Advanced numerical methods for Earth System Modelling

Nils Wedi
European Centre for Medium-Range Weather Forecasts (ECMWF)

Thanks to many colleagues at ECMWF and in particular …

Further reading:

TECHNICAL MEMORANDUM

760

The modelling infrastructure of the Integrated Forecasting System: Recent advances and future challenges

N.P. Wedi, P. Bauer, W. Deconinck, M. Diamantakis, M. Hamrud, C. Kühnlein, S. Malardel, K. Mogensen, G. Mozdzynski, P.K. Smolarkiewicz

Research Department

November 2015

Special topic paper presented at the 44th session of ECMWF's Scientific Advisory Committee, Reading,UK

European Centre for Medium-Range Weather Forecasts
Europäisches Zentrum für mittelfristige Wettervorhersage
Centre européen pour les prévisions météorologiques à moyen

An all-scale, finite-volume module for the IFS

PIOTR SMOLARKIEWICZ, WILLEM DECONINCK, MATS HAMRUD, CHRISTIAN KÜHNLEIN, GEORGE MOZDZYNSKI, JOANNA SZMELTER, NILS WEDI

ECMWF hosts the European Research Council-funded project PantaRhei, which explores novel numerical methods to complement existing, highly optimised numerical weather prediction (NWP) models. The need for such innovation stems from the fact that state-of-the-art global NWP models using the spectral transform method may become computationally inefficient at very fine resolutions due to the communication overhead associated with global spectral transformations.

As a first step, we have developed an autonomous, all-scale numerical module which uses the finite-volume method (Box A) to supplement ECMWF's Integrated Forecasting System (IFS). This module is compatible with emerging energy-efficient, heterogeneous hardware for high-performance computing (HPC), and it is able to represent elements of real weather on a large range of scales, including cloud-resolving scales.

Motivation

The advance of massively parallel computing in the 1990s and beyond has encouraged finer grid intervals in NWP models. This has improved the spatial resolution of weather systems and enhanced the accuracy of forecasts, while stimulating the development of global non-hydrostatic models. Today many operational NWP models include non-hydrostatic options either for regional predictions or research. However, to date no NWP model runs globally in operations at resolutions where non-hydrostatic effects are important (*Wedi & Malardel*, 2010; *Wedi et al.*, 2012). Such high resolutions are still computationally unaffordable and too inefficient to meet the demands of the limited time window for distributing global forecasts to regional NWP recipients and, ultimately, the public.

Efforts to ensure the computational affordability of global non-hydrostatic forecasts face a twofold difficulty. On fine grids the spectral transform method becomes computationally inefficient because of the required global data-rich inter-processor communications (*Wedi et al.*, 2013). Therefore, simply scaling up the number of processors would be unaffordable, not least due to the huge increase in electric power consumption this would entail.

At the same time, replacing hydrostatic primitive equations (HPE) that have been central to the success of weather and climate prediction exacerbates the efficiency problem. In particular, with the simulated vertical extent of the atmosphere thin compared to its horizontal extent, the vertically propagating sound waves supported by the non-hydrostatic Euler equations, from which HPE derive, impose severe restrictions on the numerical algorithms. The hydrostatic balance assumption underlying HPE conveniently filters out vertically propagating sound waves, therefore permitting large time steps in the numerical integration. Moreover, HPE imply the separability of horizontal and vertical discretisation, thus facilitating the design of effective flow solvers, such as the semi-implicit semi-Lagrangian (SISL) time stepping combined with the spectral-transform spatial discretisation that is used today. Such separability does not apply in non-hydrostatic models.

While NWP strives to extend its skill towards finer scales, non-hydrostatic research models endeavour to extend their realm towards the global domain. The two routes of development must meet, but the way to merge the different areas of expertise is far from obvious. Altogether, NWP is at a crossroads. Although massively parallel computer technology promises continued advances in forecast quality, the latter cannot be achieved by simply applying the existing apparatus of NWP models to ever finer grids.

A new way forward

Recognising the predictive skill of the IFS, we seek to address the challenges outlined above by supplying a complementary non-hydrostatic dynamical module with the capabilities of a cloud-resolving model, concurrently driven by large-scale IFS predictions based on the HPE. The first step towards this paradigm is the development of an autonomous, global, finite-volume, non-hydrostatic dynamical module capable of working on the IFS's reduced Gaussian grid and, in principle, on any horizontal mesh.

Partial differential equations (PDEs) require the calculation of differential operators. In the IFS, spatial differentiation is conducted in spectral space, and there are no practical means of calculating derivatives locally in the physical

Finite-volume method — A

The finite-volume method is an approach to the approximate integration of partial differential equations (PDEs) describing natural conservation laws. Similar to the finite difference method or finite element method, solutions are calculated at discrete places on a meshed geometry.

'Finite volume' refers to the small volume surrounding each node point on a mesh. In PDEs, integrals of the divergence terms over these finite volumes are converted into surface integrals using the Gauss divergence theorem. On a discrete mesh, these surface integrals are then evaluated as a sum of all fluxes through individual surfaces bounding each finite volume (Figure 1).

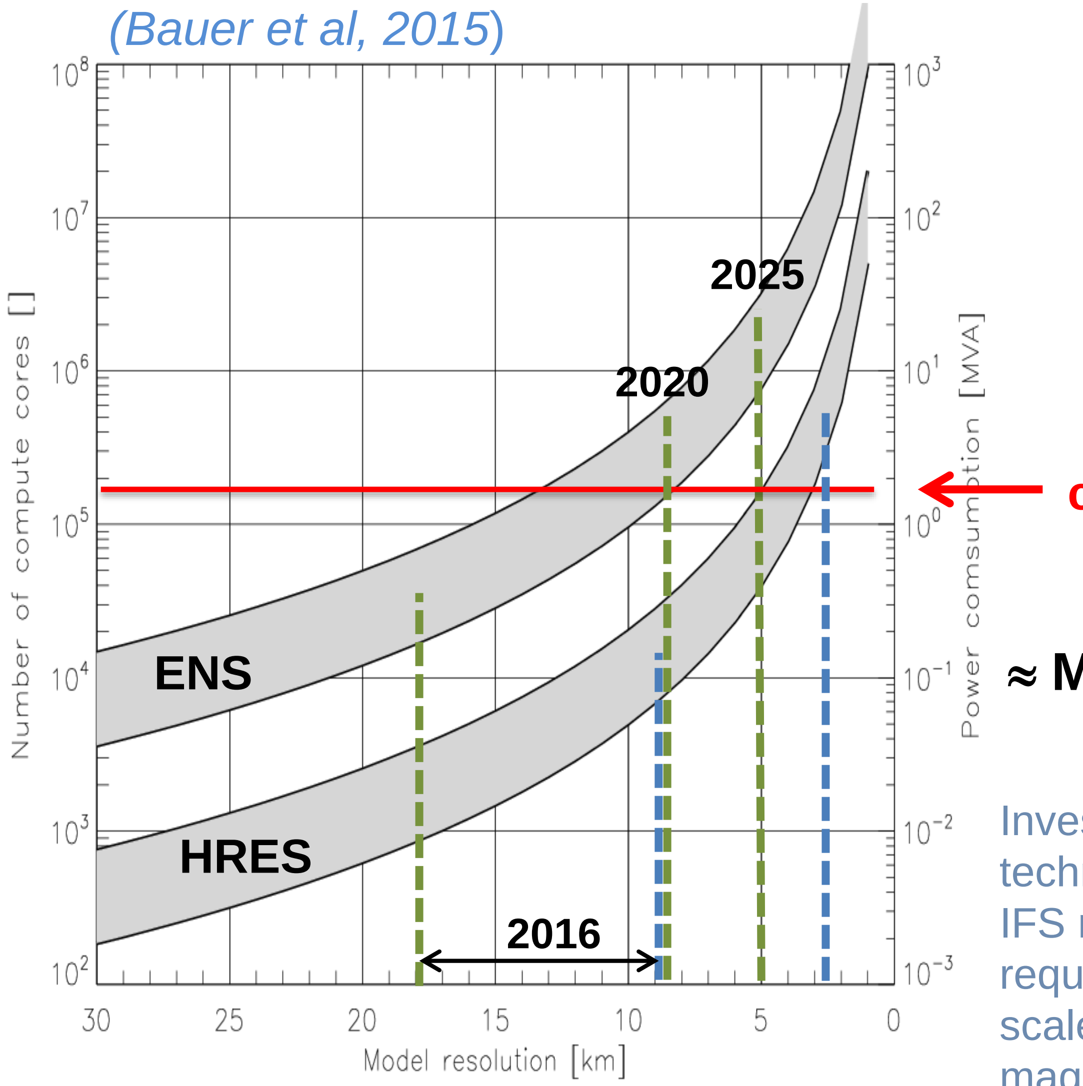
Because the flux entering a given volume is identical to that leaving the adjacent volume, these methods are conservative. An important advantage of the finite-volume method is that it can easily be formulated for unstructured meshes.

Fit for the future …

- Research and development into
  - novel and adaptive numerical methods
  - physics & dynamics coupling
  - quantifying uncertainty
  - Earth-System complexity
- A flexible support infrastructure for European NWP science and services to
  - maximise energy-efficiency
  - assure the best possible use of the available computing resources

(ECMWF Tech Memo 760, Wedi et al. 2015)

Affordability – the art and cost of computing



≈ M€ electricity/year

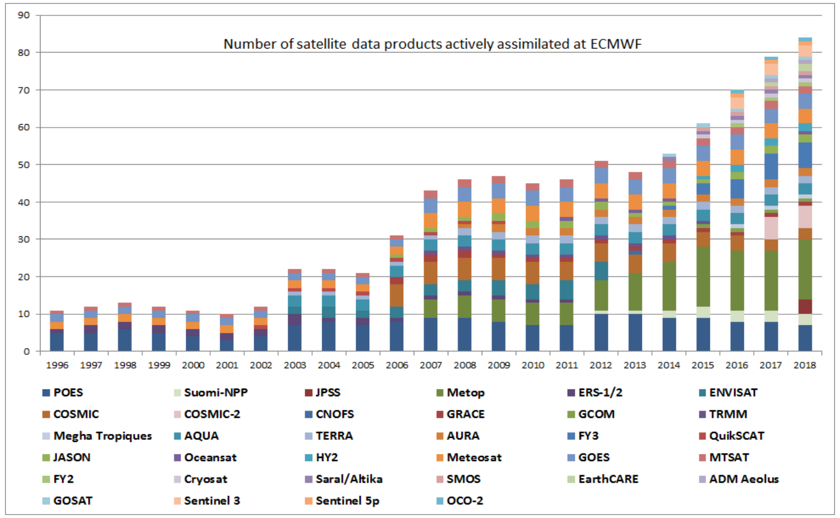
Investment in new hardware technology and continuous IFS model development required to move the energy scale down by one order of magnitude! *(see Aurora 180PF, HPCwire April 2015)*

Big Data challenge: 3Vs

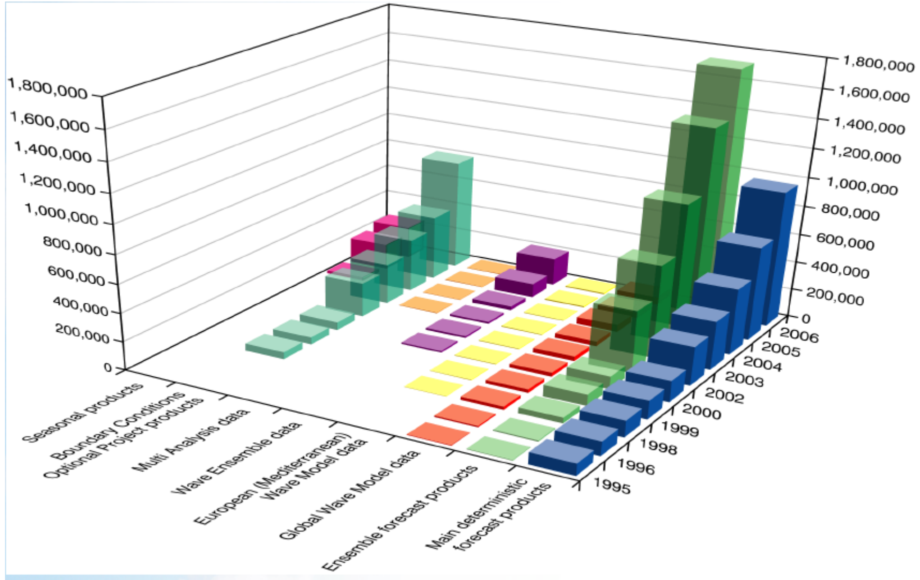
- high **V**olume, high **V**elocity, high **V**ariety

Velocity: Exponentially growing data archive: *1995: 14 Tbytes / year*
2015: 100 Tbytes / day

Volume:
Satellite observations 1996 – 2018 (projected)



Variety: Increase in products



Outline

- ECMWF's integrated forecasting system (IFS)
  - *Assets and the need for development*
  - *Developments at other Centres*
- A flexible, scalable and sustainable model infrastructure
  - *Scalability, time & spatial discretization and numerical methods, equations, transport of species, physics-dynamics coupling, (tangent linear and adjoint model), uncertainty quantification, Earth-System complexity*
- A sneak peek into the future: ~1.3 km global simulations with IFS

Schematic description of the *spectral transform method* in the ECMWF IFS model

No grid-staggering of prognostic variables

Grid-point space
- -semi-Lagrangian advection
- -physical parametrizations
- -products of terms

FFT

Fourier space

LT

Spectral space
- -horizontal gradients
- -semi-implicit calculations
- -horizontal diffusion

Inverse LT

Fourier space

Inverse FFT

FFT: Fast Fourier Transform, LT: Legendre Transform

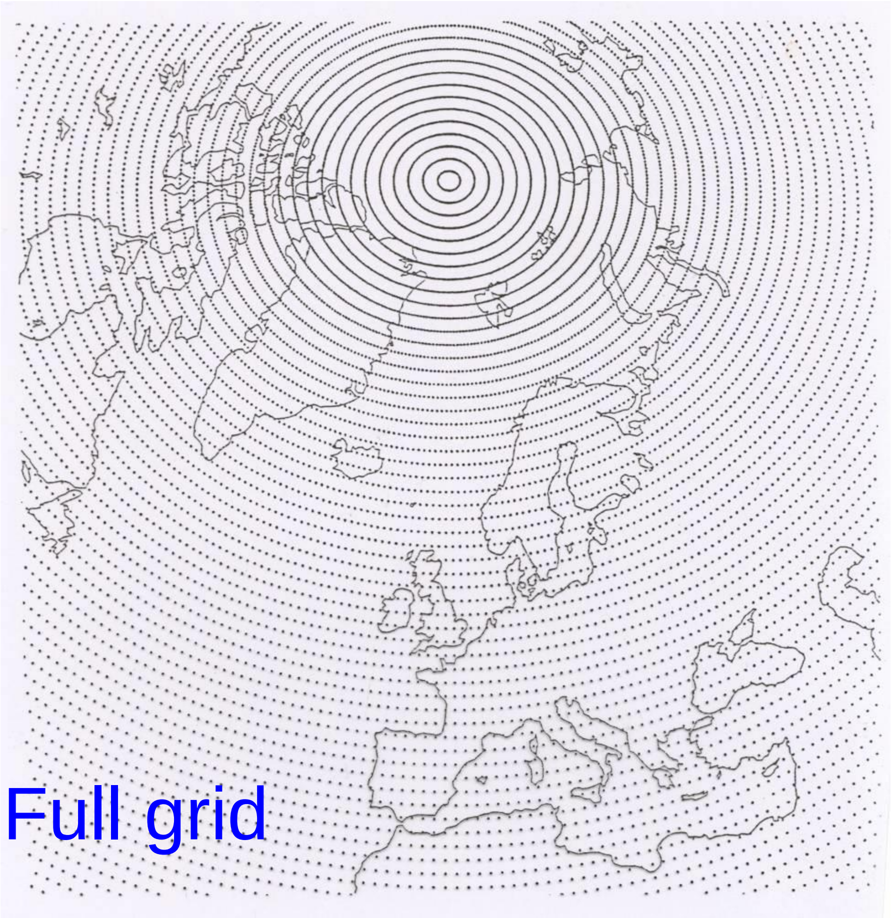
A fast Legendre transform (FLT)

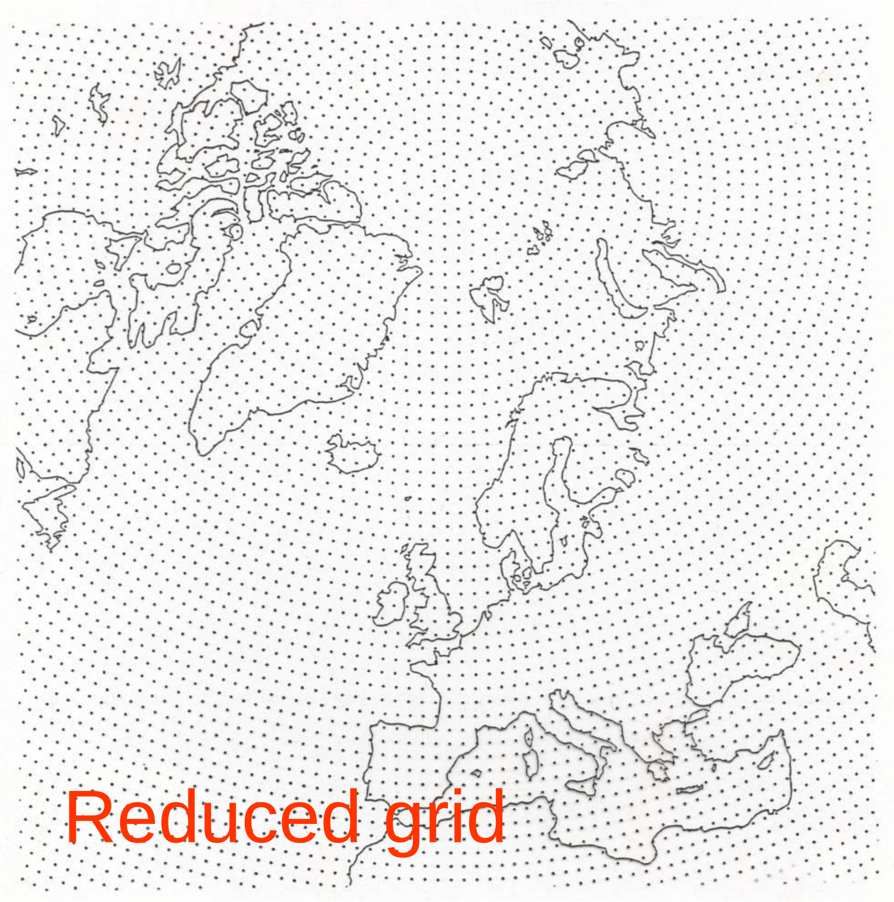
(O'Neil, Woolfe, Rokhlin, 2009; Tygert 2008, 2010)

- **The computational complexity of the ordinary spectral transform is O(N^3) (where N is the truncation number of the series expansion in spherical harmonics) and it was therefore believed to be *not computationally competitive with other methods at very high resolution***
- **The FLT is found to be O(N^2 log N^3) for horizontal resolutions up to T7999** *(Wedi et al, 2013)*

The Gaussian grid

About 30% reduction in number of points

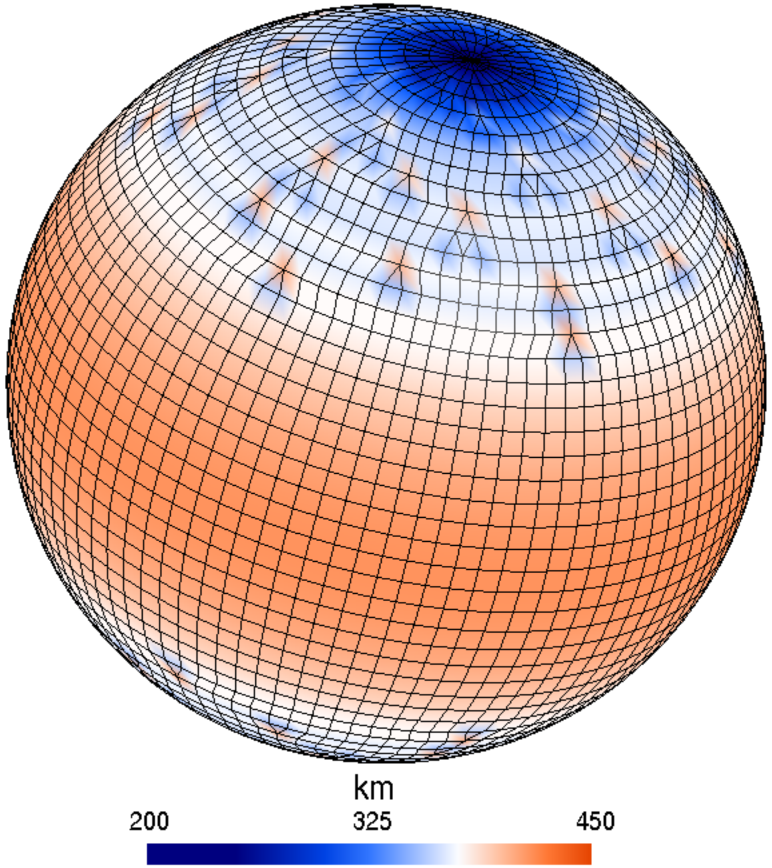




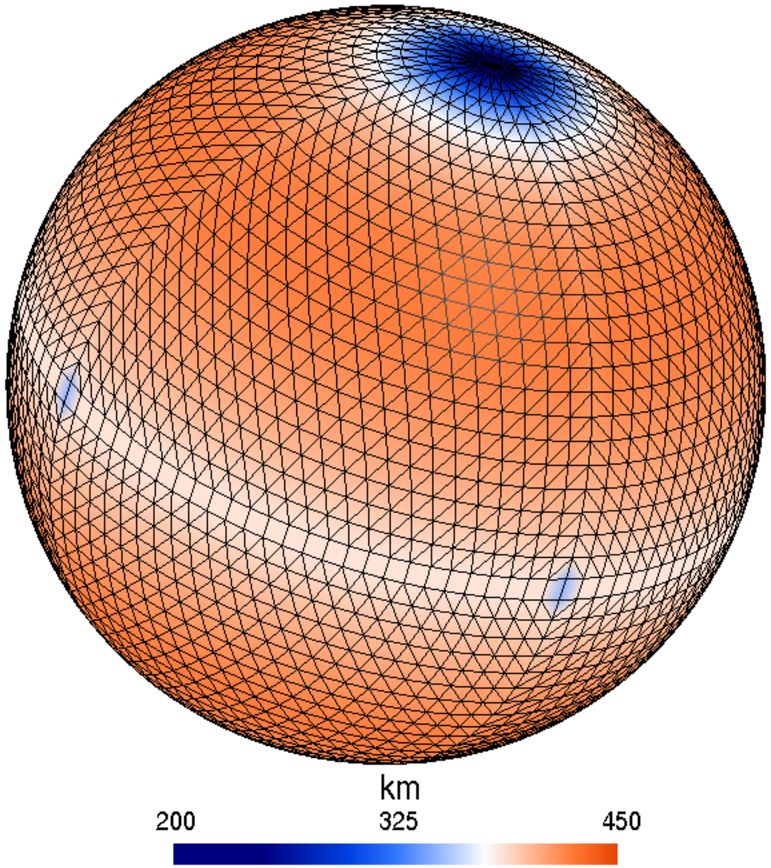
Hortal, M. and Simmons, A. J. (1991)

A new grid for ECMWF

A further ~20% reduction in gridpoints
=> ~50% less points compared to full grid



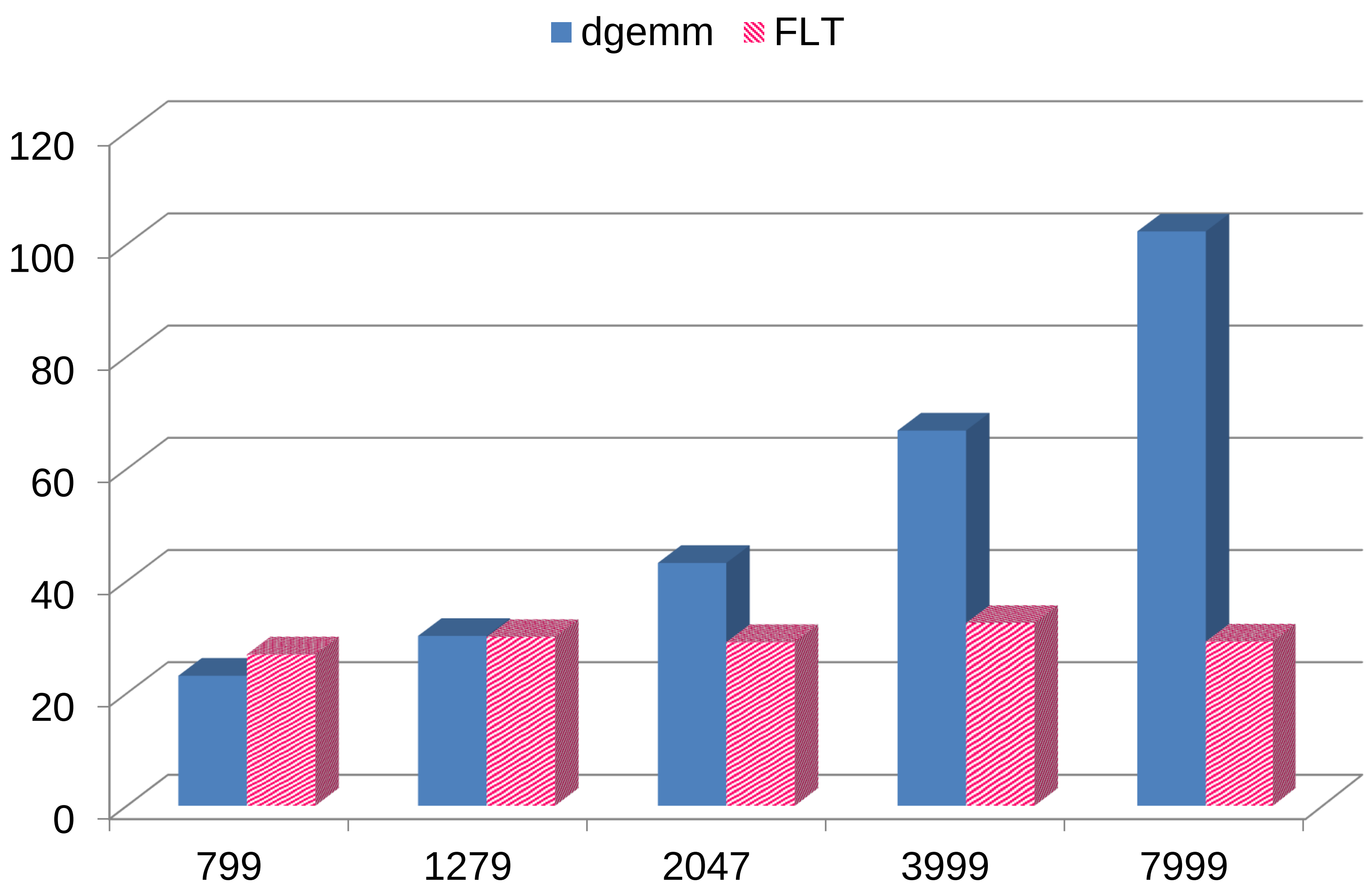
N24 reduced Gaussian grid



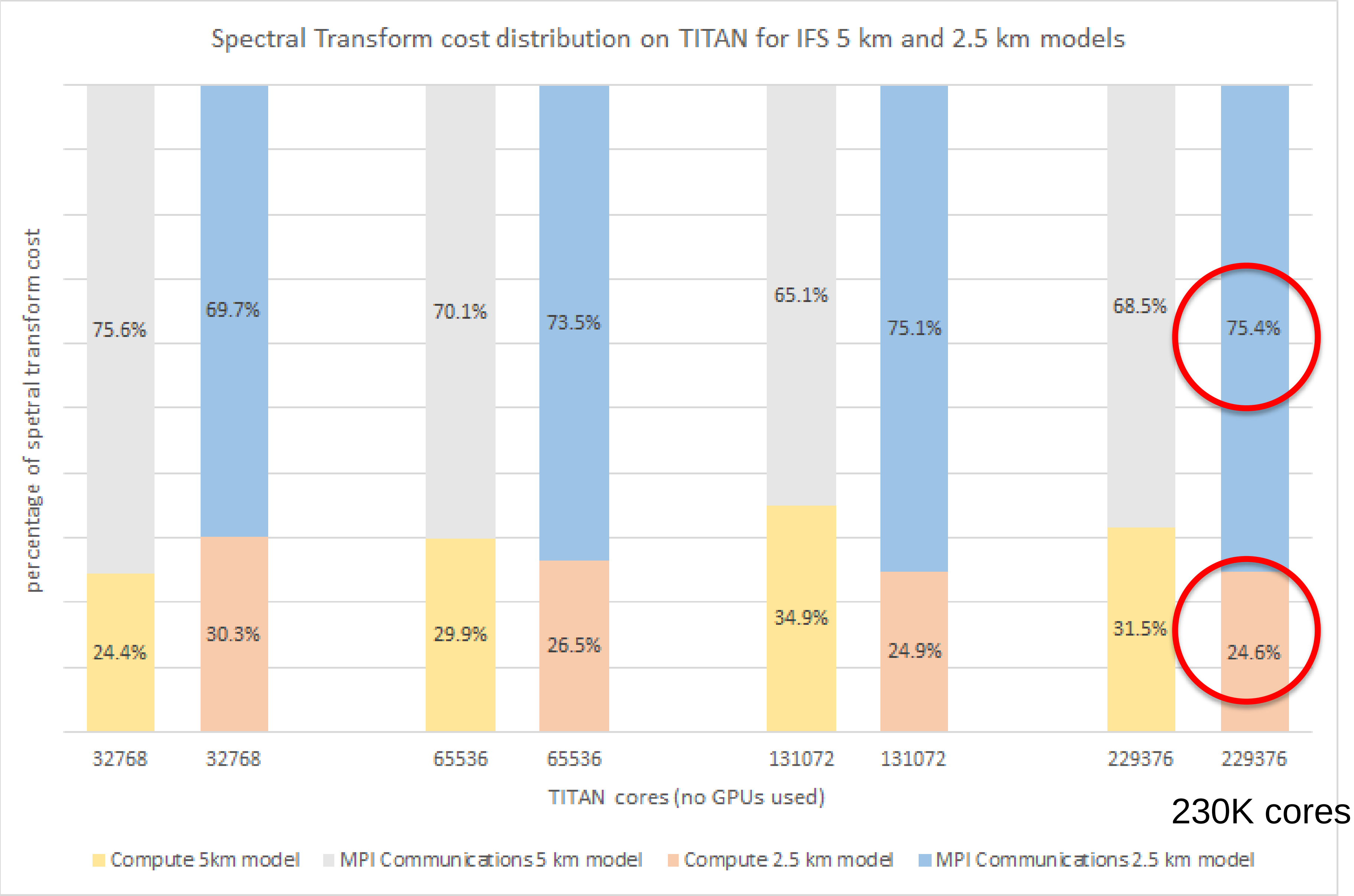
N24 octahedral Gaussian grid

(Wedi et al, 2015)

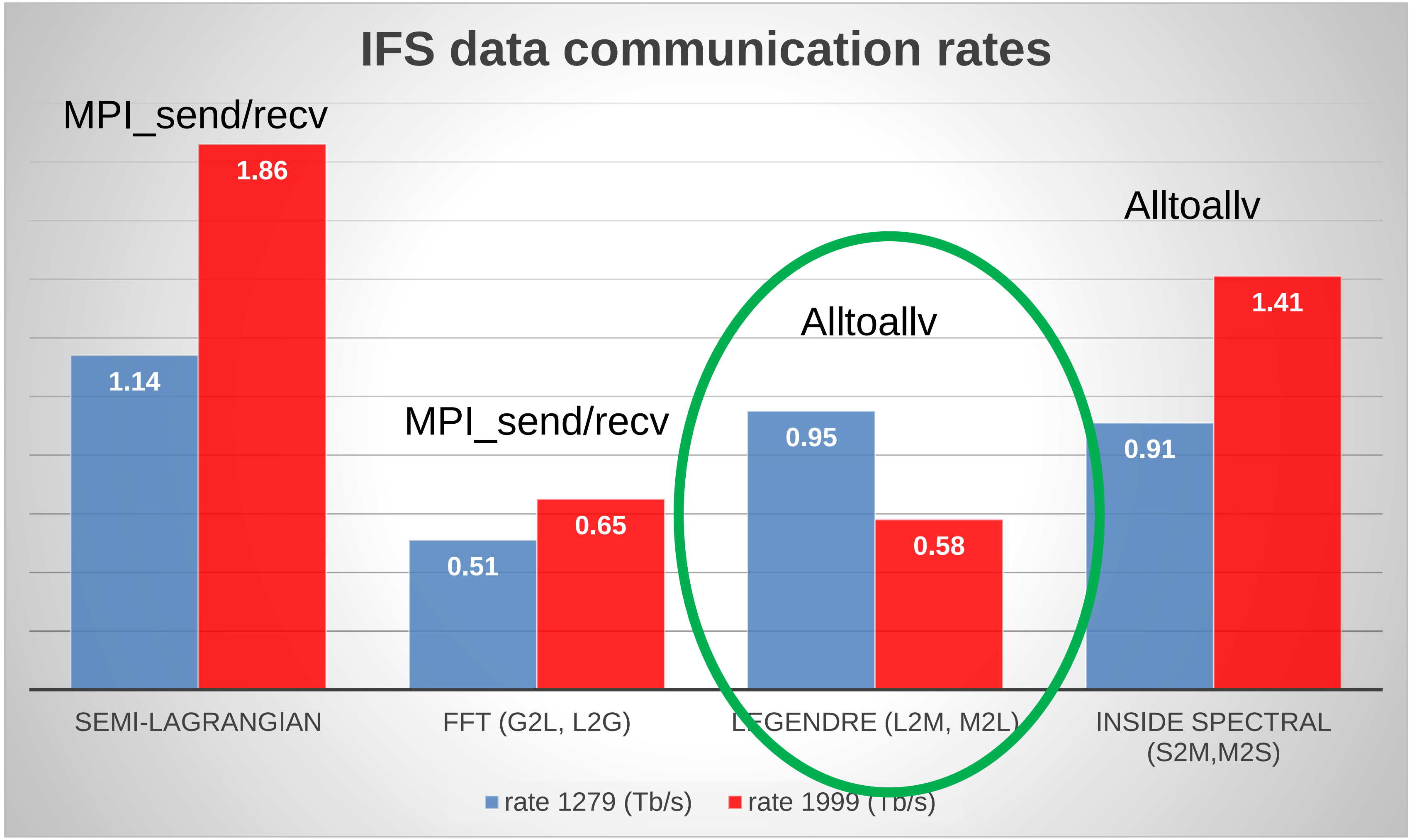
Number of floating point operations for direct or inverse spectral transforms of a single field, scaled by $N^2 log^3 N$



MPI communication cost at large core counts …

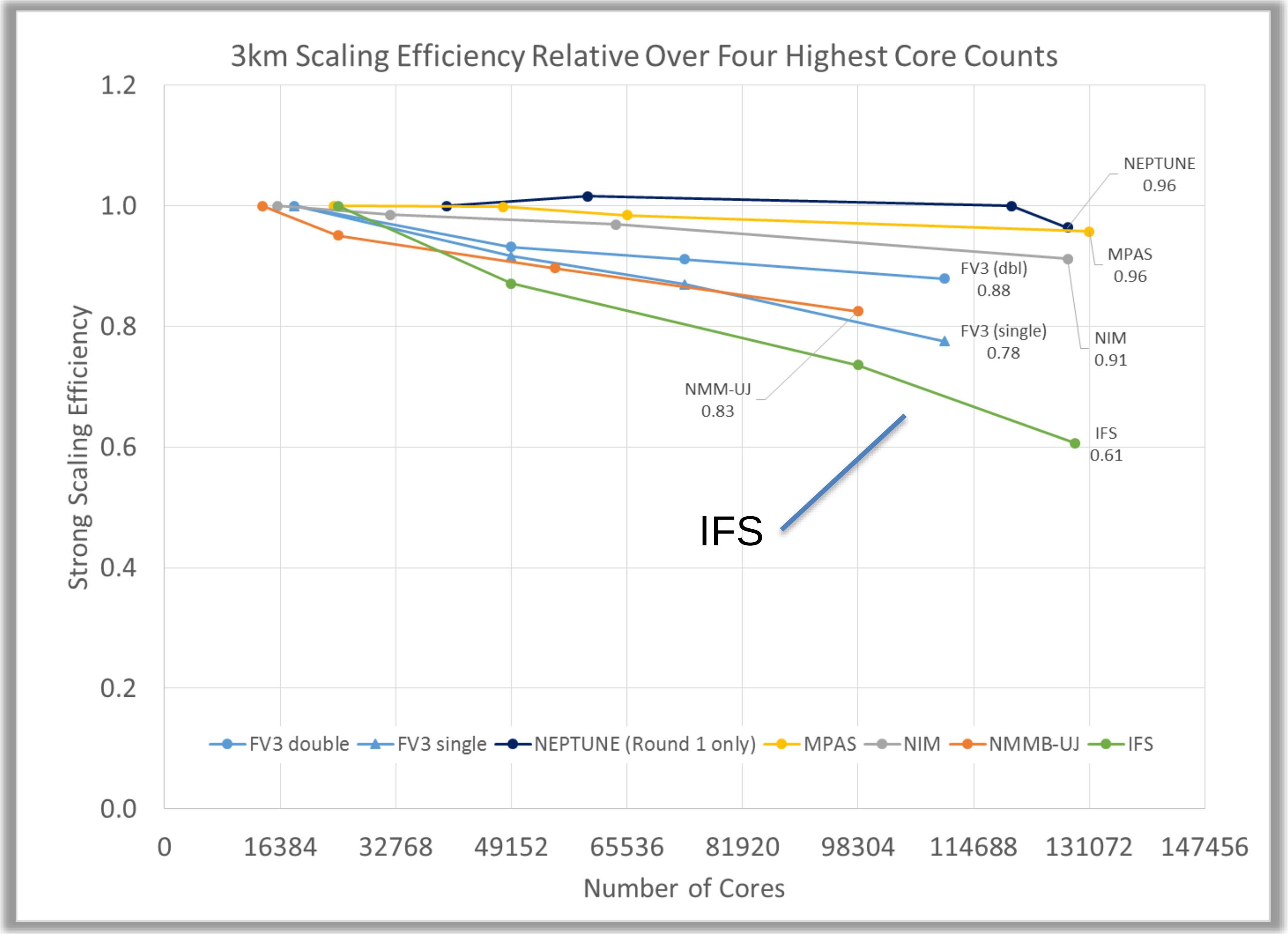


Application performance XC30



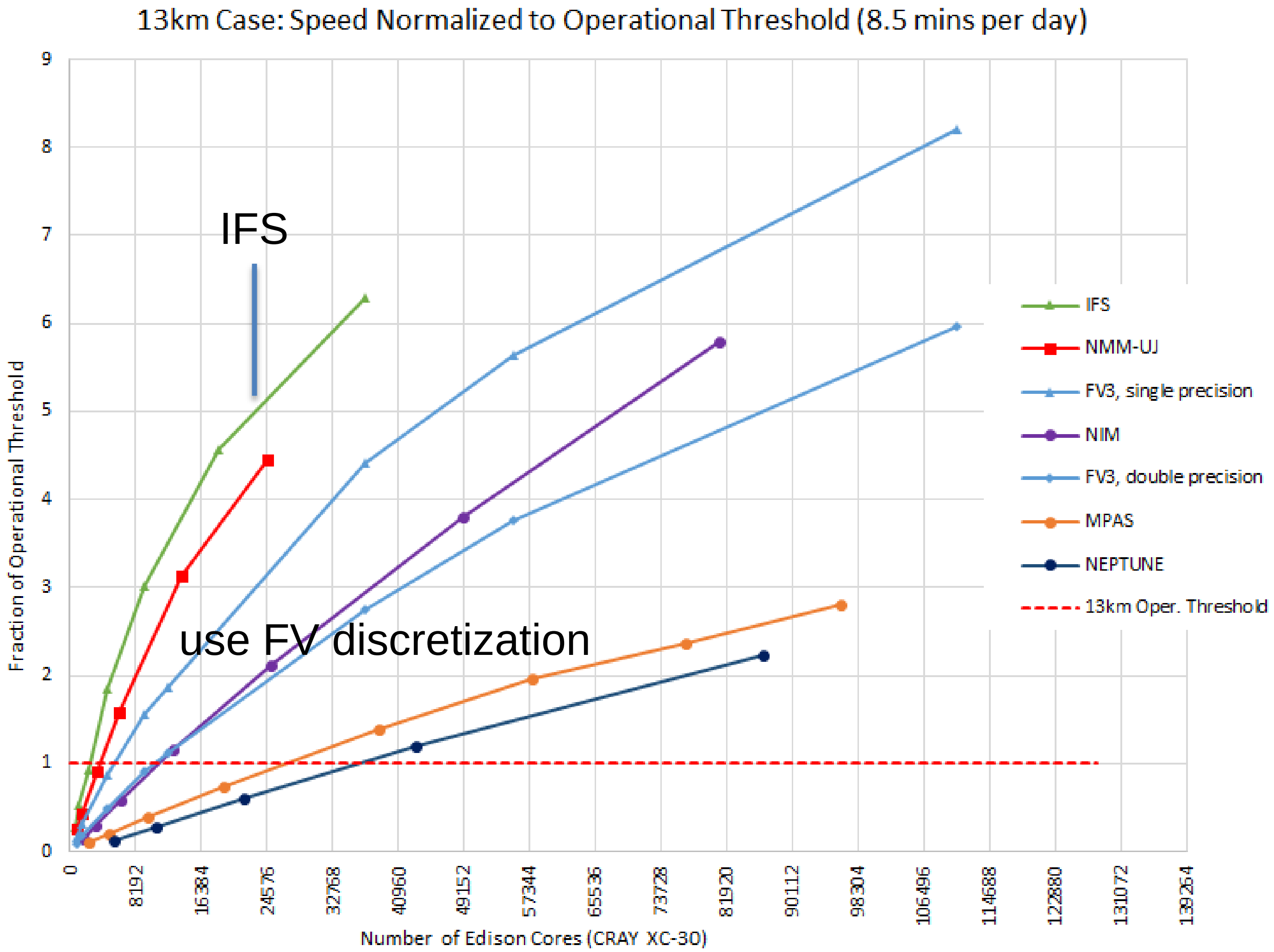
TCo1279 on 360 nodes; TCo1999 on 720 nodes ; dt=450s; 48h forecast

Scaling efficiency at 3km



(Michalakes et al, NGGPS AVEC report, 2015)

Time-to-solution at 13km

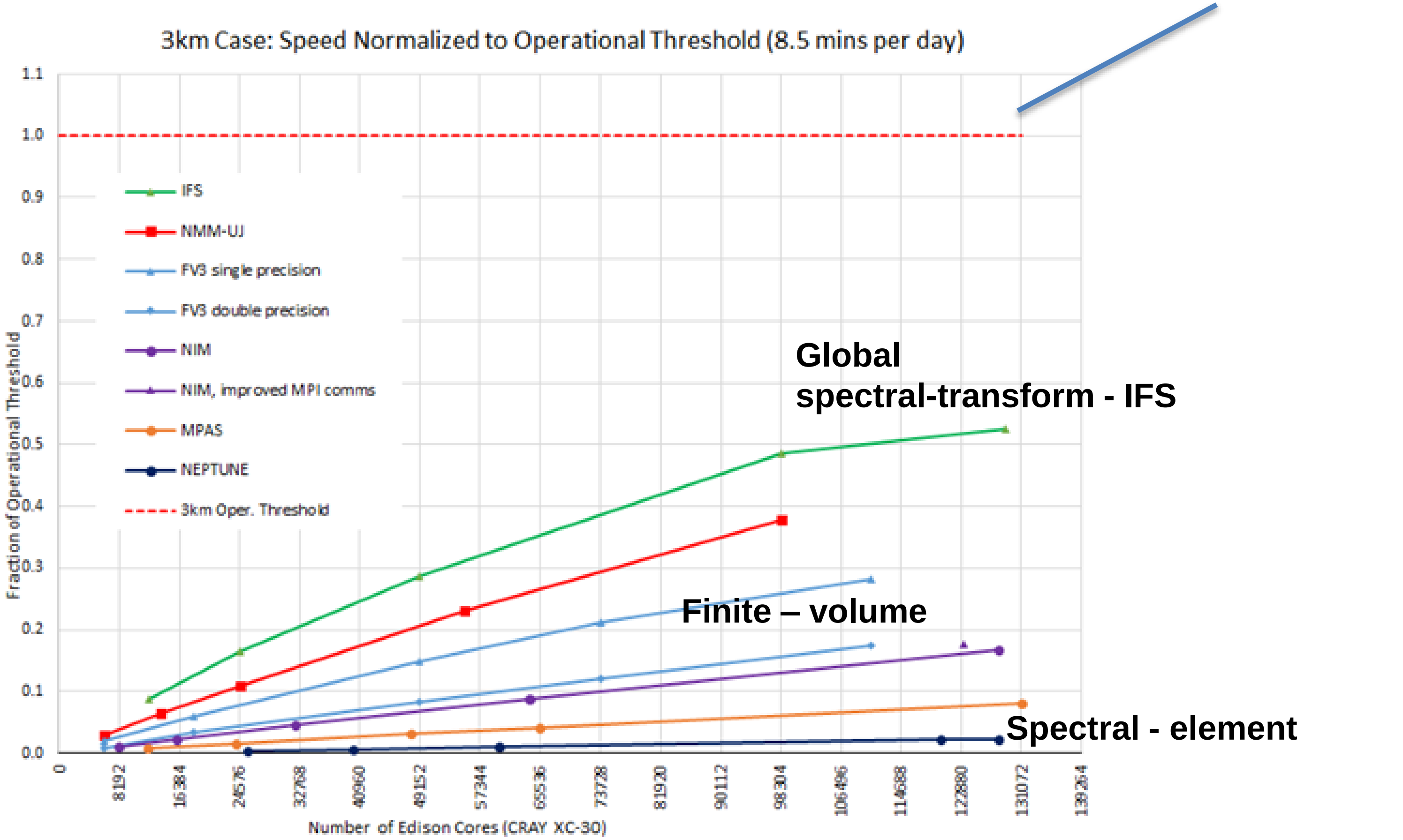


(Michalakes et al, NGGPS AVEC report, 2015)

Comparison to alternative methods

Time-to-solution as a function of CPU cores at ~ 3km

Operational need!

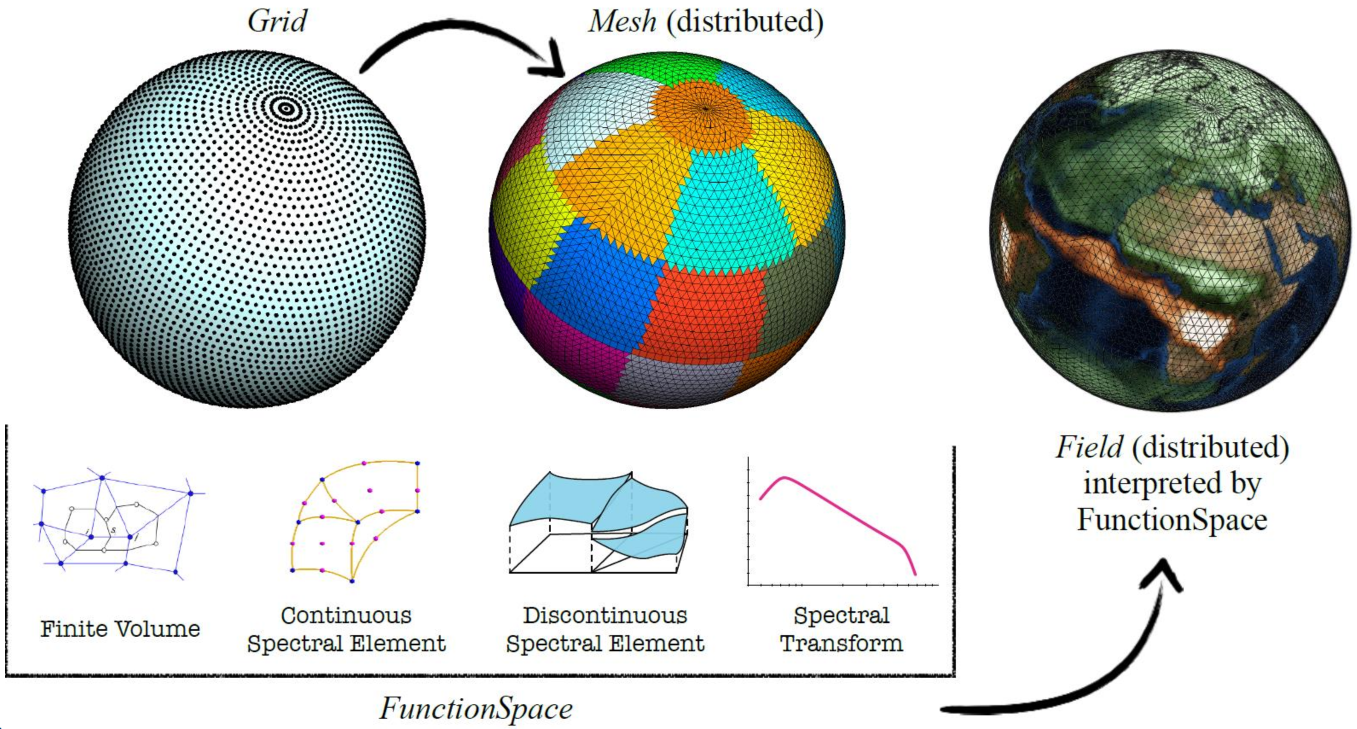


(adapted from Michalakes et al, NGGPS AVEC report, 2015)

Algorithmic flexibility

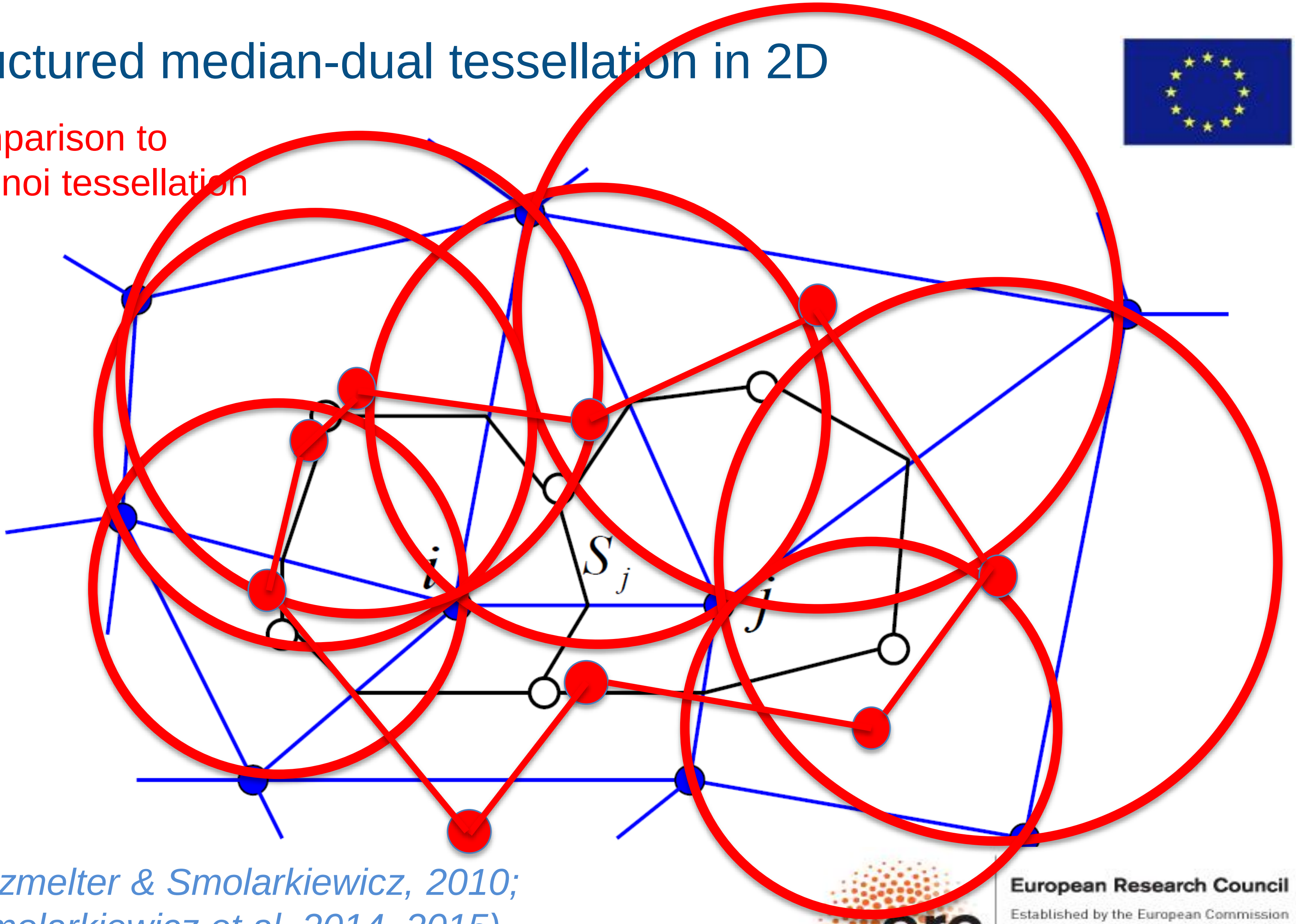
- Global (pseudo-)spectral transform
- Local low-order (<=2) methods (Finite-volume)
- Local higher-order (>2) methods (spectral element, DG, flux reconstruction)
  - Can these be efficiently maintained in the same numerical modelling framework ?
  - A grid-choice that facilitates hybrid use of several of above simultaneously ?
  - Storage layer abstraction for emerging HPC architectures ?

⇒ *Atlas* framework
(Deconinck et al, 2016)



An unstructured median-dual tessellation in 2D

Comparison to
Voronoi tessellation

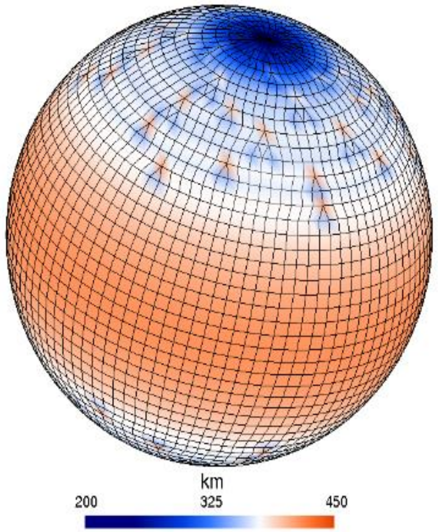


(Szmelter & Smolarkiewicz, 2010; Smolarkiewicz et al, 2014, 2015)

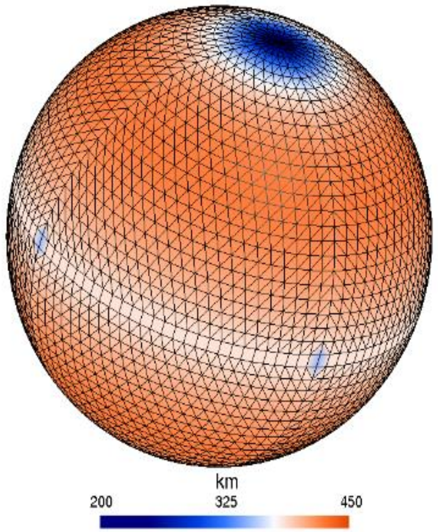
European Research Council
Established by the European Commission
Supporting top researchers from anywhere in the world

Algorithmic flexibility

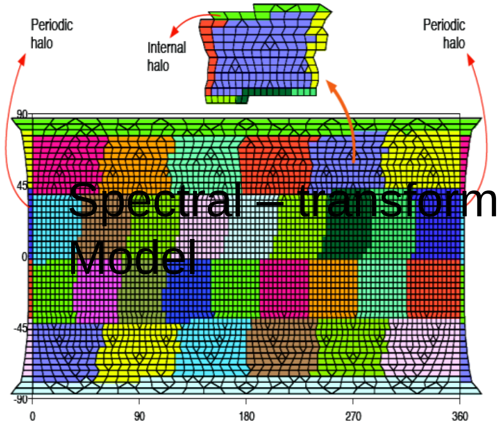
A new grid for ECMWF



N24 reduced Gaussian grid



N24 octahedral Gaussian grid

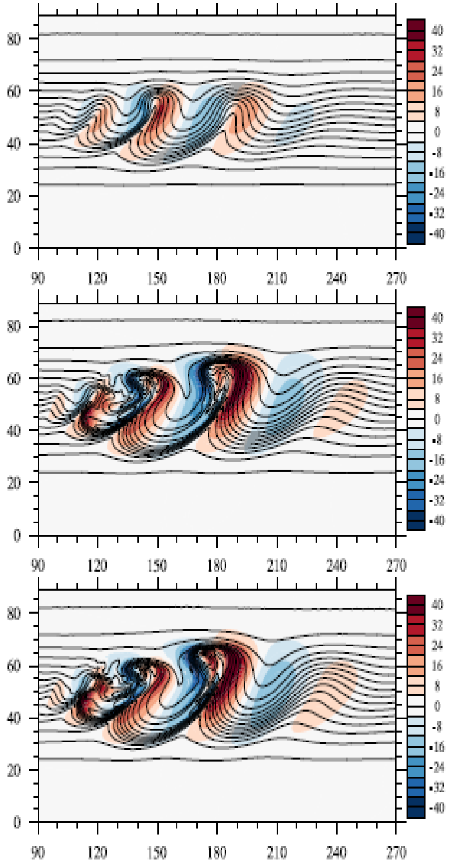


Spectral – transform Model

Local communication

IFS-FVM finite-volume module **on the same grid**!

Baroclinic instability benchmark with the FVM (anelastic, pseudo-incompressible, compressible)



Equal area (MPI) parallel decomposition (1600 tasks)

6,599,680 points x 137 levels x 10 variables at ~9km just above
9 billion points

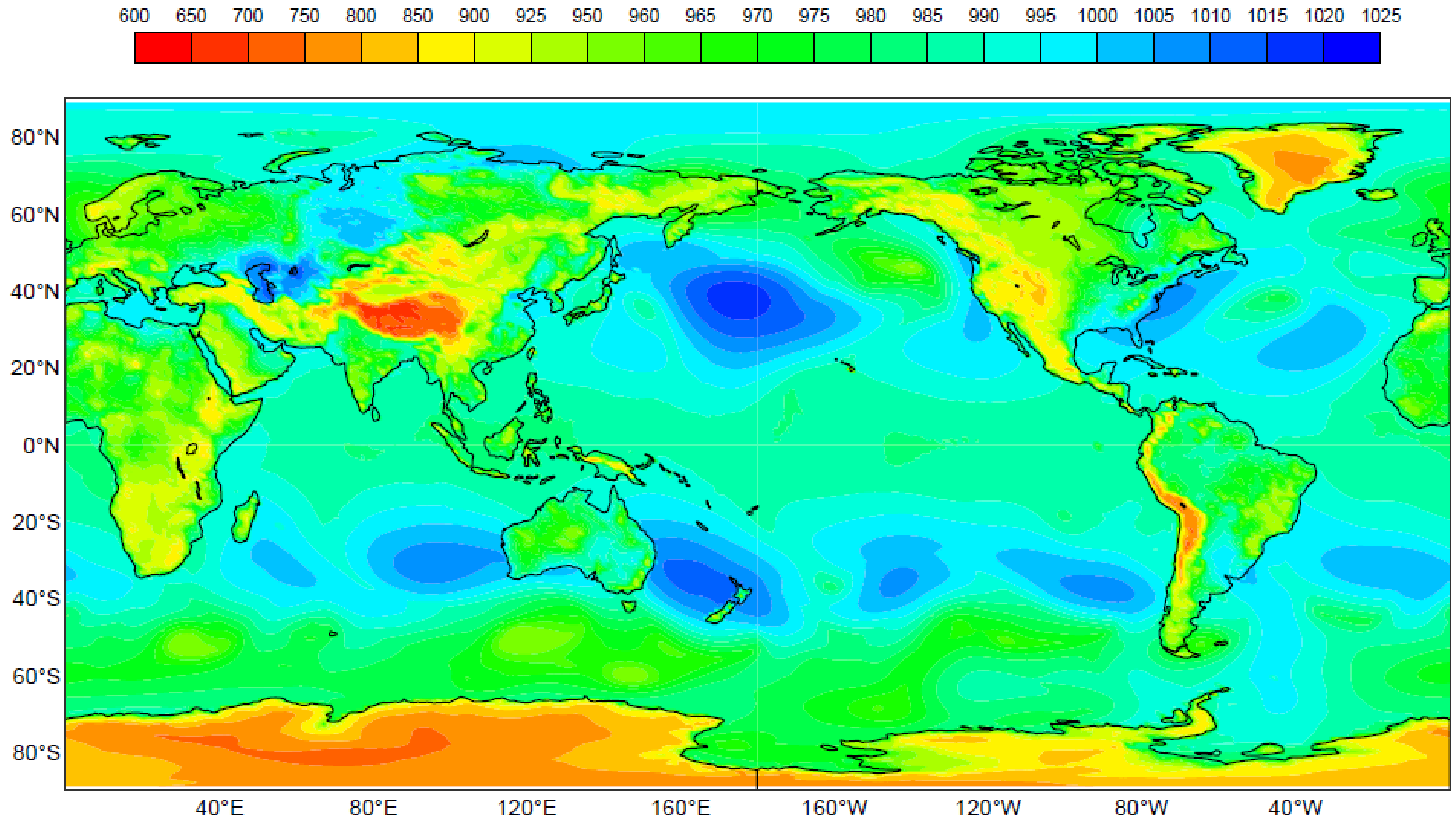
256,800,000 points x 137 levels x 10 variables at ~1,3km just below
352 billion points
FLAGSHIP SIMULATION

(Szmelter & Smolarkiewicz, 2010; Smolarkiewicz et al, 2014,2016)

European Research Council
Established by the European Commission
Supporting top researchers from anywhere in the world

ECMWF

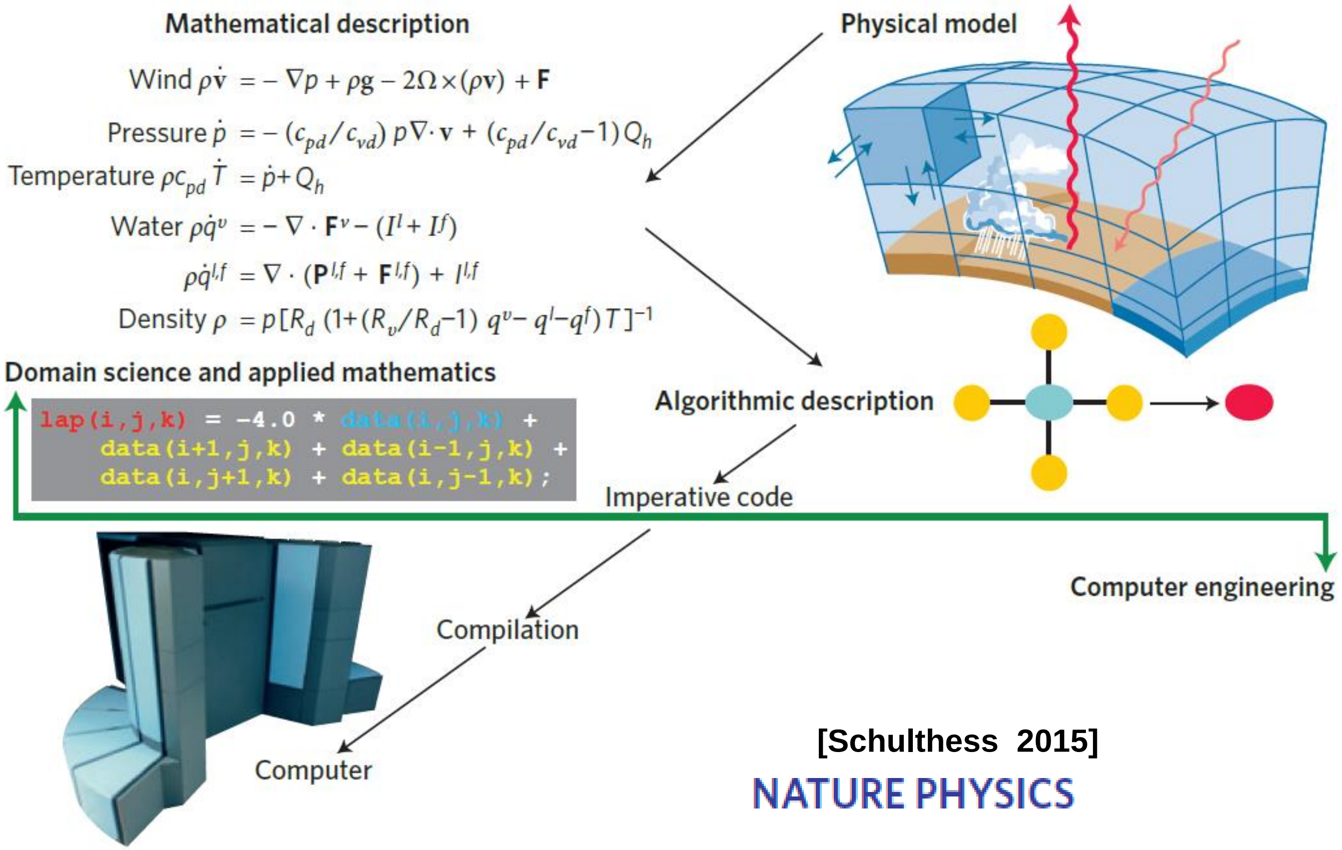
FVM simulation of a global circulation with realistic orography



O640 - Held-Suarez with real orography:
Surface pressure after 50 days of simulation

European Research Council
Established by the European Commission
Supporting top researchers from anywhere in the world

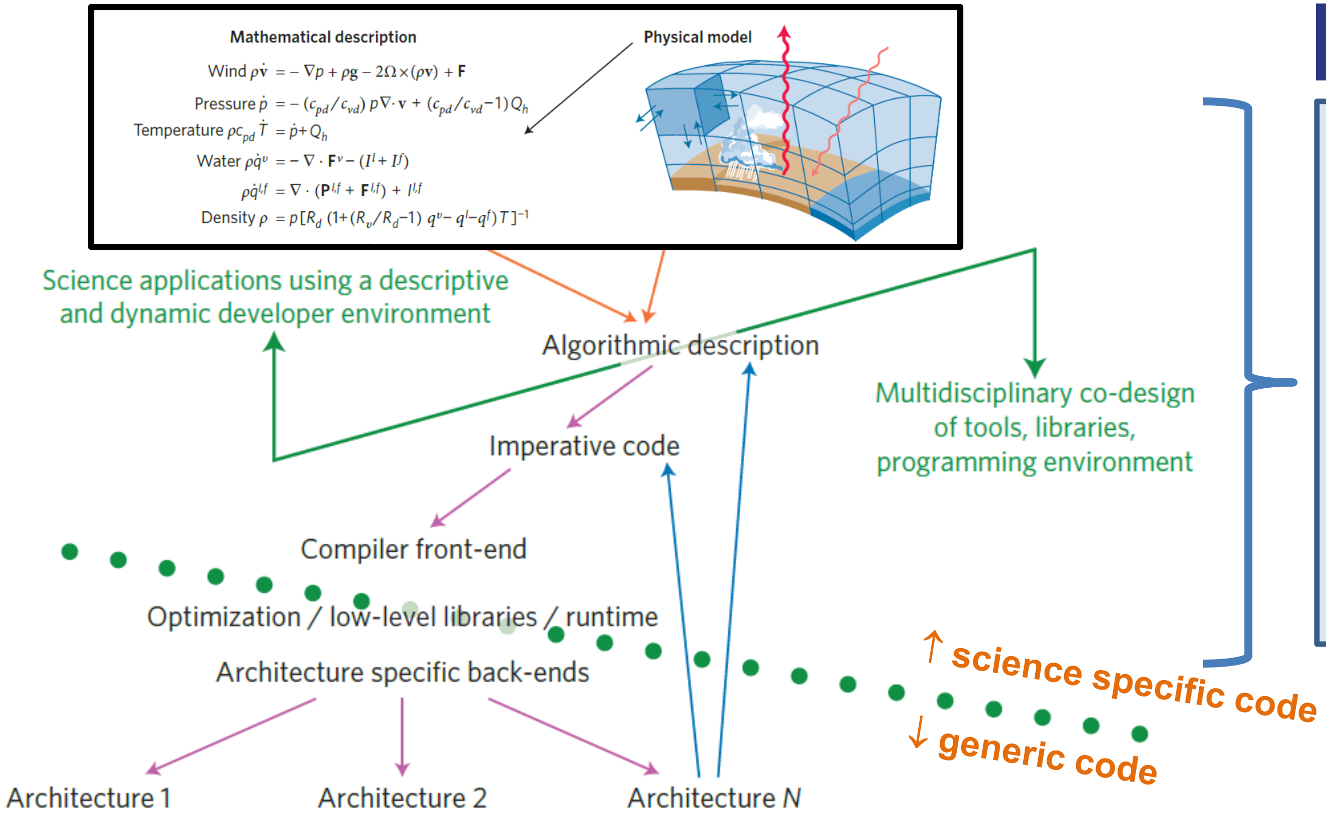
Traditional science workflow



[Schulthess 2015]

NATURE PHYSICS

Future science workflow



[Schulthess 2015]
NATURE PHYSICS

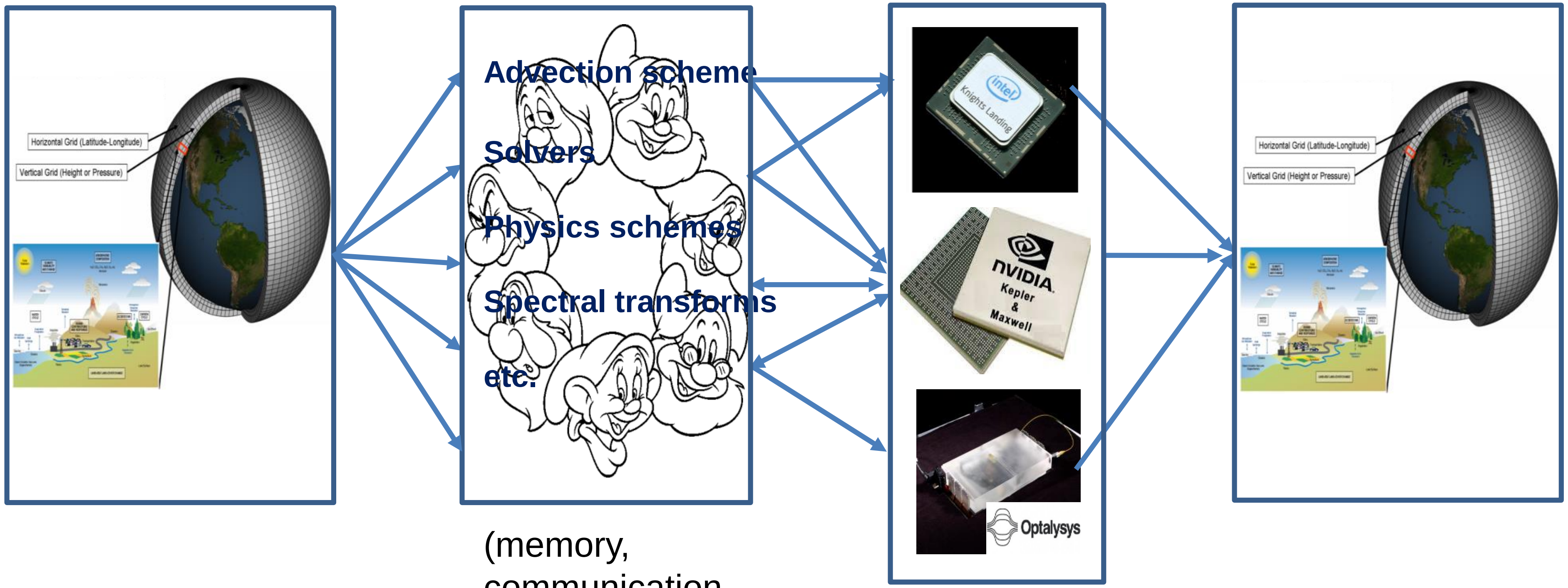
This project is funded by the European Union

<u>E</u>nergy efficient <u>SC</u>alable <u>A</u>lgorithms for weather <u>P</u>rediction at <u>E</u>xascale
→ **www.hpc-escape.eu**

ESCAPE

Energy efficient **SC**alable **A**lgorithms for weather **P**rediction at **E**xascale

Disassemble global … NWP model → Extract, redesign… key components → Optimize for energy… efficiency on new hardware → … Reassemble global NWP model



(memory, communication, compute intensity)

Essential elements of an efficient & reliable Earth-System forecast (beyond an accurate dynamical core)

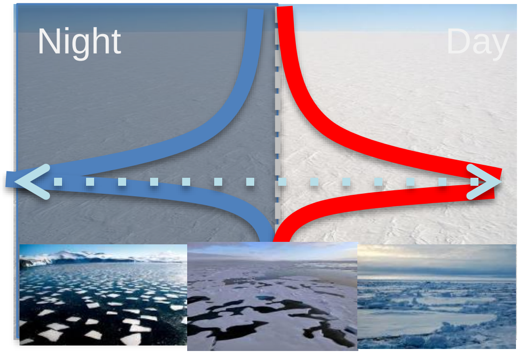
- Use of stochastic perturbations to represent model uncertainty
- Continuously adjust according to parameterized versus permitted convection
- Enable complex and fast interactions of land-surface, radiation, ocean, wave, sea-ice, and chemical processes

Coupled Processes at the surface interface: What are the challenges?

- The processes that are most relevant for near-surface weather prediction are also those that are the most interactive and exhibit positive feedbacks or have a key role in energy partitioning

Over Land

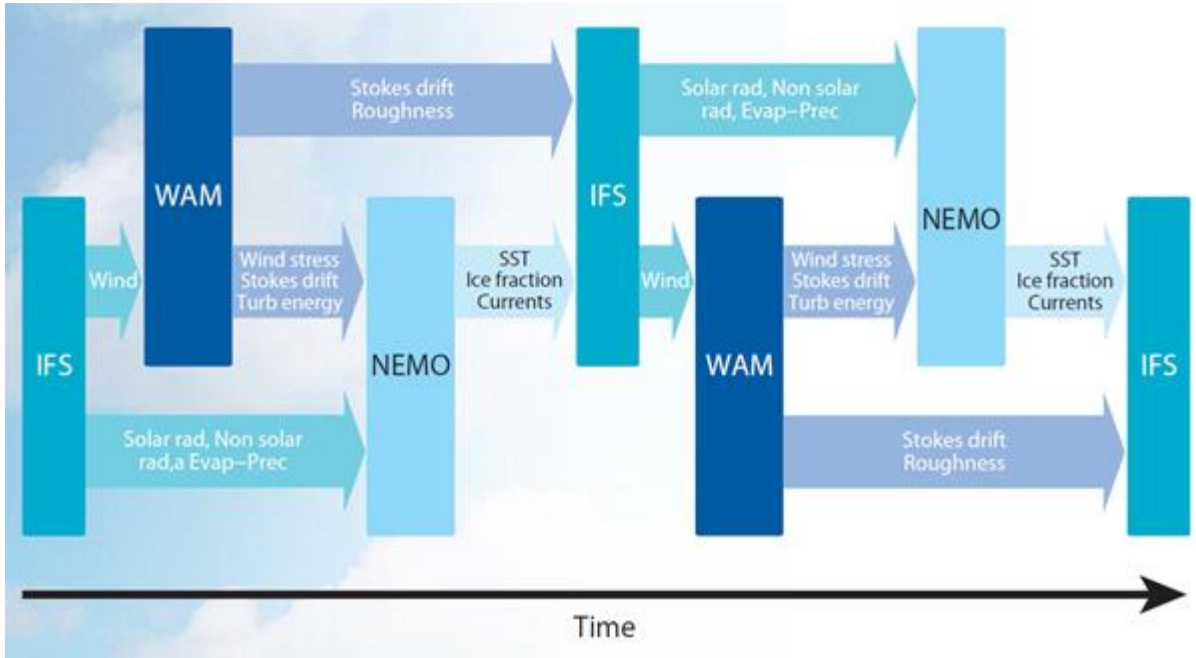
- Snow-cover, ice freezing/melting are in a positive feedback as increased snow/ice cover changes the albedo
- Vegetation growth and variability and interaction with turbulence
- Vertical heat transport in soil/snow



Over Ocean/Cryosphere

- Transition from open-sea to ice-covered conditions
- Sea-state dependent interaction wind induced mixing/waves
- Vertical transport of heat

Example: Ocean-sea-ice-wave-coupling



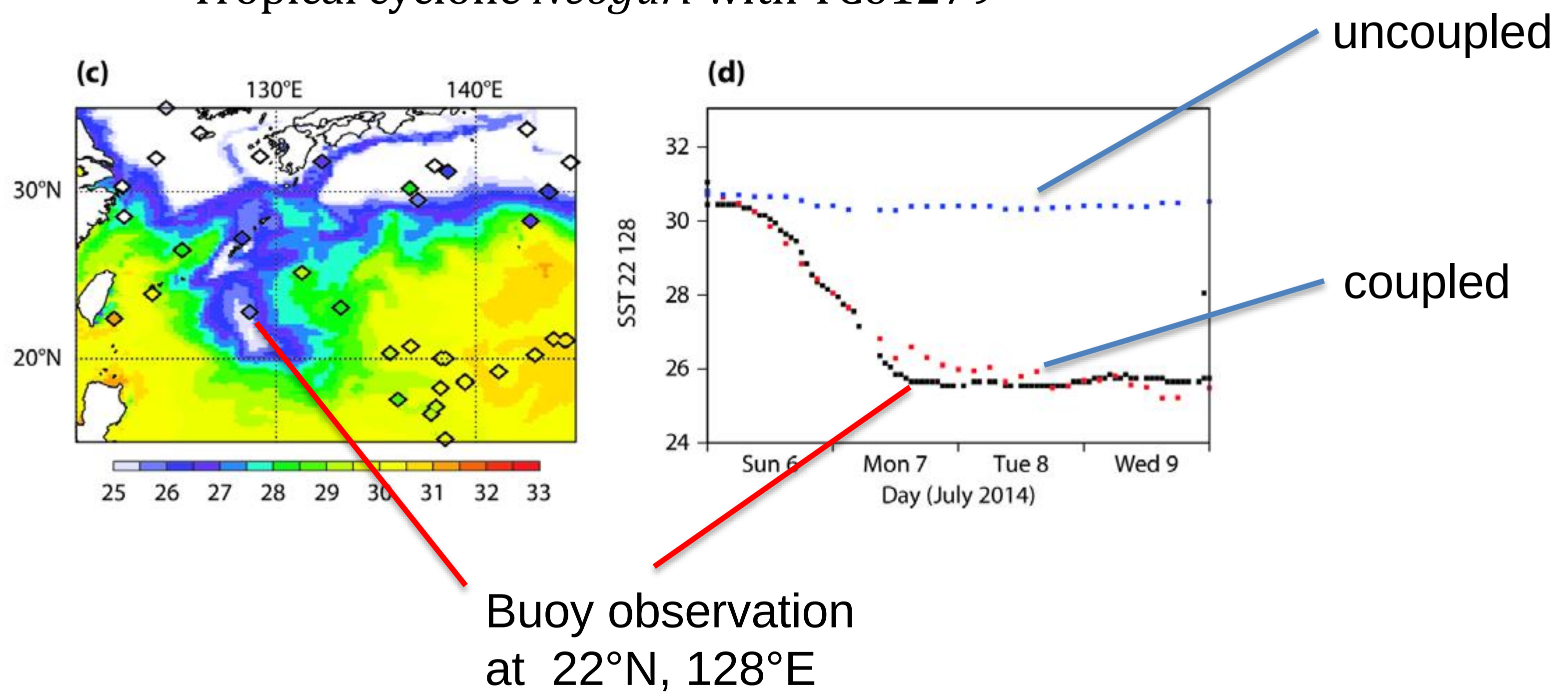
Issues: Initialization

Complexity to balance potential increase in predictive skill

Sequentiality of processes

Coupled ocean vs uncoupled simulation

Tropical cyclone *Neoguri* with TCo1279

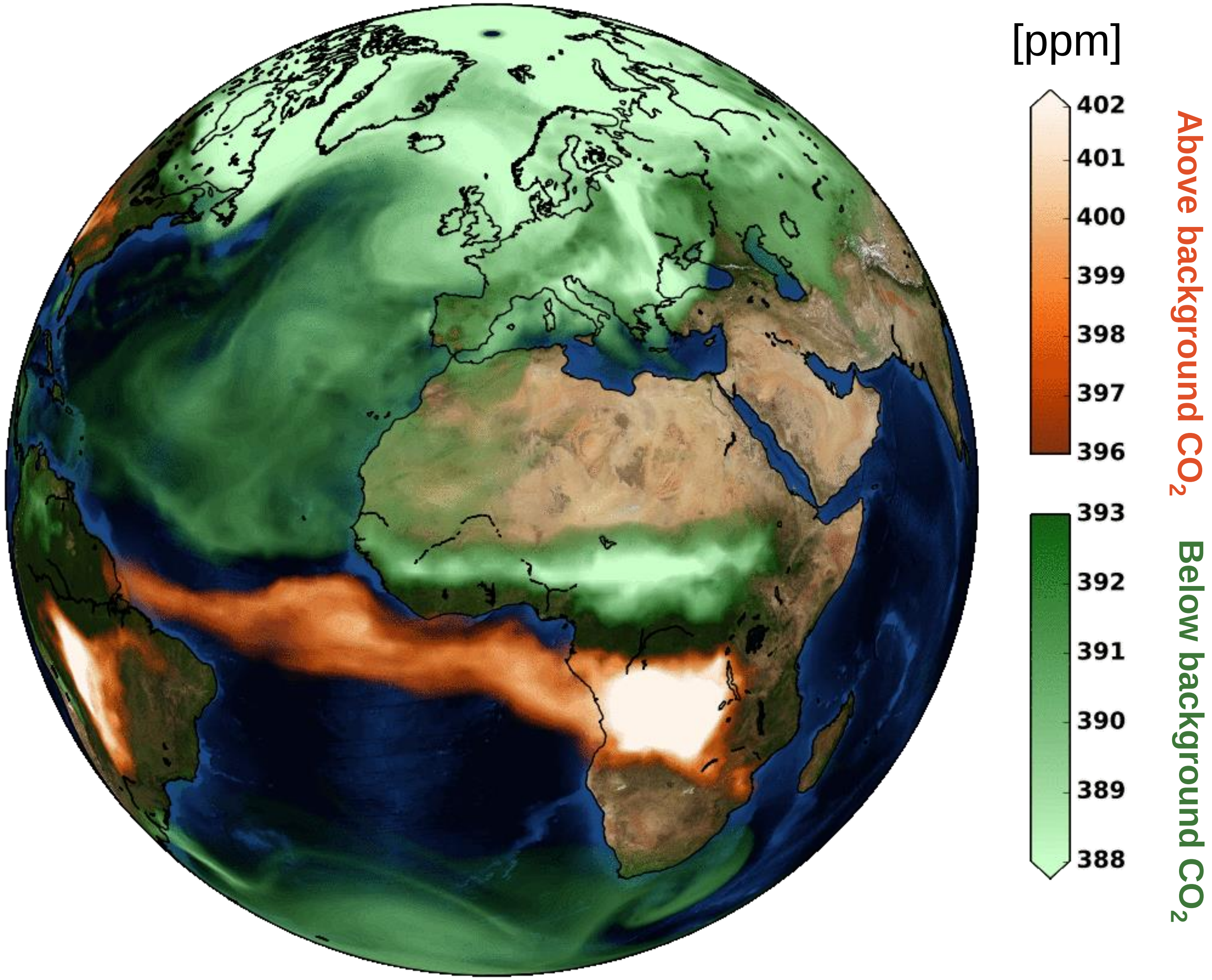


4-day forecast SSTs from the coupled forecast initialised at 0UTC on 6 July 2014 at the location of a buoy with approximate position 22°N, 128°E.

(Rodwell et al, ECMWF Technical Report 759, 2015)

Atmospheric composition: high resolution CO_2 forecast

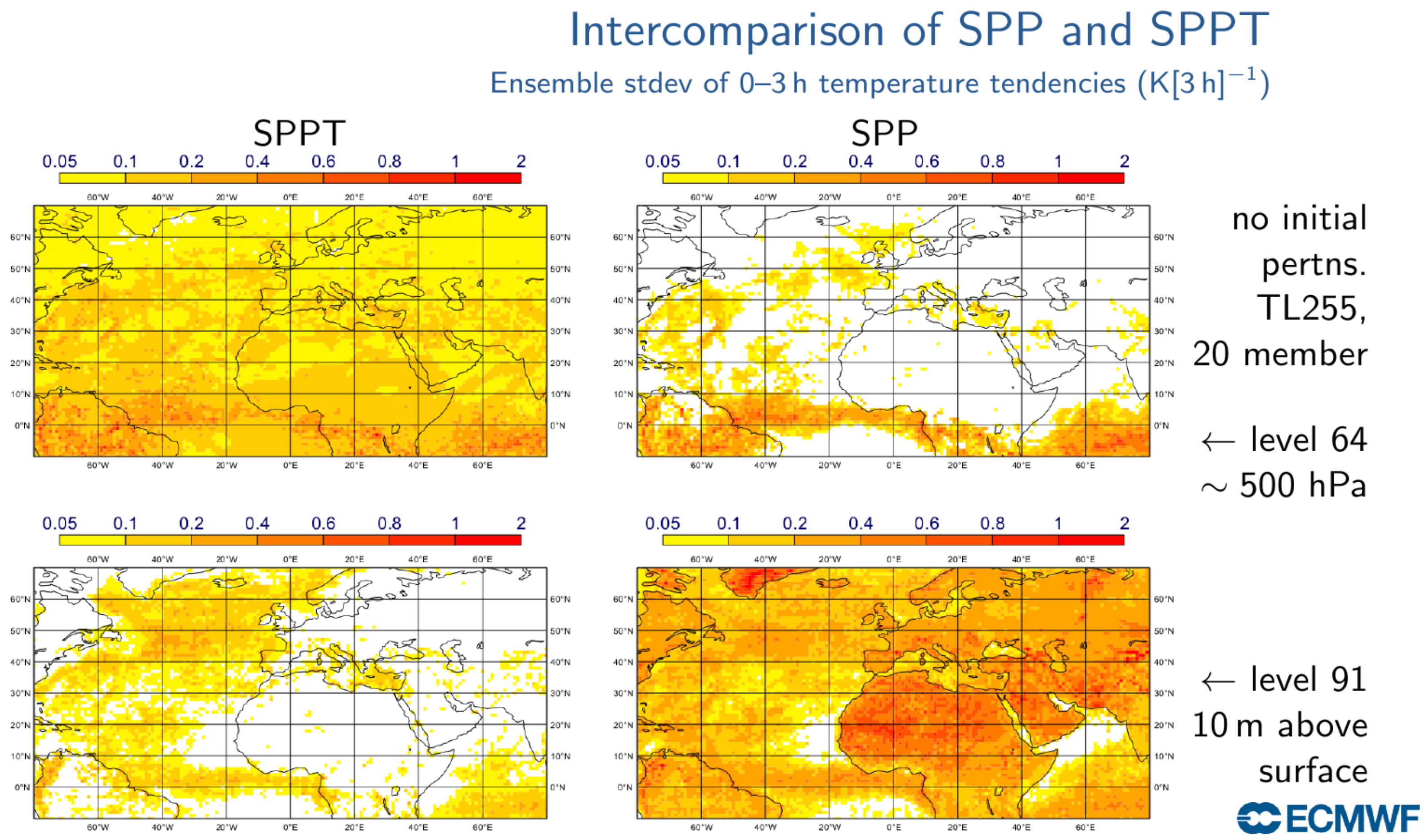
Interacting global atmosphere, land-surface, radiation, ocean and chemical processes



Total column average atmospheric CO_2: September 2013

the work of Anna Agusti-Panareda

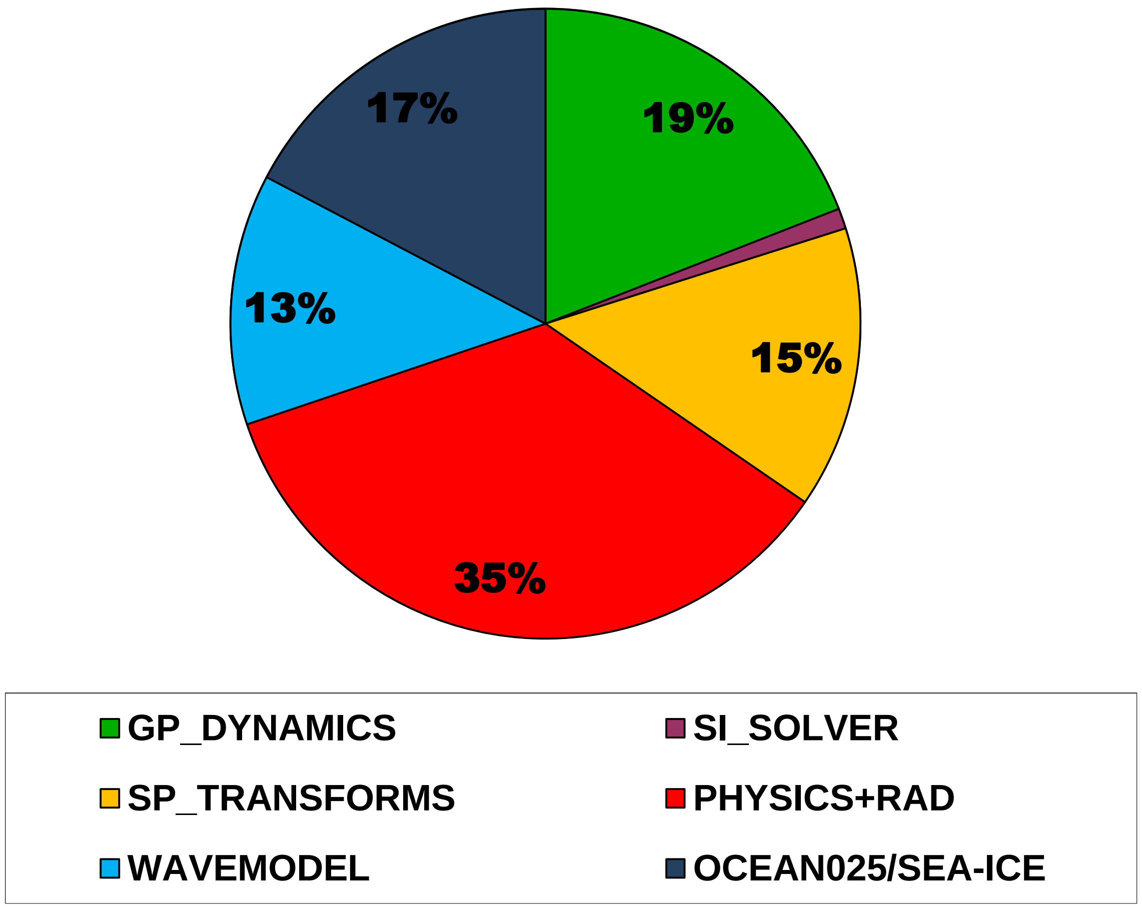
Model uncertainty representation



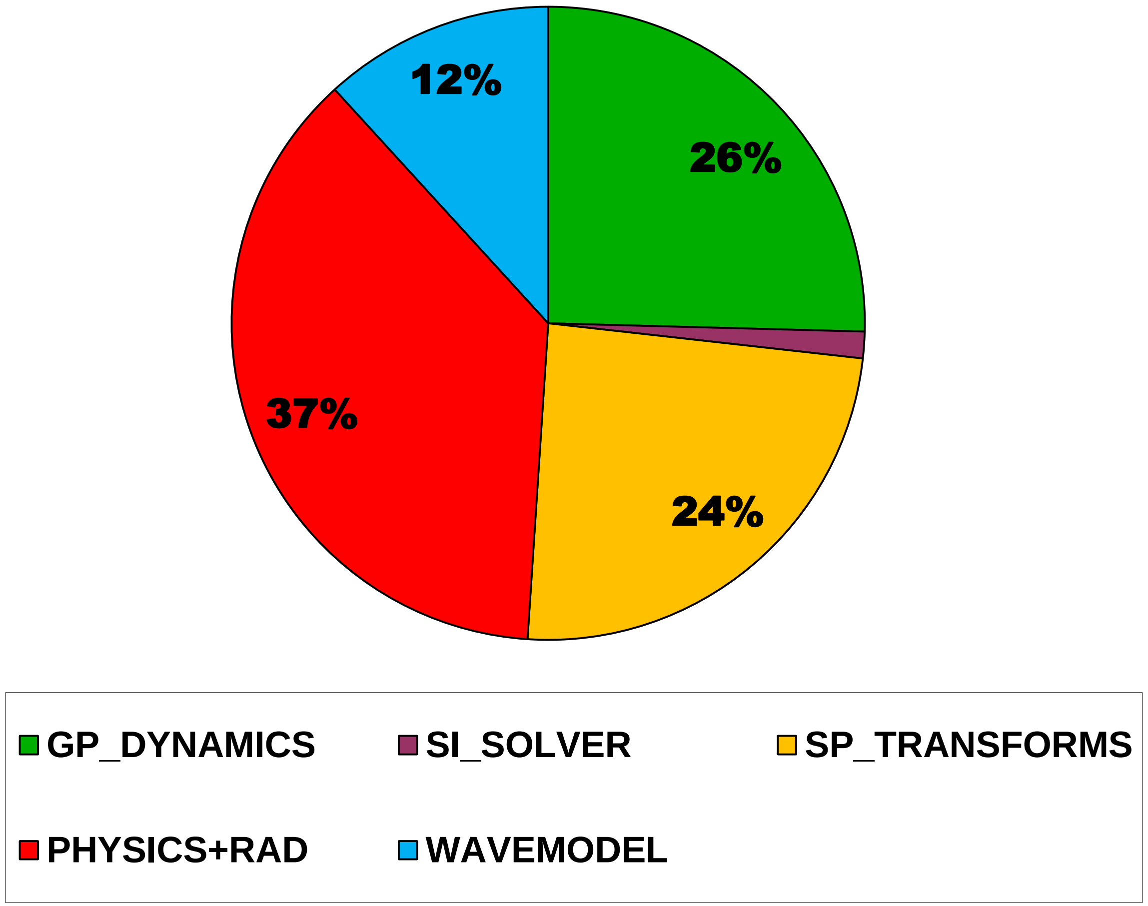
(SAC special topic paper on representation of model uncertainties, Leutbecher et al. 2016)

Cost distribution with increased complexity

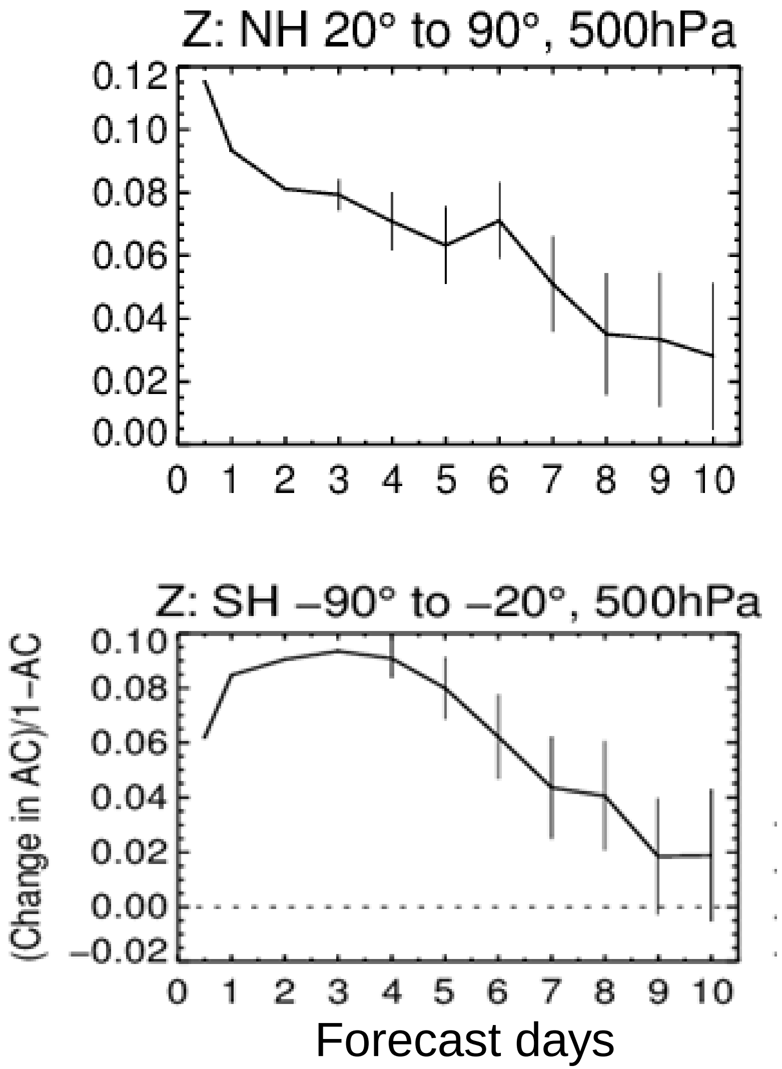
Increased complexity: TCo639 L91 (~18km)
ENS (ORCA025_Z75/LIM2 sea-ice model)



Classic example: TCo1279 L137 (~9km)
HRES

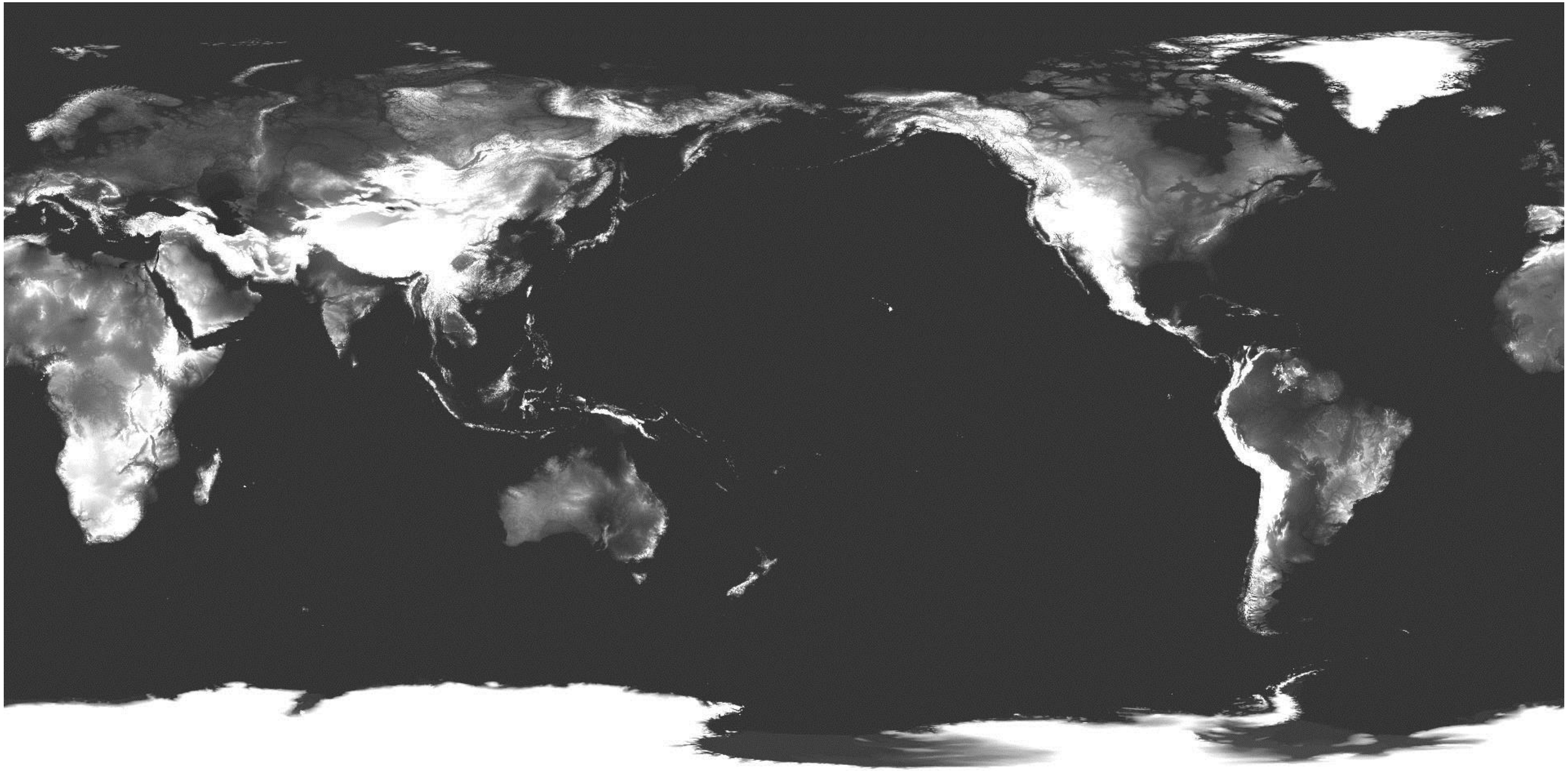


Increasing atmospheric (effective) resolution makes a difference!



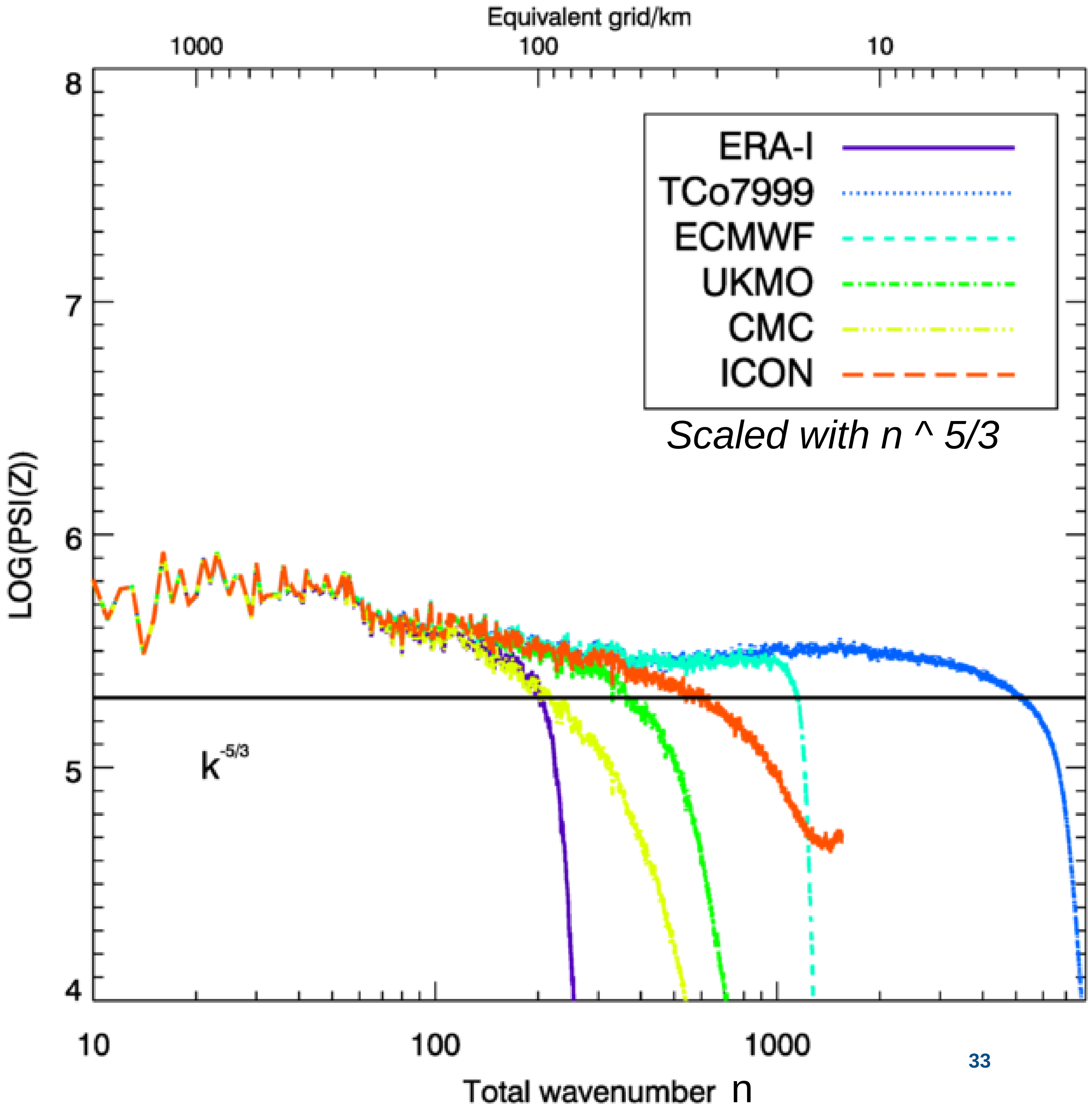
Significant change in anomaly correlation for southern hemisphere and northern hemisphere 500hPa, respectively, for 6 months of winter and summer cases, comparing the new analysis and forecast system TCo1279 (~9km) (TL399/TL319/TL255) cycle 41r2 with current operations at TL1279 (~16km)

TCo7999 (~1.3km) orography



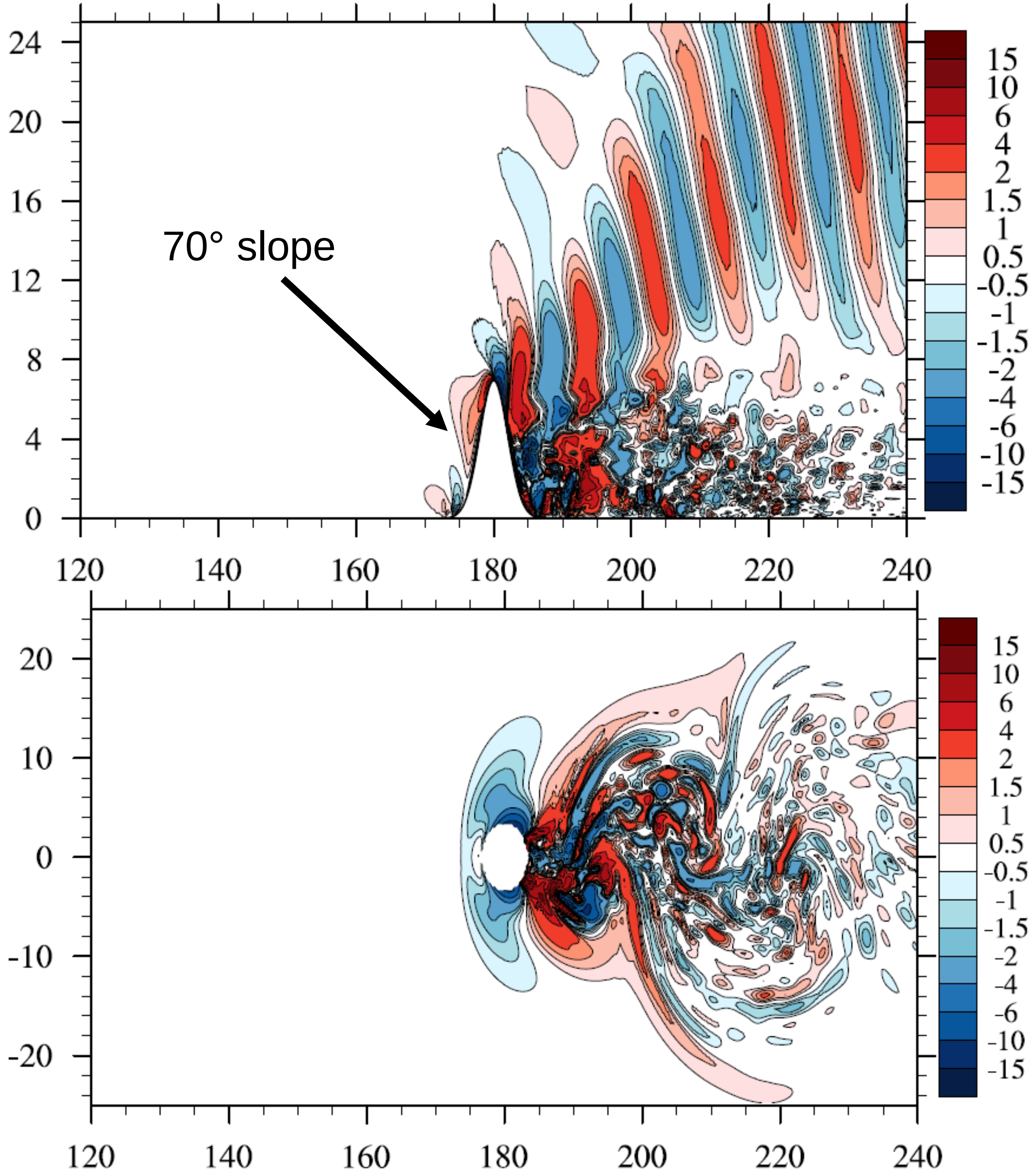
Mean orography spectra at different NWP centres today

Sylvie Malardel, Irina Sandu,
Ayrton Zadra, Simon Vosper,
Annelize van Niekerk, Daniel Klocke



Steep slopes

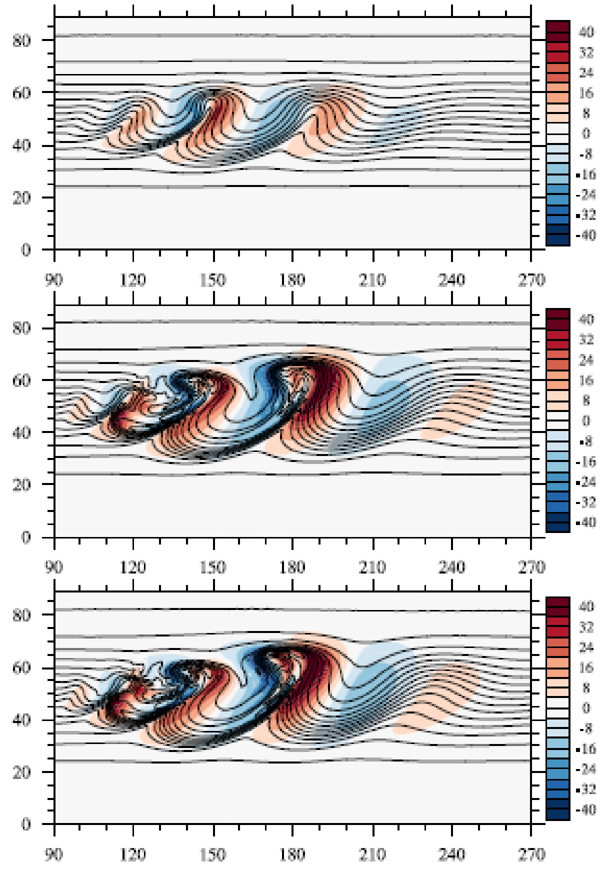
Flexible terrain-following coordinate with optionally static or dynamic adaptivity



Different approaches:
- time-independence in computational space
- height in contrast to current pressure coordinates

Equations beyond the hydrostatic system

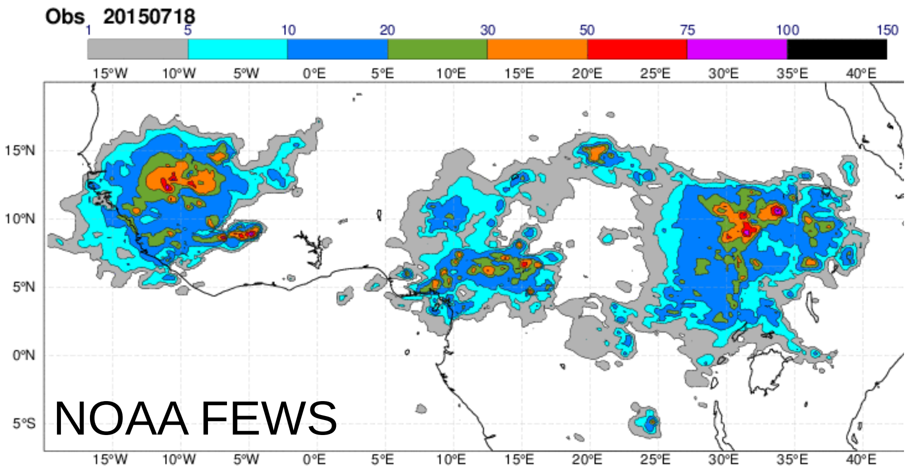
Baroclinic instability benchmark with the FVM (anelastic, pseudo-incompressible, compressible)

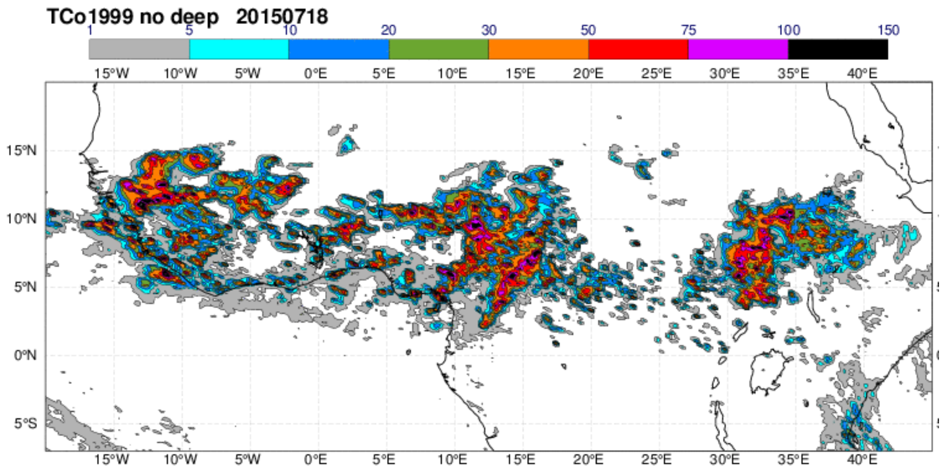


Non-hydrostatic options: anelastic, pseudo-incompressible, Arakawa-Konor, compressible Euler

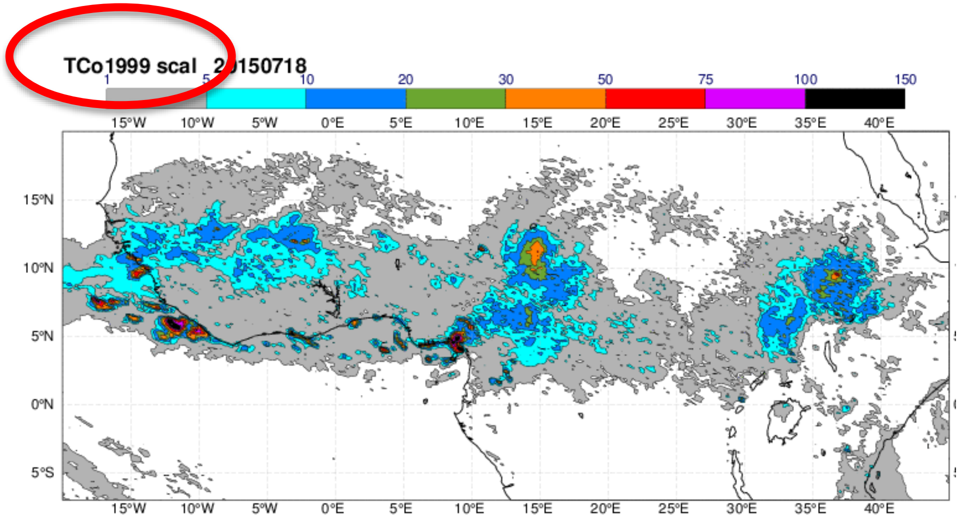
Compressible equations provide the most efficient solution as well as flexibility on the solution procedure in time

Parameterized versus permitted convection



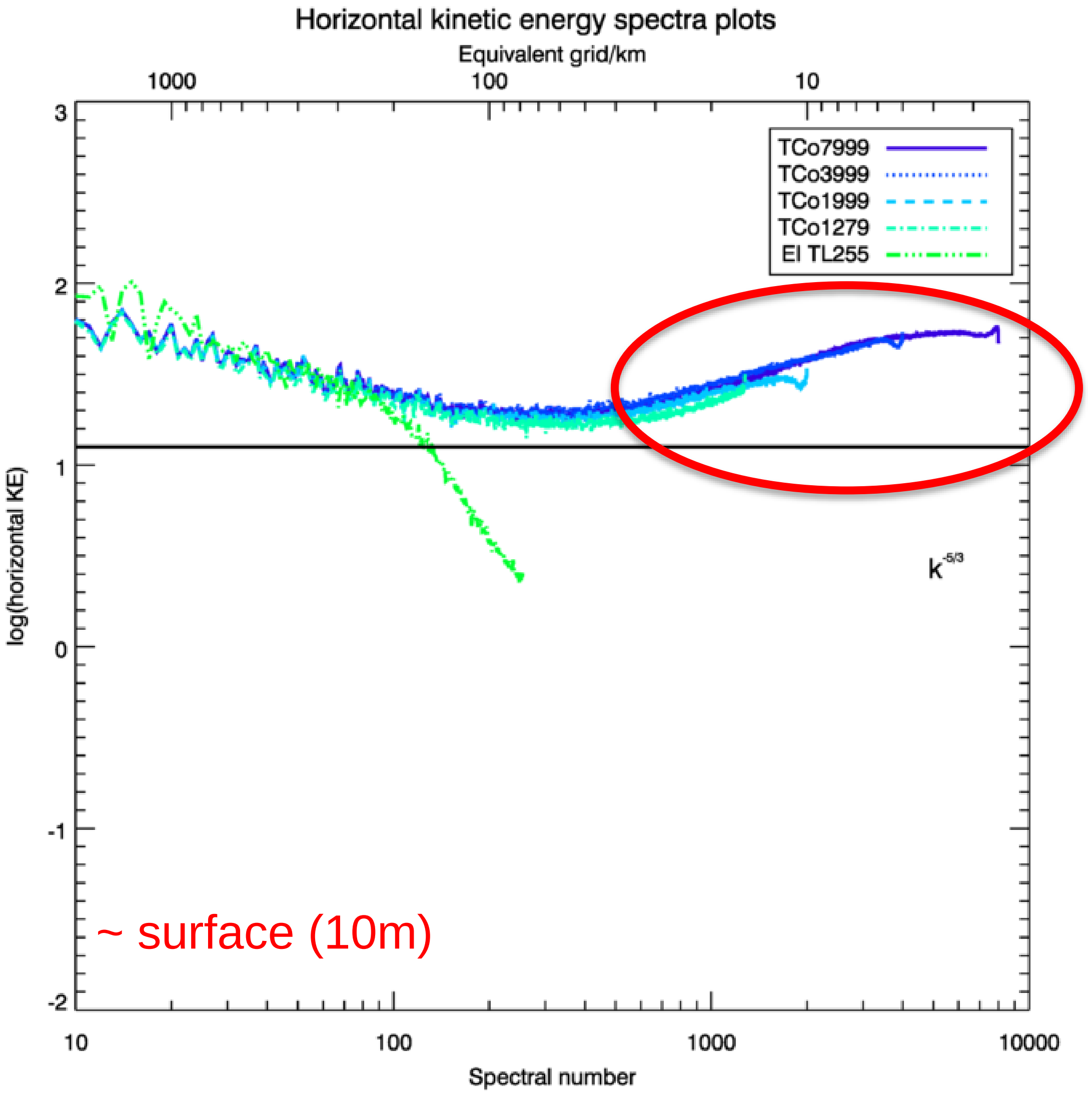


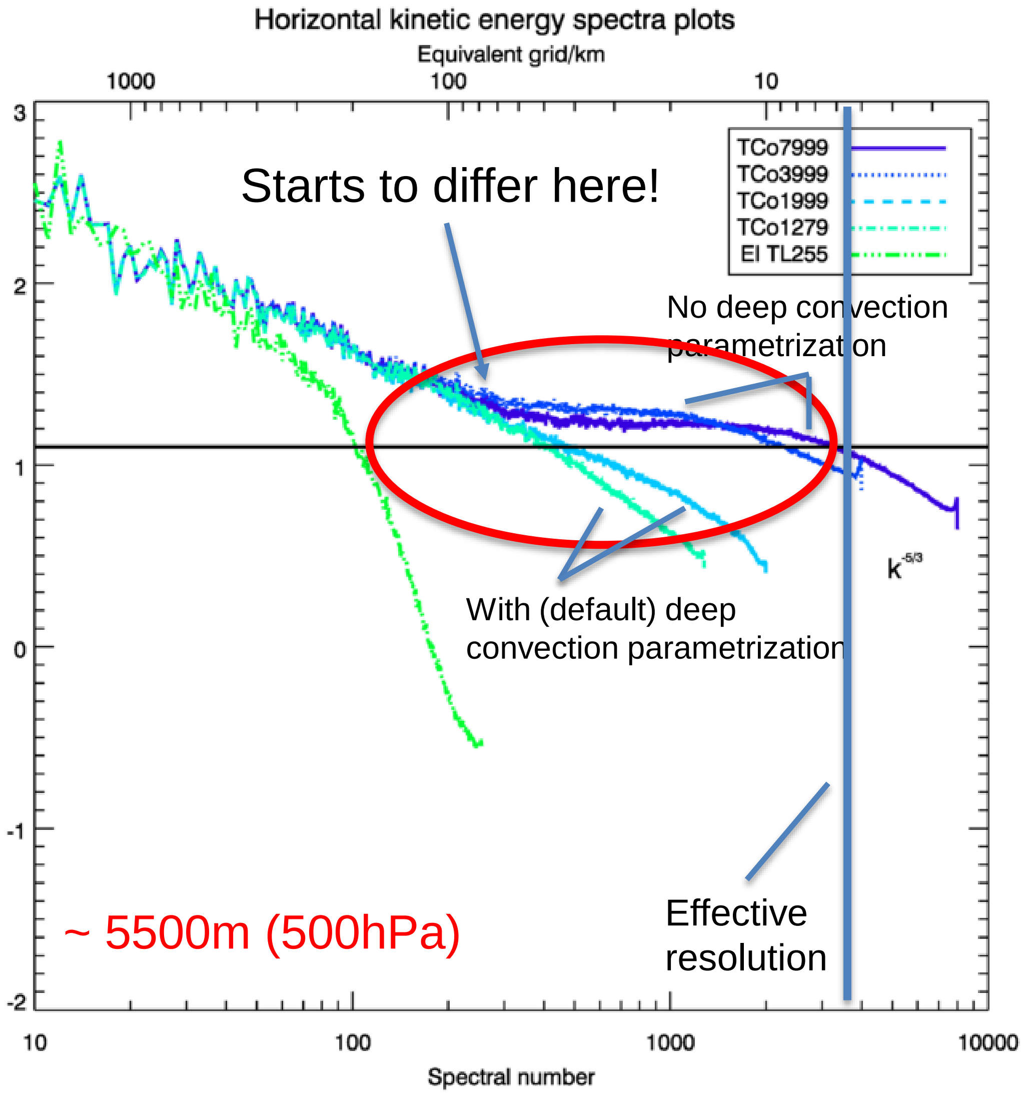
P. Bechtold in collaboration with DWD presented in ECMWF's Annual Seminar on physical processes in present and future large-scale models, 2015



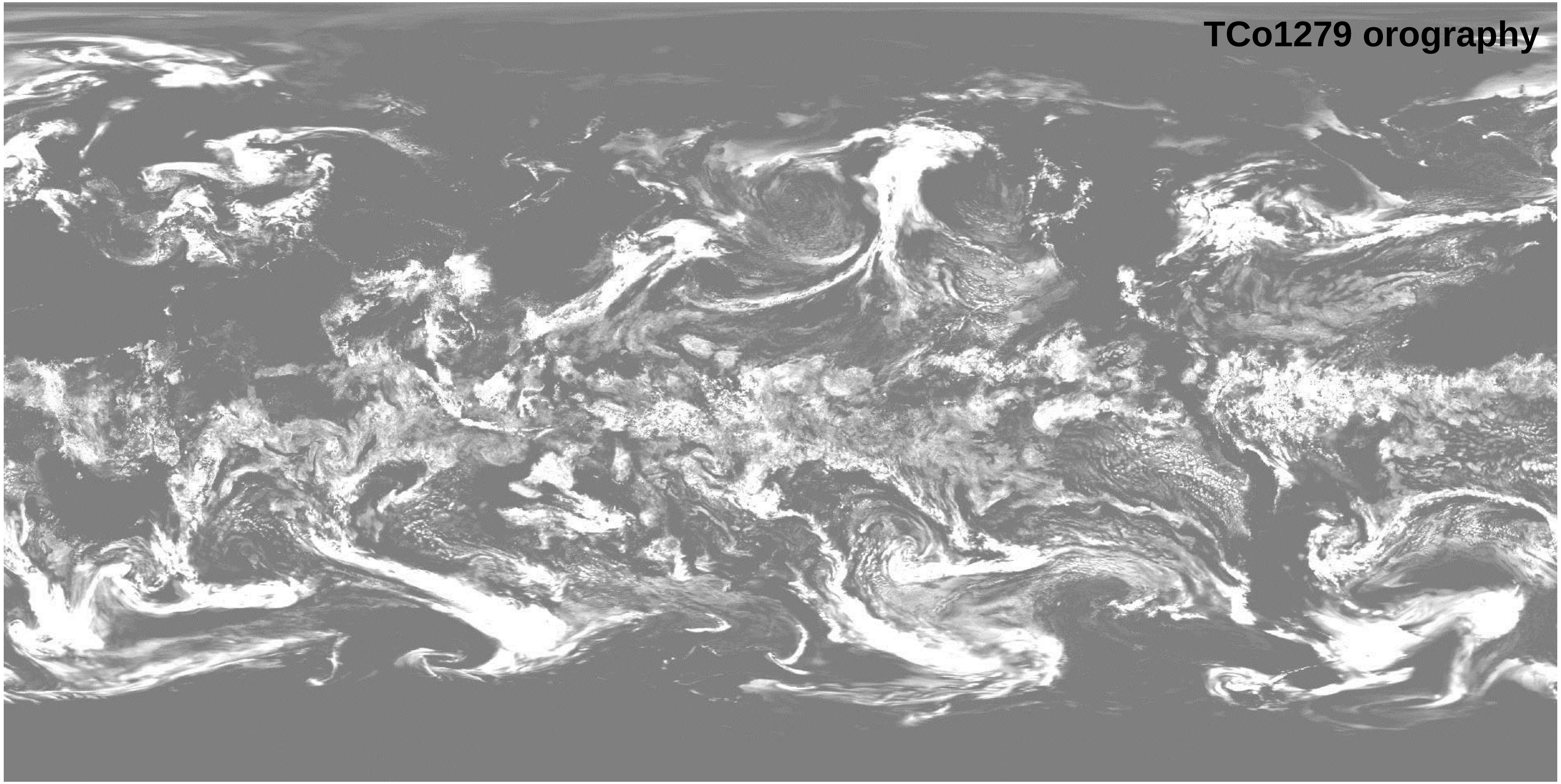
5km simulation with rescaled mass-flux parameterization

Global Kinetic Energy (scaled by n^5/3) in 1.3km global simulations TCo7999 L62



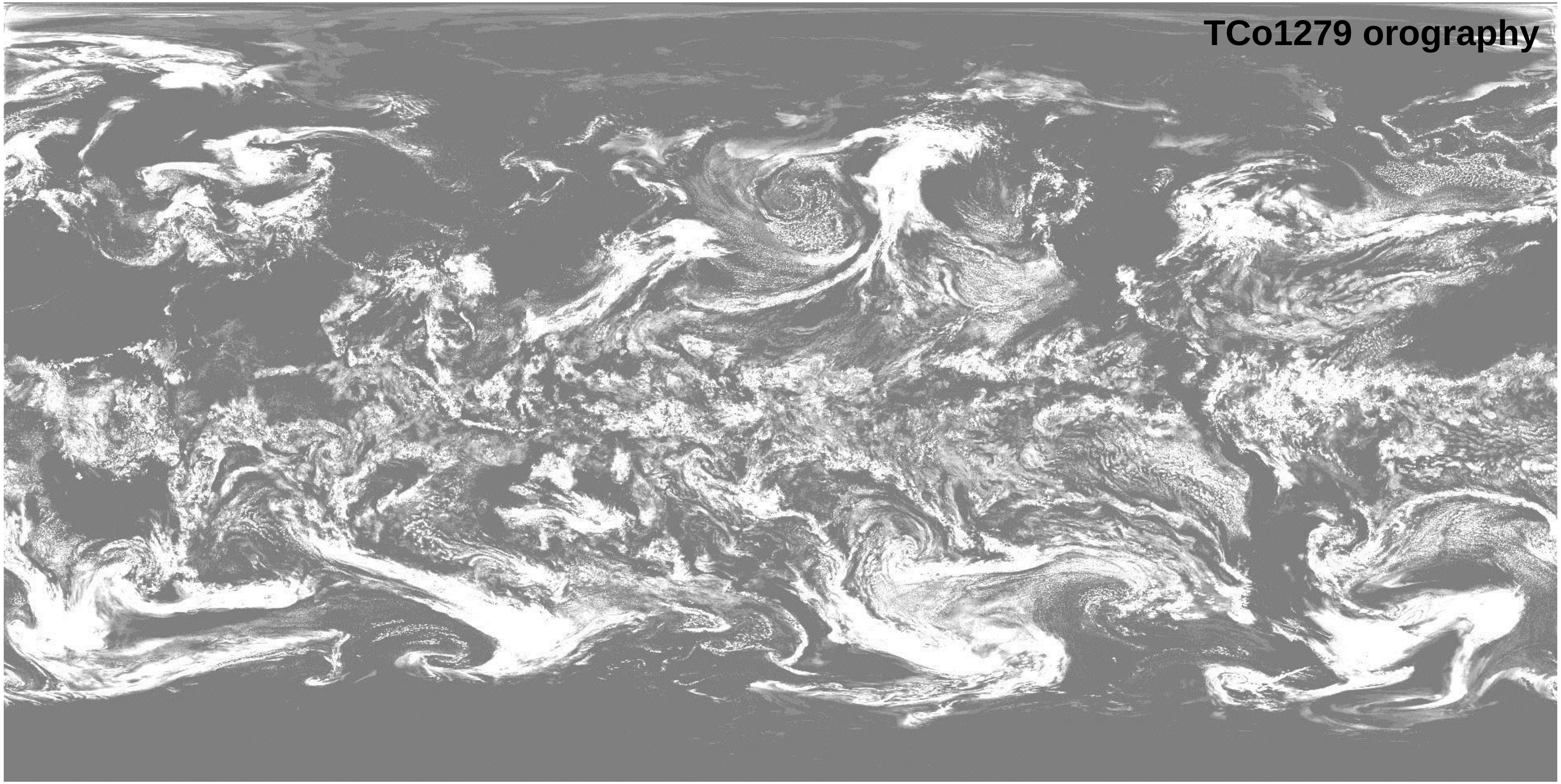


TCo1279 (~9km) total column liquid water (after 12h of simulation)



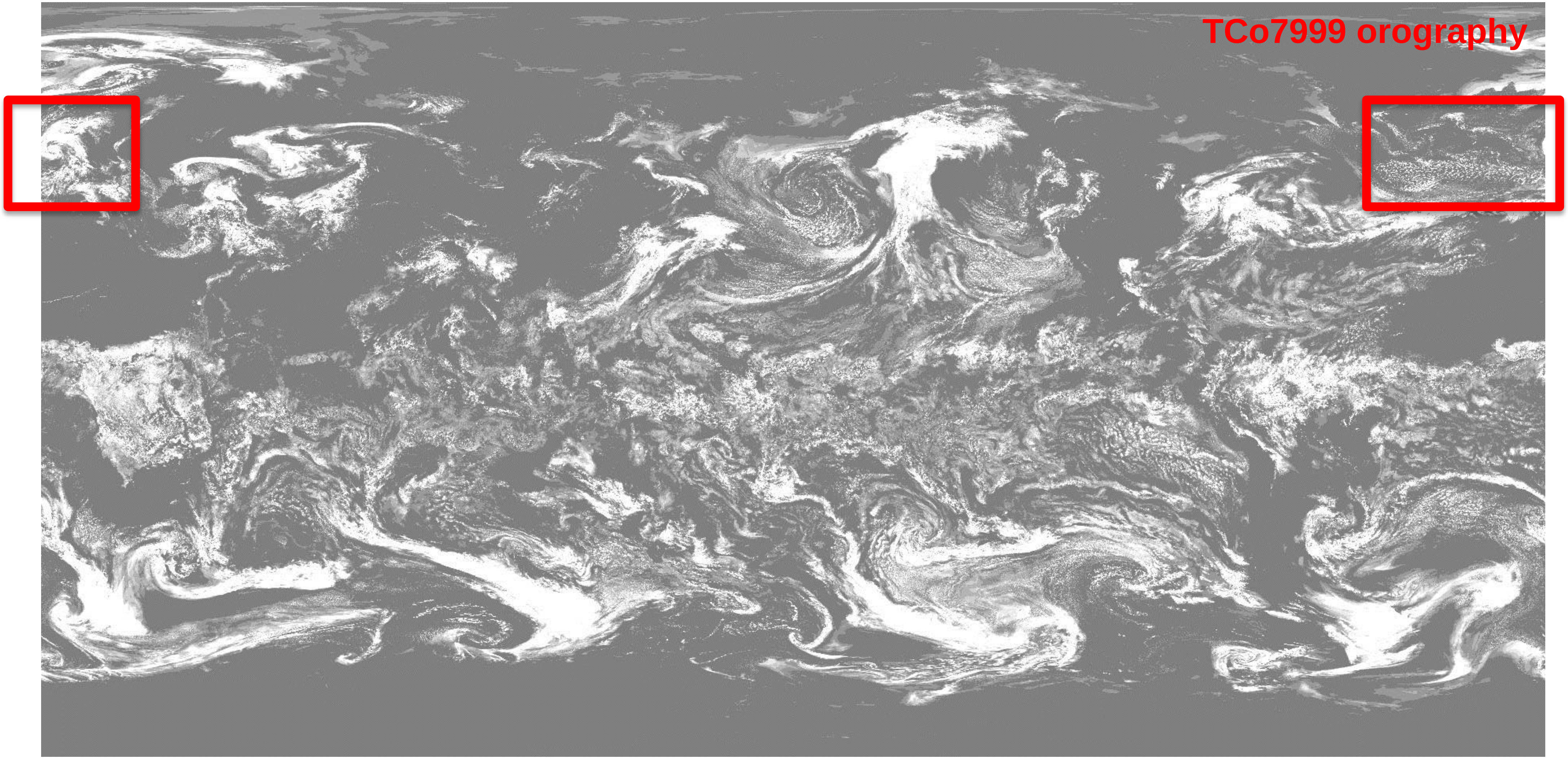
(*hydrostatic*, with deep convection parametrization, 450s time-step, 240 Broadwell nodes, ~0.75s per timestep)

TCo7999 (~1.3km) total column liquid water (after 12h of simulation)



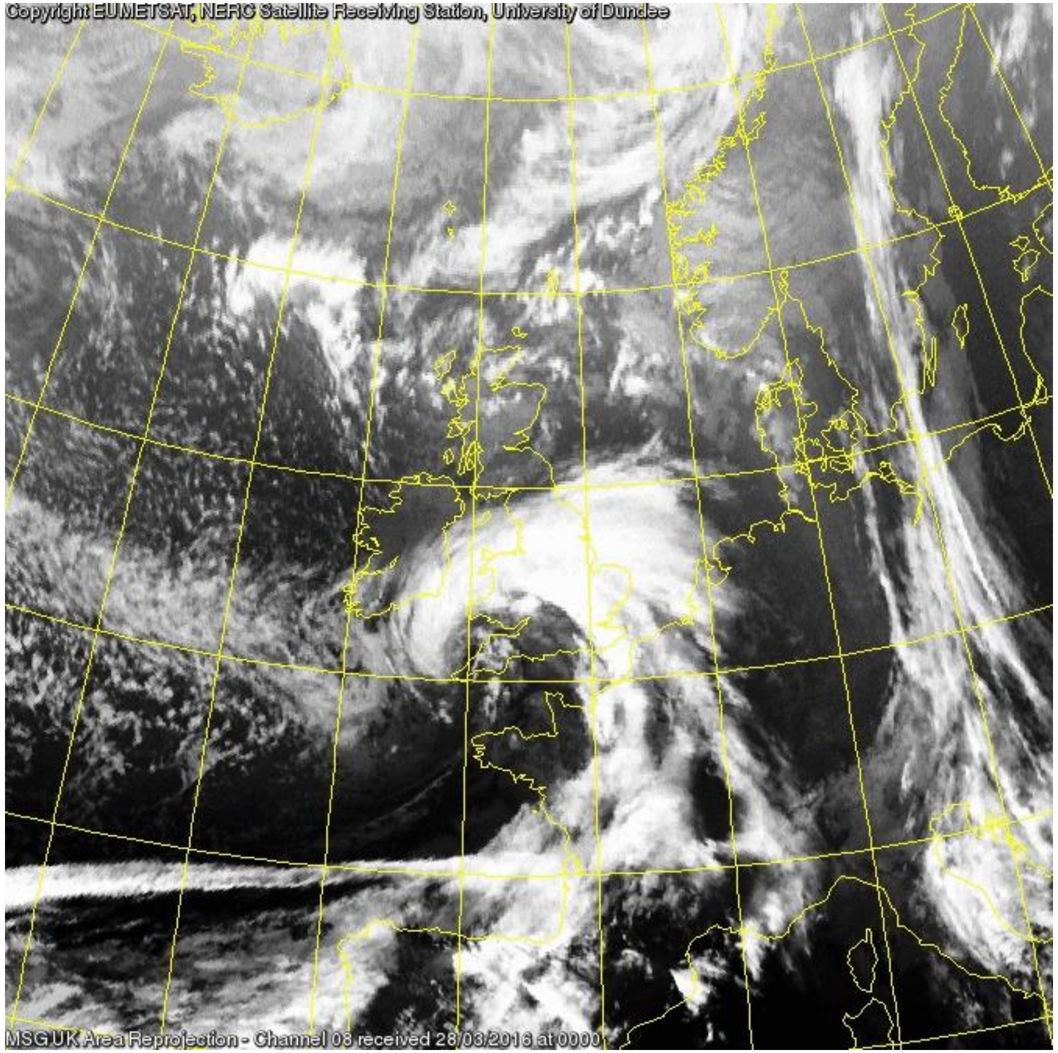
(*non-hydrostatic*, no deep convection parametrization, 120s time-step, 960 Broadwell nodes, ~30s per timestep)

TCo7999 (~1.3km) total column liquid water (after 12h of simulation)

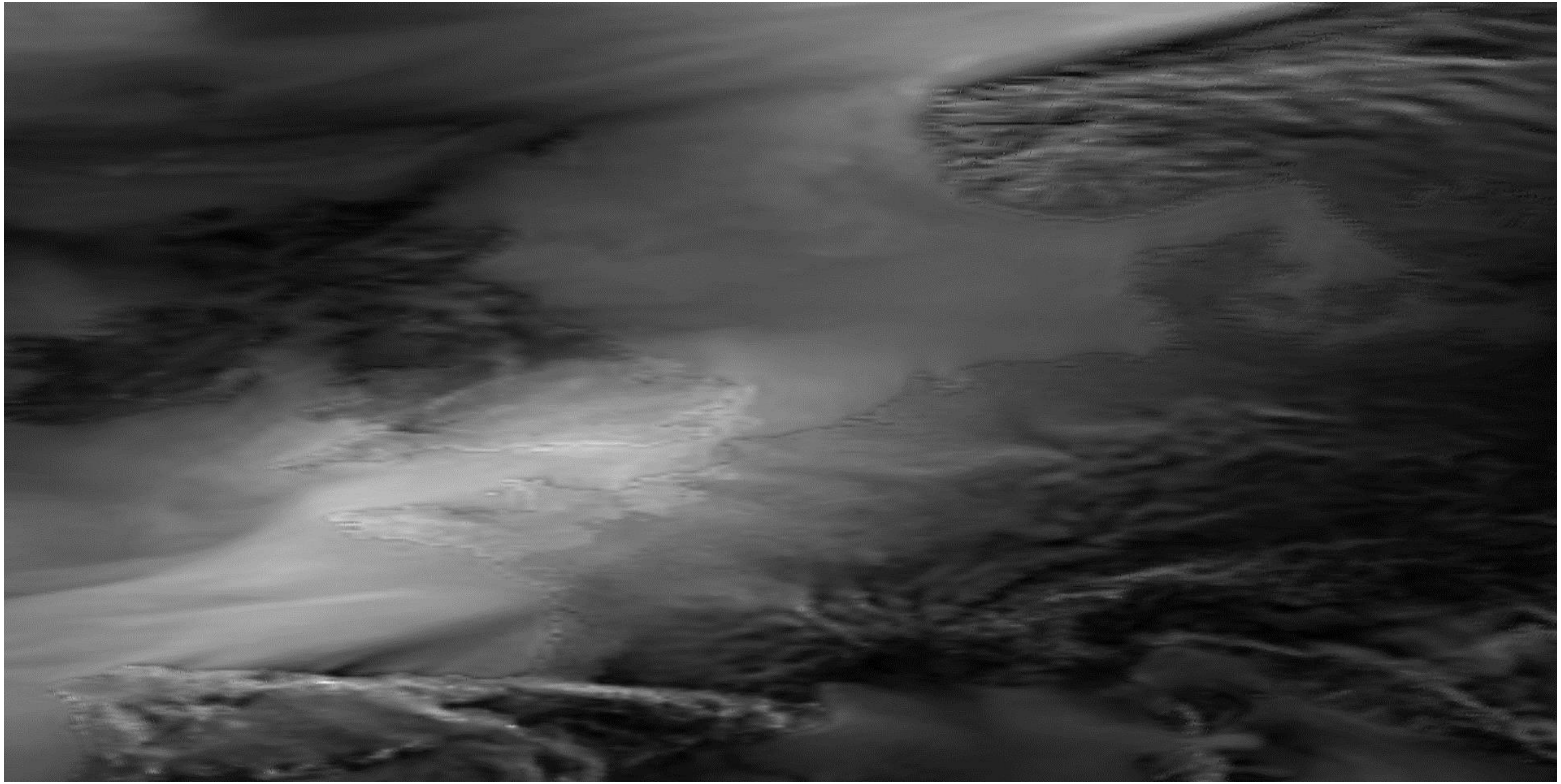


(*hydrostatic*, no deep convection parametrization, 120s time-step, 960 Broadwell nodes, ~10s per timestep)

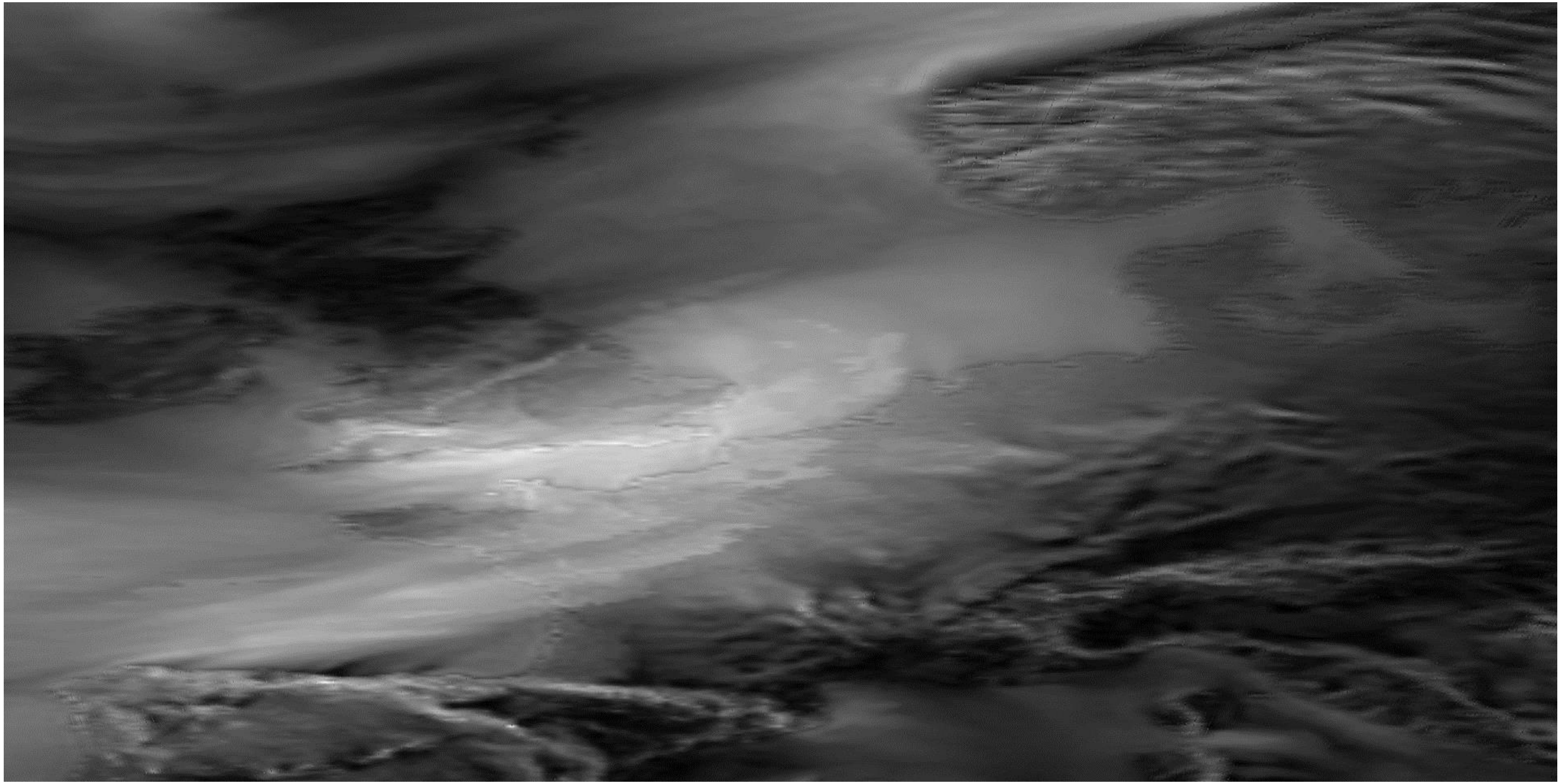
Windstorm Katie Easter 2016



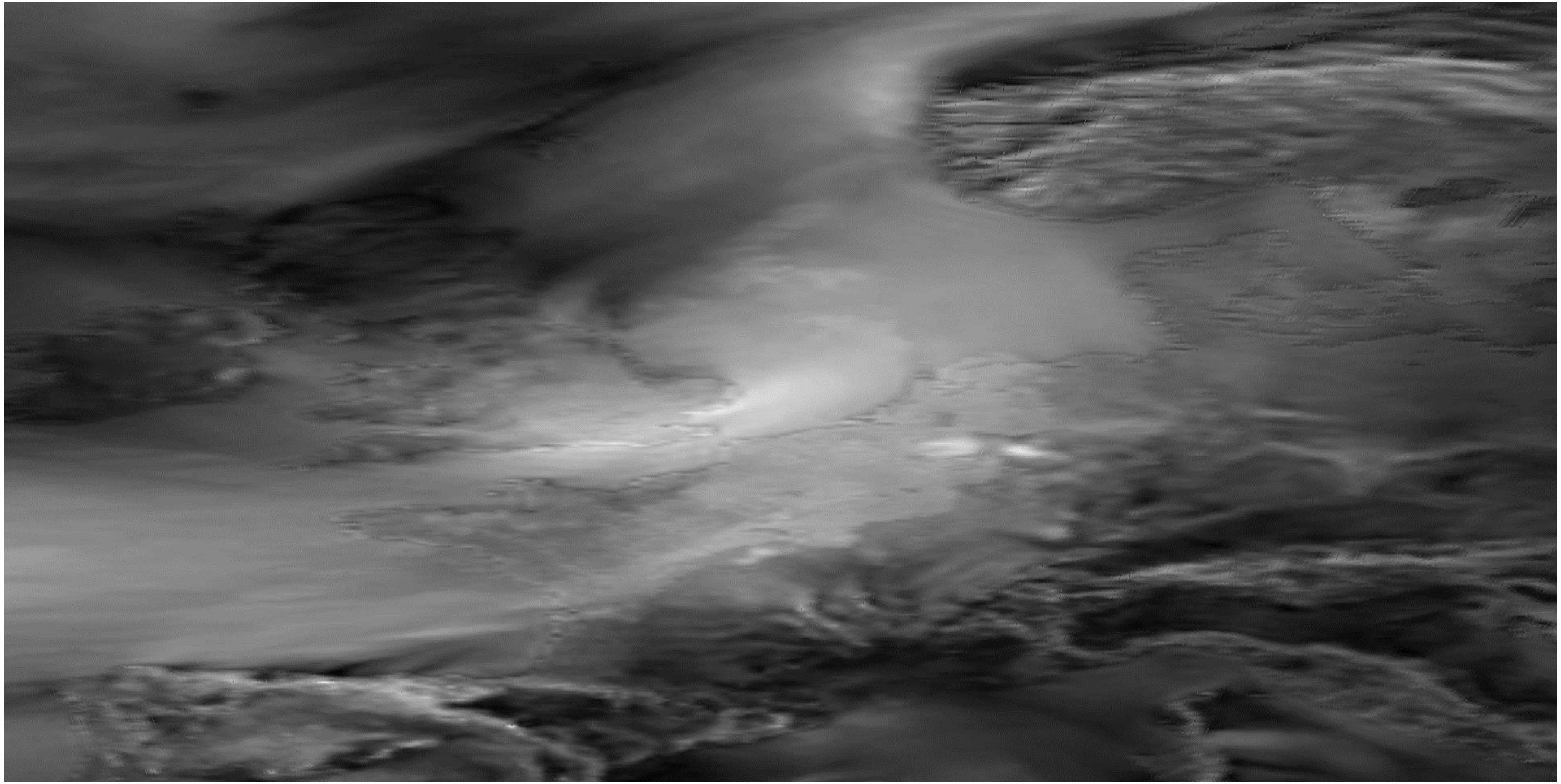
Windgusts TCo1279 (~9km) 03UTC



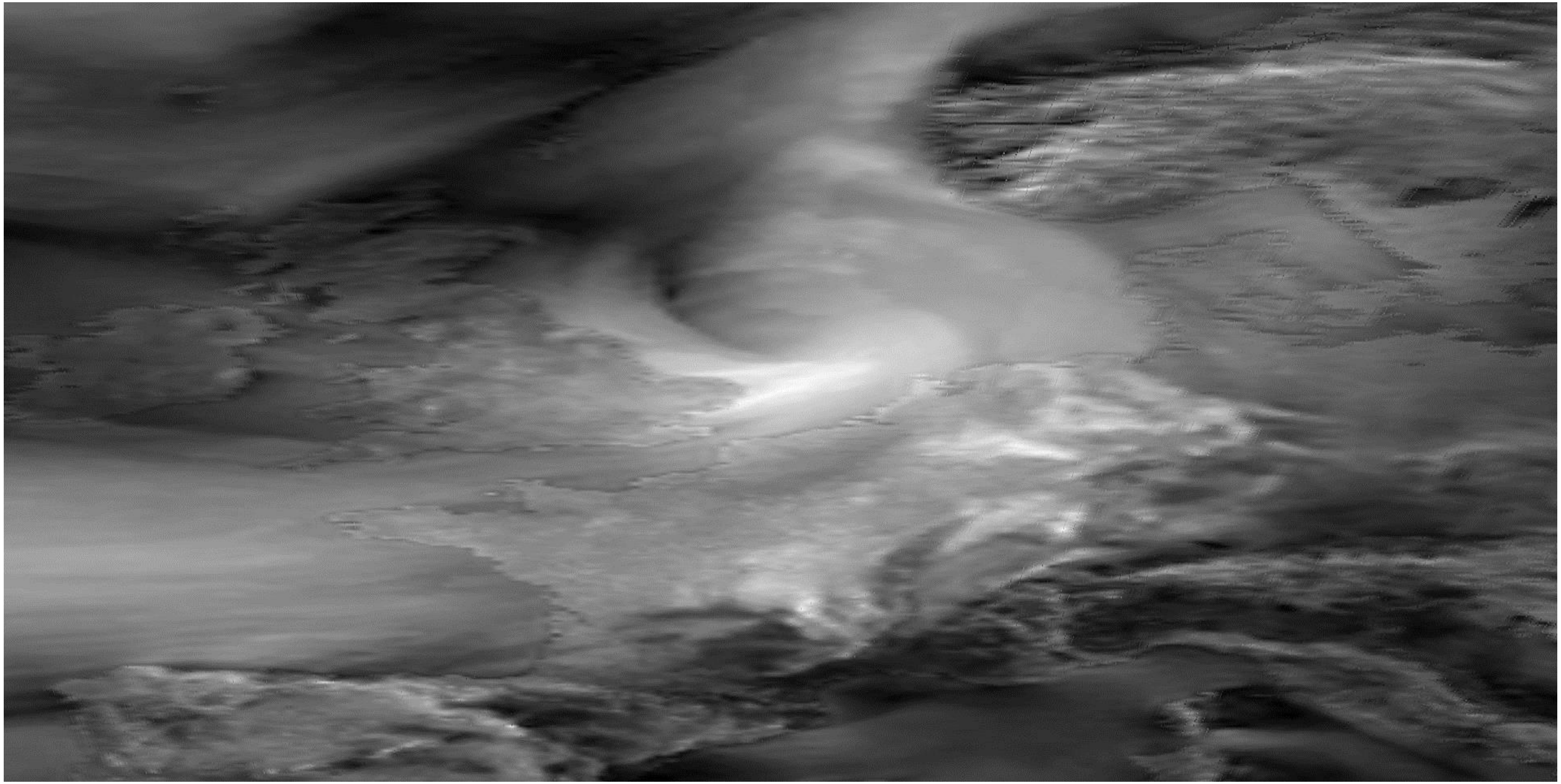
Windgusts TCo1279 (~9km) 06UTC



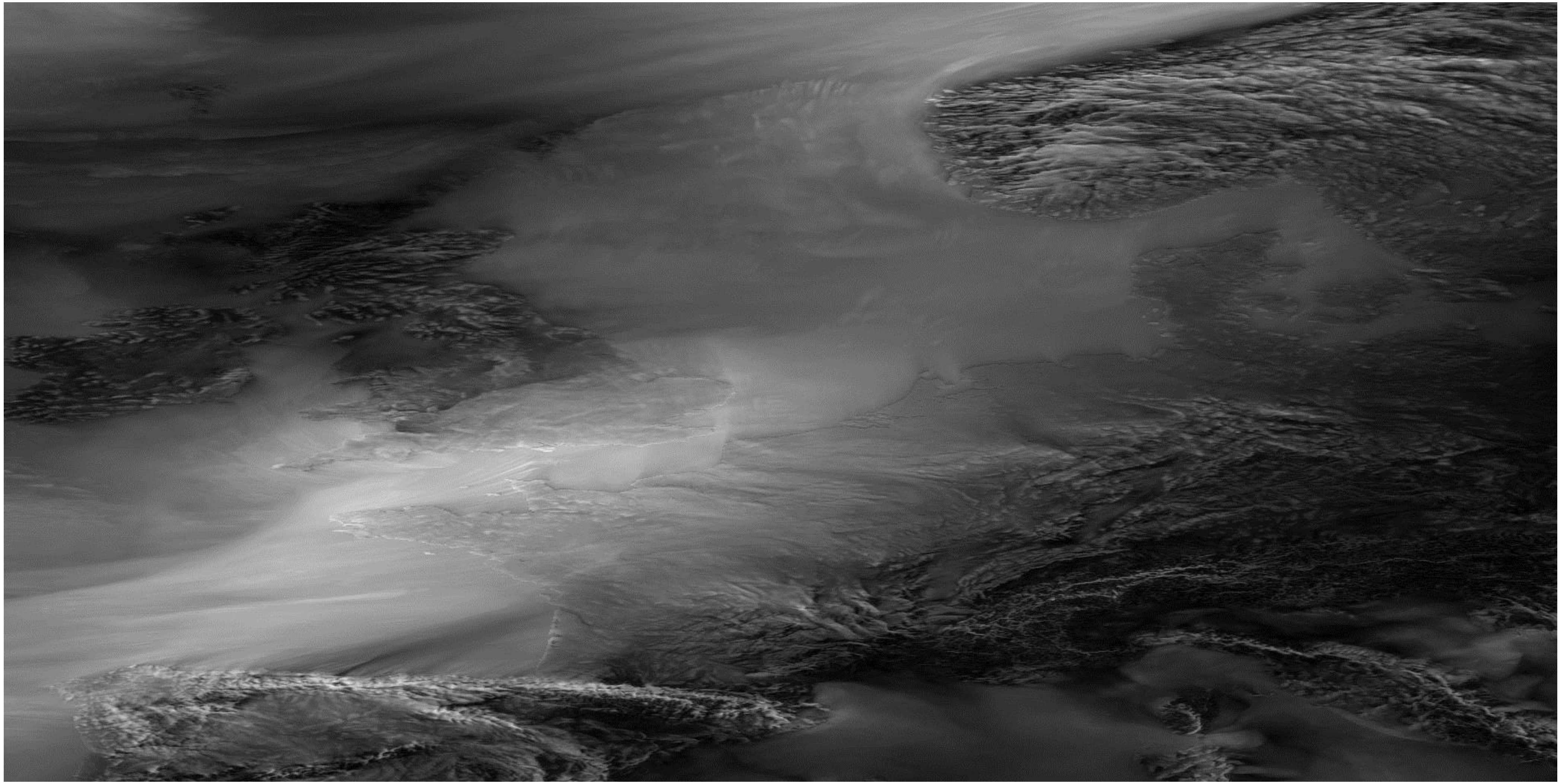
Windgusts TCo1279 (~9km) 09UTC



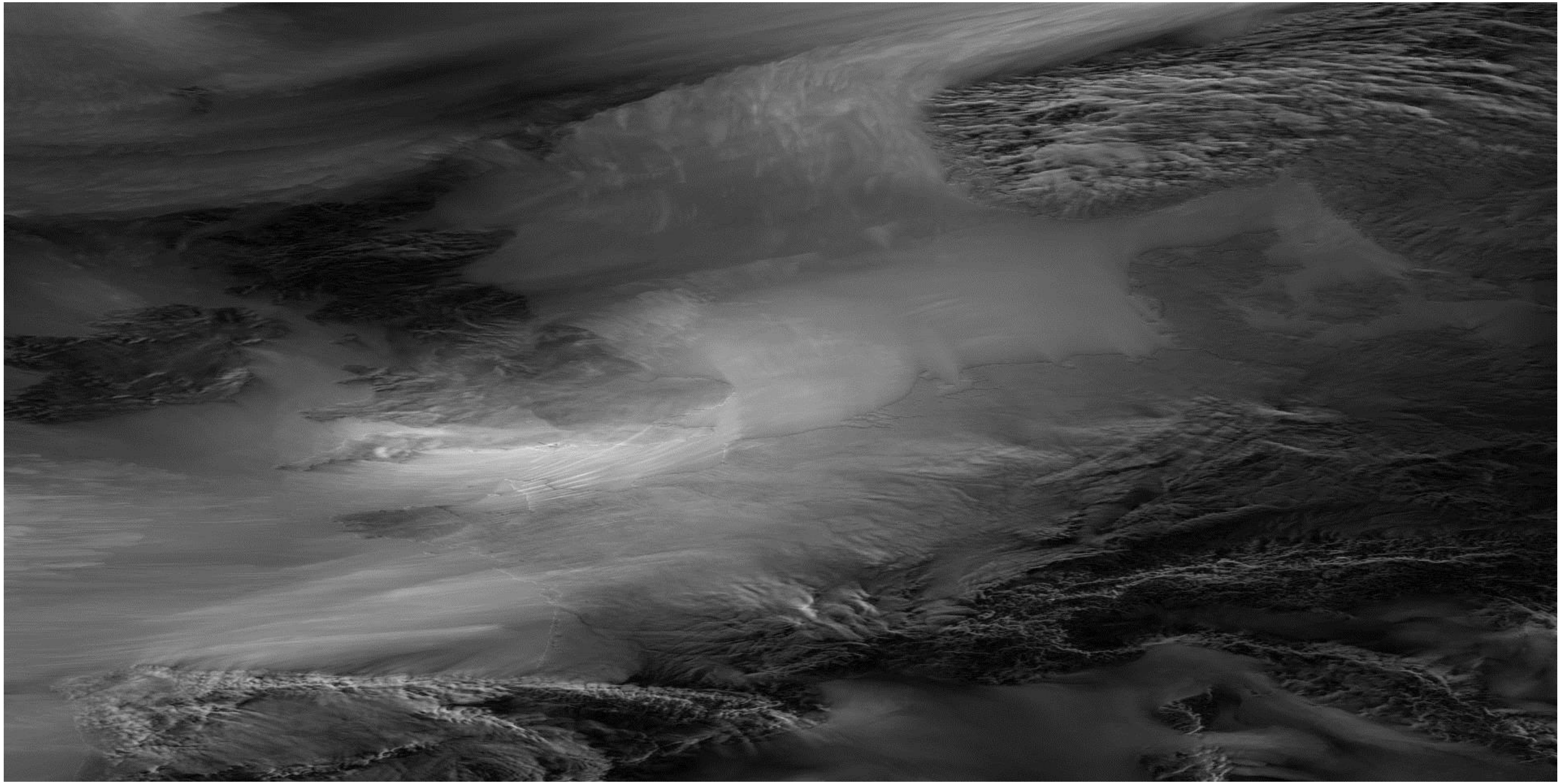
Windgusts TCo1279 (~9km) 12UTC



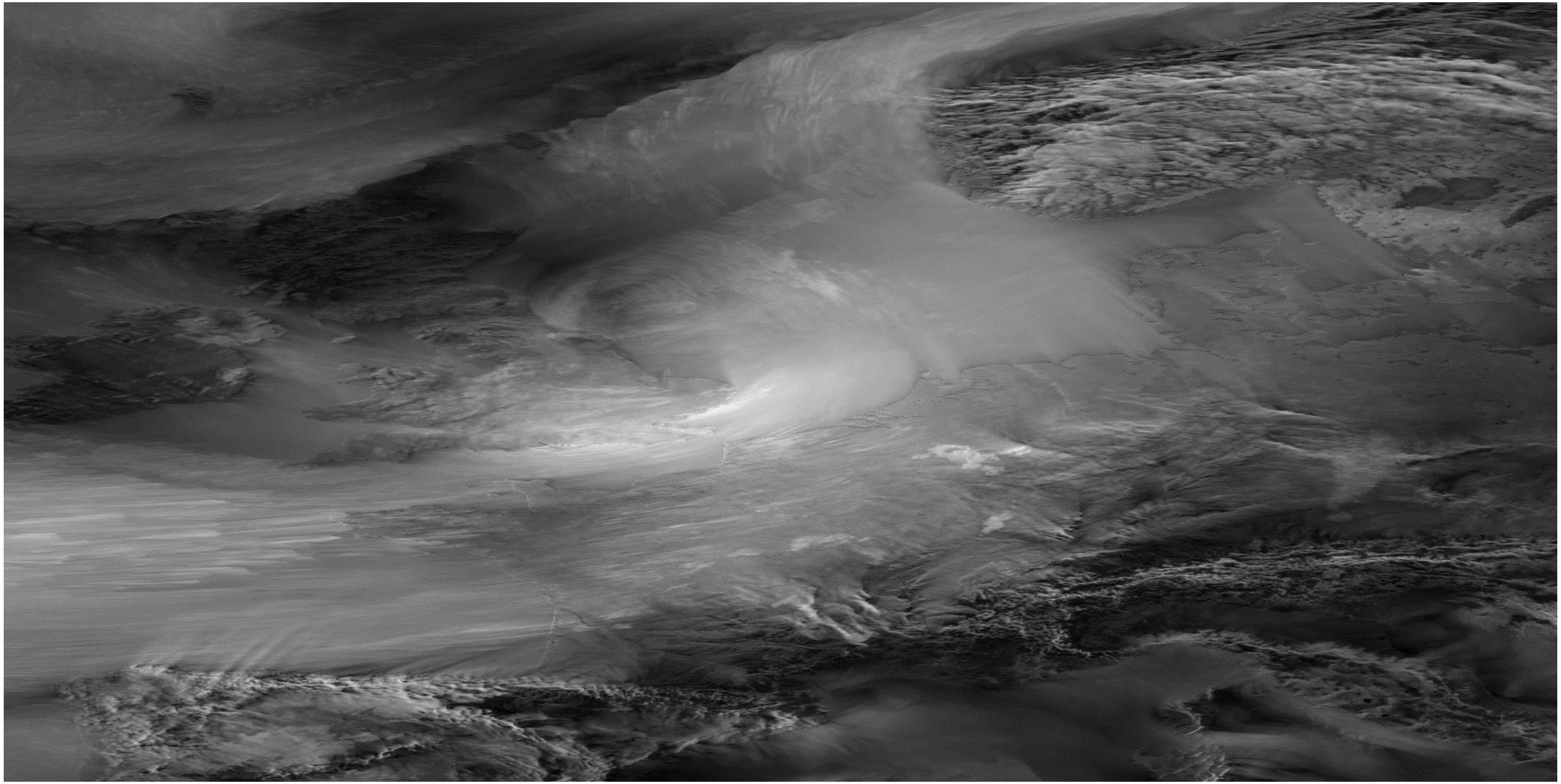
Windgusts TCo7999 (~1.3km) 03UTC



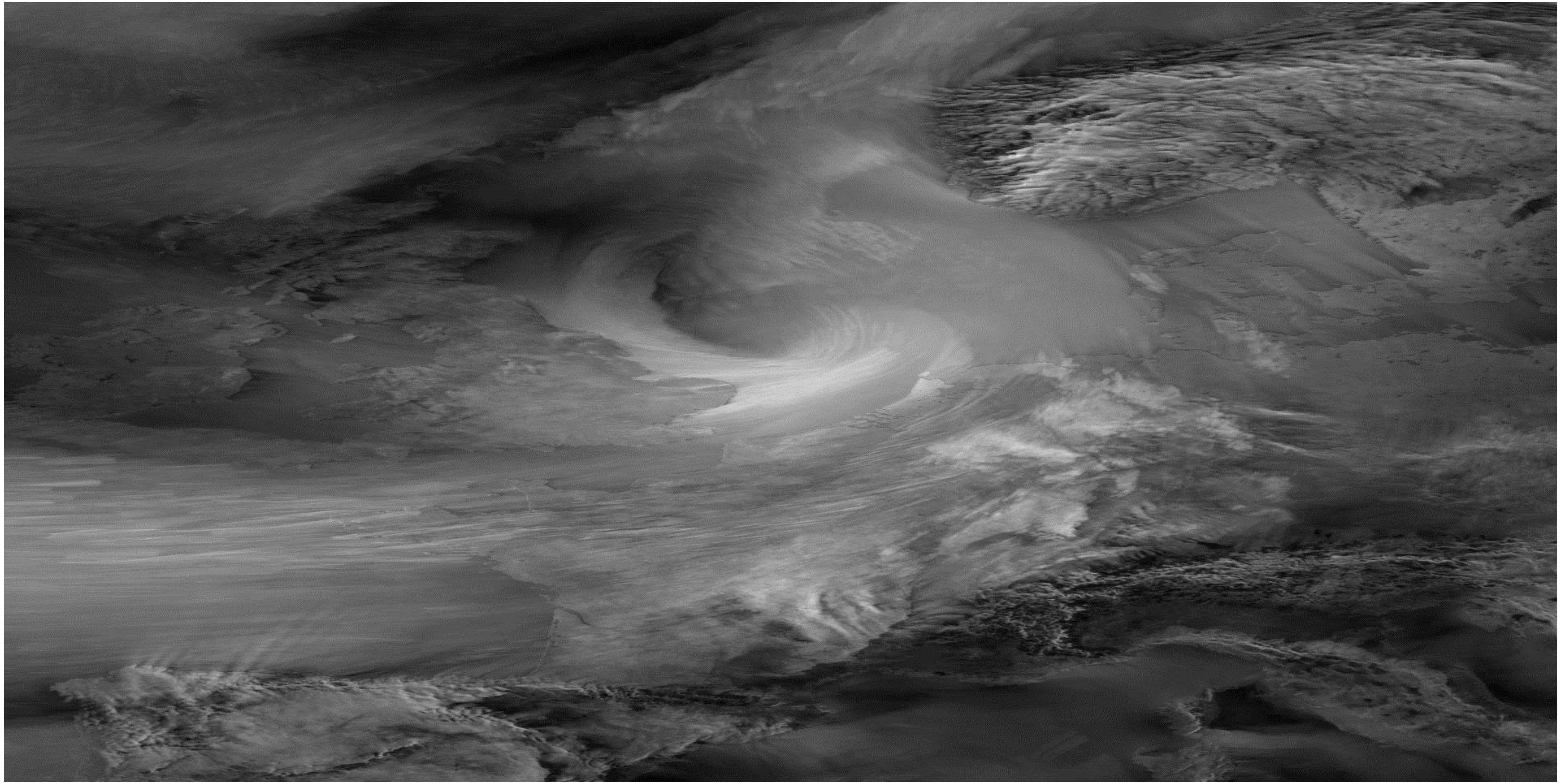
Windgusts TCo7999 (~1.3km) 06UTC



Windgusts TCo7999 (~1.3km) 09UTC



Windgusts TCo7999 (~1.3km) 12UTC



Conclusions

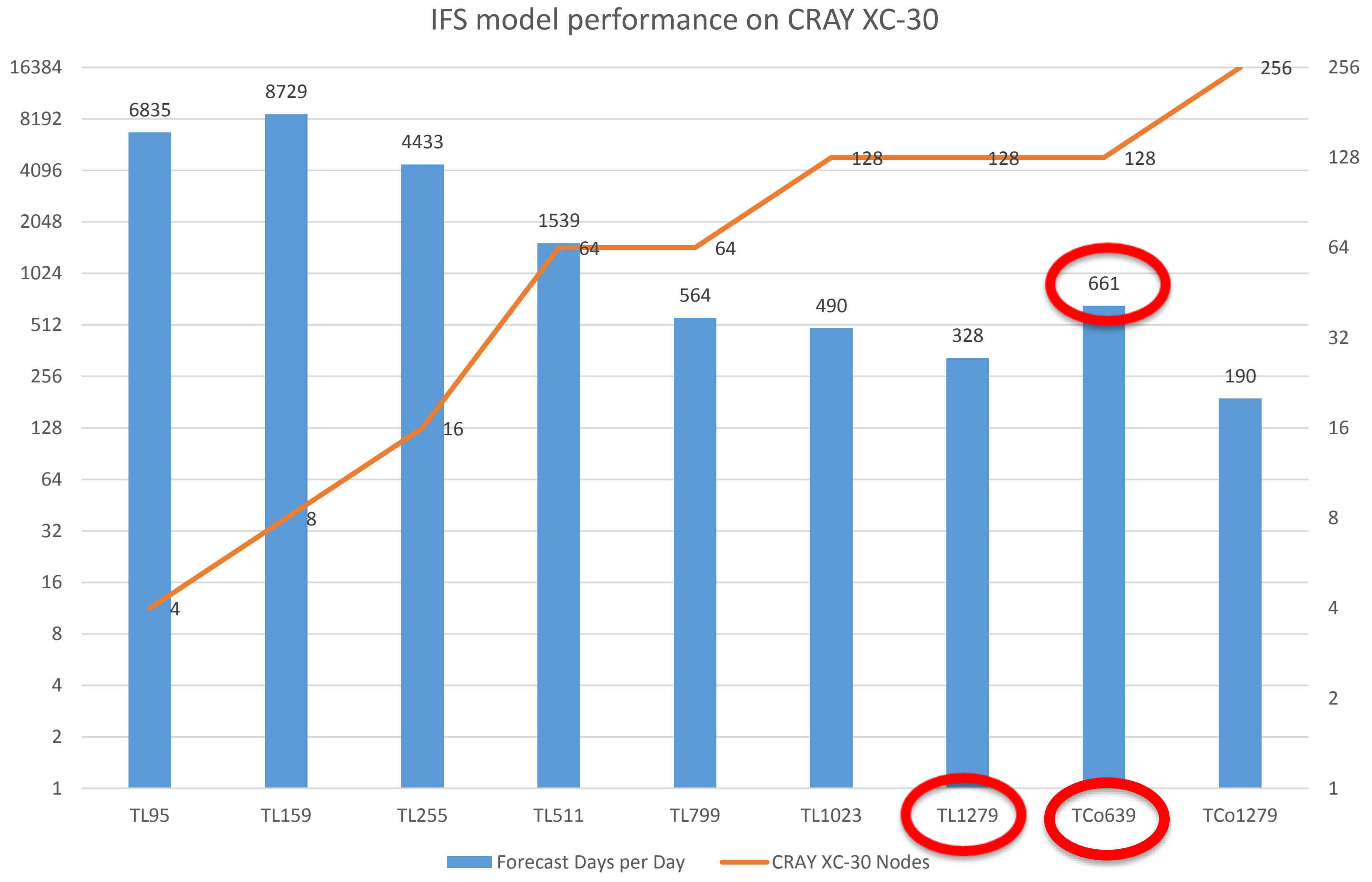
- High level of optimization in ECMWF's IFS spectral transforms
- ~1km global simulations almost within reach (~ x20 speed-up for operational use (240FC/day) needed ☺)

... the global spectral transform method is not dead, but need to prepare to

- *Increase* the flexibility in discretization choices (and/or hybrid solution procedures)
- *Develop* alternative algorithms and methods that reduce data movement, as well as communication and synchronization
- *Add* numerical and structural (code) flexibility in the effort towards full Earth-System complexity

Additional slides

RAPS14 performance



Global mass conservation in IFS