Neuralio A.I.

How WRF-Data Assimilation benefits from HPC

EuroCC@Greece HPC Training Series - Course 13

Paraskevi Vourlioti

neuralio
ARTIFICIAL INTELLIGENCE

EURO
Greece

28/4/25

OUTLOOK

Our company at a glance
Our HPC Journey
WRF-DA and HPC
- ➢WHY WE NEED DATA ASSMILATION AND WHY HPC IS VITAL
- ➢WHO IS USING DA?
- ➢ WRF-WRFDA
  - We have compiled– now what do we expect?
  - Challenges and tips

Future of WRF and HPC – AI challenge

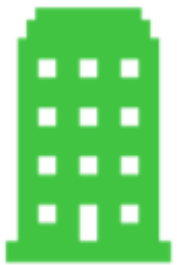
neuralio
ARTIFICIAL INTELLIGENCE

EURO
Greece

Our company at a glance

Year of Establishment

Established in 2022, Thessaloniki, Greece
Estonian Branch in 2023



Personnel

8 persons, gender balance 50/50

Expertise

AI/ML, Atmospheric Modeling, Remote Sensing, HPC

Projects

Public Funded projects: 3 closed/ 2 ongoing
Private Projects:
2 Closed/ 3 Ongoing

Services

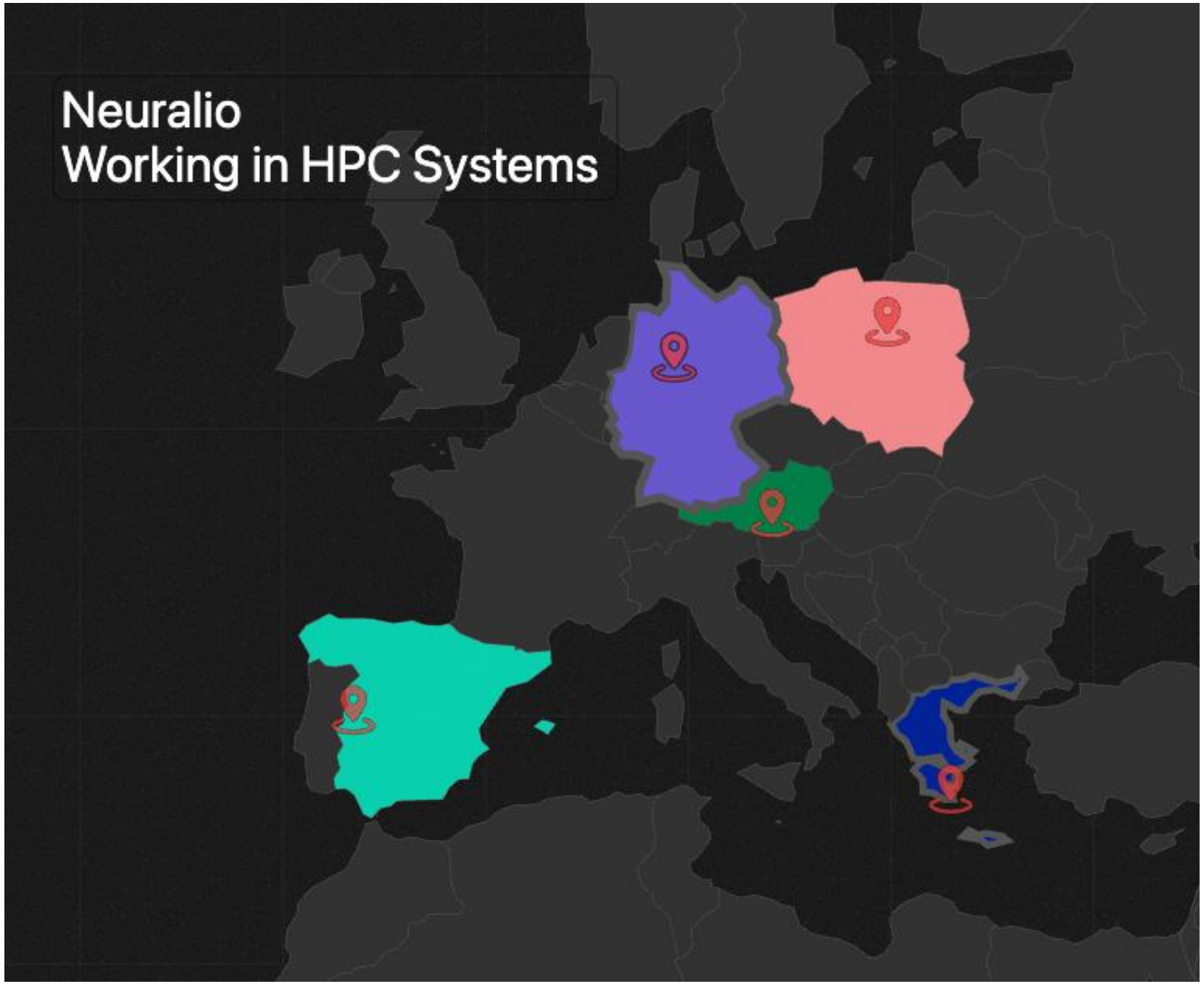
Weather Intelligence
Enviromental Foorprint
Land Evaluation
R&D as a Service

neuralio
ARTIFICIAL INTELLIGENCE

OUR HPC JOURNEY

Many thanks to EuroCC network for providing resources also to SMEs.

neuralio
ARTIFICIAL INTELLIGENCE


Neuralio
Working in HPC Systems

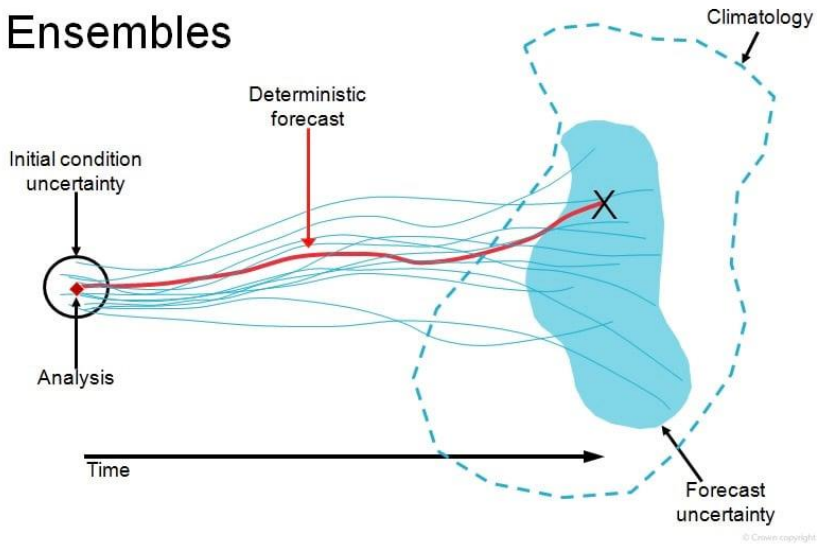
WRF-DA and HPC

WHY WE NEED DATA ASSMILATION AND WHY HPC IS VITAL?

Why Data Assimilation: The atmosphere is a chaotic system, where the accuracy of short and medium-range forecasts depends critically on the precision of the initial state.

What is Data Assimilation: Our best effort to combine mathematically our model with observations to get the closest to the true atmospheric state

WHY WE NEED DATA ASSMILATION AND WHY HPC IS VITAL?



Source: https://opensnow.com/news/post/ensemble-vs-deterministic-weather-models

Computational power to produce:

- ✓ High Spatial Resolution forecasts
- ✓ **Data assimilation to improve our starting point.**
- ✓ Fast production

  → Example in CESGA HPC: 3.5 hours on 336 cores for a 2km spatial resolution forecast and 24 hours forecast horizon

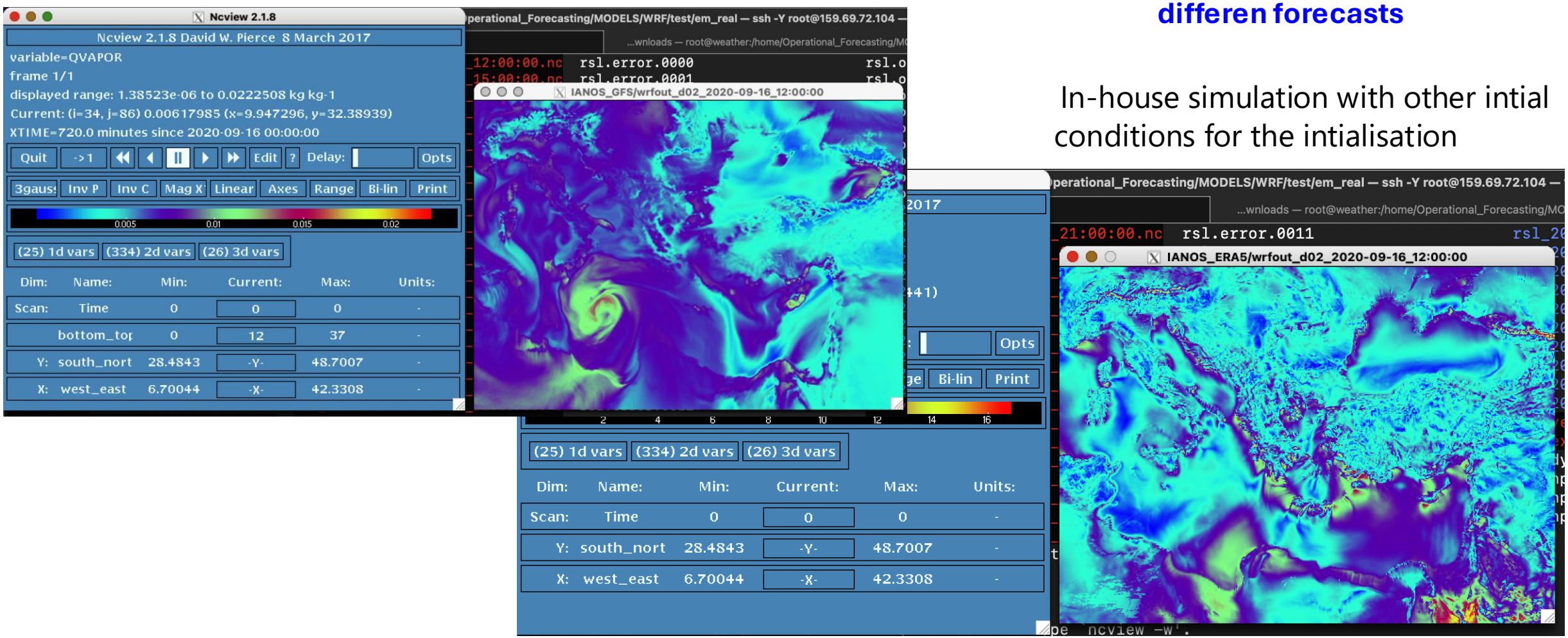
  -> Example in CESGA HPC: also 3.5 hours to assimilate (with 4DVAR)!!!

WHY WE NEED DATA ASSMILATION AND WHY HPC IS VITAL?

In-house prediction with WRF model and intial conditions from GFS system

An example of how different intial conditions can lead to differen forecasts

In-house simulation with other intial conditions for the intialisation



WHY WE NEED DATA ASSMILATION AND WHY HPC IS VITAL?
..... WHO IS USING DA?

Center	Operational DA?	Approx. daily obs assimilated (millions)	Main data sources assimilated	Main DA system type
ECMWF (European Centre)	Yes	~60 (2024)	~90% satellite measurements (radiance sounders, scatterometer winds, GPS-RO, etc.), plus all available in-situ data (surface stations, radiosondes, aircraft reports)ecmwf.int.	4D-Var (four-dimensional variational) with an ensemble of data assimilations for hybrid background error covariancesecmwf.int.
NOAA/NCEP (U.S. GFS)	Yes	~34 (2024)	Vast majority satellite data (e.g. radiances from polar orbiters, geostationary AMVs, GNSS-RO) and conventional observations (radiosondes, dropsondes, buoys, METAR surface reports, aircraft AMDAR/ACARS)repository.library.noaa.gov.	Hybrid 4D EnVar (four-dimensional ensemble-variational assimilation using 80-member EnKF for background errors)nco.ncep.noaa.gov.
JMA (Japan Meteorological Agency)	Yes	~10–15 (est.)	Satellite radiances (NOAA/MetOp ATOVS, Aqua AMSU-A, Himawari AMVs, etc.), ground-based observations (SYNOP/ship/buoy surface reports, radiosondes/pilots),	4D-Var (incremental variational) – JMA's global analysis employs 4D-Var (now with some hybrid ensemble contribution in recent upgrades)jma.go.jp.

WHY WE NEED DATA ASSMILATION AND WHY HPC IS VITAL?
WRF-WRFDA

- When you download WRF, before you compile it, make a copy of the source code and rename it WRFDA-vX.X.X. ---- EXECUTABLE: da_wrfvar.exe
- Use the same libraries, compiler you used for WRF.
- Depending on the DA methodology you will need different compilation strategy.
- Simpler DA that can run without HPC: 3DVAR -- straightforward compilation
- For WRFDA 4DVAR, it is necessary to have WRFPLUS installed first
  - HPC needed
    and a form of this methodology is used by Weather forecasting Centres
- Start with the test data available in the tutorial starting with a simple 3DVAR

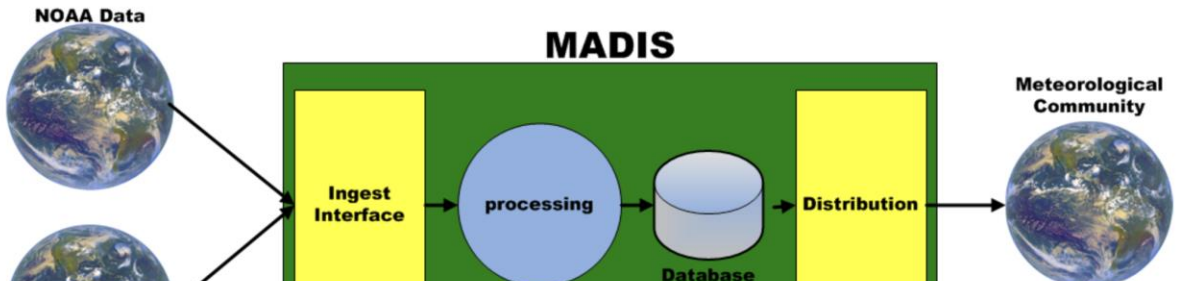
WHY WE NEED DATA ASSMILATION AND WHY HPC IS VITAL?
WRF-WRFDA

- Start with the test data available in the tutorial starting with a simple 3DVAR.
- User guide --- tutorial --- free input data for DA
  https://www2.mmm.ucar.edu/wrf/users/wrfda/
- Pain point: If you have your own data to include in the DA, you have to prepare the little-R file on your own: https://www2.mmm.ucar.edu/wrf/users/wrfda/OnlineTutorial/Help/littler.html
- Use only data that have passed quality control and you trust. Quality-controlled weather stations:

WHY WE NEED DATA ASSMILATION AND WHY HPC IS VITAL?
WRF-WRFDA

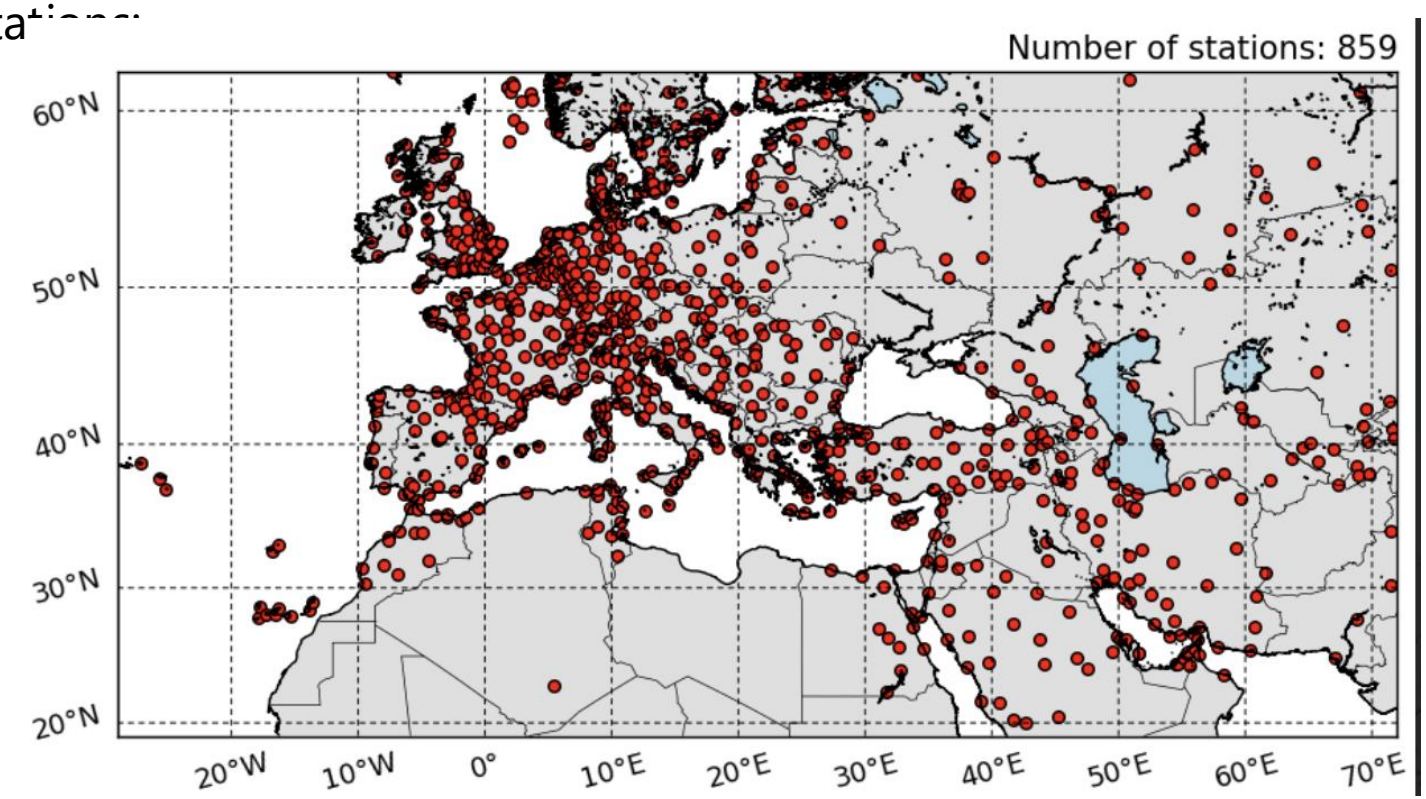
- Start with the test data available in the tutorial starting with a simple 3DVAR.
- User guide --- tutorial --- free input data for DA
  https://www2.mmm.ucar.edu/wrf/users/wrfda/
- Pain point: If you have your own data to include in the DA, you have to prepare the little-R file on your own: https://www2.mmm.ucar.edu/wrf/users/wrfda/OnlineTutorial/Help/littler.html
- Use only data that have passed quality control, and you trust. Quality-controlled weather stations:

MADIS is a meteorological observational database and data delivery system that provides observations that cover the globe.



WHY WE NEED DATA ASSMILATION AND WHY HPC IS VITAL?
WRF-WRFDA

- Use only data that have passed quality control, and you trust. Quality-controlled weather stations:

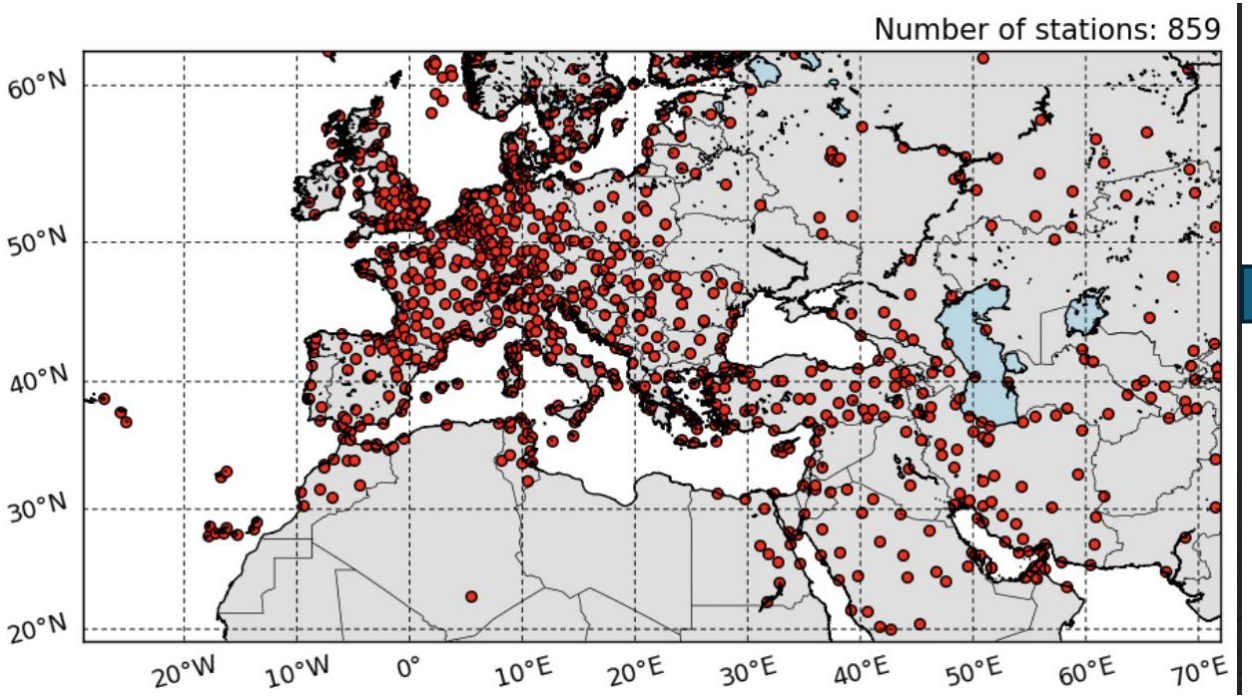


WHY WE NEED DATA ASSMILATION AND WHY HPC IS VITAL?
WRF-WRFDA

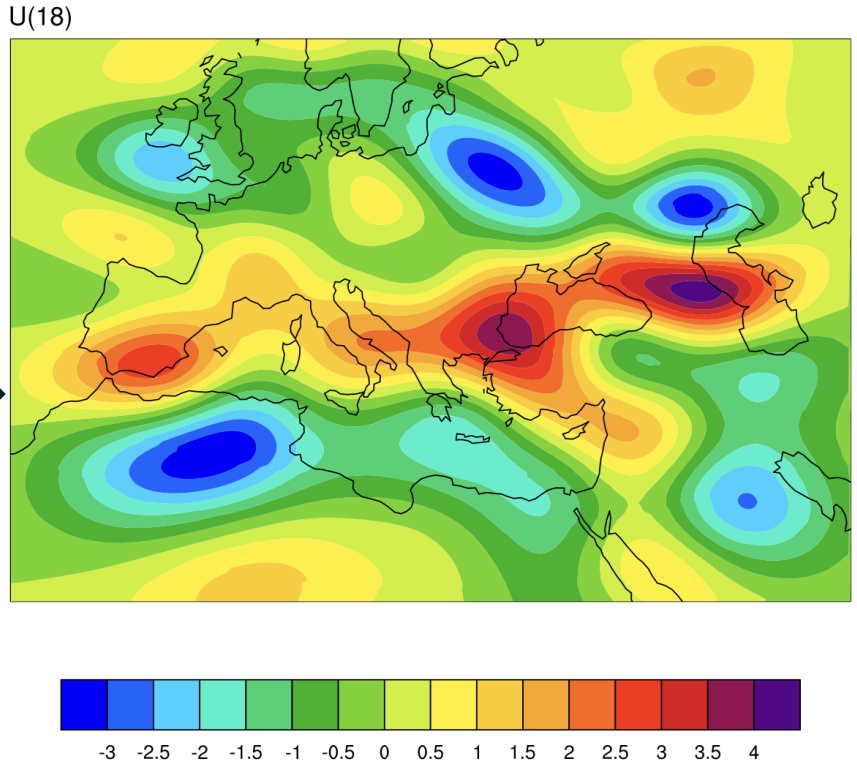
- Start with the test data available in the tutorial starting with a simple 3DVAR.
- User guide --- tutorial --- free input data for DA
https://www2.mmm.ucar.edu/wrf/users/wrfda/
- Pain point: If you have your own data to include in the DA, you have to prepare the little-R file on your own: https://www2.mmm.ucar.edu/wrf/users/wrfda/OnlineTutorial/Help/littler.html
- Use only data that have passed quality control, and you trust. Quality-controlled weather stations.
- Tools to convert the data for MADIS to little-R:
https://www2.mmm.ucar.edu/wrf/users/wrfda/download/madis.html

WHY WE NEED DATA ASSMILATION AND WHY HPC IS VITAL?
WRF-WRFDA
We have compiled– now what do we expect?

We expect to get the "analysis", our new intial state to start our forecasts.



Corrected wind fields we can see from the analysis-minus-background*.



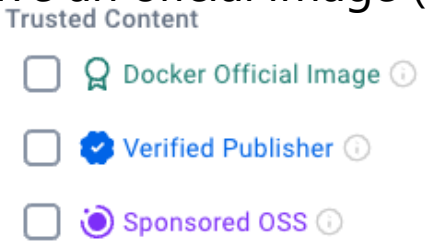
*** backround= intial field estimation**

https://github.com/wrf-model/WRFDA_TOOLS

Challenges and tips

Containerisation in the HPC

- HPC stuff can pull docker images from Docker Hub registry and convert them into singularity.
- What does this mean?
- The user can just “load” the module from command line.
- Ask always the technical team if they can do this for you --- Especially if you have an oficial image (e.g MET)

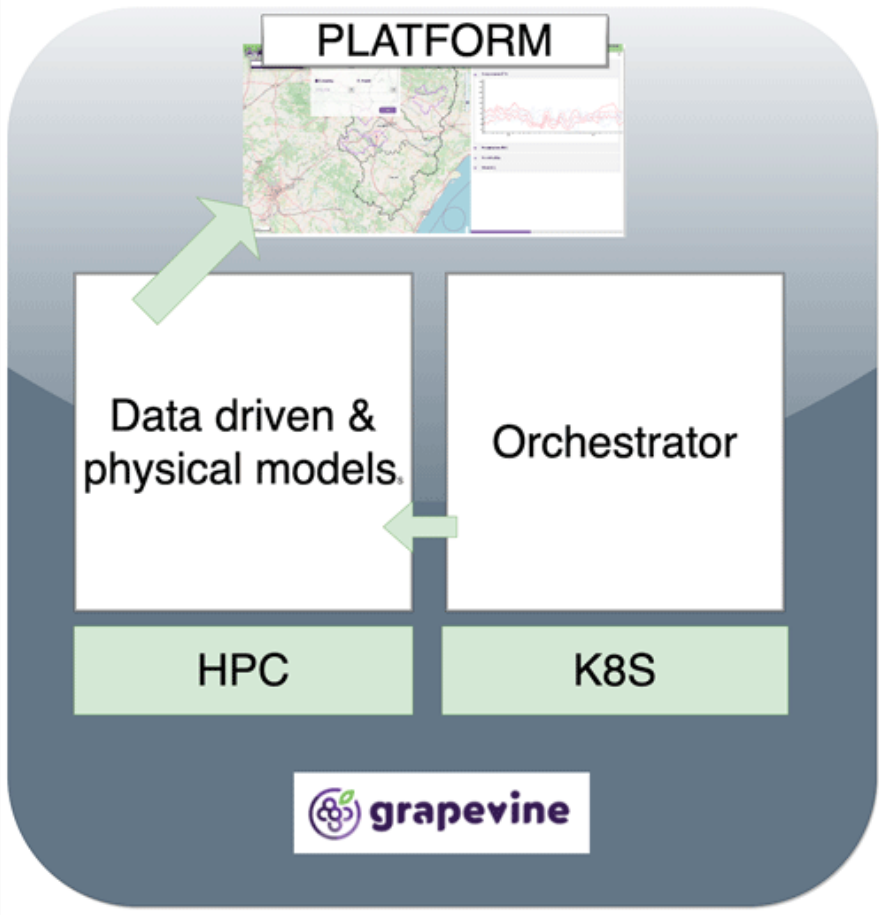


HPC

singularity build <output_image.sif> docker://<docker_image>

Challenges and tips

Operational set-ups with HPCs



Source:
https://asr.copernicus.org/articles/20/1/2023/

1. A need to orchestrate the job submisions to the HPC.
2. Orchestrators like cloudify can do this by connecting the cloud (k8s) with the HPC through croupier plugin.
3. The job-submisions are linked to a user with a certain amount of credits in the HPC.
4. The job can be submited but it cannot start if the user has not reserved the computation time in the HPC.

? Automated reservations/configurations for oprational set-ups?
Should large and small jobs grouped under one reservation? -→ Not optimal for the HPC usage as we reserve the max amount of cores but for some time these resources remain idle.

Future of WRF and HPC – AI challenge

Key Examples:

- **GraphCast (DeepMind):** Outperforms ECMWF HRES for 10-day forecasts, near real-time.
- **FourCastNet (NVIDIA/MIT):** High-res (0.25°) forecasts in <2 seconds.
- **Pangu-Weather (Huawei):** Transformer-based, cheaper and faster than classical NWP.
- **ECMWF AI4Weather:** Integrating AI into operational workflows (hybrid future).

Future of WRF and HPC – AI challenge

Challenges:

- AI struggles to fully replicate complex physics (e.g., convection, microphysics).
- Data assimilation workflows (4D-Var, EnKF) hard to integrate with pure AI.
- AI generalization and reliability for extreme events still uncertain.
- Training requires massive datasets and HPC resources.
- Trust and interpretability issues with black-box AI models.

Future of WRF and HPC – AI challenge

Opportunities:

- AI surrogate models can accelerate expensive WRF physics (e.g., radiation).
- AI improves post-processing: bias correction, downscaling, uncertainty quantification.
- Hybrid WRF + AI systems: physical consistency + AI speedup.
- Reduced computational cost → high-res forecasts and large ensembles feasible.

Neuralio A.I.

Thank you!!

AI will not replace WRF overnight — but it will change how we forecast.
HPC remains vital to support both next-gen NWP and AI models.

Contact Us

www.neuralio.ai

info@neuralio.ai

neuralio
ARTIFICIAL INTELLIGENCE