



Anemoi

an open-source framework for developing data-driven
weather forecasts

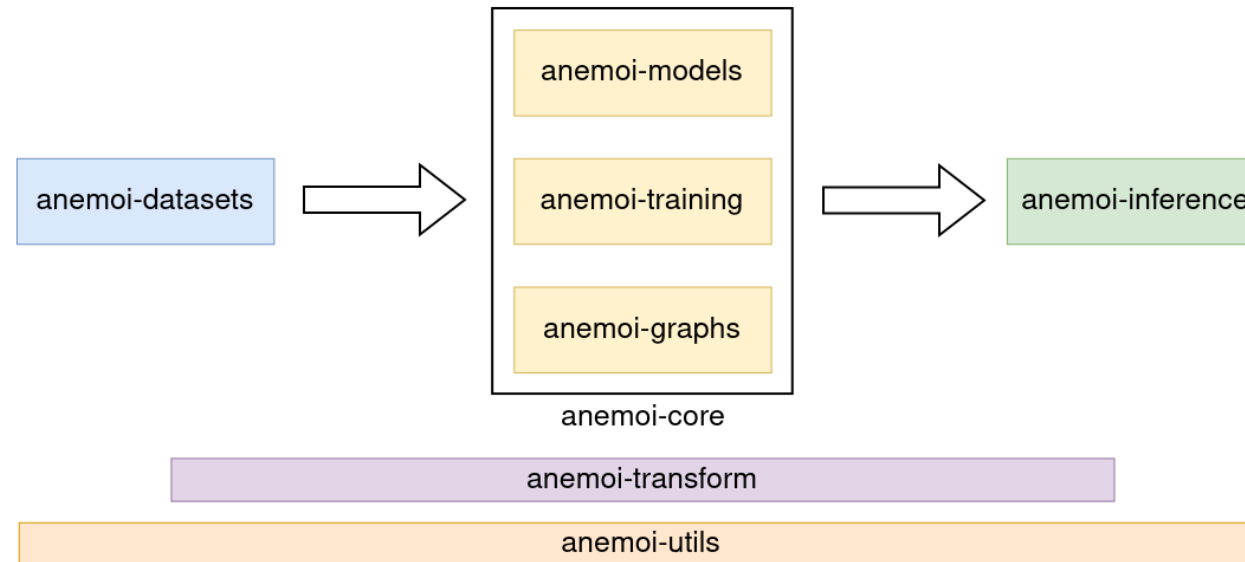
*Aaron Hopkinson – but presenting work developed by many colleagues
across ECMWF and the Anemoi community*

`aaron.hopkinson@ecmwf.int`



What is Anemoi?

- Open-source Python framework for building ML weather forecasting models
- Supports the whole workflow (datasets, training, inference)
- Modular, divided into several packages – designed to work well together
- Builds on top of established community tooling, e.g. PyTorch, Hydra, Zarr, earthkit
- Users do not need to write code: YAML configuration driven



What is Anemoi for?

Now:

- Global & local **data-driven weather forecast models** – initialised from physics-based data assimilation.
 - AIFS = An Anemoi configuration

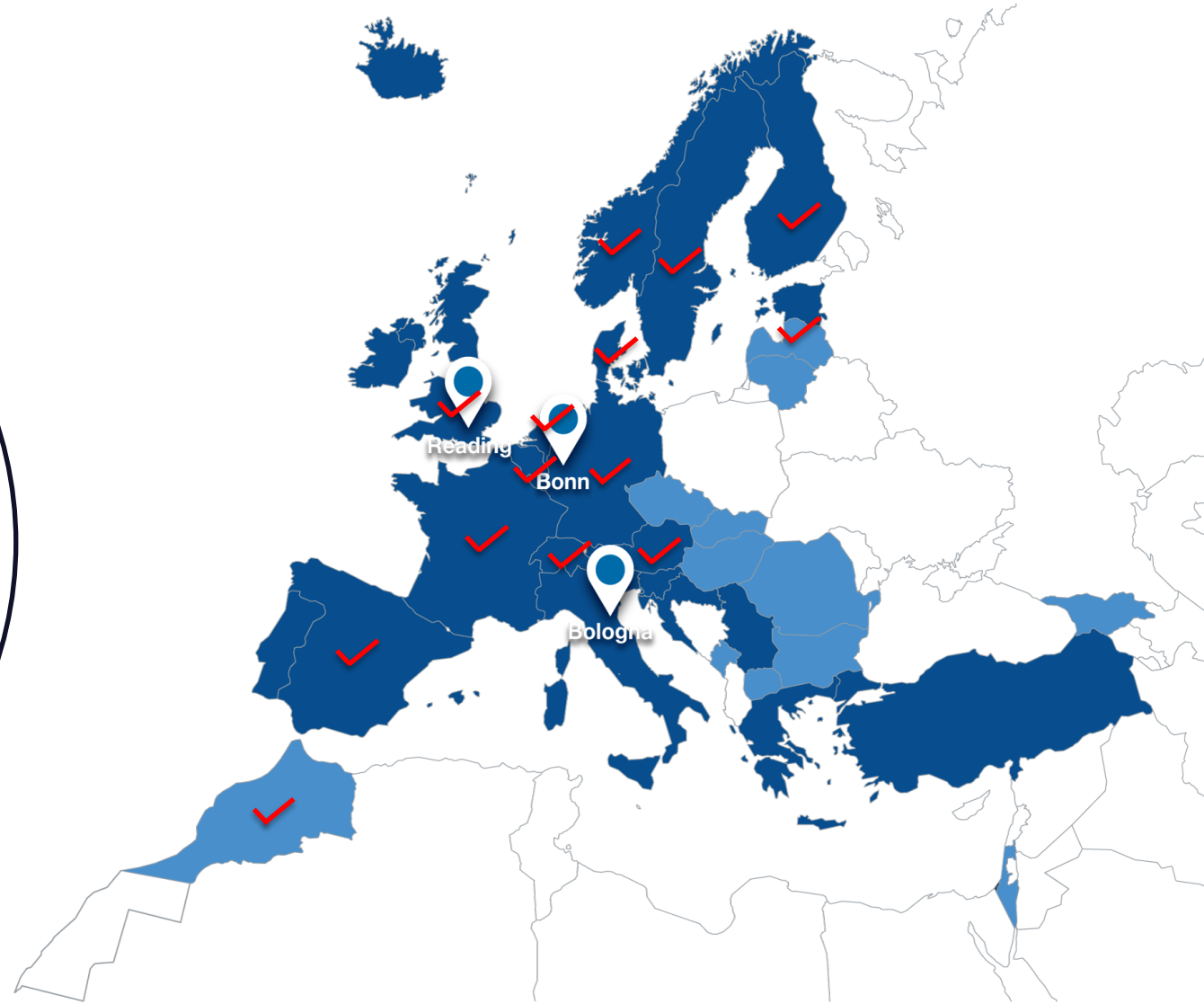
In development:

- End-to-end data-driven modelling (directly from observations)
- Other weather/climate problems with strong spatial complexity, e.g. downscaling.

What **isn't** Anemoi for?

- Arbitrary ML problems in weather and climate.
 - e.g. problems that are highly local in space.

Who is using Anemoi?



How do we work together?

- Development through GitHub: e.g. issues to plan, pull requests for implementations.
- An Anemoi technical group with representatives from across the contributing organisations meet weekly to steer Anemoi evolution, e.g. should contributions be merged?
- Regular (fortnightly) community meetings to share results, discuss development.
- Focussed meetings on specific topics as required, e.g. discussing observations data.
- Including regular in-person meetings.

Member State Pilot project

- 14 Member States and ECMWF.
- Working together across 4 work packages.
- Sharing technical and scientific developments on:
 - Data-driven modelling.
 - Data-driven ensembles.
 - Data-driven data-assimilation.
 - MLOps and infrastructure

Part of:

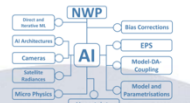
EUMETNET E-AI initiative

“Artificial Intelligence and Machine Learning for Weather, Climate and Environmental Applications” (E-AI) Optional Programme

High level objectives:

- To enhance the collaboration of European NMHSs and external partners in the area of AI/ML in weather, climate and environment.
- To share the developments which take place under E-AI using a commonly-used permissive open-source licence.

Coordinating Member for E-AI: DWD (Roland Potthast)



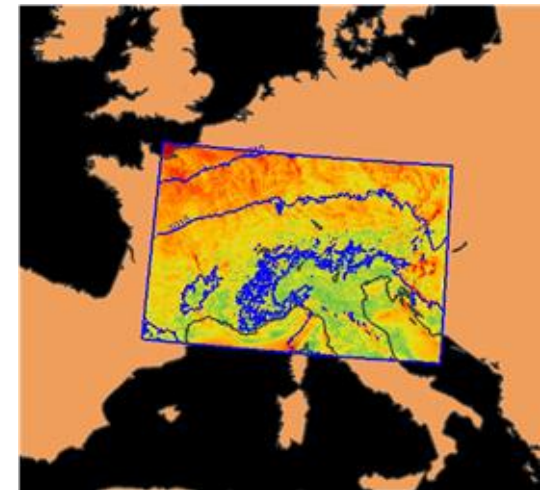
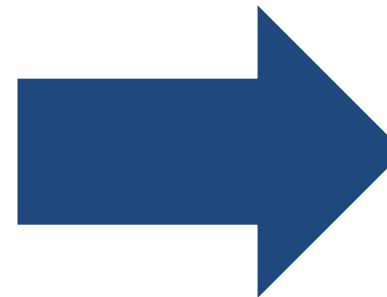
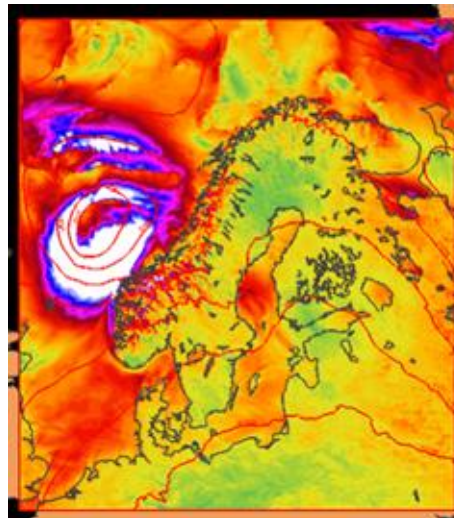
EUMETNET



ECMWF ML Pilot Partners

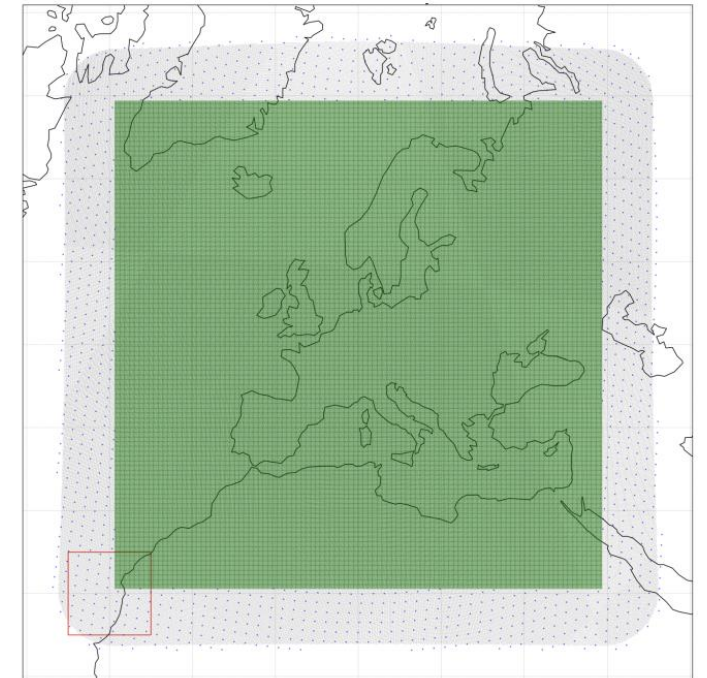
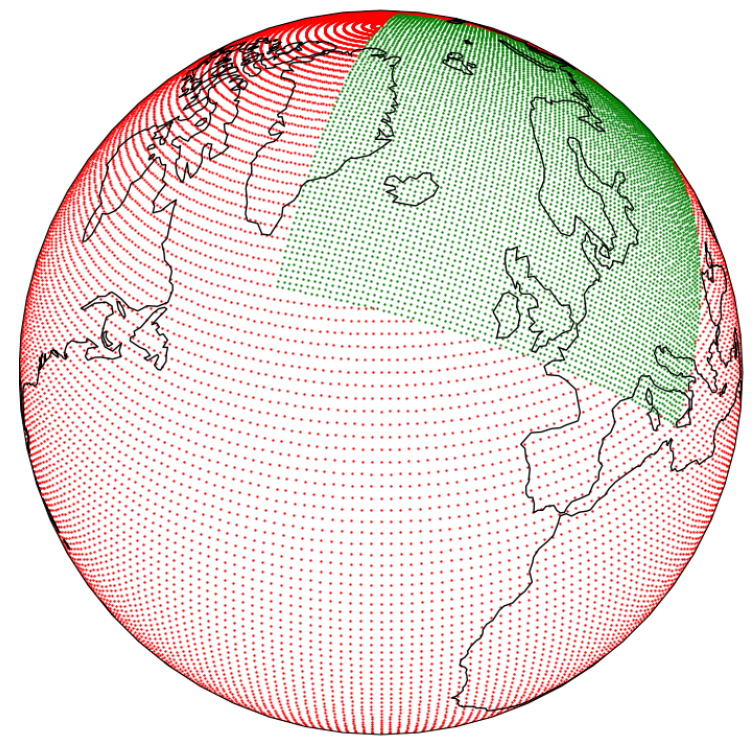


Example of cross-org collaborations
Testing the ability of Bris (trained on Nordics)
to forecast over Switzerland



What are Member States doing with Anemoi?

- Most exploring **developing regional models**, using a mixture of CERRA (5km European reanalysis) and national datasets (e.g. operational analysis or national reanalysis).
- Started with **deterministic** and now starting to explore **ensemble** models.
- **All** using variants of GraphTransformer for each of the Encoder/Processor/Decoder.
- Multiple paths to regional modelling:
 - **Stretched-grid** - global model with more resolution over a regional domain
 - Requires global initial conditions but no boundary forcing.
 - Used by most Member States.
 - Data-driven **LAM** – fully regional model.
 - Requires boundary forcing and high-resolution initial conditions.



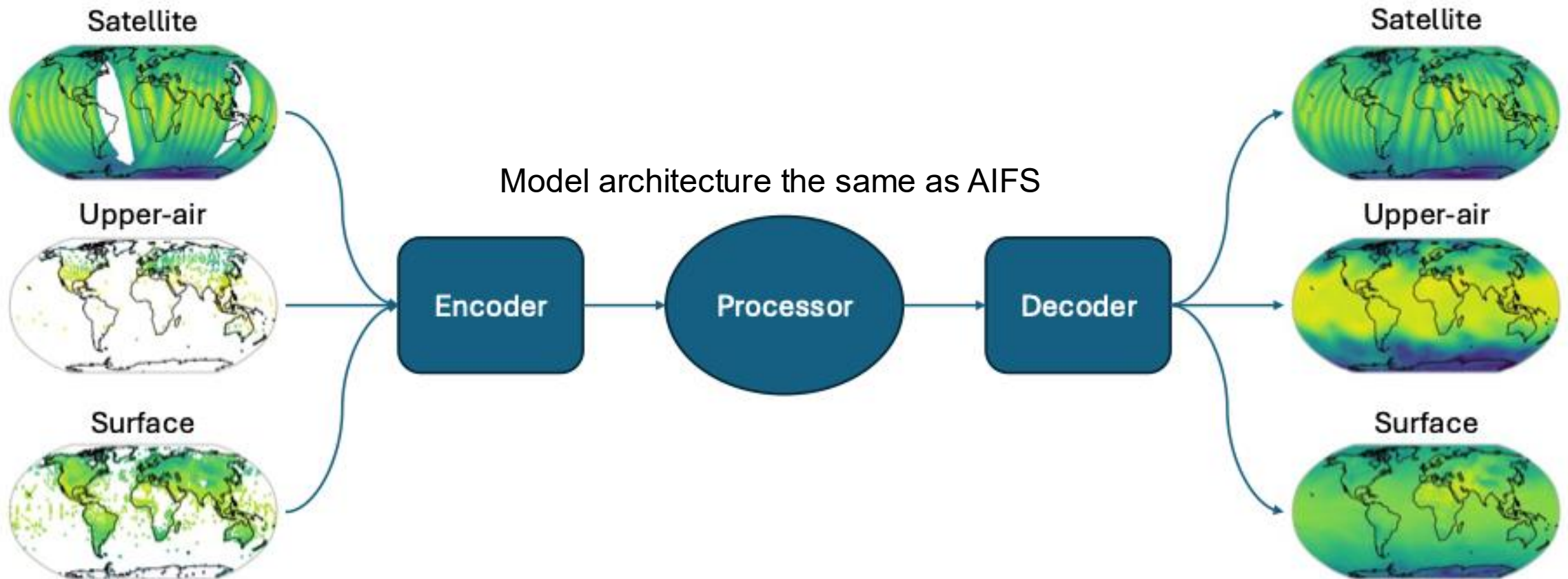
Some recent work...

Supporting observations in Anemoi

- Begun to bring support for training directly using observations datasets (rather than reanalysis)
- anemoi-datasets recently added support for ingesting irregular observations in a “tabular” format
- Resultant dataset is a Zarr, with I/O optimised for model training.
- Support different data sources, e.g. BUFR, CSV, ODB
- Being used in two different strands of work:
 1. Training as if a “regular” gridded dataset (static graph) – see next slides.
 2. Keeping as irregular/unstructured data – adding support for “dynamic graphs” in modelling framework (GraphDOP approach) – earlier stage of development.

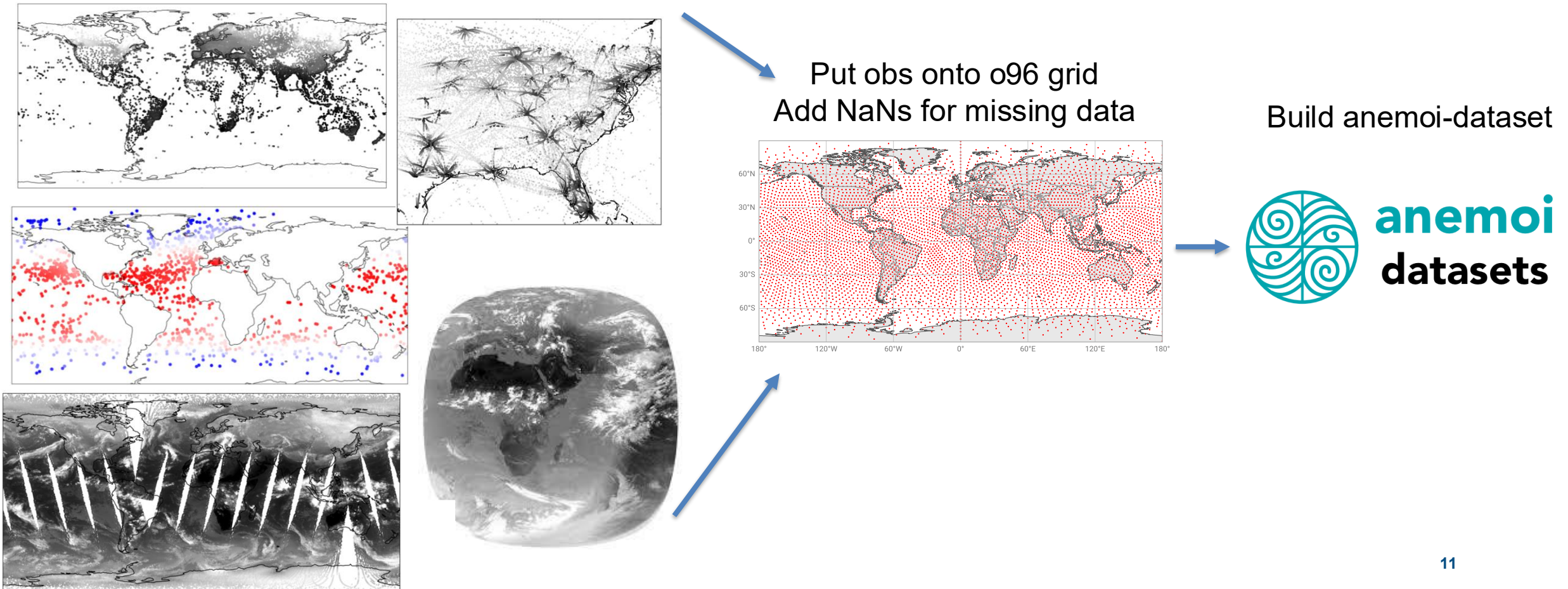
First examples of training with observations in Anemoi

- Training ML models directly from observations using Anemoi
- Architecture for this example very similar to the AIFS with some tweaks

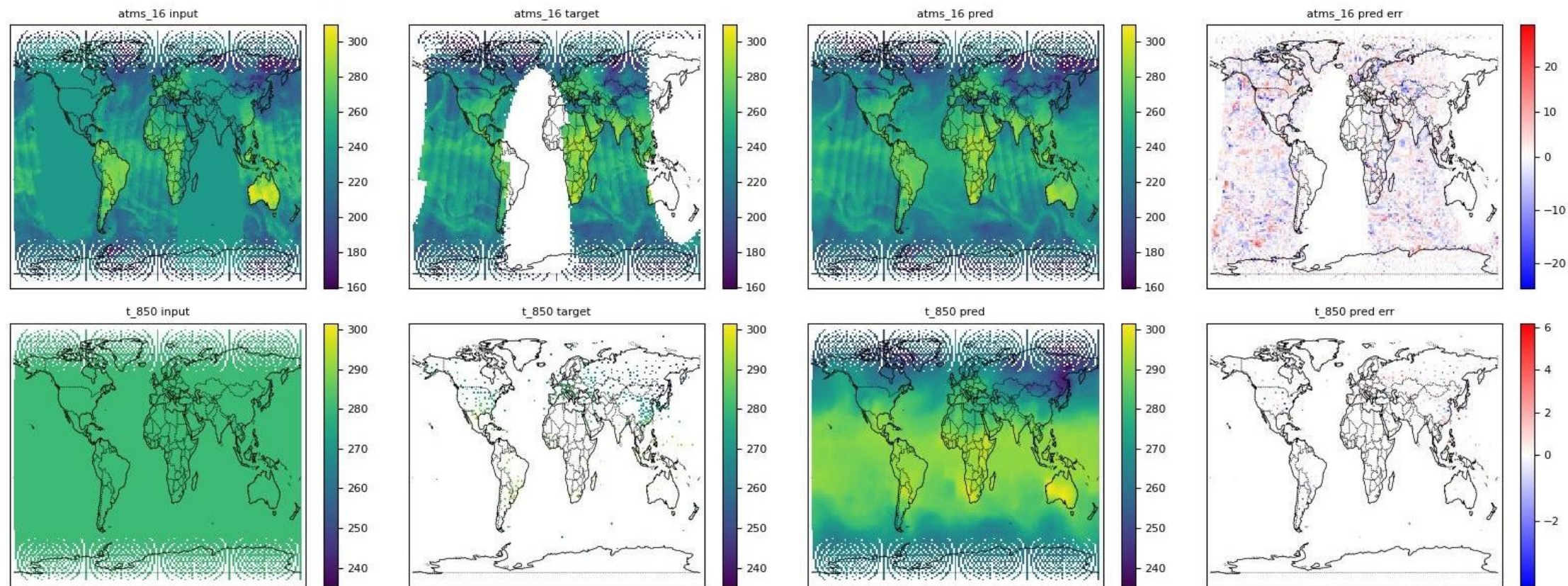


First examples of training with observations in Anemoi

- Putting obs on the same resolution grid in 6-hour slices using filters in anemoi-transform
- Treat them as additional variables under dataloader
- Slight updates to NaN-handling to allow training/generalising in space (and new “InputOnlyImputer”)



Training Example



- Example model trained with small number of datasets

Summary

- Anemoi is supporting collaborative development of data-driven weather models across European meteorological organisations
- Open-source project – open to contributions from any organisation or individual
- Being used operationally at ECMWF for AIFS
- Lots of exciting features under development
- Learn more:
 - anemoi.readthedocs.io
 - github.com/ecmwf/anemoi