Model Uncertainty – MIP

Constraining stochastic parametrisations using high resolution simulations

Hannah Christensen

hannah.christensen@physics.ox.ac.uk

Atmospheric, Oceanic and Planetary Physics, Univ. Oxford

ECMWF Workshop on Model Uncertainty, 9 – 12 May 2022

Why do we do what we do*?

*Regarding model uncertainty representation

Why do we do what we do*?

*Regarding model uncertainty representation

- 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?

Why do we do what we do*?

*Regarding model uncertainty representation

- 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 parametrisations/models span model uncertainty?
  - How much of structural errors can we attribute to poor tuning?
  - How can stochastic approaches complement (replace?) 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

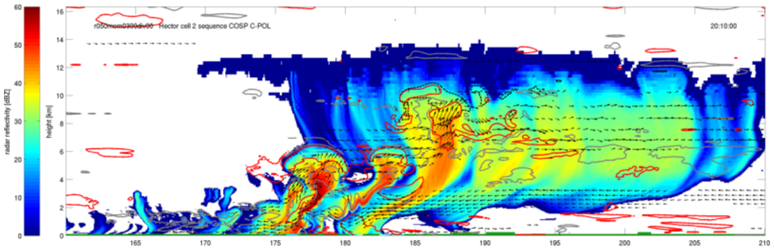
Model Uncertainty - MIP

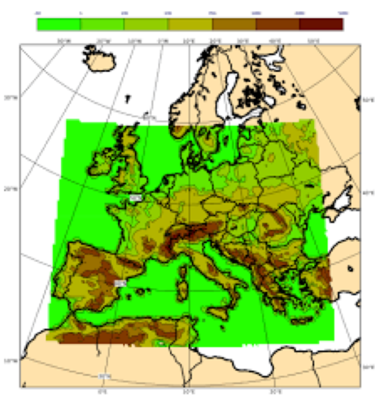
- Joint initiative of WWRP's PDEF and WCRP's WGNE
  - Primary joint interest of the two working groups is model error identification
  - Aim: intercomparison of random and systematic error characteristics across many models
- Will provide new database of model error

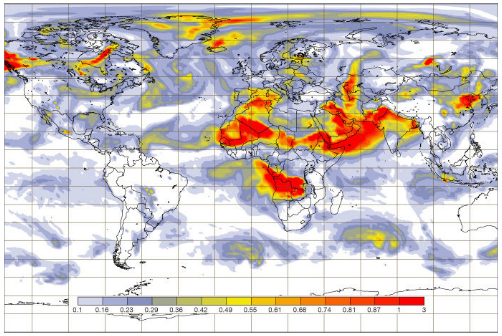
https://mumip.web.ox.ac.uk

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







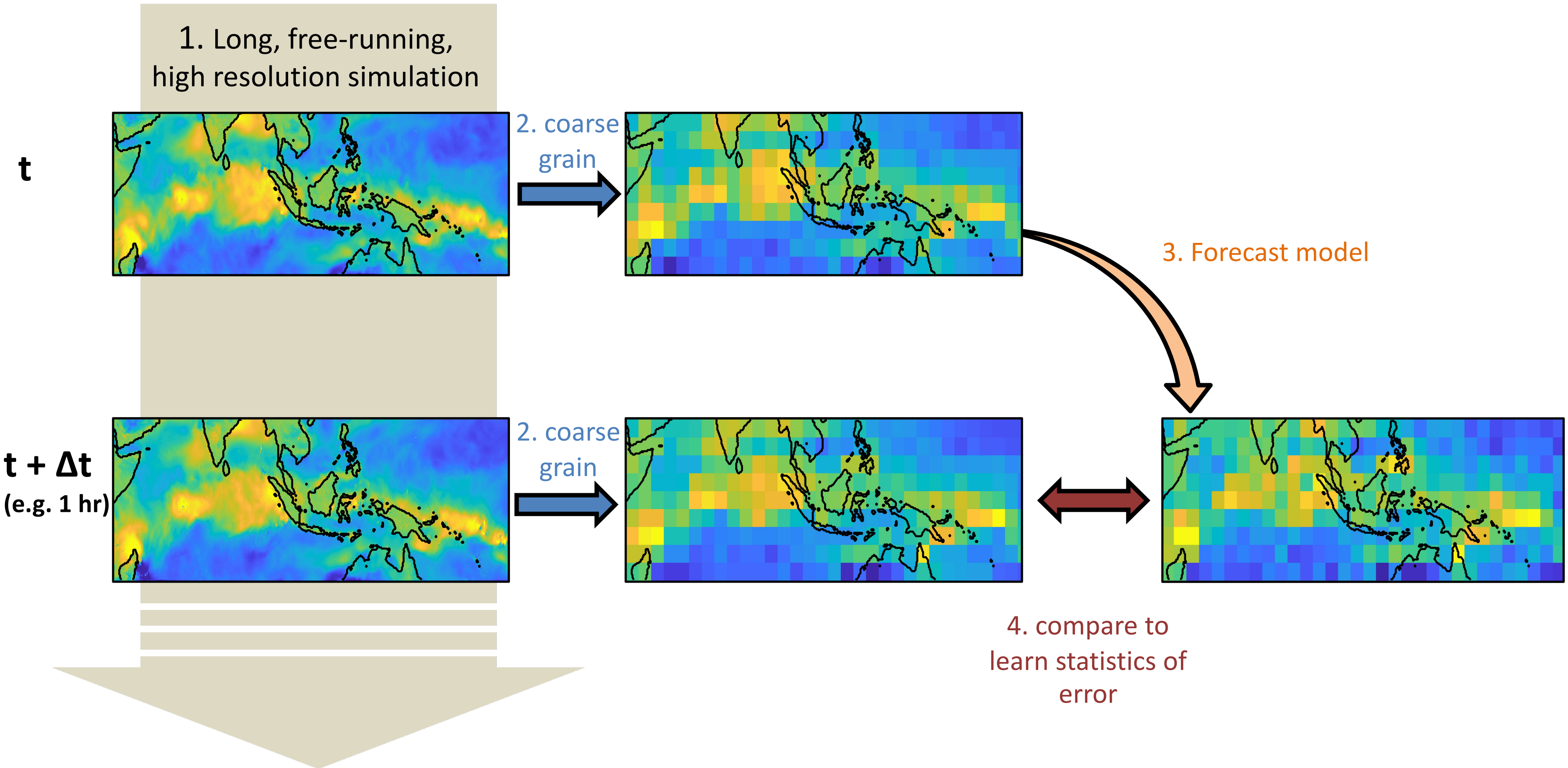
E.g. LES → inform stochastic parametrisation for LAM

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

Image credits: (L) M. Herzog (C) COSMO (R) ECMWF

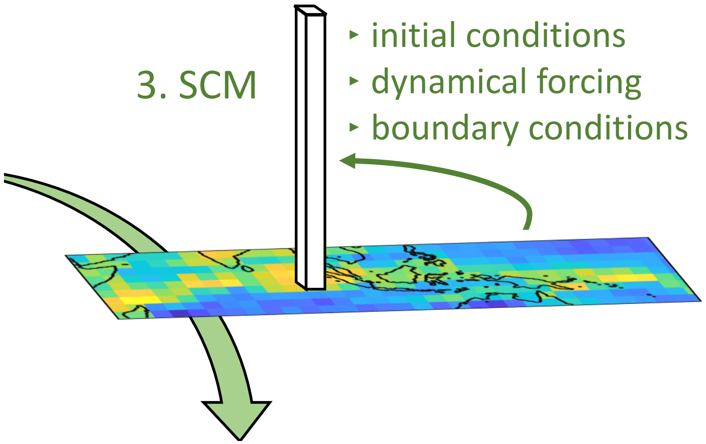
Use a high resolution simulation as ‘truth’



Christensen et al, 2018, JAMES.
Christensen, 2020, QJRMS

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

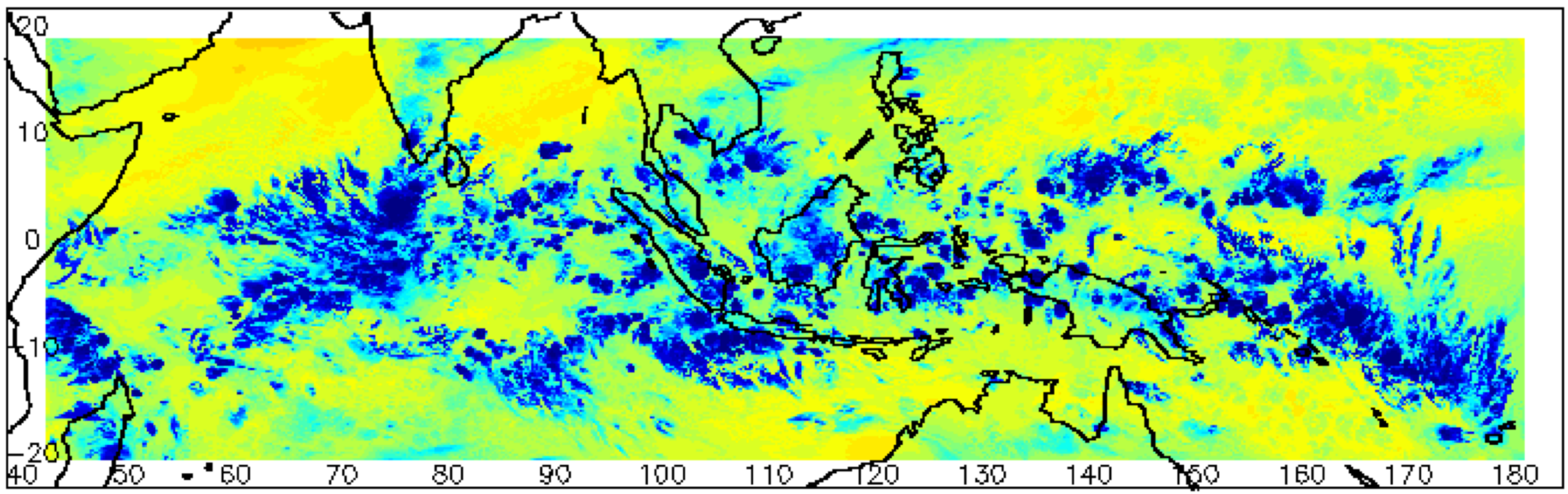


Christensen, Dawson and Holloway, 2018, JAMES

e.g. Pilot study: UKMO LAM "cascade" sim (4km) and IFS SCM

thanks to **Chris Holloway, U. Reading**

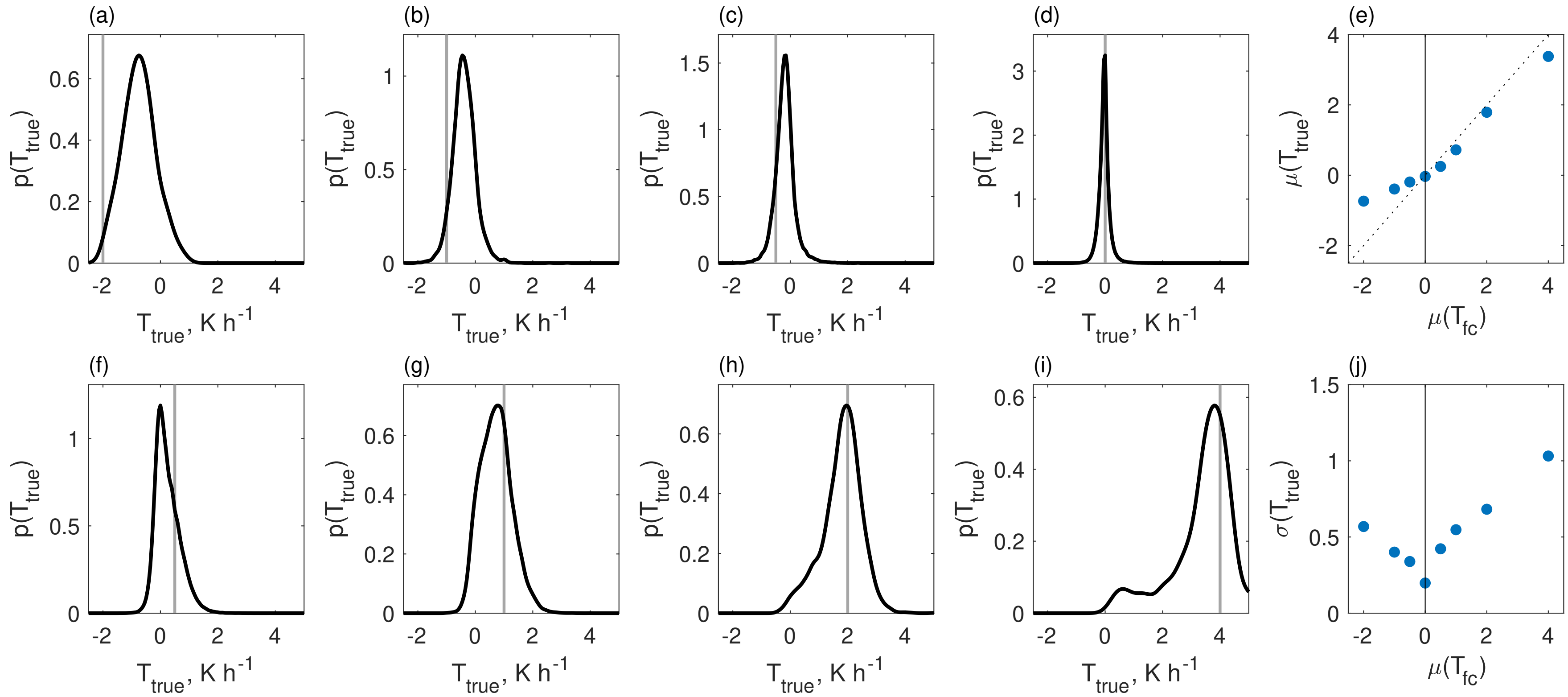
- UK Met Office atmospheric model setup
- Semi-Lagrangian, non-hydrostatic dynamics, 4km resolution
- Large tropical domain (15,500 km x 4,500 km), 9 days of data. Hourly dumps.
- Prescribe observed SST; boundary conditions from ECMWF 25 km analysis



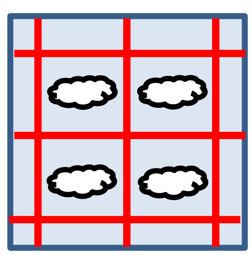
Holloway et al, 2012; 2013
Christensen et al, 2018; Christensen 2020

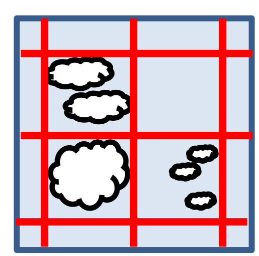
SCM provides deterministic prediction, to compare to ‘truth’

T850 parametrised tendencies



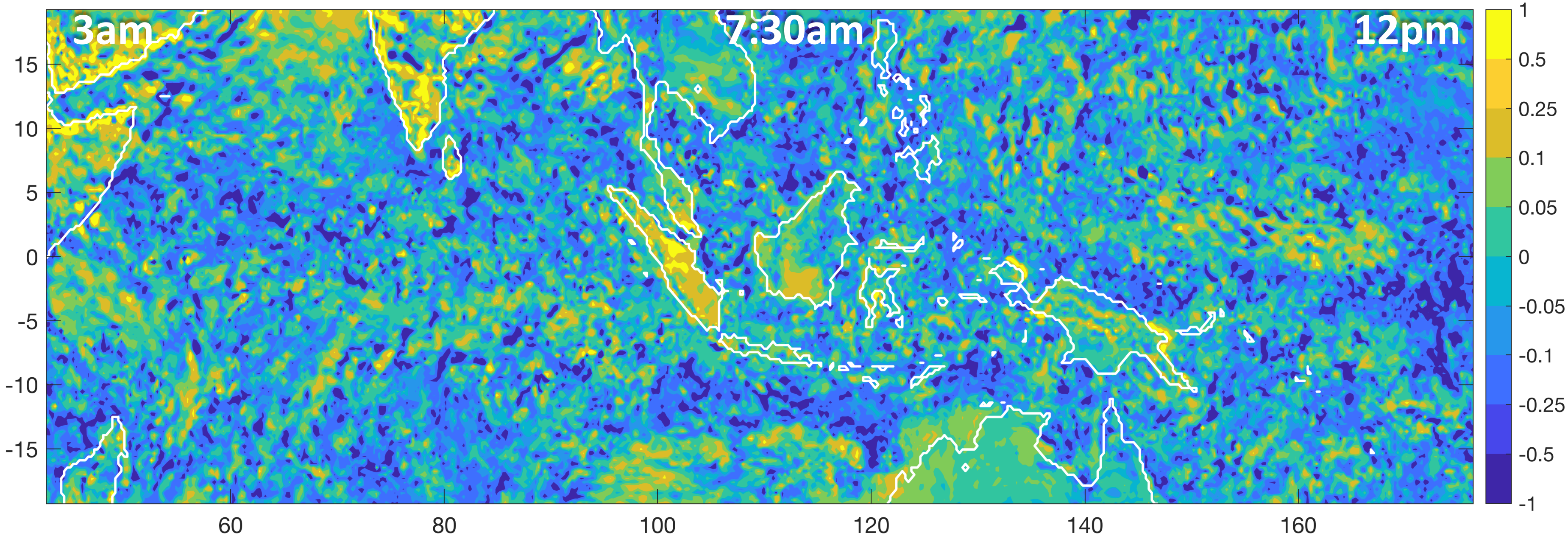
compare
deterministic
tendency
from SCM



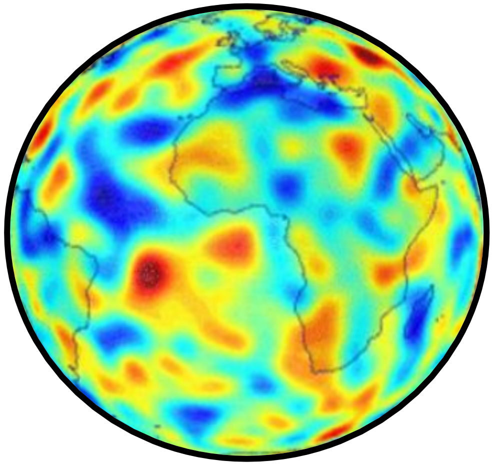


with observed
realisation from
high-res LAM

Snapshot of optimal SPPT '*e*' perturbation



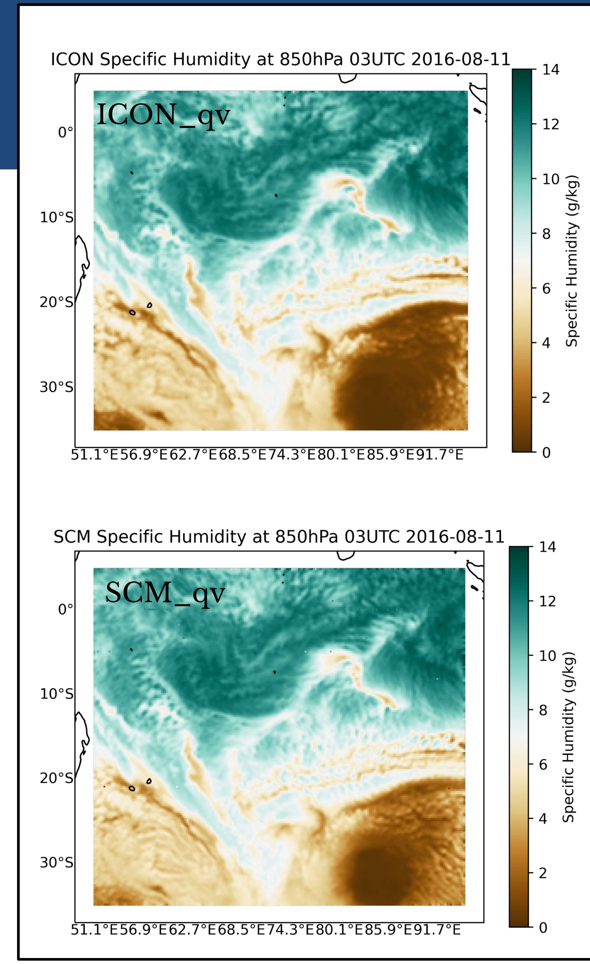
	Operational SPPT			Fitted SPPT		
σ_i	0.52	0.18	0.06	0.35	0.17	0.10
L_i (km)	500	1000	2000	32	370	-
τ_i	6 h	3 d	30 d	1.2 h	4.3 d	-



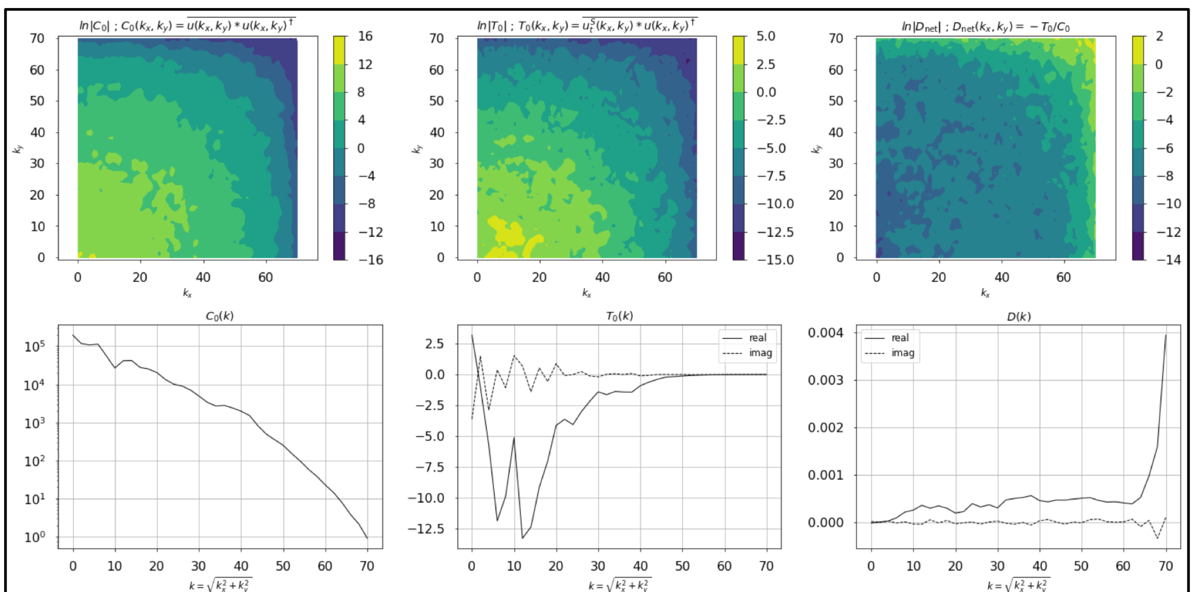
Christensen et al, 2018; Christensen 2020

Initial MU-MIP progress

- Protocol v1 agreed on
- Funding applied for/secured
- DEPHY-isation of SCMs
- SCM input files available, testing underway
- Analysis tools under development



↑Xia Sun, Kathryn Newman, Mike Ek, and Ligia Bernardet, NOAA/NCAR DTC



← Vassili Kitsios, CSIRO

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

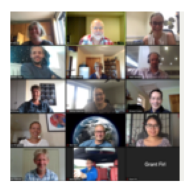
Introduction

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:

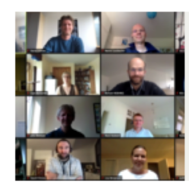
- How should we best represent model uncertainty/random error using stochastic approaches?
- To what extent should this representation be model specific or a fundamental property of atmospheric models?
- Are current approaches justified? How can they be improved?
- Can a coarse-graining approach be used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

News



Technical discussion meeting

26 August 2021

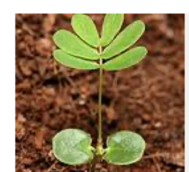


MUMIP Meeting 2

23 June 2021

Second MUMIP Team meeting scheduled

20 May 2021



Developmental Testbed Center funding for MUMIP work

24 March 2021

MUMIP Launch

Hello world

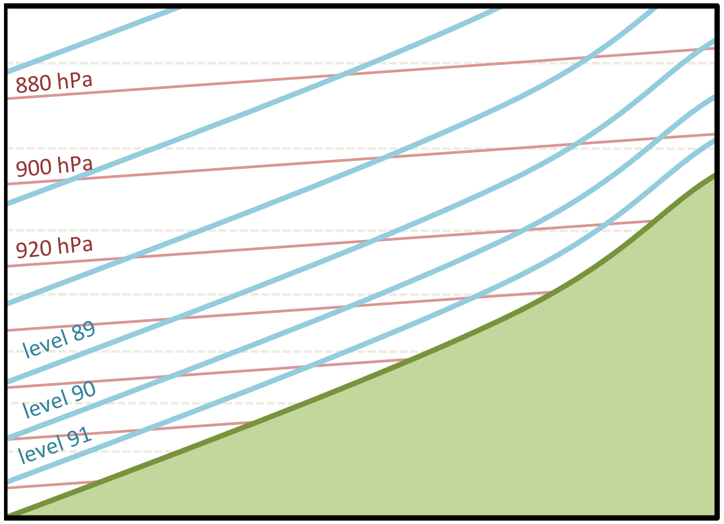
Thanks for listening

Extra Slides

Coarse graining details

1. Local area averaging for coarse graining

$$\overline{\psi}_{n,k} = \sum_{i}^{I} W_{n,i}\psi_{i,k}$$



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

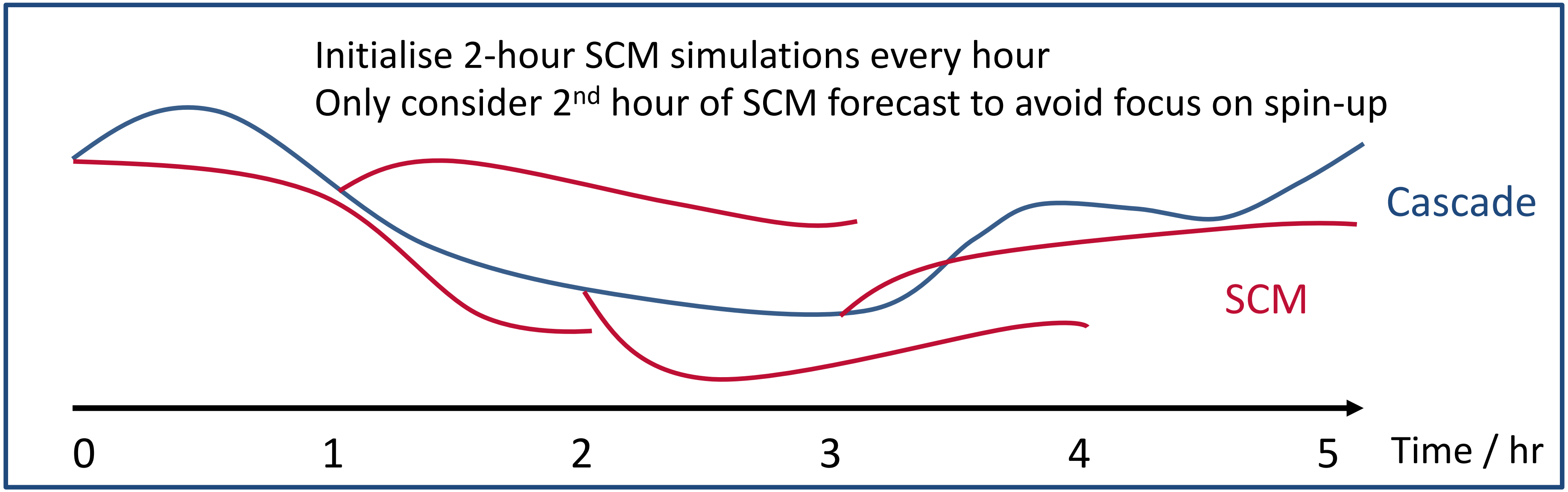
$$\text{adv}(\psi)|_{n,k} = -\overline{\mathbf{u}}_{n,k} \cdot \overline{\nabla}_k(\overline{\psi_{n,k}})$$

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

Christensen et al, 2018, JAMES.

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



What information do we have?

- ✓ **Total change in (T, q, U, V) in high-resolution Cascade** over 1hr time interval as a function of model level, location and forecast start time
- ✓ **Change in (T, q, U, V) in IFS SCM** over 1 hr, **decomposed into dynamics and individual parametrized tendencies**, as a function of model level, location and forecast start time

Cf. existing approaches to identify model error

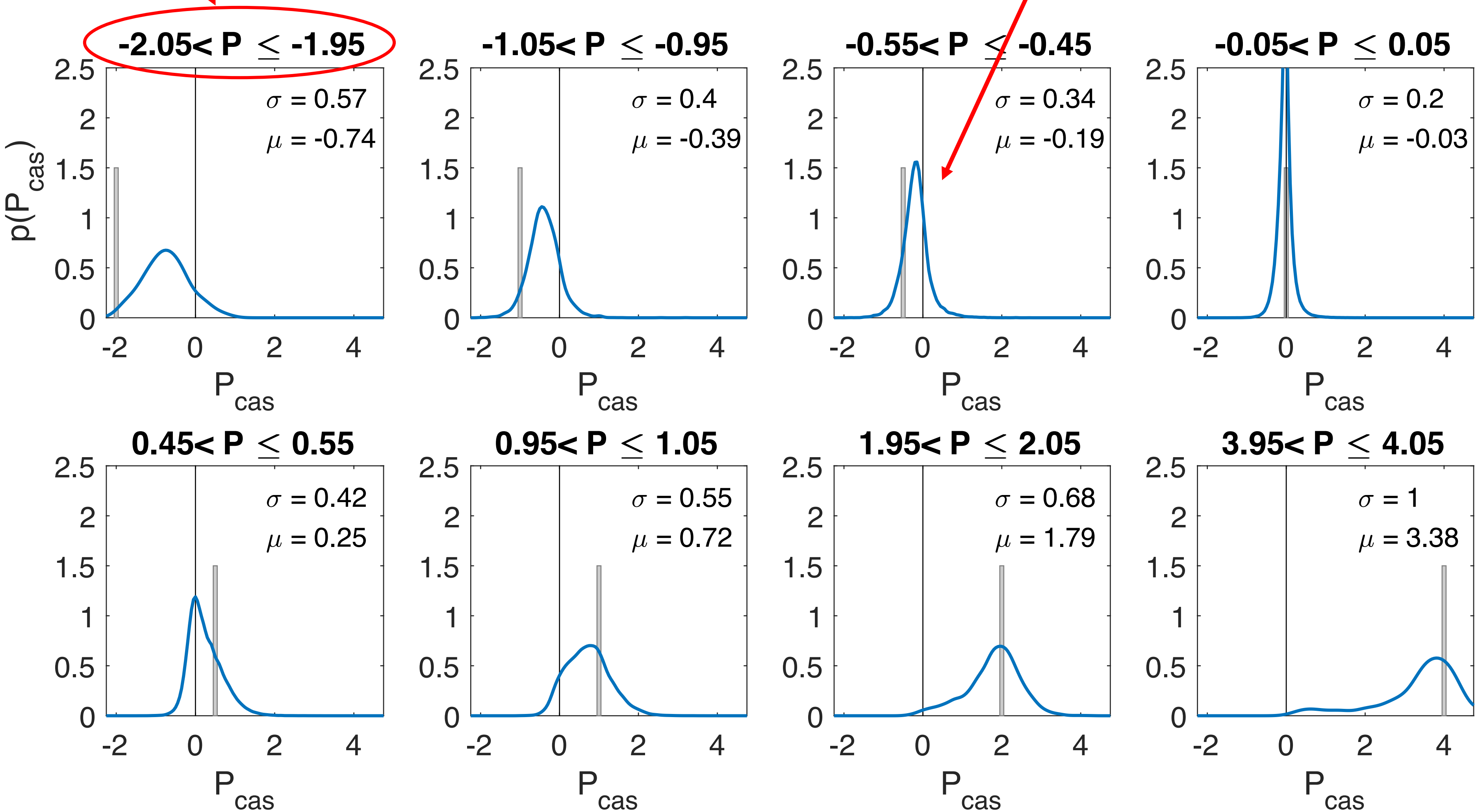
- **E.g. Initial tendency approach** in which physics tendencies in data assimilation cycle are compared to the analysis
- **E.g. Transpose AMIP** in which climate models are run in weather forecasting mode from common initial conditions

	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		☹	☹

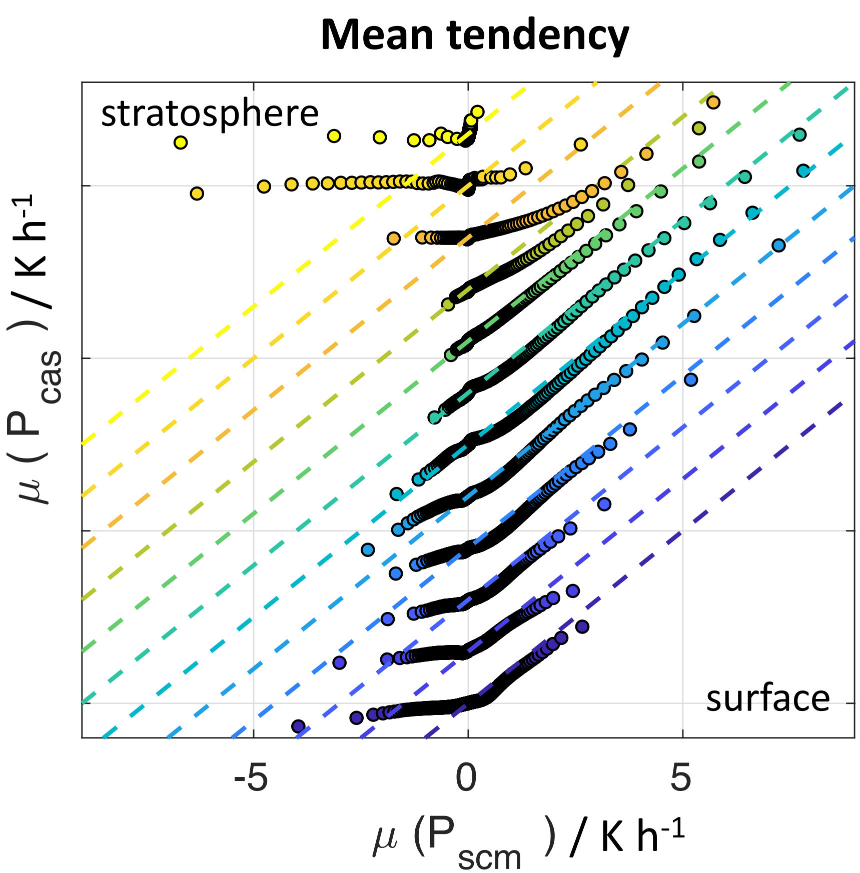
Consider T850 tendency (/ K h^{-1})

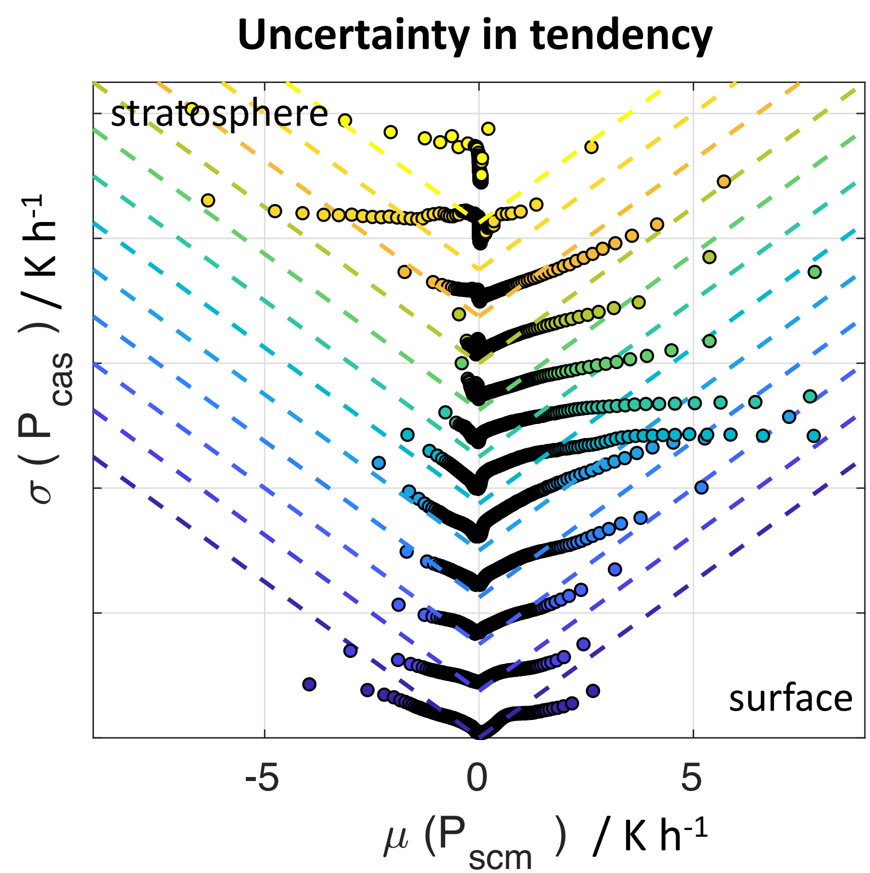
Sort data by predicted SCM tendency

Plot PDF of 'true' tendency



Consider T tendency



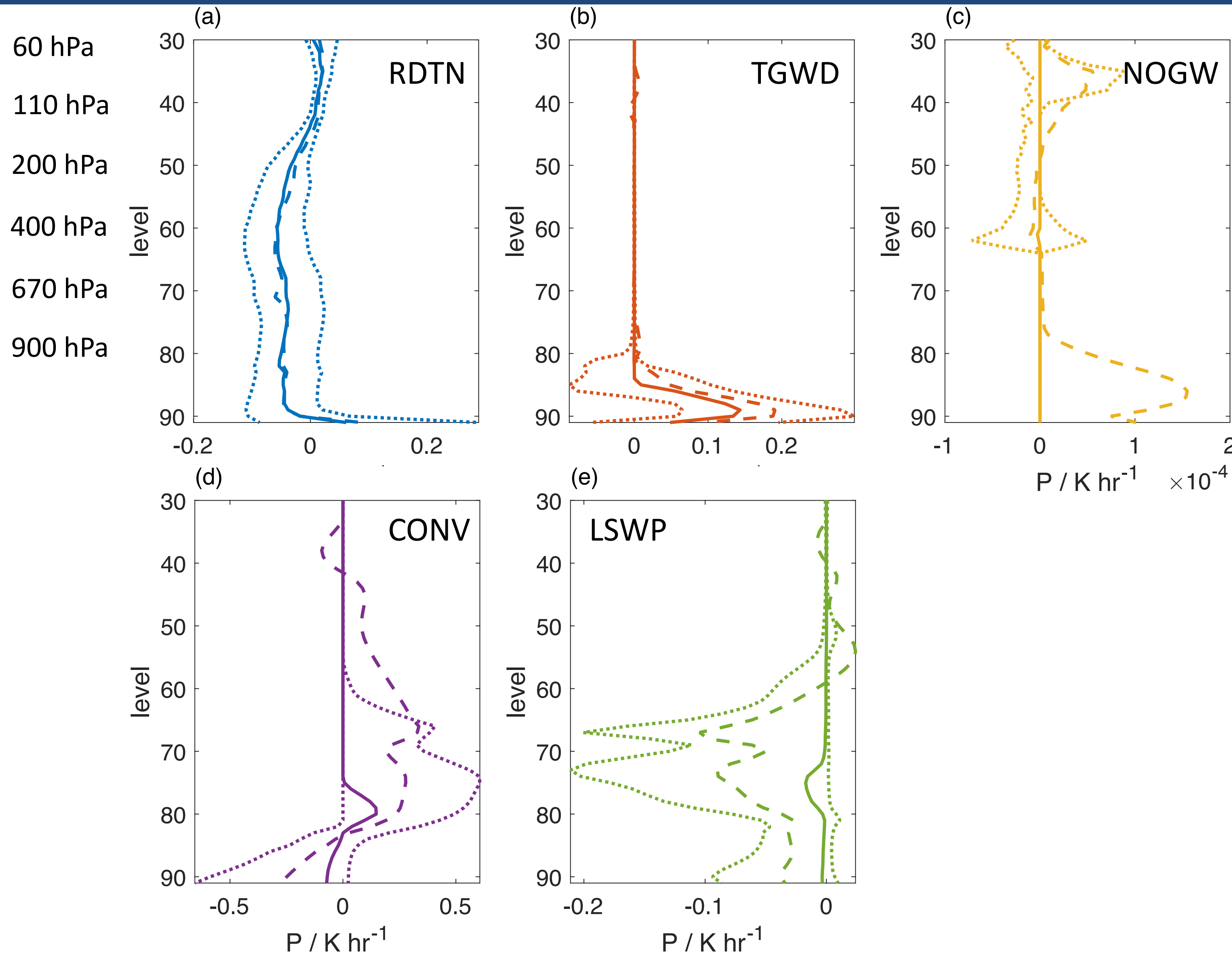


Data grouped by level.

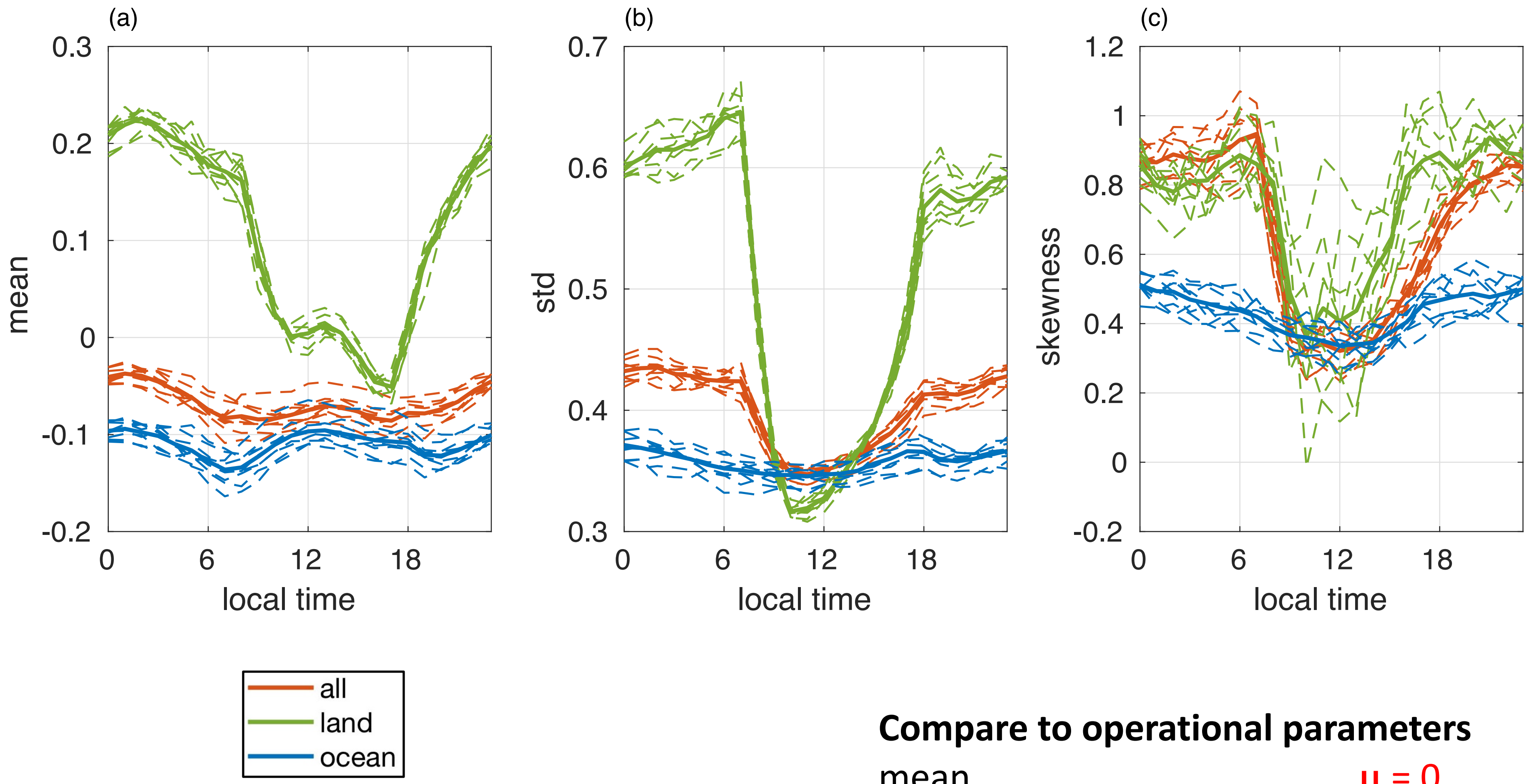
Dark blue: levels 91—87 (ground—995 hPa)

Yellow: levels 32—36 (86—60 hPa)

Where are the different schemes active?



Characteristics of ‘e’



Compare to operational parameters

mean	$\mu = 0$
standard deviation	$\sigma = 0.55$
skewness	$\gamma = 0$

Correlation scales of ‘e’

- Model temporal and spatial correlation scales as arising from a sum over several scales

$$e(t) = \sum_{i=1}^{n} X_i(t),$$

<= e.g., in time

$$X_i(t) = \phi_i X_i(t-1) + \sigma_i (1 - \phi_i^2)^{\frac{1}{2}} \xi$$

- Iteratively fit each scale, long to short

$$\sigma_e^2 = \sum_{i=1}^{n} \sigma_i^2$$

$$\rho_e = \frac{\sum_{i=1}^{n} \sigma_i^2 \phi_i^T}{\sum_{i=1}^{n} \sigma_i^2}$$

<= plot log(autocorrelation) and perform linear fit

Optimising SPPT

SPPT:

$$T = D + (1 + e)\sum_i P_i$$

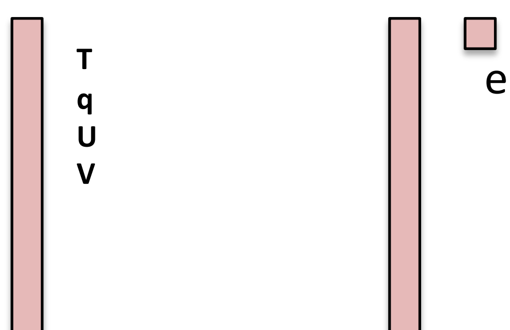
Calculate 'true' total tendency from CASCADE

Low resolution dynamics estimated from Cascade

Consider error in SCM physics tendencies

SOLVE

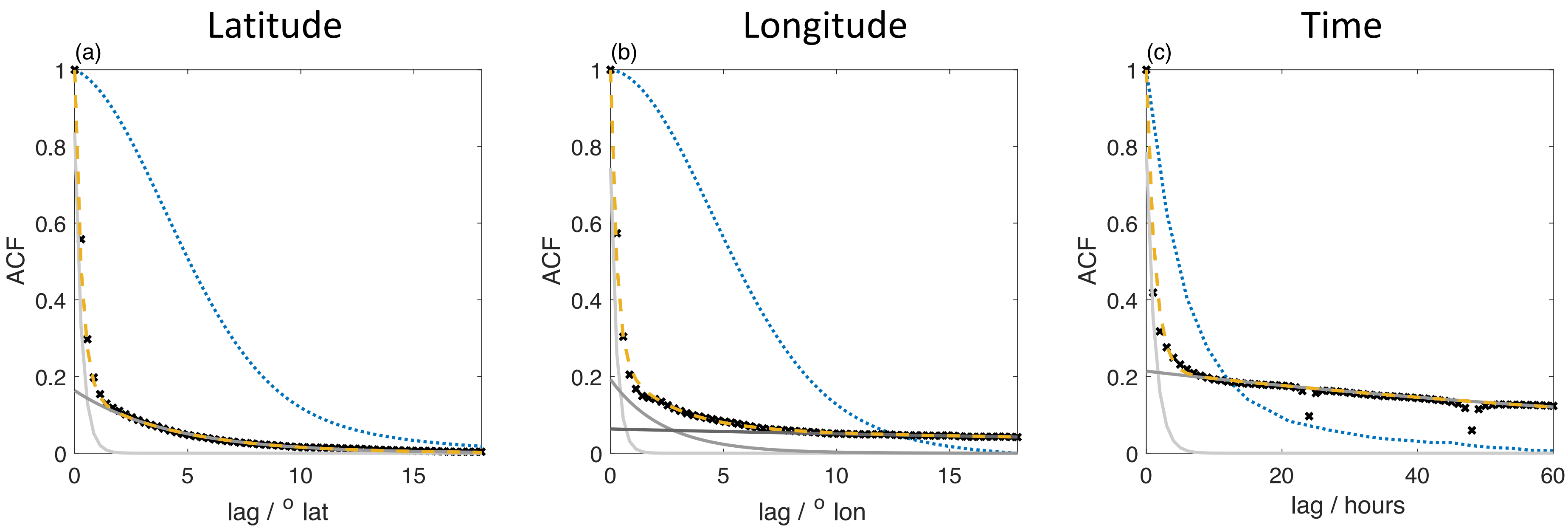
$$T - D - \sum_i P_i = e\sum_i P_i$$



Do not use data from BL or stratosphere (tapered)

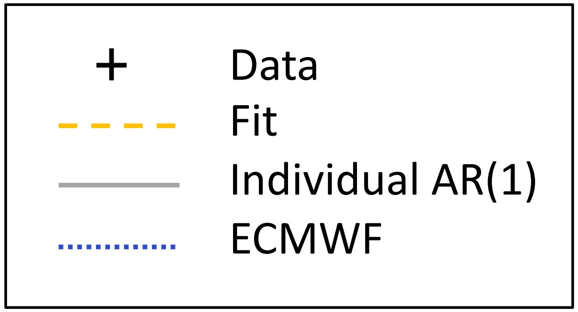
i.e.
Following the assumptions of SPPT, can we measure the statistical characteristics of the perturbation ***e***

Spatio-temporal correlations



- Model spatio-temporal correlations as a sum over *n* AR(1) processes with different scales

	Operational SPPT			Fitted SPPT		
σ_i	0.52	0.18	0.06	0.35	0.17	0.10
L_i (km)	500	1000	2000	32	370	-
τ_i	6 h	3 d	30 d	1.2 h	4.3 d	-



New optimal parameters for SPPT in IFS?

- Averaging over the variance ratios for the latitude, longitude and temporal correlations

	Operational SPPT			Fitted SPPT		
$\mu(e)$	0.0			-0.07		
$\sigma(e)$	0.55			0.40		
skew(e)	0.0			0.6		
σ_i	0.52	0.18	0.06	0.35	0.17	0.10
L_i (km)	500	1000	2000	32	370	-
τ_i	6 h	3 d	30 d	1.2 h	4.3 d	-

2. Beyond SPPT?

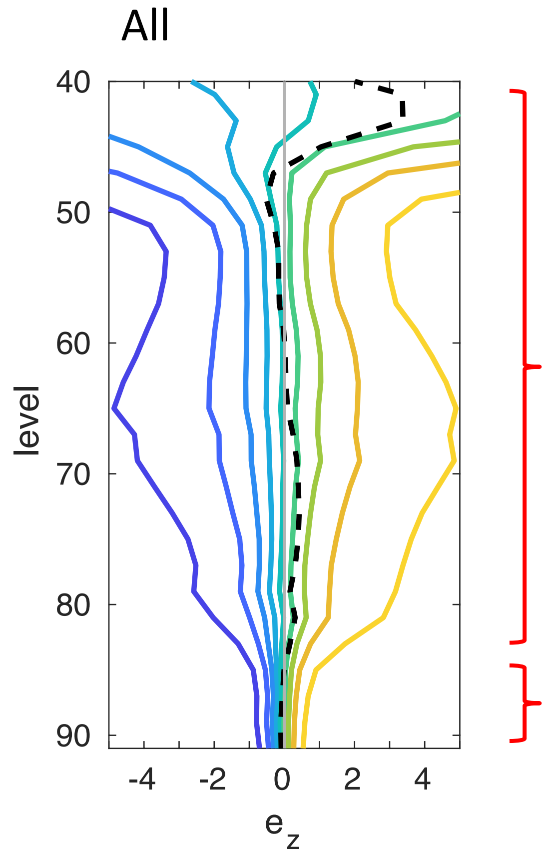
- SPPT is not a perfect representation of uncertainty in the IFS – can we improve on it?
- Have not yet assessed other assumptions made in SPPT – are these valid?
- Simple approach:
  - Relax each assumption in turn and fit new 'optimal e'
  - If the fitted 'e' is constant in dimension of interest then we should indeed hold the perturbation constant for that dimension

e.g. height,
e.g. variable,
e.g. parametrisation

Q. Vertical coherency of perturbations?

$$T_z = D_z + (1 + e_z) \sum_i P_{i,z}$$

- Fit separate e_z at each vertical level
- Consider pdf as a function of height summarised by deciles



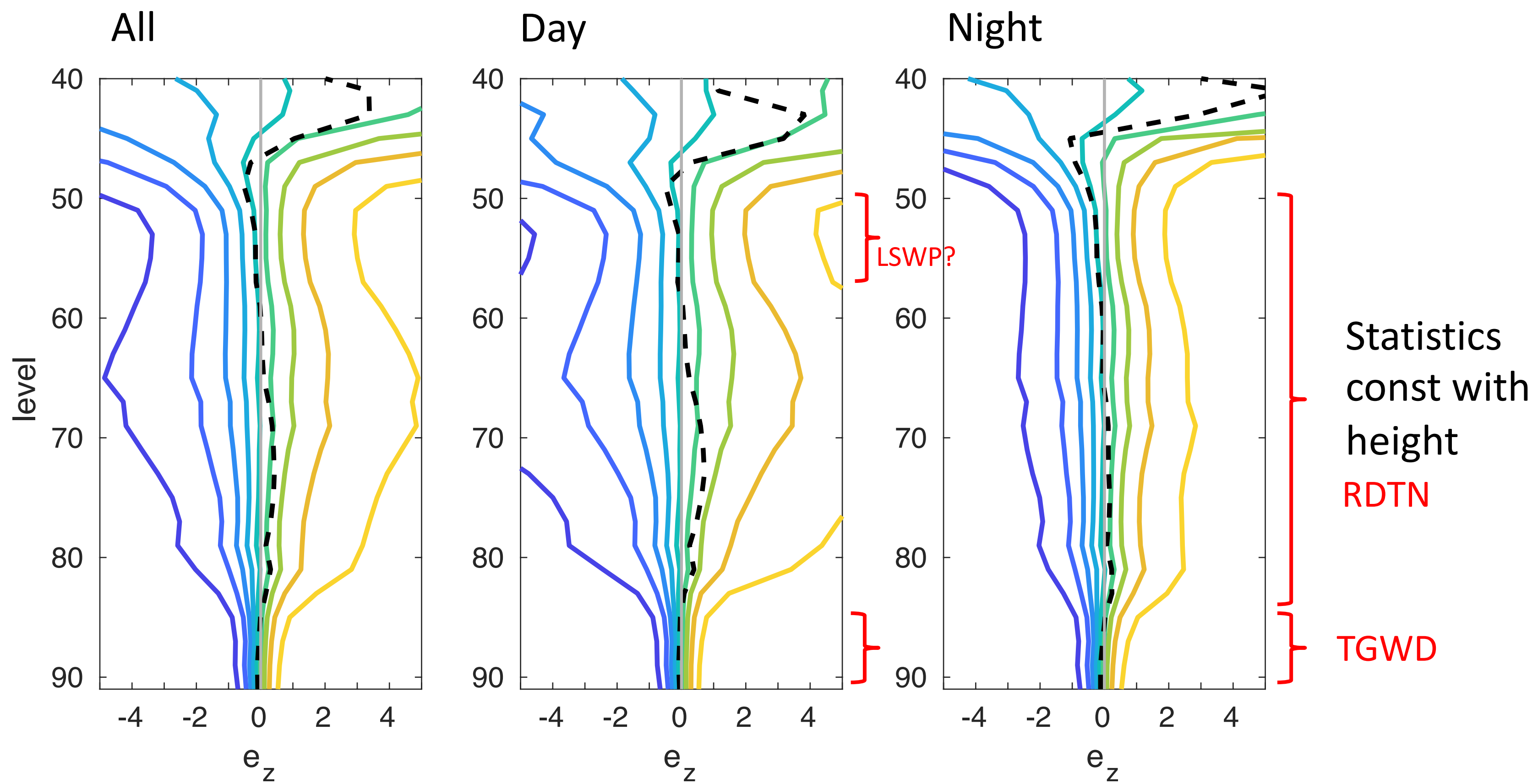
Statistics vary with height

Statistics constant with height in turbulent BL

Q. Vertical coherency of perturbations?

$$T_z = D_z + (1 + e_z) \sum_i P_{i,z}$$

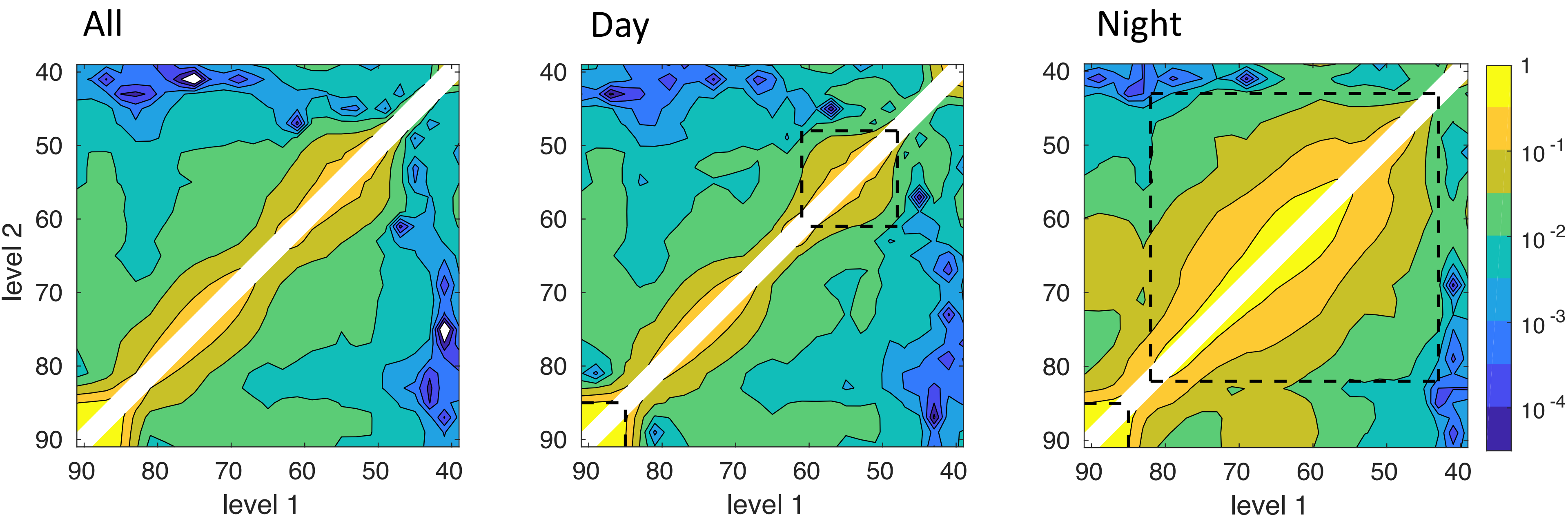
- Fit separate e_z at each vertical level
- Consider pdf as a function of height summarised by deciles



Q. Vertical coherency of perturbations?

$$T_z = D_z + (1 + e_z) \sum_i P_{i,z}$$

- Fit separate e_z at each vertical level
- Correlation between e_z fitted to different model levels



Enhanced correlations correspond to levels where one scheme dominates
High in BL.
Some evidence of enhanced correlations aloft at night.

Q. One perturbation for all tendencies? (T, q, U, V)

$$T_X = D_X + (1 + e_X) \sum_i P_{i,X}$$

- Fit separate e_X for each prognostic variable
- Assess statistics of e_X and correlation between different variables

	T			q			U			V		
μ(e)	-0.06			-0.02			-0.37			-0.52		
σ(e)	0.70			0.65			1.7			1.9		
σ_i	0.66	0.17	0.13	0.6	0.22	0.1	1.6	0.47	0.18	1.8	0.54	0.18
L_i (km)	39	400	-	33	430	-	38	270	-	26	290	-
τ_i	0.6 h	3.5 d	-	1.2 h	4.3 d	-	1.2 h	3.8 d	-	1.2 h	4.2 d	-

T, q statistics similar
Correlation = 0.35

U, V statistics similar
But low correlation = 0.08

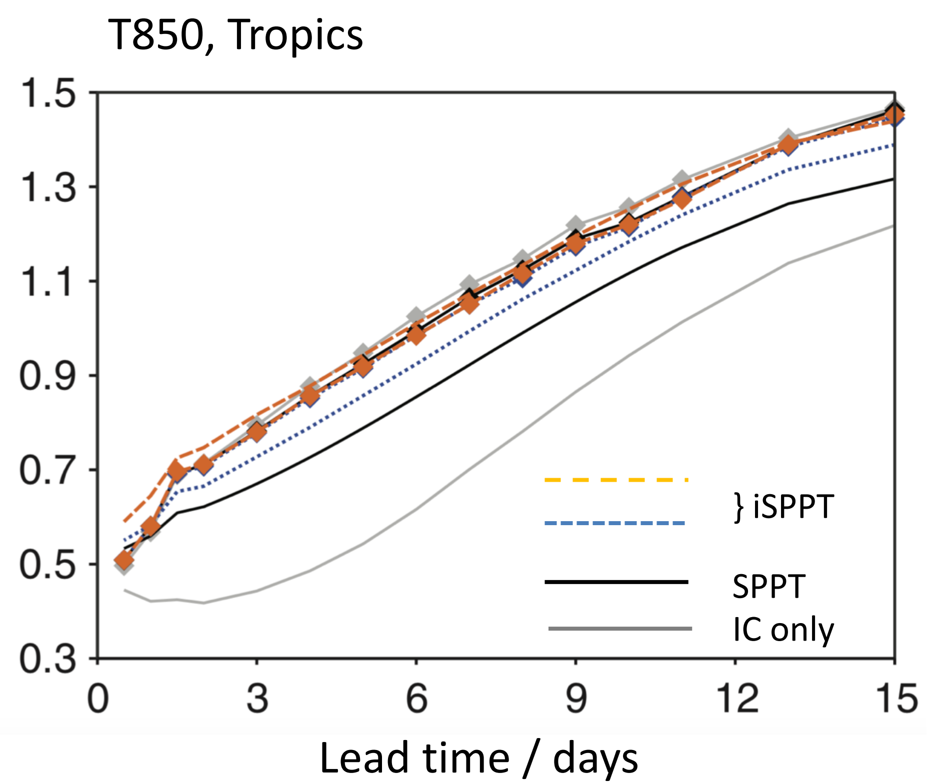
All other correlation pairs < 0.1

Q. One perturbation for all parametrisations?

- 'independent SPPT'

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

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



- Tested in IFS and found to benefit forecast reliability in the tropics

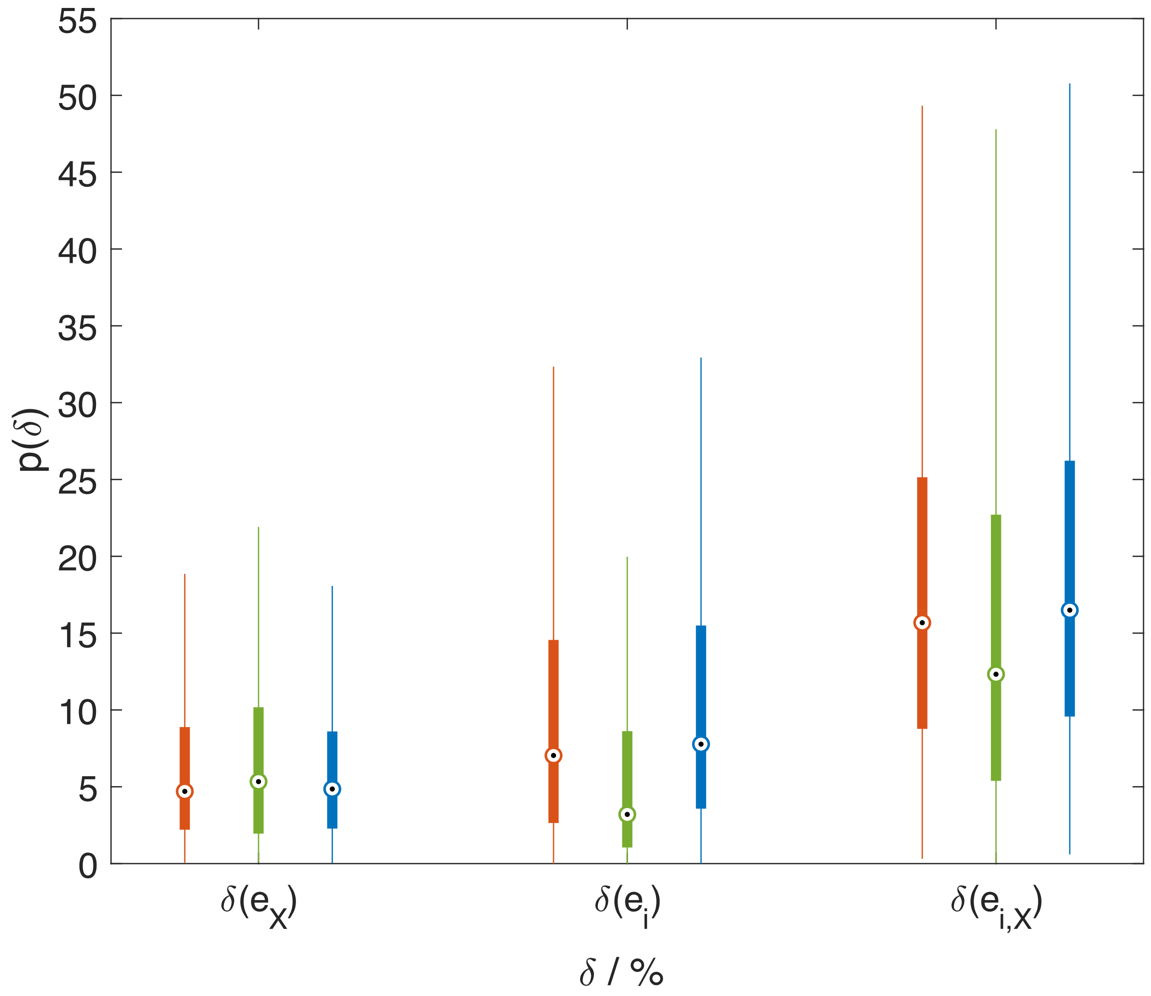
Christensen, Lock, Moroz and Palmer, 2017, QJRMetS

Q. One perturbation for all parametrisations?

- ‘independent SPPT’ seems to account for many results shown
  - Low correlation measured between perturbations fitted to different schemes
  - Perturbations to different schemes show very different noise characteristics
  - Measured correlation in the vertical is limited to within parametrisations
  - Measured correlations between perturbations applied to different variables are due to the physical relationship between those variables, as represented by the parametrisation schemes
  - Approach would enable multiplicative noise to be easily replaced by an alternative approach if desired, e.g. for convection

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

Fractional variance explained



$$\delta = 100 \cdot \frac{MSD_{SPPT} - MSD_{new\ SPPT}}{MSD_{SPPT}}$$

For Mean Square Difference (MSD) between measured and modelled error