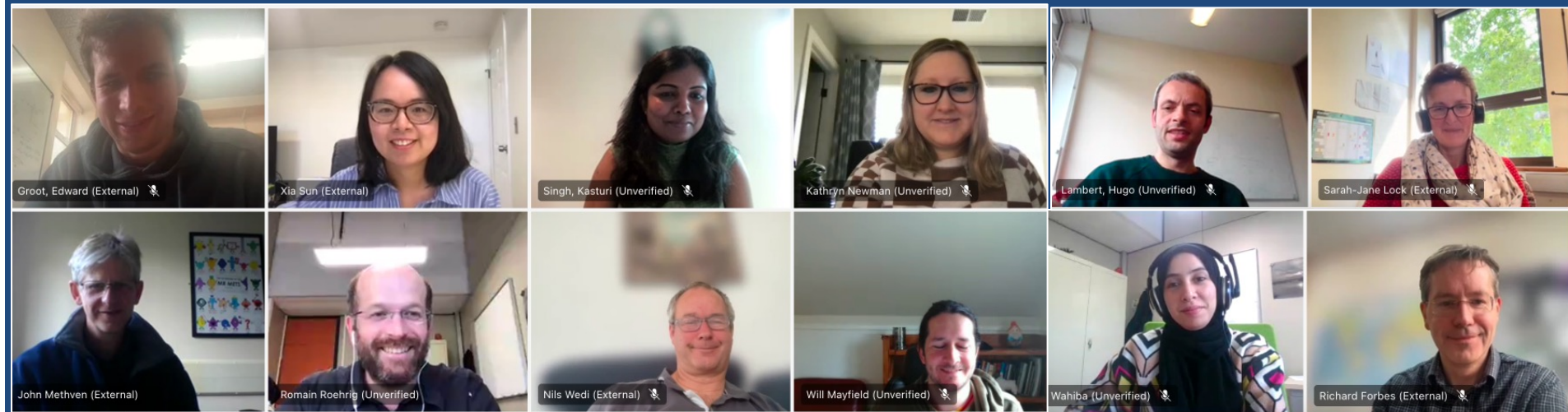


Model Uncertainty – MIP

Constraining stochastic parametrisations using high resolution simulations

Hannah Christensen, University of Oxford



How should we represent model uncertainty?

How should we represent model uncertainty?

- Why MY chosen **stochastic parametrisation**?
 - Practical performance vs. theoretical underpinning
 - Should every model use the same scheme?
 - Different modeling assumptions
 - Different resolutions
 - Different regional foci
 - To what extent can one scheme 'mop up' all uncertainty
 - Do we need multiple schemes or will this lead to double counting?

How should we represent model uncertainty?

- Why MY chosen **stochastic parametrisation**?
 - Practical performance vs. theoretical underpinning
 - Should every model use the same scheme?
 - Different modeling assumptions
 - Different resolutions
 - Different regional foci
 - To what extent can one scheme ‘mop up’ all uncertainty,
 - Do we need multiple schemes or will this lead to double counting?
- What about the climate modelling community?
 - Do **multiple models**/parametrisations span model uncertainty?
 - How much of structural errors can we attribute to poor **tuning**?
 - How can **stochastic** approaches complement multi-physics?

How can we begin to answer these questions?

→ **Require a large database of model error**

- For different models
- For different global regions
- For different seasons
- For different model resolutions

Ideally accompanied by

- Information on model parametrised tendencies

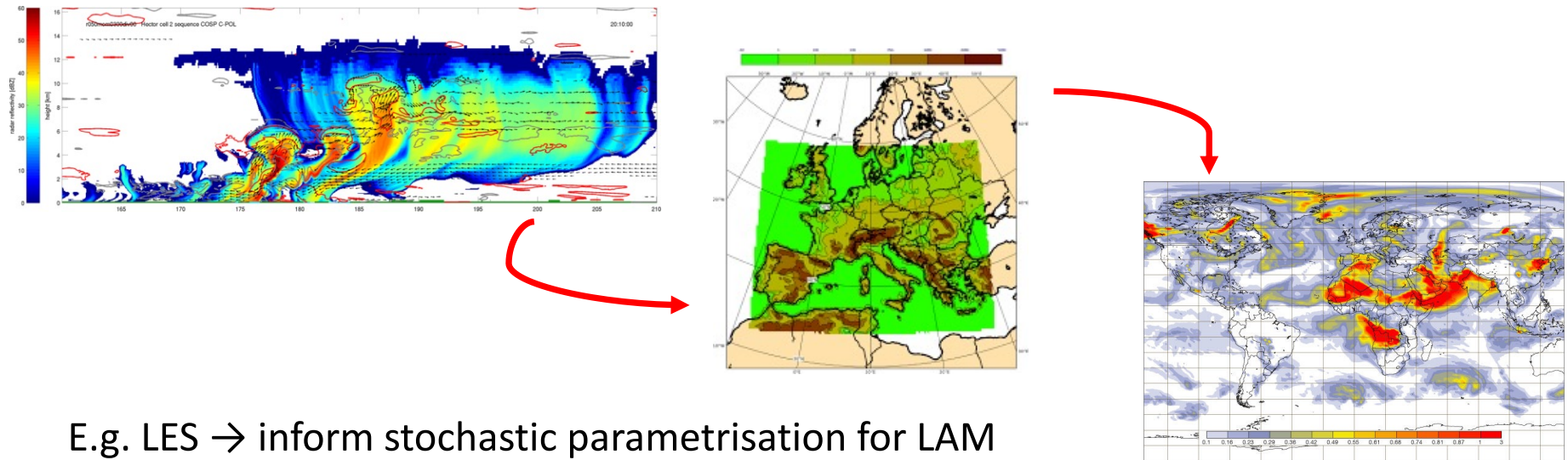
Model Uncertainty - MIP

- Joint initiative of WWRP's PDEF and WCRP's WGNE
 - model error identification
 - Aim: intercomparison of random and systematic error characteristics across many models
- Will provide new database of model error
- **Funding secured to support work**
 - **NCAR/NOAA DTC** – June 2021-June 2025
 - **Leverhulme Trust**: Oxford (with ECMWF), Exeter (with UK Met Office), Météo France – Sept 2023-Sept 2026



Use a high resolution simulation as ‘truth’

- Use high resolution simulation, that resolves process of interest, to inform parametrisation at lower resolution
- “nature run” should be extended in space and time -> learn spatio-temporal correlations needed by stochastic parametrisations



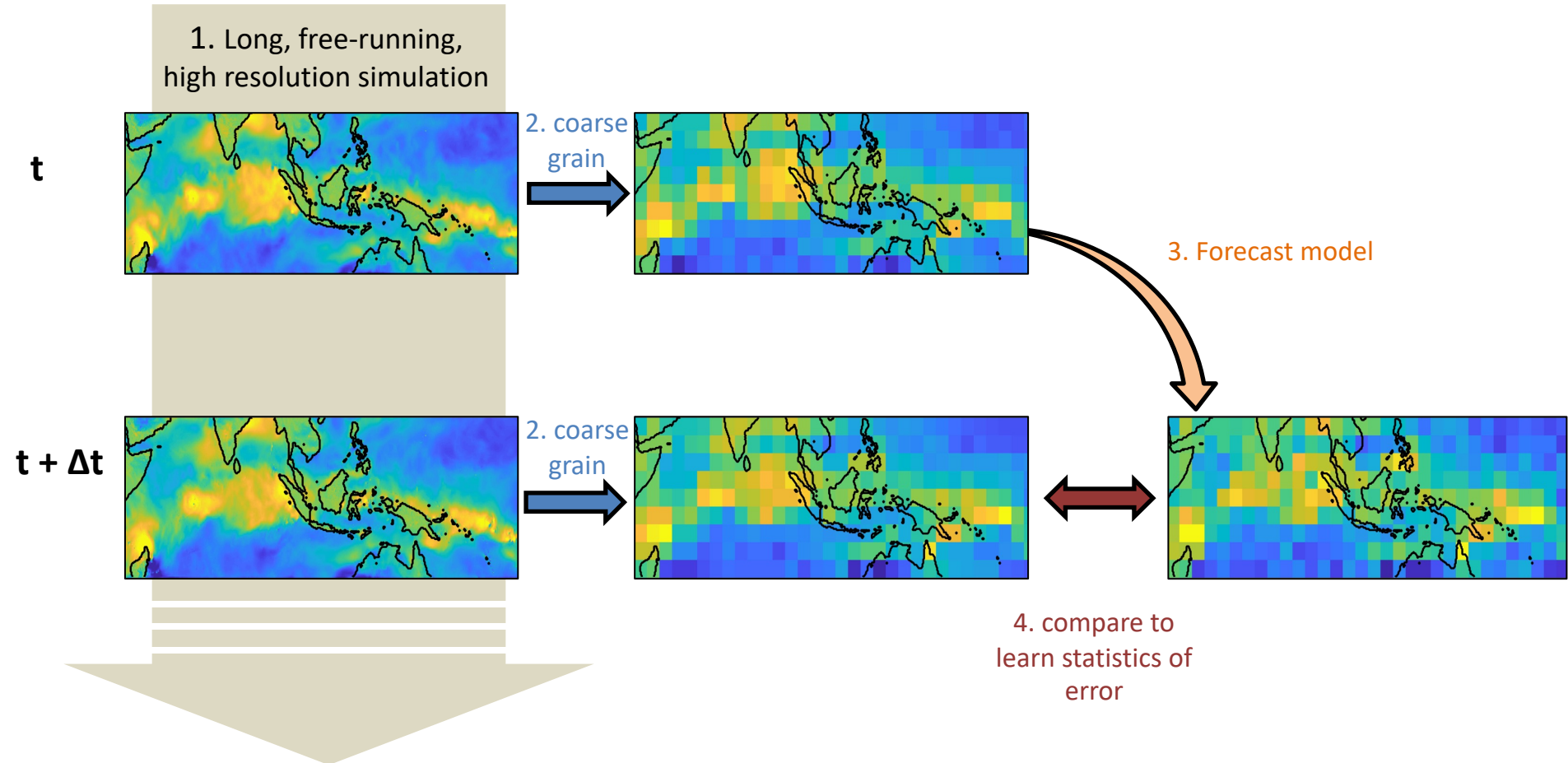
E.g. LES → inform stochastic parametrisation for LAM

E.g. convection permitting LAM → inform stochastic parametrisation for global EPS

- **Caveat: the high resolution benchmark is not the real atmosphere**

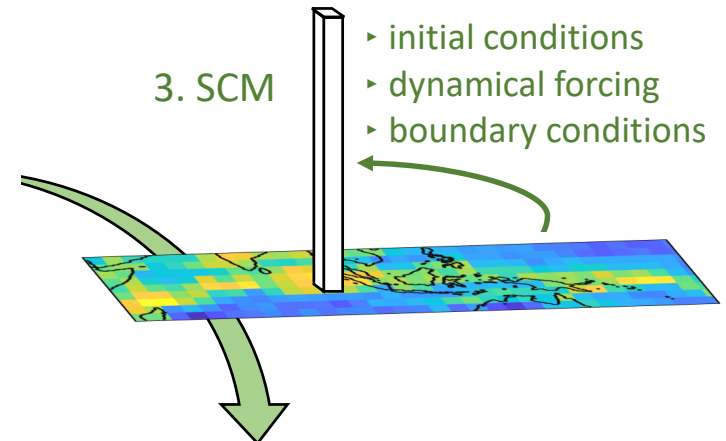
- Compare model to multiple benchmarks, including those from other models

Use a high resolution simulation as 'truth'



Single Column Model (SCM) as Forecast Model

- A SCM consists of:
 - subgrid parametrisations from parent model
 - forced with dynamical tendencies
- How do we use an SCM?
 - Use coarsened high-res simulation to prescribe **Initial conditions, Dynamical forcing** and **Boundary conditions**
- Benefits of using SCM?
 - Supply dynamical tendencies targets uncertainty in the parametrisation schemes
 - SCM portable and cheap
 - Tile many SCM to cover domain
 - New DEPHY format facilitates sharing of SCM driver files



First MUMIP experiment: Indian Ocean domain

- Last 30 days of **ICON 2.5km Dyamond Summer** simulation
- Data archived 3 hourly
- CG to resolution of 0.2 degrees (~22 km)
- Domain in Indian Ocean: (51-95E, 5N-35S)
- See website: mumip.web.ox.ac.uk

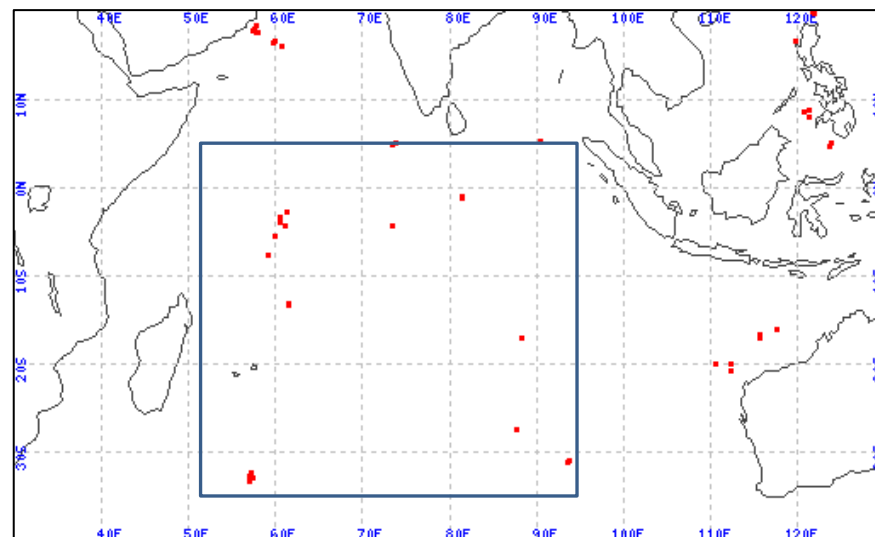
Rapid progress over recent months!

Complete datasets from four models!

- CCPP SCM + Global Forecast System (GFSv17) physics (X. Sun, NOAA/DTC, K. Newman, NCAR/DTC)
- CCPP SCM + Rapid Refresh (RAP) physics (X. Sun NOAA/DTC, K. Newman, NCAR/DTC)
- ECMWF OpenIFS CY48r1 (E. Groot, U. Oxford)
- Météo France ARPEGE-Climat SCM (W. Lfarh, Météo France)

UK Met Office runs in production, available shortly

(K. Singh, Univ Exeter)



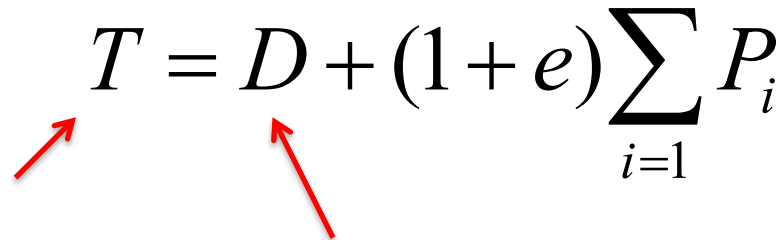
Analysing the data: multiplicative noise?

SPPT:

$$T = D + (1 + e) \sum_{i=1} P_i$$

Analysing the data: multiplicative noise?

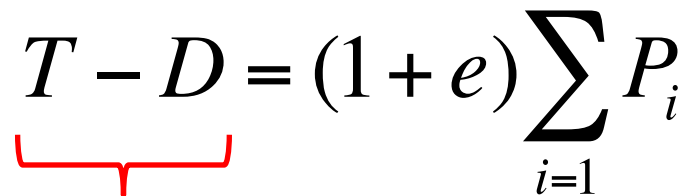
SPPT:

$$T = D + (1 + e) \sum_{i=1} P_i$$


Calculate 'true' total
tendency from coarsened
ICON

Assume SCM dynamics
tendency is 'correct'

Assess error in SCM
physics tendencies

$$T - D = (1 + e) \sum_{i=1} P_i$$


Compare 'true'
physics tendency ...

... to parametrised
physics tendency

Analysing the data: multiplicative noise?

SPPT:

$$T = D + (1 + e) \sum_{i=1} P_i$$

Calculate 'true' total tendency from coarsened ICON

Assume SCM dynamics tendency is 'correct'

Assess error in SCM physics tendencies

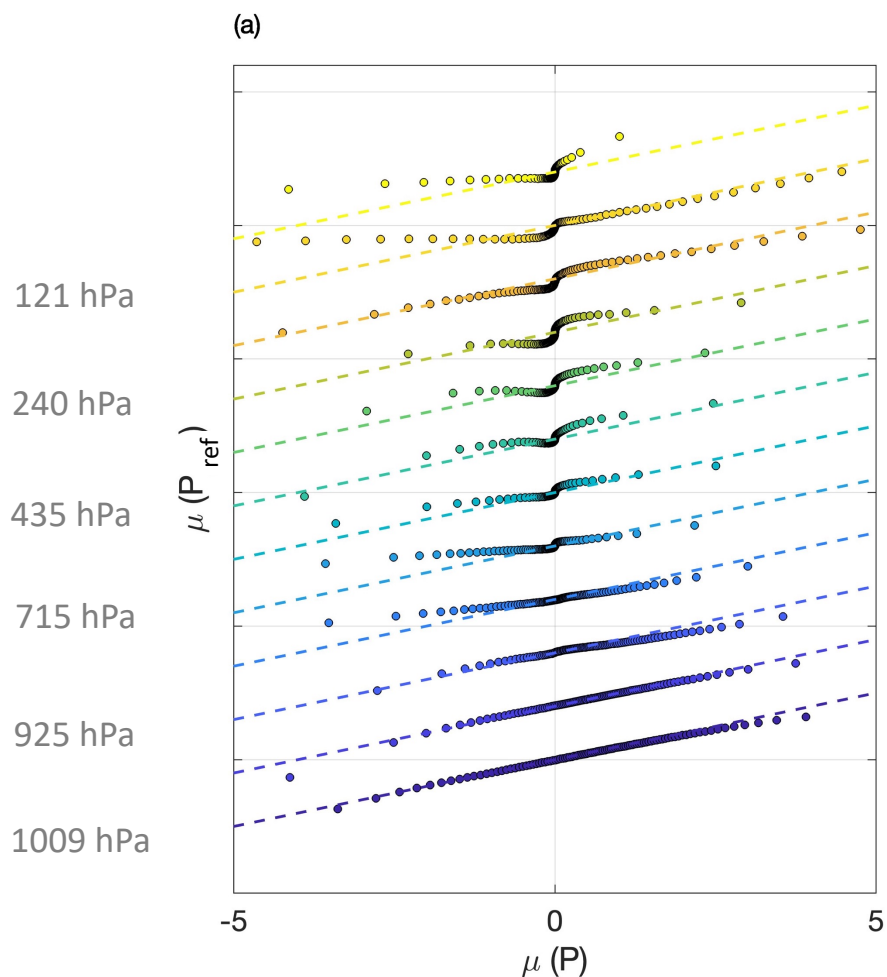
$$T - D = (1 + e) \sum_{i=1} P_i$$

Compare 'true' physics tendency ...

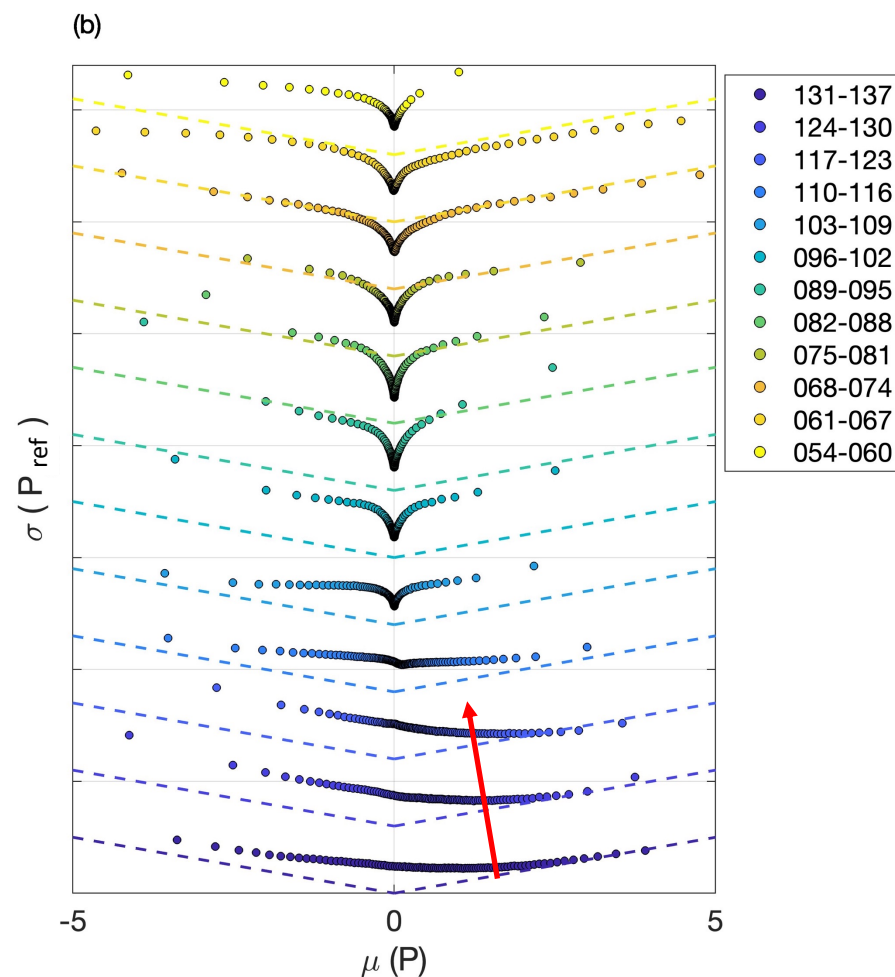
... to parametrised physics tendency

Very preliminary analysis so far!
- 3 models
- 5 days simulation considered:
= 1.8 million simulations

mean and standard deviation of reference conditioned on SCM prediction

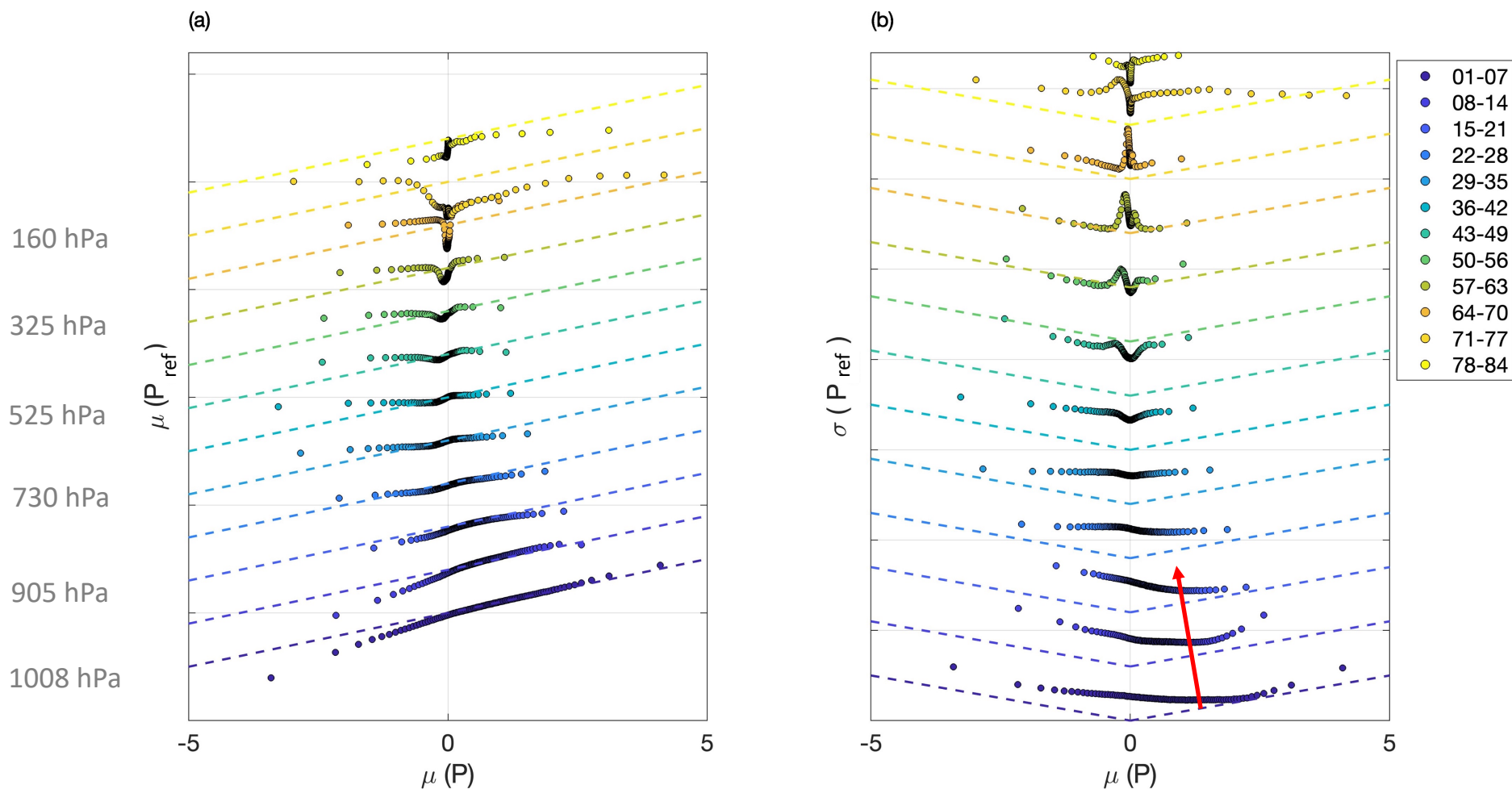


Well calibrated, especially near surface



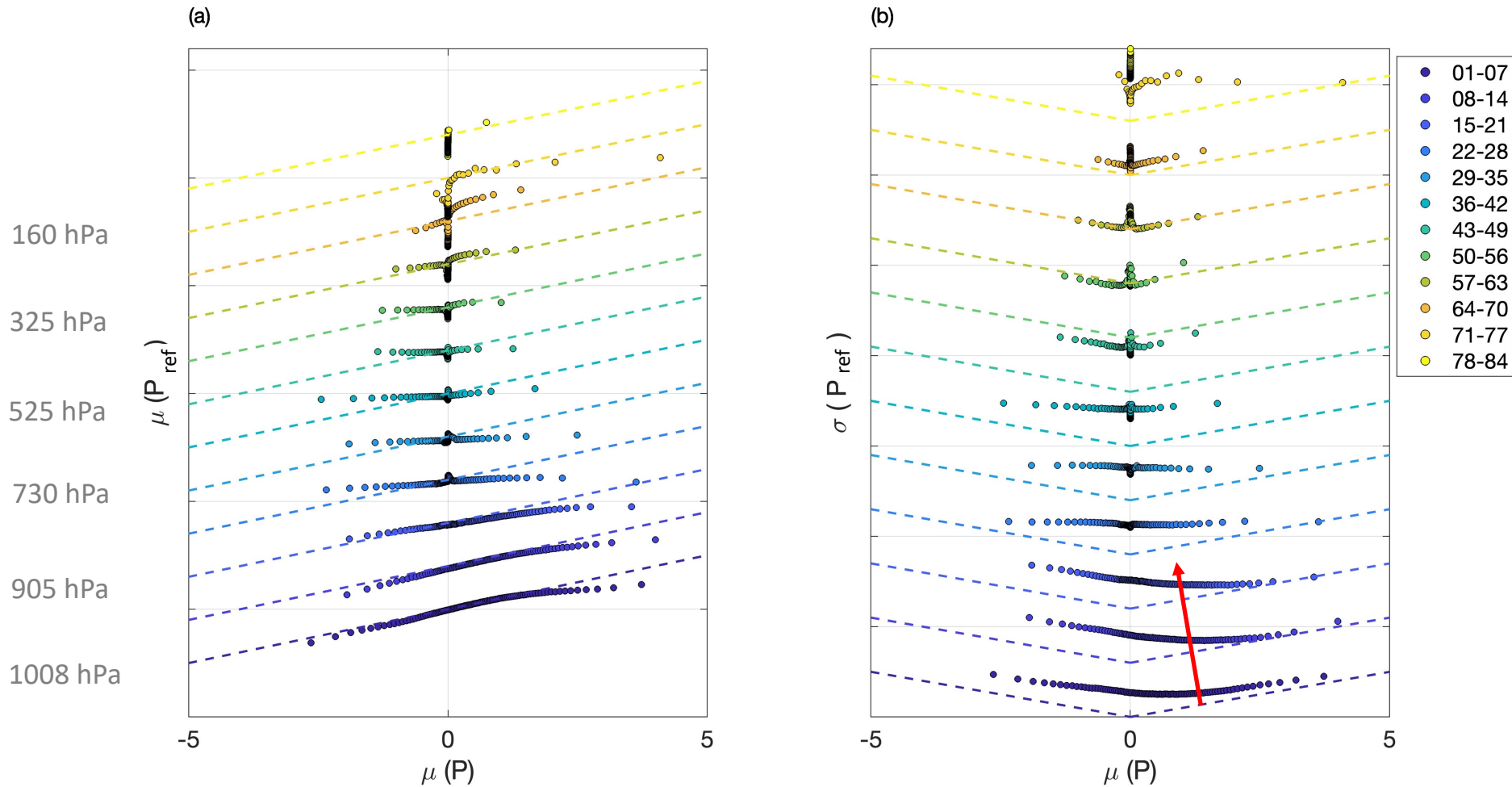
Uncertainty minimum

mean and standard deviation of reference conditioned on SCM prediction



Uncertainty minimum

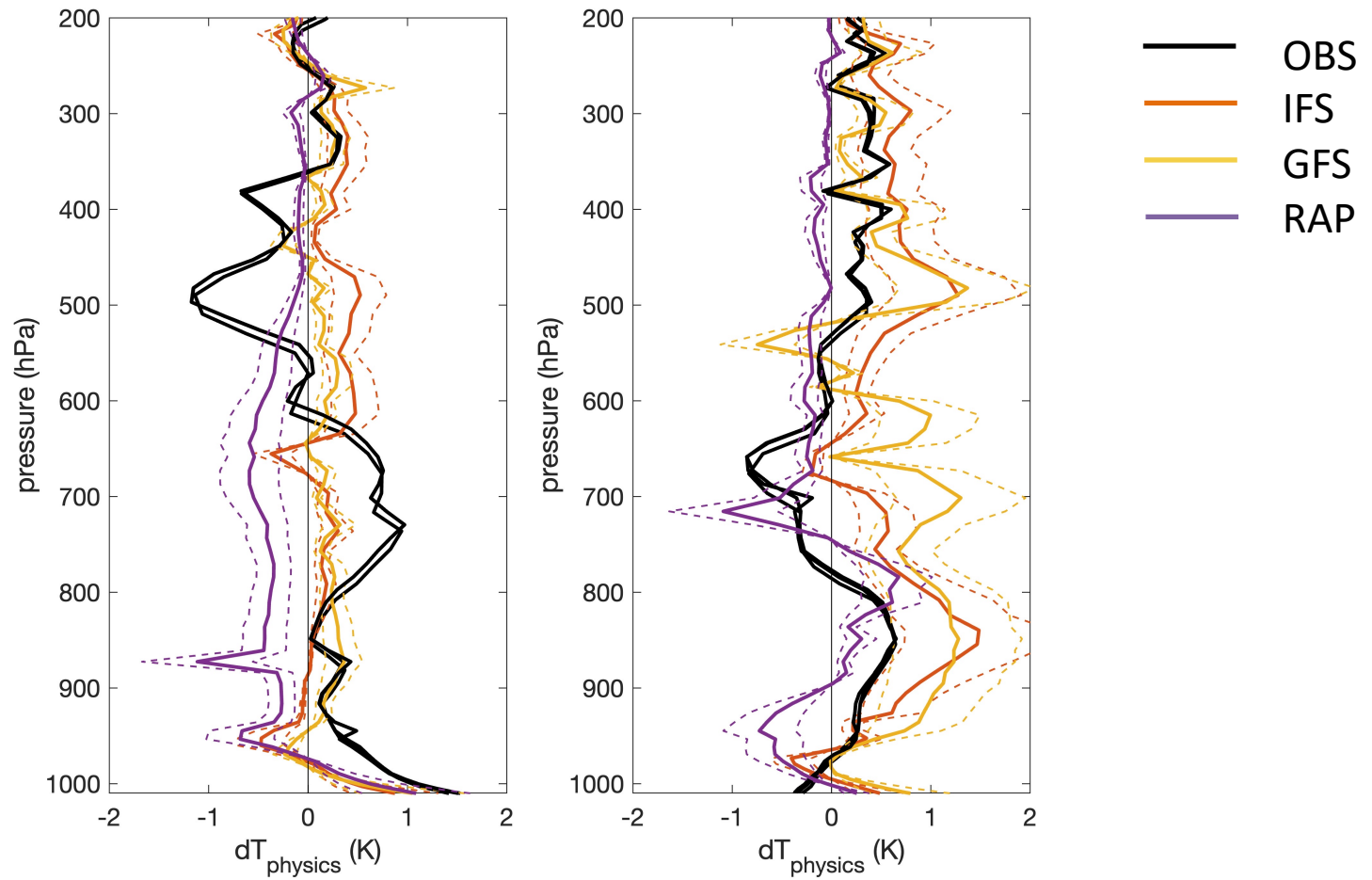
mean and standard deviation of reference conditioned on SCM prediction



Uncertainty minimum

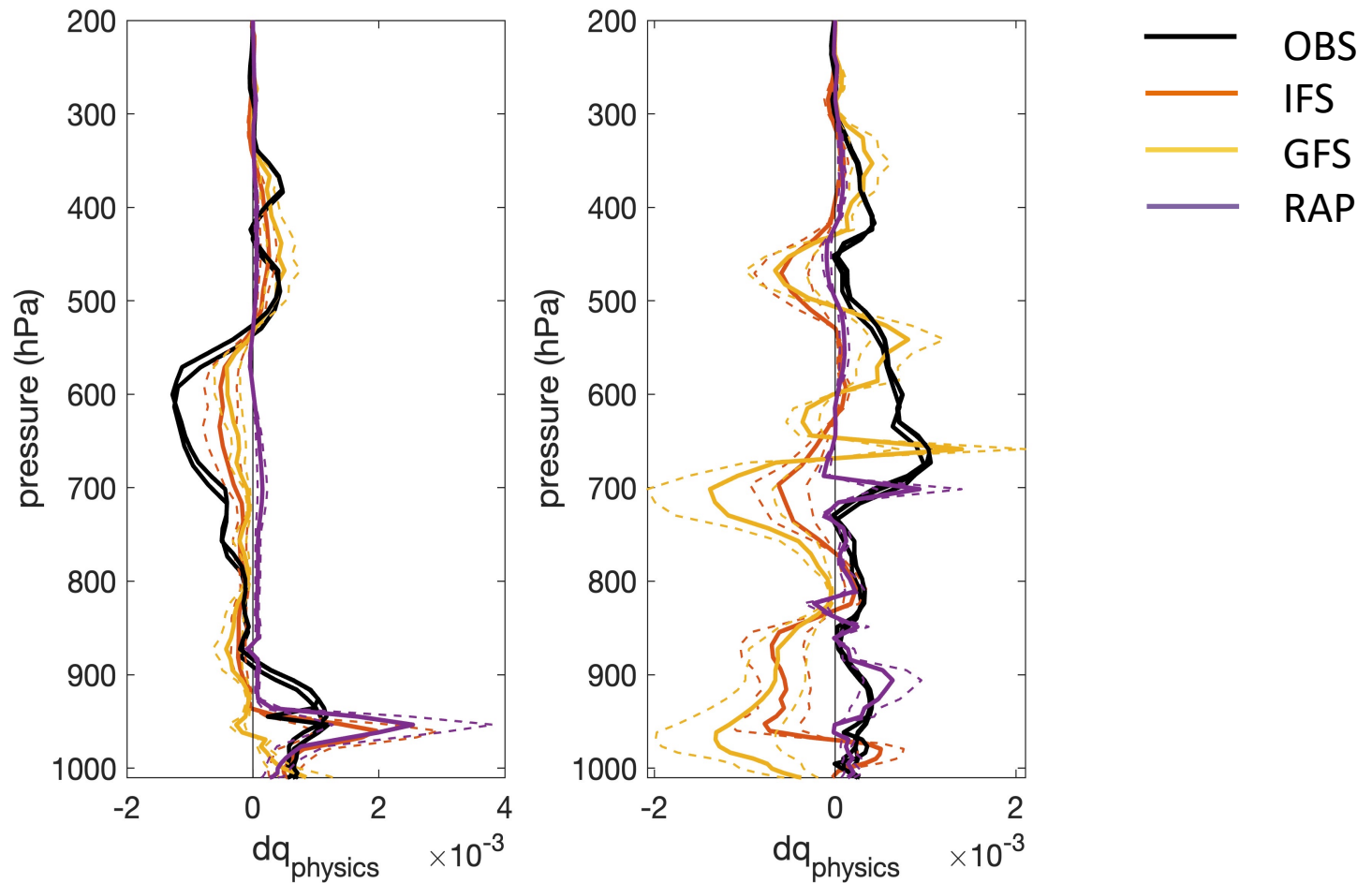
Multi-model vs SPPT

- Consider vertical profiles of physics T tendency in two sample columns



Multi-model vs SPPT

- Consider vertical profiles of physics q tendency in two sample columns



MU-MIP plans

- Assess structural error across different models
 - Exeter lead: H. Lambert and K. Singh
- Assess parametric uncertainty
 - Meteo France lead: R. Roehrig and W. Lfarh
- Assess random error, foundations of stochastic parametrisations
 - Oxford lead: H. Christensen and E. Groot

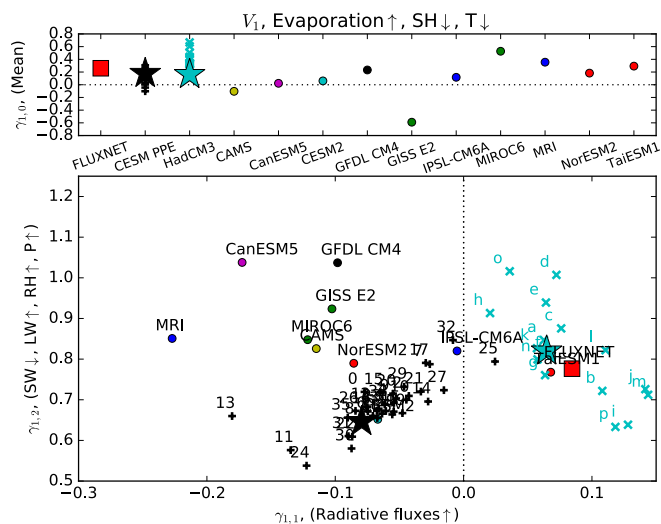


Fig: Hugo Lambert

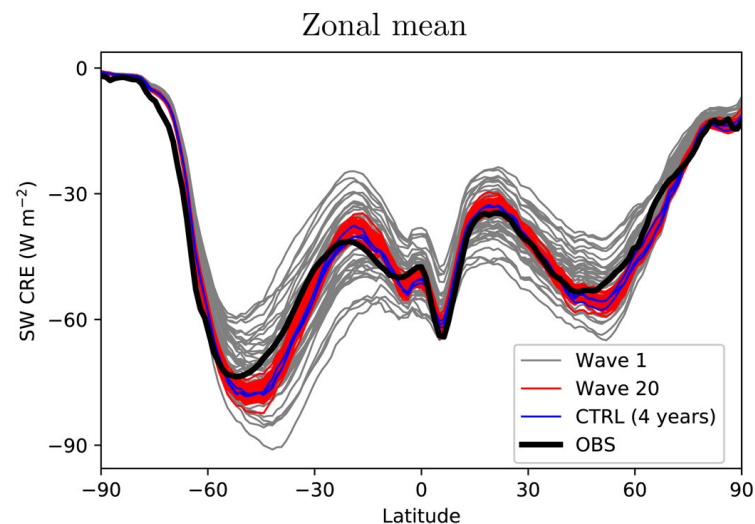


Fig: Hourdin, ..., Roehrig, et al, 2020, JAMES

Please see Edward Groot's poster for more analysis


Model Uncertainty-Model Intercomparison Project (MU-MIP)

A dataset for model physics intercomparison and model error

Linking CAPE and precipitation rate non-stationarity during spin-up across NWP physics packages

Edward Groot¹, and MU-MIP core team: Hannah Christensen¹, Xia Sun², Kathryn Newman², Wahiba Lfarh³, Kasturi Singh^{4,5}, Romain Roehrig³, Hugo Lambert⁴, Jeff Beck² and Keith Williams⁵

1: University of Oxford; 2: Developmental Testbed Center; 3: Meteo France; 4: University of Exeter; 5: UK MetOffice




Contact: edward.groot@physics.ox.ac.uk

mump.web.ox.ac.uk

Introduction

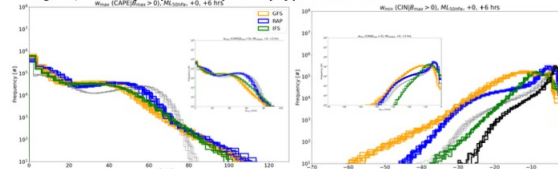
MU-MIP is an intercomparison project for model uncertainty in which we intercompare the physics parameterization suites used in numerical weather and climate modelling. Each physics suite consists of a package of parameterizations, e.g. for turbulence, convection, radiation, surface exchange with land/ocean and cloud processes. These are thought to be the dominant contribution to model uncertainty across all GCMs and numerical weather prediction models.

We run the simulations with parameterization suites by utilizing the single-column version of operational models (SCM) over the Indian Ocean domain about ten million times. To ensure fixed and representative dynamical constraints, we assume a ground truth derived from DYAMOND simulations and insert its dynamics as initial and boundary conditions in the SCMs. One month of 2016 is covered based on the storm-resolving ICON ($\Delta X = 2.5$ km) and driven by three-hourly archived dynamics.

After re-gridding to 0.2 degrees, we currently carry out an ocean-only intercomparison over a subdomain of 44,000 tiles. Two parameterization packages from the Developmental Testbed Center at NCAR have been utilized: RAP and GFS (version 17).

The OpenIFS-SCM dataset with cycle 48 physics is near completion (as of August 2024) and MeteoFrance and UK MetOffice/University of Exeter will follow.

Throughout, we strive for optimal comparability of parameterization suites.

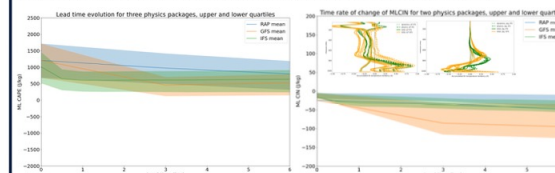


PDF of Mixed-layer (ML) values of $\sqrt{2}$ CAPE and $\sqrt{2}$ CIN from three parameterization suites on a log-axis at 6-hour lead time (enclosed: 3 hours). The lines indicate PDFs across variation across the diurnal cycle. Grey: ground truth (ICON 2.5km-derived conditions prescribed as initials) Right, black: same ground truth MLCIN following slightly different IFS levels define the ML (further investigation needed)

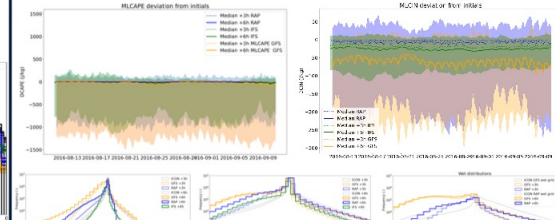
We intercompare conditional PDFs of tendencies and the model state to learn about multi-model uncertainty, eventually at benefit of stochastic perturbation schemes (Christensen, 2020, QJRMS).



How can we learn about model physics, and what the numerical models do?



Left: mixed-layer CAPE as function of lead time for the ITCZ band; right: mixed-layer CIN as a function of lead time for the entire domain, with enclosed net physics and dynamics tendencies of IFS across the diurnal cycle. Below: time evolution of 3 and 6-hour change of mixed-layer CAPE (left) and CIN (right) over the full domain.



(Conditional) PDFs of mixed-layer CIN change (left) and CAPE change (center) at 6 (3) hours lead time across various suites; left and center: full PDF; right: conditioned PDF for mixed-layer CAPE over precipitating parts of the grid (at least 1.7 mm per gh). GFS has much more of such precipitating cells than RAP.

Convective adjustment from a model's non-native regime is linked to precipitation intensity, which could link to non-stationarity in parameterized precipitation rates manifested by parameterized deep-convection in ERA5 (Buschow, 2024, QJRMS).



mumip.web.ox.ac.uk



Model Uncertainty - MIP



[Home](#) [About](#) [People](#) [Resources](#) [Stochastic Parametrisation](#)

Home

Welcome to the Model Uncertainty - Model Intercomparison Project (MUMIP)

An initiative of the WCRP Working Group for Numerical Experimentation and the WWRP Predictability, Dynamics and Ensemble Forecasting Working Group

Introduction

MU-MIP is an international project which seeks to characterise systematic and random component of model error across many different climate models. This is the first coordinated intercomparison of random model error, and will be used to inform stochastic parametrisation development.

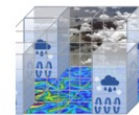
Some key questions:

- How should we best represent model uncertainty/random error using stochastic approaches?
- To what extent should this representation be model specific or a fundamental property of atmospheric models?
- Are current approaches justified? How can they be improved?
- Can a coarse-graining approach be used to validate and compare high-resolution simulations and their behaviour across scales?

Contact

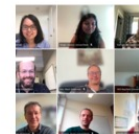
The MU-MIP team consists of scientists from 10+ institutes spanning three continents. Please get in touch by emailing Hannah Christensen on hannah.christensen 'at' dhvscs.ox.ac.uk if

News



New MUMIP data available for analysis

29 August 2024



MUMIP Meeting 5

29 April 2024



Leverhulme Trust Research Project team meeting 1

12 January 2024



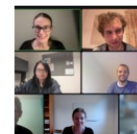
Oxford Workshop on Model Uncertainty, 23-26 Sept

10 June 2024



Leverhulme Trust Research Project Team Meeting 2

21 February 2024



MUMIP Meeting 4

4 October 2023

Thanks for listening

Hannah.Christensen@physics.ox.ac.uk

Extra Slides

Coarse graining details

1. Local area averaging for coarse graining

$$\bar{\psi}_{n,k} = \sum_i W_{n,i} \psi_{i,k}$$

2. Linearly interpolate in time

3. Vertical interpolation

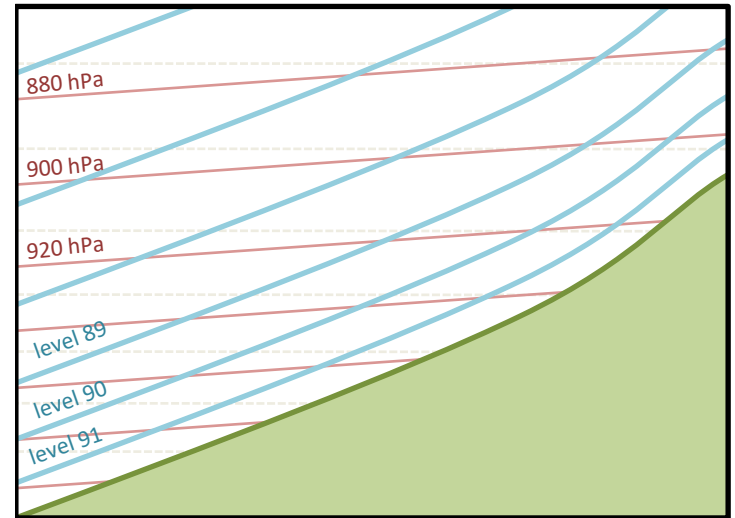
- Evaluate coarse-scale grid box mean p_{sfc}
- Coarse-grain other fields along model levels
- Interpolate from native model levels to target model levels

4. Above high-resolution model top, pad data using ECMWF analysis

5. Advective tendencies estimated from the coarsened fields

$$\text{adv}(\psi)|_{n,k} = -\bar{\mathbf{u}}_{n,k} \cdot \bar{\nabla}_k(\bar{\psi}_{n,k})$$

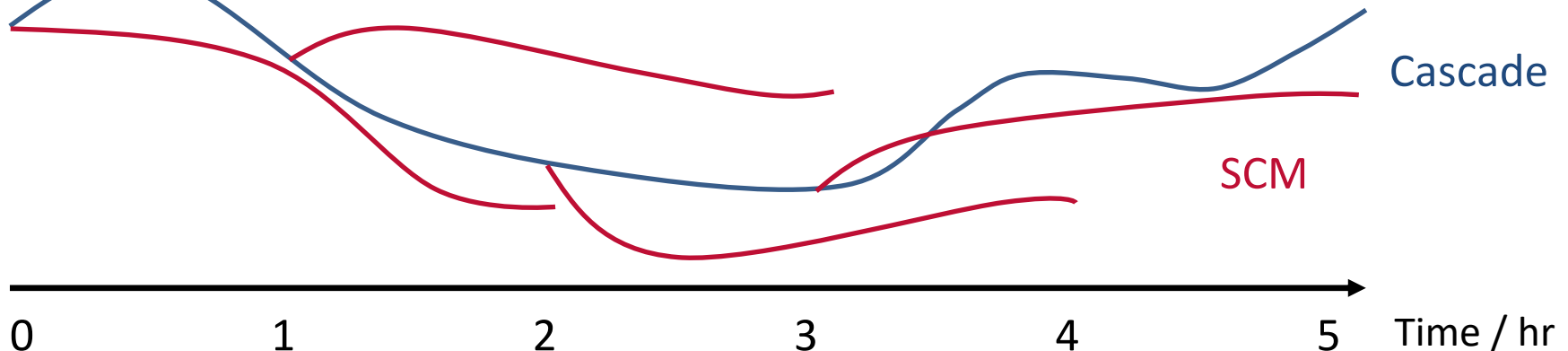
6. Specify sensible and latent heat fluxes from high-resolution dataset, but take static boundary conditions from operational ECMWF model at T639



What we do








- Coarse-grain Cascade to T_L639
- Run an independent SCM simulation, initialised every hour, from every lat-lon point ($>68,000$) in the coarse-grained domain
- Run each SCM simulation for two hours, discard the first hour to avoid focus on spin up
- Repeat for entire 9-day Cascade simulation

Initialise 2-hour SCM simulations every hour
Only consider 2nd hour of SCM forecast to avoid focus on spin-up

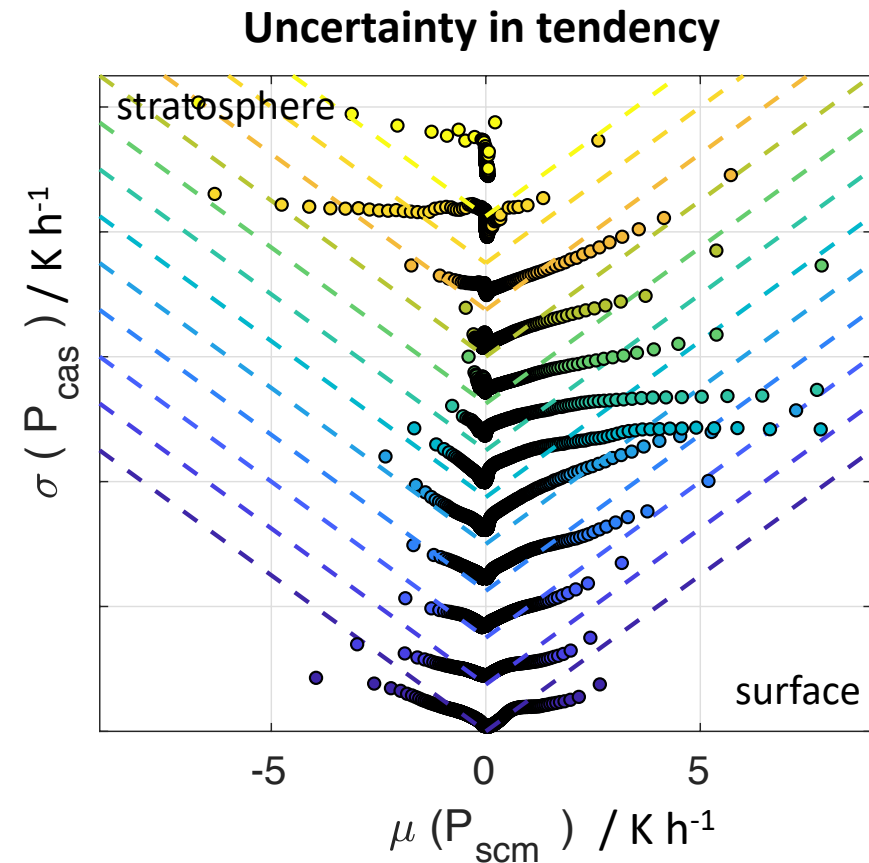
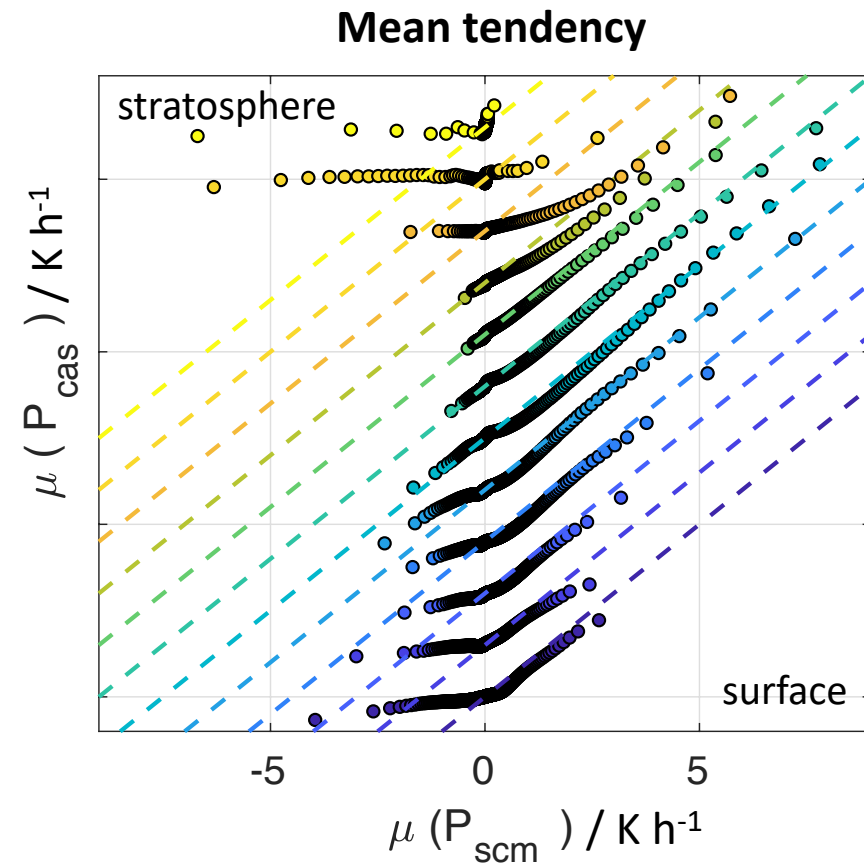


Cf. existing approaches to identify model error

- **E.g. Initial tendency approach** in which physics tendencies in data assimilation cycle are compared to the analysis
- **E.g. Transpose AMIP** in which climate models are run in weather forecasting mode from common initial conditions

	Initial tendency	Transpose AMIP	My SCM approach
Decompose model evolution (& error) into single processes			
No data assimilation capabilities needed to evaluate forecast model			
Comparison of model with its native analysis may mask errors			
Inconsistencies in IC can lead to systematic drifts			

Consider T tendency



Data grouped by level.

Dark blue: levels 91—87 (ground—995 hPa)

Yellow: levels 32—36 (86—60 hPa)