Norwegian
Meteorological
Institute

Bris

A high-resolution data-driven machine learning model for weather forecasting

Thomas Nipen, Ivar Seierstad, Håvard Homleid Haugen, Magnus Sikora Ingstad, Even Nordhagen, Aram Salihi, Paulina Tedesco, John Bjørnar Bremnes

5th ACCORD ASW, 31 March - 4 April, 2025, Zalakaros, Hungary.

Basics of initial-state machine learning modelling

Objective: Learn a time step

- Mapping from current/previous atmospheric state(s) to next state
    - mapping specified by neural network
    - size of state variable: ∽100 million
    - unknown parameters in neural networks: ∽100 million
    - size of time step: 6 hours (standard for global medium range models)

Basics of initial-state machine learning modelling

Objective: Learn a time step

- Mapping from current/previous atmospheric state(s) to next state
  - mapping specified by neural network
  - size of state variable: ~100 million
  - unknown parameters in neural networks: ~100 million
  - size of time step: 6 hours (standard for global medium range models)

- Unknown parameters learned by solving an optimisation problem
  - need a mapping that is good for all initial states
  - use of long datasets like 40 years of ERA5
  - mean square error (weighted) commonly applied loss function
  - stochastic gradient descent

Deep learning

Basics of initial-state machine learning modelling

Objective: Learn a time step

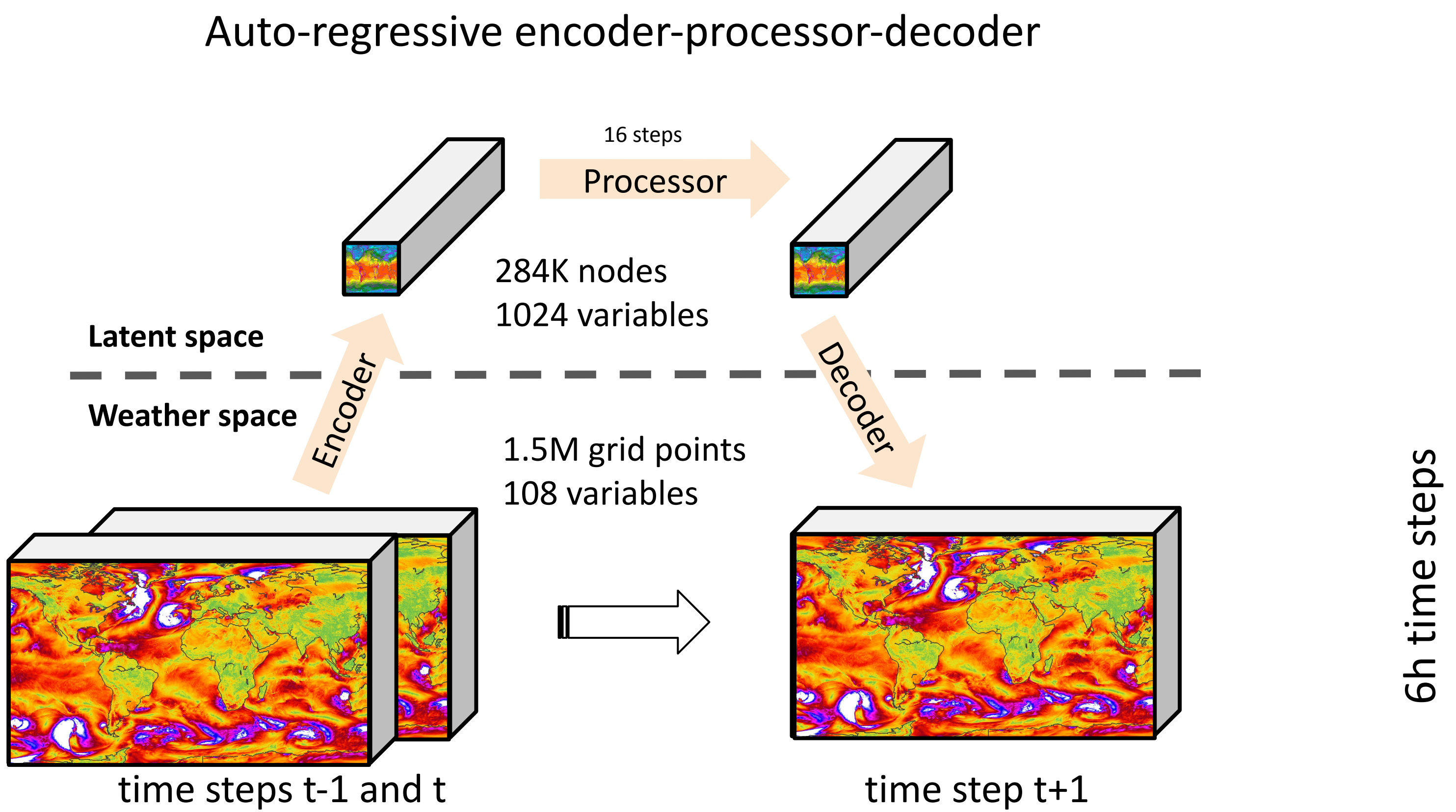
- Mapping from current/previous atmospheric state(s) to next state
  - mapping specified by neural network
  - size of state variable: ~100 million
  - unknown parameters in neural networks: ~100 million
  - size of time step: 6 hours (standard for global medium range models)

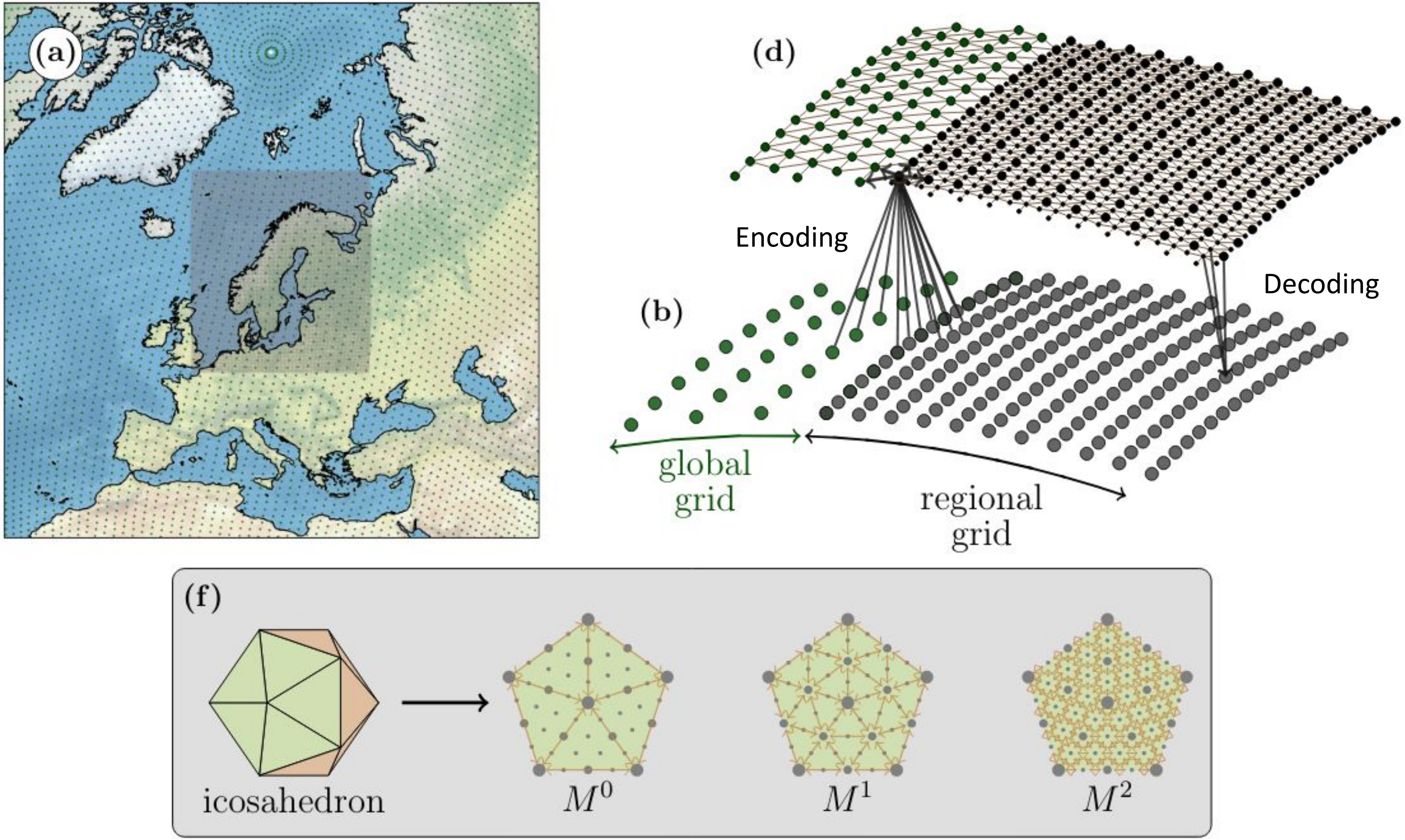
- Unknown parameters learned by solving an optimisation problem
  - need a mapping that is good for all initial states
  - use of long datasets like 40 years of ERA5
  - mean square error (weighted) commonly applied loss function
  - stochastic gradient descent

Deep learning

- Learned time step applied to generate longer forecasts (auto-regression)

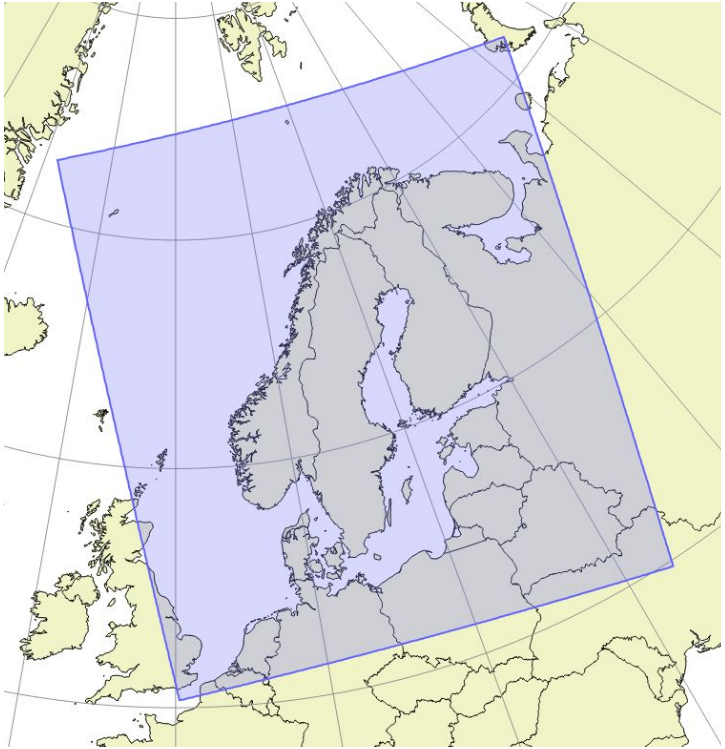
Bris model architecture




(a)
(d)
Encoding
Decoding
(b)
global
grid
regional
grid
(f)
icosahedron
M^0
M^1
M^2

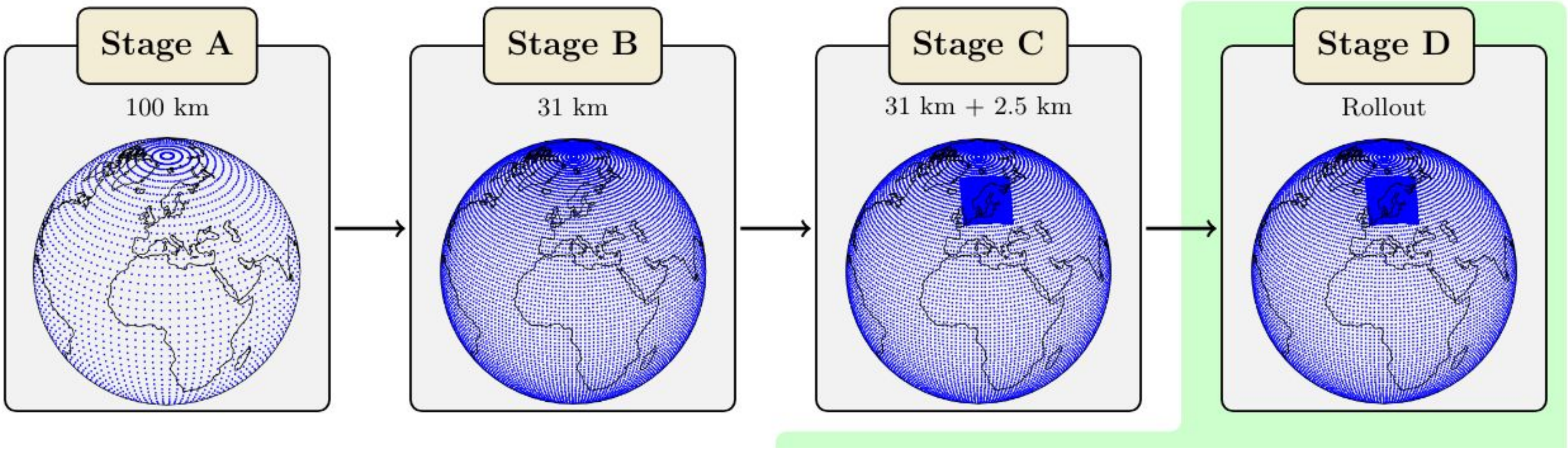
Training data: ERA5 (40 years) + MEPS analyses (3 years)

6-hour time resolution



- Input and output variables
  - T, U, V, W, Z, Q at 13 pressure levels
  - 13 single level fields (T0, T2, Td2, U10, V10, MSLP, surface pressure, cloud area fraction, …)
- Additional diagnostic/output variables
  - precipitation, wind gust, visibility, fog
- Forcing input variables
  - solar insolation
  - elevation, land-sea mask
  - longitude, latitude
  - time of year and day

Training in four stages



Training and prediction

Training:

250 million trainable parameters

10 TB of training data

Global pre-training: approx 4000 GPU hours
High resolution fine-tuning: approx 2000 GPU hours

Generating forecasts:
NVIDIA H200 approx 2 min 10 day forecast
NVIDIA A100 approx 4 min 10 day forecast

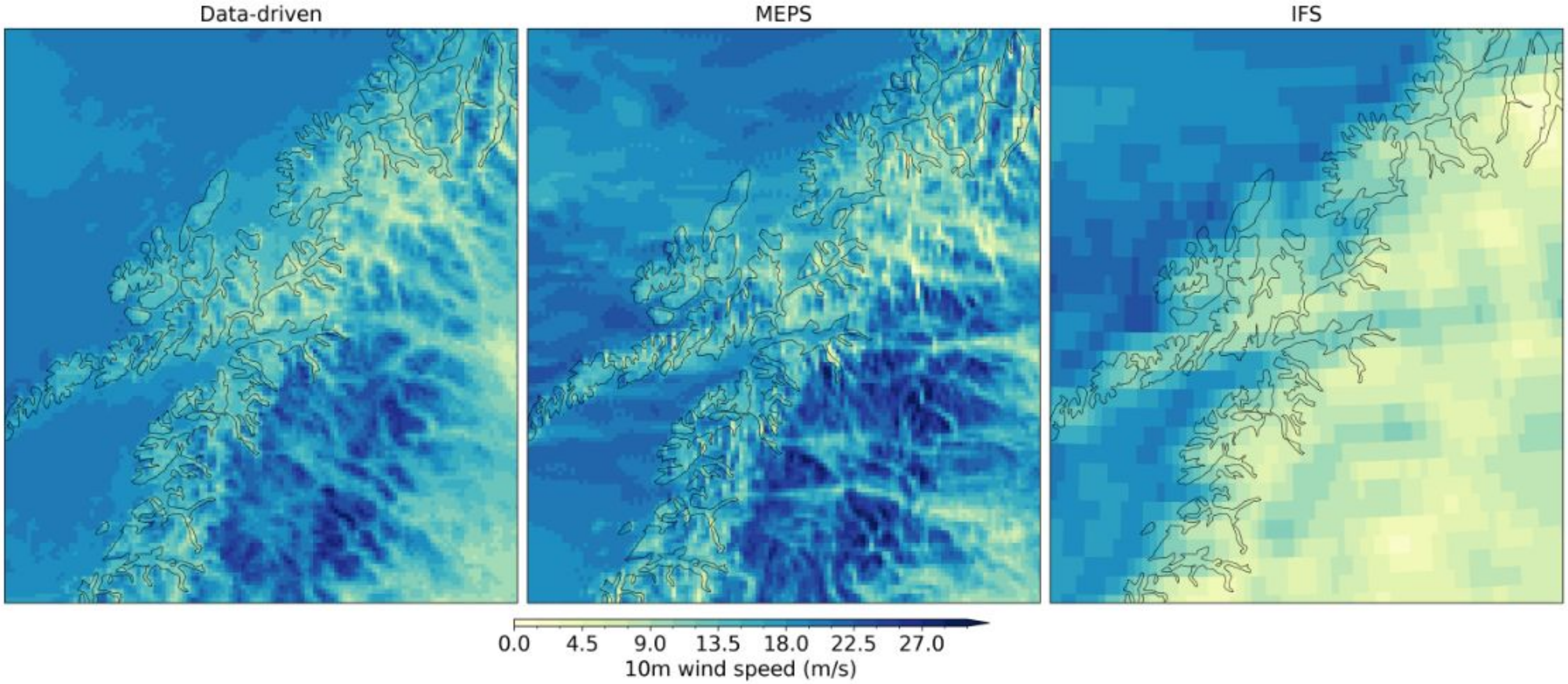
LUMI supercomputer

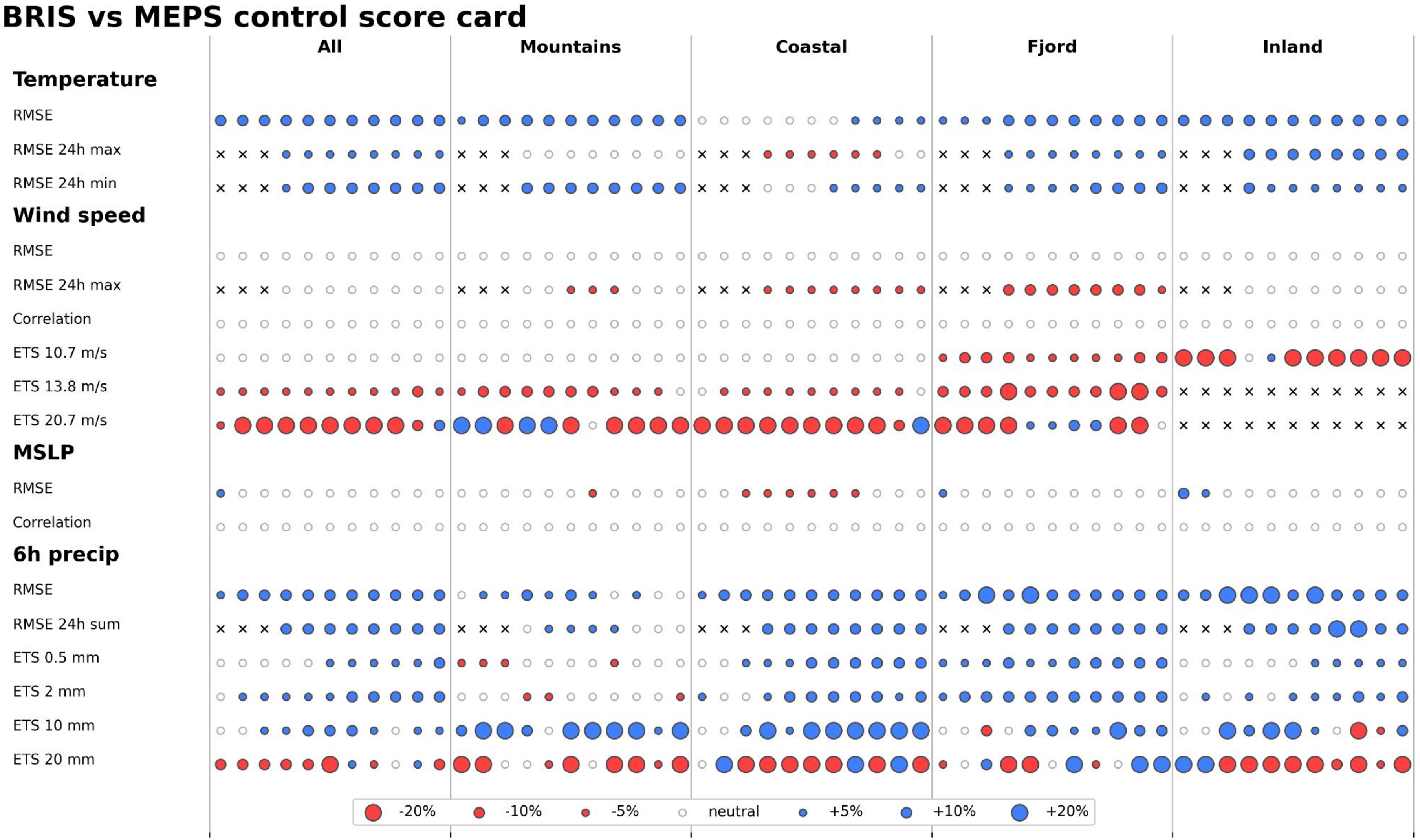
What can we expect based on statistical knowledge?

Optimising with respect to square error implies

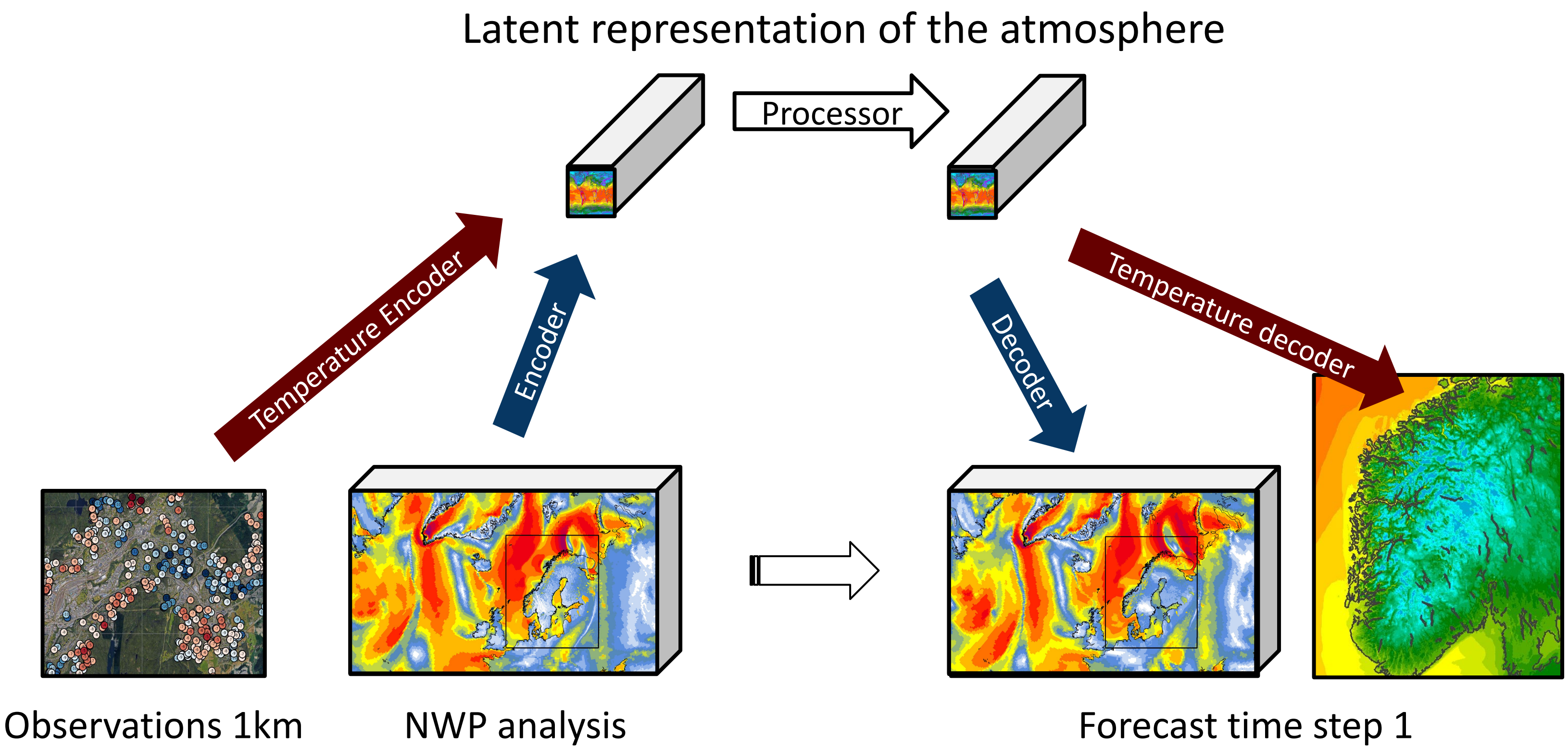
- model predicts/estimates distribution means, not a realisations
    - less variation than the reference data (here reanalyses)
    - extremes will on average be under-estimated
    - fields are spatially smoother than the reference
- interpretation as distribution means only for the first 6h-time step
- for longer lead times variability is larger than distribution means
    - with increasing lead time
        - distribution mean → climatological mean value
        - model has about the same variability as for the first time step(?)

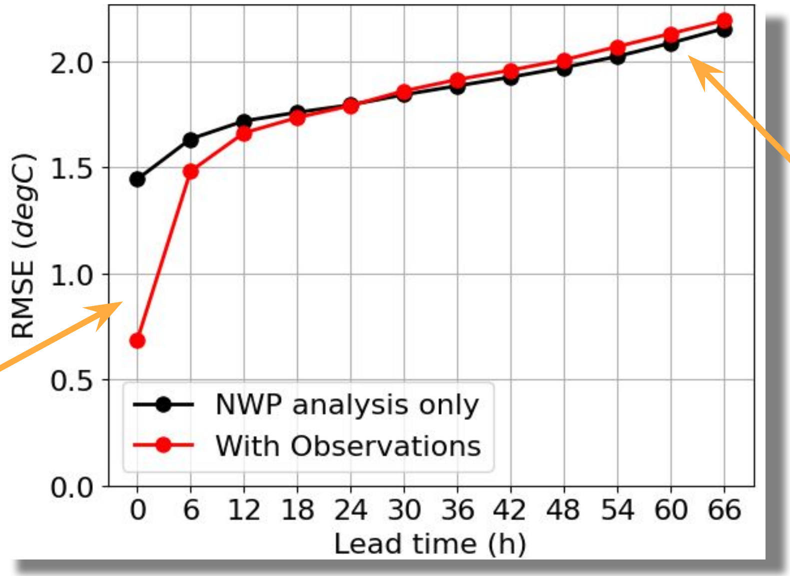
Wind speed example




BRIS vs MEPS control score card
All
Mountains
Coastal
Fjord
Inland
Temperature
RMSE
RMSE 24h max
RMSE 24h min
Wind speed
RMSE
RMSE 24h max
Correlation
ETS 10.7 m/s
ETS 13.8 m/s
ETS 20.7 m/s
MSLP
RMSE
Correlation
6h precip
RMSE
RMSE 24h sum
ETS 0.5 mm
ETS 2 mm
ETS 10 mm
ETS 20 mm
-20%
-10%
-5%
neutral
+5%
+10%
+20%

Integrating additional gridded fields



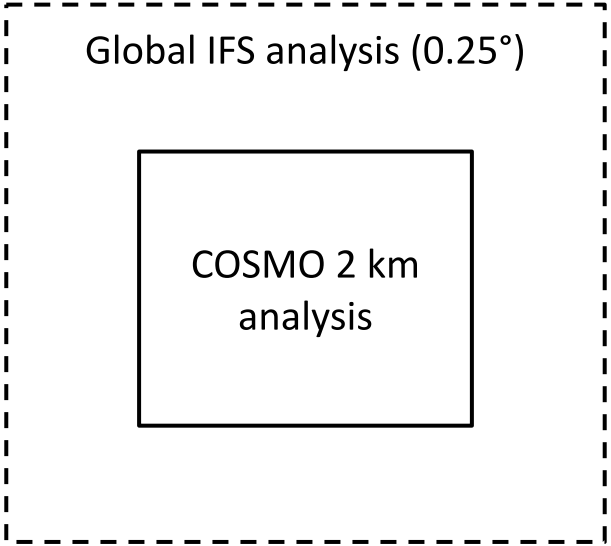


Improvement from fusing observations

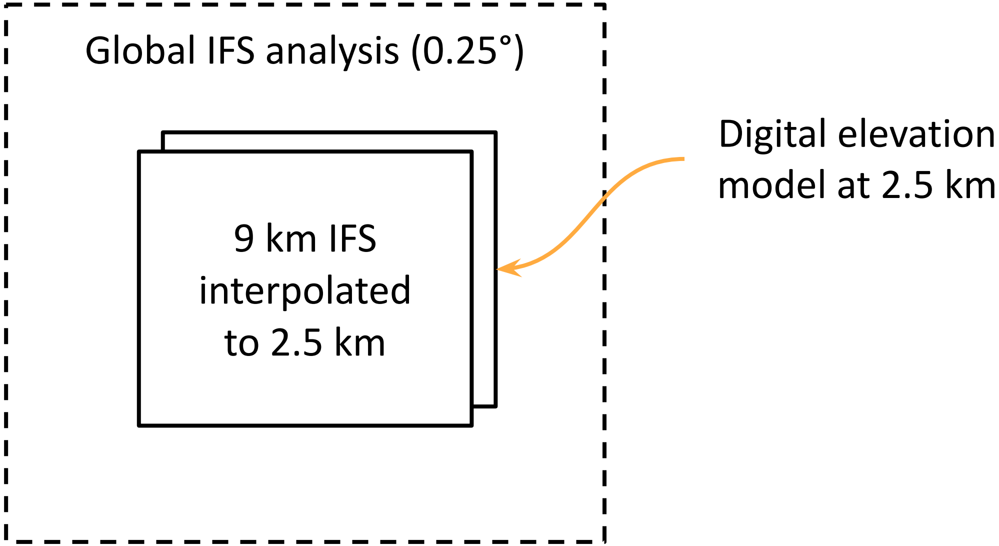
Still some work to do…

Does Bris generalise to other regions?

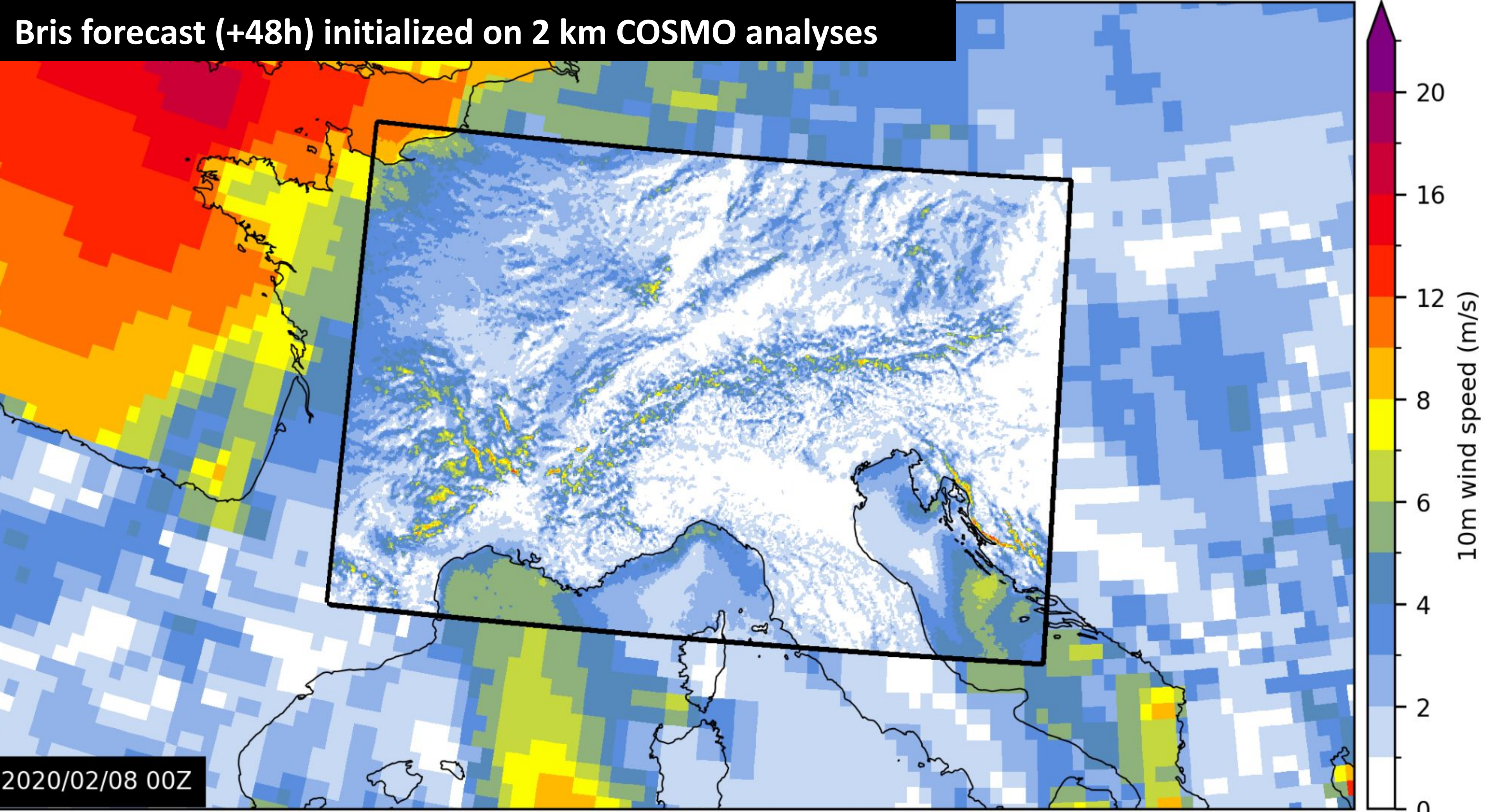
Experiment 1: Bris initialized from MeteoSwiss analyses



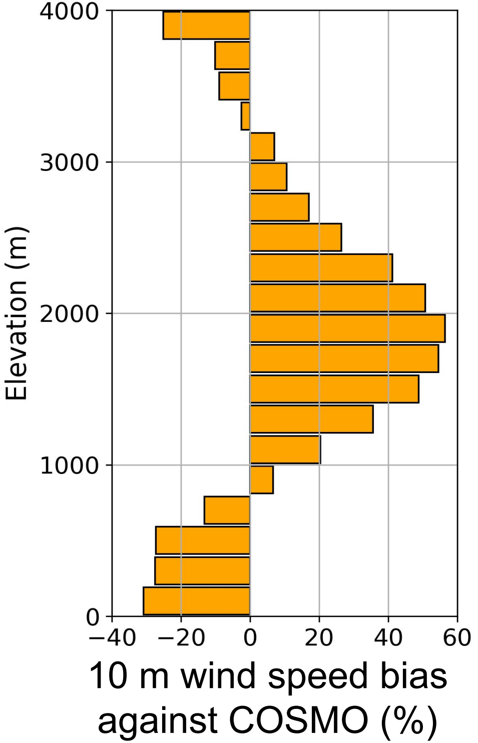
Experiment 2: Bris initialized entirely from IFS analyses

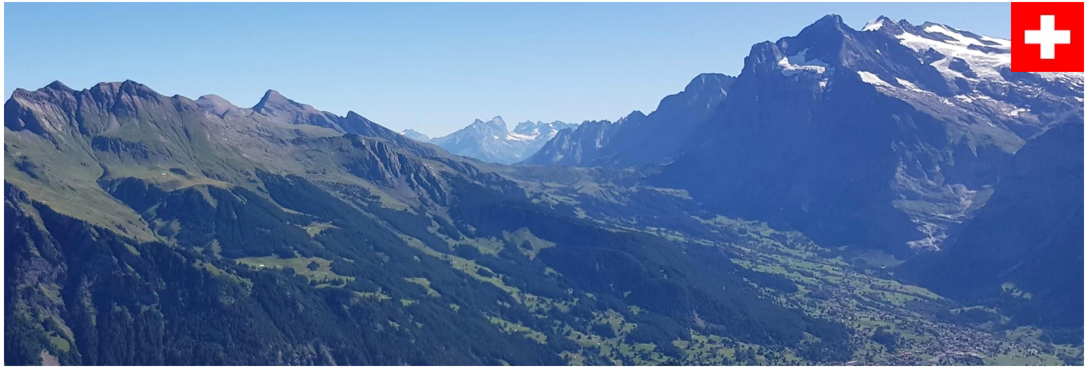


Bris forecast (+48h) initialized on 2 km COSMO analyses



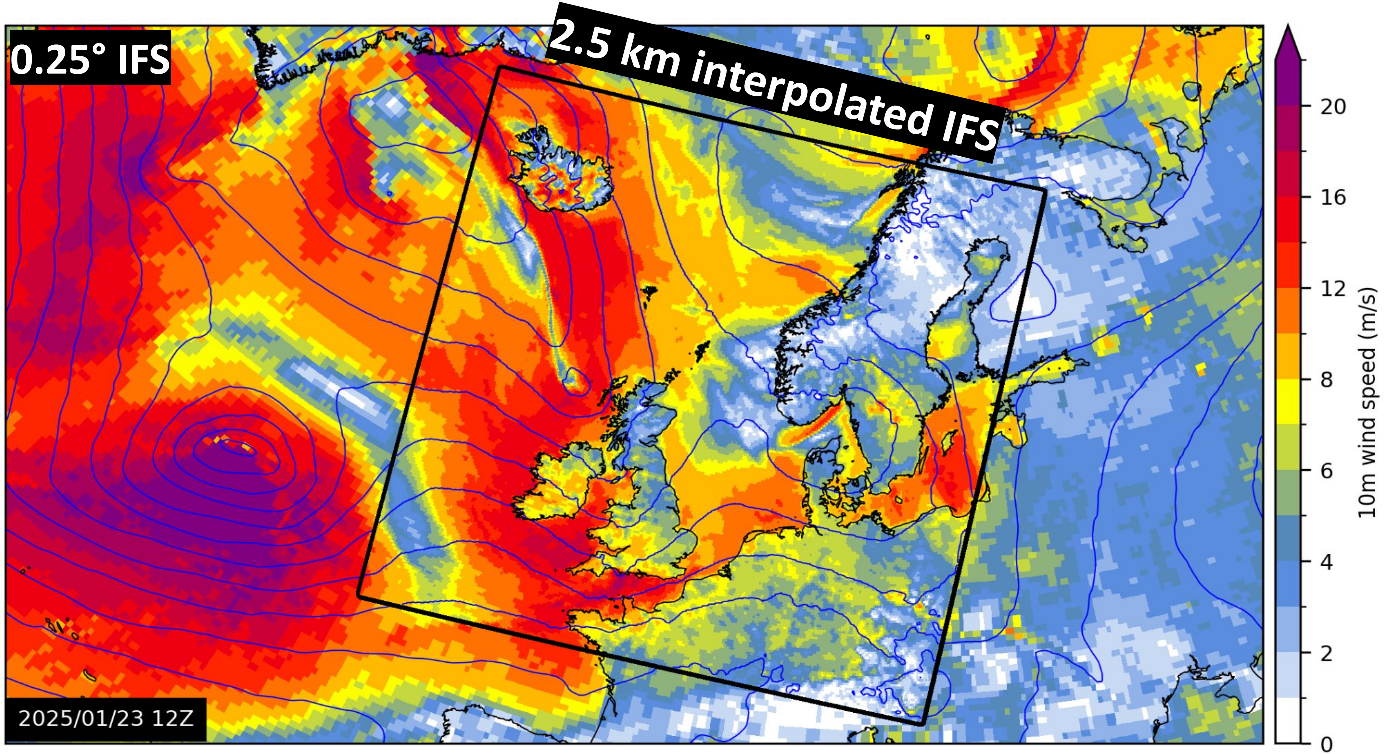
*Thanks to MeteoSwiss for providing data

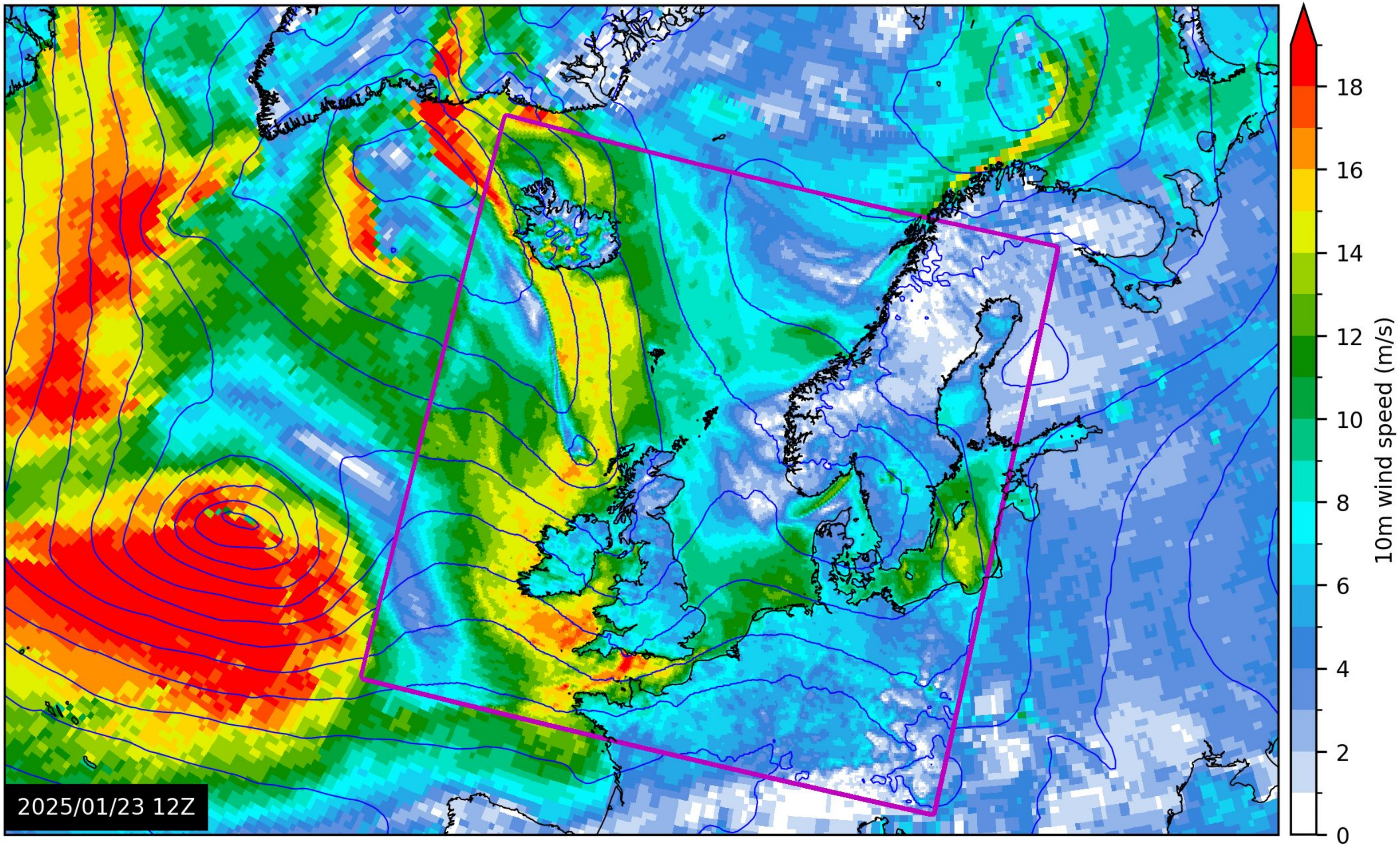




*Thanks to MeteoSwiss for providing data

EOWYN - initialised with downscaled IFS





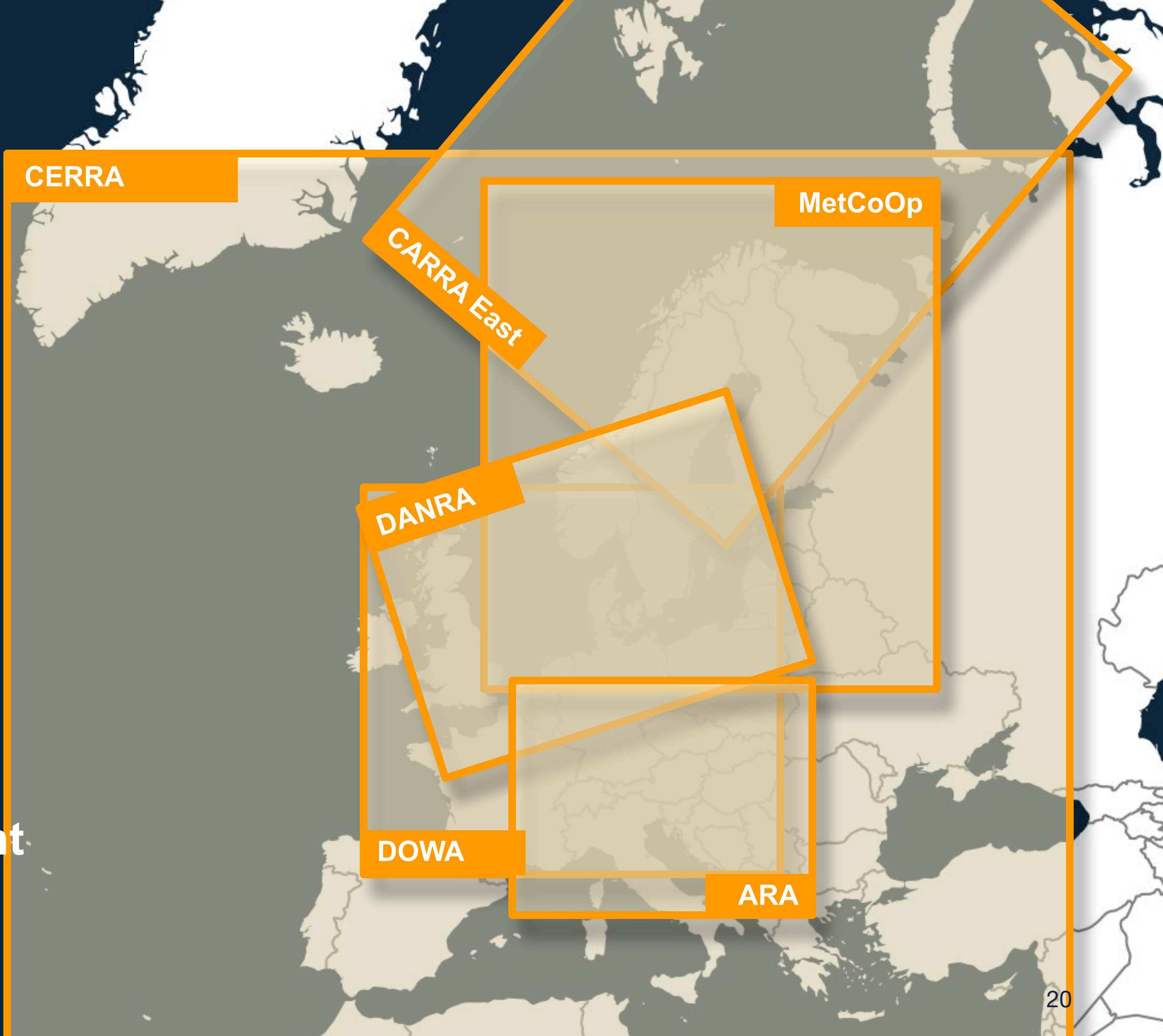
2025/01/23 12Z
10m wind speed (m/s)
0
2
4
6
8
10
12
14
16
18

Training data

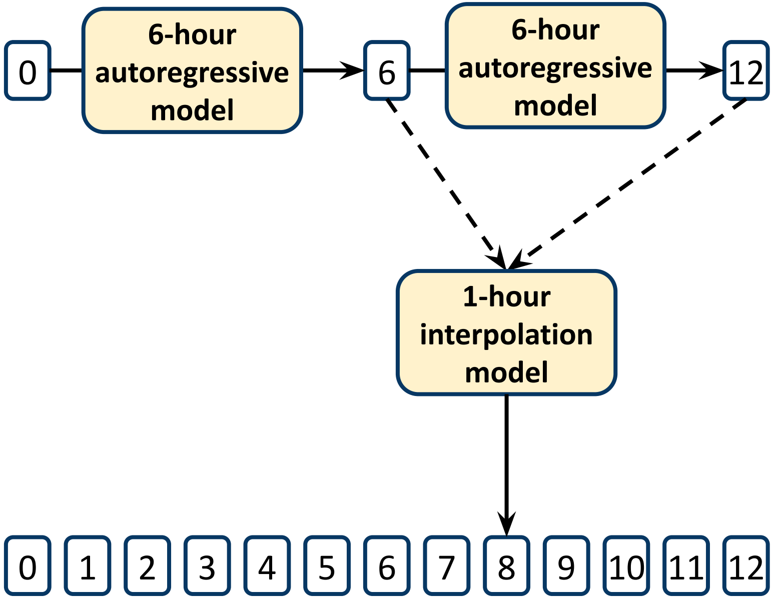
1 **European reanalyses** 5.5 km

5 **Regional reanalyses** 2.5 km

400+ **DE330 extreme event simulations** 200-750m

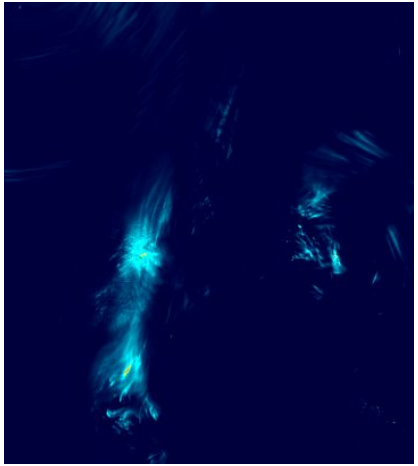


Hourly resolution by time interpolation model

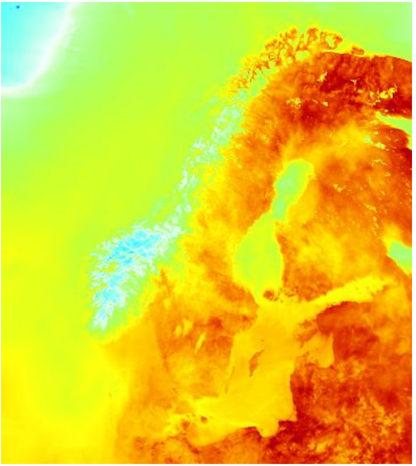


Trained on forecast data, uses full state of model for 0 and 6 hours

6h precip



2m temperature



Available as a feature in anemoi-core

Magnus Sikora

Ongoing work and next steps

- Set up a Bris model for **tests** in downstream applications (yr.no) before summer
  - include all variables needed
  - hourly resolution
  - temperature on 1-km grid
  - output in netcdf format (as for MEPS, IFS)

- Ensemble generation
  - optimisation based on CRPS
  - scenarios less smooth than in deterministic Bris as expected
  - probabilistic verification scores already good
  - challenges with spatial coherence

- More verification
  - assessment of model features

- More high-resolution training data
  - CARRA2

- Multi-domain training
  - joint work in DE330

- Data-assimilation
  - joint work in Machine Learning Pilot Project

https://arxiv.org/pdf/2409.02891