



EuroHPC
User Days

Dublin Ireland
22→24
September 2026



AI4Land: A Scalable Deep Learning Framework on MareNostrum5 for High-Resolution Land Surface Boundary Condition Reconstruction

Amirpasha Mozaffari, Marina Castaño, Stefano Materia, Etienne Tourigny, Oscar Molina-Sedano, Jordi Varela-Agrelo, Dario Garcia-Gasulla, Miguel Castrillo Melguizo, Mario Acosta, Amanda Duarte



**Barcelona
Supercomputing
Center**
Centro Nacional de Supercomputación



Amanda Duarte¹



Amirpasha
Mozaffari¹



Marina Castaño¹



Ferran Mirabent
Rubinat¹



Joan Vedri
Casino¹



Miguel Castrillo
Melguizo¹



Stefano Materia¹



Etienne Tourign¹



Junyan Ding¹



Dario Gracia
Gasulla²



Oscar
Molina-Sedano²



Jordi Varela
Agrelo²

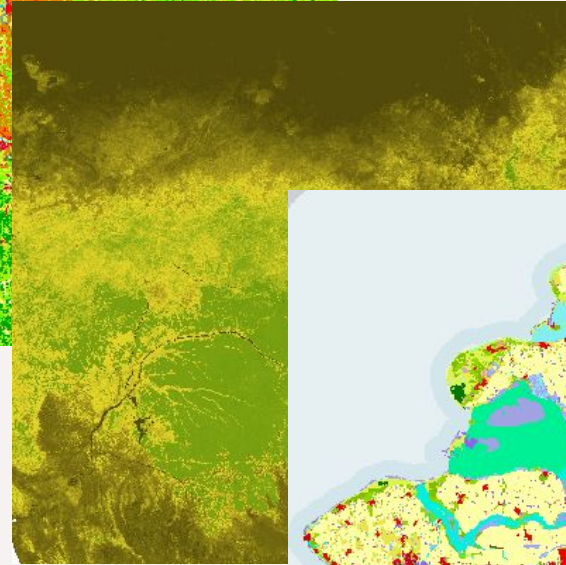
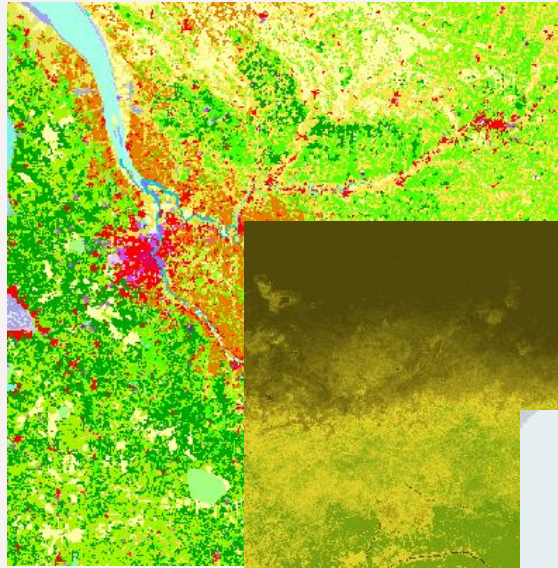
1: Earth Department, BSC

2: AI Institute, BSC

Motivation

Lacking consistent, **high-resolution datasets** for historical periods (pre-EO) and future projections.

We aim at improving the representation of **land surface parameters** to ensure more realistic fluxes of water, energy, and carbon towards the atmosphere.



Goals

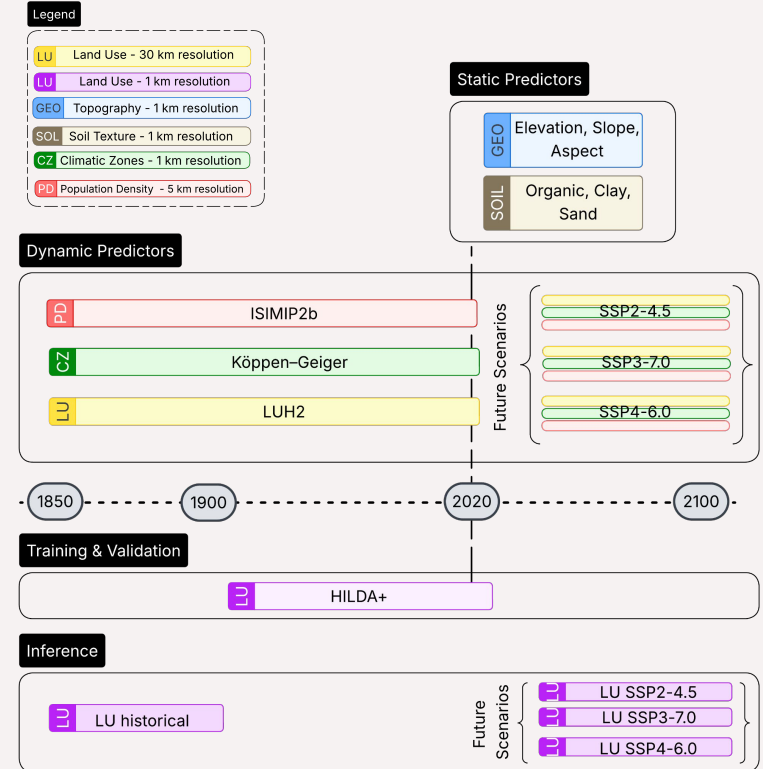


To produce:

- ✓ High-resolution global map of **land use** change (LU) covering the historical and the future scenario period (1850-2100)
- High-resolution dataset of annually-varying **land cover** (LC)
- High-resolution dataset of weekly-varying **leaf area index** (LAI)
- And couple the produced models into weather-climate frameworks to serve as **real-time land surface emulators** for Digital Twins of the Earth

Land Use

- **Dynamic Predictors:** coarse-resolution ~30km LU data (LUH2h), and high-resolution 1km data for population (ISIMIP2b) and climatic zones (Köppen-Geiger)
- **Static Predictors:** high-resolution 1km data for topography and soil (GEBCO, PNNL)
- **Training Target:** high-resolution 1km LU data (HILDA+)
- **Final Output:** 1km resolution LU dataset spanning 1850-2100



In collaboration with AI Factory, we already have a working prototype pipeline to generate a unified ARCO* Zarr store for LU reconstruction task.

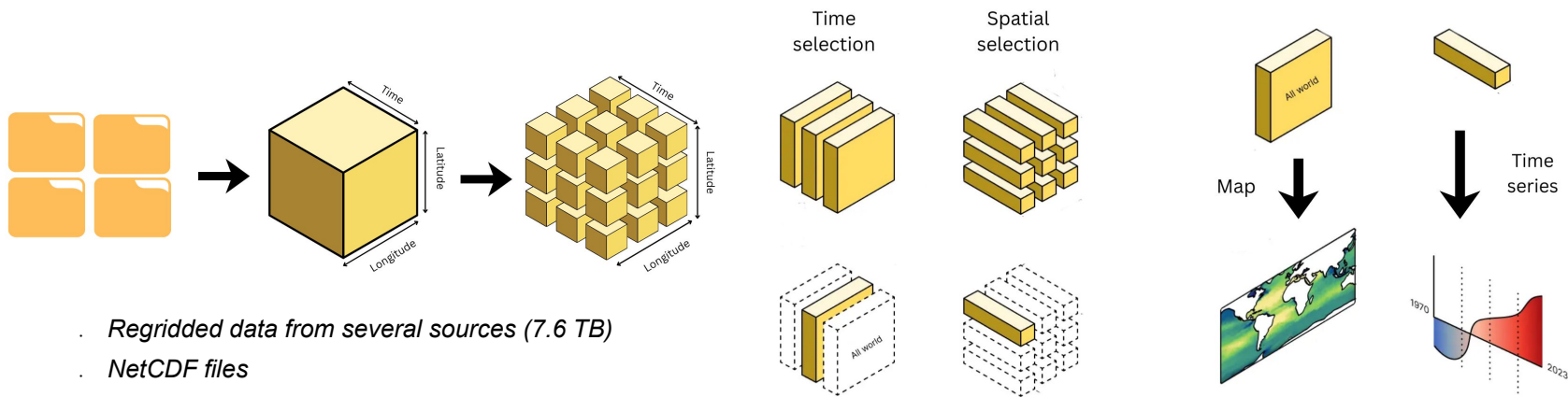
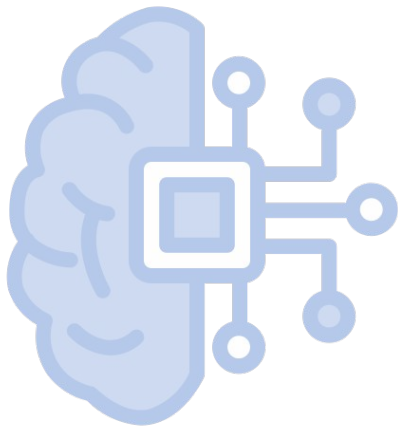


Image: [Lobelia.earth](https://lobelia.eearth.org)

* Analysis-Ready, Cloud-Optimized

Training

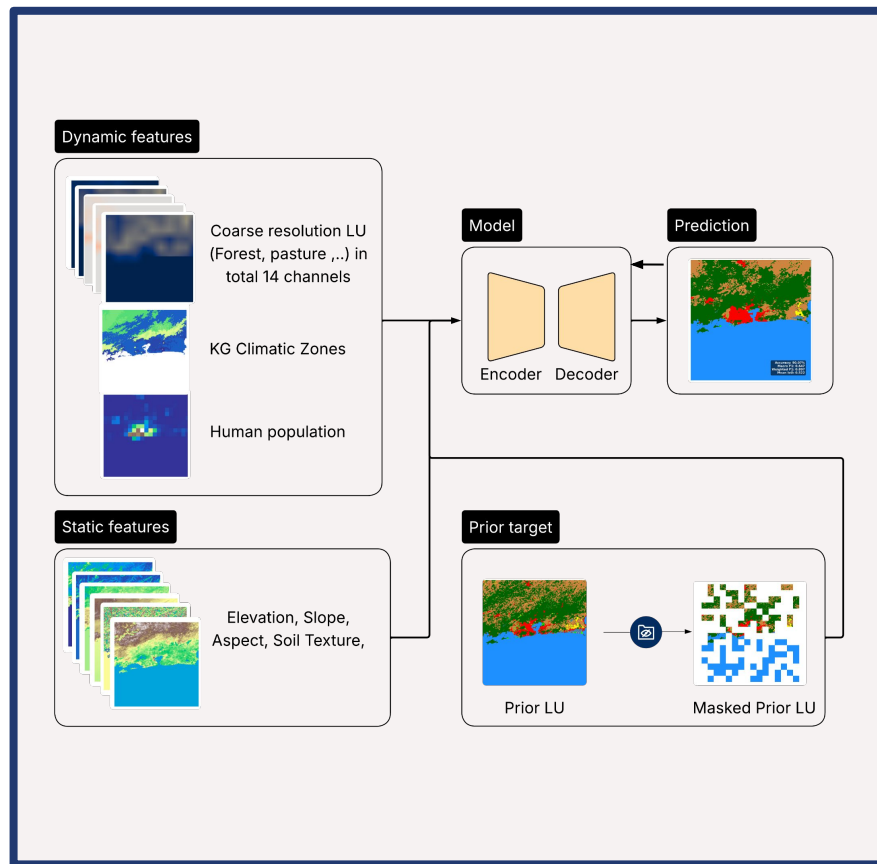


UNet: convolutional neural network suited for spatial downscaling

- > **Inputs:** 256x256-pixel images
- > **Optional embeddings:** For an optimal representation of categorical data (climatic zones, HILDA prior)
- > **Accelerate:** Distributed training using single-node multi-GPU or multi-node parallelization

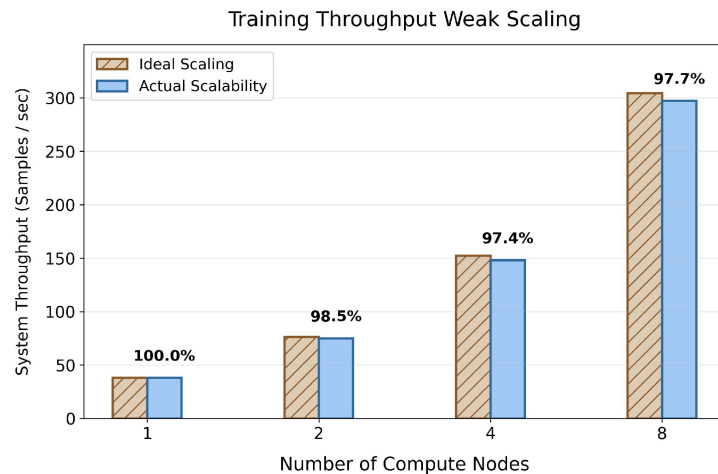
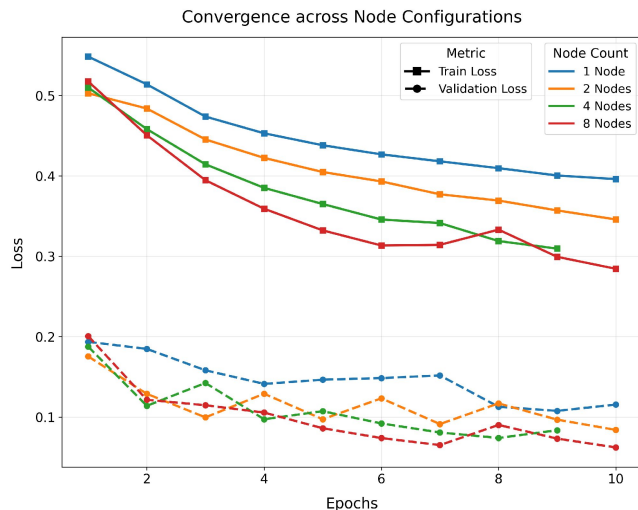
Model R-Unet

- **R-UNet** with 256 x 256 pixel images
- **Inputs**
 - Coarse-resolution Land Use
 - Elevation
 - Climatic zones
 - Population
 - Soil texture
 - Masked prior
- **Embeddings**
 - Categorical variables: climatic zones, prior
- **Target**
 - High-resolution Land Use



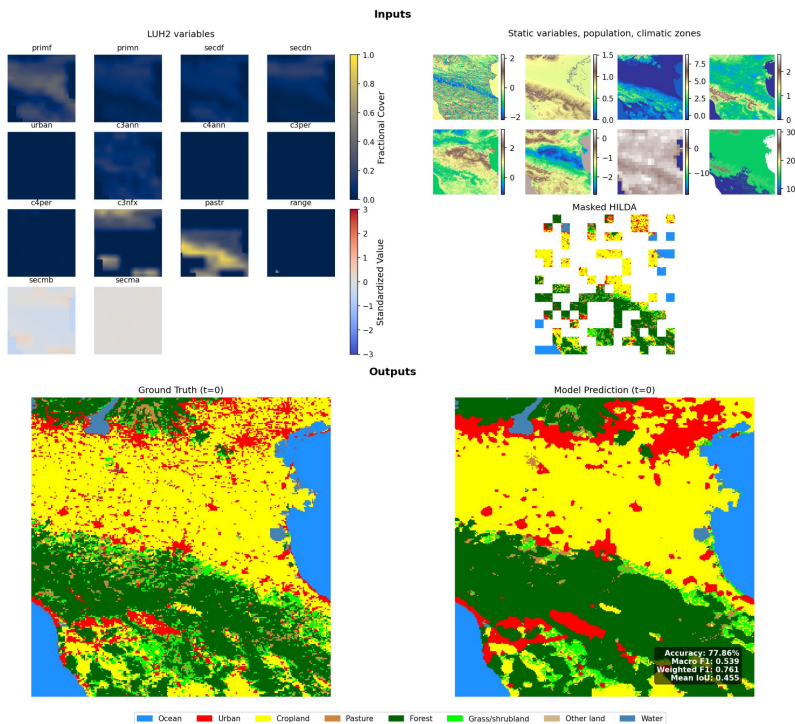
Scalability on MN5

- DDP training on MN5, scaled 1→8 nodes (4→32 H100 GPUs)
- Near-linear weak scaling
- Negligible inter-node communication overhead



Sample Output

Prediction for Bologna (Year 2014) | Model Epoch: 8



Quantitative Results

Metric	First Year		15-Year Average	
	Overall	LU Change	Overall	LU Change
Accuracy	0.8826	0.4796	0.8270	0.4513
Macro F1	0.8072	0.3182	0.7110	0.2661
Weighted F1	0.8802	0.4818	0.8156	0.4175
Mean IoU	0.6983	0.2089	0.5894	0.1748

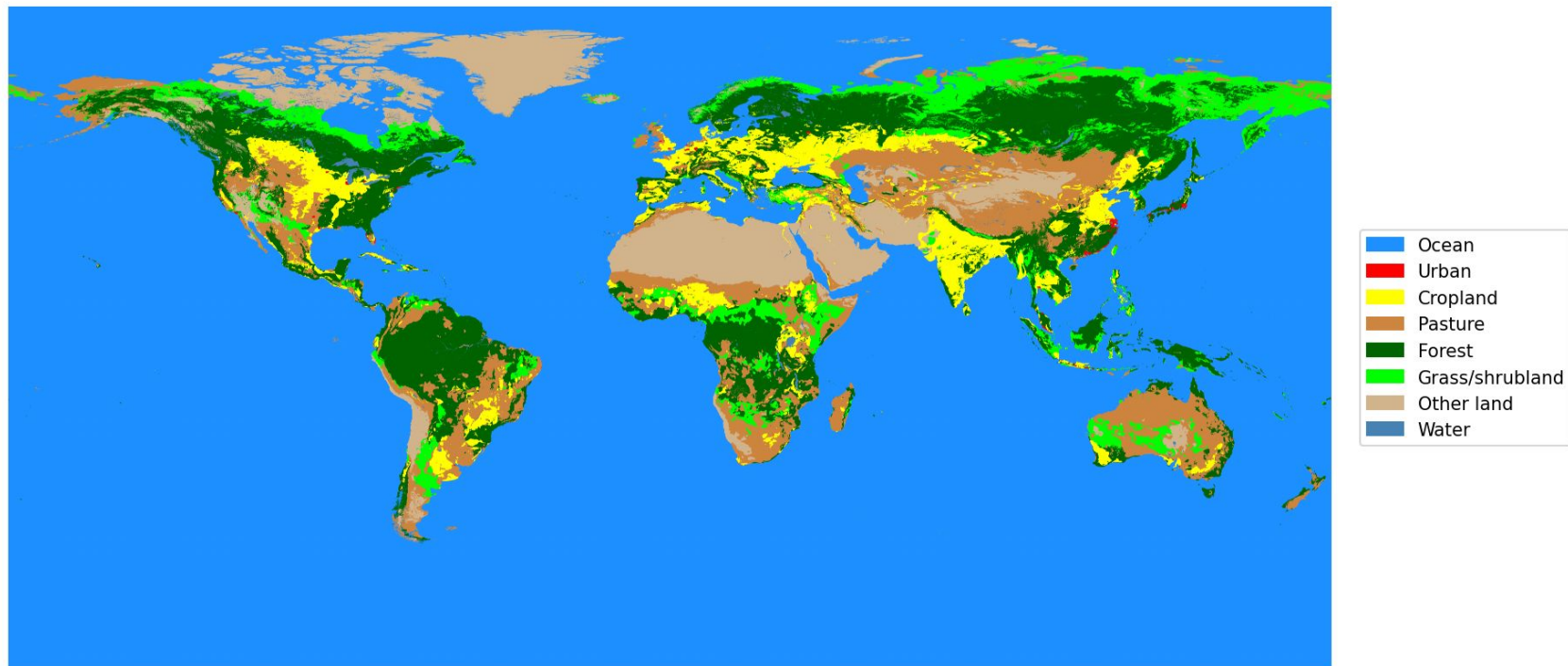
Inference



A distributed inference pipeline that processes the globe in overlapping 256x256 images

- **Patch Predictions:** The world is divided in "tiles" to cover it using 256x256 images
- **Overlapping Patches:** To avoid edge artifacts
- **Gaussian Window:** Prioritizes center pixels over edges during aggregation
- **Distributed Reduction:** Class probabilities are computed across multiple GPUs and then aggregated before taking the maximum probability

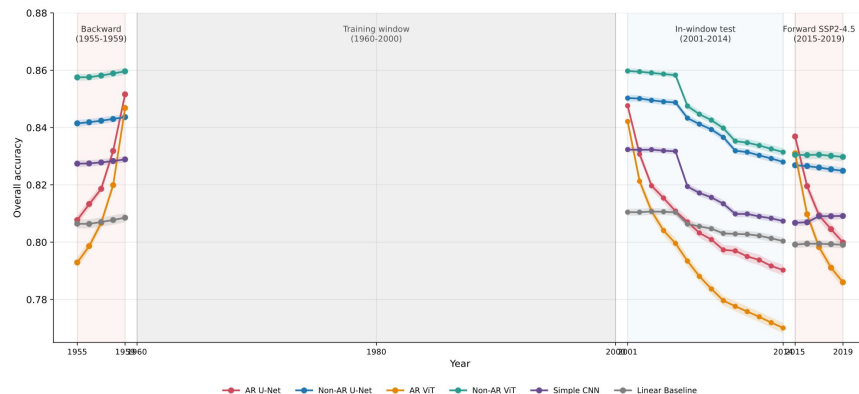
Output



TerraNostra - Bench

Model	Overall		Land-use change	
	Acc.	mIoU	Acc.	mIoU
Rolling persistence	0.9947	0.9890	0.0000	0.0000
Frozen persistence	0.9704	0.9419	0.1630	0.0883
Linear	0.8115	0.3994	0.4388	0.1904
Simple CNN	0.8302	0.4938	0.4594	0.2006
AR U-Net	0.8083	0.4216	0.5046	0.2385
non-AR U-Net	0.8464	0.5333	0.4815	0.2264
AR Swin-ViT	0.7968	0.4532	0.5030	0.2380
non-AR Swin-ViT	0.8607	0.5480	0.4143	0.1893

Model	Backward (1955–1959)		Forward (2015–2019)	
	Overall Acc.	Change Acc.	Overall Acc.	Change Acc.
AR U-Net	0.8246	0.3207	0.8140	0.4212
non-AR U-Net	0.8425	0.3559	0.8259	0.3923
AR Swin-ViT	0.8130	0.3010	0.8032	0.4452
non-AR Swin-ViT	0.8583	0.3312	0.8303	0.4131
Simple CNN	0.8280	0.2987	0.8082	0.4188
Linear	0.8072	0.3361	0.7993	0.3663

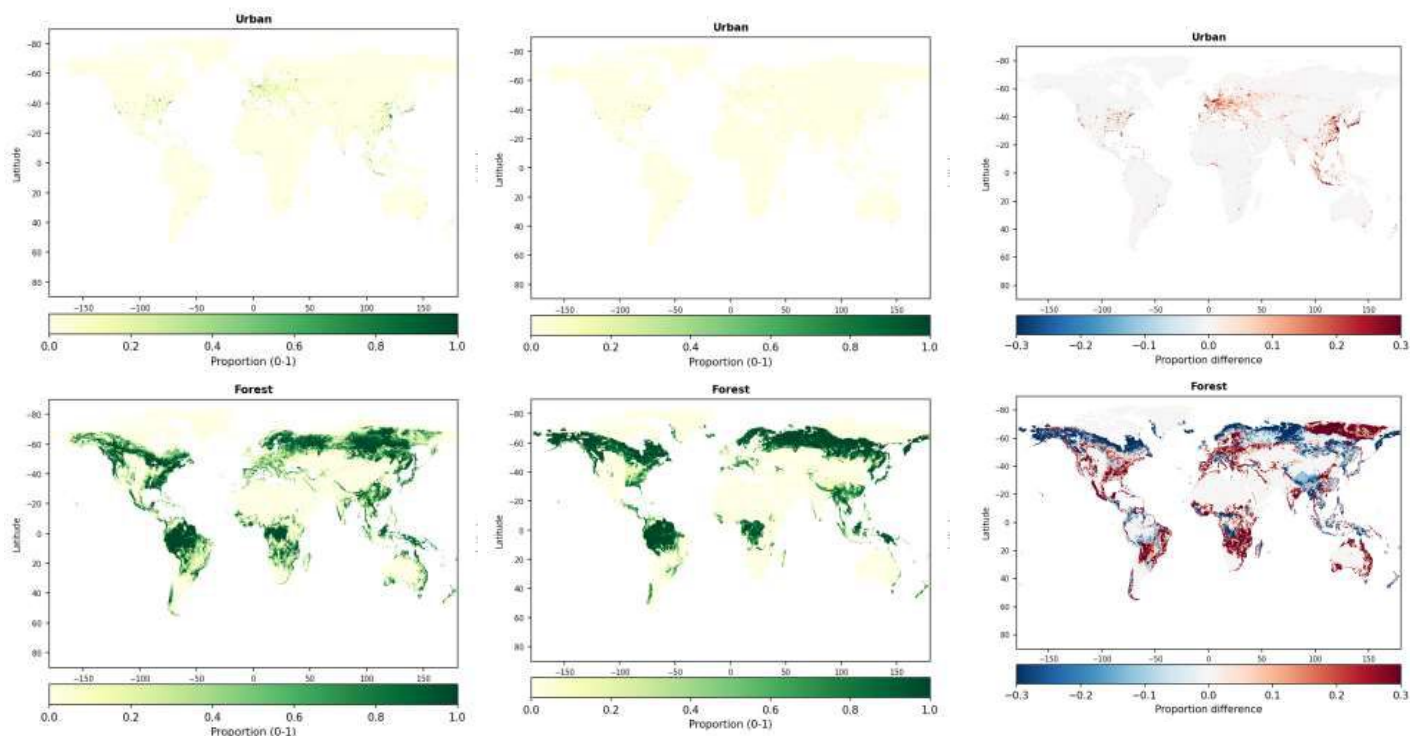


TerraNostra-Bench: Deep-Learning Models and a Benchmark Dataset for High-Resolution Land-Use Reconstruction and Projection , in prep.

Output

Comparison of upscaled LU data and LUH2 (0.25°), which is the current standard for ESMs' land use boundary conditions.

Fraction of urban (top) and forest (bottom) land in each gridbox



Summary and Future Steps

We have build a deep learning framework to generate high-resolution land surface boundary conditions for Land Use, Land Cover, Leaf Area Index.

Next:

- Publishing **TerraNostra - LU benchmark dataset**
- Complete **LC and LAI emulator and datasets**



AM acknowledges the Grant JDC2023-051208-I, funded by MICIU/AEI/10.13039/501100011033 and, as appropriate, by "ERDF A way of making Europe", by "ERDF/EU", by the "European Union", or by the "European Union NextGenerationEU through PRTR". AD acknowledges her AI4S fellowship within the "Generación D" initiative by Red.es, Ministerio para la Transformación Digital y de la Función Pública, for talent attraction (C005/24-ED CV1), funded by NextGenerationEU through PRTR. This work received funding from the European Union's Horizon Europe Framework Programme through the project CONCERTO (Grant Agreement 101185000). This work, as part of TerraDT -- Digital Twin of Earth System for Cryosphere, Land Surface and related interactions, has received funding from the European Union's Horizon Europe Framework Programme (HORIZON) under Grant Agreement no.\ 101187992. This work was supported by the European Union's Horizon Europe research and innovation programme under grant agreement No 10214398 (ELLIOT).

Backup slides



2025 — 2028

Historical
reconstruction & future
prediction of LU/LC &
LAI 1850–2100 for
weather and climate



2025 — 2028

Historical
reconstruction, future
prediction & emulator
of LU/LC & LAI for
Digital Twins

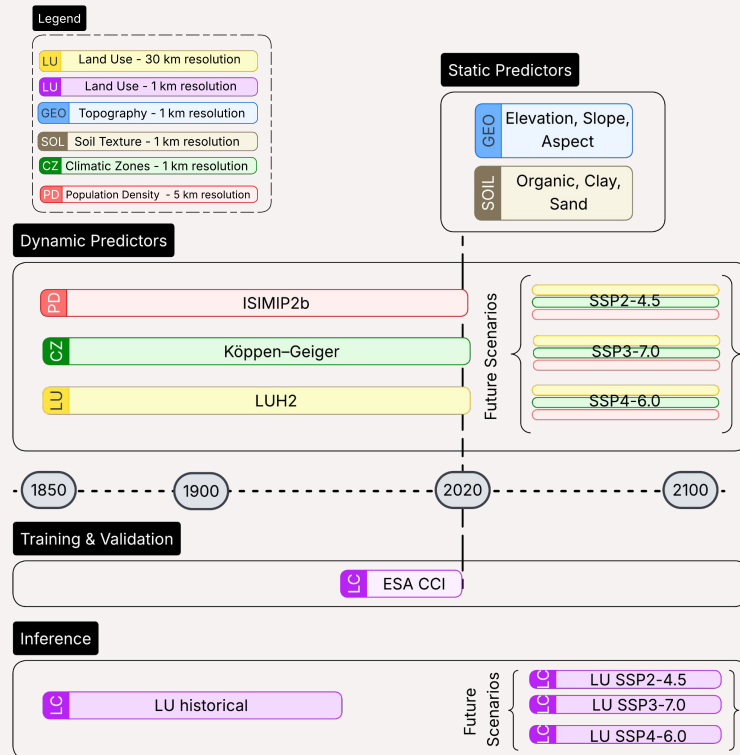


2025 — 2029

Fine-tuning of
foundation model for
LU/LC & LAI emulation

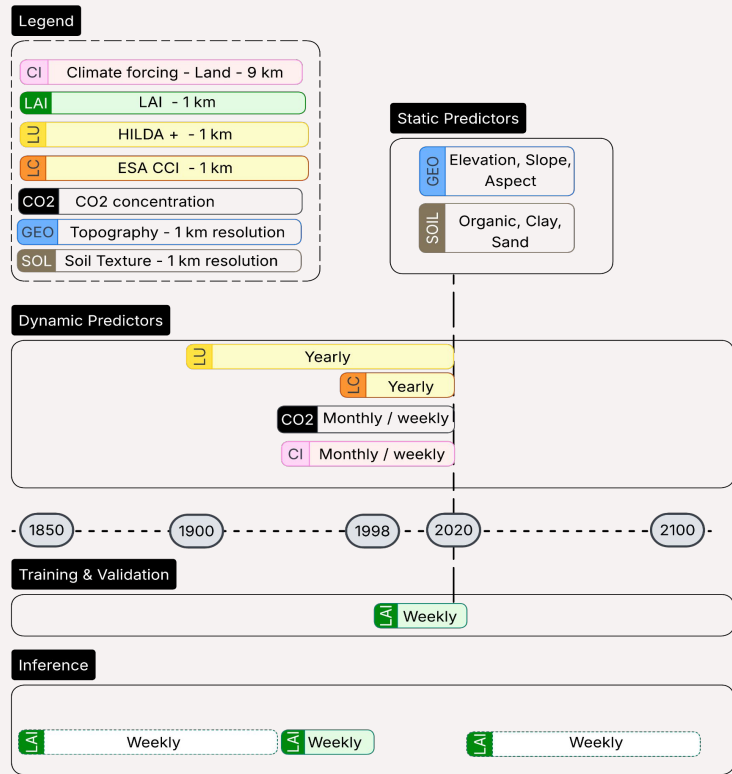
Land Cover

- **Dynamic Predictors:** coarse-resolution ~30km LU data (LUH2h), and high-resolution 1km data for population (ISIMIP2b) and climatic zones (Köppen-Geiger)
- **Static Predictors:** high-resolution 1km data for topography and soil (GEBCO, PNNL)
- **Training Target:** high-resolution 1km LC data (ESA CCI)
- **Final Output:** 1km resolution Land Cover dataset spanning 1850-2100



Leaf Area Index

- **Dynamic Predictors:** high-resolution 9km climate forcing (ERA5-Land)
 - To be included: CO₂, LC
- **Static Predictors:** high-resolution 1km data for topography and soil (GEBCO, PNNL)
- **Training Target:** high-resolution 1km LAI data (Copernicus Land or ESA-CCI)
 - Tests are being performed using the LAI produced for the project CONFESS LAI
- **Final Output:** 1km resolution LAI dataset spanning 1850-2100





EuroHPC
User Days

Dublin Ireland
22→24
September 2026



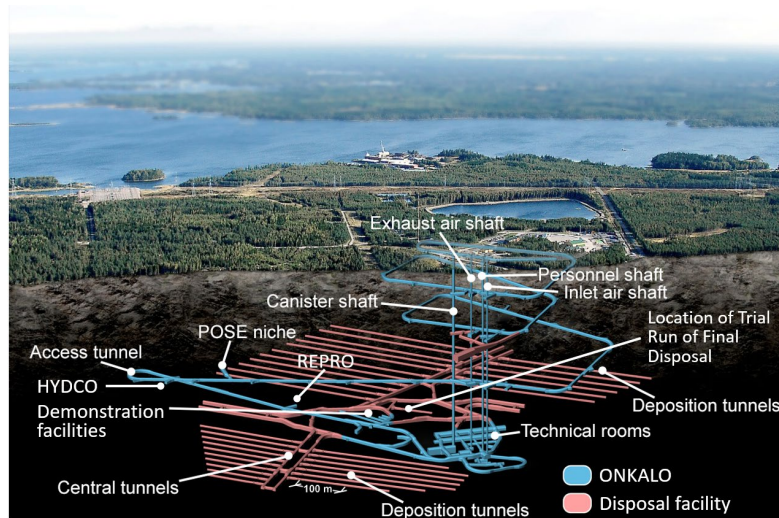
High-Performance Computing for subsurface flow and transport modelling of Olkiluoto

Albert Costa-Solé, Emilie Coene, Gisela Martí, Olga Riba, Orlando Silva, Tiina Lamminmäki, Lasse Koskinen, Antti Poteri, Antti Joutsen, Veli-Matti Pulkkanen, Tuomo Karvonen, Albin Nordström, Guido Deissmann

AMPHOS²¹ **Posiva**
an RSK company

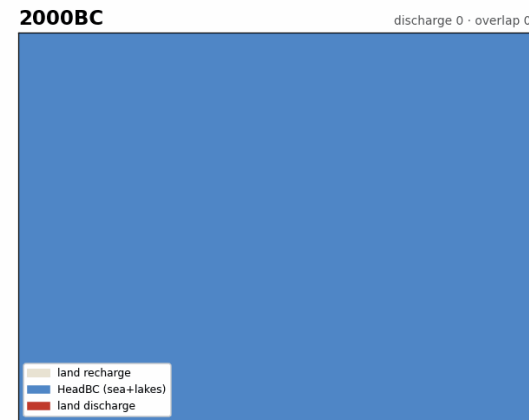
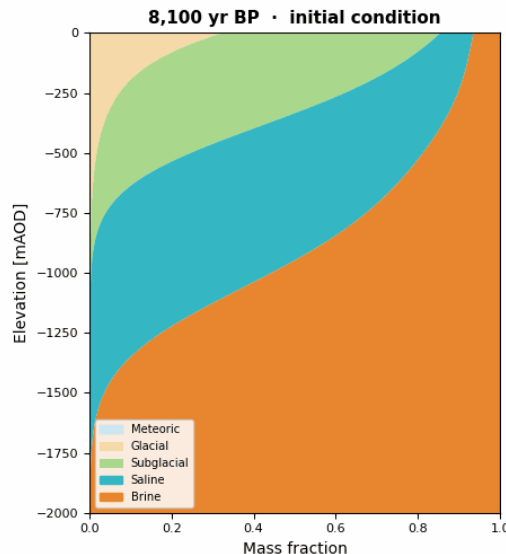
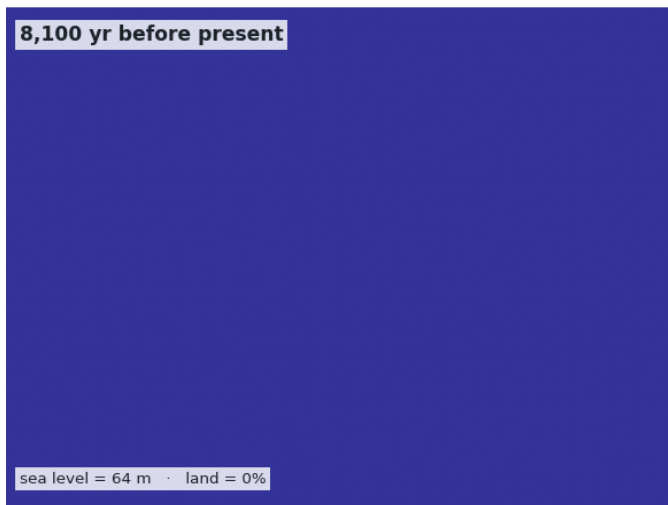


Olkiluoto and the ONKALO® repository



ONKALO® is the world's first deep repository for spent nuclear fuel.

Paleoevolution. 8100 years reconstruction



Post-glacial rebound lifted the site out of the Baltic Sea.

Sea level fell, the shoreline retreated, the island emerged.

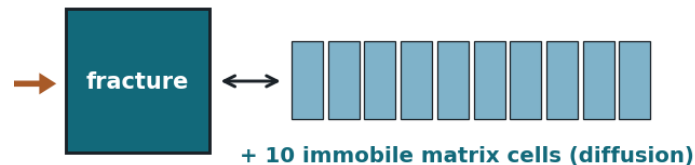
Meteoric recharge began a freshwater lens grew over deep brine.

Numerical model



Cl ⁻	Br ⁻	SO ₄ ²⁻	HCO ₃ ⁻	Na ⁺	Ca ²⁺
K ⁺	Mg ²⁺	H ₄ SiO ₄	H ⁺	Fe	+ tracer

DUAL CONTINUUM (our choice)



- ✓ captures matrix diffusion & retention
 - ✓ reproduces the observed matrix pore-water
 - ✓ physically faithful
- cost: x11 cells per location → memory-bound

Matrix exchange

diffusion into
the rock matrix

Anion exclusion

charged mineral
surfaces

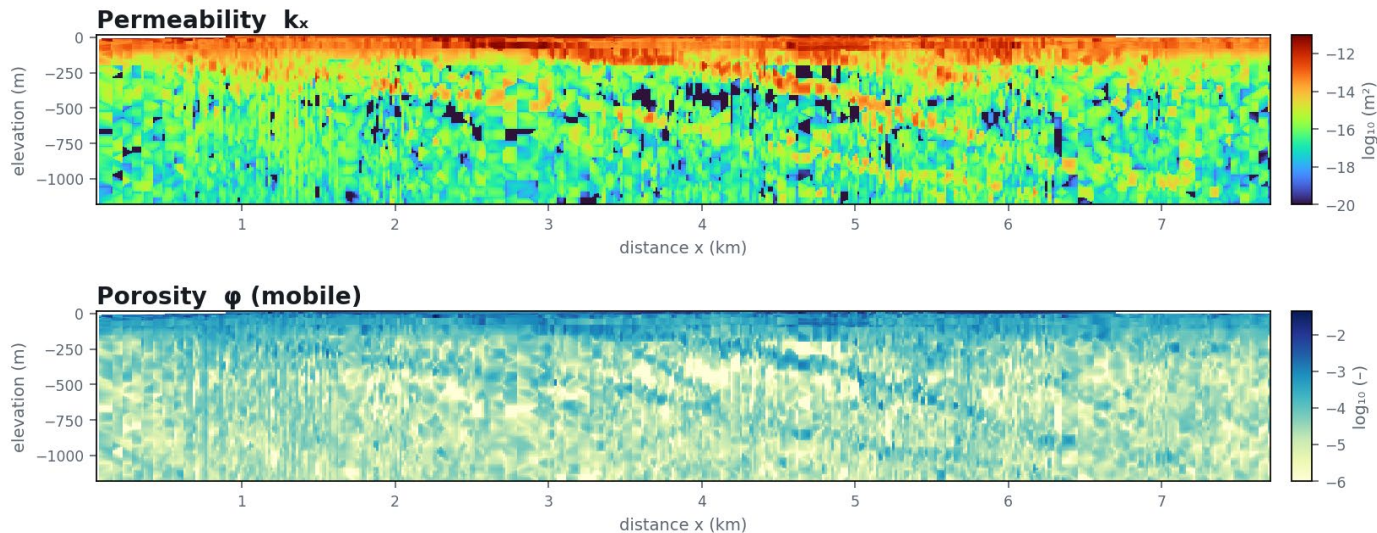
Mineral reactions

dissolution ·
precipitation · pH

Density-driven flow

ρ depends
on salinity

Numerical model



Moreover, highly heterogeneous materials fields

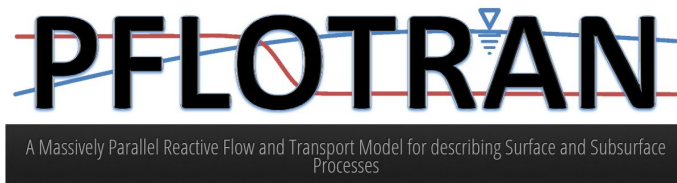
Numerical model

$$\begin{array}{|c|} \hline \mathbf{2.9\ M} \\ \hline \text{fracture cells} \\ \hline \end{array} \times \begin{array}{|c|} \hline \mathbf{\times\ 10} \\ \hline \text{matrix cells} \\ \hline \end{array} \times \begin{array}{|c|} \hline \mathbf{15} \\ \hline \text{species} \\ \hline \end{array} = \begin{array}{|c|} \hline \mathbf{\approx\ 50\ M} \\ \hline \text{degrees of freedom} \\ \hline \end{array}$$

Need for HPC



Numerical model



open source, massively parallel, for subsurface flow

Flow related capabilities:

General, Richard's,
Density Driven Flow

Reactive transport capabilities:

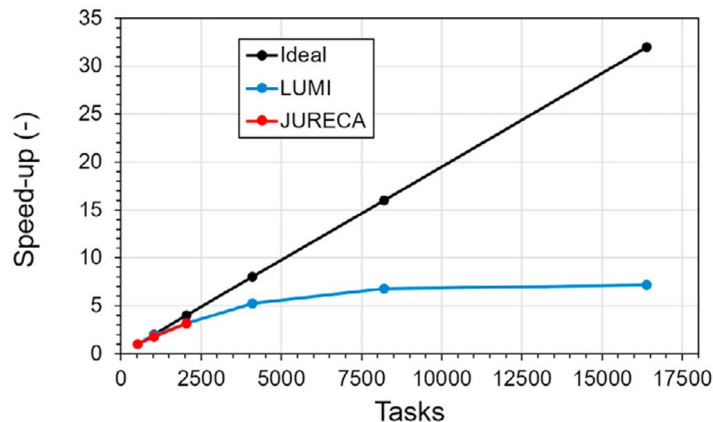
Multiple continuum,
Nernst-Plank

Other capabilities:

Structured and unstructured meshes
Implicit temporal scheme

Lichtner, P. C., Hammond, G. E., Lu, C., Karra, S., Bisht, G., Andre, B., ... & Kumar, J. (2015). PFLOTRAN user manual: A massively parallel reactive flow and transport model for describing surface and subsurface processes (No. LA-UR--15-20403).

HPC. PFLOTRAN parallelization



Nodes	Tasks	LUMI	JURECA	DOFs	Efficiency
4 nodes	512	✓	✓	97 656	1.00
8 nodes	1 024	✓	✓	48 828	0.94
16 nodes	2 048	✓	✓	24 414	0.79
32 nodes	4 096	✓	×	12 207	0.65
64 nodes	8 192	✓	×	6 104	0.42
256 nodes	16 384	✓	×	3 052	0.22

4,096 tasks / 32 nodes
the production configuration

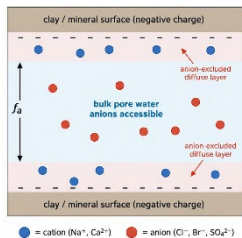
≈ 65 %
efficiency

128 tasks/node is the optimal configuration spot.

Beyond 4,096, communication dominates.

More nodes finish sooner but waste core hours.

Calibration is required

**PFLOTRAN run**

24–48 h on LUMI-C

Parameters $\phi_m \cdot \text{CEC} \cdot f_a$ **Simulated**

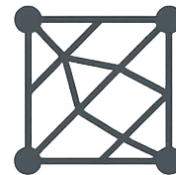
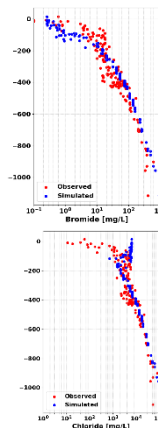
concentrations

*one loop =
one simulation***Optuna**

proposes next params

Compare

observed vs simulated



PYDELLING

A preprocessor for modelling



OPTUNA

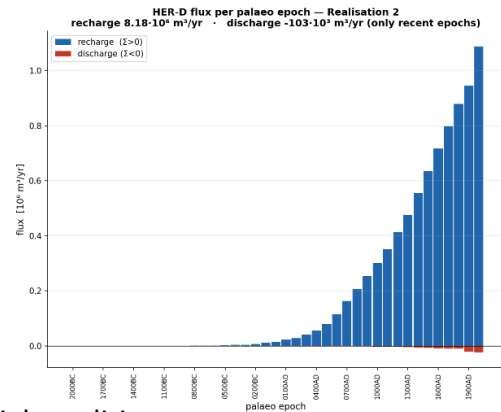
1 run = ~24 h (2M mesh).

Calibration & sensitivity = HUNDREDS of runs.

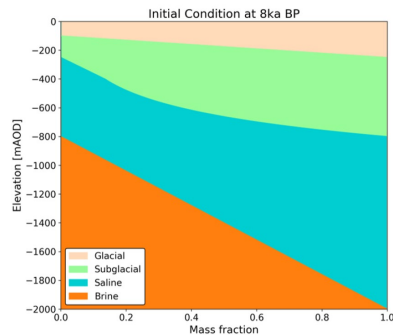
The allocation lets us close the loop enough times to trust the model.

Results

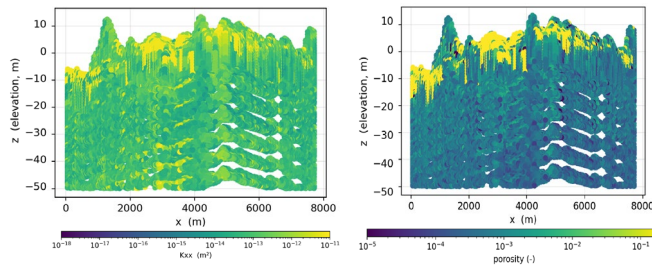
Recharge condition



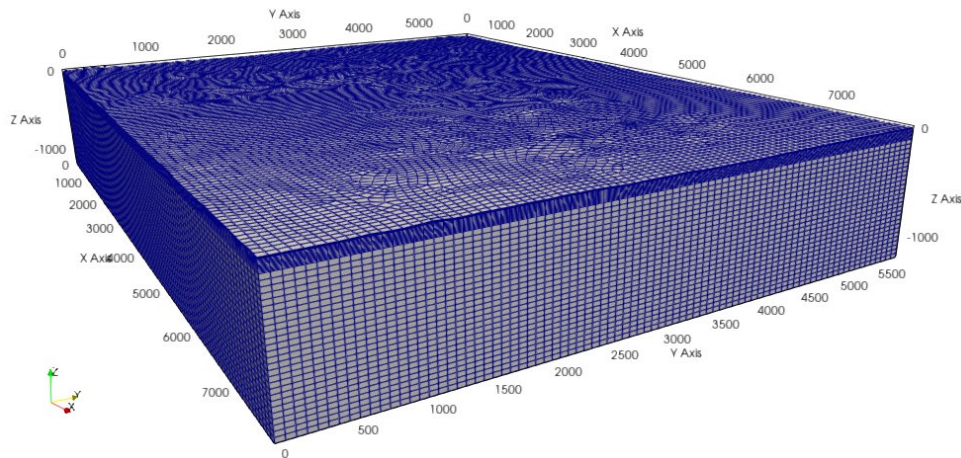
Initial condition



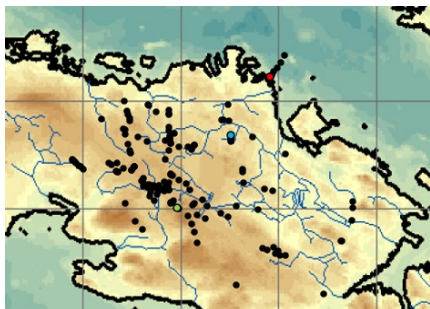
Material structures



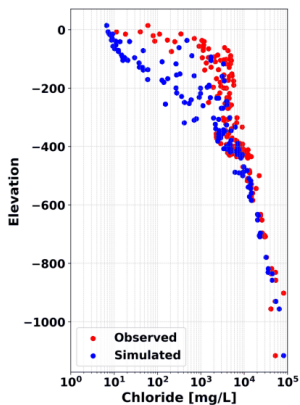
Mesh (560.793 cells)



Results



- Other BH
- ONK-PH2_1204
- OL-KR46_82_0908
- OL-KR47_111_0708
- OL-KR47_116.8_0309_2

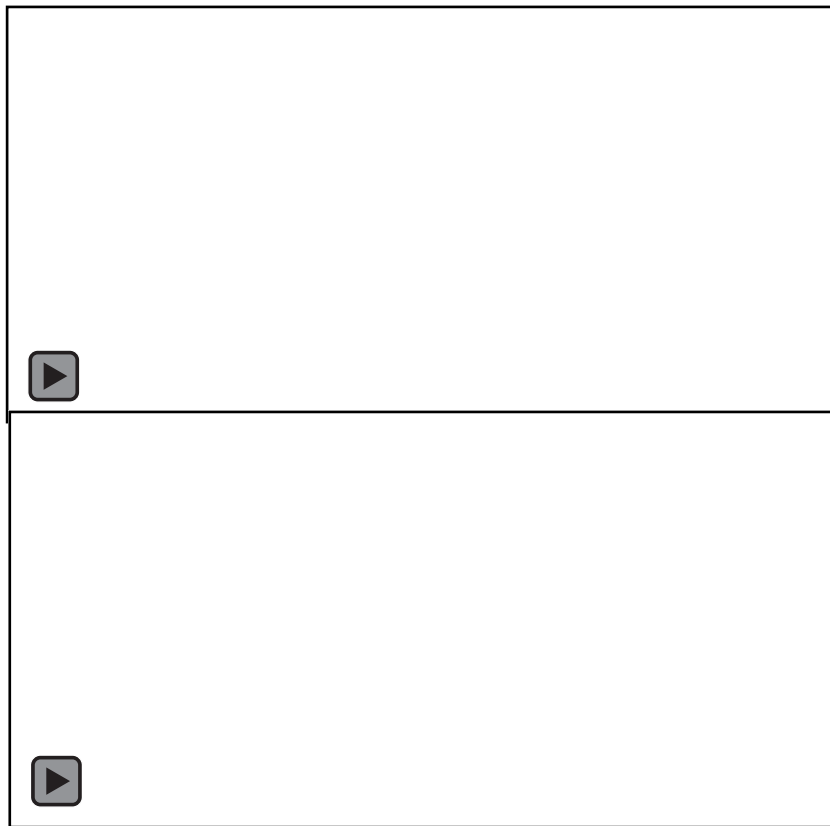


OBSERVED vs SIMULATED

Chloride

(anion)

- observed
- simulated

final conservative model
matrix deactivated in soil
Realisation 2 - 8,100 y

Reduced case: 560.793 cells. ~100 hours for the whole simulation (8100 years)

Conclusions

≈50 M

DOFs

4,096

MPI ranks

24–48 h

per run

50–70×

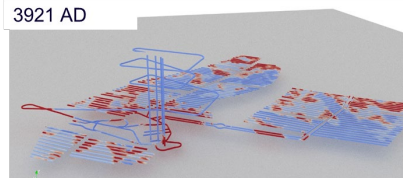
algorithmic

100s

of runs

- Strong physical model with promising computational capabilities
- HPC enables the throughput required for testing, calibration, and model improvement
- Further optimisation is needed to improve efficiency and simulation quality
- **Next step:** maximise scalability, performance, and scientific value
- ... and Reactive transport (mineral dissolution-precipitation, biotic processes) implies even more unknowns!
- ... include the nuclear repository facility

3921 AD



Acknowledgments

Supercomputing resources:

- EuroHPC JU for awarding the project ID:
EHPC-REG-2025R02-204, EHPC-REG-2022R03-225, EHPC-EXT-2025E02-059
- JURECA at Forschungszentrum Jülich (Germany) under grant no. JIEK63.

Project is funded by Posiva Oy

Thank you



EuroHPC
User Days

Dublin Ireland
22→24
September 2026



WeatherGenerator applications

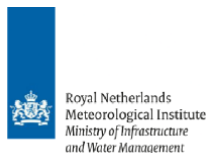
*Using a foundation model for weather
and climate applications*

Cristian Lussana (MET Norway)
on behalf of the WeatherGenerator team





Weather Generator



What is the WeatherGenerator?

Data

- ☐ Reanalysis
- ☐ Satellites
- ☐ Radar
- ☐ In-situ observations
- ☐ ...

Model

Adapting the technology
behind LLMs to Earth
System Science

Applications



Where we are now

O(100 TB)

700 million

22 applications

What we are aiming for

O(1-10 PB)

100 billion

A generic tool for all

What is the Weather generator?

This presentation

Data

- ☐ Reanalysis
- ☐ Satellites
- ☐ Radar
- ☐ In-situ observations
- ☐ ...

Model

Adapting the technology
behind LLMs to Earth
System Science

Applications



Where we are now

O(100 TB)

700 million

22 applications

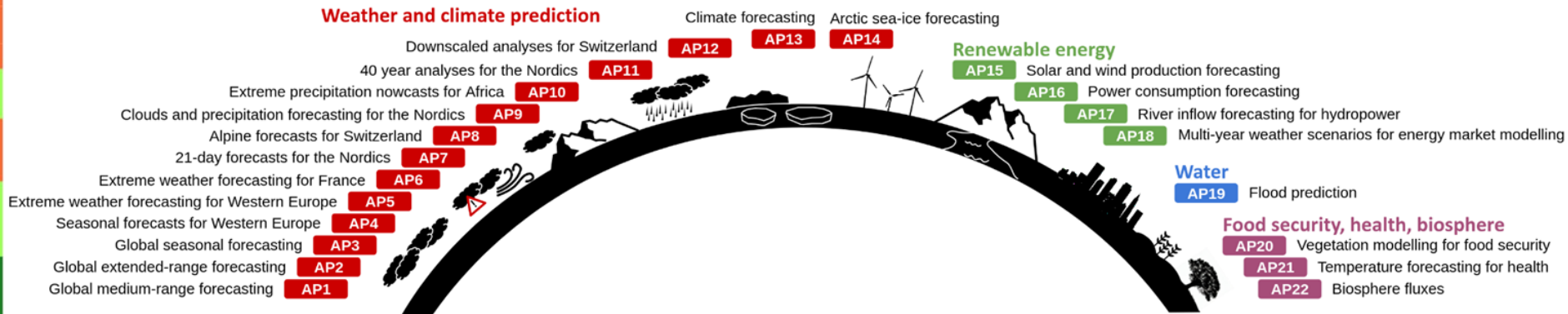
What we are aiming for

O(1-10 PB)

100 billion

A generic tool for all

Our applications



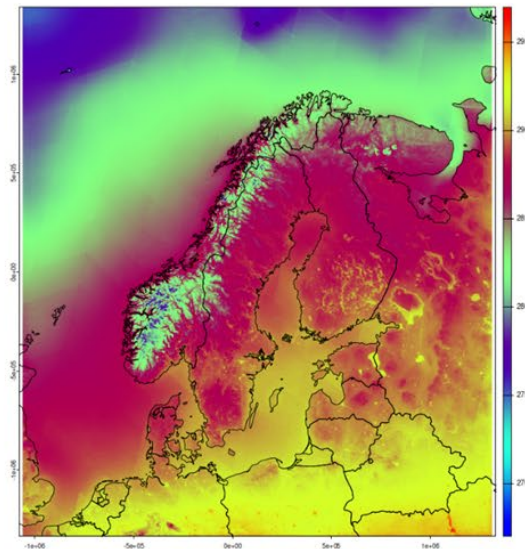
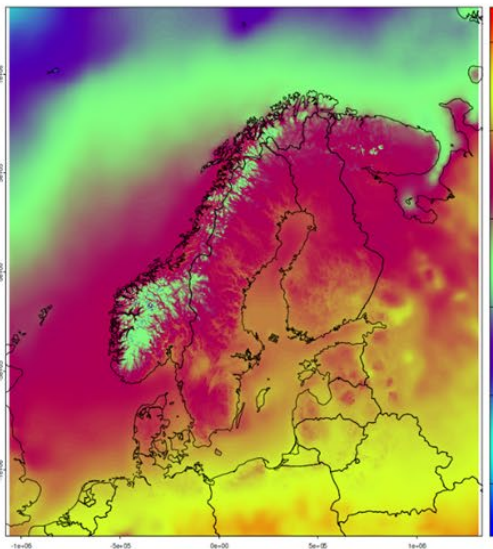
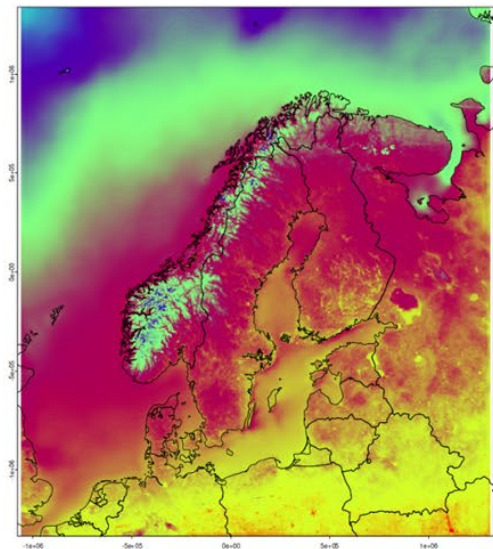
Downscaling with

WeatherGenerator

TARGET (MEPS)

BASELINE (ERA5*)

WG downscaling

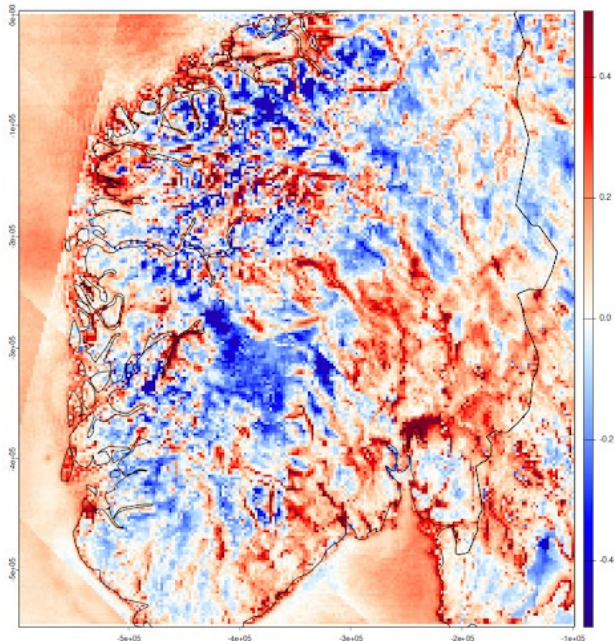


(*) ERA5 31km elevation adjustment

Downscaling with

WeatherGenerator

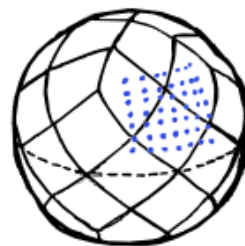
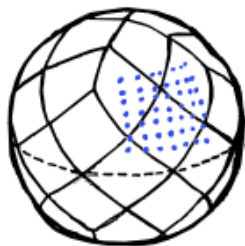
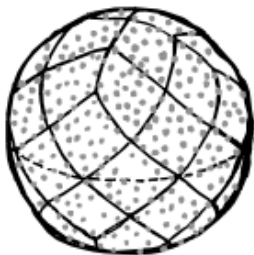
Richer geographical info consistently improves reconstruction skill



Model	Bias (°C)	RMSE (°C)
Baseline	0.30	1.215
WeatherGenerator	-0.21	1.197
+ elev, land mask	-0.057	1.074
+ urban, nature, land cover, ...	-0.013	1.027

Project progress

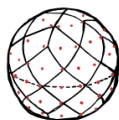
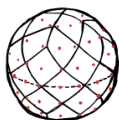
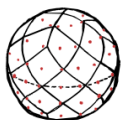
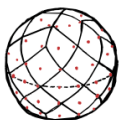
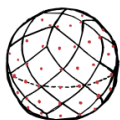
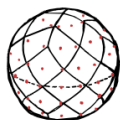
**Global
Reanalysis**



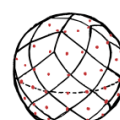
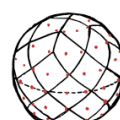
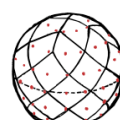
**In-situ
Observations**



The input data are compressed into the latent space...



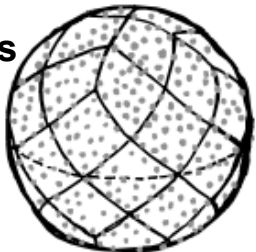
...



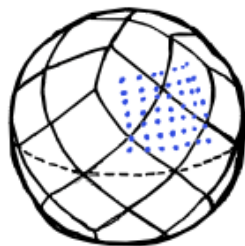
**Latent
Space**

... Then, decoded into the original data streams

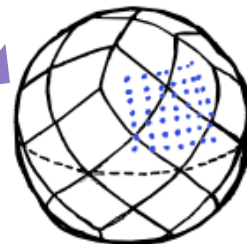
**Global
Reanalysis**



**Regional
Reanalysis**

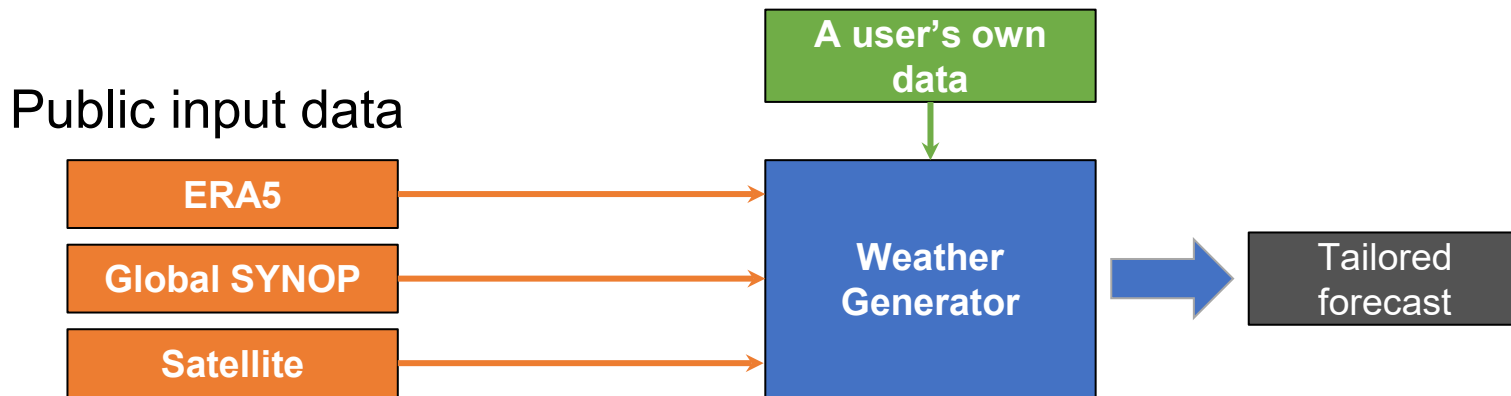


**In-situ
Observations**



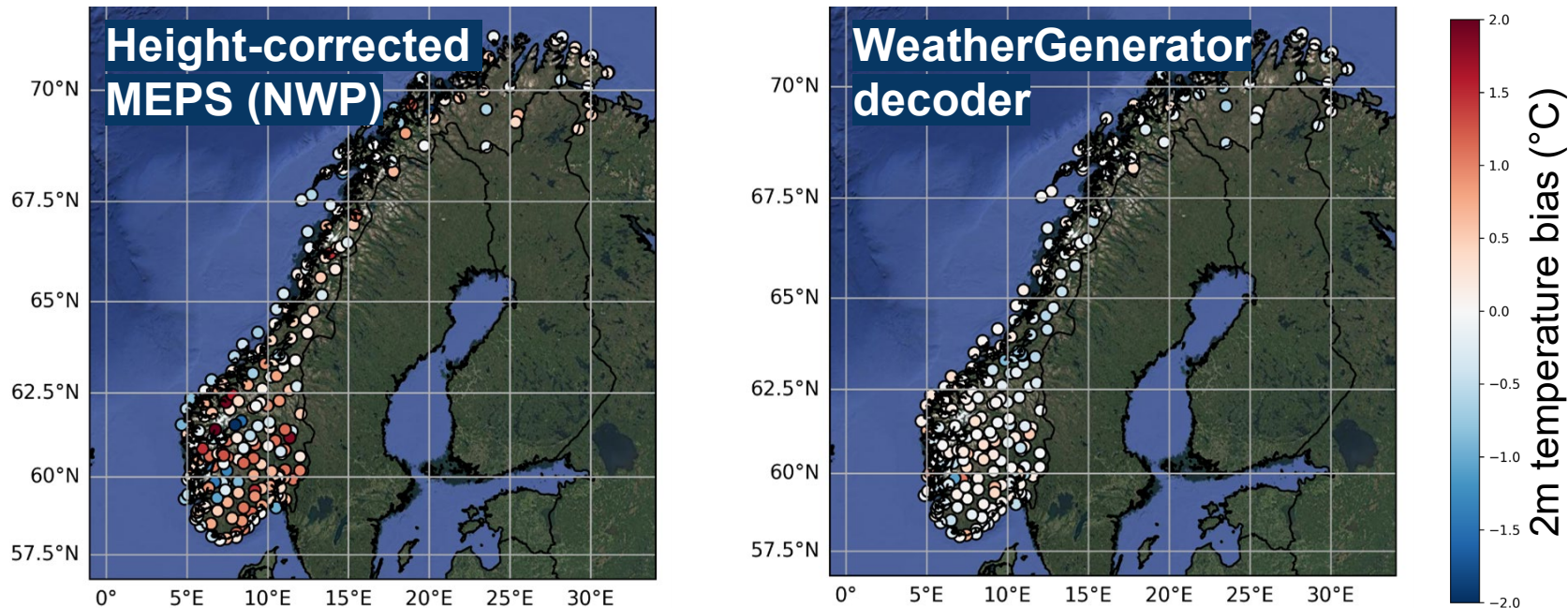
Fine-tuning on local data

- Many users (e.g. businesses) want to exploit their own proprietary domain-specific data
- The WeatherGenerator offers fine-tuning functionality for



Fine-tuning on local data

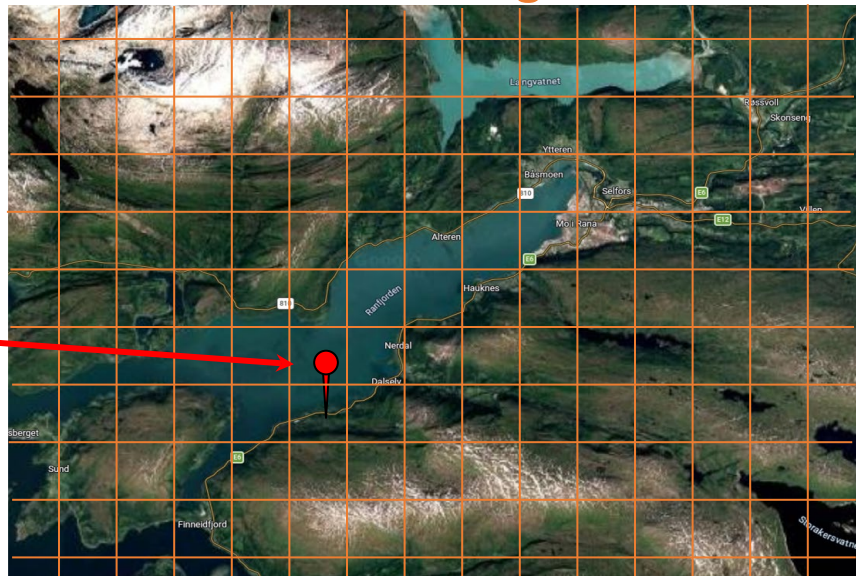
Biases greatly reduced at each station location



Fine-tuning on local data

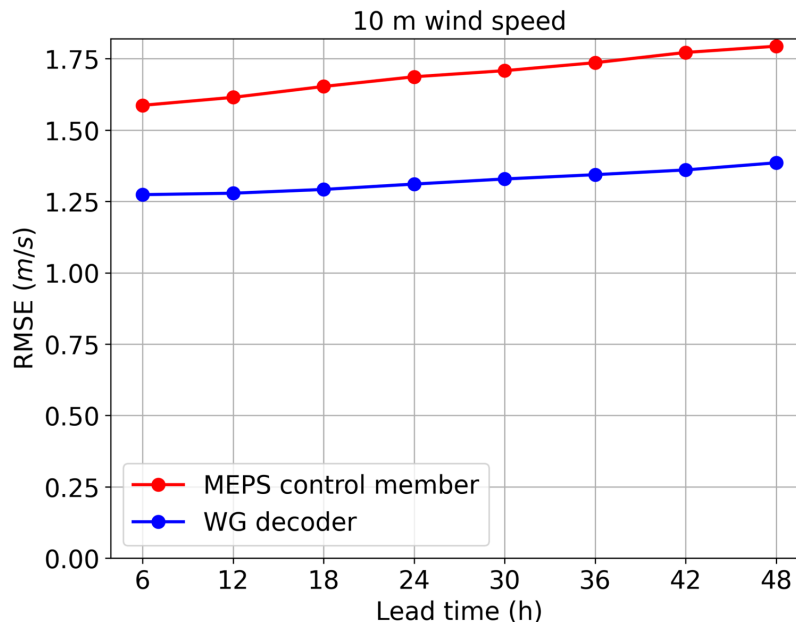
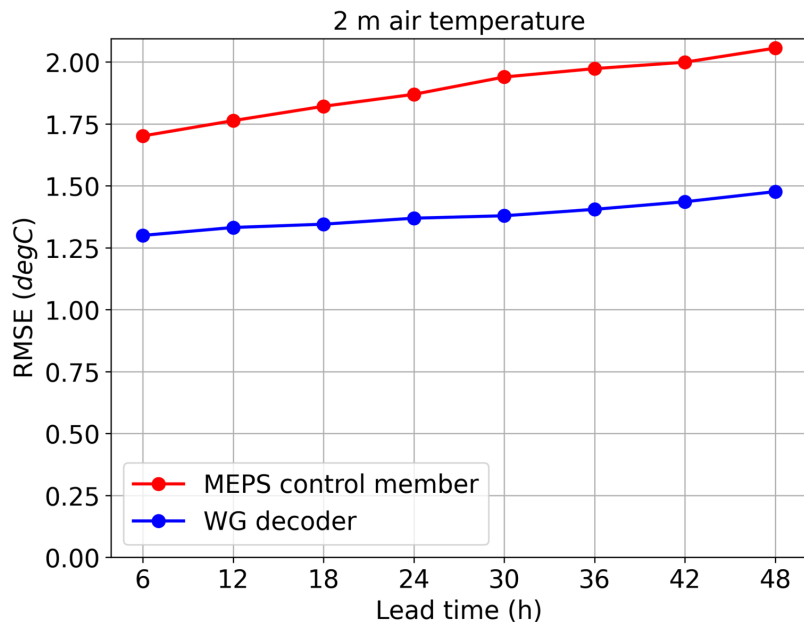
Biases greatly reduced at each station location

NWP grid resolution



Fine-tuning on local data

Overall forecast accuracy increases for multiple parameters



EuroHPC experiences

- AI Factory allocation on Leonardo
- An excellent way to enable 16 partner organizations to collaborate on a project like the WeatherGenerator
 - Large amount of shared data
 - Running many experiments in parallel
 - Rapidly iterating towards a solution

Final words

Visit weathergenerator.eu for more information:

- Blogs on the science behind the WeatherGenerator
- Applications described in detail
- Sign up for public hackathons

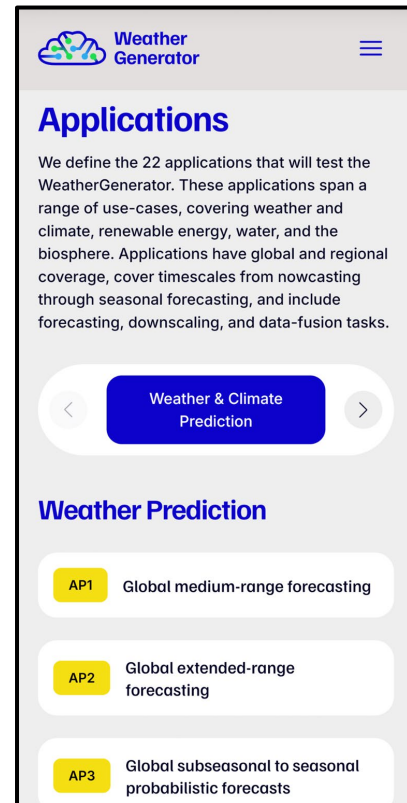
Weather
GeneratorFunded by
the European Union

Coordinated by



The WeatherGenerator project (grant agreement No101187947) is funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or EREA. Neither the European Union nor the granting authority can be held responsible for them.

We acknowledge EuroHPC JU for awarding the project ID EHPC-AIF-2025SC03-154 access to Leonardo hosted by CINECA. We acknowledge EuroHPC JU for awarding the project ID EHPC-AIF-2026SC02-173 access to JUPITER Booster module hosted by Jülich Supercomputing Centre.





EuroHPC
User Days

Dublin Ireland
22→24
September 2026



Mixed-Precision GPU Acceleration of RTE-RRTMGP-CPP

Stijn Heldens¹, Alessio Sclocco¹, Gijs van den Oord¹, Ben
van Werkhoven², Chiel van Heerwaarden³, Erwan Raffin⁴,
Xavier Yepes-Arbós⁵

¹Netherlands eScience Center,

²Leiden University,

³Wageningen University & Research,

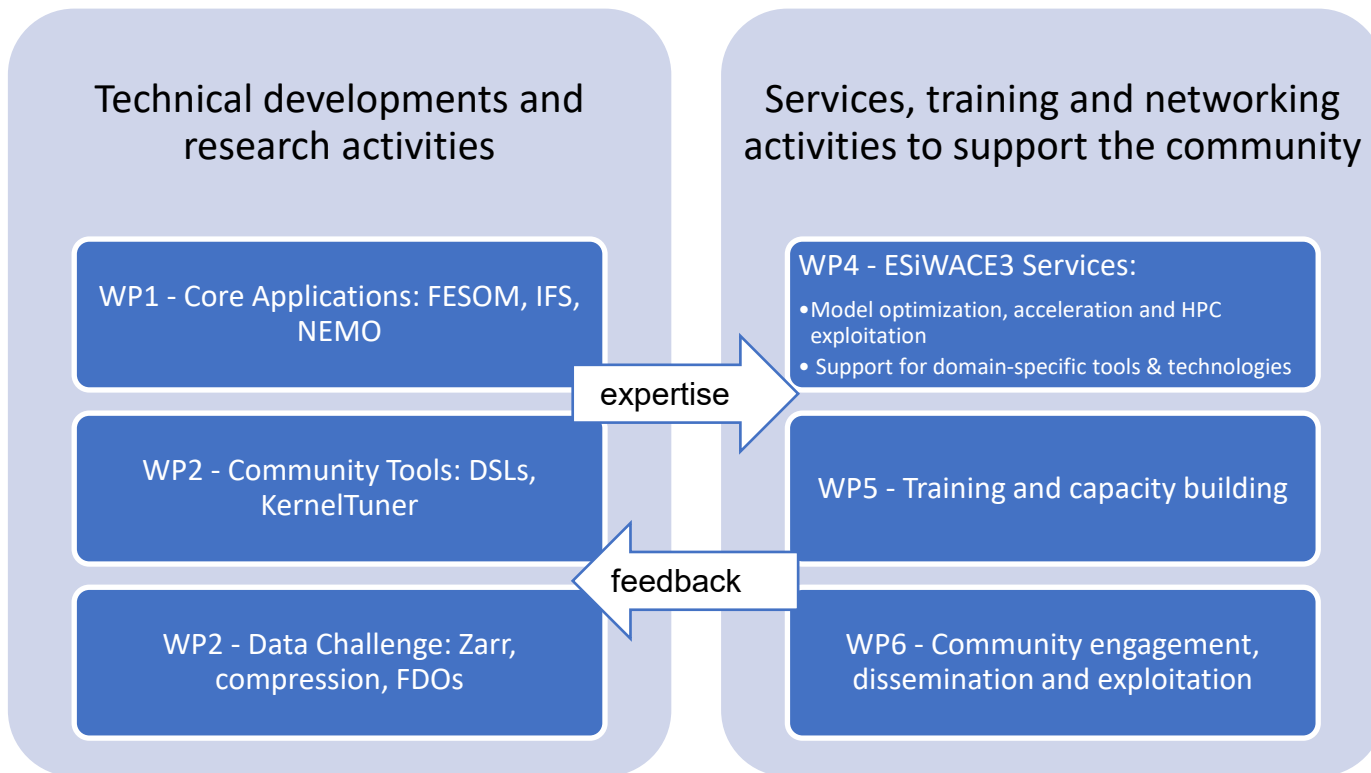
⁴Center for Excellence in Performance Programming (CEPP),

⁵Barcelona Supercomputing Center

The European Center of Excellence for HPC in Weather & Climate



ESiWACE Project Pillars



ESiWACE3 Services to support domain-specific tools & technologies

- 2 person months of effort from ESiWACE experts
- Open call for requests from European earth system modeling community
- Projects to support adoption of ESiWACE3 tools

 **Kernel Tuner****AutoRFE****Automatic Performance
Profiling (APP)**

RTE+RRTMGP-C++

github.com/microhh/rte-rrtmgp-cpp

- C++ library for simulation of **atmospheric radiation** in ESM
- Port of original **Fortran** widely-used code, adds supports for **GPUs** using **CUDA**
- Very computationally expensive!

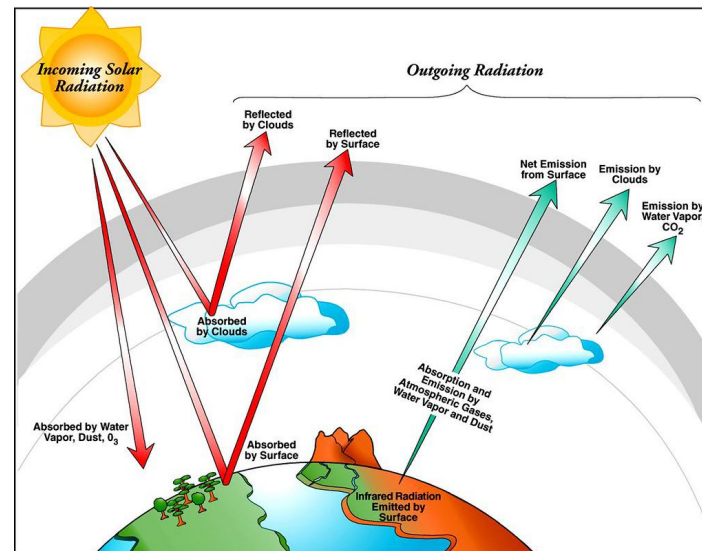
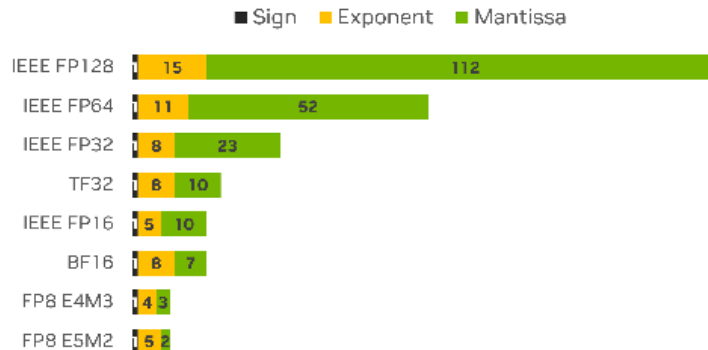


Image courtesy of the Department of Energy Office of Science

Floating-point Precision

- RTE **currently** supports
 - double (64 bit) or float (32 bit)
 - New GPUs support **low-precision**:
 - 16 bit, 8 bit, 6 bit, 4 bit...
 - **Higher** performance
 - **Loss** in accuracy is unacceptable
- Idea: Could we **mix** precision levels?



GPU	FP64	FP32	FP16	FP8	FP4
A100	9.7	19.5	312	-	-
H100	34	67	989	1979	-
B100	30	60	1750	3500	7000
Rubin	33	130	4000	17500	35000

Implementation

- Identify **four** most expensive GPU kernels of RTE
 - Total 68.8% of run-time
- Make their **data types** compile-time configurable.
- **Select** optimal data types
 - Trade-off: runtime versus error

```
template <
    typename float_type,
    typename tau_type,
    typename source_type,
    typename surface_type,
    typename flux_type,
    typename compute_type,
    typename intermediate_type
> __global__ void lw_solver_kernel(
```

GPU Tools

- **Kernel Float**

- **Header-only library** used for writing **portable** mixed-precision GPU code
- One codebase, **supported** across NVIDIA/AMD, data types, vector sizes, ...



Kernel Float

github.com/KernelTuner/kernel_float

- **Kernel Tuner**

- **Automatic performance tuning** for GPU code
- Searches data types, block size, vector size, ...
- Finds optimal configuration within **error threshold**



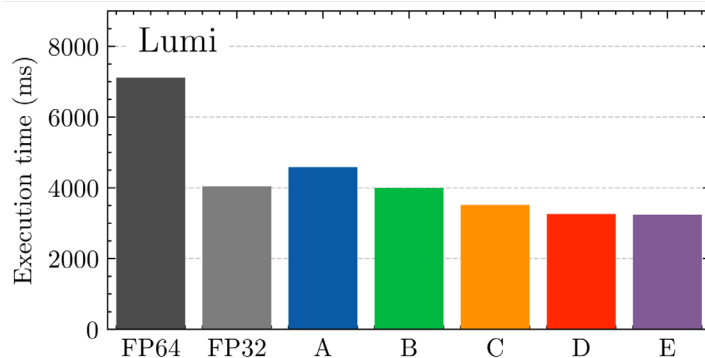
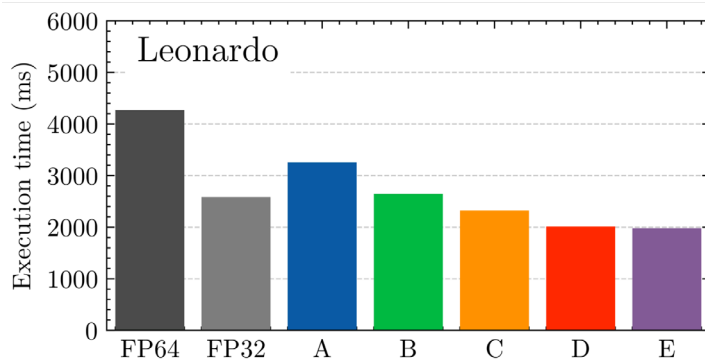
Kernel Tuner

github.com/KernelTuner/kernel_tuner

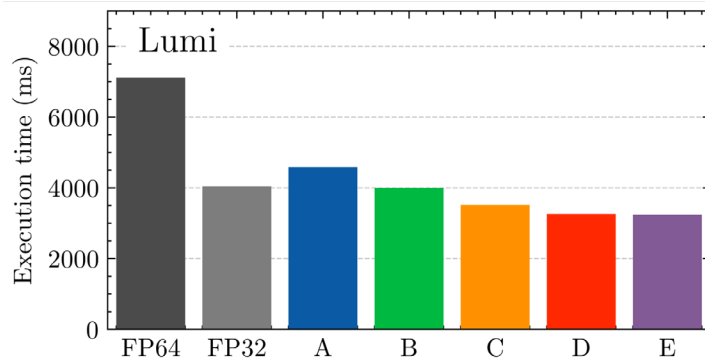
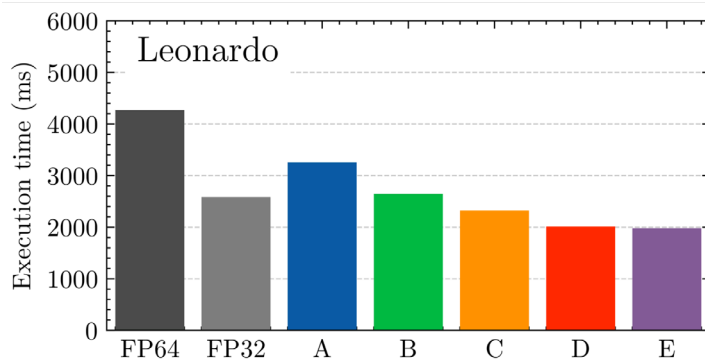
Experimental setup

- Tuned for **five** error thresholds:
 - A=0.00001%, B=0.001%, C=0.1%, D=1%, E=10%
 - Kernels tuned individually, then integrated into code
- Benchmarks on **two** EuroHPC systems:
 - Leonardo: **NVIDIA** A100
 - LUMI: **AMD** MI250X (*ported to **HIP***)
- **Dataset:** 384 x 384 at 100m resolution

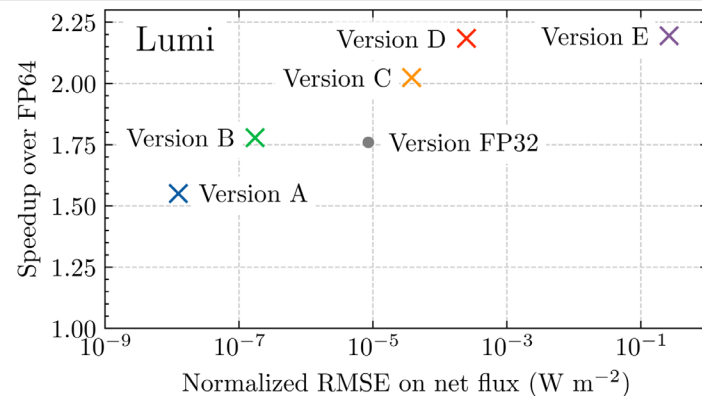
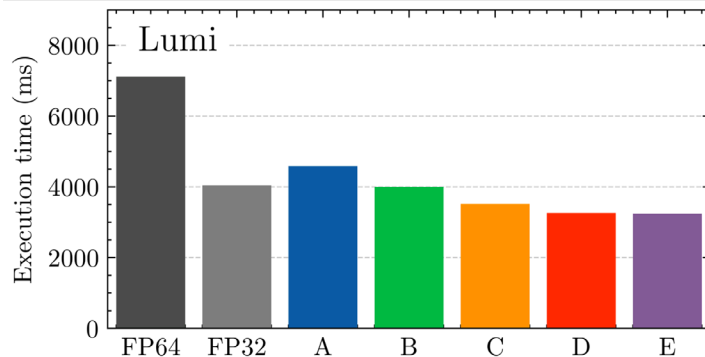
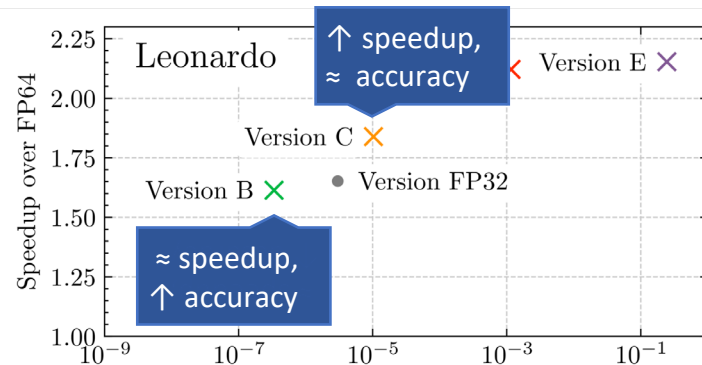
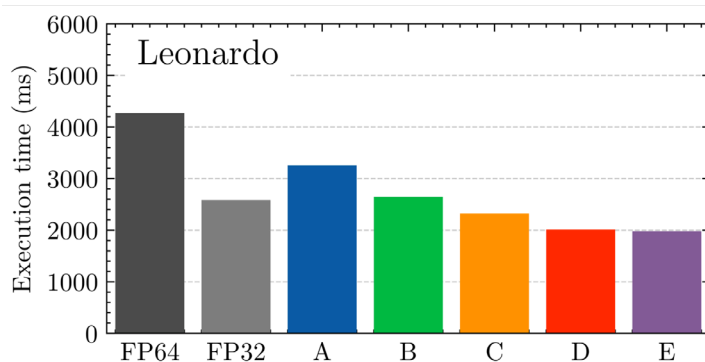
Performance Results



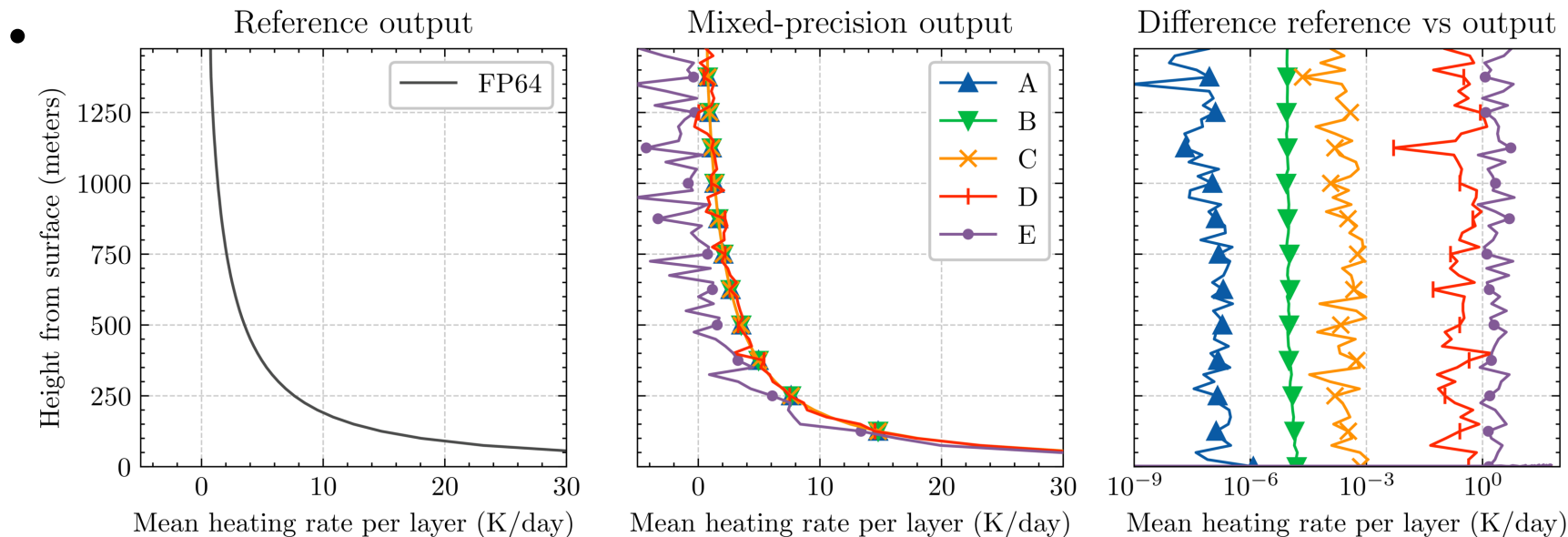
Performance Results



Performance Results



Accuracy Results



Conclusions

- **Mixed precision** gives speedups over uniform precision
 - Up to 2.1× (Leonardo) and 2.2× (LUMI) over FP64
 - Higher accuracy **or** higher performance over FP32 possible
 - No single optimum: error-runtime trade-off
- Portable to AMD and NVIDIA

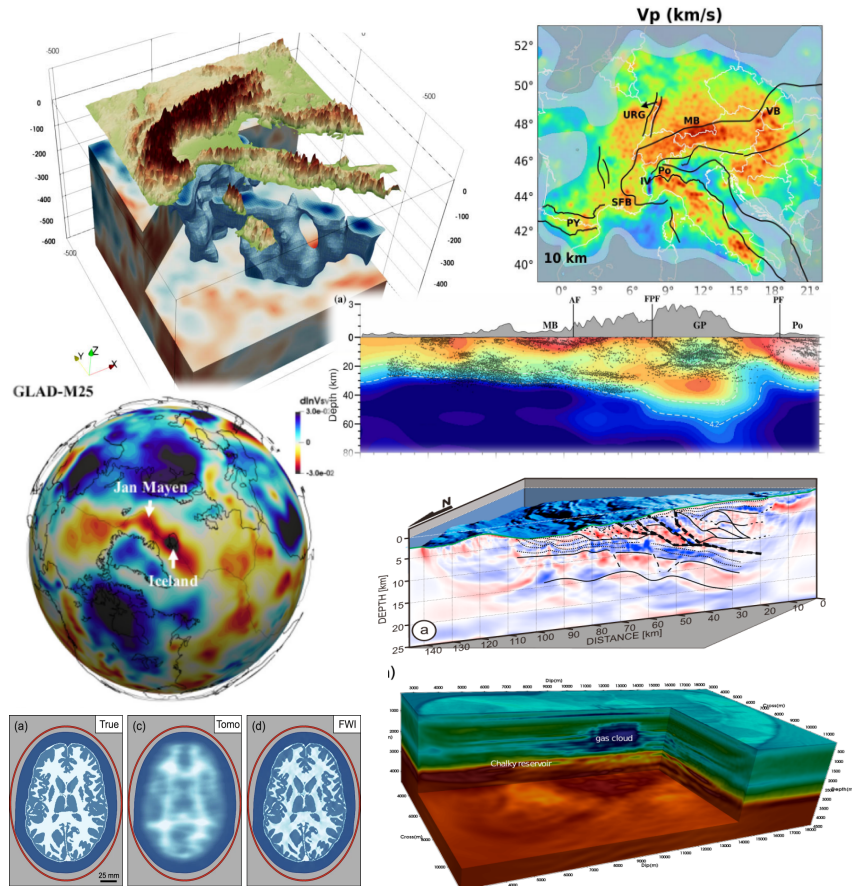
Elastic anisotropic Full-Waveform Inversion of multi-component Ocean Bottom Node data

The Gorgon case study (4C-EFWI)

S. Beller, R. Caiza-Grivalja, L. Combe and S. Operto

WIND project, Geoazur, Université Côte d'Azur, CNRS, IRD, OCA

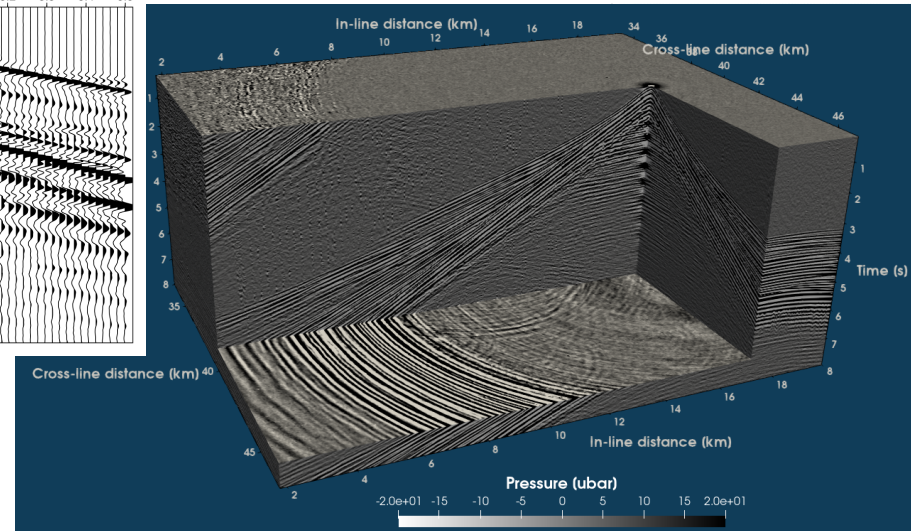
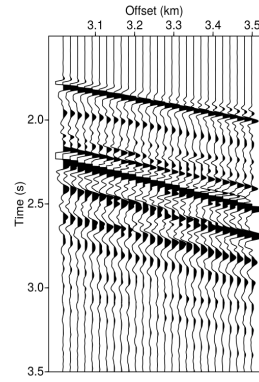
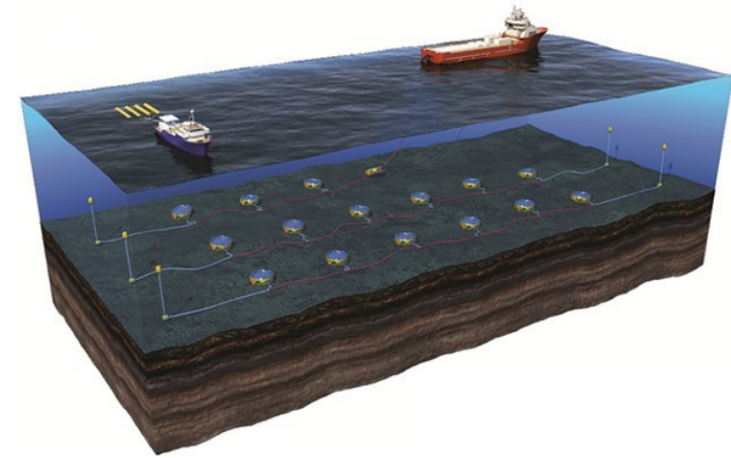
Seismic imaging: a quest for resolution at all scales



- **Seismic imaging applied at various scales**
→ global, continental, crustal, reservoir, samples
- **Structural information and physical properties**
→ non-destructive control of materials (engineering, archaeology)
→ resources prospecting: oil/gas, H_2O , H_2 , mining
→ reservoir monitoring (CO_2 sequestration)
→ geodynamics / geo-hazards
- **Methodological bridges**
→ medical imaging, planetary science, GPR, etc.
- **Never-ending need for higher resolution images**

Seismic exploration configuration

- Shooting vessel triggers **pressure sources** at dense and regular intervals in the water on top of the geological target
- Acoustic and elastic **waves propagate** inside the medium and **interact with heterogeneities** before being recorded by **receivers at the seabed**
 - dense grid of autonomous 4C nodes (water pressure + 3 ground velocities)
- **Time recordings** of these waves
 - indirect measurements of **subsurface properties** (seismic wavespeed, density)
- **Seismic imaging**
 - draw inferences about the subsurface from these measurements



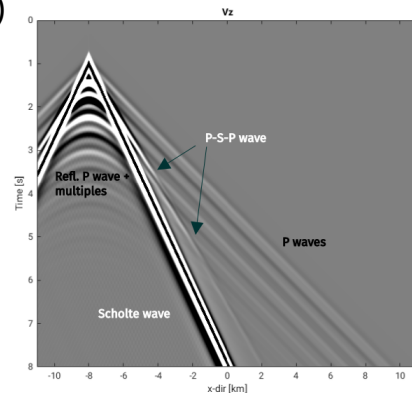
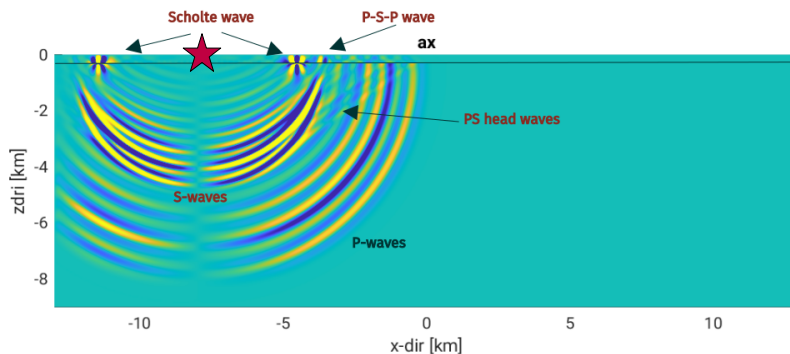
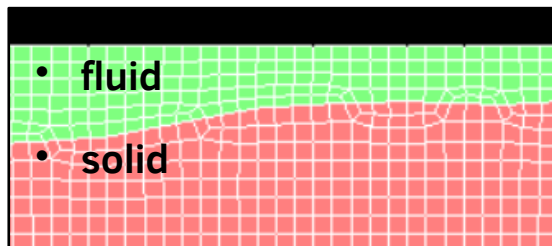
Simulation of wave propagation with SEM

- Near seabed wave propagation is governed by the **acoustic** and **elastodynamics** PDEs

$$\kappa^{-1}(\mathbf{x})\partial_t^2 P(\mathbf{x}, t) - \nabla \cdot (\rho^{-1}(\mathbf{x})\nabla P(\mathbf{x}, t)) = S(\mathbf{x}_s, t) \longrightarrow \text{fluid}$$

$$\rho(\mathbf{x})\partial_t^2 \mathbf{u}(\mathbf{x}, t) - \nabla \cdot [\mathbf{c}(\mathbf{x}) : \nabla \mathbf{u}(\mathbf{x}, t)] = \mathbf{f}(\mathbf{x}_s, t) \longrightarrow \text{solid}$$

- Both discretized with the **Spectral Element Method** and coupled at the fluid-solid boundary
 - **high accuracy** using high-degree polynomials (4th to 7th order)
 - **unstructured hexahedral** elements meshes (*h*-adaptivity)
 - accurate representation of boundaries (**bathymetry**, free surface) and handle **full anisotropy**
 - diagonal mass matrix = **explicit time domain solver**
 - **very efficient** on parallel architectures (data locality, domain decomposition)



Full-waveform inversion (FWI)

- A **non-linear** and **ill-posed parameter estimation** problem solved as a local optimization problem
- **Goal:** Adjust model parameters $\mathbf{m}$ iteratively to minimize the discrepancy $\mathcal{C}(\mathbf{m})$ between recorded seismic waves $\mathbf{d}$ and those simulated $R(\mathbf{u})$ from $\mathbf{m}$

$$\mathbf{m}^* = \arg \min_{\mathbf{m}} \mathcal{C}(\mathbf{m}) \quad \text{where} \quad \mathcal{C}(\mathbf{m}) = \frac{1}{2} \sum_s \sum_r \sum_t [\underbrace{\mathcal{R}_r \mathbf{u}_{s,t}(\mathbf{m}) - \mathbf{d}_{s,r,t}}_{\Delta \mathbf{d}(\mathbf{m})}]^2$$

- **Linearization** and **optimality** condition:

$$\mathcal{C}(\mathbf{m} + \Delta \mathbf{m}) \approx \mathcal{C}(\mathbf{m}) + \nabla \mathcal{C}(\mathbf{m}) \Delta \mathbf{m} + \frac{1}{2} \Delta \mathbf{m} \nabla^2 \mathcal{C}(\mathbf{m}) \Delta \mathbf{m}$$

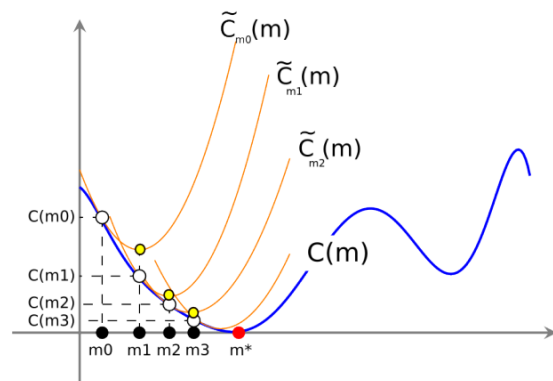
$$\Delta \mathcal{C}(\mathbf{m}) = \nabla \mathcal{C}(\mathbf{m}) \Delta \mathbf{m} + \frac{1}{2} \Delta \mathbf{m} \nabla^2 \mathcal{C}(\mathbf{m}) \Delta \mathbf{m} = 0$$

$$\longrightarrow \nabla^2 \mathcal{C}(\mathbf{m}) \Delta \mathbf{m} = -\nabla \mathcal{C}(\mathbf{m})$$

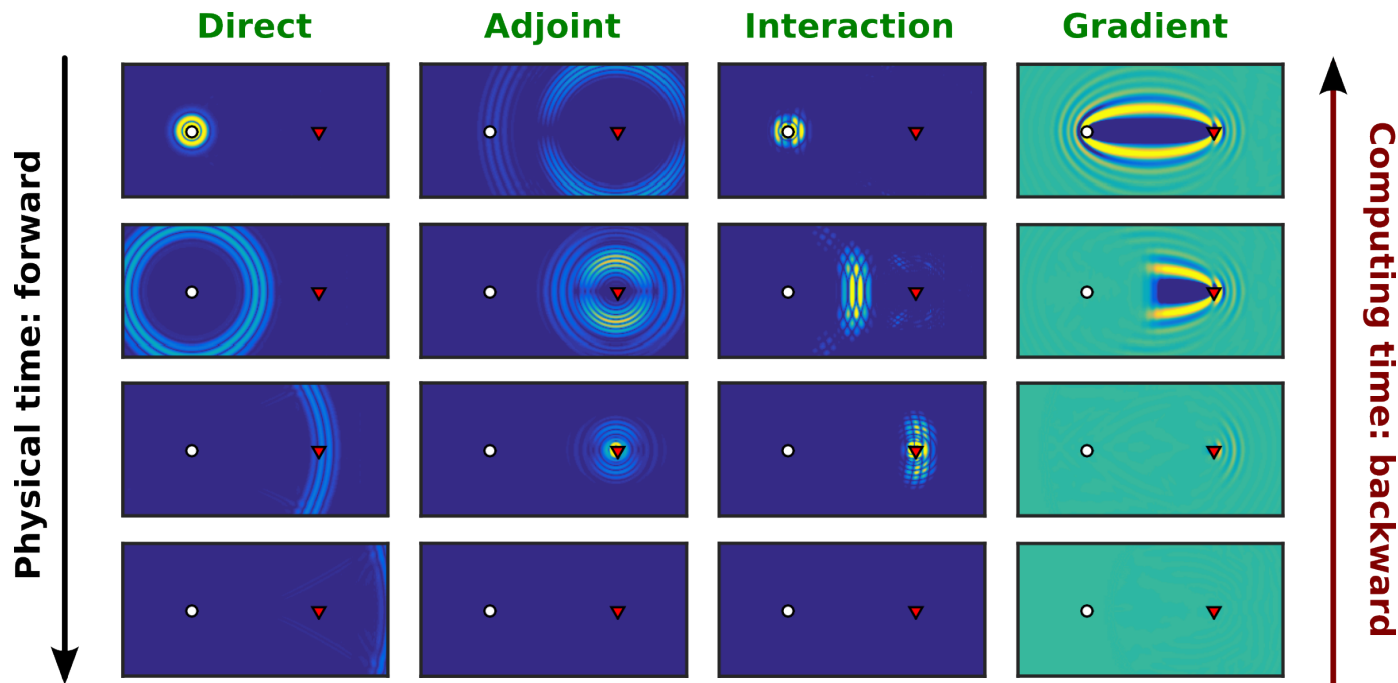
- At each iteration k , the **update** is given by

$$\Delta \mathbf{m}^{(k)} = \left[\nabla^2 \mathcal{C}(\mathbf{m}^{(k)}) \right]^{-1} \nabla \mathcal{C}(\mathbf{m}^{(k)})$$

and solved with a **quasi-Newton algorithm** (l-BFGS) until convergence



Efficient gradient computation (ajdoint state method)

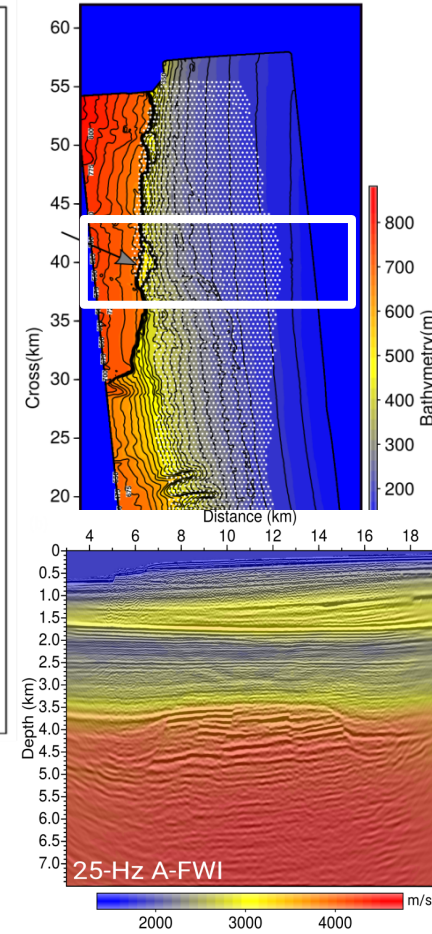
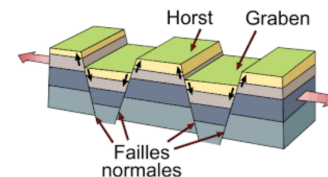
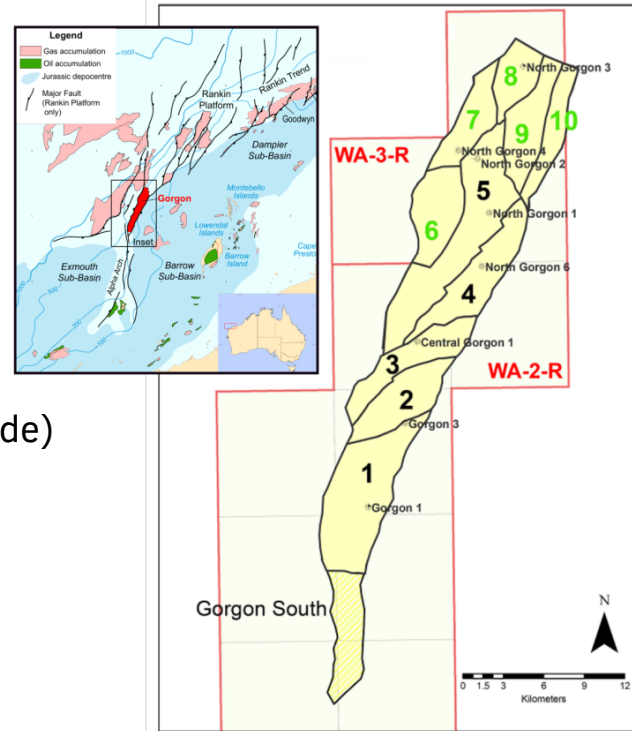


- An **FWI gradient** evaluation **needs 3 wavefields simulations**
 - **forward propagation** from the source for misfit evaluation and wavefield snapshots
 - **adjoint simulation** for backward propagation of data residuals from receivers
 - **reverse time propagation** of the forward wavefield for zero-lag correlation

The Gorgon field, western continental shelf, Australia

4C-EFWI:

- **Target:** Gorgon horst
→ partitioned into fault blocks
- **Reservoir:** Mungaroo formation
→ sand-rich packages alternating with shales
- **Depth:** 3600-4300 m subsea
- **Bathymetry:** 160-760m (submarine landslide)
- **Fault shadows:** Large contrasts between Gearle/Muderong above fast Barrow group
- **Data:** 2015 OBN survey (open)
→ hydrophone + 3C geophones per OBN
→ 386 OBN in the area of interest
- **Innovation:** apply 3D Elastic FWI from the 4C



Elastic FWI, scientific and HPC challenges

- **Motivation**

- Industry mostly uses mono-parameter acoustic FWI (recovering V_p) in marine seismics
- Elastic FWI is needed for accurate imaging and modeling of near seabed complex wave propagation
- Gives access to other quantities (V_s) to better constrain rocks properties (fluid characterization)
- Mandatory when processing geophone components

- **Methodological issues**

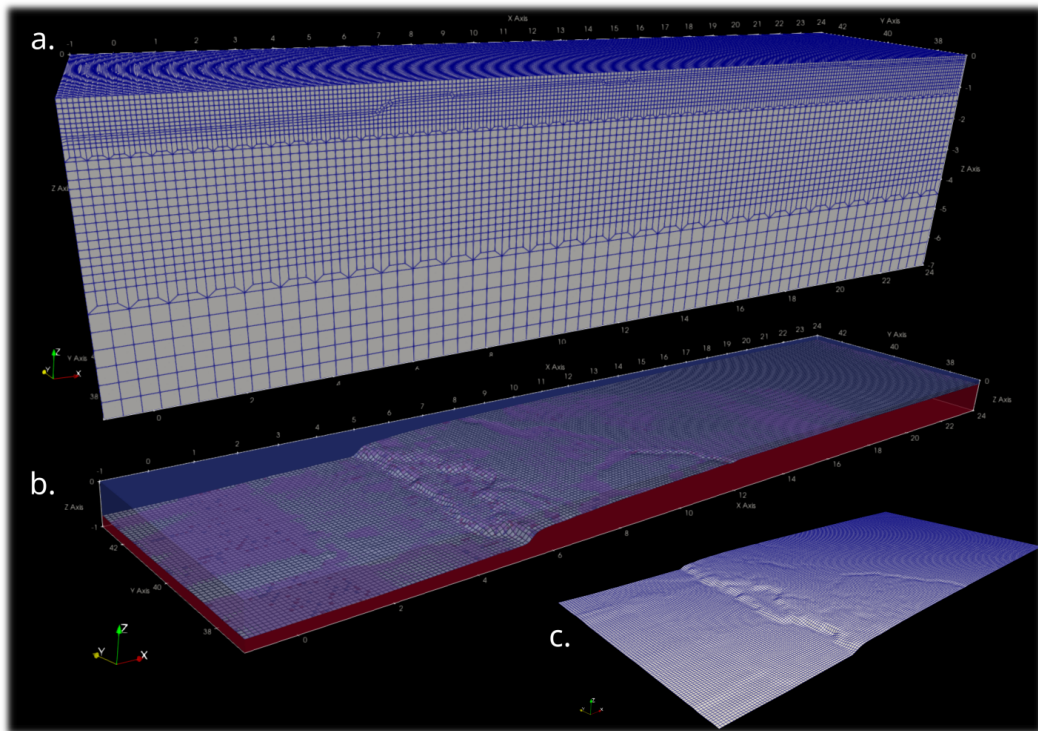
- Initial V_s wave macromodel building, crosstalk when updating several parameters (ρ , V_p , V_s)
- Horizontal components of geophones are often noisy
- Which workflow (combination of components, parameter class and scaling) to perform reliable multiparameter elastic FWI?

- **HPC challenge**

- Elastic FWI costlier than acoustic FWI (10x to 100x) and is more memory demanding
- 3D time domain simulation has N^4 complexity (doubling resolution $\sim 16x$ costlier)
- Time-domain FWI scales linearly with the number of source (and component)
- For convergence, several tens of iterations are required

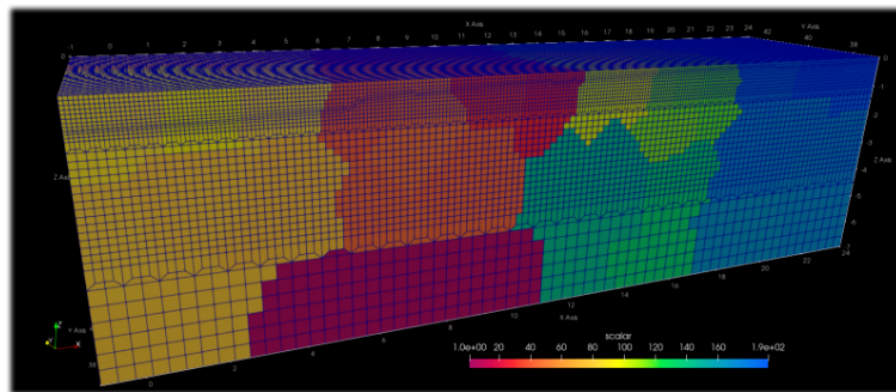
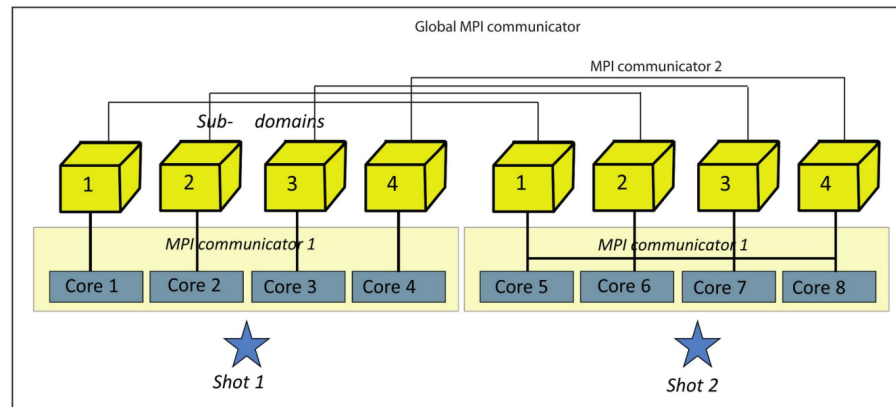
The Gorgon field, meshing

- **Upper-part** needs accurate representation of bathymetry
 - overlay grid method Cubit's Sculpt
 - small elements to handle low V_s
- **Lower-part** needs mesh coarsening to adapt element's size to local wavelength
 - in house cartesian mesher with doubling or tripling layers
- **Final mesh** contains 279 882 P4 elements
 - ~ 25% acoustic and ~ 26 millions dofs
 - adequate for waves with $f_{\max} \approx 7$ Hz
- **Time integration**
 - CFL time step ~ 0.3 ms
 - 30 000 iterations for 9s of simulation



ImEarth, MPI parallelism

- Multi-source simulations with **2 levels of MPI parallelism** → 2 MPI communicators
- **1st level: source parallelism**
 - distribute sources & their simulations across compute nodes
 - “embarrassing” parallelism
 - a few communications once per iteration to gather source contributions
- **2nd level: domain decomposition parallelism**
 - to compute simulation + gradient in parallel over several subdomains
 - communications at domain boundary at each simulation time step
 - tuned for memory usage, parallel efficiency and reasonable time to solution

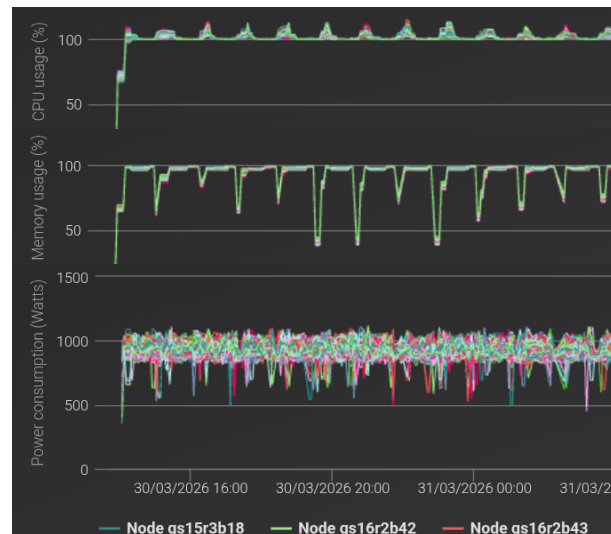
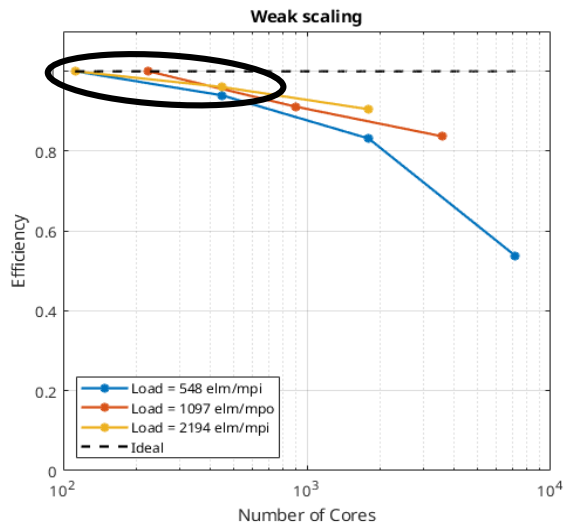
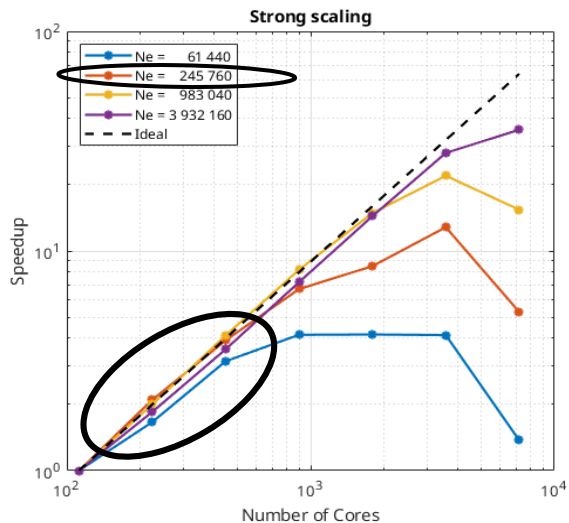


EuroHPC allocation

- **EuroHPC allocation:**
 - 570 000 node hours on MareNostrum 5 GPP (BSC-CNS)
- **Benchmark** performed on MN5
 - **scaling** of forward modeling and gradient **for a single source**
 - reasonable time to solution, good parallel efficiency and full memory occupation using a **single node per source** (112 cores)
 - ~20 min for a simulation and ~60 min for an iteration

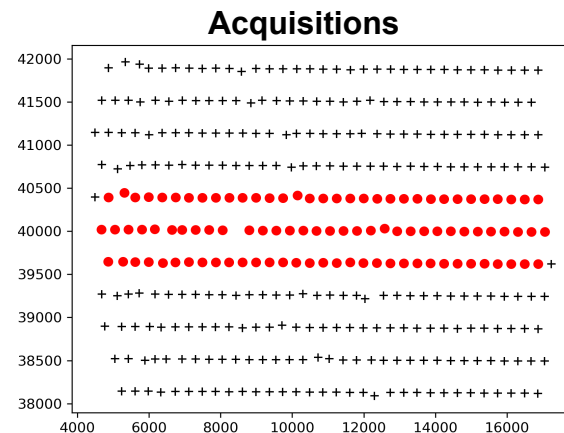


**Barcelona
Supercomputing
Center**
Centro Nacional de Supercomputación



EuroHPC allocation

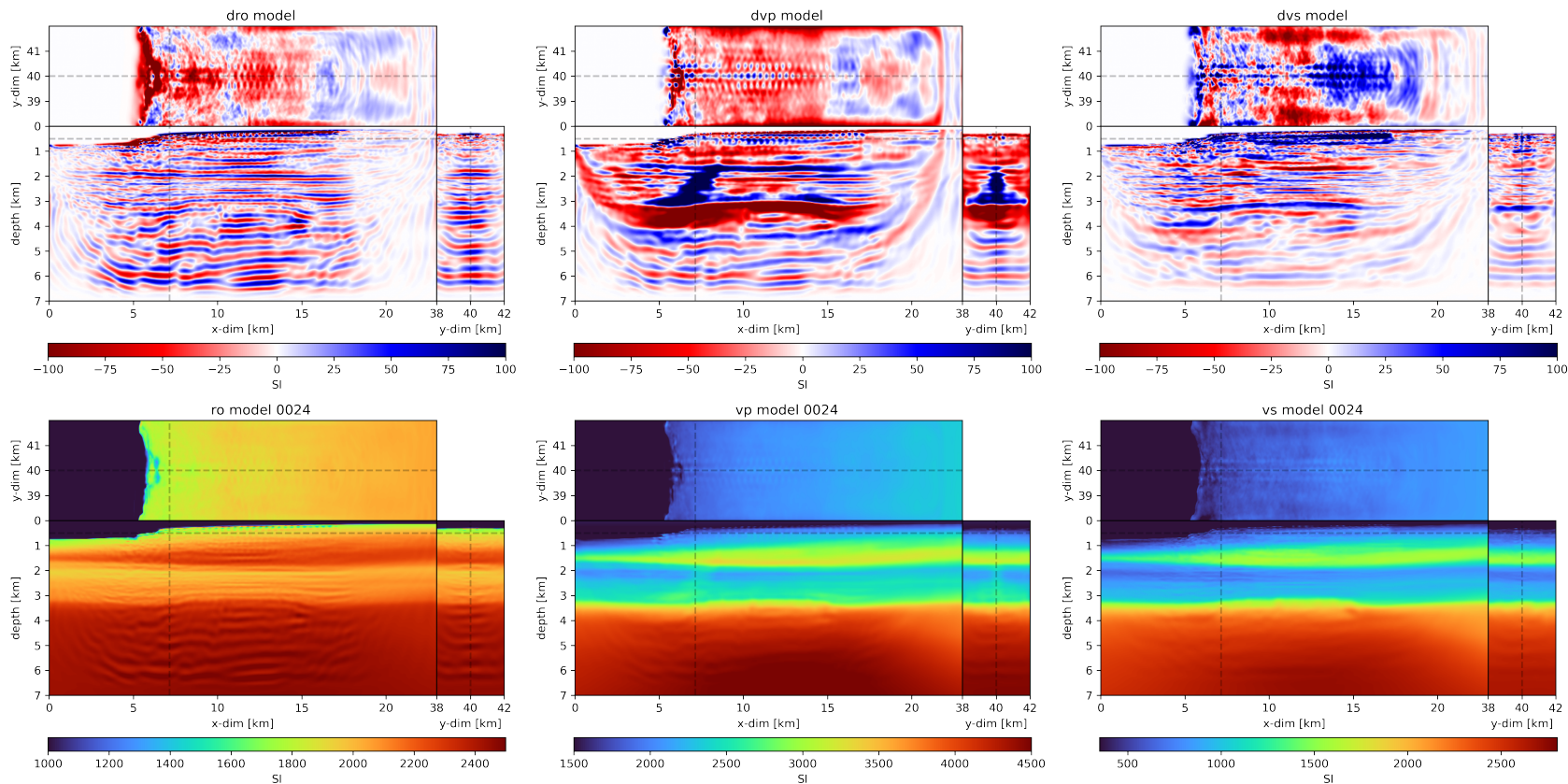
- We performed **several runs** to assess:
 - individual and mixed contribution of components and wave types
 - scaling of parameters (mostly density)
 - test workflow scenarios (P, 4C, P+4C, P+XY)
 - parametric analysis performed on a subset of sources (100 OBN)
 - anisotropy is taken into account but only V_P , V_S & are updated
- When possible, **all sources were treated in parallel**
 - 386 nodes for full acquisition / mono-component FWI (P, Z, X and Y components independently)
 - 100 nodes for subset of acquisition (mono-component)
 - 200/300 nodes for subset of acquisition (mixed-components)
 - 400 nodes for subset of acquisition and 4C
- **For biggest runs:** full acquisition, 4 components jointly
 - 1544 sources dispatched on 736 nodes
 - two sequential simulations per node
 - 70 iterations ~ 110 000 node hours



Total consumption:

- **Node hours:** 513 000 (~90%)
- **Energy:** $4.962 \times 10^{12} \text{ J}$
- **Equivalent CO₂ footprint:** 64.676 t

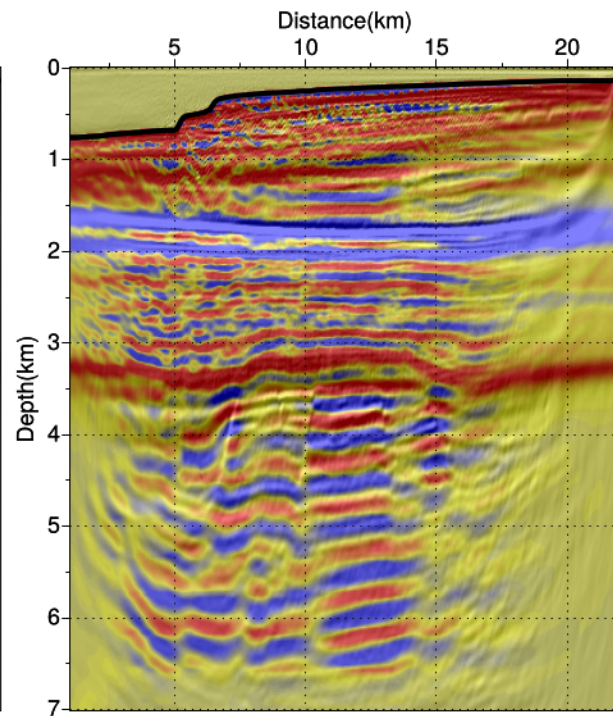
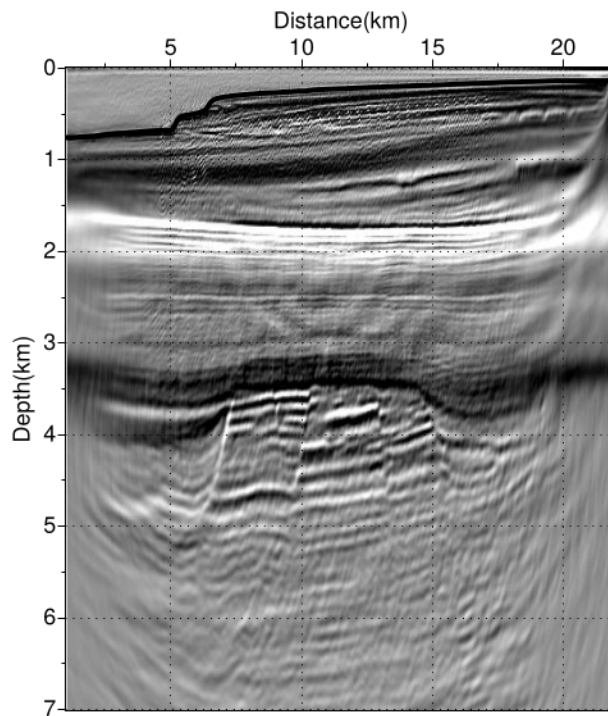
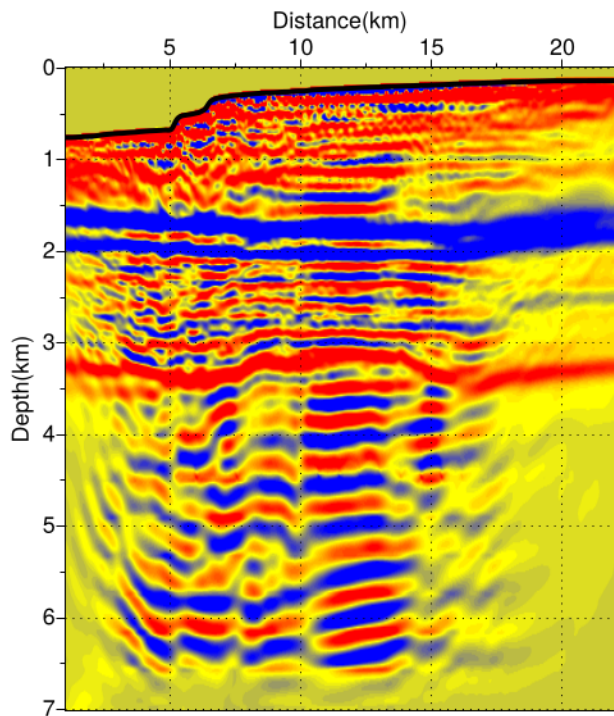
Example of FWI results – P followed by 4C

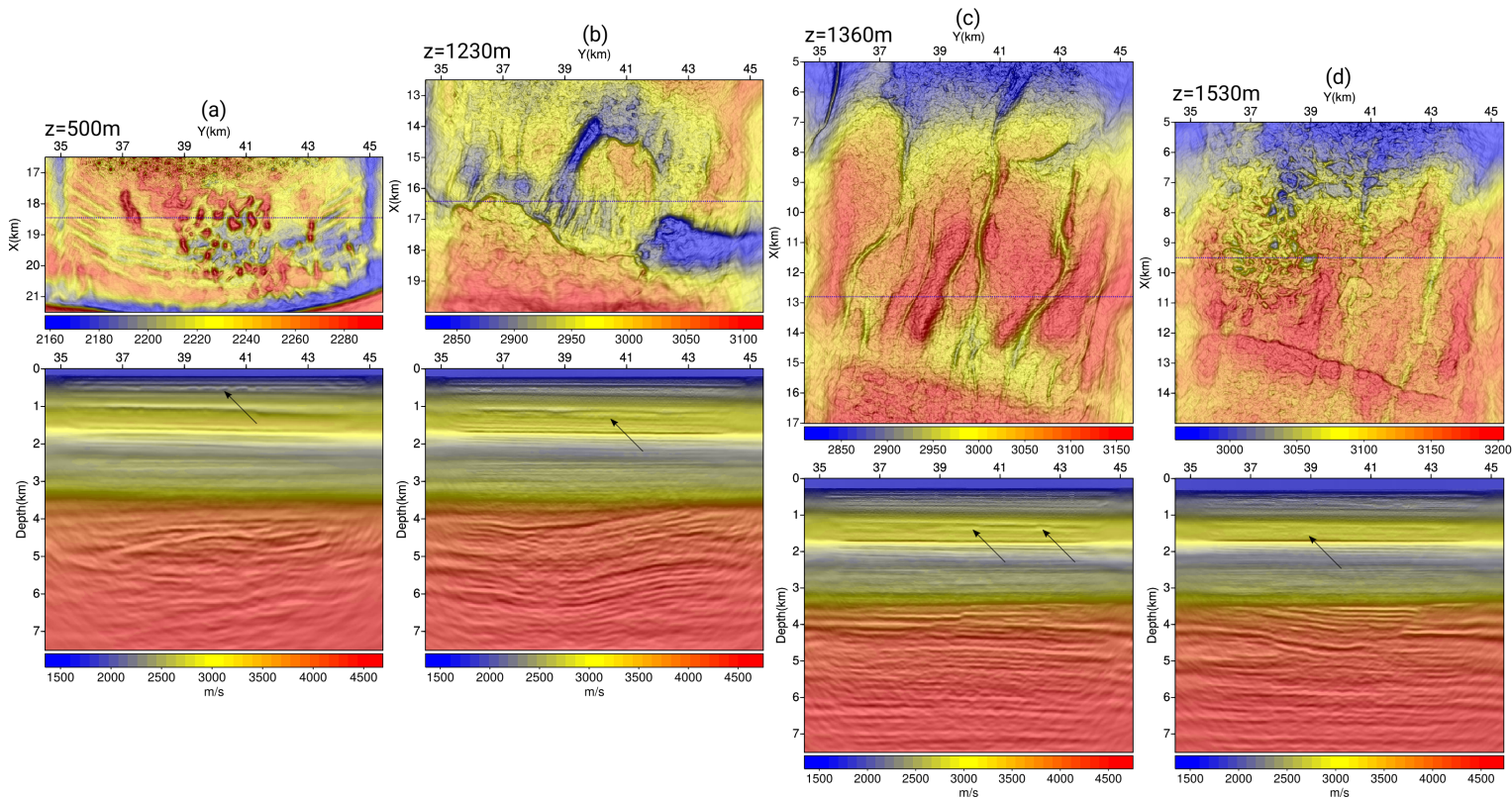


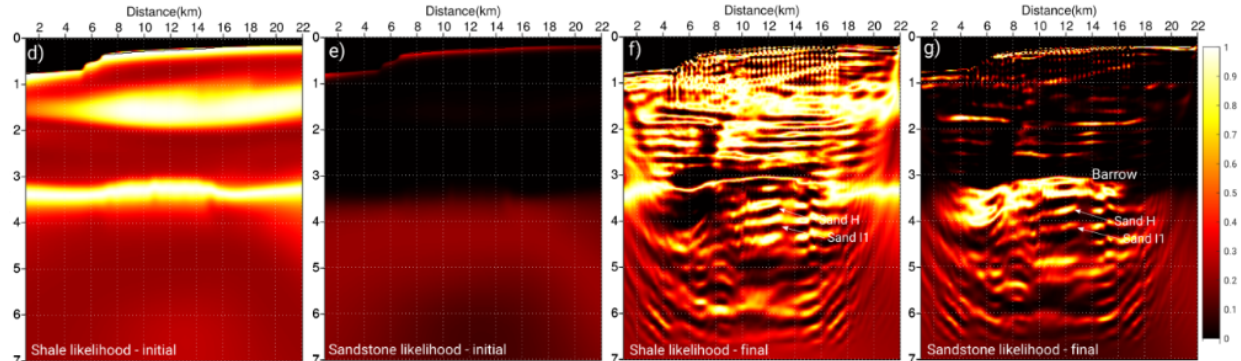
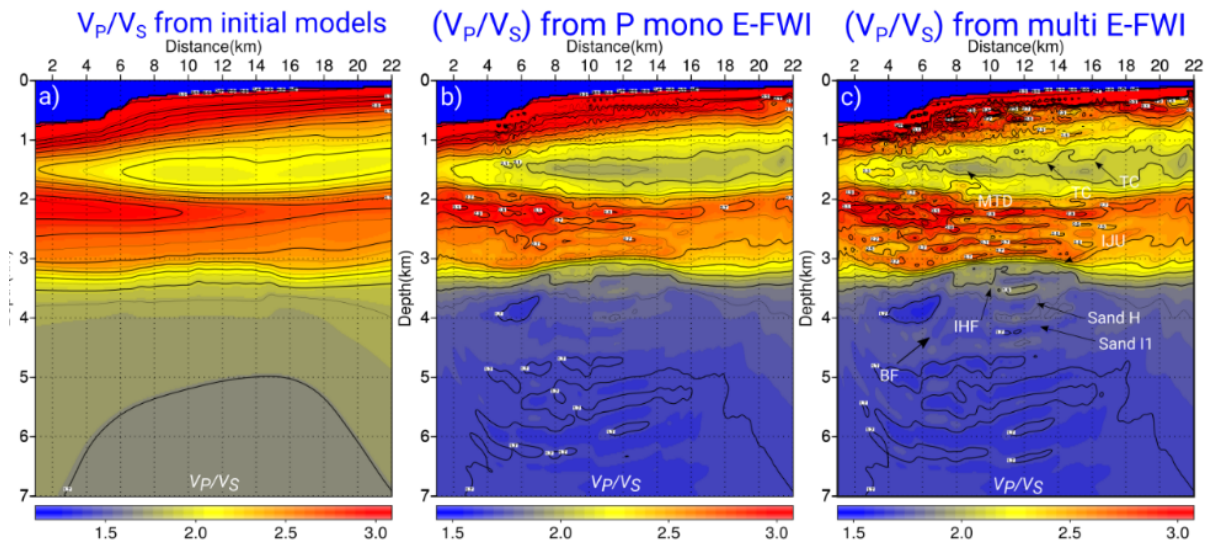
Conclusion & perspectives

- This allocation showed the **feasibility** of multi-parameter 4C E-FWI from an industrial dataset
- It showed that careful tuning of 4C-EFWI involves
 - multi-parameter and multi-component scalings
 - careful QC of estimated 4C source signature
- **Quality control of E-FWI results is still on-going** against high-frequency pure A-FWI, available well-logs and lithology using empirical relationships relating V_p and V_s
- Added-value provided by 4C E-FWI relative to P-EFWI needs to be further assessed with
 - more representative wide-azimuth datasets (increasing surface coverage)
 - higher frequencies
- **Next steps involves:**
 - adding seismic attenuation and anelasticity as optimization parameters (on-going)
 - adding more constraints on the multiparameter inversion (penalizing deviations to empirical relationships)

Thanks for your attention!





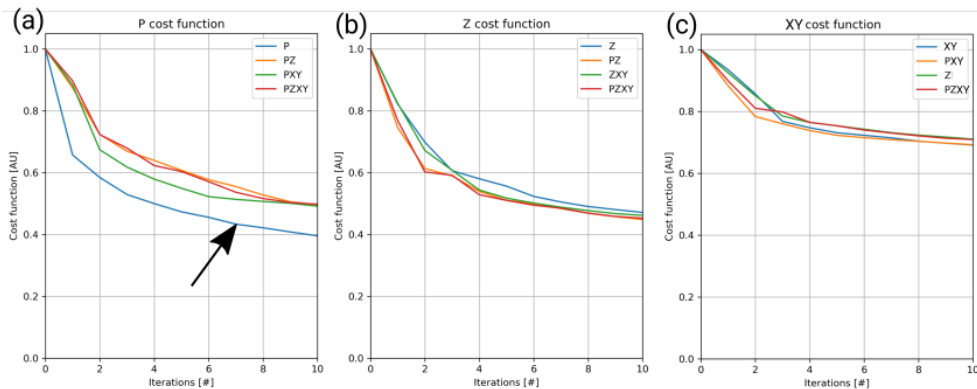


Initial models

multi E-FWI models

Shale and sandstone likelihood (multi E-FWI)

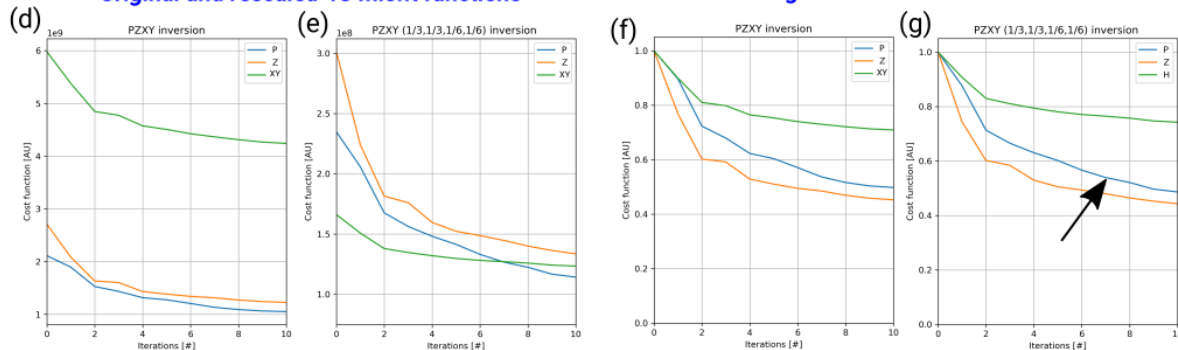
Normalized P, Z, XY misfit functions for single- and multi-component E-FWI



On the scaling of 4C misfit function

Original and rescaled 4C misfit functions

Normalized original and rescaled 4C misfit functions



Component-wise misfit functions

(a) Convergence of P misfit function hampered by other components. (b,c) Conversely, better convergence of Z & XY misfit functions when combined with P. (c) Slower convergence of XY misfit function (noisy data).

4C E-FWI misfit functions

(d) XY components dominate.

(e) Re-scaling the 4C balances their contribution.

(f,g) Normalized misfit functions: re-scaling hampers

convergence of XY misfit function at

benefit of P, Z counterparts. However,

convergence of P misfit function

remains affected by other components

in 4C E-FWI (arrows).



EuroHPC
User Days

Dublin Ireland
22→24
September 2026



EuroHPC
Joint Undertaking

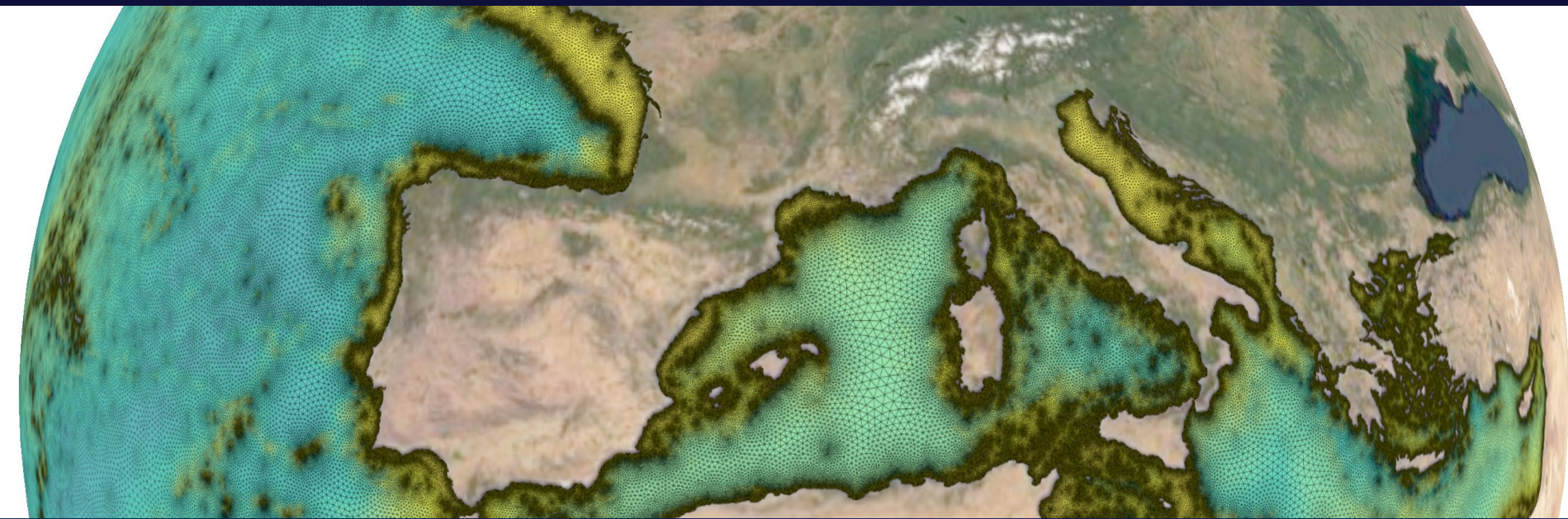


ÉIRE2026.EU
IRELAND2026.EU

Seareport

Modeling the global coastal ocean

Context



Our mission in DG JRC E.1.:

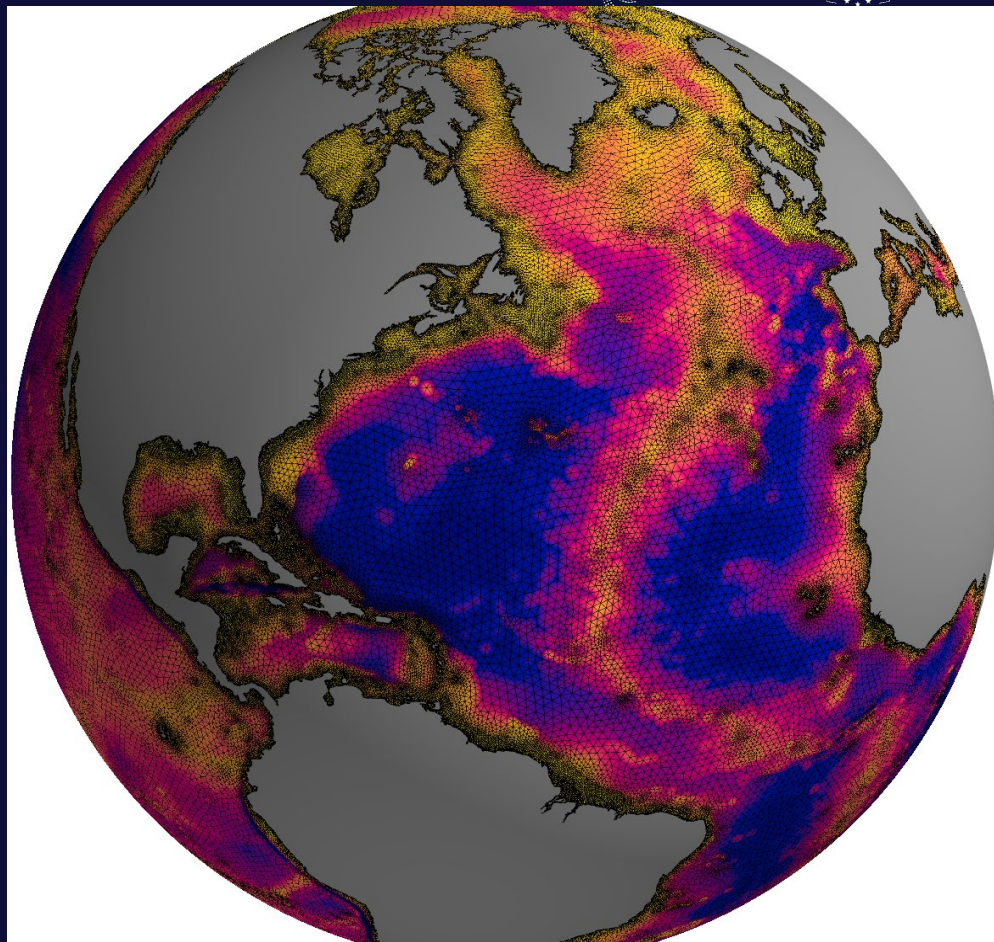
- Conduct research, provide situational awareness related to coastal risk
- Build research excellence, expertise and legacy
- Develop a pre-operational global 3D baroclinic model
- Provide support & data for Member States - through GDACS/ERCC

Our scope:

- Agile development (Robustness, Reproducibility, DevOps)
- Address frontier problems with state of art tools and methods
- Python best practice approach on the ocean modeling chain
- Cultivate our international collaborations

Semi-automatic generation of global models

- Using coastlines / bathy information
- Automatic mesh generation
- Automatic pre-processing (input preparation)



Semi-automatic generation of global models

- Using coastlines / bathy information
- Automatic mesh generation
- Automatic pre-processing (input preparation)
- Global simulations with high coastal resolutions
 - Tides and surges
 - Temperatures & Salinity

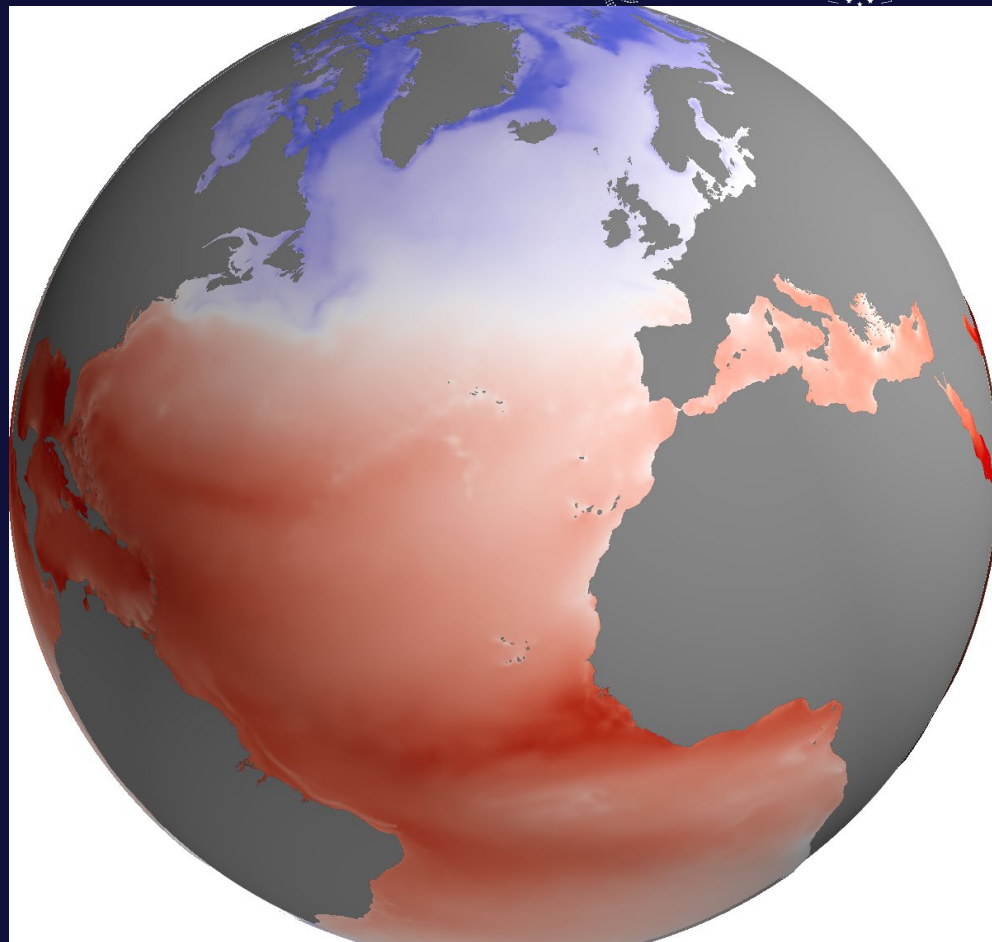


Semi-automatic generation of global models

- Using coastlines / bathy information
- Automatic mesh generation
- Automatic pre-processing (input preparation)
- Global simulations with high coastal resolutions
 - Tides and Surges
 - Temperatures and Salinity

Running on EuroHPC instances

- Meluxina Regular Scale – 110 000 CPU nh
- Marenstrum Extreme Scale – 240 000 CPU nh



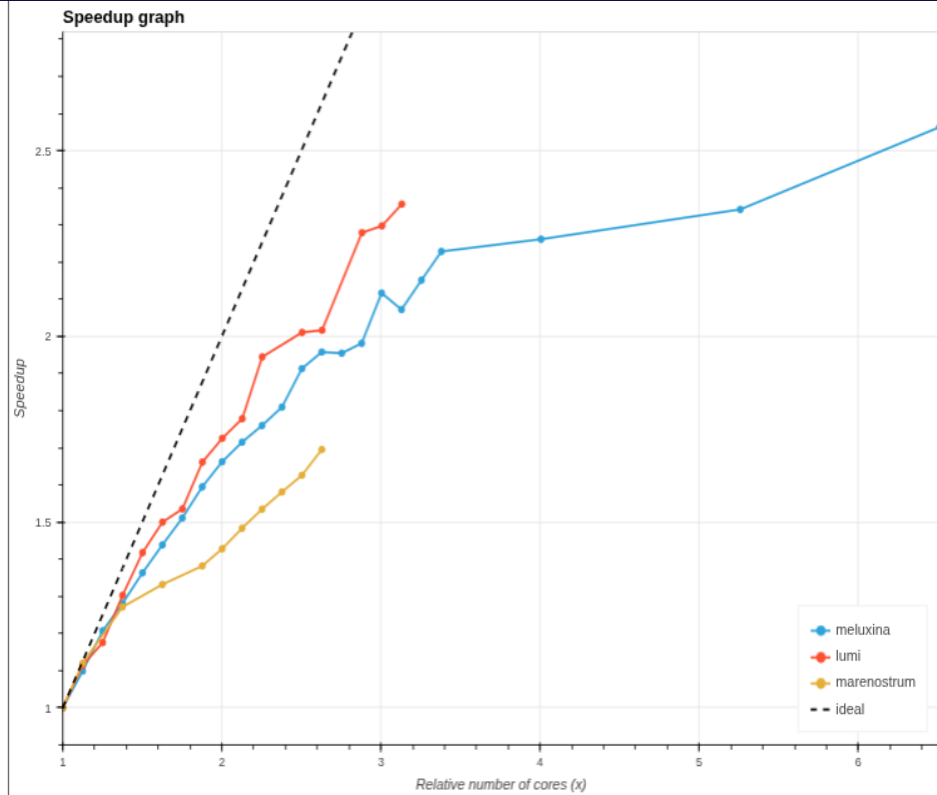
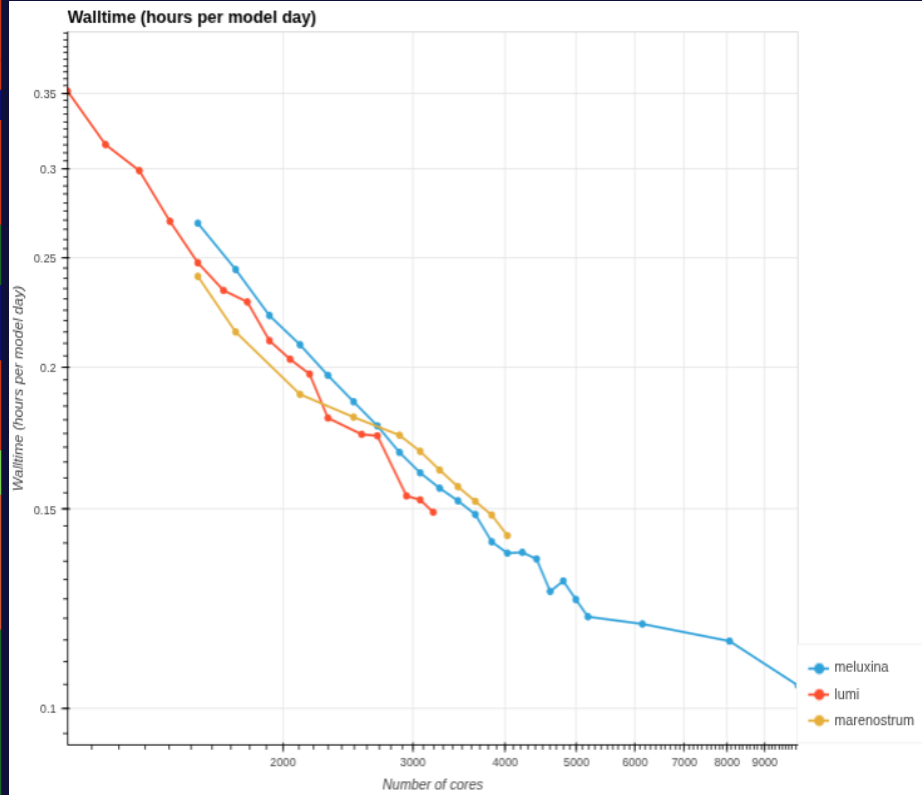
Details:

- Mesh (current version: r06)
 - 11M nodes / 21M elements
 - 33 vertical layers
 - From 50m to 10km resolution
- Model
 - Runs on SCHISM master
 - 3D baroclinic
 - Forced by ERA5 (ECMWF)

Zoom in Irish sea



Scaling curves



Scaling towards Extreme resolutions & EPICURE support:

- **Challenge:** major bottleneck beyond approximately 1024 MPI tasks
- **Help:** EPICURE diagnosed the bottleneck in our solver and integrated NVIDIA's HPC-X toolkit
- **Breakthrough:** we can now scale beyond 1024 MPI tasks, which enables
 - Full 3D baroclinic physics
 - Increased spatial resolution
 - Run large historical simulations (hindcasts)
- **What's next:**
 - 50+ years hindcasts
 - Add physics (waves, compound flooding)
- **About us:**
 - [EuroHPC Project page](https://eurohpcsupport.eu) (eurohpcsupport.eu)
 - [EPICURE blog post](https://epicure-hpc.eu) (epicure-hpc.eu)



Feedbacks:

1. Diverse technical/operational choices between EuroHPC centers creates an interoperability barrier.
2. The EuroHPC Federation Platform is a step in the right direction but still some way to go.
3. Big Data sharing between centers is challenging.
4. A Simplified way to renew/update a long-term project would be appreciated.

NB: Our applications have benefited from the specific allocation for public administrations.

Acknowledgements:

- EuroHPC consortium
 - The Meluxina Support Team
 - The EPICURE support Team
 - The Peer-Review Team
- Our Team:
 - Thomas Saillour
 - Panagiotis Mavrogiorgos
 - George Breyiannis
- Our collaborators:
 - Joseph Zhang (VIMS)
 - NOAA OCS: Saeed Moghimi, Sorrosh Mani, Felicio Cassalho, Jack Reeves-Eyre, Atieh Alipour, Daniel Codiga
 - University of Notre Dame: Albert Cerrone, Joannes Westerink, Damrongsak Wisaraet
 - Many others: EC-JRC, [Oceanmodeling](https://github.com/oceanmodeling) github organisation, TELEMAC Consortium, HEREON



github.com/oceanmodeling

