ETH zürich

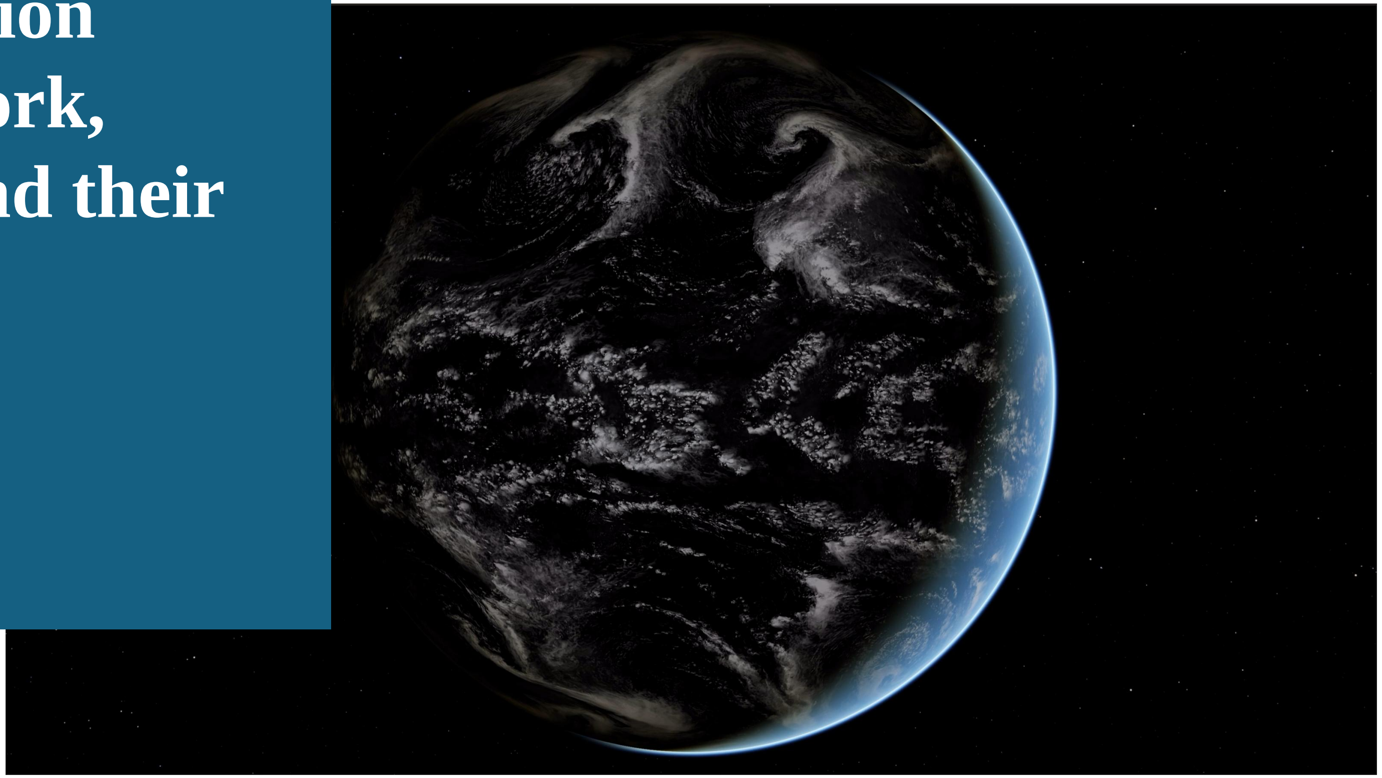
Institute for Atmospheric and Climate Science
Center for Climate System Modeling (C2SM)

Python-Based High-Resolution Global Simulation Framework, Global Monsoon Systems and their Processes.

Praveen Kumar Pothapakula, Andreas Prein, Anurag Dipankar

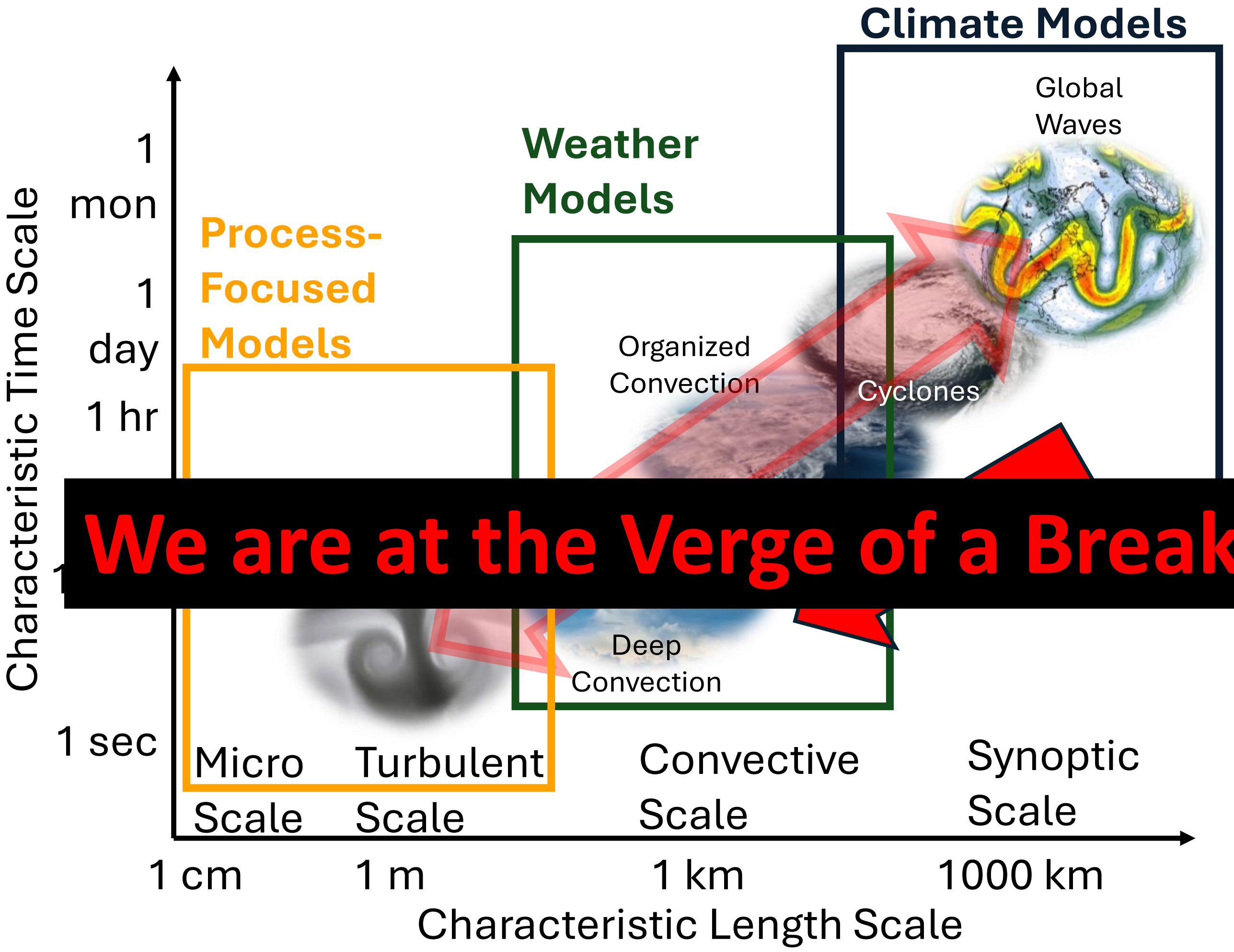
Institute for Atmospheric and Climate Science
ETH Zürich

*Presented at Eighth WMO International Workshop on Monsoons, IITM Pune, India



In collaboration with:
MPI-M/DWD/KIT
COSMO/ICON-CLM Community, Germany.
Software Engineers at EXCLAIM (ETH Zürich)

We need models with local granularity to resolve mesoscale processes globally

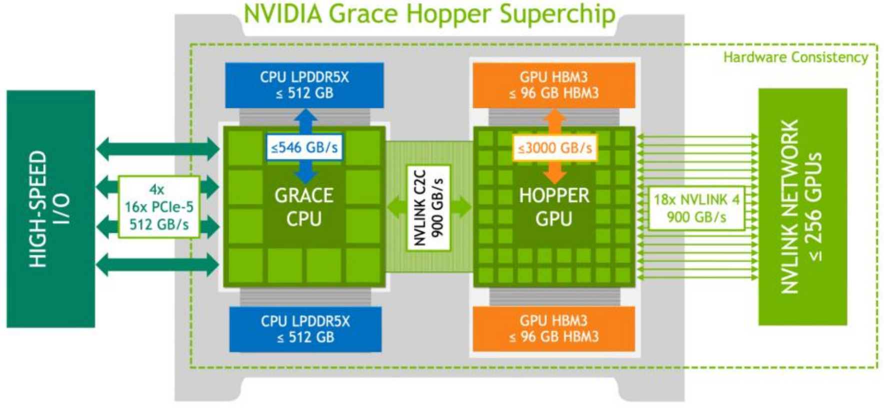


We are at the Verge of a Breakthrough in Climate Modeling!

Kilometre Scale Global Climate Models & EXCLAIM Project at ETH Zürich

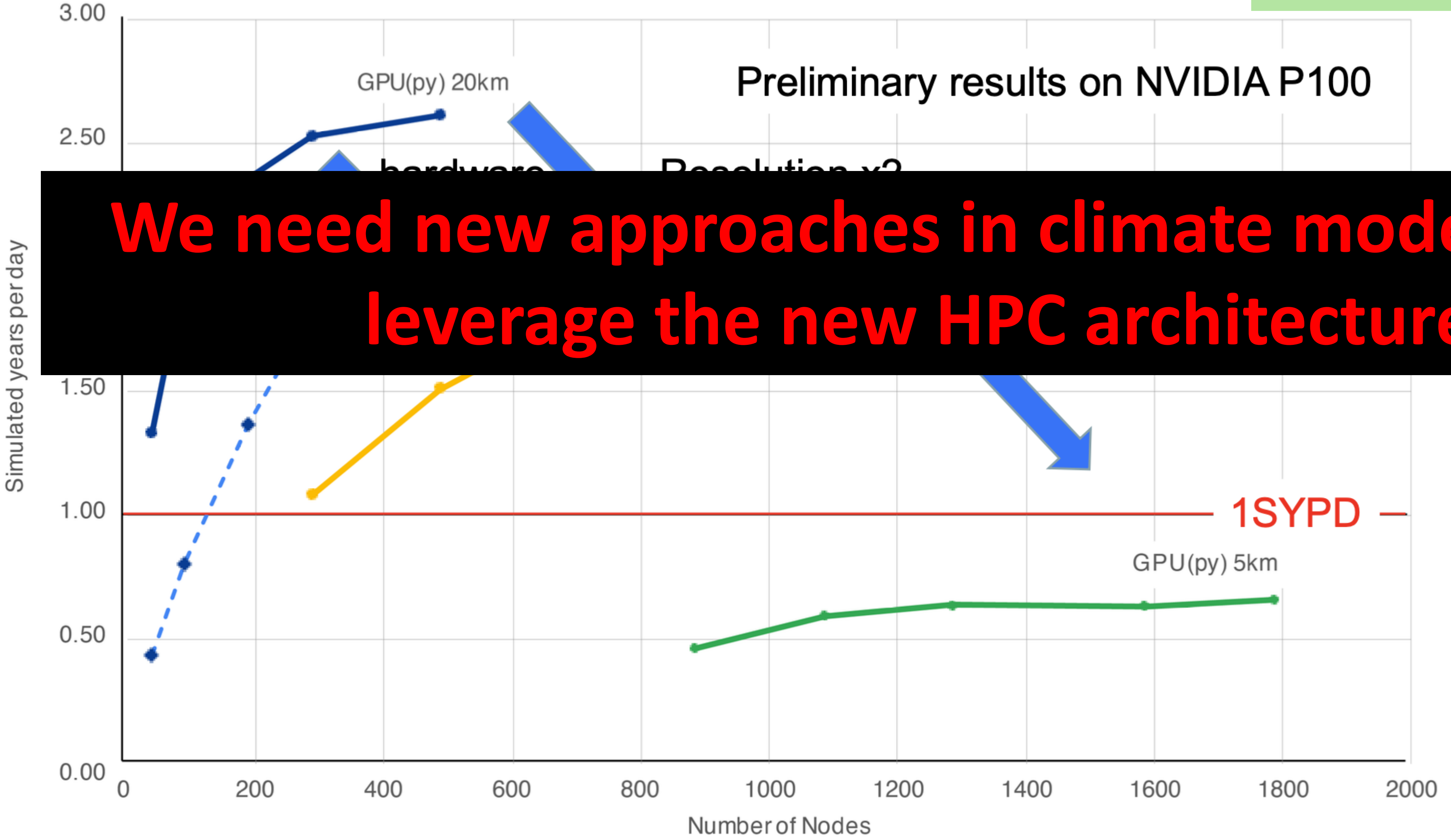
Nodes Overview

# of nodes	# of sockets per node	Total # of sockets	Processor(s)	Specifications	TFlops
2,688	4	10,752	NVIDIA Grace-Hopper	72 ARM cores, 128 GB LPDDR 5X RAM, H100 GPU with 96 GB HBM3 memory	n/a
1,024	2	2,048	AMD EPYC 7742 CPU (Rome)	2x64 cores, 256/512 GB DDR RAM	4,719



Scaling on Pitz Daint (Old) System

GPUs are more energy efficient (approx. > 50%) !



We need new approaches in climate model codes to leverage the new HPC architectures!

The preliminary scaling results show the large advantage of using

are not yet where we want to be. A lot of optimization is still ahead of us…

We need to rethink how we develop climate model codes for new architectures

$$\underline{\nabla}_{\underline{n}}\psi(e) = \frac{\psi(c_1(e)) - \psi(c_0(e))}{\hat{l}}$$

(very) straight forward implementation
"actual science" + mesh

```
DO jk = slev, elev
 DO je = i_startidx, i_endidx
   grad_norm_psi_e(je,jk) =
      (psi_c(iidx(je,2),jk)-psi_c(iidx(je,1),jk))/lhat(je)
 ENDDO
END DO
```

Now code needs to target another architecture (eg., GPUs)

```
#ifdef _OMP
!$OMP ....
#else
!$ACC ....
#endif
DO jb = i_startblk, i_endblk
 CALL get_indices_e(ptr_patch, ...)
 #ifdef __LOOP_EXCHANGE
 DO je = i_startidx, i_endidx
   DO jk = slev, elev
 #else
   DO jk = slev, elev
     DO je = i_startidx, i_endidx
 #endif
     grad_norm_psi_e(je,jk,jb) =  &
       ( psi_c(iidx(je,jb,2),jk,iblk(je,jb,2)) -
         psi_c(iidx(je,jb,1),jk,iblk(je,jb,1)) )
       / ptr_patch%edges%lhat(je,jb)
   ENDDO
 END DO
END DO
#ifdef _OMP
!$OMP ...
#else
!$ACC ...
#endif
```

Scientist and performance engineers work on the same code!

GT4Py – A Domain Specific Language (DSL) code which separates the concerns and converts the python code into C++

```
@field_operator
def _grad_norm(psi: Field[[CellDim, KDim], float], lhat: Field[[EdgeDim], float],
) -> Field[[EdgeDim, KDim], float]:
    return (psi(E2C(1)) - psi(E2C(0)))/lhat
```

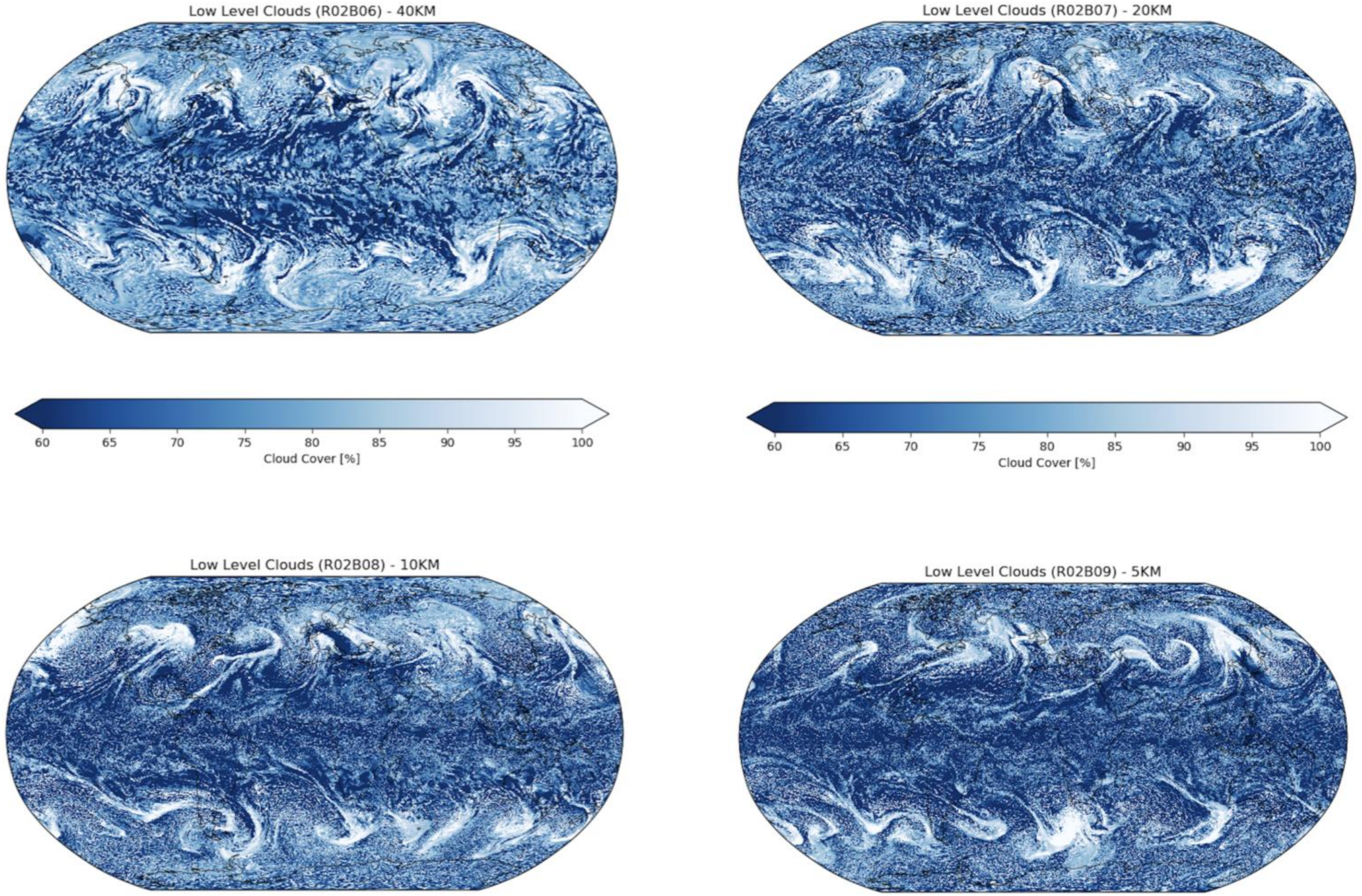
```
grad_norm(psi, lhat, grad_norm_psi_n, offset_provider = {"E2C": e2c_offset_provider})
```

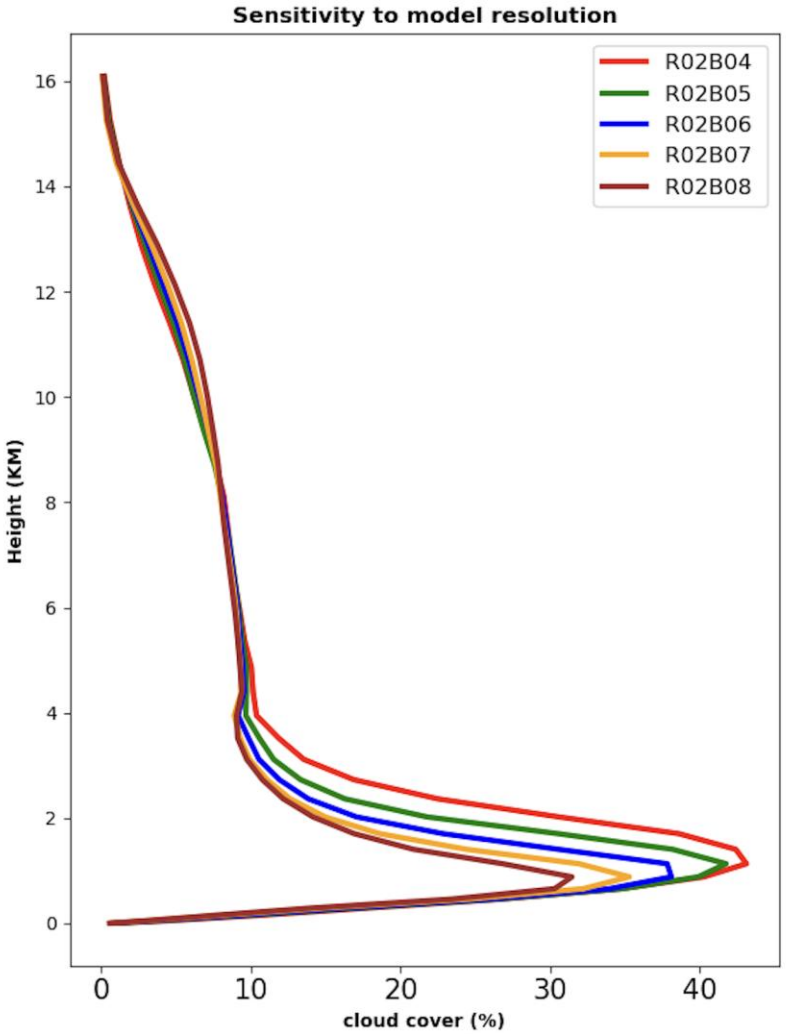
EXCLAIM targets to refactor the ICON code with a DSL

DSL efficient codes are essential for increasingly complex hardware architectures!

Figure 1: Planned evolution of ICON from Fortran/OpenACC (blue) to Python/GT4Py (orange). The steps are: (1) explore the use of Python/GT4Py with one or a few parameterizations; (2) replace the dynamical core of ICON with a Python/GT4Py-based implementation; (3) to make available the whole (or most) of ICON in Python.

Low level clouds – Aqua Planet SImulations with ICON DSL dycore





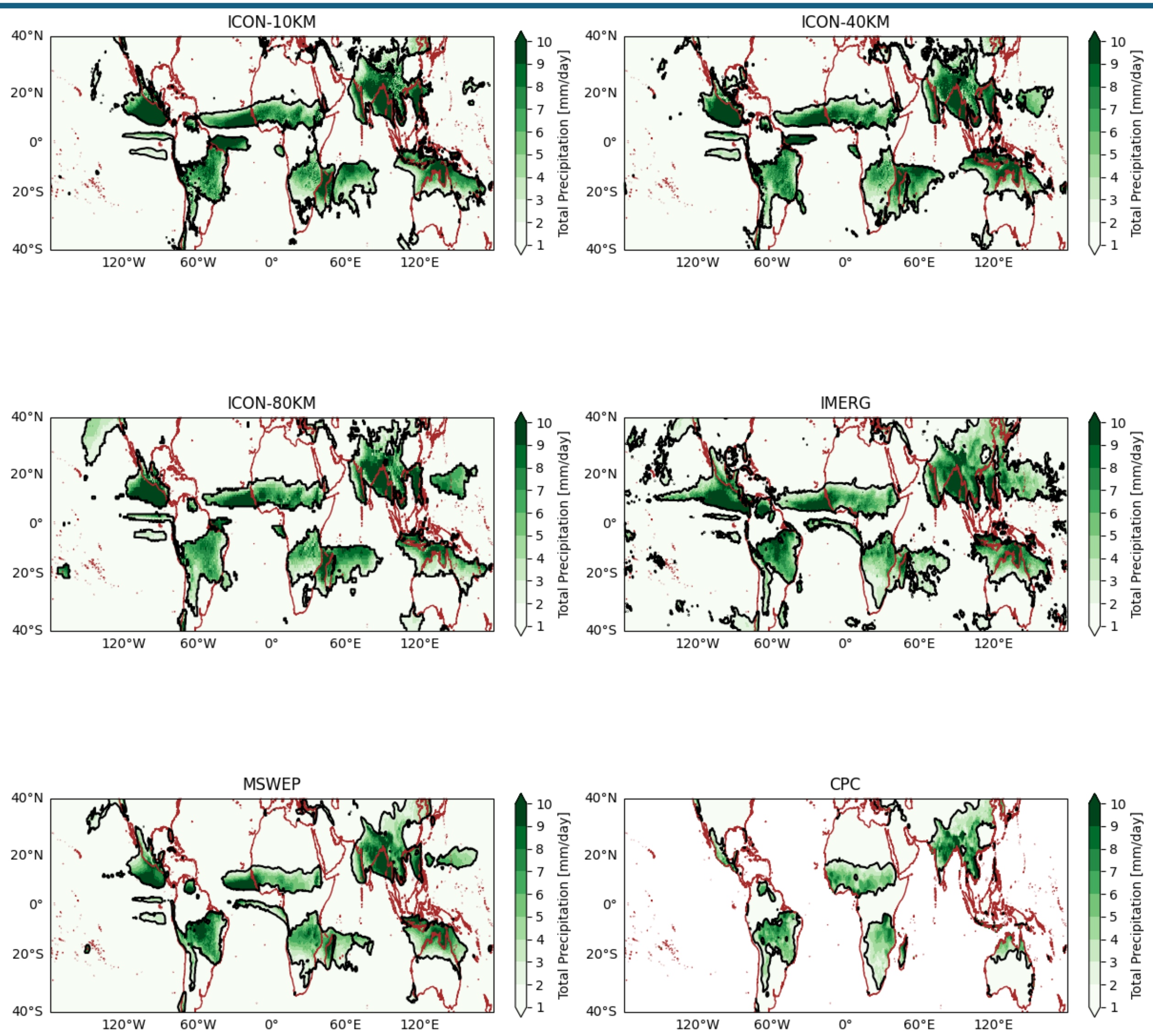
ETH zürich

→ As we increase the resolution ↑ the low level cloud cover ↓ decreases

Prescribed SST test simulations with ICON DSL code at 10km resolution for 10 years.

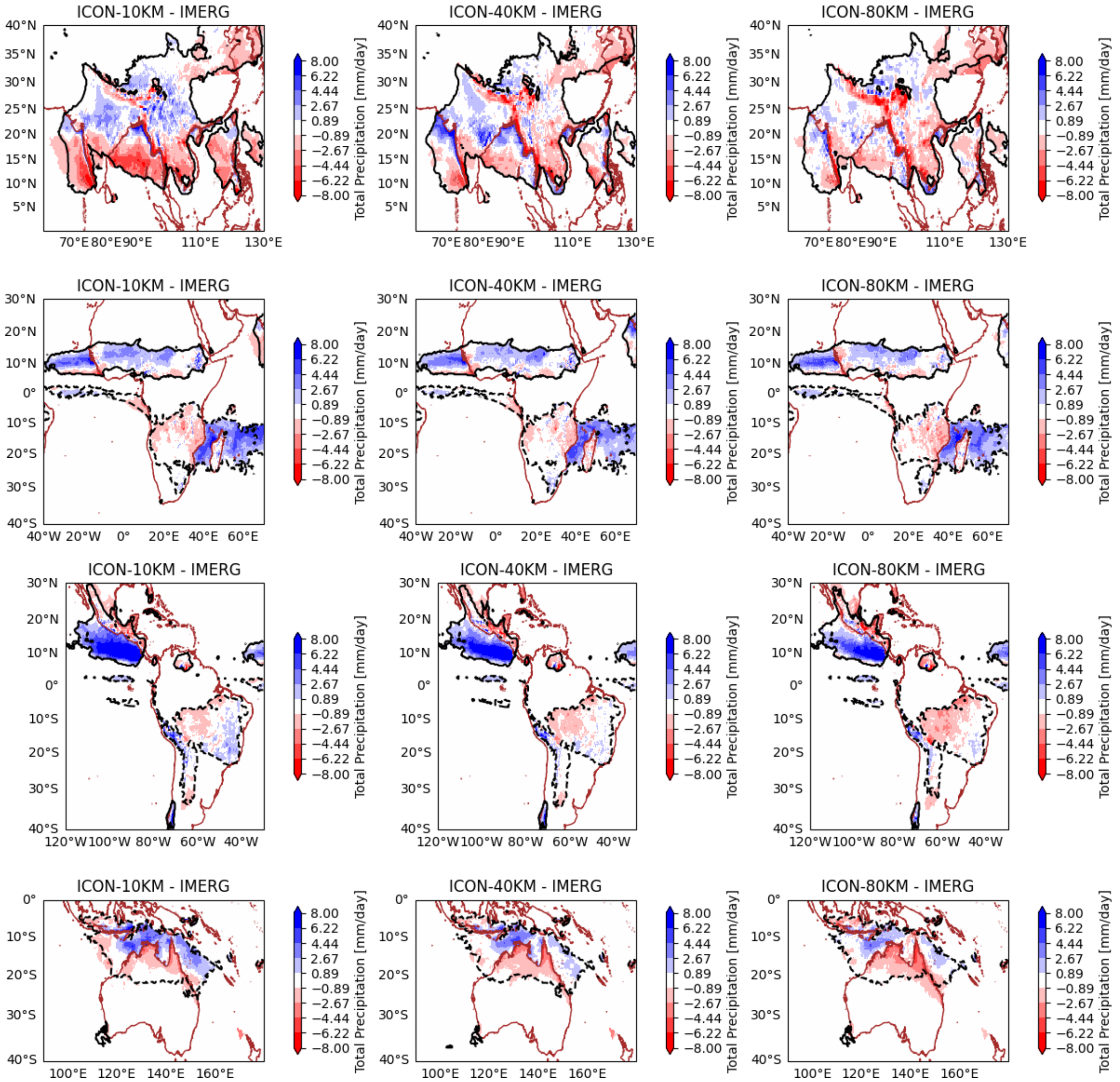
- ICON-NWP + TERRA configuration with resolutions 80km, 40km & 10km (R02B05/06/08).
- Investigating years: 2007-2016 (10 Years).
- Intitialized by ERA5 state on 2005/11/01, with prescribed monthly SST
- Convection is switched on in all the simulations.

Global Monsoon Domains - Mean Precipitation



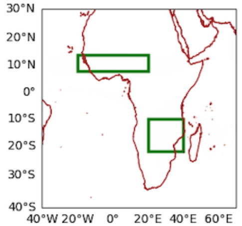
The contours indicate regions where local summer minus winter precipitation exceeds 2 mm/day and summer accounts for at least 55% of the annual total precipitation (cf. Wang & Ding, 2008; Wang et al., 2014).

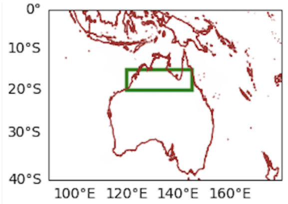
Global Monsoon Domains - Mean Bias of Precipitation

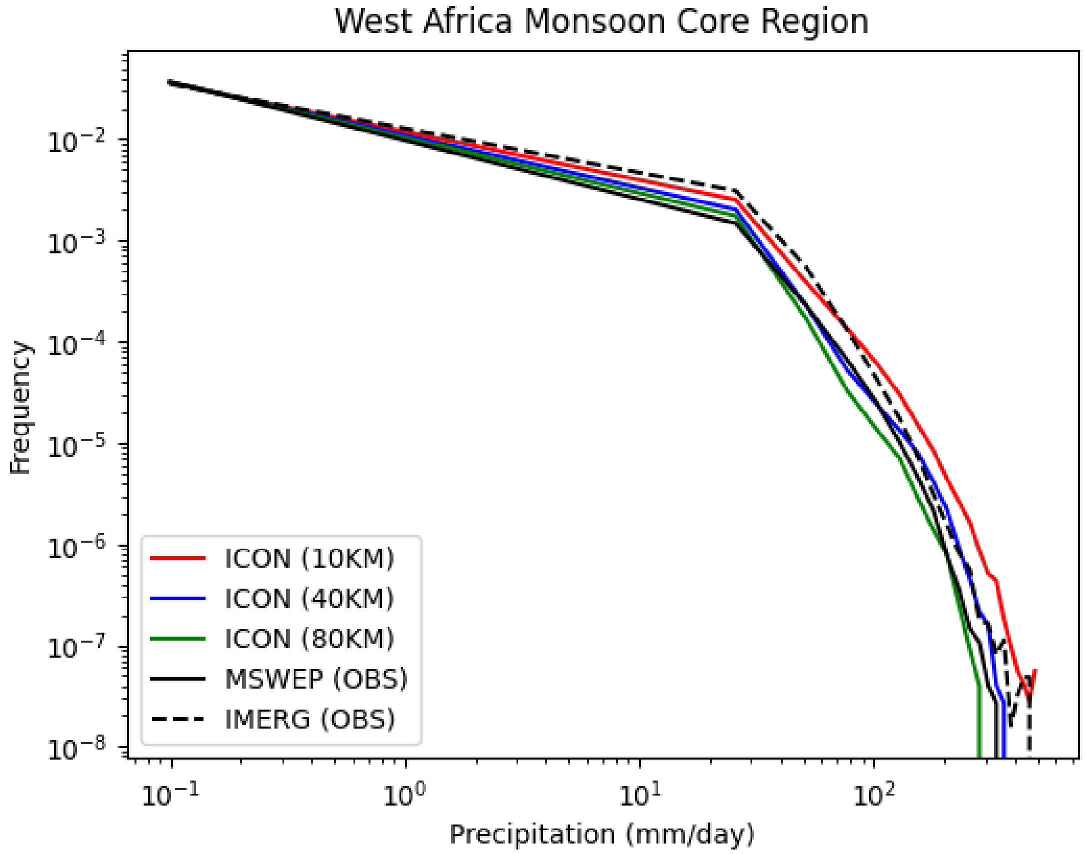


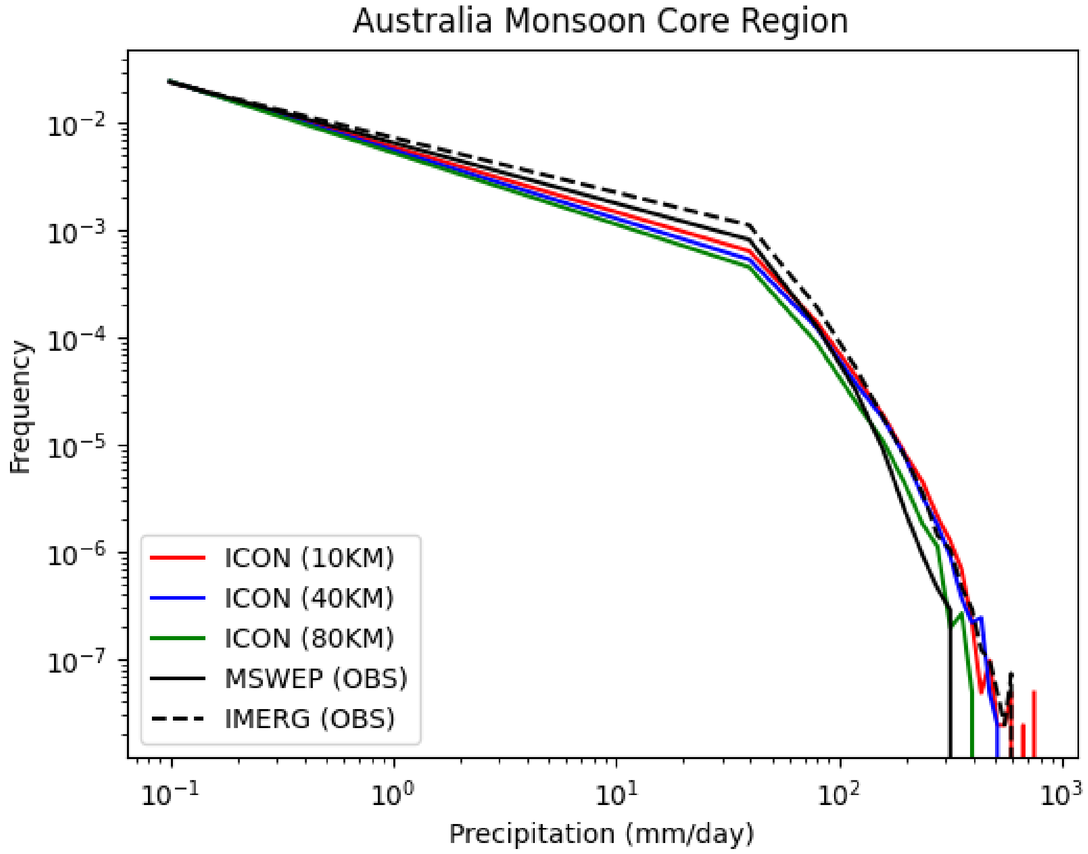
On land: Wet Bias in NH, Dry Bias over SH with exceptions indicating non systematic region dependent bias.

West Africa/Australia Monsoon Core region - Probability Distribution



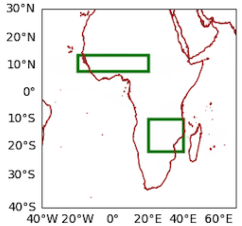


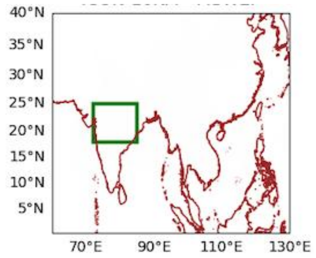


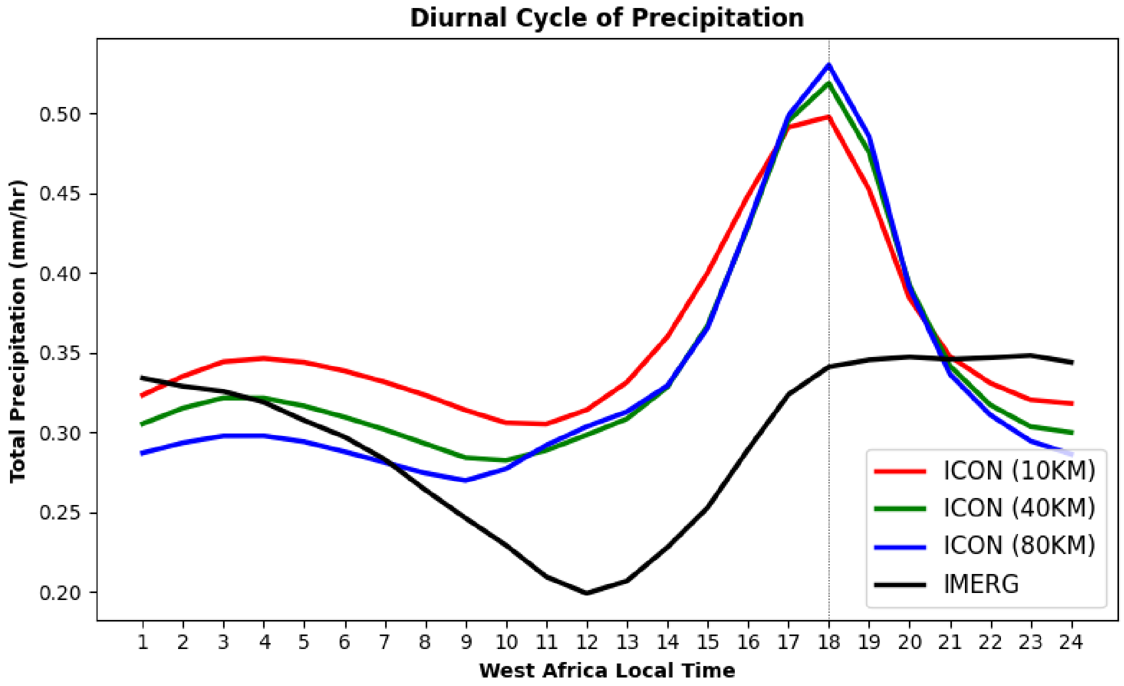


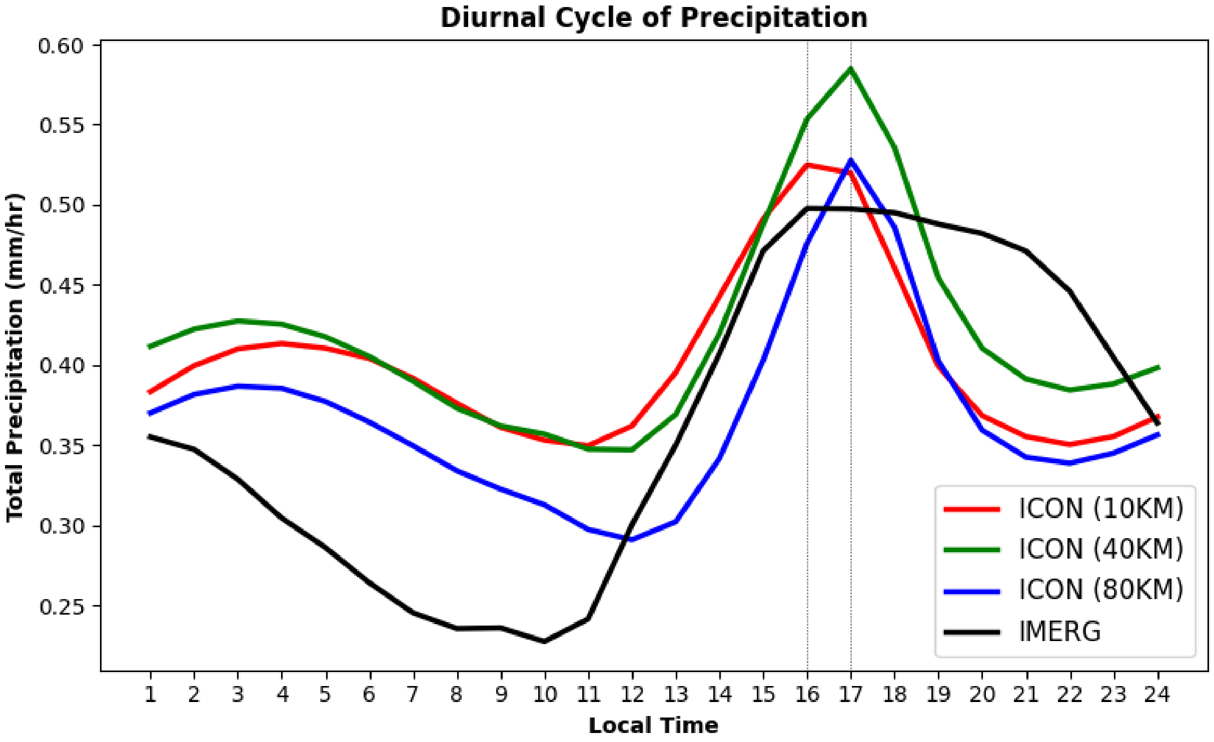
ICON high resolution simulation shows higher probability of extreme precipitation but does not necessarily agree with the observations. (keep in mind the observational uncertainty)

Diurnal Cycle of Hourly Precipitation in ICON compared with Observation (IMERG) in JJAS



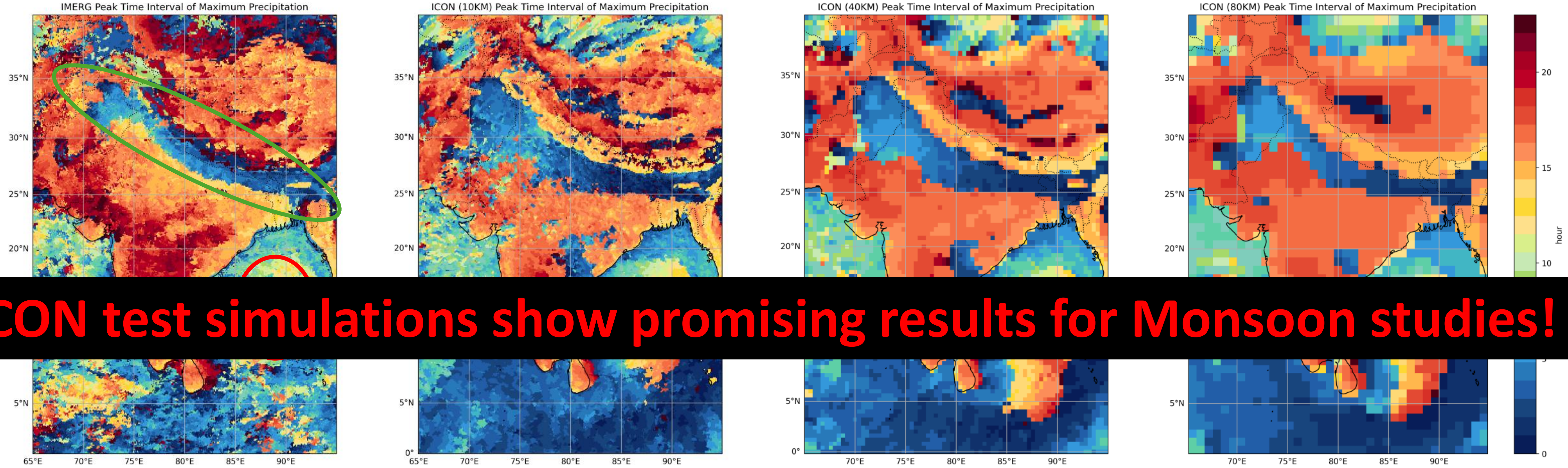






The dominant pattern of diurnal variability is a maximum in rainfall over land during the afternoon/evening in response to solar heating of the surface in general is well captured in ICON

Peak Time of Diurnal Cycle in ICON and Observations over South Asia (IST) in JJAS



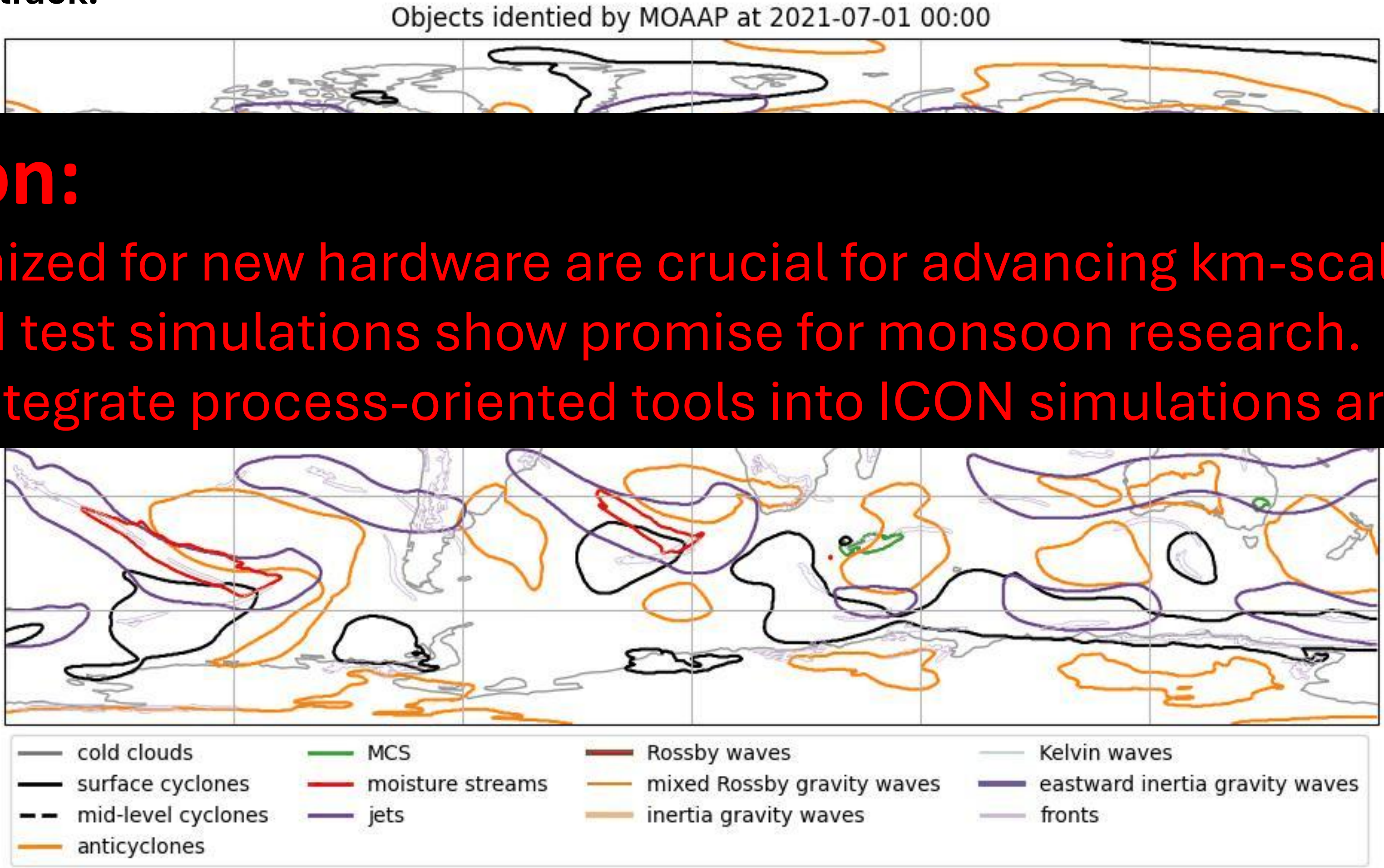
ICON test simulations show promising results for Monsoon studies!

➡ Morning maximum over the oceans, and early afternoon peak over land; however, there are important exceptions to this pattern over both land and ocean.

MOAAP
Multi-Object Analysis of Atmospheric Phenomena

Uses 12 standard model output variables to identify and track:

1. Surface cyclones
2. Mid-level cyclones
3. C
4. T
5. J
6. A
7. A
8. N
   C
   Systems
9. Fronts
10. Equatorial Rossby waves
11. Kelvin Waves
12. Eastward inertia gravity waves
13. Inertia gravity waves
14. Mixed Rossby gravity waves

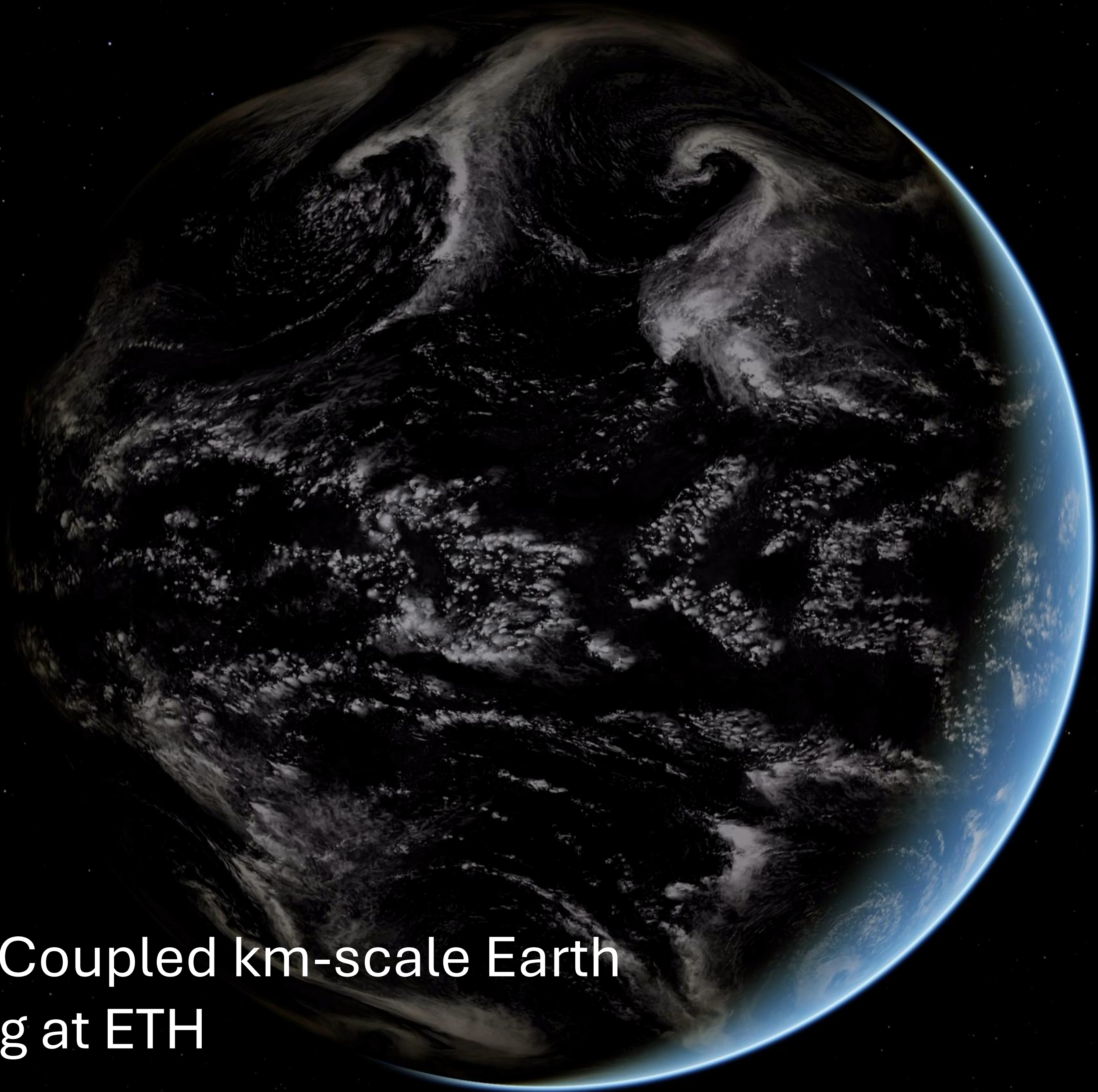


GitHub

Conclusion:

- DSLs optimized for new hardware are crucial for advancing km-scale models
- Initial ICON test simulations show promise for monsoon research.
- Efforts to integrate process-oriented tools into ICON simulations are in progress.

Prein et al. (2023; Earths Future)



Thank you!

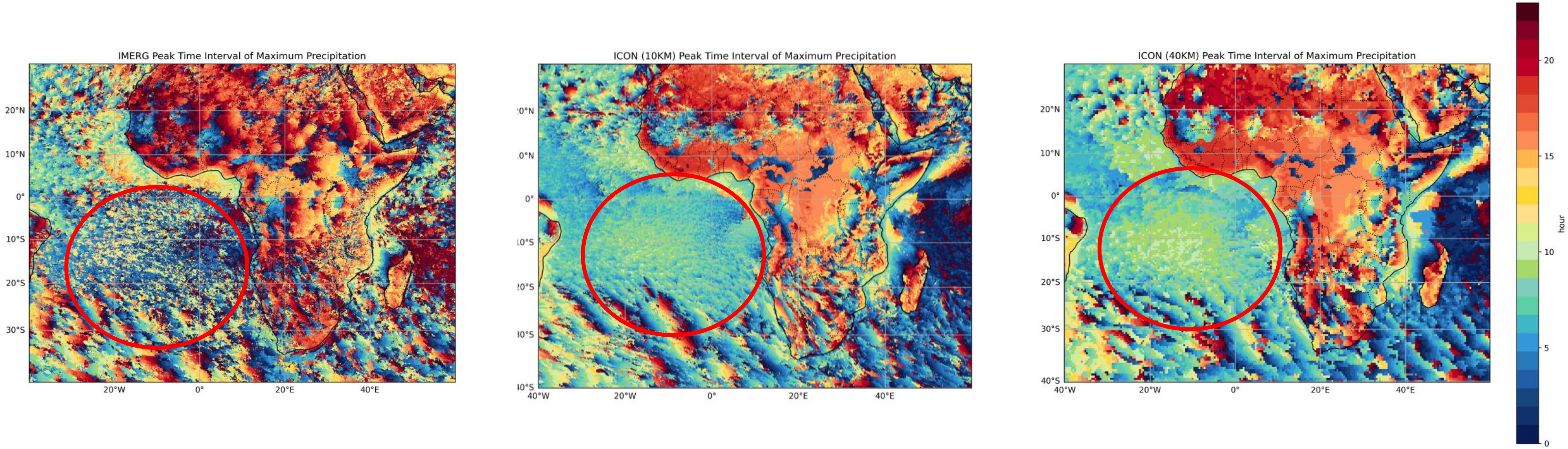
Towards Global Coupled km-scale Earth System Modeling at ETH

ETH zürich

XCLAIM

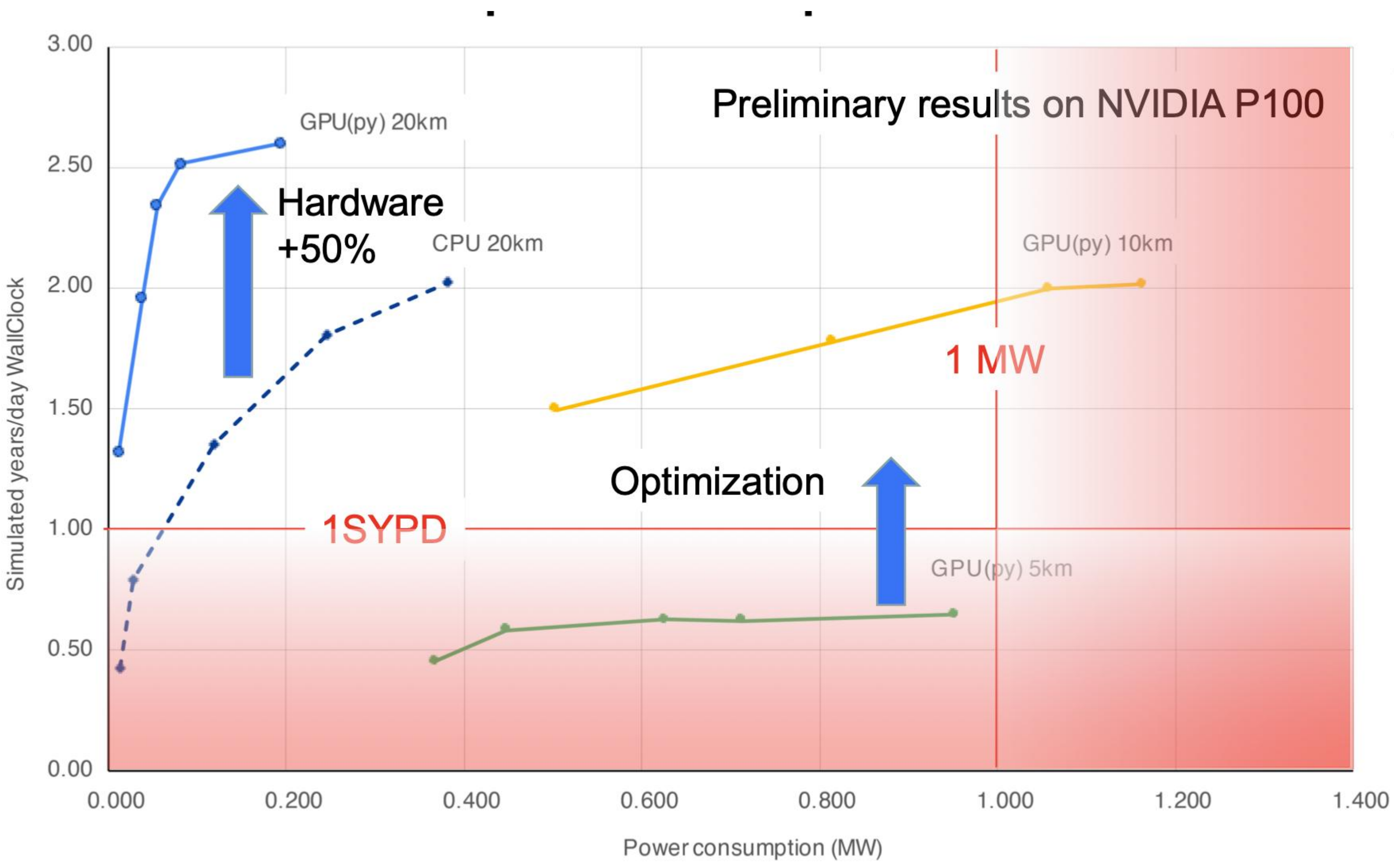
NVIDIA

Peak Time in Diurnal Cycle in ICON and Observations over Africa (Local time) in JJAS



Over the ocean we see systematic differences, we need to understand why?

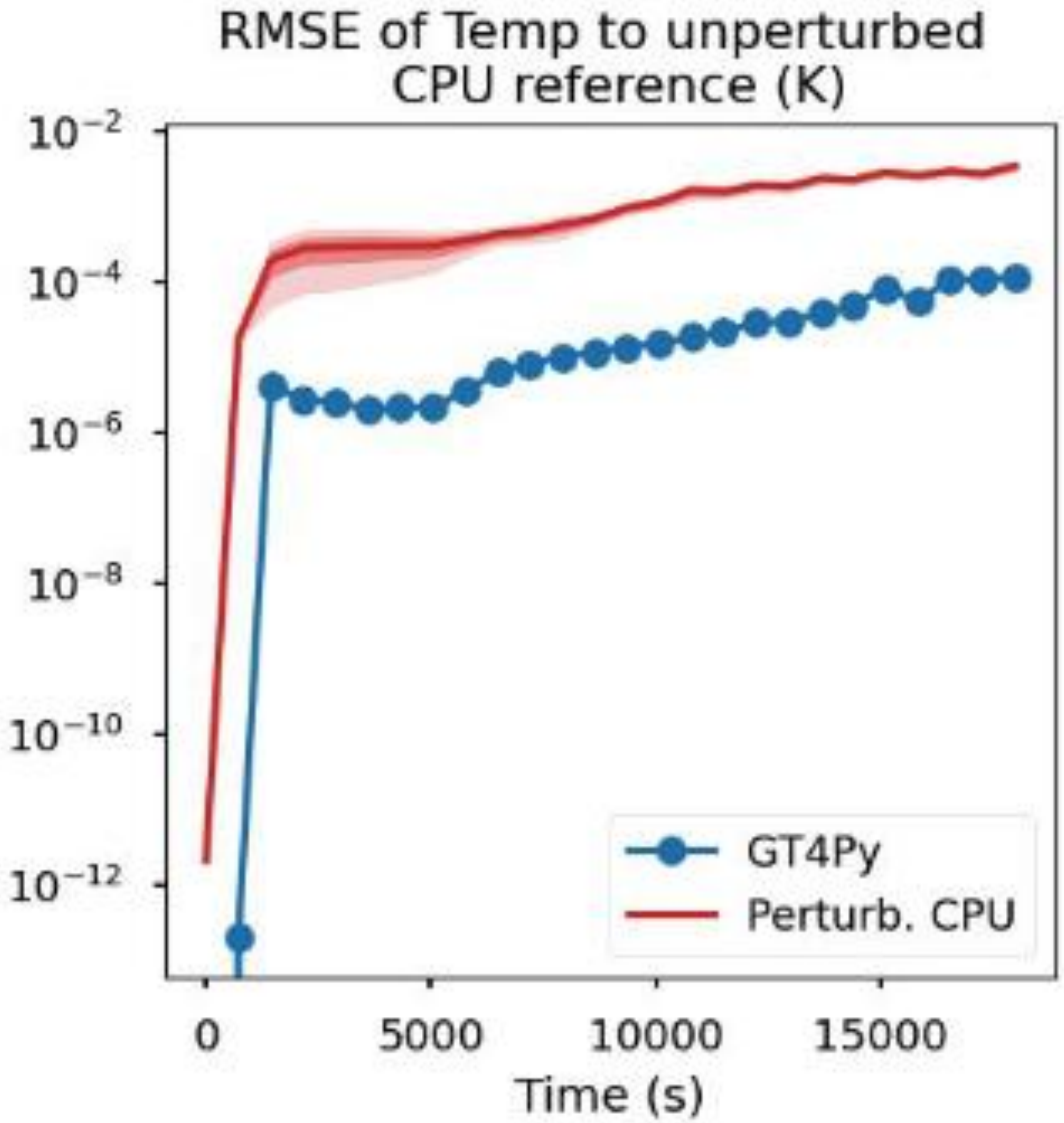
Power consumption on Old Pitz Daint System



The power consumption of the simulations is increasingly becoming a strong constraint on the design and optimization of our modeling systems. For example 1 MW constraint is equivalent to 2000 to 4000 t CO_2 emissions per year (assuming 100% fossil fuel). Our expected optimization of a factor of 10 in the next year make a 2 km run feasible.

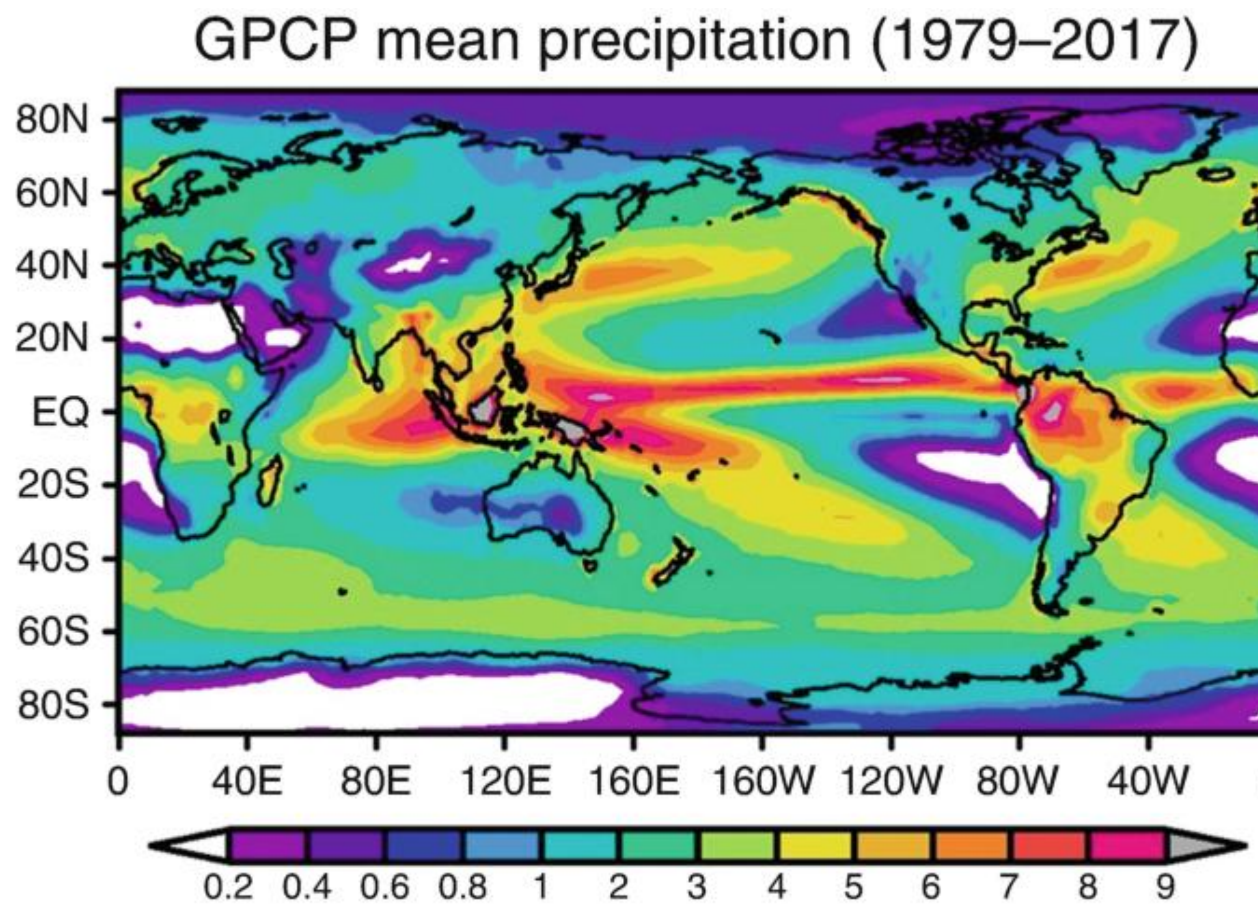
Simulation	Setup	Resolution	Duration
Aqua Planet	Global atmosphere only, no land	10 km, 5 km, 2.5 km, 1 km	2 years
Global Uncoupled	Global atmosphere and land, prescribed SSTs	10 km, 5 km, 2.5 km	5-10 years
Global Coupled	Global, ocean, sea-ice, land, atmosphere	10 km, 5 km, 2.5 km	3 decades to century
Regional Climate Europe (Scenarios CH202X)	Europe (CORDEX domain)	12 km 1 km	century

Still long way to go ...



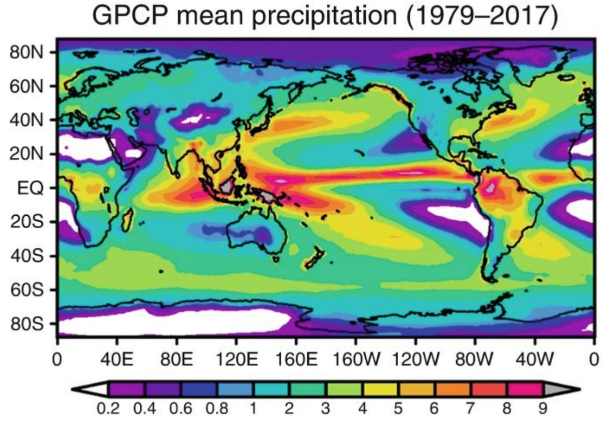
- Take a reference CPU run
- Pertrub 20CPU ensembles where intial perturbations of O(10^{-14})
- Calculate the RMSE of the perturbed CPU runs with reference CPU
- Calculate the GT4Py run RMSE comapred to reference CPU.
- Ideally the RMSE of GT4Py should be less than the purturbed test.

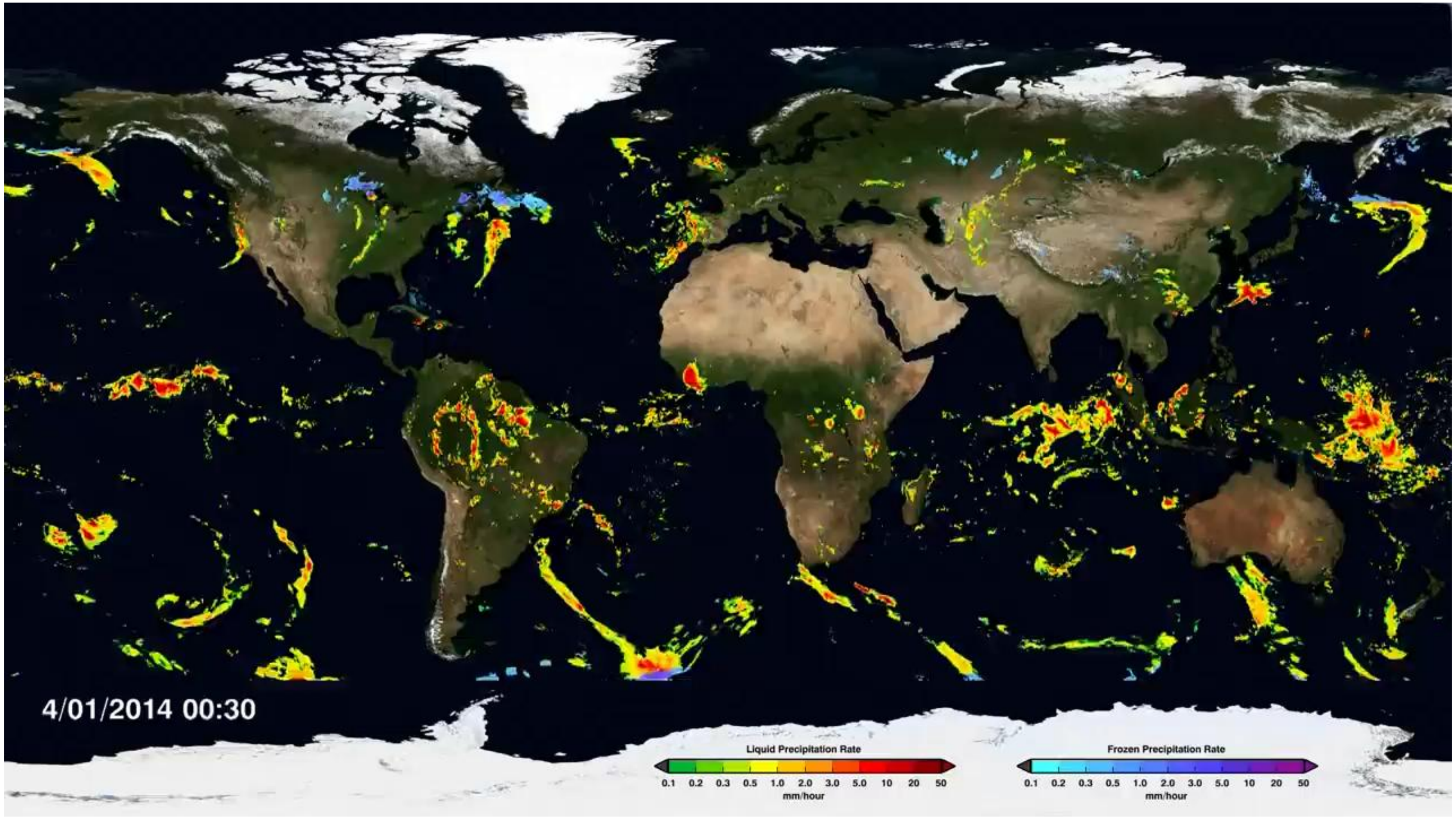
Why do we want to use high resolution (km-scale) Models for Earth System Predictions?



Why do we want to use km-scale Models for Earth System Predictions?

Weather and Climate Interface





For better CPU cache performance, add blocks

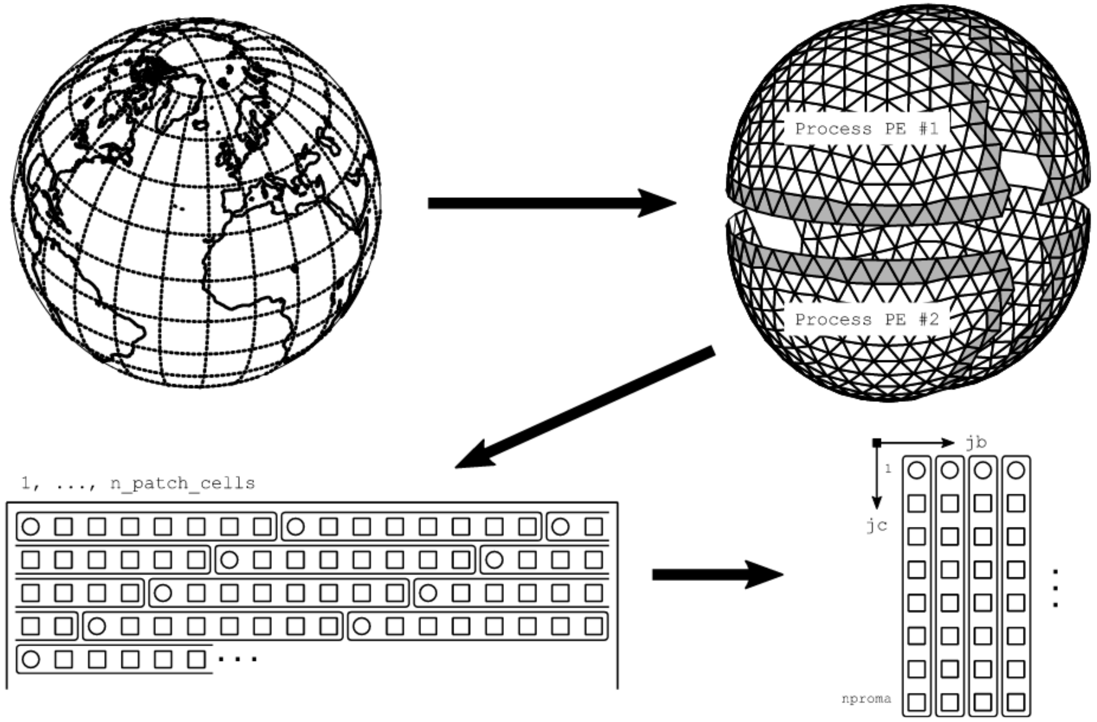


Figure 9.1.: Illustration of the 2D field representation. The original spherical domain is decomposed (light-gray: halo region). Afterwards, the long vector of grid cells is split into several chunks of a much smaller length **nproma**.

```
DO jb = i_startblk, i_endblk
 CALL get_indices_e(ptr_patch, jb, i_startblk, i_endblk, &
                    i_startidx, i_endidx, rl_start, rl_end)
 DO jk = slev, elev
   DO je = i_startidx, i_endidx
     grad_norm_psi_e(je,jk,jb) =  &
       ( psi_c(iidx(je,jb,2),jk,iblk(je,jb,2)) -
         psi_c(iidx(je,jb,1),jk,iblk(je,jb,1)) )
       / ptr_patch%edges%lhat(je,jb)
   ENDDO
 END DO
END DO
```

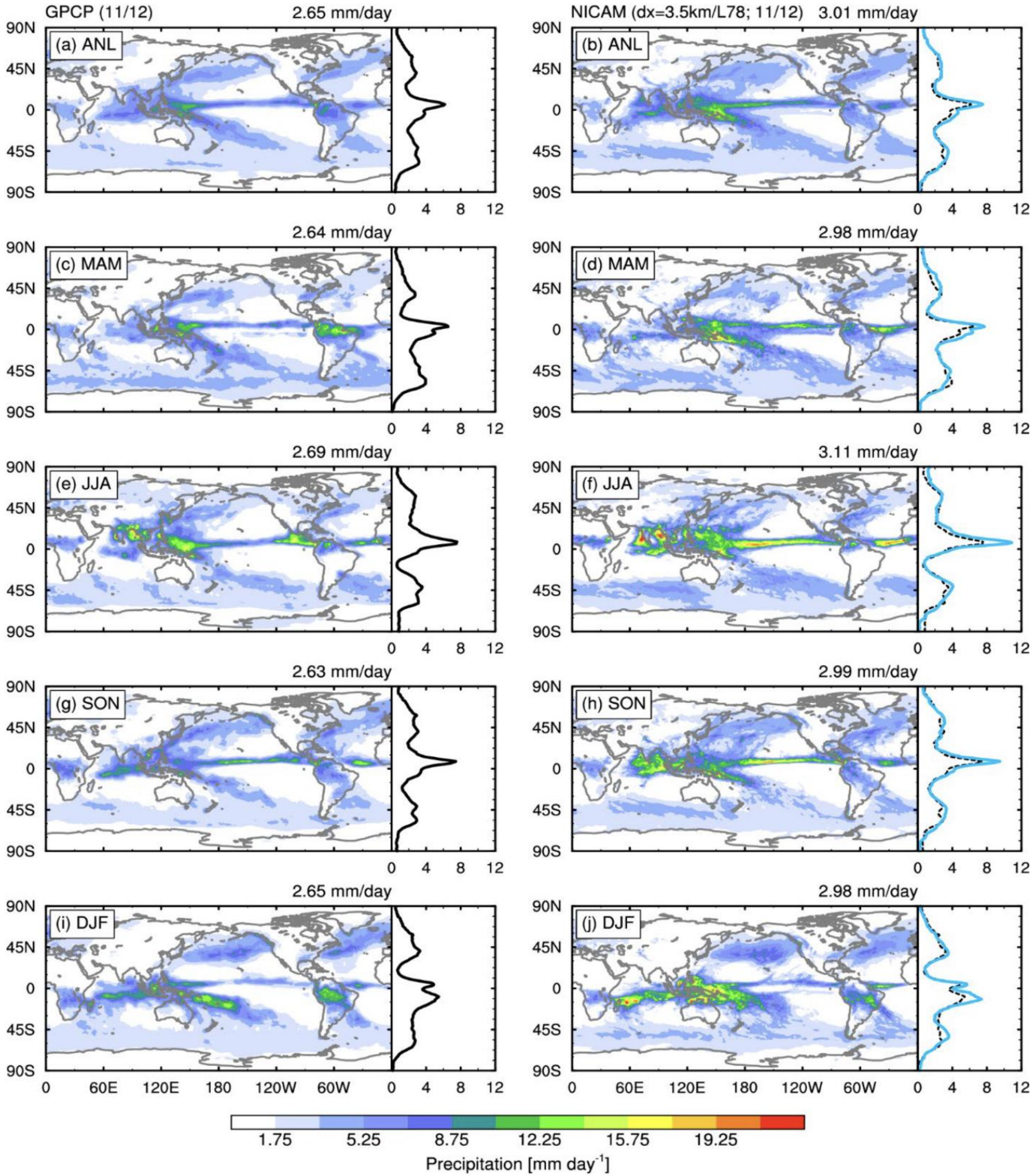
To conclude ....

1. International efforts such as CORDEX, DYAMOND bring communities together and strive for multi-model studies.

2. Information Theory has potential in unravelling interactions in sub-components, but a lot of work ahead!

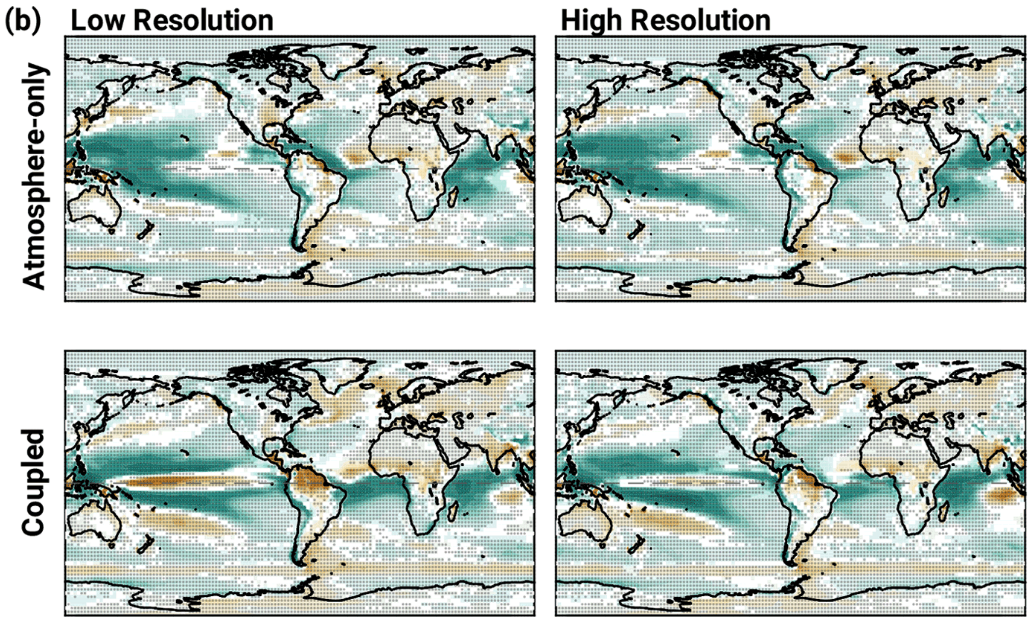
3. Kilometre Scale Climate Models are the future! Still lot of challenges ahead.

Climate modeling & analysis is very interdisciplinary in nature, and even more in the coming years and decades!

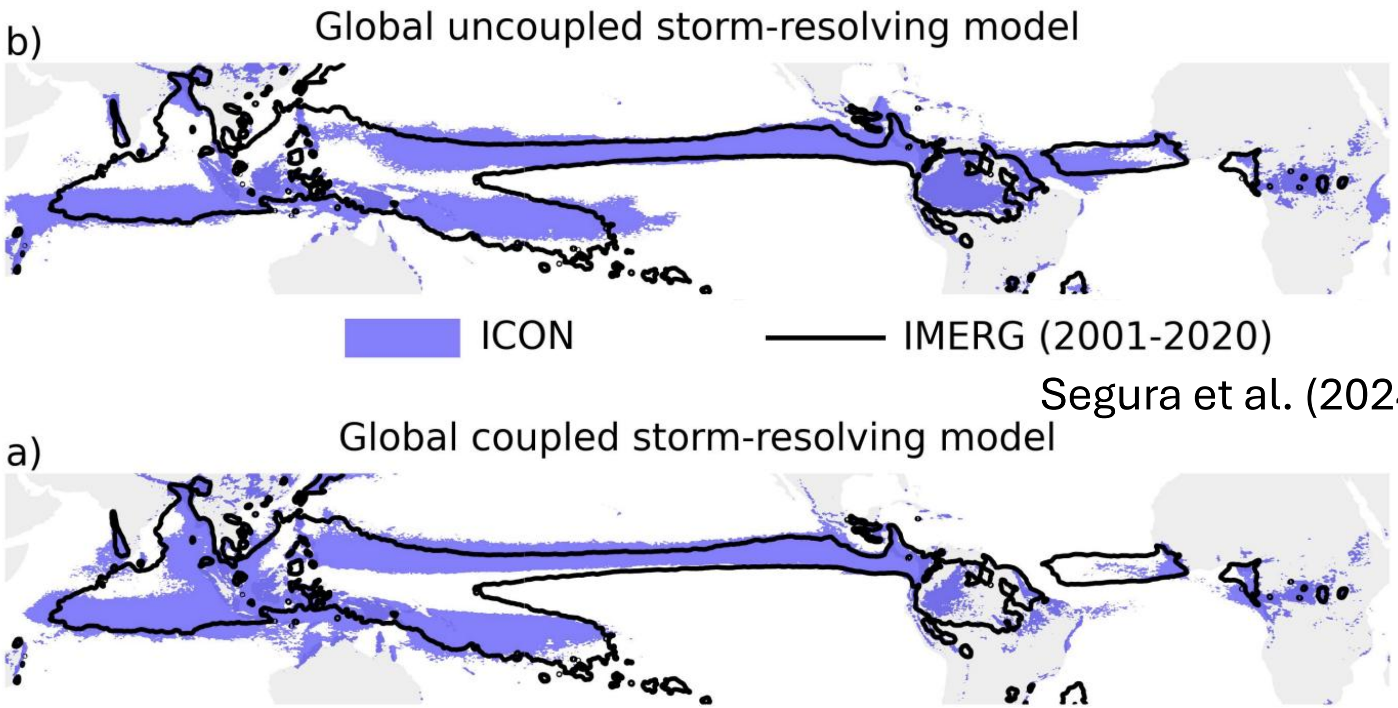


GPCP (11/12)
2.65 mm/day
NICAM (dx=3.5km/L78; 11/12)
3.01 mm/day
(a) ANL
(b) ANL
2.64 mm/day
2.98 mm/day
(c) MAM
(d) MAM
2.69 mm/day
3.11 mm/day
(e) JJA
(f) JJA
2.63 mm/day
2.99 mm/day
(g) SON
(h) SON
2.65 mm/day
2.98 mm/day
(i) DJF
(j) DJF
90N
45N
0
45S
90S
0
60E
120E
180
120W
60W
0 4 8 12
1.75
5.25
8.75
12.25
15.75
19.25
Precipitation [mm day-1]

Km-Scale global models are not solving the double ITCZ bias



Moreno-Chamarro et al. (2022)



Segura et al. (2024)