



# Model Uncertainty Intercomparison Project WGNE Update

**Hannah Christensen**

[Hannah.christensen@physics.ox.ac.uk](mailto:Hannah.christensen@physics.ox.ac.uk)

Atmospheric, Oceanic and Planetary Physics,  
University of Oxford

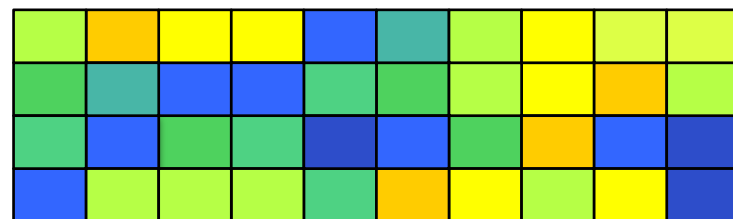
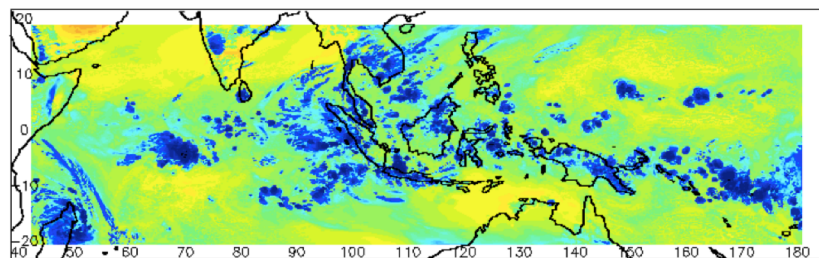
**WGNE35, 2-5 November 2020**

# Background

- Joint initiative of Predictability, Dynamics and Ensemble Forecasting (PDEF) working group and Working Group on Numerical Experimentation (WGNE)
- A primary joint interest of the two working groups is model error identification and its representation in ensemble forecasts
  - Systematic and random error
- At the joint WGNE/PDEF meeting in Tokyo, October 2018, a coordinated activity was proposed to evaluate model error across a number of forecast models
- Some key questions:
  - How should we best represent model uncertainty (random error)?
  - 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 we design scale-aware schemes?
  - What data or measurements are available to address these questions?
  - ...

# Summary of protocol: use high-resolution dataset as 'truth'

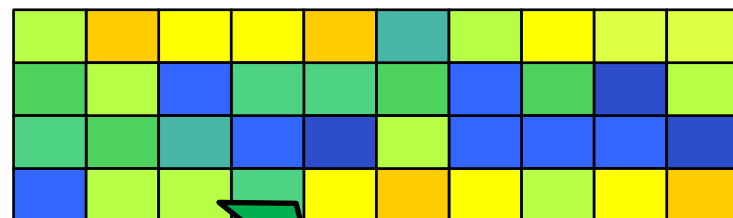
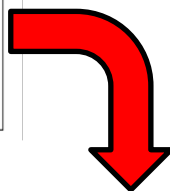
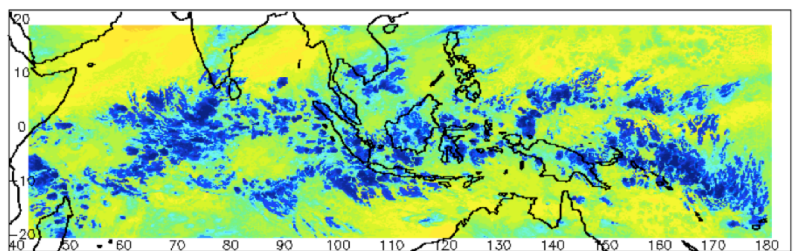
1. Coarse grain high resolution dataset to forecast model grid



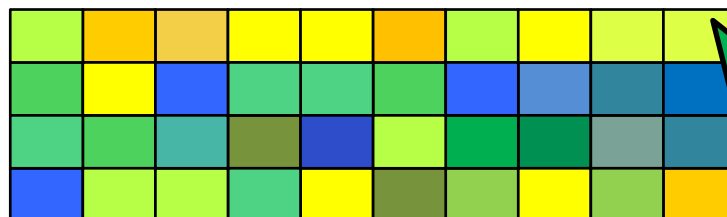
High resolution model

2. Use forecast model to step forward coarse-grained fields

Forecast model



3. Compare at later time



# Use SCM as forecast model

- Use coarse-grained high resolution simulation to prescribe
  - Initial conditions
  - Forcing: advective tendencies, geostrophic winds, vertical velocity
  - Boundary conditions: Surface sensible and latent heat fluxes, Skin temperature

## Why use the SCM?

- Supply dynamical tendencies → target uncertainty in the parametrization schemes
- The SCM is more portable than the full model, and is cheap to run. Potential to run SCM on computer where high-res data is stored
- (Spectral models cannot be run over a limited domain, but we can tile many independent SCM to cover the limited domain.)

# 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                 |                                                                                      |  |  |

# What information do we have?

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

*For examples of analysis that can be carried out with this data,  
please see Christensen, 2020, QJRMetS*

*Case study using UKMO limited area high-res simulation and OpenIFS SCM*

# DEPHY common SCM format

- New standardised SCM protocol has been proposed by a group of French researchers involved in the High Tune and DEPHY communities.
  - standardises the format of input/output files needed to run an SCM.
- Many SCM groups participated at an interactive workshop in June 2020
  - protocol was discussed, and groups each began to implement this protocol
- Ideally, all SCM participating in this intercomparison will use DEPHY format

DEPHY\_FormatCommun\_Roehrig\_v1

File Edit View Insert Format Tools Add-ons Help

100% Normal text Arial 18

DEPHY – Common format for SCM simulations  
Version 1 - On construction

R. Roehrig, F. Hourdin, M.-P. Lefebvre, F. Couvreux, C. Rio,  
E. Bazile, Y. Bouteloup, J.-F. Guérémy, P. Marquet

For each case, 2 netCDF files will be made available:

- A file "REF", which defines the case as close to its reference definition (literature, intercomparison project) as possible;
- A file "SCM" similar to the file "REF", but with a common vertical axis common to all variables (high vertical resolution, e.g., 10 m, in order to ensure a quasi convergence of profiles applied to any SCM), a time axis common to all forcing variables. The file will also contain anything required to initialize and force a model which uses T or  $\theta$ ,  $q_v$ ,  $q_i$ ,  $r_v$ , or  $r_i$  as state variables. Therefore interpolation/extrapolation and variable conversion will be handled by shared tools when creating the "SCM" file from the "REF" file.

1. **Formatting common to both files:**

All netCDF files should have **NETCDF3** format.  
All netCDF variables are of type *double*. Name and unit conventions are defined in Appendix

1. Each file contains a series of global attributes which define the forcing type of the case. This series of attributes is defined in Appendix 2.

Formatting common to both files...

File "DEF"

Initial profiles:

Forcing variables:

File "SCM"

initial profiles:

Forcing :

Appendix 1: Conventions for v...

Appendix 2: Global attributes

adv\_\$X\$ = 0 unactivated / 1 a...

rad\_temp = 0/1/"adv"

forc\_omega = 0/1

forc\_geo = 0/1

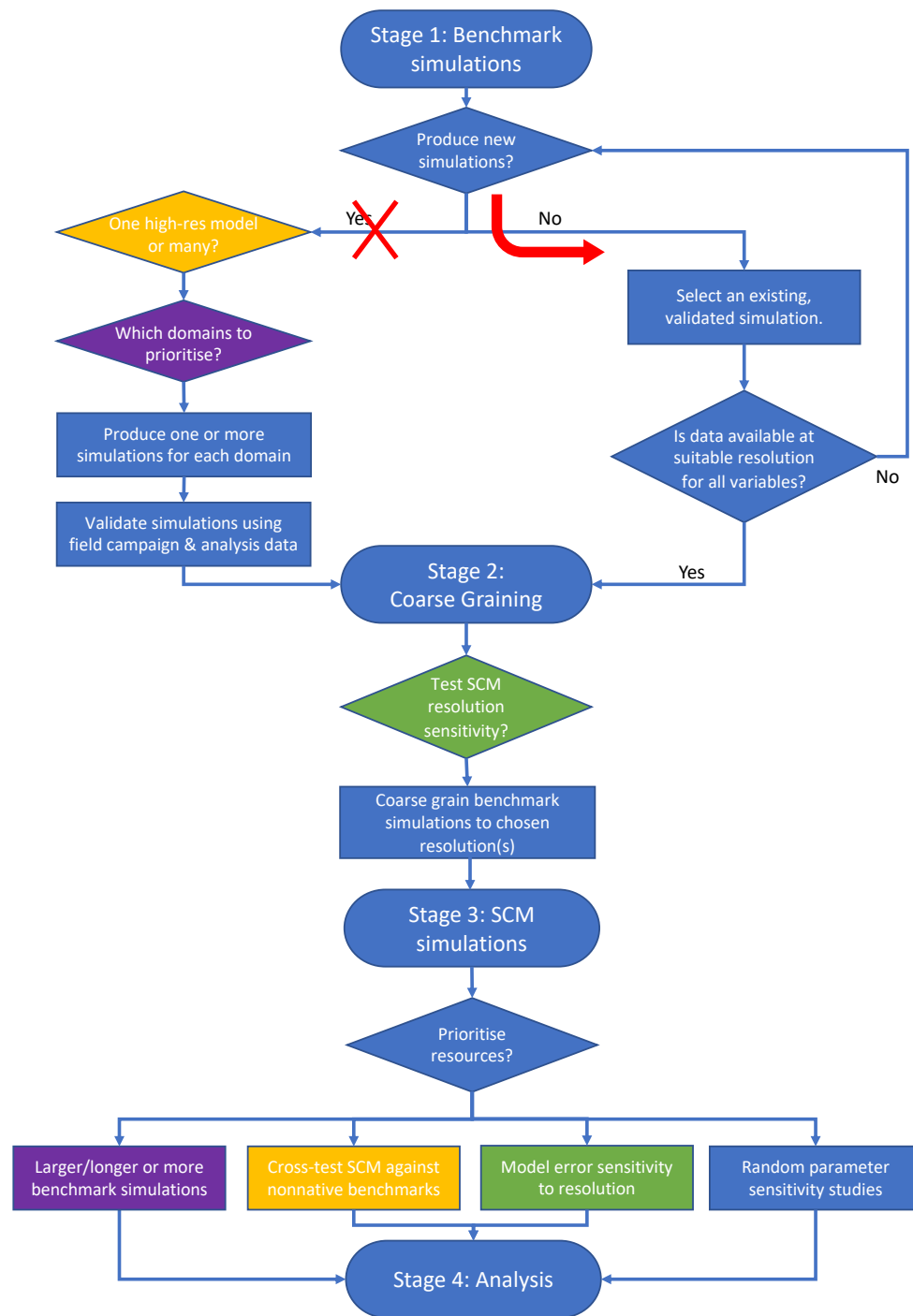
Steven Böing  
9:51 AM Jun 4

In the lower troposphere at least. For some of the ERA-based cases, we are thinking of including profiles high into

Show more

Romain Roehrig  
3:56 PM Jun 8

Agreed. The vertical resolution should





# Kick-off meeting, 22 September 2020

- **Attendees:**

- Lisa Bengtsson (NOAA)
- Ligia Bernardet (NOAA),
- Judith Berner (NCAR)
- Hannah Christensen (U. Oxford)
- Grant Firl (NCAR)

- Daniel Klocke (HErZ, DWD)
- Martin Leutbecher (ECMWF)
- Mark Rodwell (ECMWF)
- Nils Wedi (ECMWF)
- Keith Williams (Met Office)

- **Apologies:** John Methven (U. Reading)

## Aims

- Find out who is interested in the project and in what capacity
- Come to a consensus on the main scientific goals and priorities of the project
- Make practical decisions regarding domains etc.

# Kick-off meeting summary

- Agreed on protocol
- Comparison of model uncertainty characteristics between models is main priority
- Chose ICON Dymond (2.5km) simulation as first benchmark
  - To be coarse grained to NWP model resolution: 10-20 km.
  - In future, aim to coarse-grain ICON to coarser resolution (say 1 degree), as well as choose a second benchmark, also coarse-grained to 10-20km.
- Will first consider a tropical ocean domain (Pacific v Atlantic TBC)
- Each party present indicated their anticipated participation in the project.
  - Coarse graining: Hannah Christensen
  - Three SCM groups hope to participate: UKMO (Keith Williams), ECMWF (Nils Wedi?), NOAA/NCAR (Ligia Bernadet/Grant Firl/Judith Berner).
  - Translation of results to inform stochastic parametrisations (Lisa Bengtsson, Martin Leutbecher)
  - Linking up existing approaches with this new analysis (Mark Rodwell).

# Looking ahead

**Follow up meeting TBC! Space for more participants!**

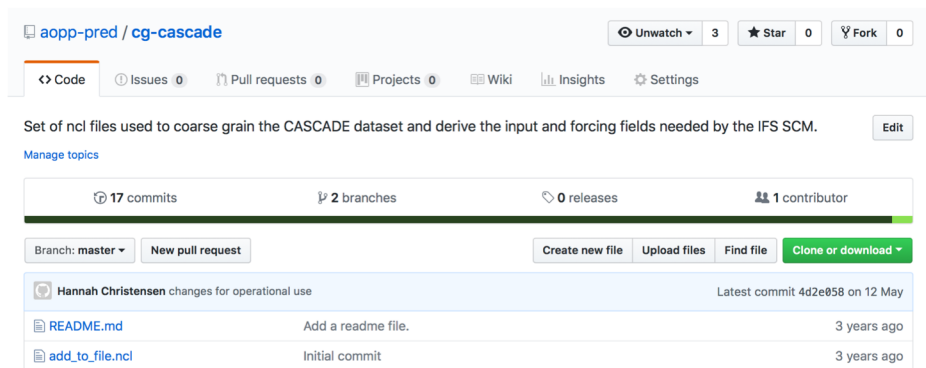
## References

Christensen, Dawson and Holloway, 2018, JAMES, 10(8) 1833-1857

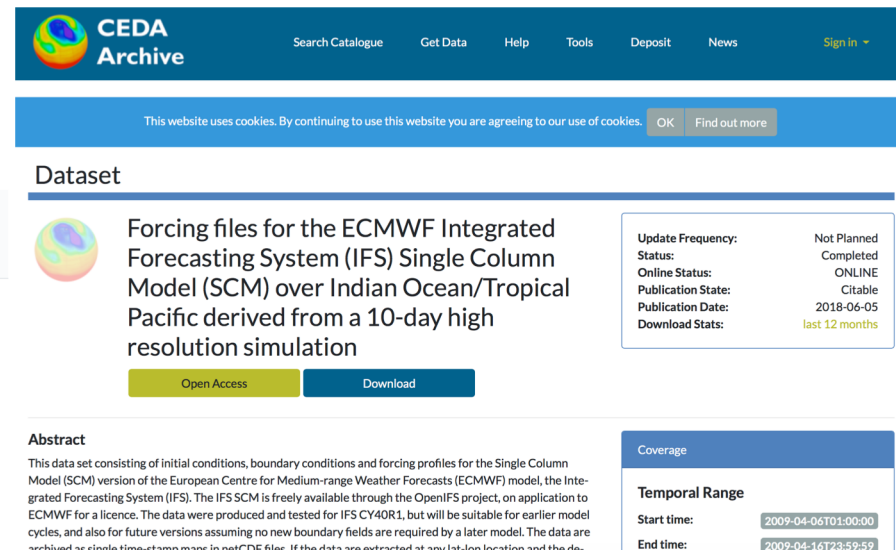
Christensen, 2020, QJRMetS, 146(727), 938-962

Coarse-grained Cascade data published on UK CEDA archive

NCL coarse graining scripts, and python  
SCM deployment scripts on github



The screenshot shows the GitHub repository page for 'aopp-pred / cg-cascade'. The repository has 17 commits, 2 branches, 0 releases, and 1 contributor. The description states: 'Set of ncl files used to coarse grain the CASCADE dataset and derive the input and forcing fields needed by the IFS SCM.' The repository is managed by Hannah Christensen, with the latest commit '4d2e058' on 12 May. The repository contains files 'README.md' and 'add\_to\_file.ncl', both committed 3 years ago.



The screenshot shows the CEDA Archive website. The header includes the CEDA Archive logo and navigation links: Search Catalogue, Get Data, Help, Tools, Deposit, News, and Sign in. A cookie notice is displayed below the header. The main content area is titled 'Dataset' and features a large image of a globe. The dataset title is 'Forcing files for the ECMWF Integrated Forecasting System (IFS) Single Column Model (SCM) over Indian Ocean/Tropical Pacific derived from a 10-day high resolution simulation'. Below the title are buttons for 'Open Access' and 'Download'. To the right of the title is a table with metadata:

| Update Frequency:  |                |
|--------------------|----------------|
| Status:            | Not Planned    |
| Online Status:     | Completed      |
| Publication State: | ONLINE         |
| Publication Date:  | Citable        |
| Publication Date:  | 2018-06-05     |
| Download Stats:    | last 12 months |

Below the dataset title is an 'Abstract' section. The abstract text reads: 'This data set consisting of initial conditions, boundary conditions and forcing profiles for the Single Column Model (SCM) version of the European Centre for Medium-range Weather Forecasts (ECMWF) model, the Integrated Forecasting System (IFS). The IFS SCM is freely available through the OpenIFS project, on application to ECMWF for a licence. The data were produced and tested for IFS CY40R1, but will be suitable for earlier model cycles, and also for future versions assuming no new boundary fields are required by a later model. The data are archived as single time-stamp maps in netCDF files. If the data are extracted at any lat-lon location and the de-'. To the right of the abstract is a 'Coverage' section with a 'Temporal Range' table:

| Temporal Range |                     |
|----------------|---------------------|
| Start time:    | 2009-04-06T01:00:00 |
| End time:      | 2009-04-16T23:59:59 |

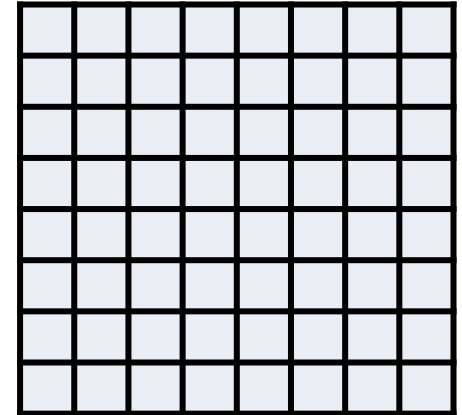
Thanks for listening

# Coarse graining details 1

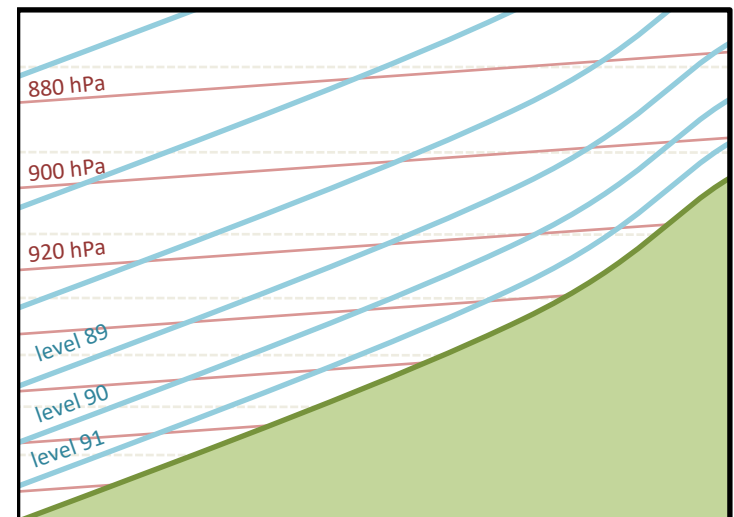
- Local area averaging for coarse graining.

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

*i*th  
box



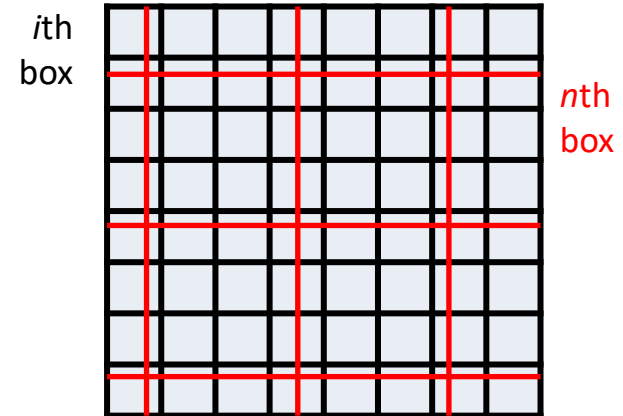
- Linearly interpolate in time.
  - High-resolution simulations not stored every timestep
  - Do we have fine enough temporal resolution for our needs?
- Vertical interpolation
  - Evaluate coarse-scale grid box mean  $p_{\text{sfc}}$
  - Coarse-grain other fields along model levels
  - Interpolate from native model levels to target model levels
  - Only an issue over orography  $\rightarrow$  propose that we focus exclusively on ocean regions.



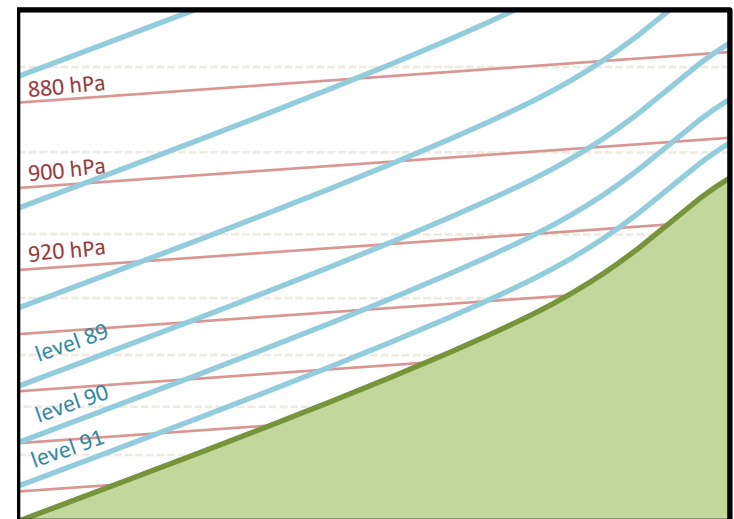
# Coarse graining details 1

- Local area averaging for coarse graining.

$$\bar{\psi}_{n,k} = \sum_i^l w_{n,i} \psi_{i,k}$$



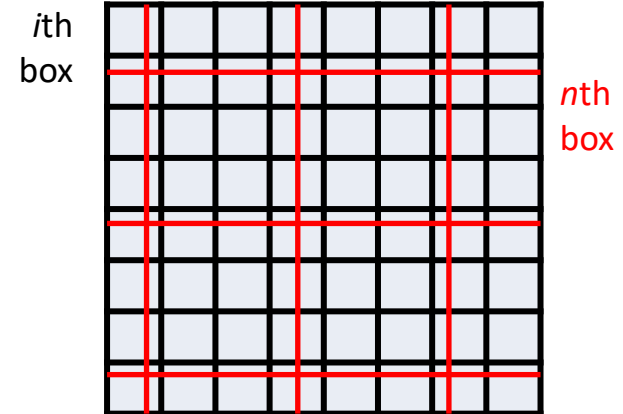
- Linearly interpolate in time.
  - High-resolution simulations not stored every timestep
  - Do we have fine enough temporal resolution for our needs?
- Vertical interpolation
  - Evaluate coarse-scale grid box mean  $p_{\text{sfc}}$
  - Coarse-grain other fields along model levels
  - Interpolate from native model levels to target model levels
  - Only an issue over orography  $\rightarrow$  propose that we focus exclusively on ocean regions.



# Coarse graining details 1

- Local area averaging for coarse graining.

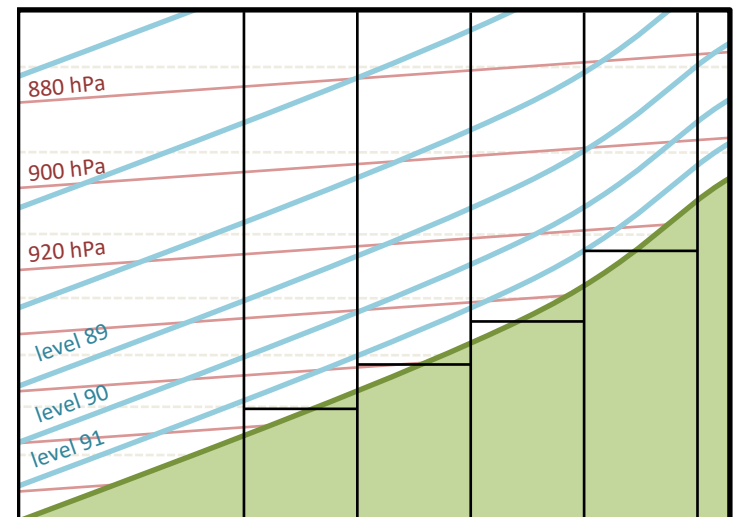
$$\bar{\psi}_{n,k} = \sum_i^l w_{n,i} \psi_{i,k}$$



- Linearly interpolate in time.
  - High-resolution simulations not stored every timestep
  - Do we have fine enough temporal resolution for our needs?

- Vertical interpolation

- Evaluate coarse-scale grid box mean  $p_{\text{sfc}}$
- Coarse-grain other fields along model levels
- Interpolate from native model levels to target model levels
- Only an issue over orography  $\rightarrow$  propose that we focus exclusively on ocean regions.



# Coarse graining details 2

- Above high-resolution model top, pad data using ECMWF analysis
- Advective tendencies estimated from the coarsened fields

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

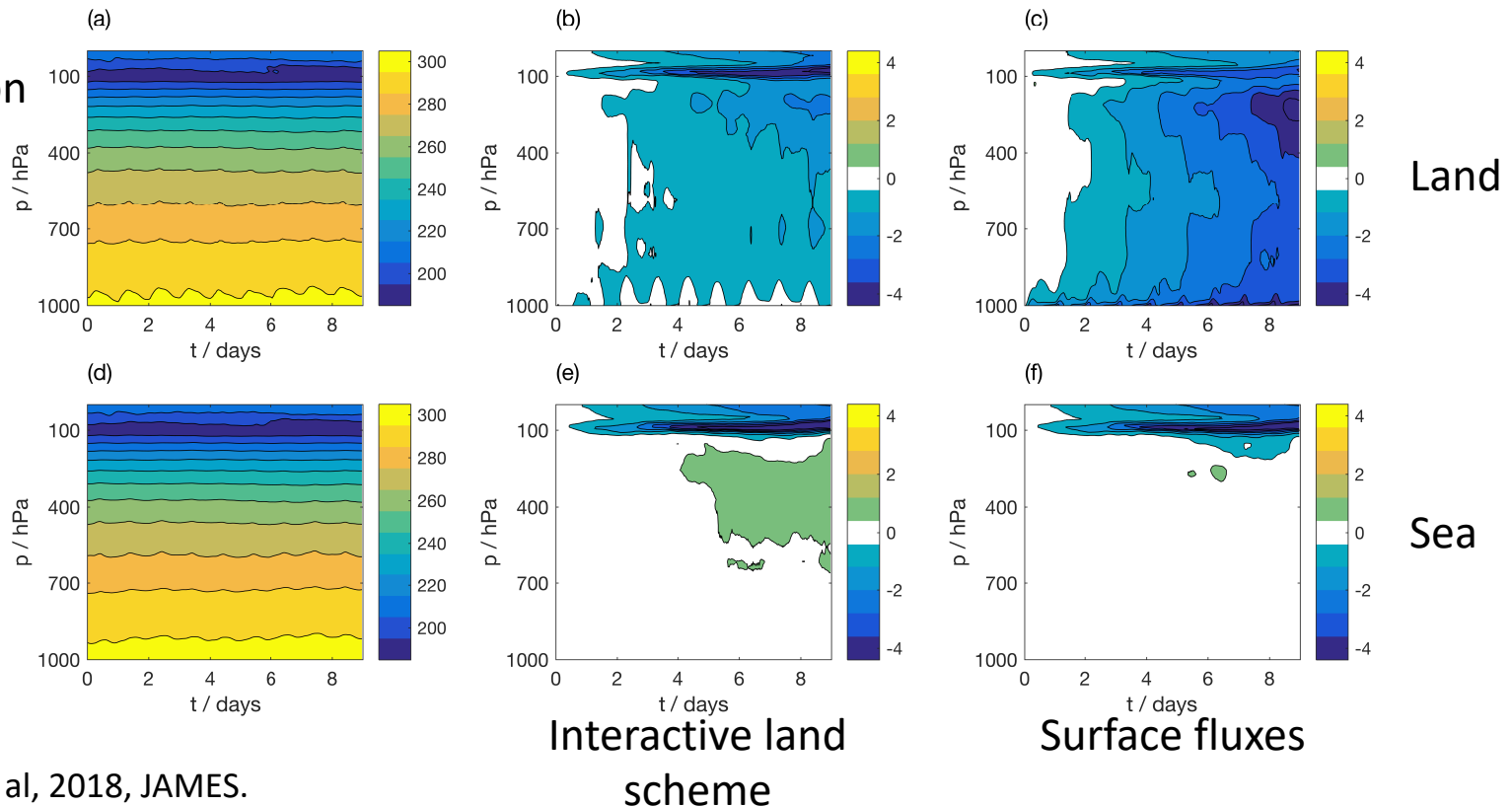
- Specify sensible and latent heat fluxes from high-resolution dataset
- Static boundary conditions (e.g. orography, land surface type) from operational model at chosen resolution



# Implementation details

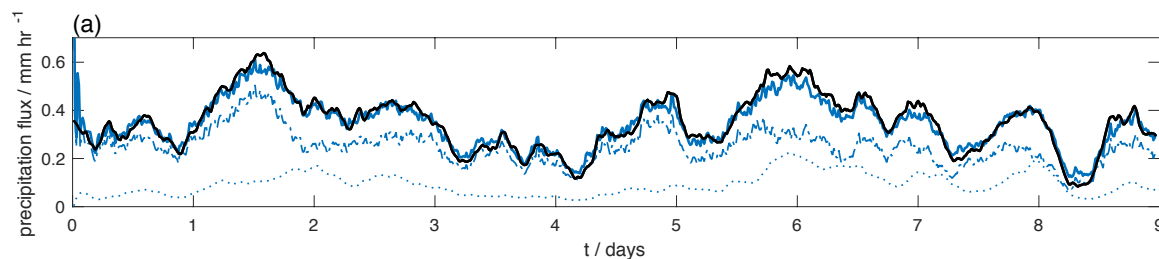
1. Verify coarse-graining procedure by taking IFS forecast data at T639
  - Linearly interpolate 1hr  $\rightarrow$  15 mins
  - Estimate advective fluxes from gridpoint fields
  - Supply sensible and latent fluxes instead of interactive land scheme
  - Interpolate from native model levels to target model levels

T: MC region

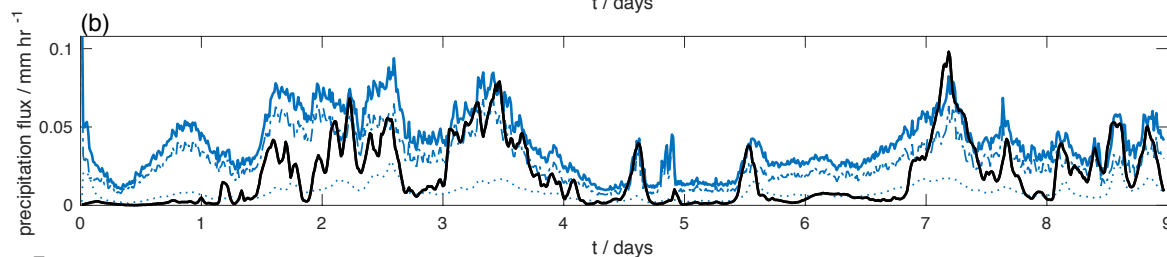


# How does the SCM compare to Cascade?

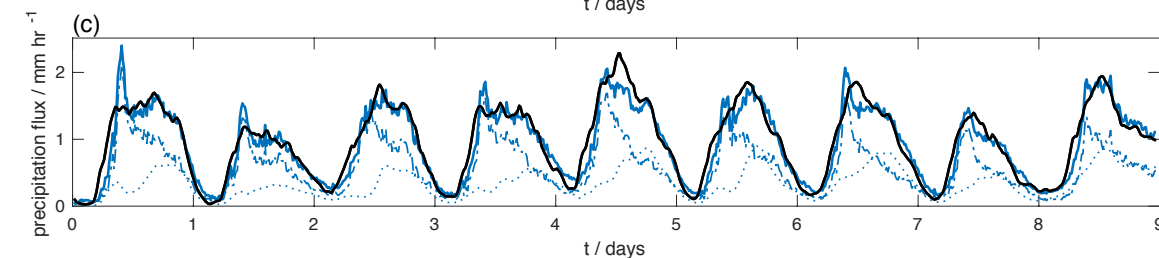
W. Pacific  
Ocean



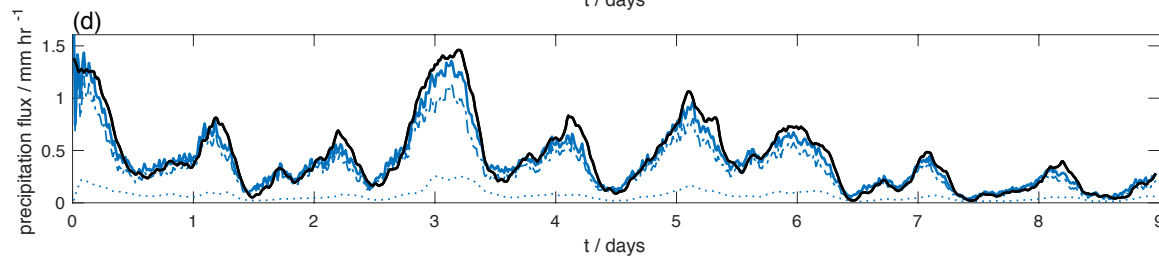
Ocean  
West of  
Australia



Maritime  
Continent  
Land



Maritime  
Continent  
Ocean



Precipitation

Total ———

Conv - - -

Strat .....

CAS ———

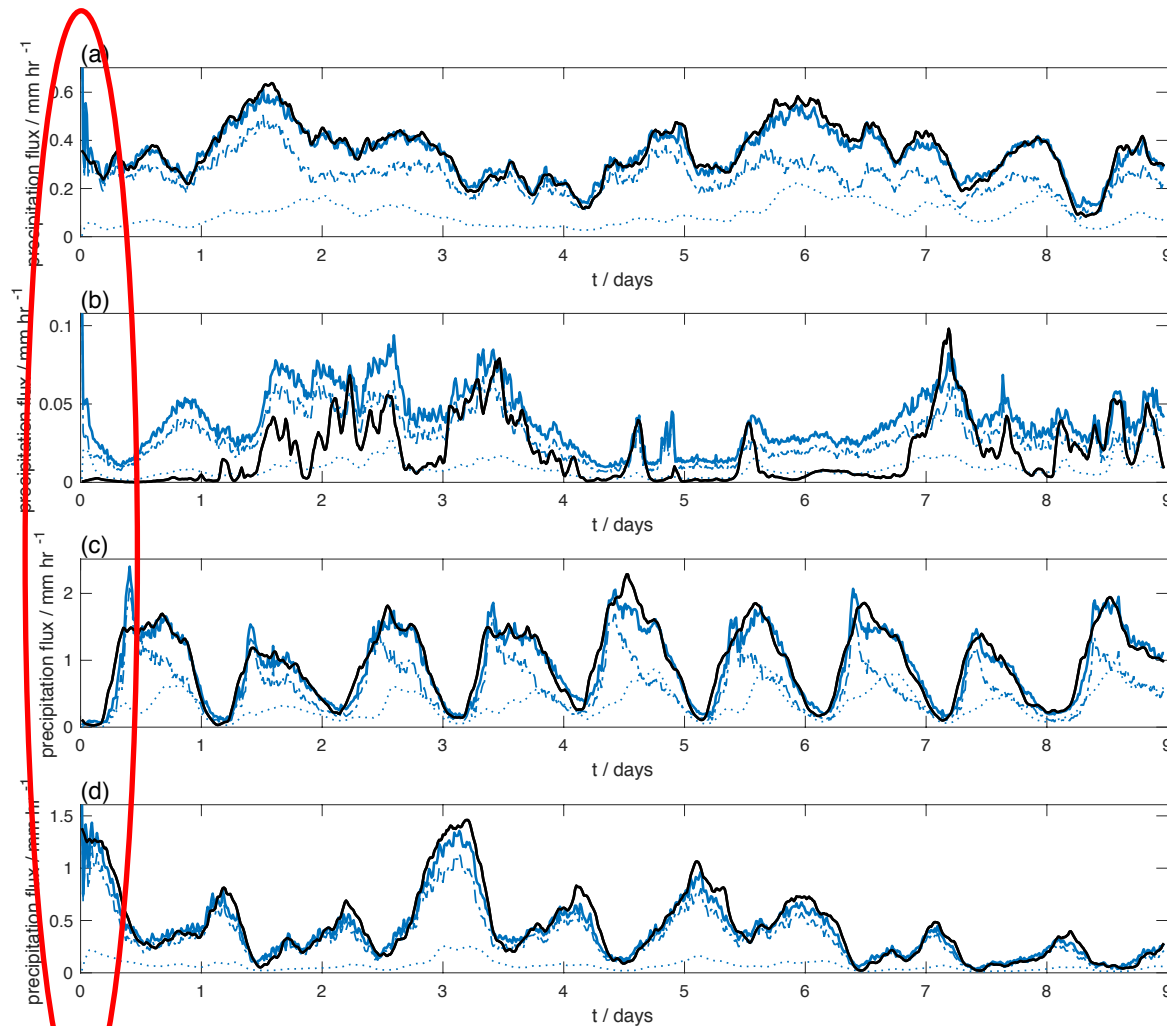
# How does the SCM compare to Cascade?

W. Pacific  
Ocean

Ocean  
West of  
Australia

Maritime  
Continent  
Land

Maritime  
Continent  
Ocean



Precipitation  
Total —  
Conv - -  
Strat ...  
CAS —

-> discard first hour of SCM, and compare evolution over 2<sup>nd</sup> hour