



Funded by the  
European Union

Co-ordinated by  **ECMWF**

# WP1: Algorithms and mathematics

# ESCAPE 2





**W.Deconinck, P.Düben, A.Müller, G.Tumolo (ECMWF)**

**M.Gillard, J.Szmelter (LU)**

**P.Ukkonen, K.P.Nielsen (DMI)**

**T.Benacchio, L.Bonaventura (POLIMI)**



**D1.1+ D1.2: A SL-DG dwarf and a SI-SL DG prototype dynamical core**

**D1.3+ D1.4 A report on identification of local data recovery approaches suitable for weather - climate prediction applications and fault tolerant implementations of GMRES**

**D1.5 Multigrid & multilevel technology-based IFS-FVM model**

**D1.6 Training & preliminary validation of ANN for the physical parametrization of radiation**

**D1.7+ D1.8 Weather and climate dwarfs extracted from participating models, complemented by some novel dwarfs**

**D5.4 Summer school**



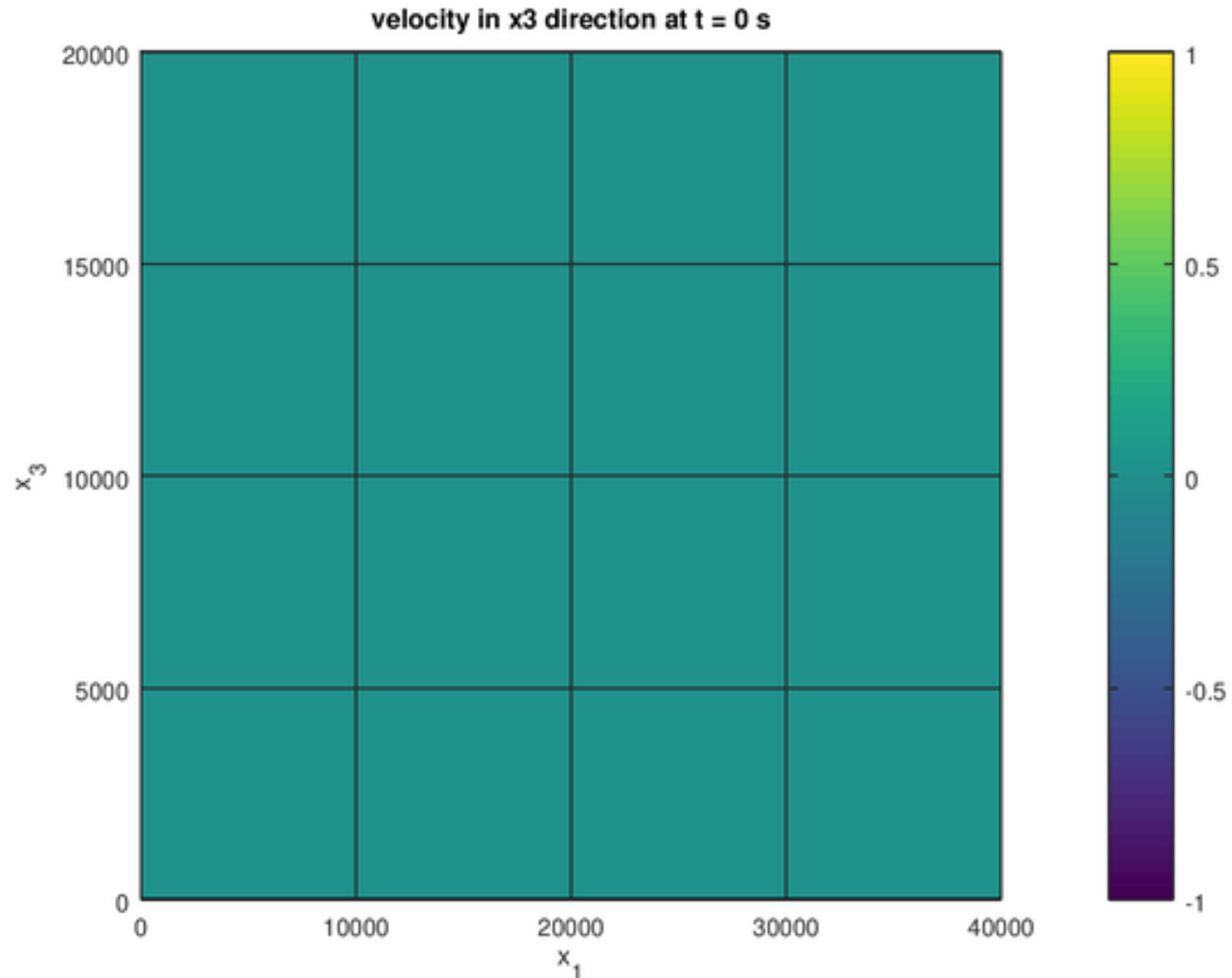
# Deliverables D1.1 and D1.2

## A prototype semi-implicit semi-Lagrangian DG dynamical core



Nonhydrostatic, nonlinear  
mountain wave test on **vertical  
slice mesh**:

- **Horizontal** resolution 200m,  
**vertical** resolution 100m
- Polynomial degrees **up to 8**,  
with **different values** in the  
vertical and horizontal  
directions
- Timestep 2s, yielding  
acoustic **Courant number**  
around **20**





- **Euler equations with rotation** in advective vector form and in **spherical geometry**

$$\frac{D\Pi}{Dt} = -(\gamma - 1) \Pi \nabla \cdot \mathbf{u},$$

$$\frac{D\mathbf{u}}{Dt} = -c_p \Theta \nabla \Pi - g \mathbf{k} - 2\boldsymbol{\Omega} \times \mathbf{u},$$

$$\frac{D\Theta}{Dt} = 0,$$

- **Reference state** is introduced  $\Pi(\mathbf{x}, t) = \pi^*(x_3) + \pi(\mathbf{x}, t), \quad \Theta(\mathbf{x}, t) = \theta^*(x_3) + \theta(\mathbf{x}, t), \quad c_p \theta^* \frac{d\pi^*}{dx_3} = -g,$
- **Metric and Coriolis terms are combined**  $\tilde{f}_1 = \frac{u_2}{\tilde{x}_3}, \quad \tilde{f}_2 = 2\Omega \cos x_2 + \frac{u_1}{\tilde{x}_3}, \quad \tilde{f}_3 = 2\Omega \sin x_2 + \frac{u_1 \tan x_2}{\tilde{x}_3},$
- **Governing equations** are then written component-wise as:

$$\frac{D\Pi}{Dt} = -(\gamma - 1) \Pi \mathcal{D}_i u_i,$$

$$\frac{Du_1}{Dt} = -c_p \Theta \mathcal{G}_1 \pi + \tilde{f}_3 u_2 - b \tilde{f}_2 u_3,$$

$$\frac{Du_2}{Dt} = -c_p \Theta \mathcal{G}_2 \pi - \tilde{f}_3 u_1 - b \tilde{f}_1 u_3,$$

$$\frac{Du_3}{Dt} = -c_p \Theta \mathcal{G}_3 \pi + g \frac{\theta}{\theta^*} + \tilde{f}_2 u_1 + \tilde{f}_1 u_2,$$

$$\frac{D\theta}{Dt} = -\frac{d\theta^*}{dx_3} u_3.$$

where, given a vector and a scalar field  $v_i, q$ :

$$\mathcal{D}_i v_i = \frac{1}{m_1 m_2 m_3} \frac{\partial}{\partial x_i} \left( \frac{m_1 m_2 m_3 v_i}{m_i} \right), \quad \mathcal{G}_i q = \frac{1}{m_{(i)}} \frac{\partial q}{\partial x_{(i)}},$$

and:  $m_1 = a \cos x_2, m_2 = a, m_3 = 1$



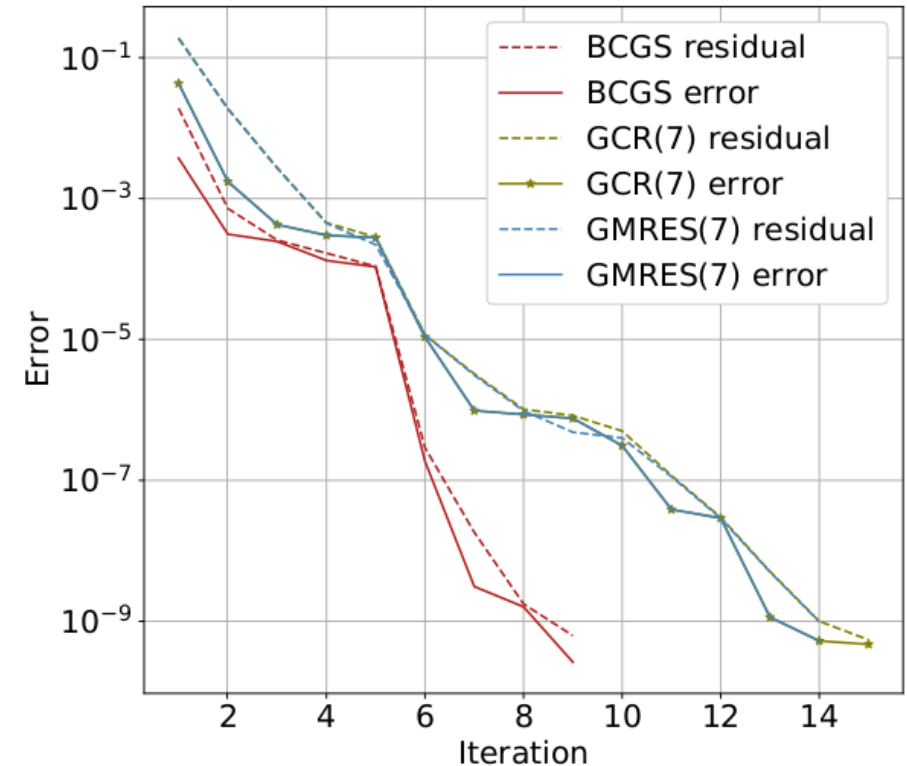
P-  
Adaptive  
Numerical  
Tool for  
High  
ordEr  
discRetizations

- **Modal** DG: tensor product of 1-D Legendre polynomials
- **Direct** addressing of *dof* as well as quadrature nodes and weights within hexahedral elements
- **Indirect** addressing of columns of elements vs. **direct** addressing of elements within columns
- **Global pointer arrays** to local column-wise data structures
- **One-sided asynchronous** communication
- Coupling with **Atlas** library for mesh generation and optimal partitioning



**Krylov space solvers Bi-CGSTAB, GCR(k) and GMRES(k) implemented and tested within the SISLDG solver, fault-tolerant options available**

**Preconditioner infrastructure and vertical diagonal/tridiagonal option implemented, full integration with solver to be completed**







# Deliverables D1.3 and D1.4

A report on identification of local data recovery approaches suitable for weather - climate prediction applications and fault tolerant implementation of Krylov solvers



# Review paper on methods and prospective applications in NWP and climate models co-authored by major field experts

Research Paper

International Journal of  
HIGH PERFORMANCE  
COMPUTING APPLICATIONS

## Resilience and fault tolerance in high-performance computing for numerical weather and climate prediction

Tommaso Benacchio<sup>1</sup> , Luca Bonaventura<sup>1</sup>,  
Mirco Altenbernd<sup>2</sup>, Chris D Cantwell<sup>3</sup>, Peter D Düben<sup>4,5</sup>,  
Mike Gillard<sup>6</sup>, Luc Giraud<sup>7</sup>, Dominik Göddeke<sup>2</sup>,  
Erwan Raffin<sup>8</sup> , Keita Teranishi<sup>9</sup> and Nils Wedi<sup>4</sup>

The International Journal of High  
Performance Computing Applications  
2021, Vol. 35(4) 285–311  
© The Author(s) 2021



Article reuse guidelines:  
[sagepub.com/journals-permissions](https://sagepub.com/journals-permissions)  
DOI: 10.1177/1094342021990433  
[journals.sagepub.com/home/hpc](https://journals.sagepub.com/home/hpc)





## Algorithm 1 FT-GCR(k):

For any initial guess,  $\phi^0$ , set  $r^0 = \mathcal{L}(\phi^0) - \mathcal{R}$ ,  $p^0 = \mathcal{P}^{-1}(r^0)$ ; then iterate:

**for**  $n = 1, 2, \dots$  until convergence **do**

**for**  $\nu = 0, \dots, k - 1$  **do**

$$\beta = -\frac{\langle r^\nu \mathcal{L}(p^\nu) \rangle}{\langle \mathcal{L}(p^\nu) \mathcal{L}(p^\nu) \rangle}$$

$$\phi^{\nu+1} = \phi^\nu + \beta p^\nu$$

$$r^{\nu+1} = r^\nu + \beta \mathcal{L}(p^\nu)$$

**if**  $\|r^{\nu+1}\|_2 \leq \epsilon$  **then**

    exit

**end if**

**if**  $\|r^{\nu+1}\|_2 \geq \|r^\nu\|_2$  **then**

$n = n - 1$

    reset  $[\phi, r, p, \mathcal{L}(p)]^0$  to  $[\phi, r, p, \mathcal{L}(p)]^*$

**else if**  $\nu = 0$  **then**

    set  $[\phi, r, p, \mathcal{L}(p)]^*$  to  $[\phi, r, p, \mathcal{L}(p)]^0$

**end if**

$$e = \mathcal{P}^{-1}(r^{\nu+1})$$

  Compute  $\mathcal{L}(e)$

$$\alpha_l = -\frac{\langle \mathcal{L}(e) \mathcal{L}(p^l) \rangle}{\langle \mathcal{L}(p^l) \mathcal{L}(p^l) \rangle} \quad \forall l = 0, \dots, \nu$$

$$p^{\nu+1} = e + \sum_{l=0}^{\nu} \alpha_l p^l$$

$$\mathcal{L}(p^{\nu+1}) = \mathcal{L}(e) + \sum_{l=0}^{\nu} \alpha_l \mathcal{L}(p^l)$$

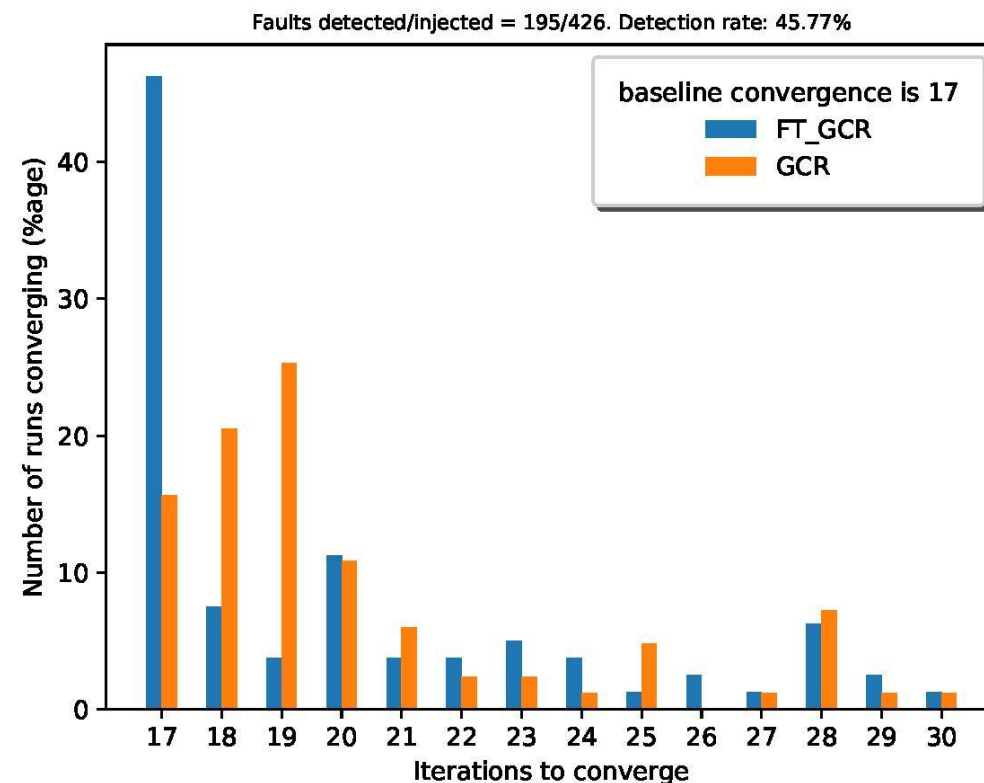
**end for**

  reset  $[\phi, r, p, \mathcal{L}(p)]^k$  to  $[\phi, r, p, \mathcal{L}(p)]^0$

**end for**

60-80% of faults detected.

Convergence delay reduced by 5 to 8 iterations compared to unprotected runs.



O80 Grid, 1% of data corrupted per fault

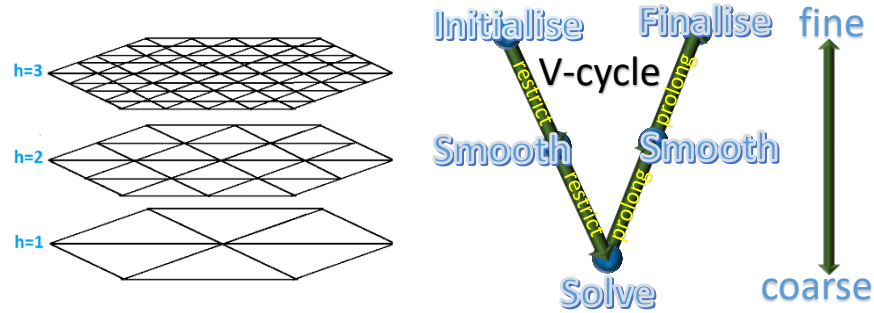


# Deliverable D1.5

## Multigrid & multilevel technology-based IFS-FVM model



## Multigrid Preconditioner for FVM



Elliptic **Helmholtz** problem, solved via preconditioned Krylov solver: **anisotropy** addressed by lagging the unstructured horizontal, using the combination of **Richardson** and **weighted Jacobi** iterations as smoothers/solvers for V-cycle **multigrid** (see below).

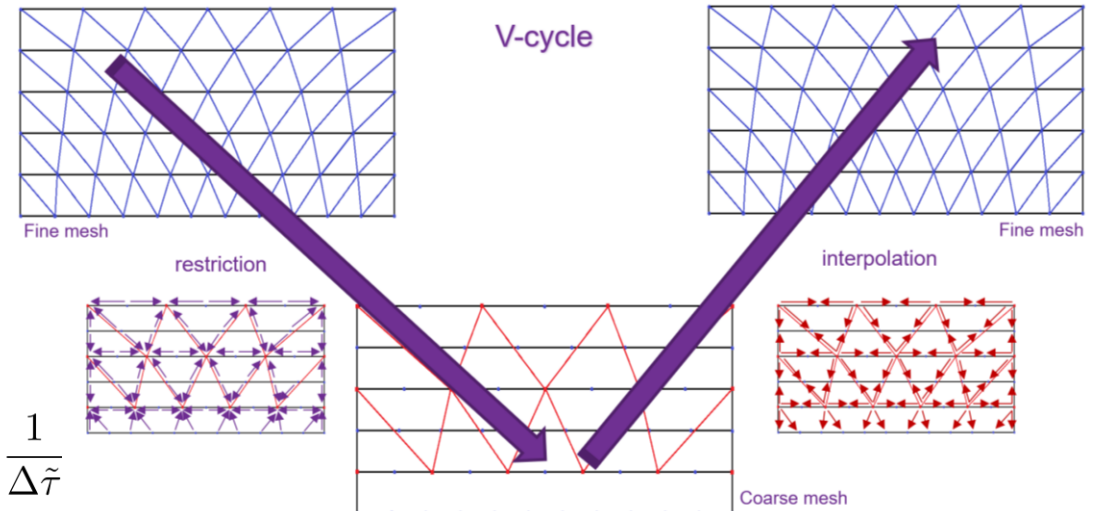
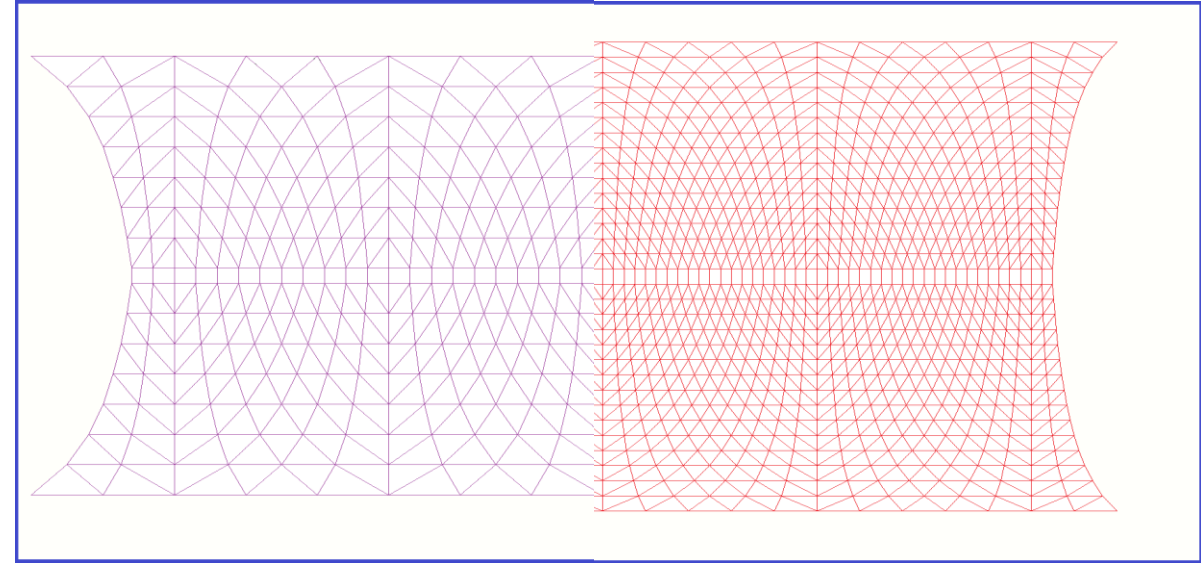
$$\mathcal{L} \approx \mathcal{P}(e) \equiv \mathcal{P}_H(e^\mu) + \mathcal{P}_z(e^{\mu+1})$$

$$0 = - \sum_{l=1}^3 \left( \frac{A_l^*}{\zeta_l} \nabla \cdot \zeta_l \tilde{\mathbf{G}}^T (\tilde{\mathbf{u}} - \mathbf{C} \nabla \varphi') \right) - B^* (\varphi' - \hat{\varphi}') \equiv \mathcal{L}(\varphi') - R$$

$$e^{\mu+1} = \omega [\mathcal{D} + \mathcal{P}_z]^{-1} (\mathcal{D}e^\mu - \mathcal{P}_H(e^\mu) - r^{\nu+1}) + (1 - \omega) e^\mu$$

$$\mathcal{D}_{k,i} = -\frac{1}{4V_i} \sum_l \frac{A_{l,k,i}^*}{\zeta_{l,k,i}} \sum_{j=1}^{nbrs} \frac{\zeta_{l,k,j}}{V_j} \left( \mathcal{S}_{x_j}^2 (\tilde{\mathbf{G}}^T \mathbf{C})_{xx_{k,j}} + \mathcal{S}_{y_j}^2 (\tilde{\mathbf{G}}^T \mathbf{C})_{yy_{k,j}} \right) \quad \text{OR} \quad \frac{1}{\Delta \tilde{\tau}}$$

## Parallel restriction, prolongation and Atlas mesh generation





## Multigrid Preconditioner performance improvement:

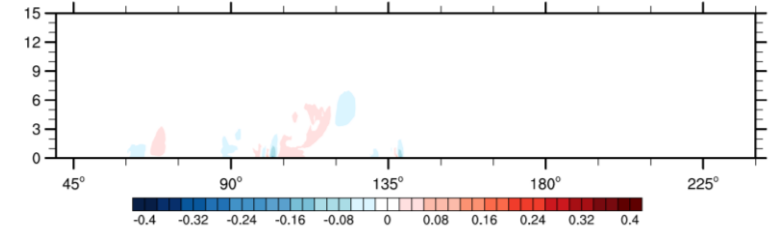
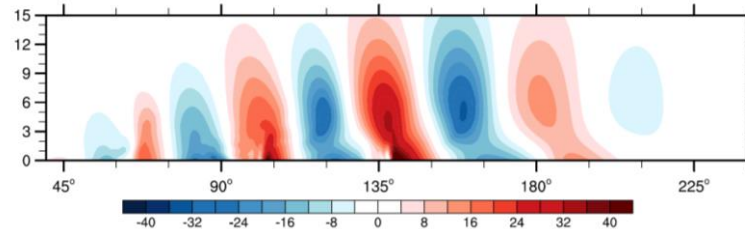
Better at high resolutions, ranging from 1.5x on the O180 grid **4.81x** on the O1800 grid in comparison to the Richardson preconditioner.

Shows excellent agreement with the previous configuration: - >

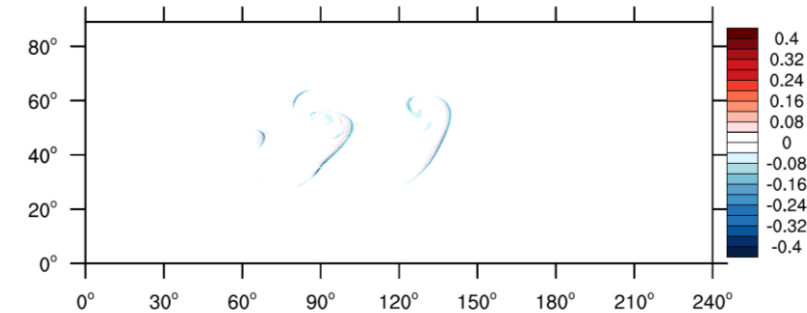
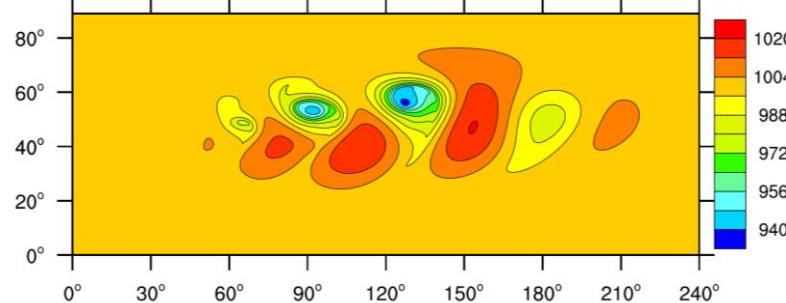
10 days, dry baroclinic instability test with 31 vertical levels, computed on the Cray XC30 at ECMWF (**Dynamical core only**)

## O360 plots for day 10 of dry baroclinic instability test: Multigrid (left), Difference to Richardson preconditioner (right).

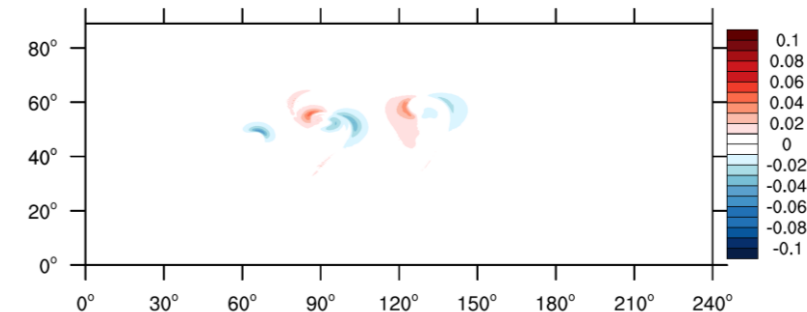
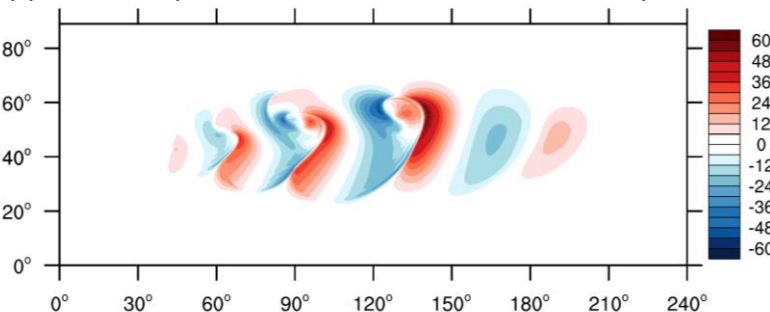
vertical velocity [m/s] at the 50°N;



Upper Hemisphere surface Pressure [hPa]



Upper Hemisphere surface meridional velocity [m/s]





## FAS multigrid with Dual time stepping (DTS) for MPDATA

$$\mathcal{L}_t \Psi^{n+1} = -R(\Psi^{n+1})$$

$$\frac{d}{d\tau} \Psi = -\tilde{R}(\Psi) \equiv R(\Psi) + \frac{3\Psi^{n+1} - 4\Psi^n + \Psi^{n-1}}{2\Delta t}$$

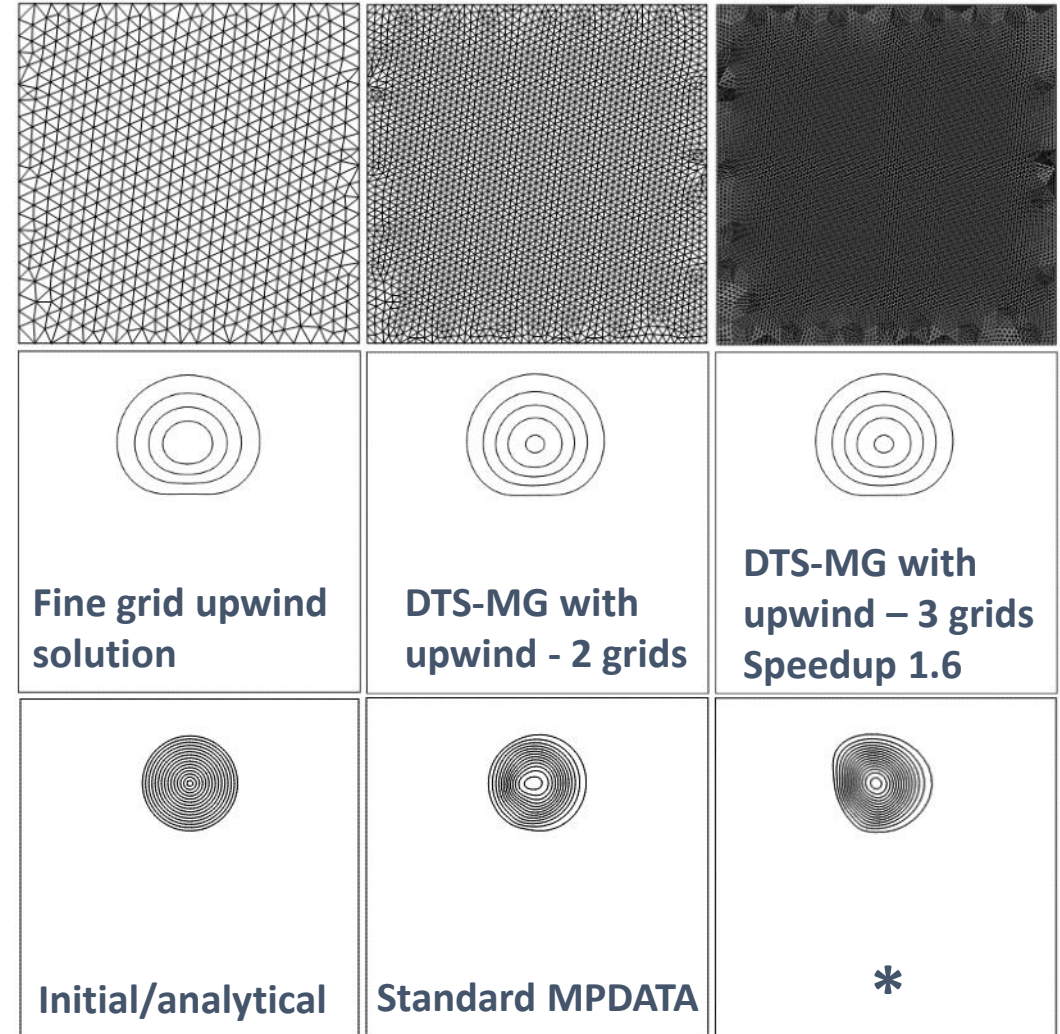
**DTS** allows for the use of a non-linear **multigrid** approach to accelerate the convergence of steady state flows for unsteady flow systems.

Initial results show some promise for a classic rotating cone test case: - >

\* DTS-MG with MPDATA (fine) , upwind (coarse), 2 grids, speedup 1.7 (time to solution)

Top: meshes coarsest (left) to finest Right.  
Middle: First order upwind as the smoother advected through 1 rotation.

Bottom: MPDATA as the smoother advected through 1 rotation.





# Deliverable D1.6

## Training & preliminary validation of ANN for the physical parametrization of radiation





- Development of an **ANN version** of the **gas optics scheme RRTMGP** to predict optical properties of the gaseous atmosphere
- **ANN** version intended as a **usable replacement for real applications** in NWP and climate: lots of training data and Hypercube sampling required
- Implementation of RRTMGP-NN in the **ECMWF radiation scheme ecRAD**
- **Comparison** of different **emulation strategies** for radiation codes



- When combined with refactoring of the radiative transfer solver, the new ANN-based radiation scheme **is 3-4 times faster** than the original code, with no loss in accuracy

- Results **published** in a paper

**JAMES** | Journal of Advances in  
Modeling Earth Systems





RESEARCH ARTICLE

10.1029/2020MS002226

**Key Points:**

- Neural networks (NNs) were trained to predict the optical properties of the gaseous atmosphere
- Training data were generated with a recently developed radiation scheme for dynamical models (RRTMGP)
- RRTMGP-NN is roughly 3 times faster than the reference code and has a similar accuracy, also in future climate scenarios

**Accelerating Radiation Computations for Dynamical Models With Targeted Machine Learning and Code Optimization**

Peter Ukkonen<sup>1,2</sup> , Robert Pincus<sup>3,4</sup> , Robin J. Hogan<sup>5</sup> , Kristian Pagh Nielsen<sup>1,5</sup>, and Eigil Kaas<sup>2</sup> 

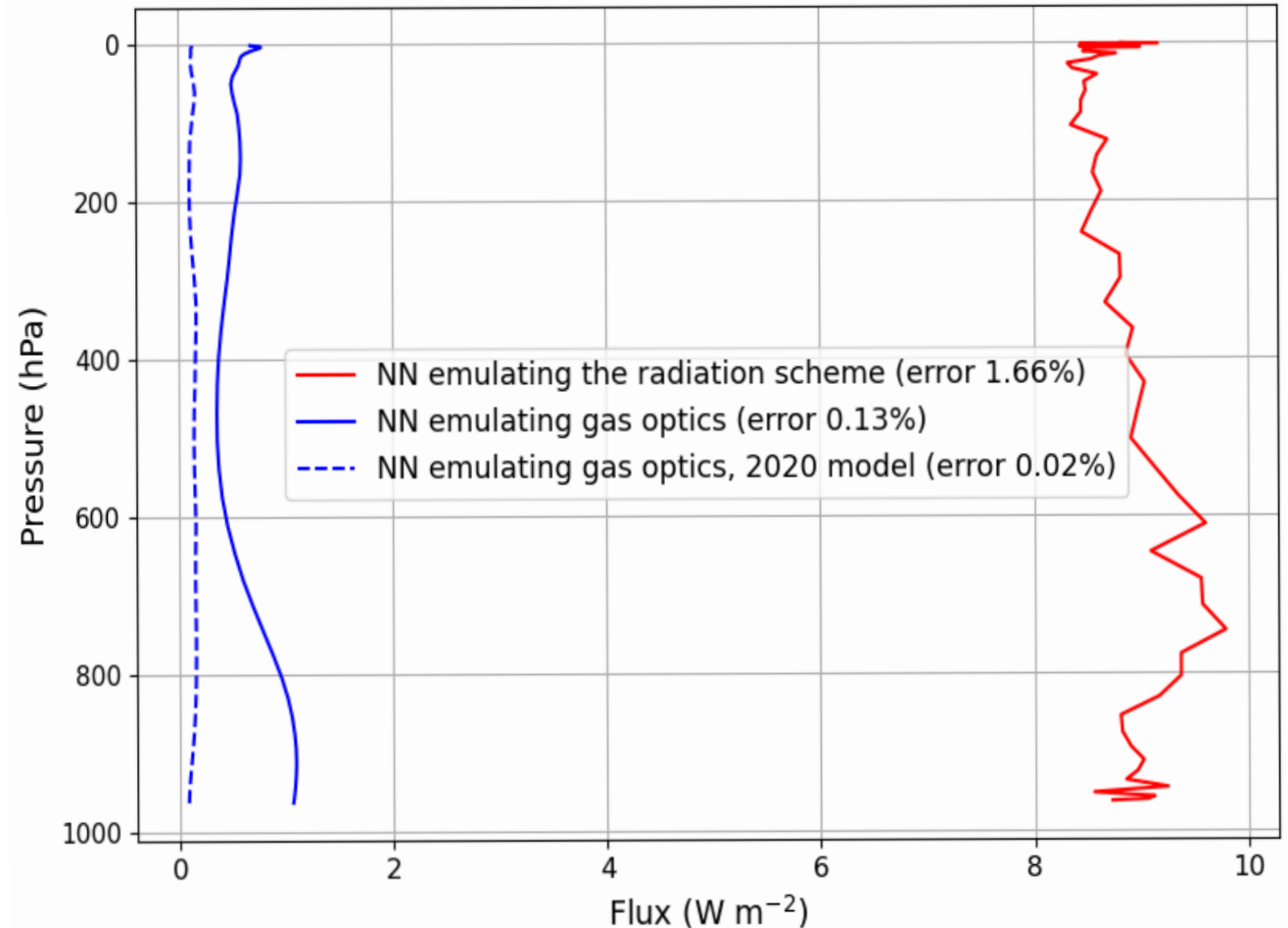
<sup>1</sup>Danish Meteorological Institute, Copenhagen, Denmark, <sup>2</sup>Niels Bohr Institute, University of Copenhagen, Copenhagen, Denmark, <sup>3</sup>Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, Boulder, CO, USA, <sup>4</sup>NOAA Physical Sciences Laboratory, Boulder, CO, USA, <sup>5</sup>European Centre for Medium-Range Weather Forecasts, Reading, UK

- RRTMGP-NN was included in the **ECMWF radiation scheme ecRAD**. In addition, CPU optimization work was performed on ecRAD (SPARTACUS solver, which is now **35-50% faster in single precision**)



- Emulating an entire radiation code leads to **>50x speedup** but comes at the cost of generalization and **accuracy**
- Focusing on a **subproblem** has the benefit of reduced dimensionality, making it **easier** to **cover the input space**

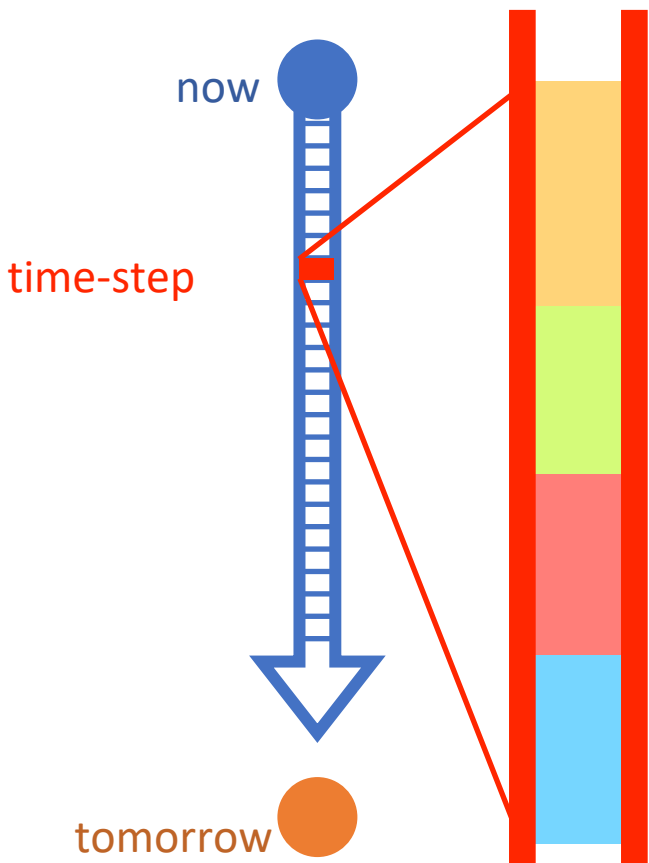
Mean absolute error in net shortwave flux, CAMS 2017





# Deliverables D1.7 and D1.8

Weather and climate dwarfs extracted from pre-ESCAPE2 models, complemented by some novel dwarfs



components: options:		ocean		atmosphere		global		regional		D1.7	D1.8	HPCW
discretisation	spectral transform*		✓	✓	✓	✓					✓	
	finite volume	✓	✓	✓	✓						✓	
	discontinuous Galerkin	✓	✓	✓	✓			✓		✓		
time-stepping	multigrid elliptic solver	✓	✓	✓	✓			✓				
	fault tolerant elliptic solver	✓	✓	✓	✓			✓				
	horizontal explicit, vertical implicit	✓	✓	✓	✓						✓	
advection	semi-Lagrangian		✓	✓	✓			✓			✓	
	MPDATA*	✓	✓	✓	✓			✓			✓	
	MUSCL	✓	✓	✓	✓			✓			✓	
physics	CLOUDSC microphysics*		✓	✓	✓			✓			✓	
	ecRad radiation		✓	✓	✓						✓	
	ACRANEB2 radiation*		✓	✓	✓			✓			✓	
	machine learned radiation		✓	✓	✓			✓		✓		

grey: work in progress, \*from ESCAPE 1

**work-steps for each dwarf:** isolation into self-contained prototype, documentation, adaptation to different hardware, maintenance of repo



# Deliverable D5.4 (for WP5)

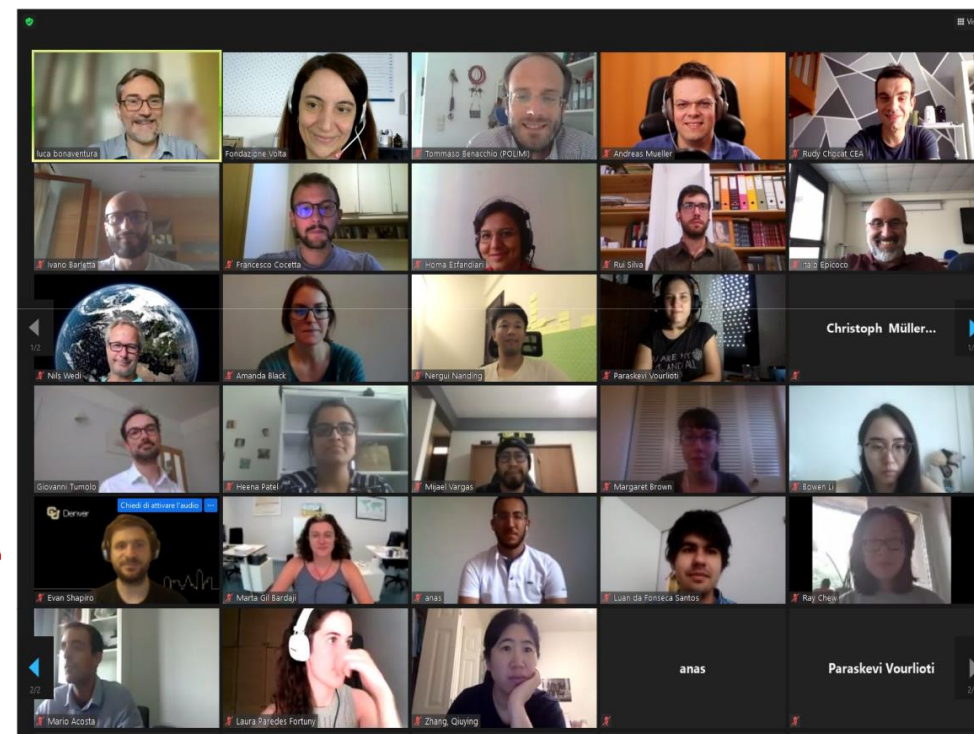
## Summer school



# Towards exascale computing for numerical weather prediction

Lake Como School of Advanced Studies, 19 - 23 July 2021

**39 participants from 18 countries with almost perfect gender balance**



ESCAPE 2

Funded by the European Union

Project work on Discontinuous Galerkin methods for Numerical Weather Prediction

Project 1: black box study

Rudy Chocat

Team members: Emilia Rizzi, Musa Ssemujju, Voula Vourlioti

Giovanni Tumolo  
giovanni.tumolo@ecmwf.eu  
European Centre for Medium-Range Weather Forecasts

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 800987



ECMWF ESCAPE 2 © ECMWF 21/07/2021

## Parallel Programming Project

Parallelization of Conjugate Gradient iterative method

Tutor: Italo Epicoco

Students: Amanda Black

Homa Esfandiari  
Nergui Nanding  
Rui Silva  
Qiuying Zhang

ESCAPE 2

Project 4 : Introduction to profiling tools and performance analysis

Project Coordinators: Daniel Beltrán Mora and Mario C. Acosta  
Project Participants: Heena, Daniele, Christos, Laura, Queen

## DSL Group Project Presentations

ESCAPE 2 Summer School

Tutors:  
Ben  
Christoph  
Matthias

Students:  
Anas, Francesco, Ivano, Nikita

ESCAPE 2