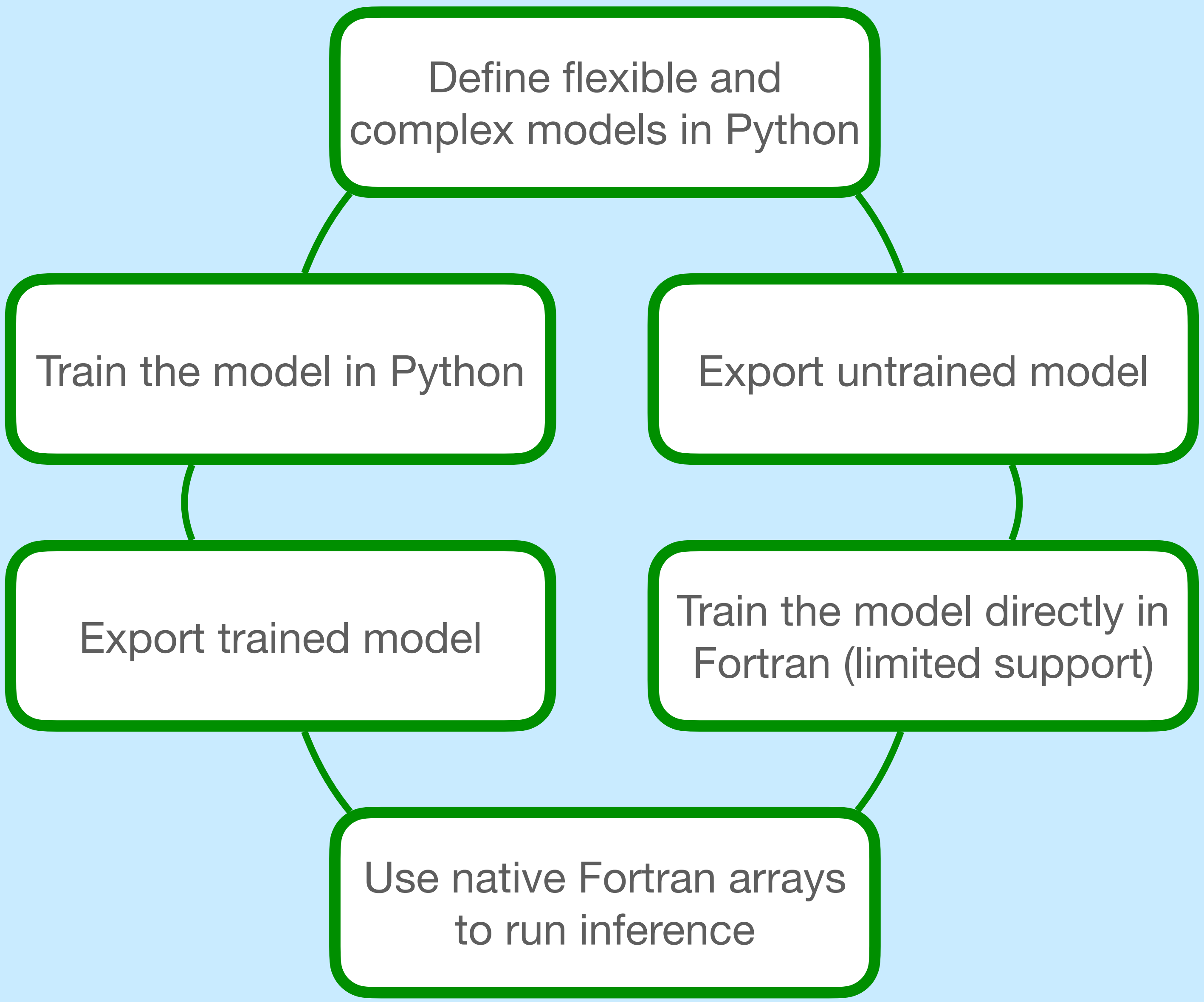
pasc22.pasc-conference.org

Guidelines for Minisymposia | PASC22

1 9

Pytorch Fortran bindings

Simplify usage of modern DL in HPC applications




```python
import torch.nn as nn
import torch.nn.functional as F


class Net(nn.Module):
```

```python
for epoch in range(2):  # loop over the dataset multiple times

    running_loss = 0.0
    for i, data in enumerate(trainloader, 0):
```

```fortran
call torch_mod%load(in_fname, flag)
call torch_mod%create_optimizer_sgd(0.1)
call in_tensor%    from_array(input)
call target_tensor%from_array(target)
```

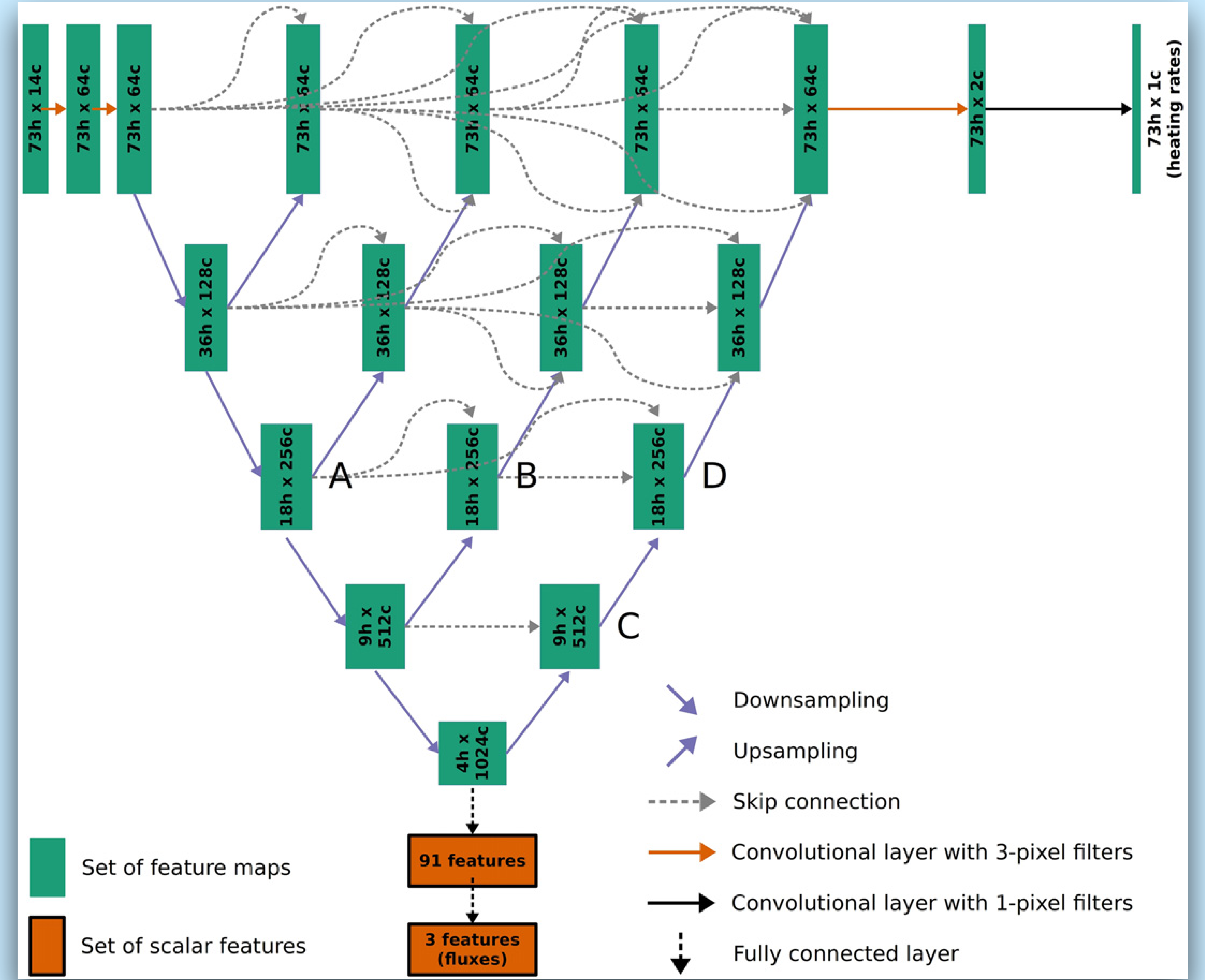
```fortran
type(torch_module) :: torch_mod
type(torch_tensor) :: in_tensor, out_tensor

real(real32) :: input(224, 224, 3, 10)
real(real32), pointer :: output(:, :)

input = 1.0
call in_tensor%from_array(input)
call torch_mod%load(filename)
call torch_mod%forward(in_tensor, out_tensor)
call out_tensor%to_array(output)
```

Machine learning for radiation has a 20+ year history

The first uses of machine learning in weather forecasting were by Chevallier and colleagues (ECMWF,1998-2000) and Krasnoplosky and colleagues (NCEP, 2004-2010).

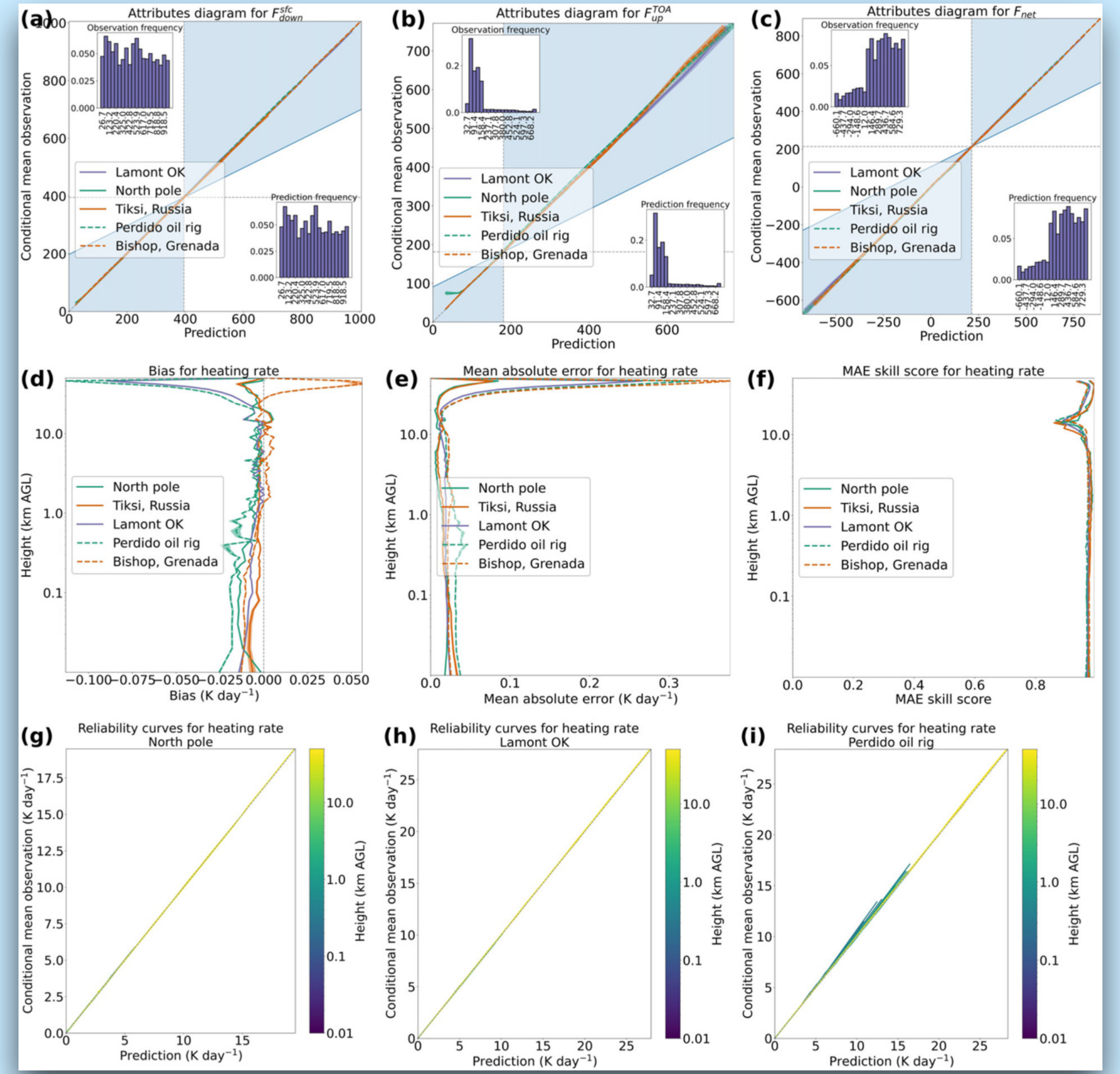
Both aimed to replace radiation calculations with artificial neural networks to increase computational speed. ANNs were trained on parameterizations and succeeded in reproducing fluxes, which were could be used in stable simulations.

People are still doing this (with important wrinkles)



Largerquist et al. 2021, doi:10.1175/JTECH-D-21-0007.1

People are still doing this (with important wrinkles)



Largerquist et al. 2021, doi:10.1175/JTECH-D-21-0007.1

No need for speed (alone)…

Computational efficiency is (IMHO) a poor motivation to “learn radiation”

removing the (current) cost isn’t transformative
existing compromises work reasonably well
computational alternatives abound

Machine learning of parameterized radiation inherits errors and simplifications. It abandons equation we know to be true - “a terrible idea”

But! Lagerquist trained on fast-*ish* calculations - slower than routine, but not necessarily heroic

Computational cost is the data-limiting factor

A hybrid approach might learn the correction to a low-resolution calculation

One might even want to build a forward model but never use it directly (at least for NWP)…

ML to enable 4Dvar data assimilation

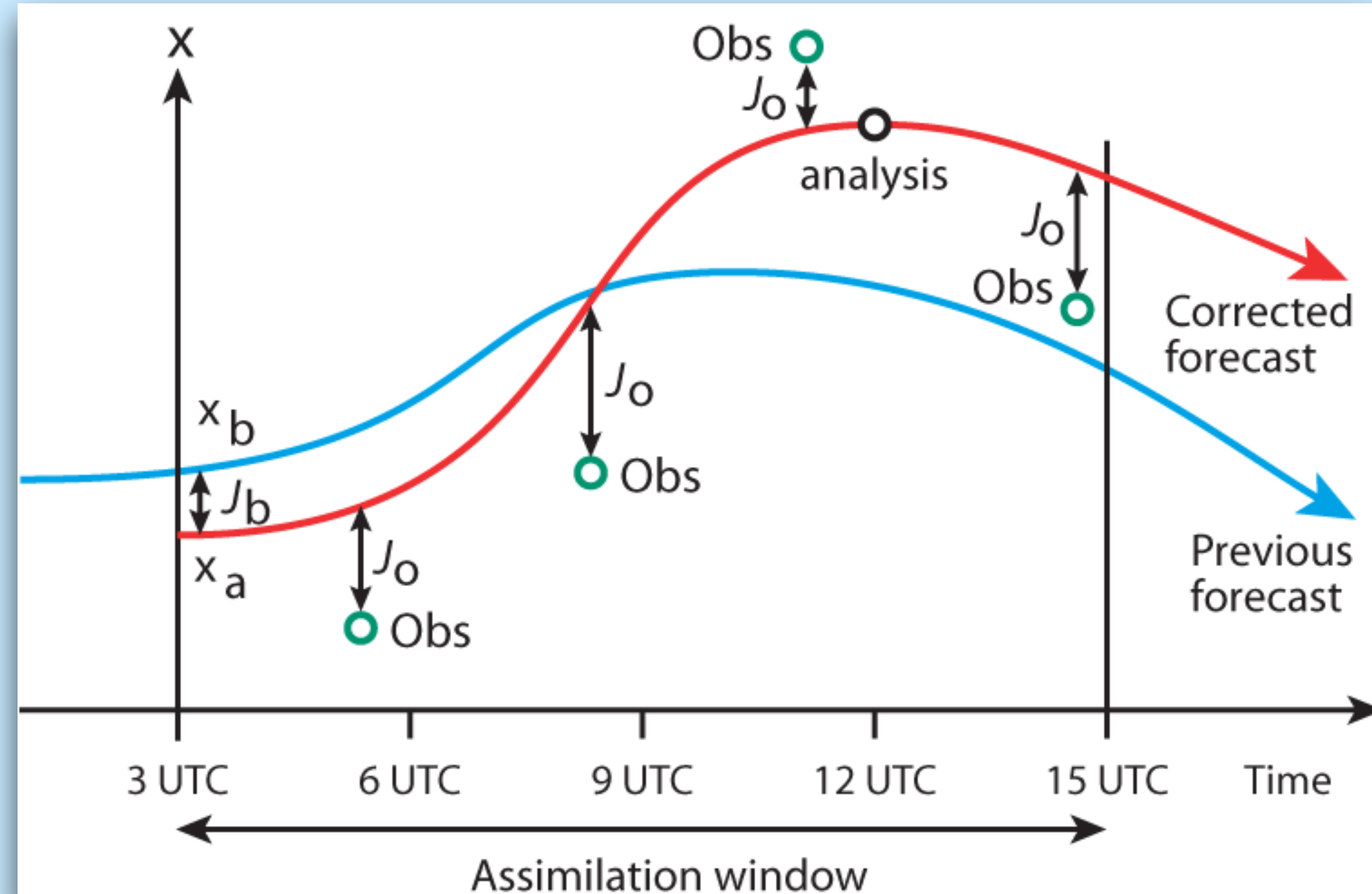
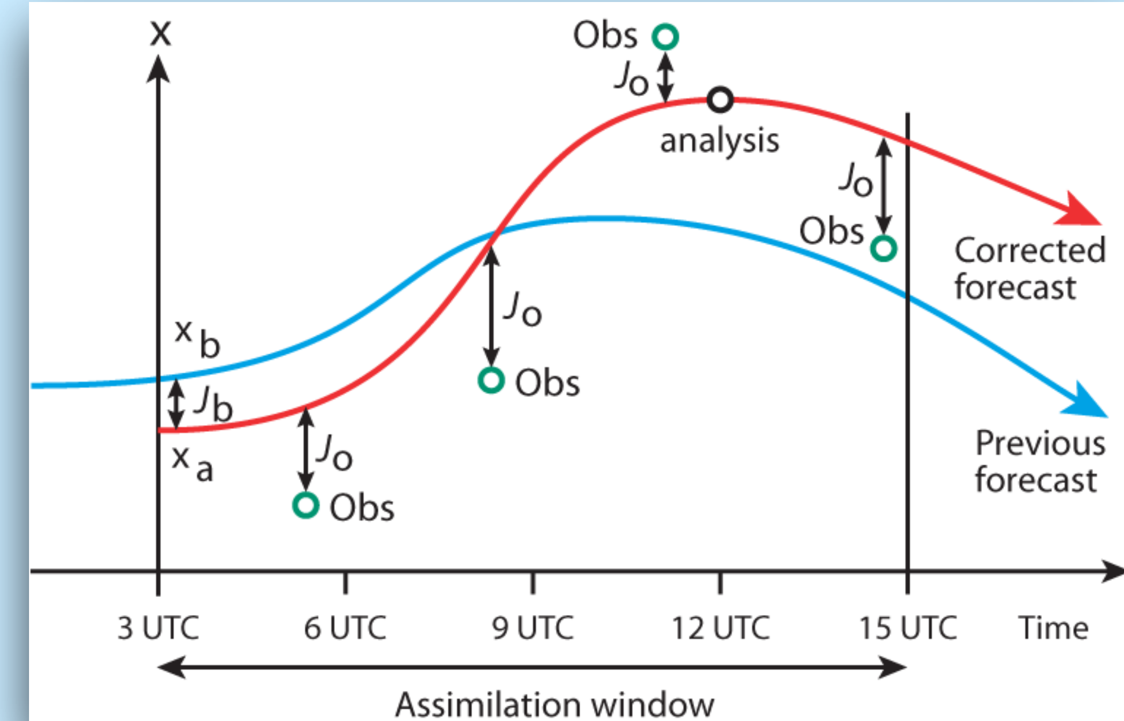


Figure: ECMWF

ML to enable 4Dvar data assimilation



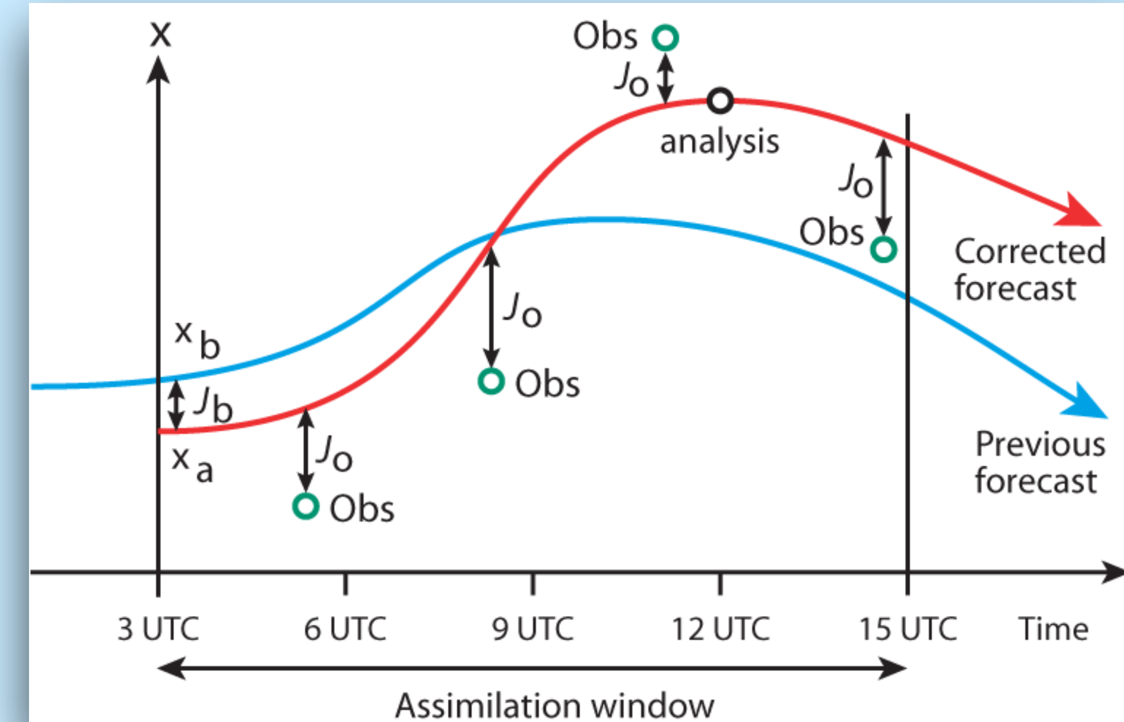
Q. J. R. Meteorol. Soc. (2004), **130**, pp. 2495–2517 doi: 10.1256/qj.03.162

A cloud scheme for data assimilation: Description and initial tests

By A. M. TOMPKINS*and M. JANISKOVÁ
European Centre for Medium-Range Weather Forecasts, Reading, UK

(Received: 21 August 2003; revised 24 February 2004)

ML to enable 4Dvar data assimilation



Q. J. R. Meteorol. Soc. (2004), **130**, pp. 2495–2517 doi: 10.1256/qj.03.162

A cloud scheme for data assimilation: Description and initial tests

By A. M. TOMPKINS* and M. JANISKOVÁ
European Centre for Medium-Range Weather Forecasts, Reading, UK

(Received: 21 August 2003; revised 24 February 2004)

JAMES Journal of Advances in Modeling Earth Systems

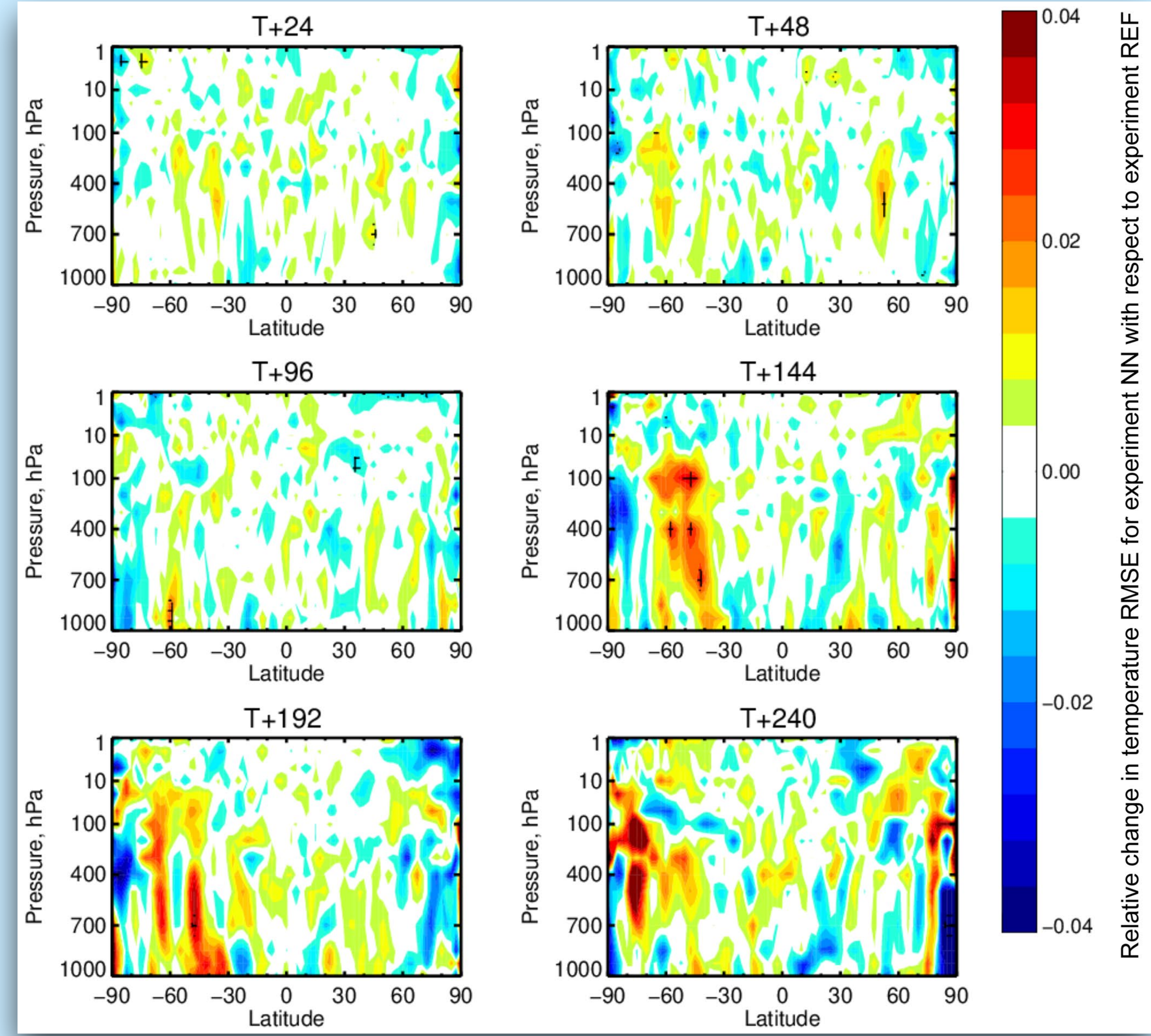
RESEARCH ARTICLE
10.1029/2021MS002521

Building Tangent-Linear and Adjoint Models for Data Assimilation With Neural Networks

Sam Hatfield[1], Matthew Chantry[2], Peter Dueben[1], Philippe Lopez[1], Alan Geer[1], and Tim Palmer[2]

[1]European Centre for Medium-Range Weather Forecasts, Reading, UK, [2]Atmospheric, Oceanic and Planetary Physics, University of Oxford, Oxford, UK

ML to enable 4Dvar data assimilation



Hatfield et al. 2021, doi:10.1029/2021MS002521

Building a radiation parameterization: theory…

From the equations describing radiation in the atmosphere

$$F^{\pm}(\mathbf{x}) = \int_0^{\infty} \int^{2\pi} \hat{n} \cdot I_\nu(\Omega, \mathbf{x}) d\Omega d\nu$$

$$\Omega \cdot \nabla I_\nu(\Omega, \mathbf{x}) = -\beta_\nu(\mathbf{x}) I_\nu(\Omega, \mathbf{x}) + S_\nu(\mathbf{x})$$

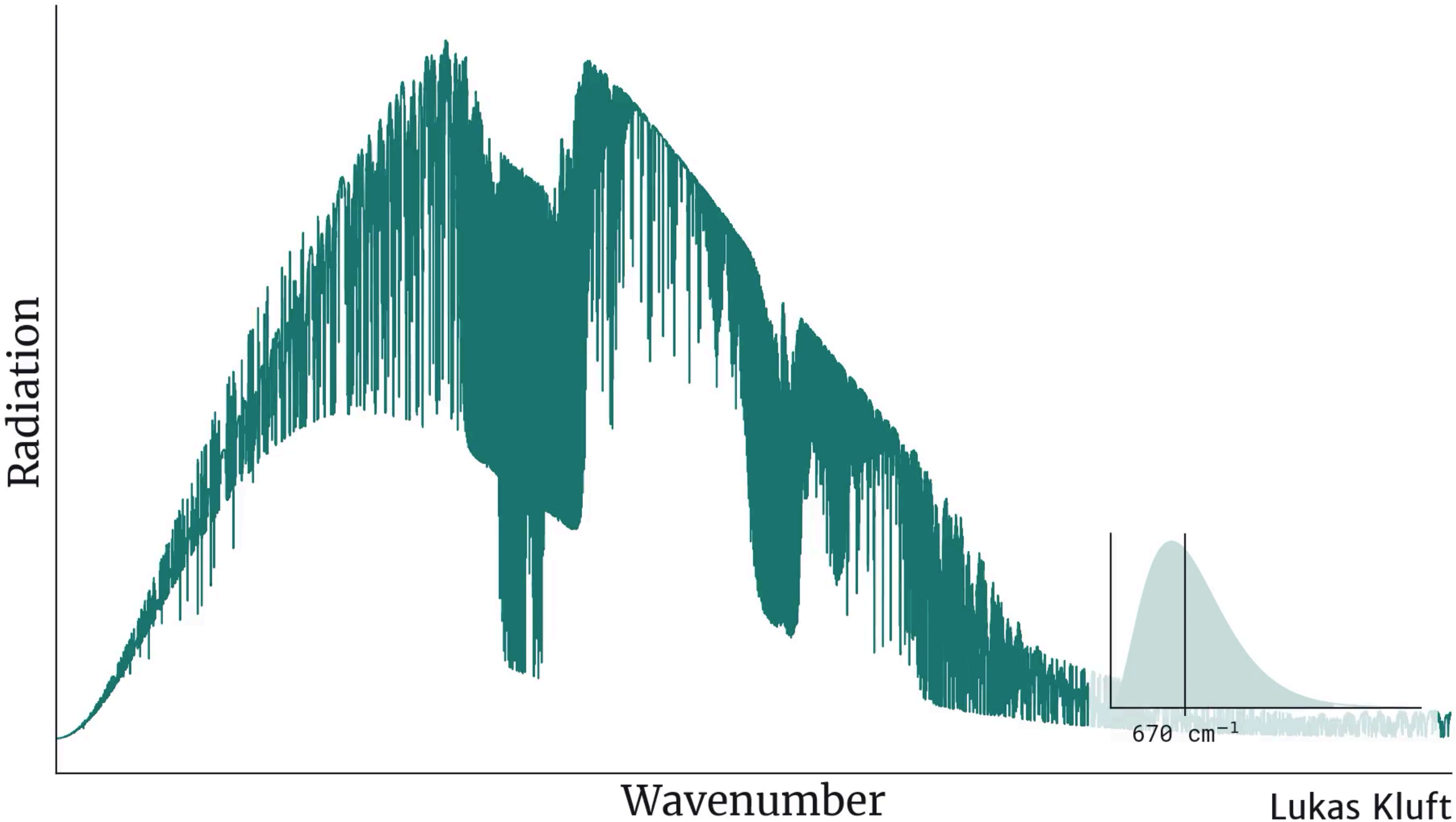
parameterizations make three approximations

plane-parallel, homogeneous *ansatz*: $u \dfrac{dI_\nu(\tau_\nu, u, \phi)}{d\tau_\nu} = -I_\nu(\tau_\nu) + S_\nu(\tau_\nu)$

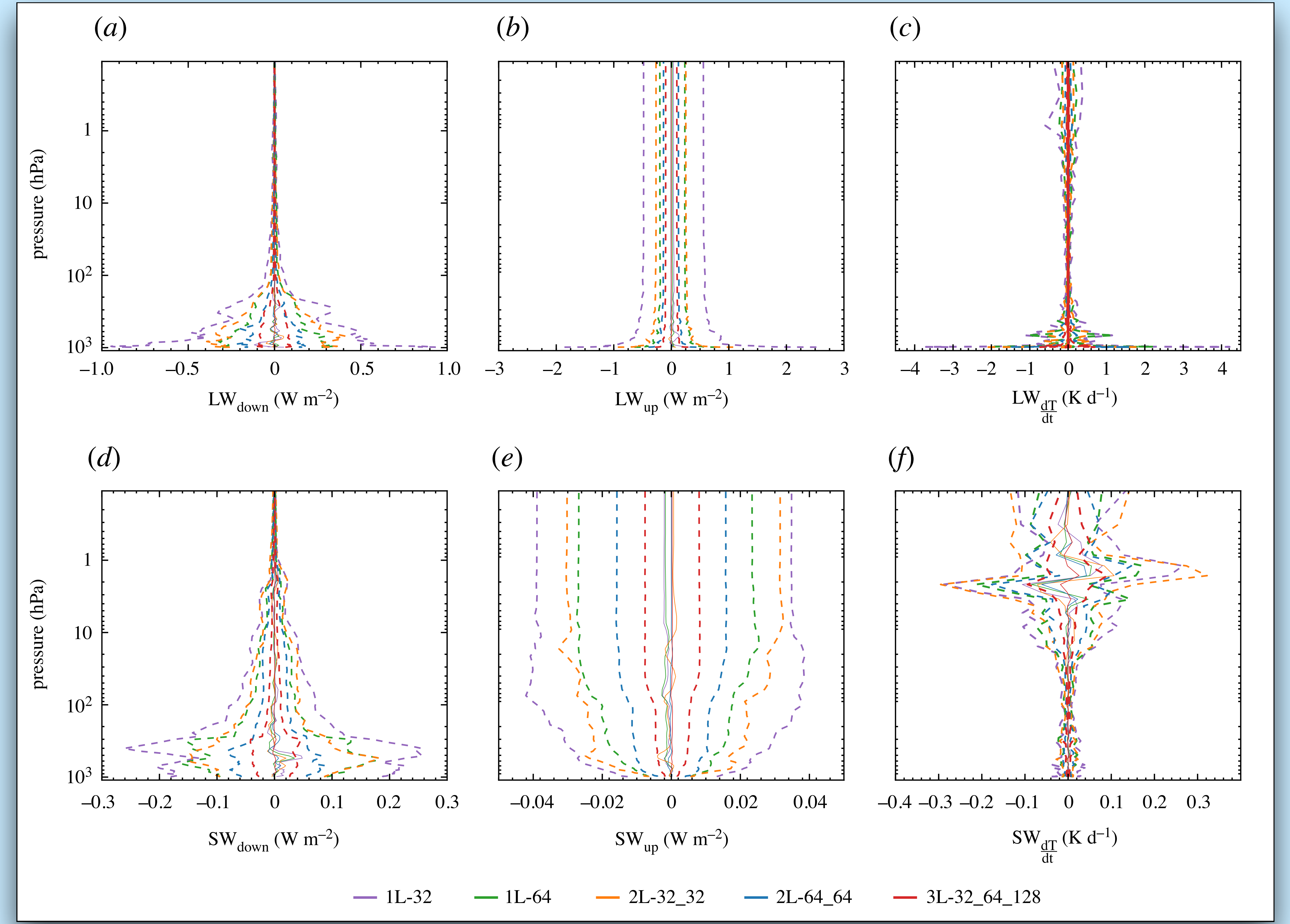
analytic angular integration i.e. two-stream: $\dfrac{dF_\nu^+}{d\tau} = \gamma_1 F_\nu^+ - \gamma_2 F_{nu}^- + S_\nu(\tau_\nu)$

spectral integration: $\displaystyle\int_0^{\infty} F_\nu^+ d\nu \approx \sum_g^G F_g$

… and a healthy dose of empiricism

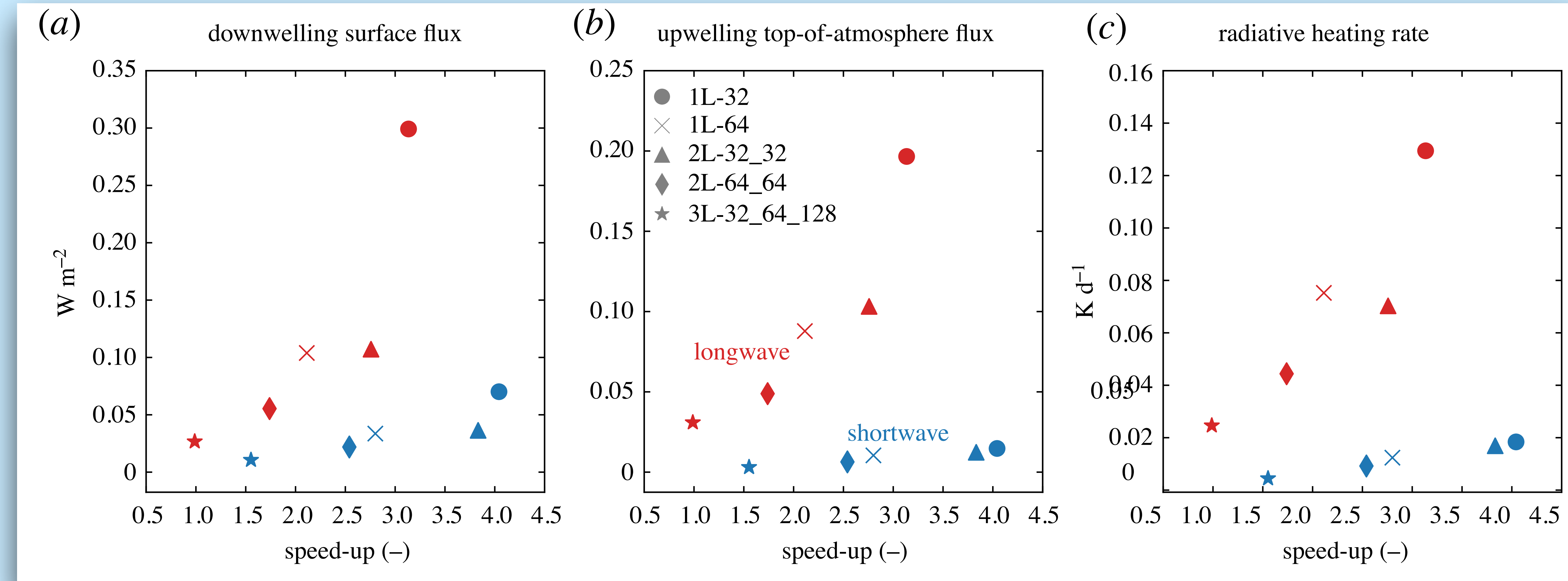


Empiricism is catnip for ML



Veerman et al. 2021, doi:10.1098/rsta.2020.0095

Empiricism is catnip for ML



Veerman et al. 2021, doi:10.1098/rsta.2020.0095

Learning our way towards benchmarks in the presence of empiricism

The main source of non-random errors in radiation parameterizations is the one-dimensional *ansatz*

This is actually several approximations

a. net horizontal transport of radiation between columns (see: advection)

b. homogeneity within columns

Treating inhomogeneity: tricks plus empiricism

The equations can be solved independently for each element of the one-point distribution of the vertical distribution of opacity, directly or with tricks

Determining this distribution requires information about

the distribution of opacity in each layer

how the distributions are related in the vertical

at scales smaller than the spatial discretization but larger than a radiative smoothing scale (O(100) m or larger)

The distribution of opacity might be available from other sub-models. Vertical relationships are prescribed and/or diagnosed empirically

There are opportunities to learn the characteristics of sub-grid homogeneity from empirical data

Learning our way towards benchmarks in the presence of empiricism

The main source of non-random errors in radiation parameterizations is the one-dimensional *ansatz*

This is actually several approximations

a. net horizontal transport of radiation between columns (see: advection)

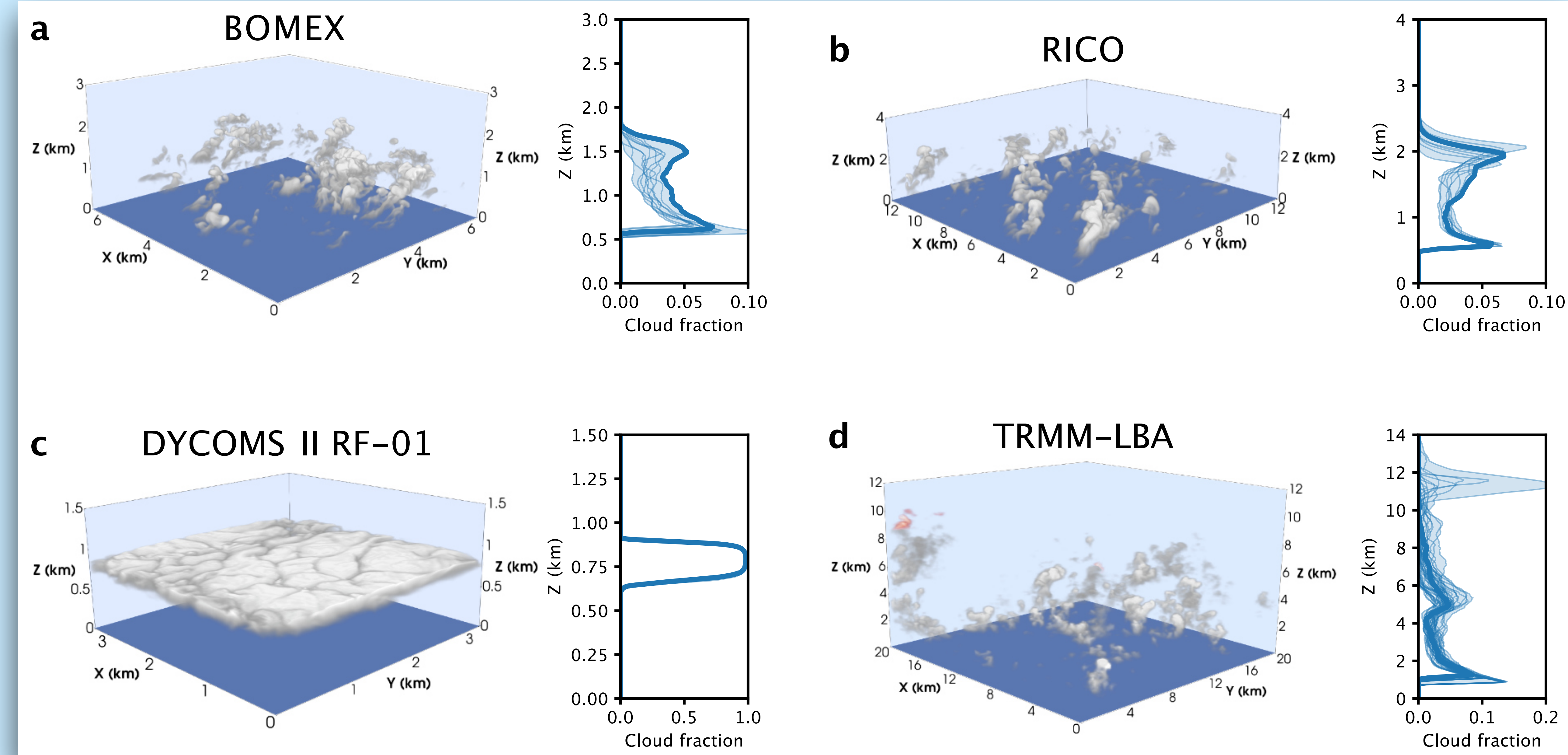
b. homogeneity within columns

c. net horizontal transport of radiation within columns (coupled with b)

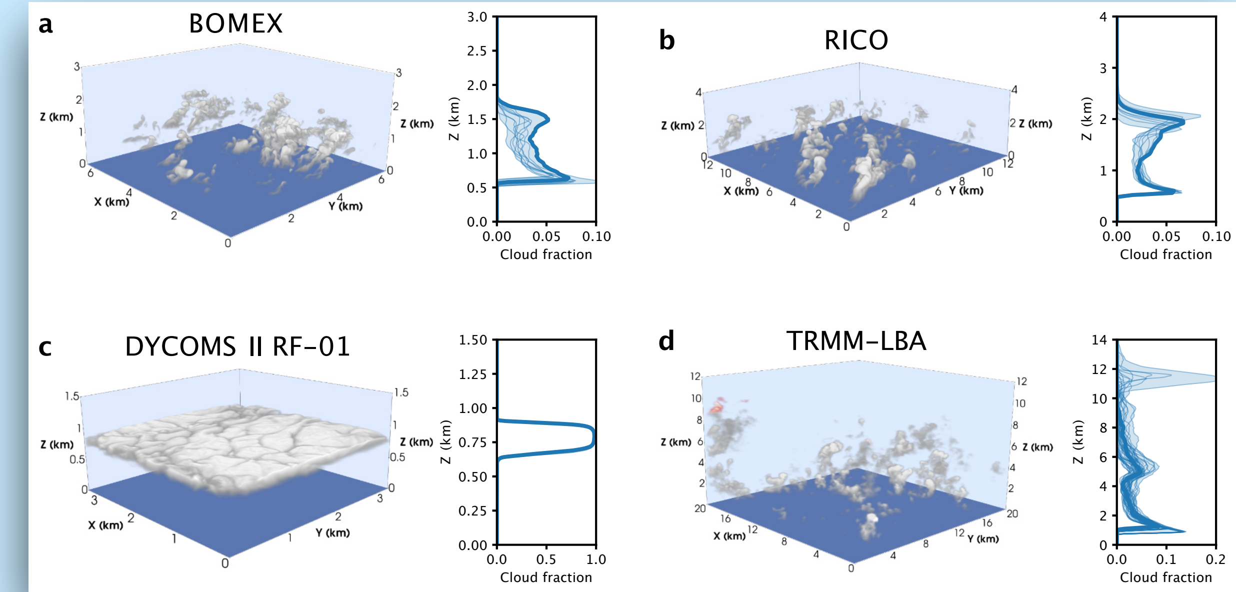
Relaxing assumption c is both computationally and conceptually hard.

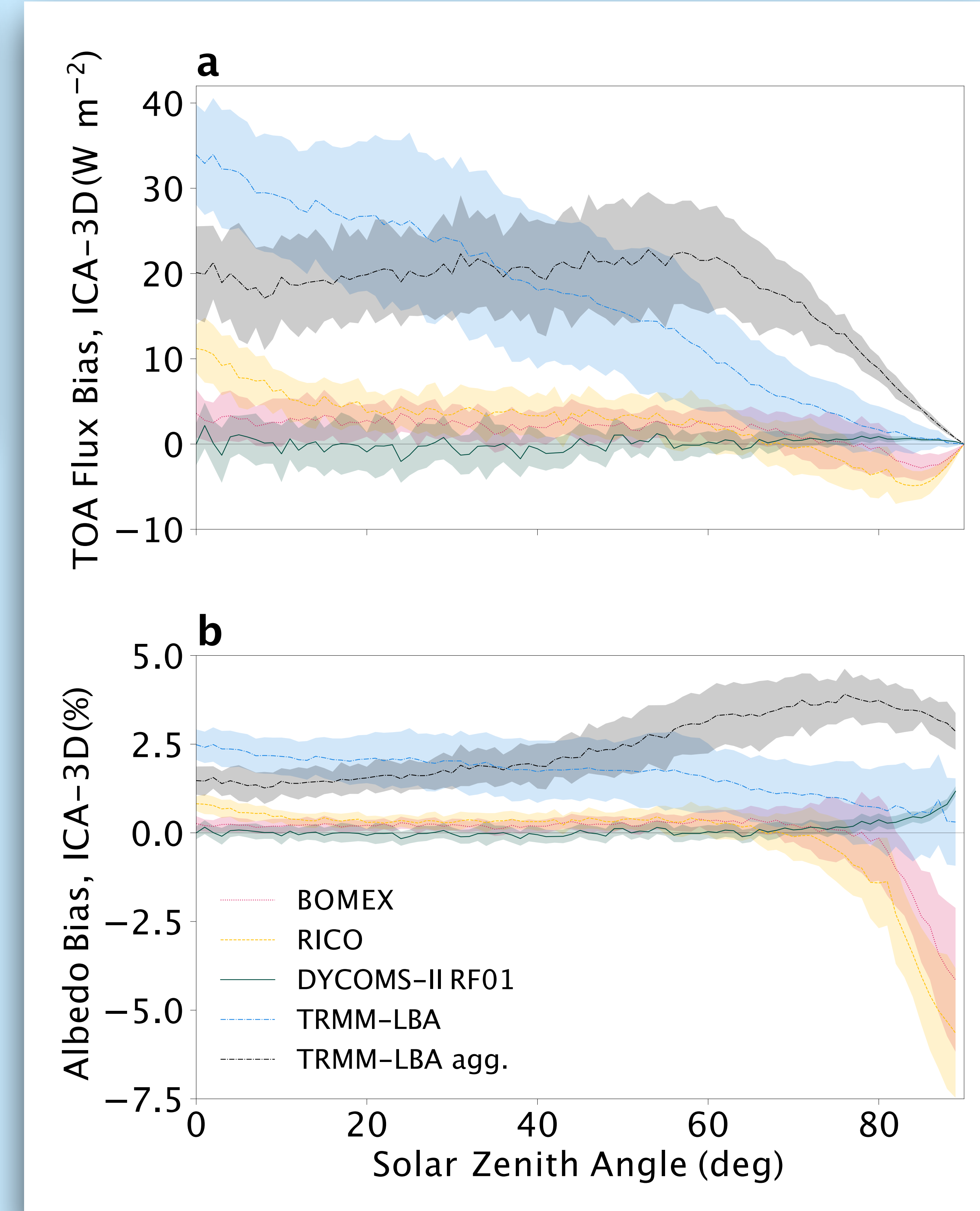
The degree to which approximations b and c impact fluxes is directly tied to the small-scale spatial distribution of the properties of the medium (most likely clouds)

Lacking a theory for this distribution we are back to empiricism

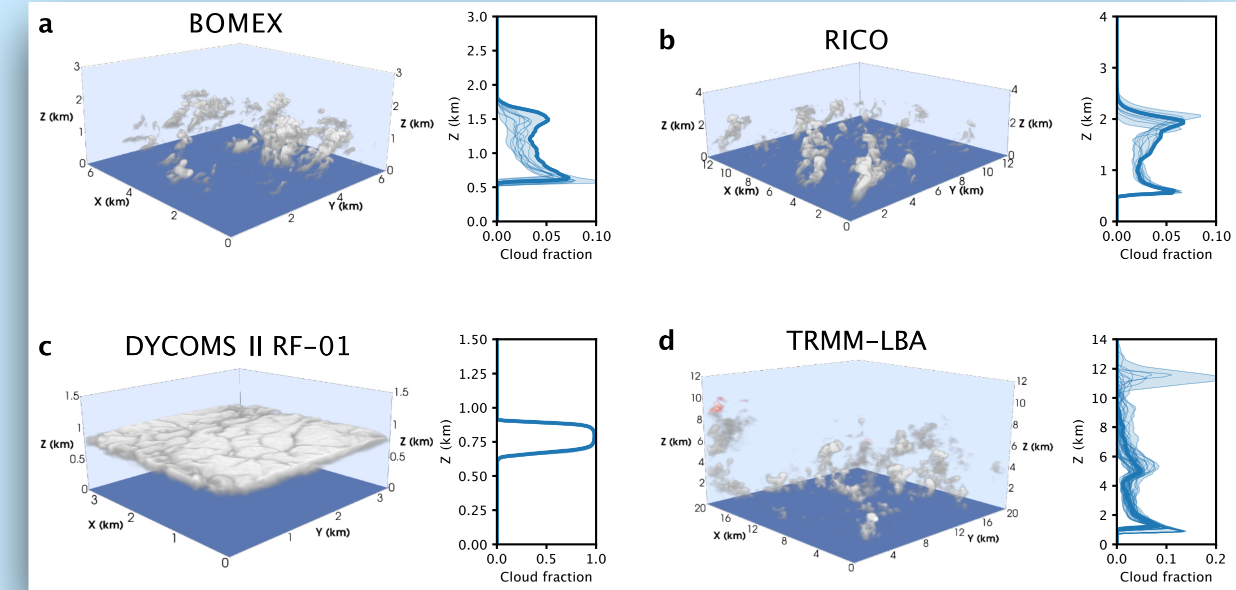


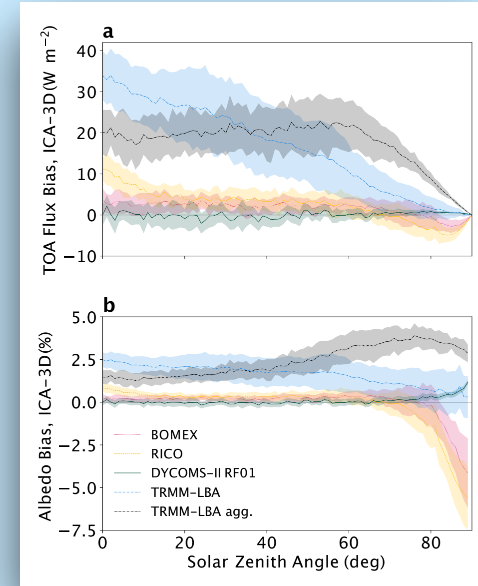
Singer et al. 2021, doi:10.1175/JAS-D-21-0032.1

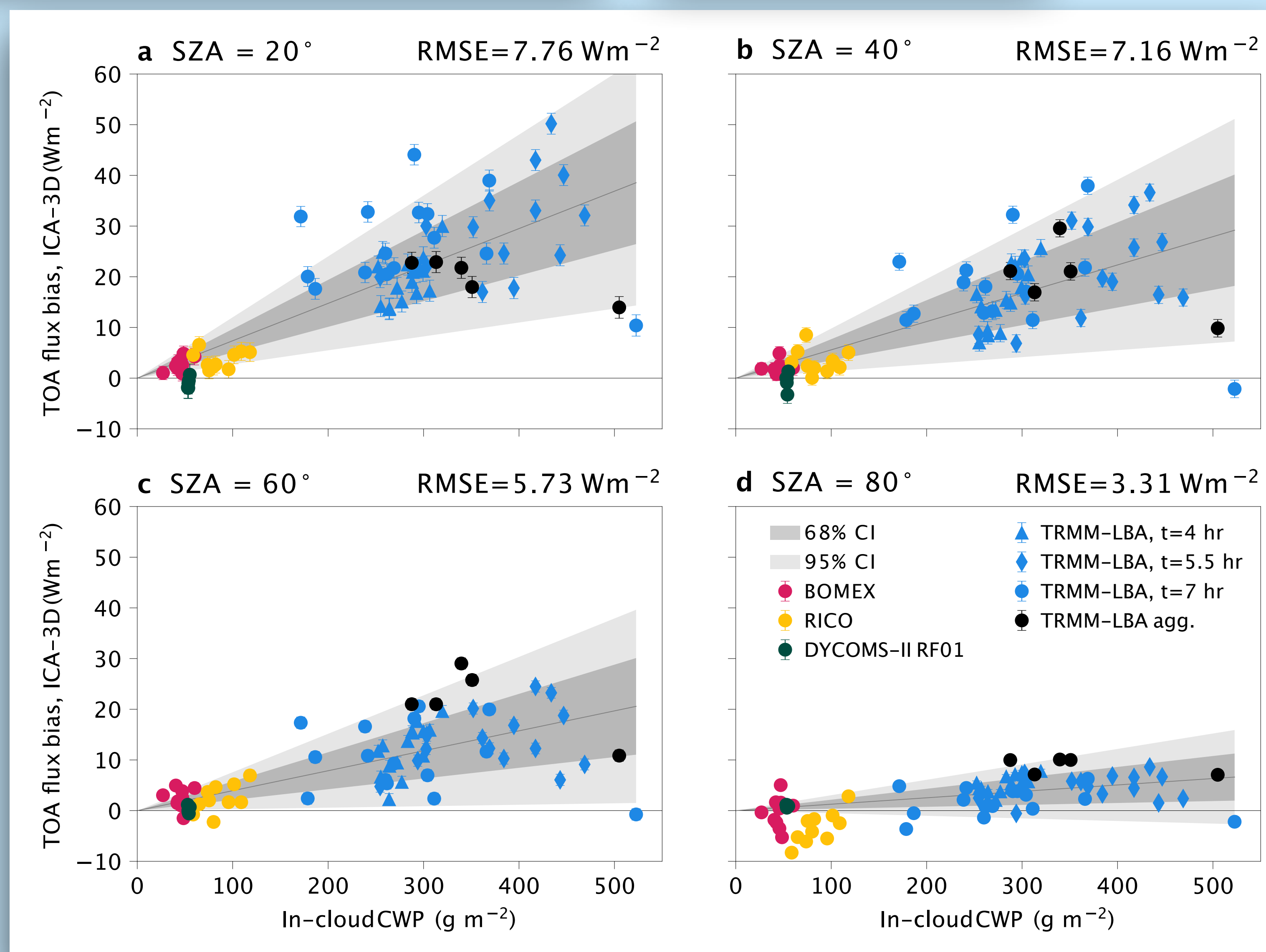




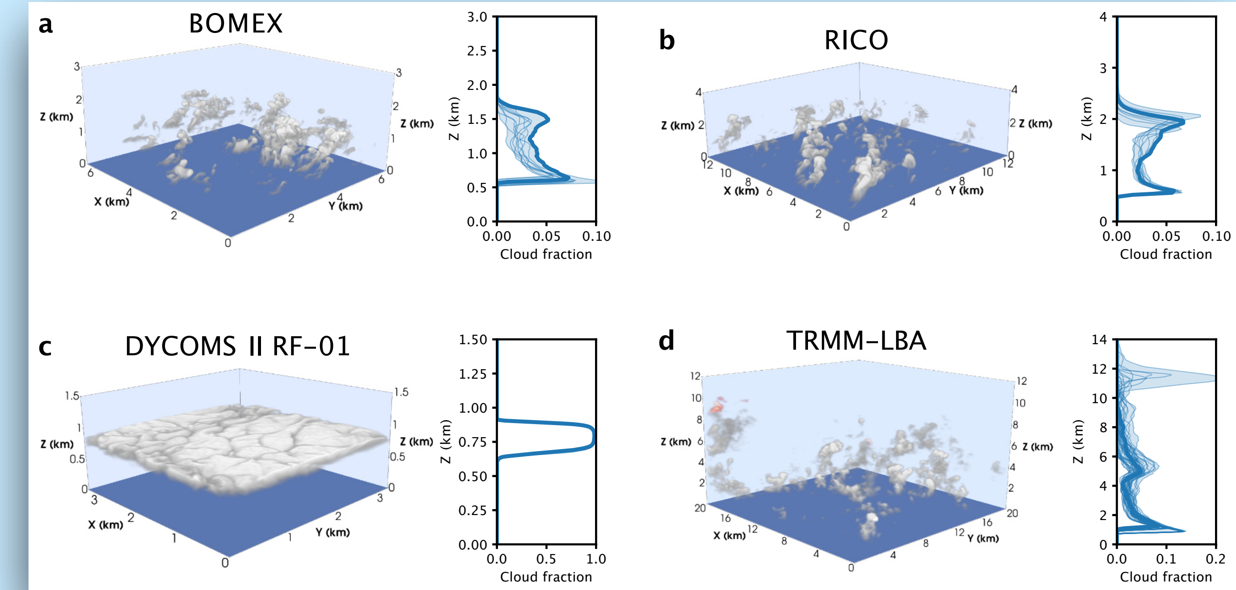
Singer et al. 2021, doi:10.1175/JAS-D-21-0032.1

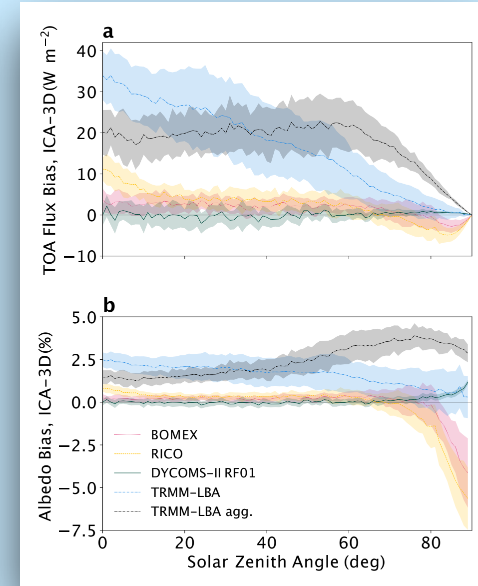


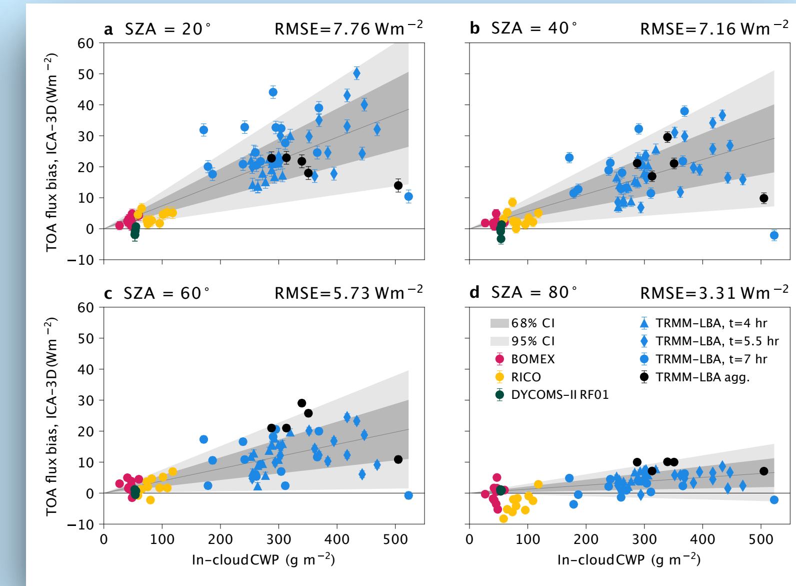


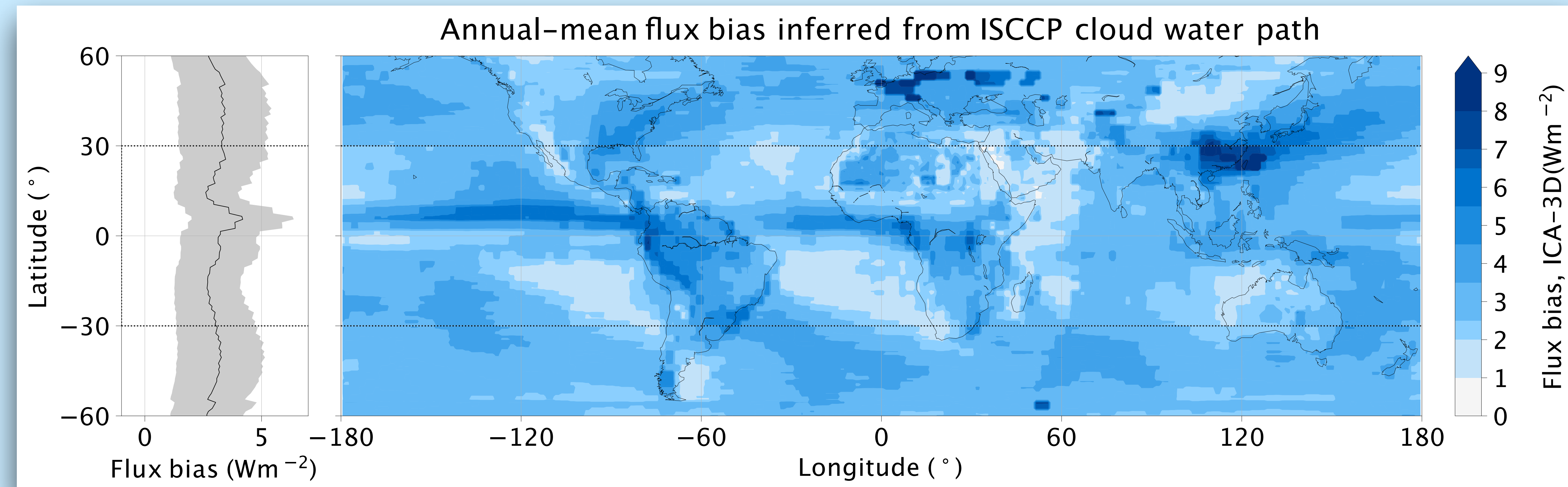


Singer et al. 2021, doi:10.1175/JAS-D-21-0032.1





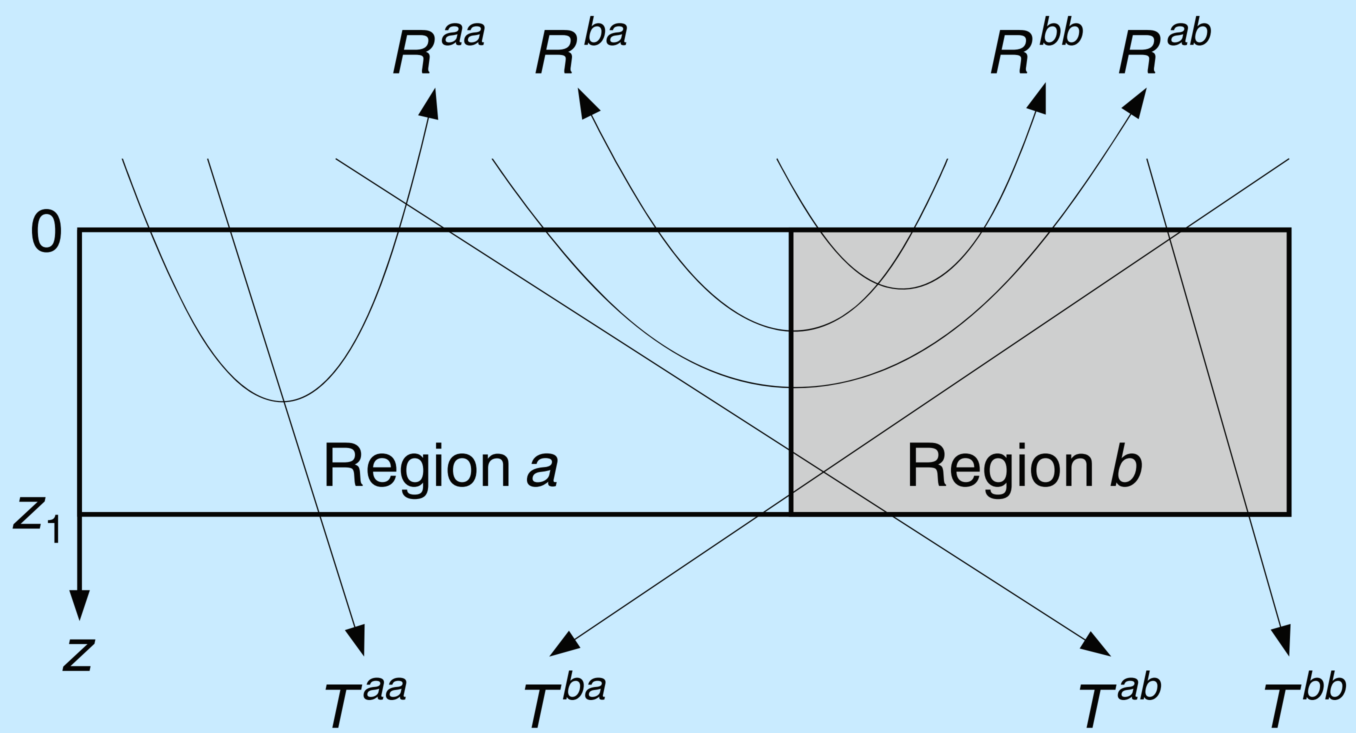




Singer et al. 2021, doi:10.1175/JAS-D-21-0032.1

Fast-ish treatments of 3D sub-grid effects

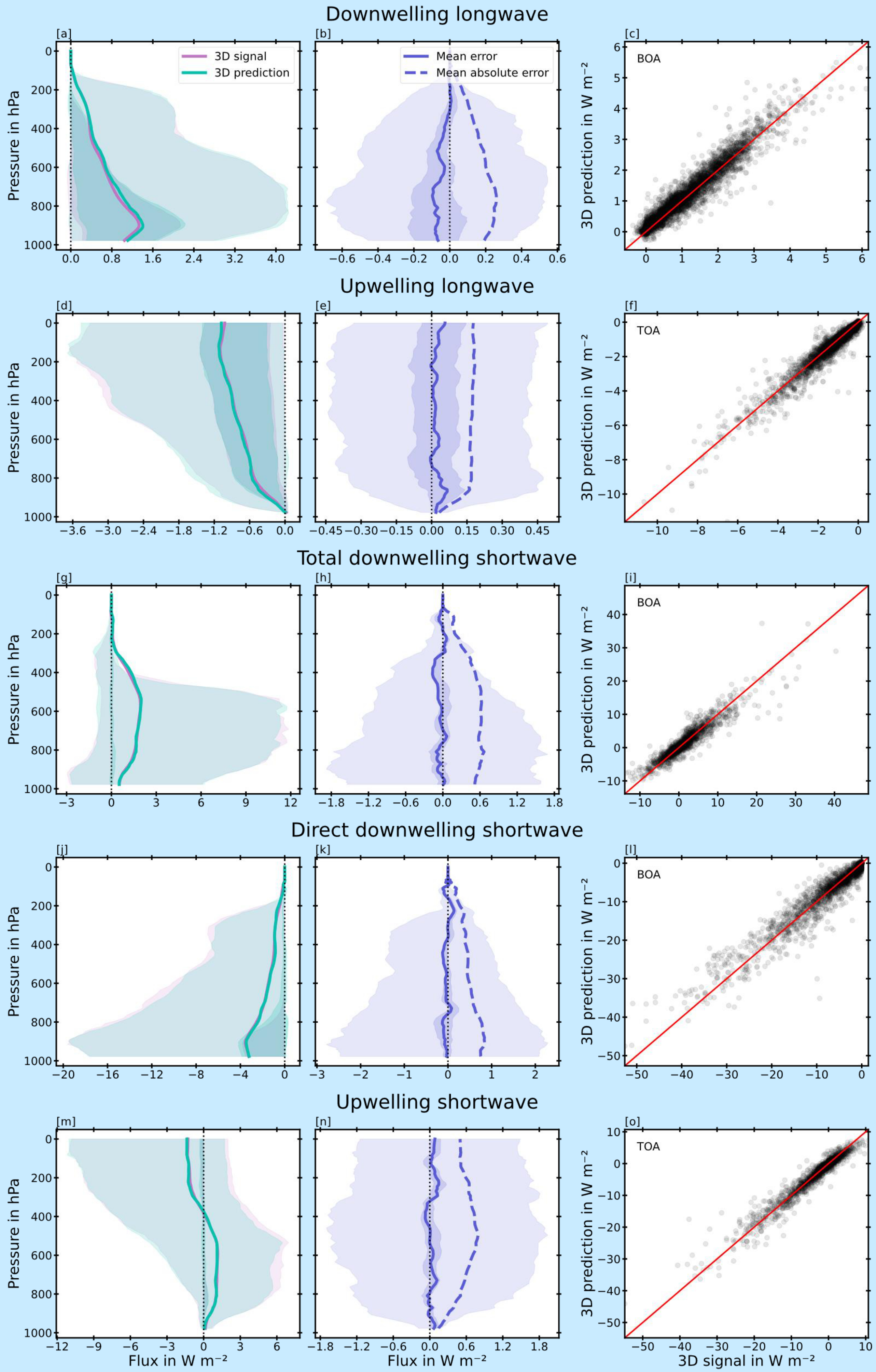
Robin Hogan's SPARTACUS is an ambitious extension of the two-stream/adding paradigm to account for net horizontal transport between subgrid-scale cloud elements.



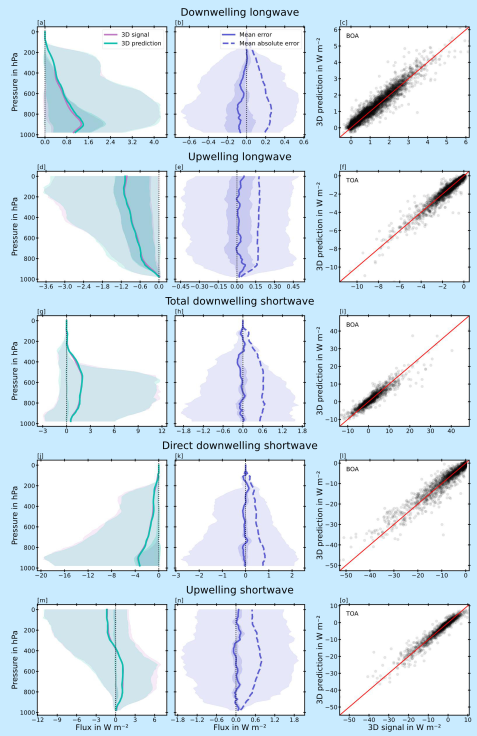
SPARTACUS treats biases, especially long wave radiation from cloud sides, and conditional errors that depend on solar zenith angle and cloud distributions.

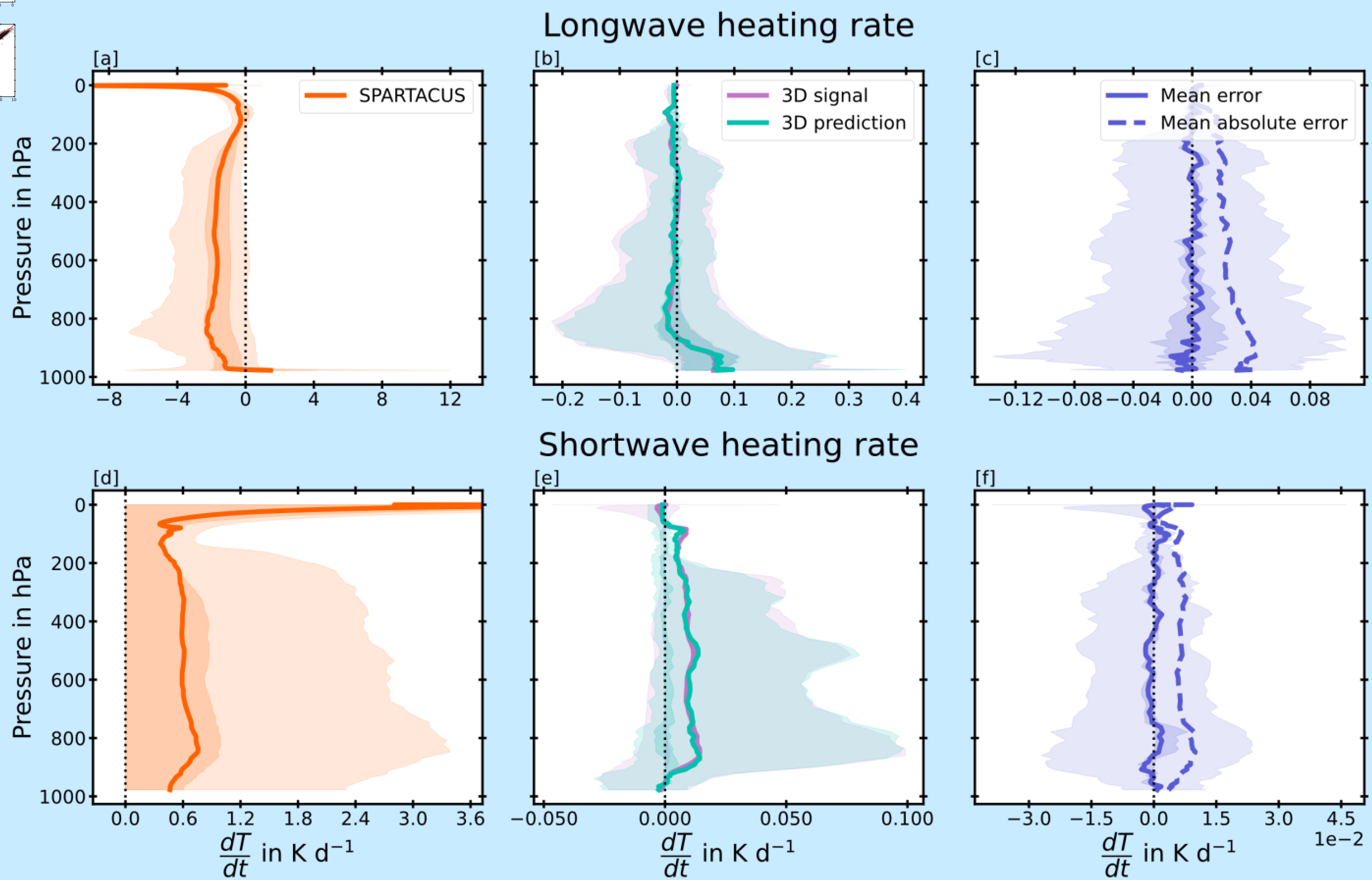
Unaffordable in routine applications, but the 3D-1D difference is quite learnable

Hogan et al 2016, 10.1002/2016JD024875



Meyer et al 2021, almost-accepted in JAMES





Meyer et al 2021, almost-accepted in JAMES

Learning corrections

This solves the affordable approximate equations and corrects for unaffordable terms

Data for the correction was generated by a fast-*ish* low-order model including O(10%) 3D impacts

To two parameters describing horizontal and vertical structure, SPARTACUS adds a single parameter to characterize 3D impacts… more empiricism, but quasi-accessible in data or high-resolution simulations

Calibration and structural error

JAMES | Journal of Advances in Modeling Earth Systems

RESEARCH ARTICLE
10.1029/2020MS002423

Process-Based Climate Model Development Harnessing Machine Learning: III. The Representation of Cumulus Geometry and Their 3D Radiative Effects

Najda Villefranque[1], Stéphane Blanco[2], Fleur Couvreux[1], Richard Fournier[2], Jacques Gautrais[3], Robin J. Hogan[4], Frédéric Hourdin[5], Victoria Volodina[6], and Daniel Williamson[6,7]

Calibration using cloud structure from LES simulations of shallow clouds, using relatively sparse benchmark calculations with variable solar zenith angles

Emulation of SPARTACUS predictions of as a function of SPARTACUS parameters

Sampling to identify the parameter values that minimize SPARTACUS errors with respect to the benchmark

Calibration and structural error

Table 3
Parameter Values for the "Best" Configurations of ecRad

Parameters	FSD	z_0 (m)	C_s (m)
Mean LES–derived	0.705	187	247
Best global	1.079	436	155
Best TOA up	1.646	493	119
Best absorption	0.102	294	821
Best surface down	1.469	374	113

Villefranque et al 2020, 10.1029/2020MS002423

Calibration and structural error

JAMES Journal of Advances in Modeling Earth Systems

RESEARCH ARTICLE
10.1029/2020MS002423

Process-Based Climate Model Development Harnessing Machine Learning: III. The Representation of Cumulus Geometry and Their 3D Radiative Effects

Najda Villefranque[1], Stéphane Blanco[2], Fleur Couvreux[1], Richard Fournier[2], Jacques Gautrais[3], Robin J. Hogan[4], Frédéric Hourdin[5], Victoria Volodina[6], and Daniel Williamson[6,7]

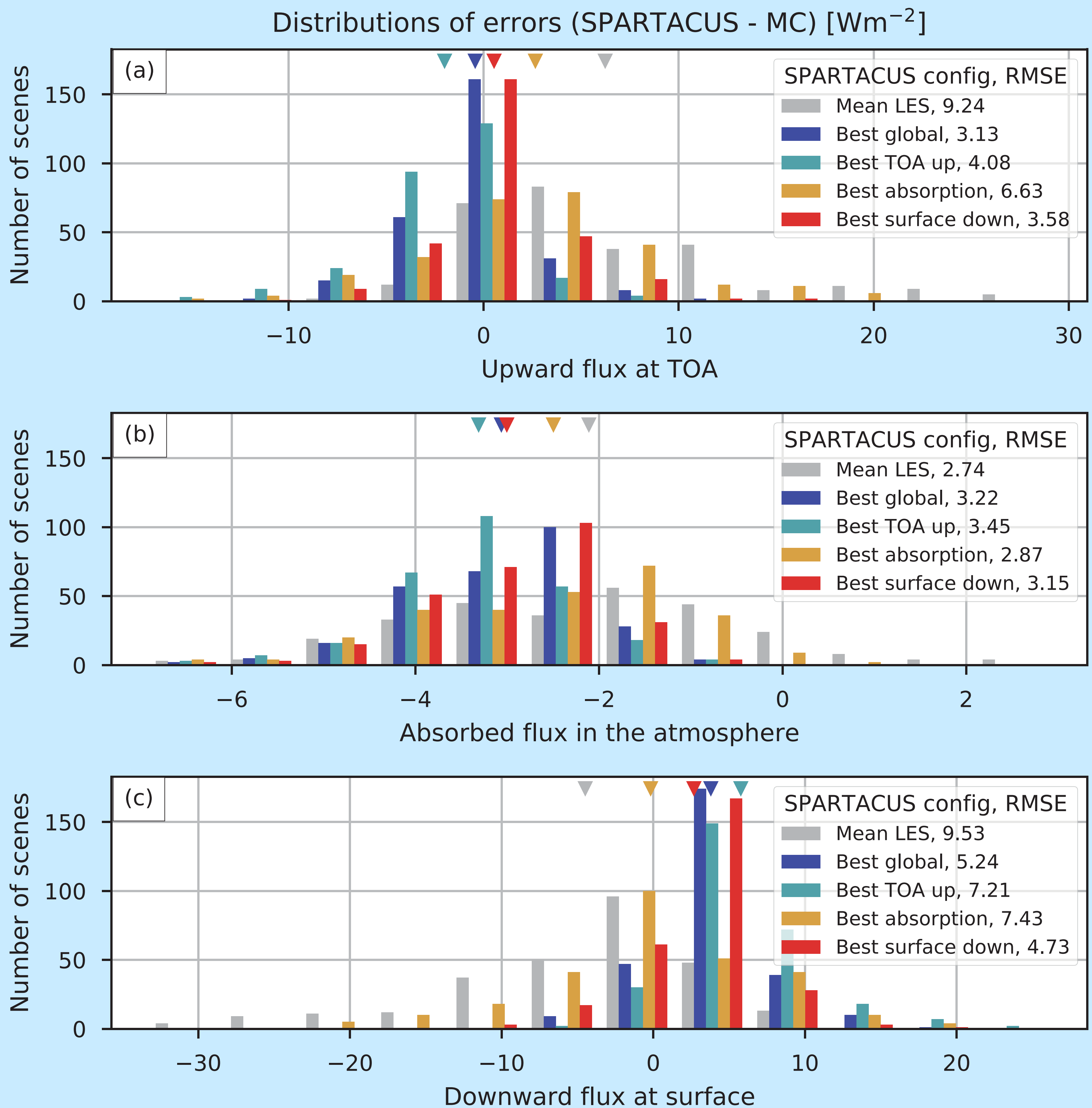
Calibration for LES simulations of shallow clouds, using relatively sparse benchmark calculations with variable solar zenith angles

Emulation of SPARTACUS predictions of as a function of SPARTACUS parameters

Sampling to identify the parameter values that minimize SPARTACUS errors with respect to the benchmark

Assessment against out-of-sample data

Calibration and structural error



Villefranque et al 2020, 10.1029/2020MS002423

Learning when data is hard to come by

There's no obvious shortcut to understanding 3D effects in place

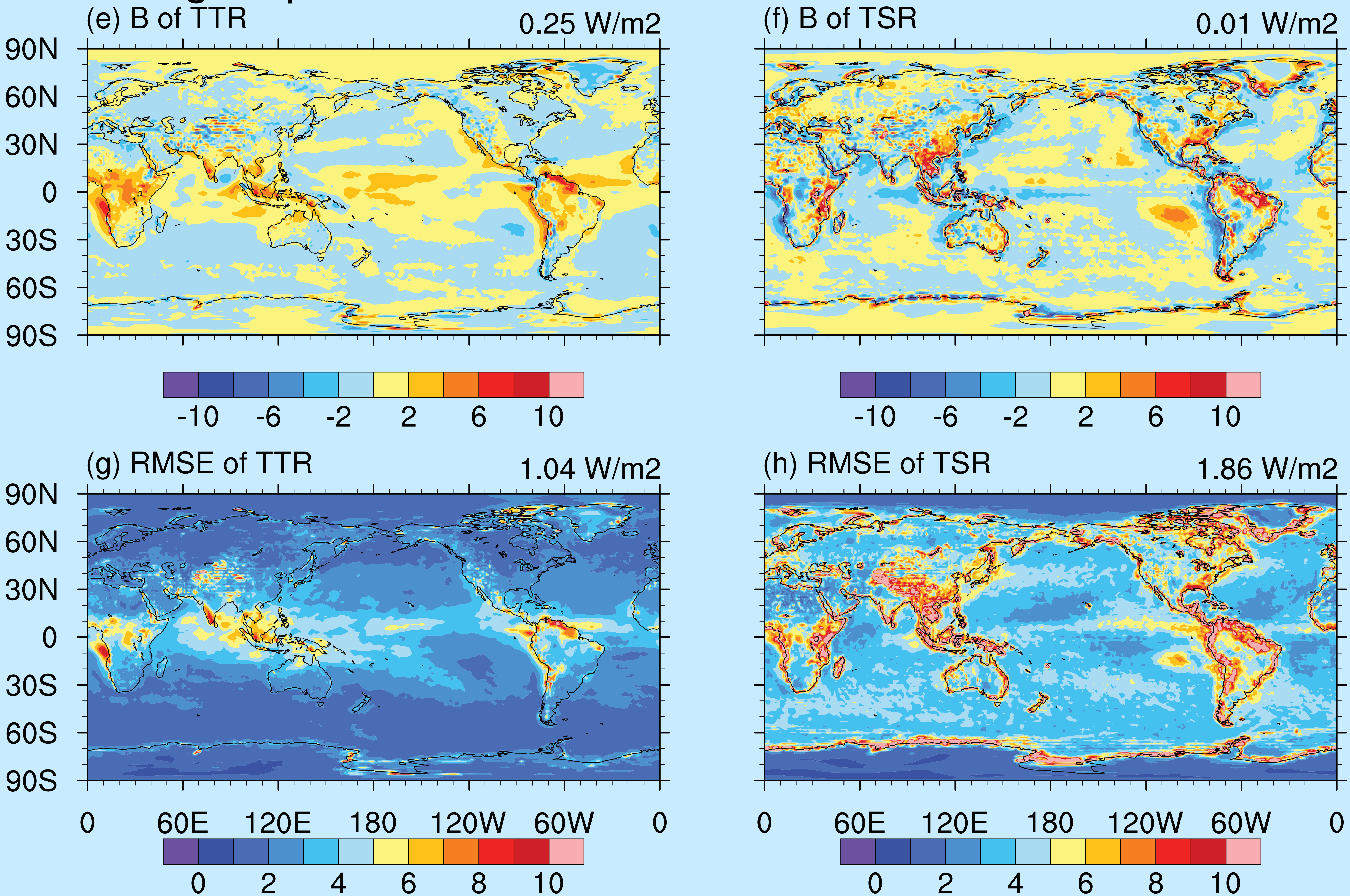
Data needed to learn the those effects directly is likely to remain sparse

expensive high-resolution clouds scenes x expensive radiative transfer simulations

Feature identification (which aspects of the scenes control the magnitude of the 3D impact?) could refine data generation strategy

But a large scale model can't know the small-scale state; predicting features controlling 3D impacts is a separate problem

Learning the problem and the solution



Zhu et al. 2019, doi:10.1029/2018JD029223

Machine learning for radiation

The opportunities for machine learning to inform the modeling of radiation include

representing empirical knowledge

providing corrections to approximations

The most challenging problem is data limited; it's not yet clear how best to frame the problem to generalize well

(For local participants: come talk to me about dimension reduction and equation discovery problems)