

Bris Powered by Forecast-in-a-Box

A WMO pilot project for Malawi

AIM4Scale Workshop, Abu Dhabi

Corentin Carton de Wiart, James Hawkes, Harrison Cook, Vojtech Tuma, Jenny Wong, Tiago Quintino ECMWF

Vegard Bønes, Thomas Nipen, Lene Østvand, Nico Budewitz et al. **MET Norway**



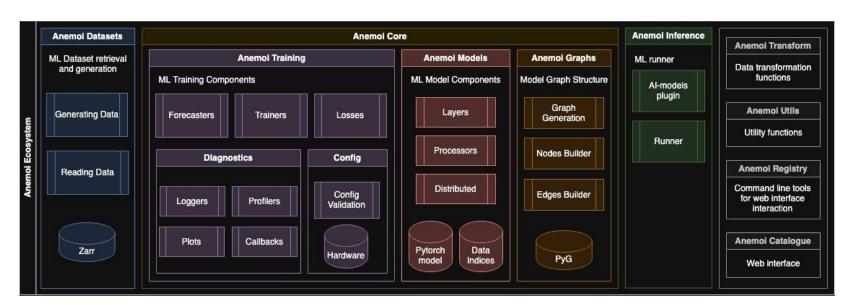




AIFS: Artificial Intelligence Forecasting System. Prediction(s) Data assimilation Modelling Machine learni Observation components parts of the 0 0 (i) (ii) (iii) (ii anemoi Input ERA5 a dataset of hourly Checks accuracy physical states of Earth since 1940 against output Output Corrects errors to improve accuracy

What is Anemoi?

- Open-source framework that provides a complete toolkit to develop data-driven weather models – from data preparation through to inference
- A highly modular framework organised into different Python packages









Who is Anemoi?

Koninklijk Nederlands Meteorologisch Instituut

∕∢ GeoSphere

Ministerie van Infrastructuur en Waterstaat

Winner of this 2025 EMS Technology Award

























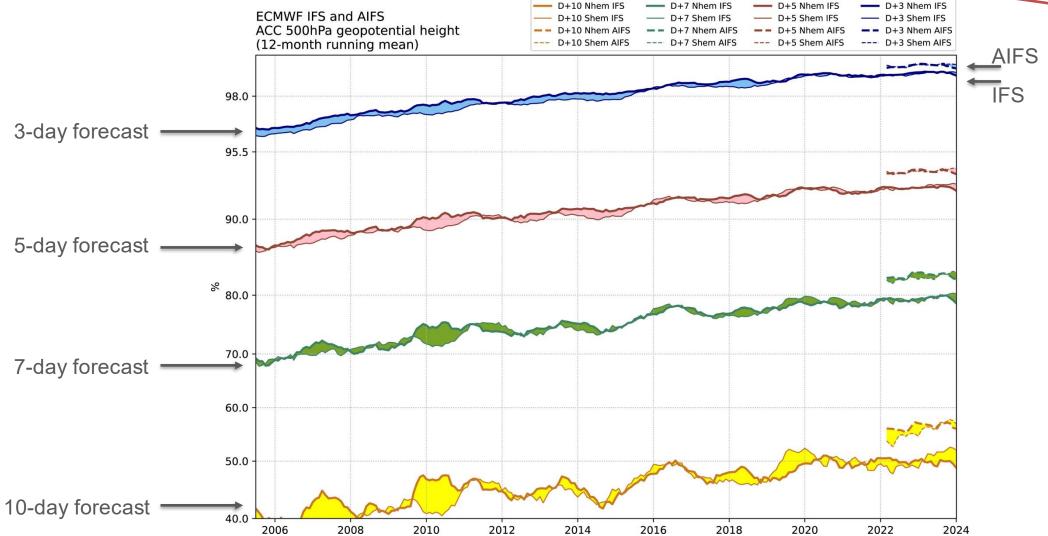
Pooling of resources without resulting in a single forecasting model

Open-source development driven by European Meteorological Centers and ECMWF



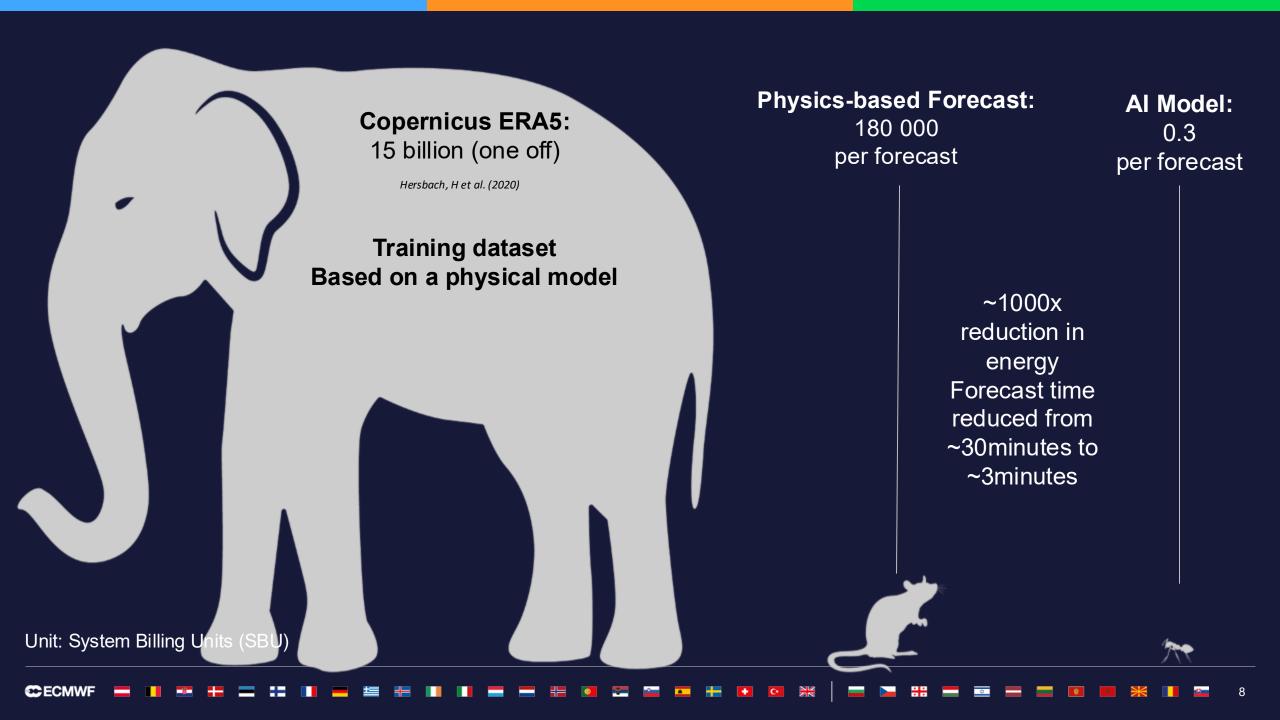
Verification of AIFS against IFS: Accuracy of Z500 forecasts



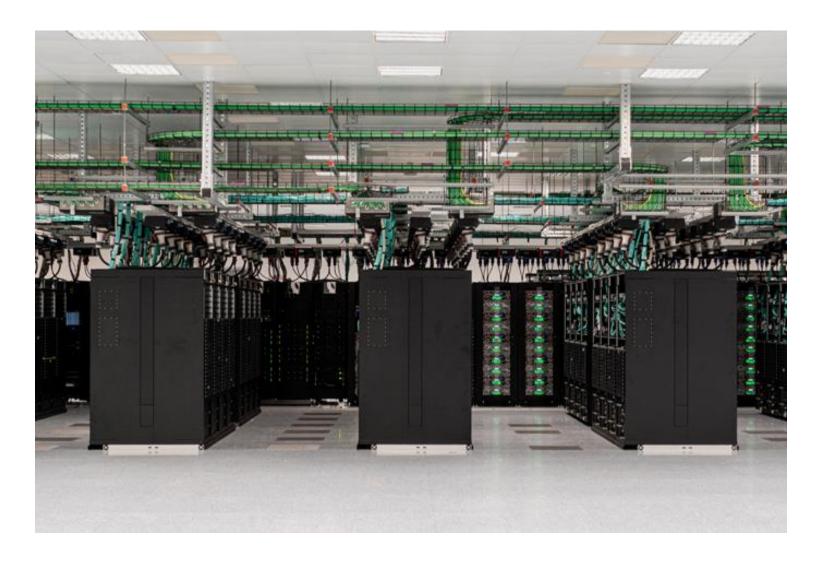




following Bauer et al. 2015



High Performance Computing at ECMWF



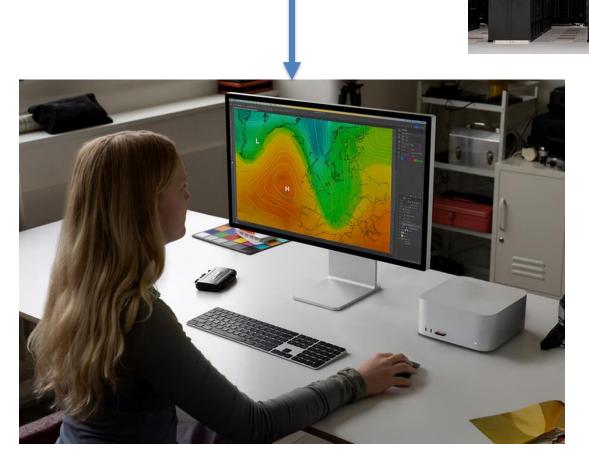


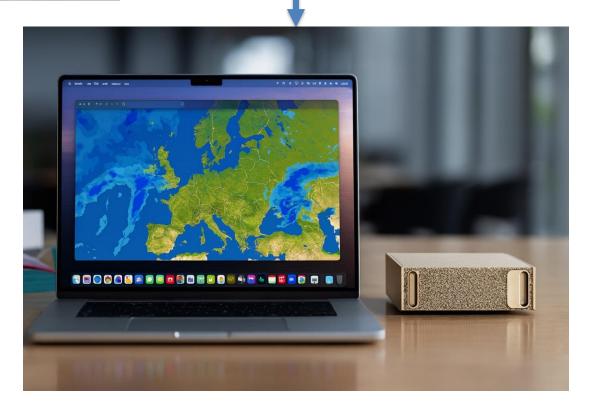


ECMWF 50-year anniversary!



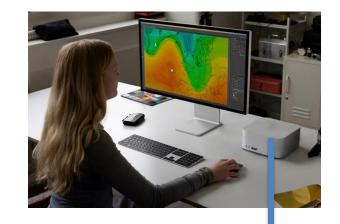
From Supercomputer to Super Convenient!







Forecast-in-a-Box: a Forecasting System on your Desk





New Paradigm:

Instead of shipping huge forecast data, we can send the model instead

Give them the "box" that produces the data

earthkit



Note: it is a virtual box, NOT hardware!



The Earthkit Ecosystem













earthkit





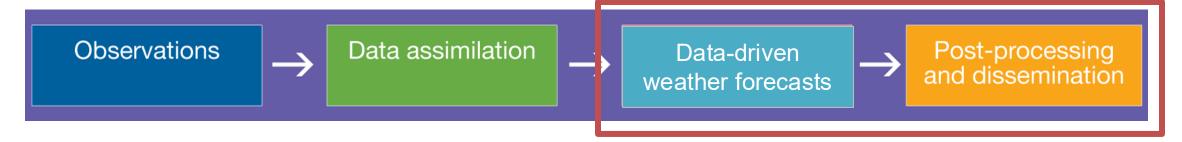








Forecast-in-a-Box - What is it?



Forecast products, not only model outputs

- Includes all the the product generation and processing capabilities
- Includes catalogue of products
- Visualization of graphical products

Packaged in a 'Box'

- Easy to Install, Run and Use
- Hardware agnostic "Whatever works for you ..."
- Workflow designed to work on resource constrained environments
- Model, products and execution environment are fully configurable

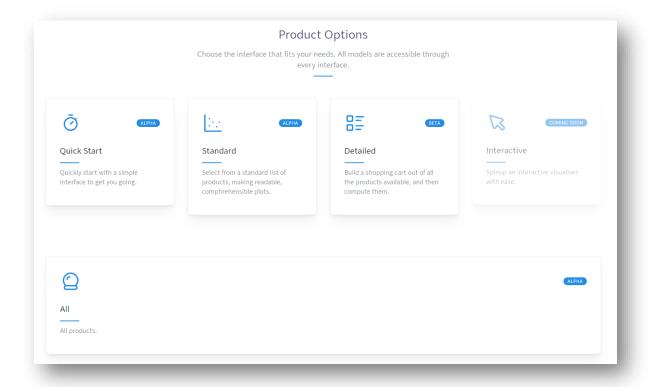








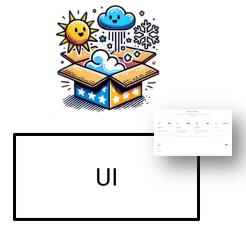
- Application containing UI + Catalogue
 - Provides control to the user, whilst keeping it accessible

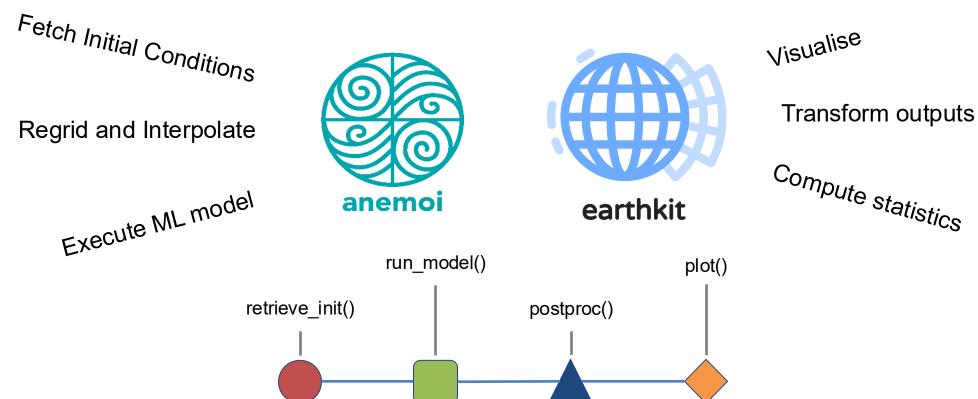




Task Definitions

Constructed from Anemoi and Earthkit functions

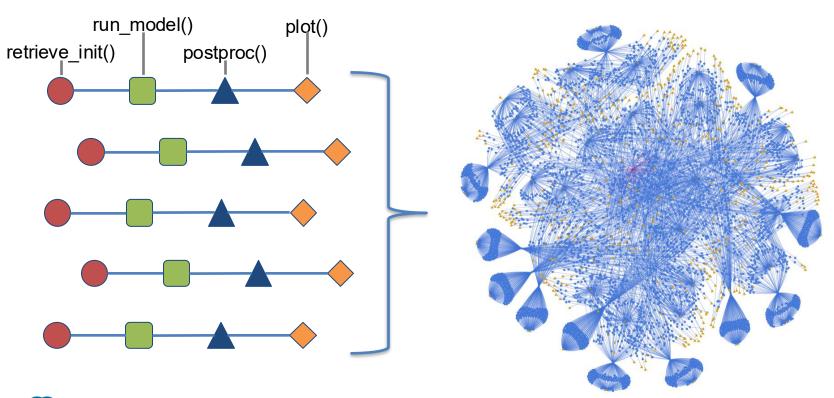


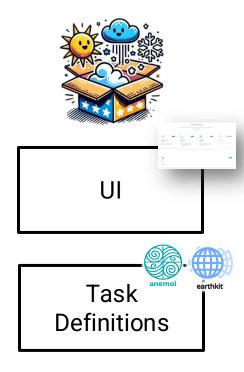




Orchestration

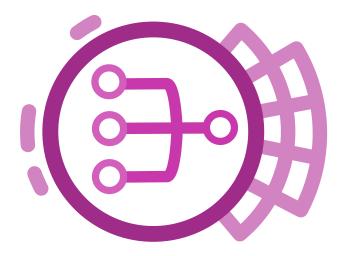
- Creation of a series of graphs from products specifications
- Construction of meta graph and scheduling

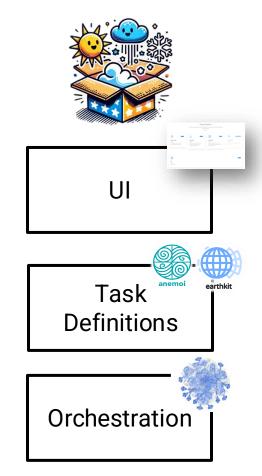




Execution Layer

- Executes the tasks graph on the computing infrastructure
- Data caching and persistence layer
- Optimizing for data colocation and transfer
- All of that in memory

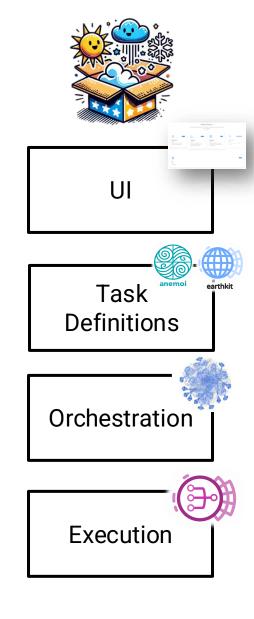






Putting it all together

- A complete forecasting system
- Capable of running any Anemoi-compatible model inference
- Runs ECMWF's full operational post-processing software
- Provides ECMWF production quality graphs and products

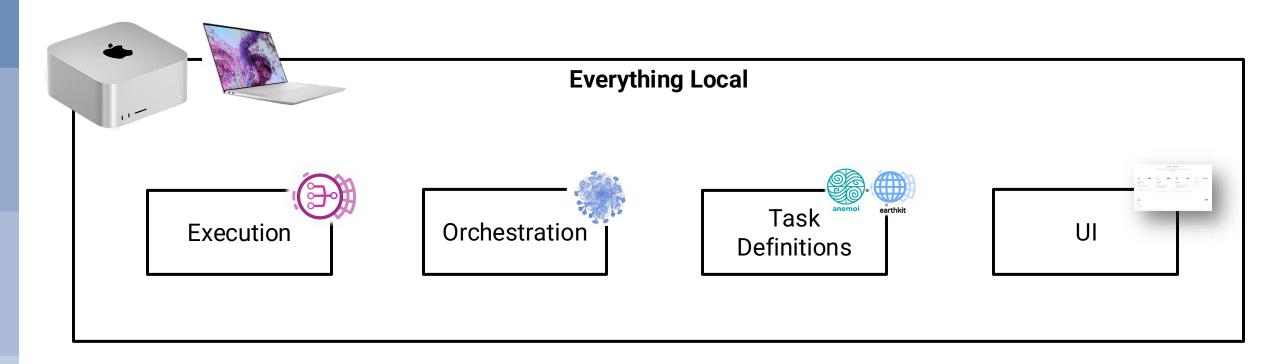




Forecast-in-a-Box – "How to run the box?"



Can be deployed in many ways...

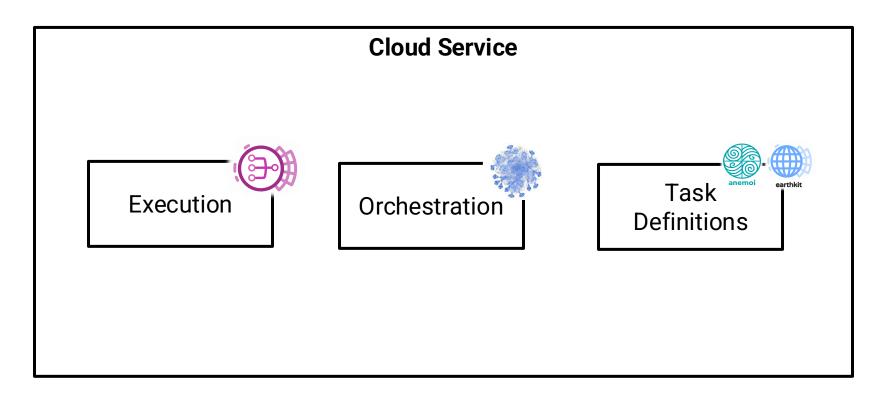


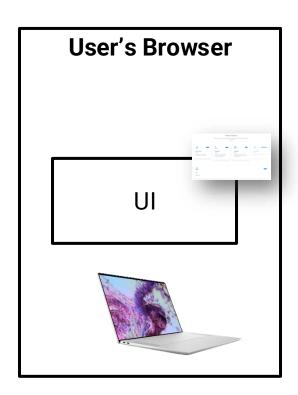


Forecast-in-a-Box – "How to run the box?"



Can be deployed in many ways...



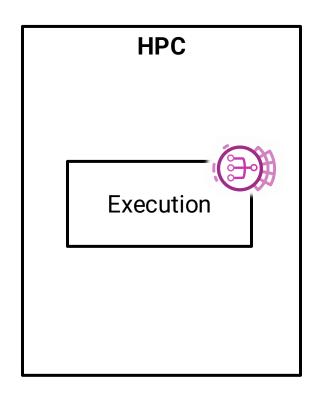


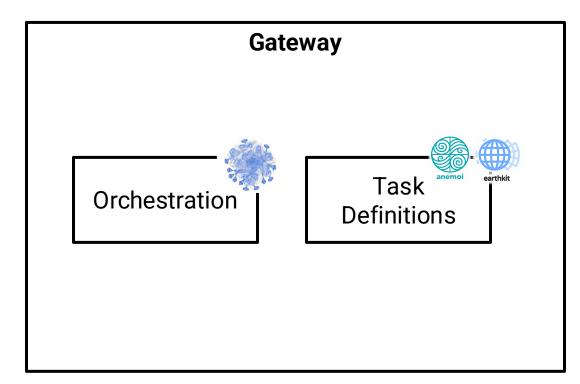


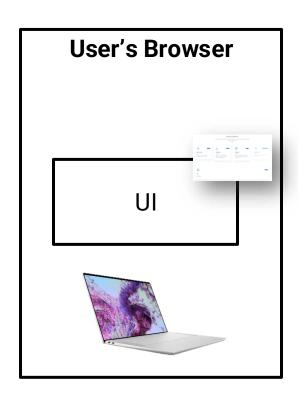
Forecast-in-a-Box – "How to run the box?"



Can be deployed in many ways...

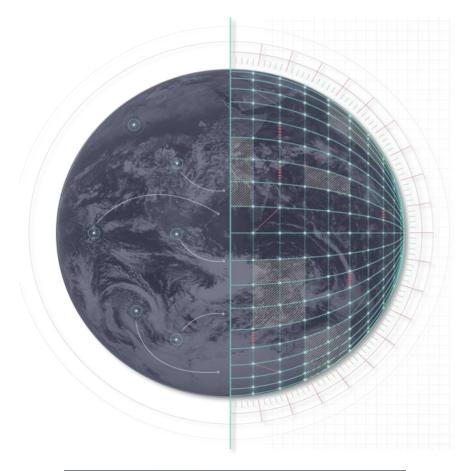








Ongoing Projects

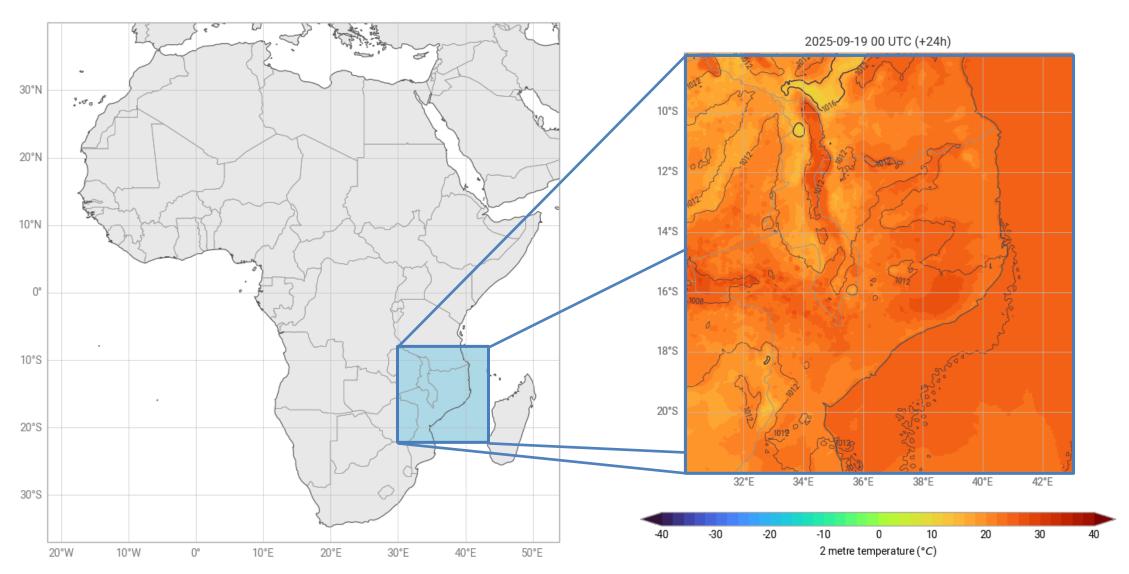








WMO/Malawi Pilot Project Lead by MetNorway







Bris in Malawi

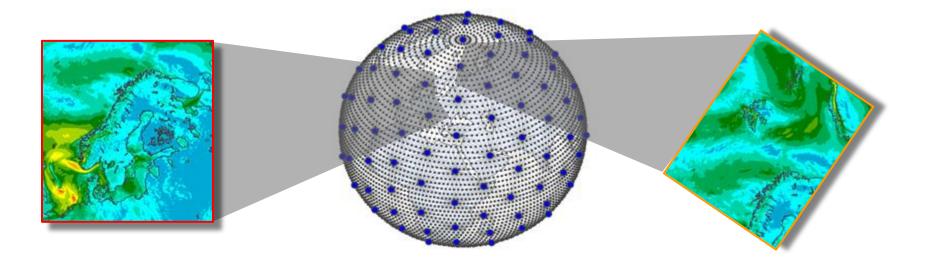
Running a localized model in Forecast-in-a-Box

24.09.2025

- Data-driven model for weather forecasting
 - Developed by Met Norway, in cooperation with several other institutions
- To be used with Met Norway's weather service yr.no

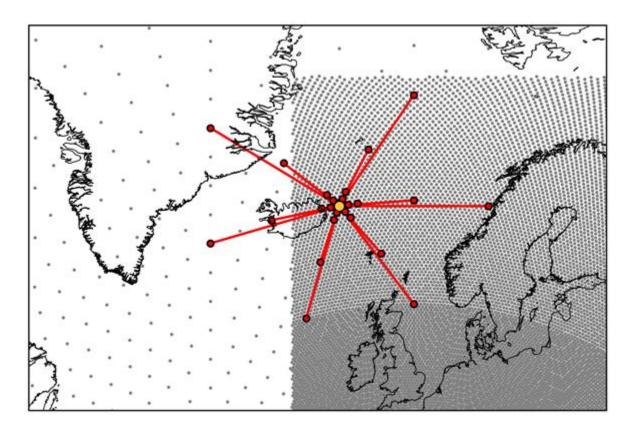


- Based on the Anemoi framwork
- A global AI-based model with high resolution over a focus area
- Any focus area is possible (eg. Malawi)



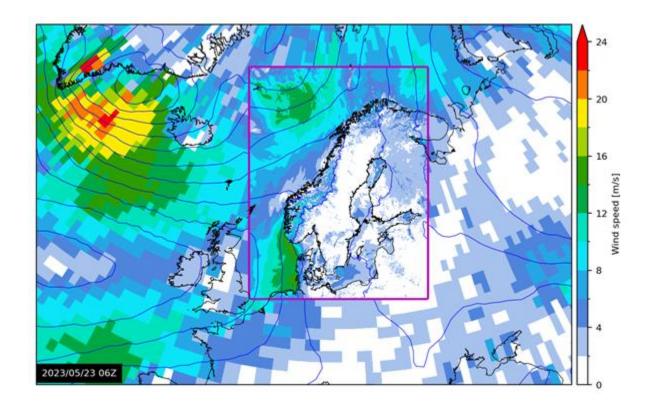


• A stretched-grid approach





• A stretched-grid approach





- 6-hour time resolution
 - WIP: interpolating to 1-hour resolution
- Training
 - 43 years of ERA5 data (31 km)
 - 3 years of MEPS data (2.5 km)
- Can run with any resolution









The vision

Deploy and evaluate AI weather prediction in Malawi, and explore how AI can strengthen early warning capabilities.

Work with WMO to scale up to additional NMHSs in LDCs and SIDS.



Bringing Bris to Malawi

- Move the domain of the Bris model to Malawi
- Run Bris in Forecast-in-a-Box
- Provide a physical box that can run all of this

This is a **pilot project**

We want to have a proof of concept before scaling up



Why should it work?



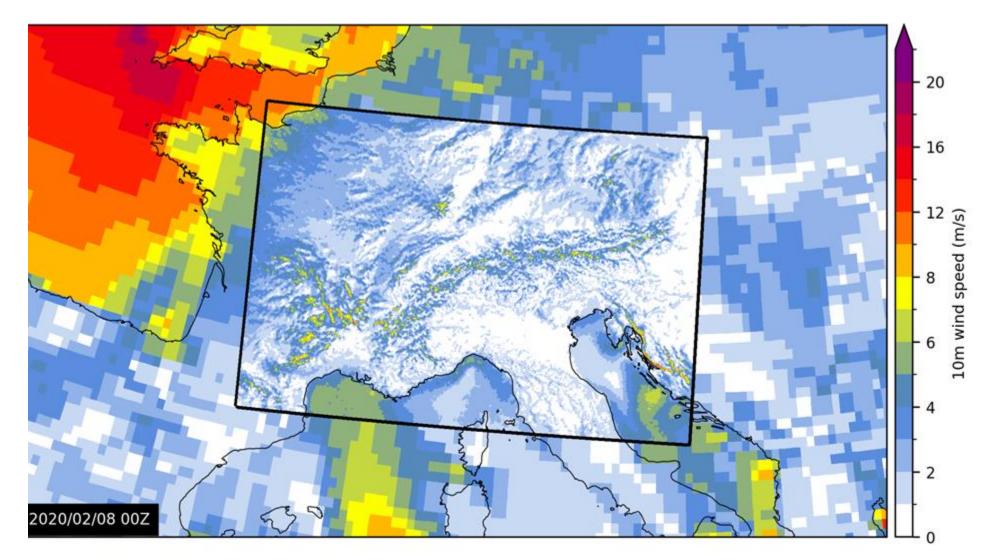


Why should it work?

- Bris has not been trained specifically for Africa, but
 - Training included global ERA5 data
 - Training for the nordics should be transferable to other parts of the world
- Moving the model has already been tested, with success



Why should it work?





Why should it work?

- What if it does not work?
 - Training with higher resolution for more areas of the world
 - Adding other parameters in training
 - Surface type (forest, city, etc)



Running it



Running it

- Physical box
 - Mac Studio M4 Max
 - 128 GB memory
 - 16 core CPU
 - 40 core GPU
 - Costs around \$5000





Running it

- Input data
 - ECMWF polytope service provides analysis data
 - Post-processed in the box
 - Downscaling
 - Adiabatic correction of surface parameters
 - In the future:
 - Locally provided analysis



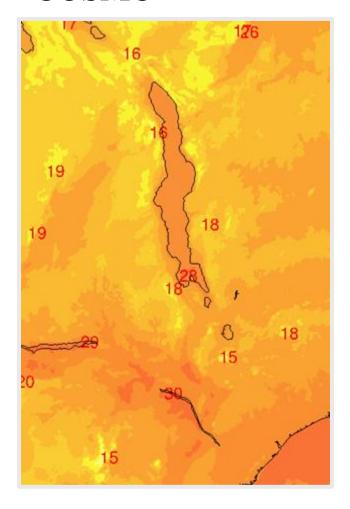
Status right now

- Working version
- Some issues remain
 - Physical box problems
 - No systematic verification of the forecast
 - Incomplete outputs

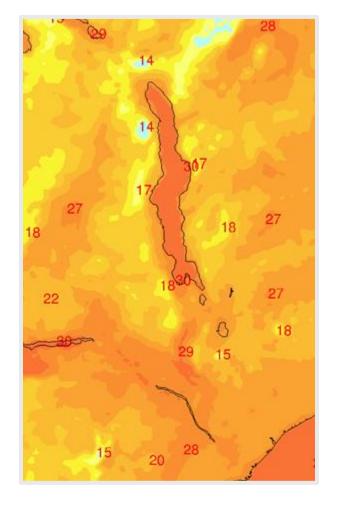


Results

COSMO

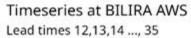


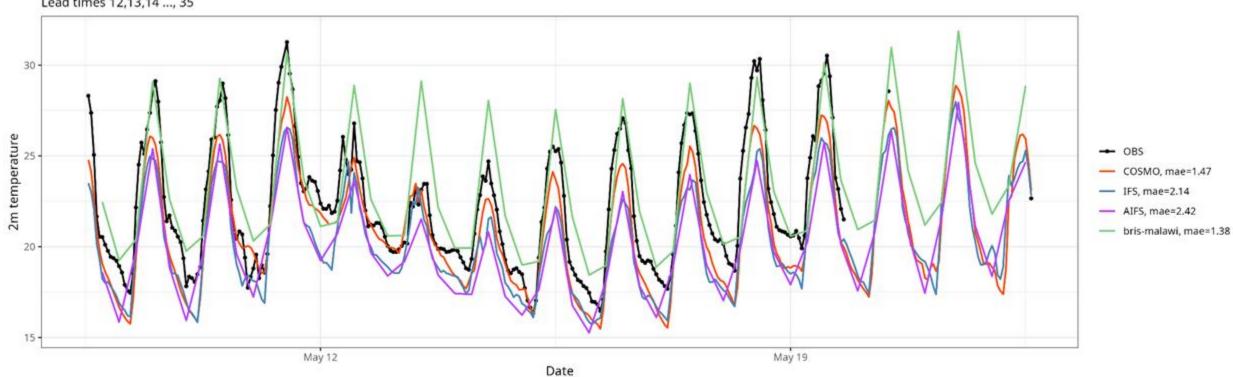
Bris





Results







Results - timings

Running a 10-day forecast for Malawi

- Mac studio
 - ~23 minutes
- NVIDIA A100, 40GB memory
 - ∘ ~20 minutes
- NVIDIA H200, 141GB memory
 - 80 seconds



Other options - cost overview

Costs presented are without VAT, in USD and do not include running costs, such as electricity, cooling, network, software etc.

Factors to consider:

- Different architectures e.g. ARM(GH200), x86_64 require different levels of adaptation.
- High density systems (>4ways systems) require water cooling
- time constraints, ensemble runs

Example air cooled systems:

Grace Hopper GH200, 5y warranty, total memory ~540GB, 100Gb network: ~USD 50k

4way system: 4xH200 nvlink, 5y warranty, 4 x144GB GPUmemory, 100Gb network: ~USD 160k

2way system: 2xH200 nvlink, 5y warranty, 2 x144GB GPUmemory, 100Gb network: ~USD 96k

2way system means: 2x H200s in one chassis, connected via nvlink.



Remaining work

- Verification
- Fixing Mac problems
- Working with DCCMS on further improvements
- Explore possibilities and benefits of using local data



The road ahead - after the pilot

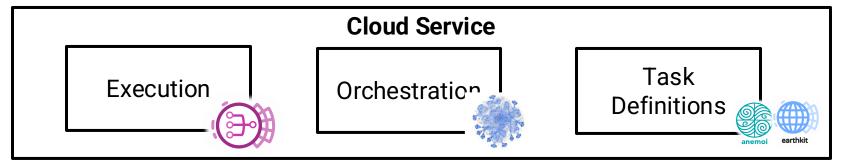
Working with WMO on making this available elsewhere

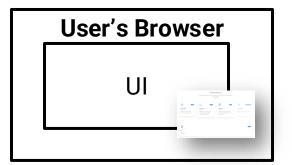


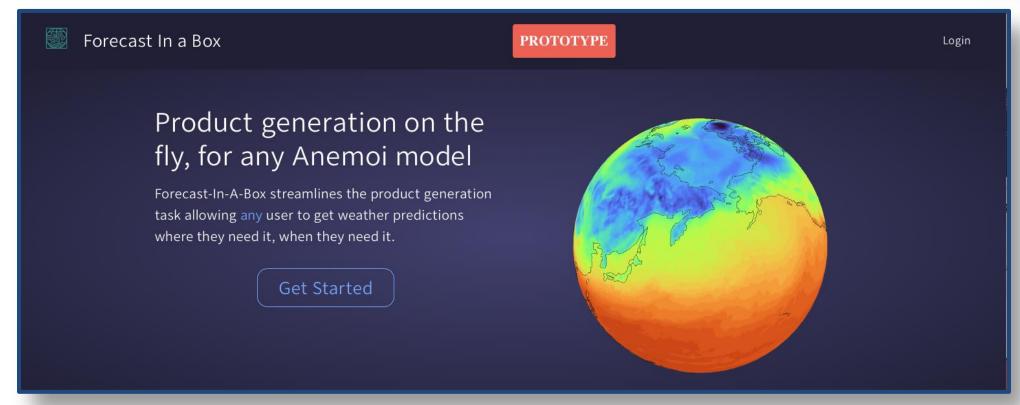


Norwegian Meteorological Institute

Demo









Future Work

- Forecast-in-a-Box
 - Contractor working on a redesigned UI
 - Scale-up the system using ensemble forecasts (Multi-GPUs, scaling, etc.)
- Improve user experience
 - Still a prototype
 - Allow users to define their own products
 - Improve Bris integration by allowing users to select directly their local domain
 - Build plugins for forecasters/verification
 - Workflow triggers
 - Push products to operational database (Zarr, Grib, NetCDF)



Thank you for your attention!







