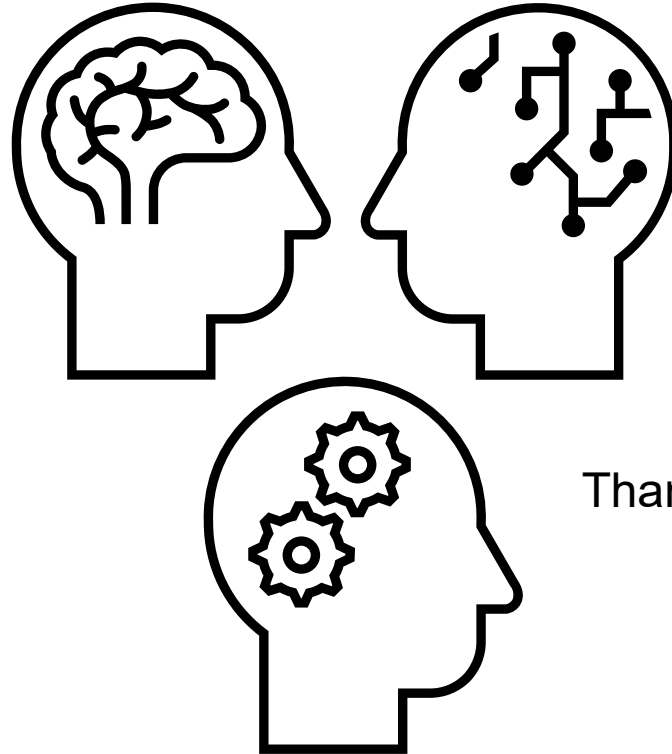


Machine Learning in NWP

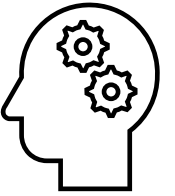
2026 edition

Linus Magnusson
ECMWF
Stockholm University



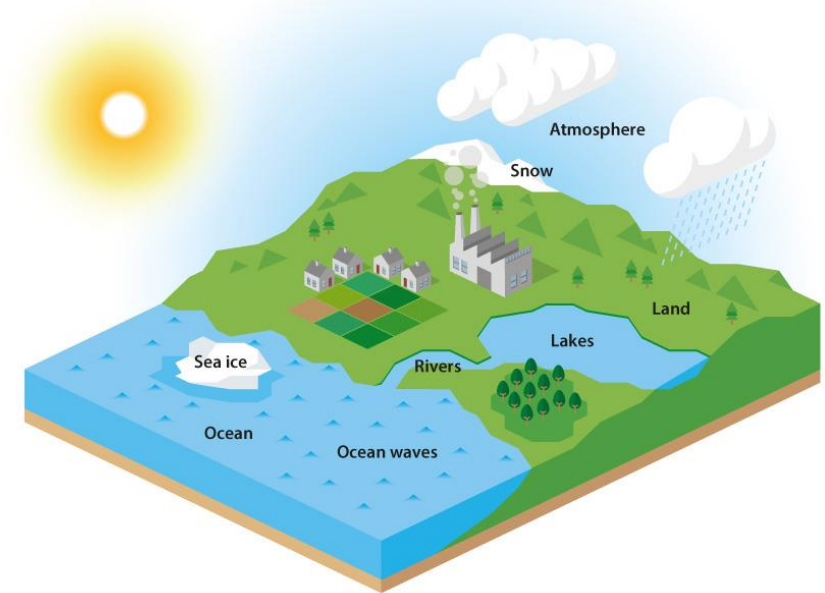
Thanks to a lot of colleagues at ECMWF!

Physics-based forecast systems



Medium-range

- Cycle 50r1
- Atmosphere: 9 km
- 15 days lead time
- 50 ensemble members + 1 unperturbed control forecasts
- Distributed around 8 hours after initialisation time



Sub-seasonal

- Cycle 50r1
- Atmosphere: 36 km
- 6 weeks lead time
- 101 ensemble members
- Initialised every day at 00UTC
- Distributed in the evening (UTC time)

Seasonal (SEAS5)

- Atmosphere: 36 km
- 7 months lead time
- 51 ensemble members
- Initialised 1st every month
- Distributed 5th every month
- SEAS6 coming soon!



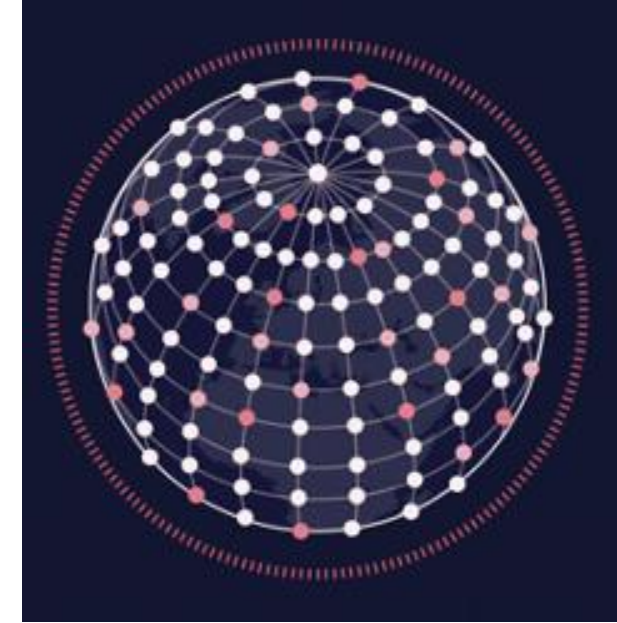
Medium-range deterministic AIFS-single v2.0

- Atmosphere: AIFS Model v2.0 using **MAE loss**, N320 (0.25°, ~25 km)
- 15 days lead time
- Initialised from physics-based analysis
- Trained on ERA5 + Operational analyses

Medium-range ensemble AIFS-ENS v2.0

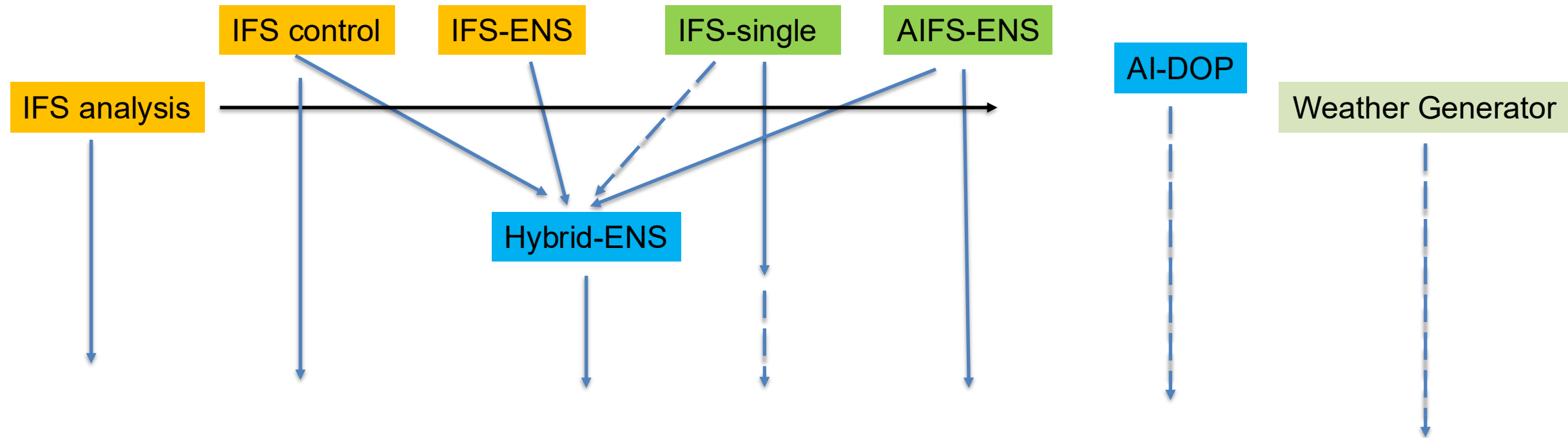
- Atmosphere: AIFS Model v2.0 using **scale-aware CRPS loss**, N320 (0.25°, ~25 km), 13 vertical levels
- 15 days lead time
- 50 ensemble members + 1 unperturbed* control forecasts
- Same initial conditions including perturbations as physics-based ensemble
- Trained on ERA5 + Operational analyses

* With the use of CRPS-loss, also the control forecast include the “noise” to simulate model uncertainty



Experimental sub-seasonal system on the way!

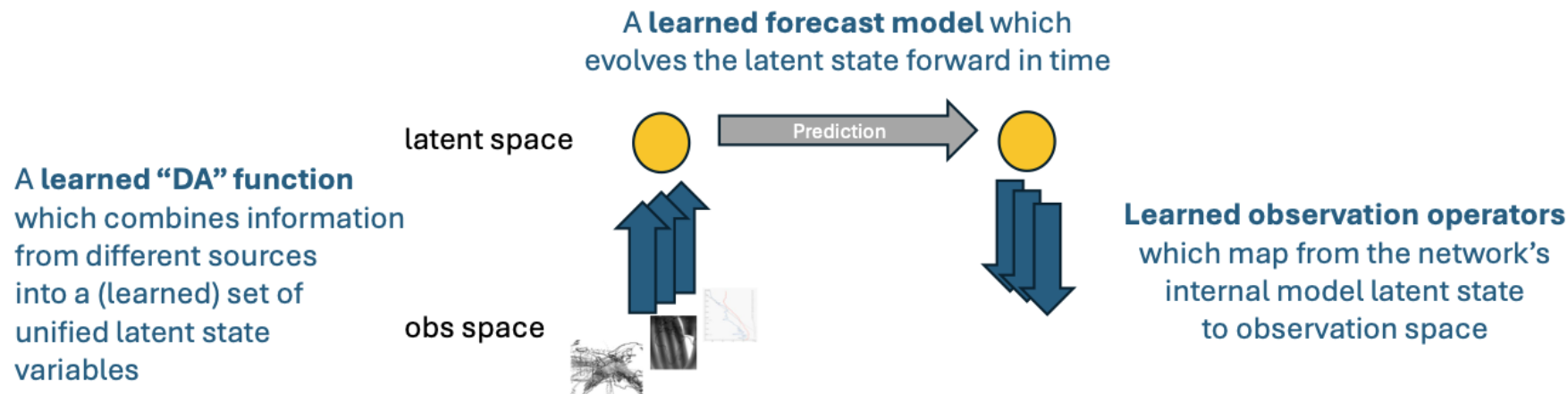
Ongoing medium-range forecast system developments and future pathways...



What is AI-DOP? (Direct Observation Predictions)

ECMWF is exploring observation-driven ML forecasting models: AI-DOP

End-to-end observations → forecast



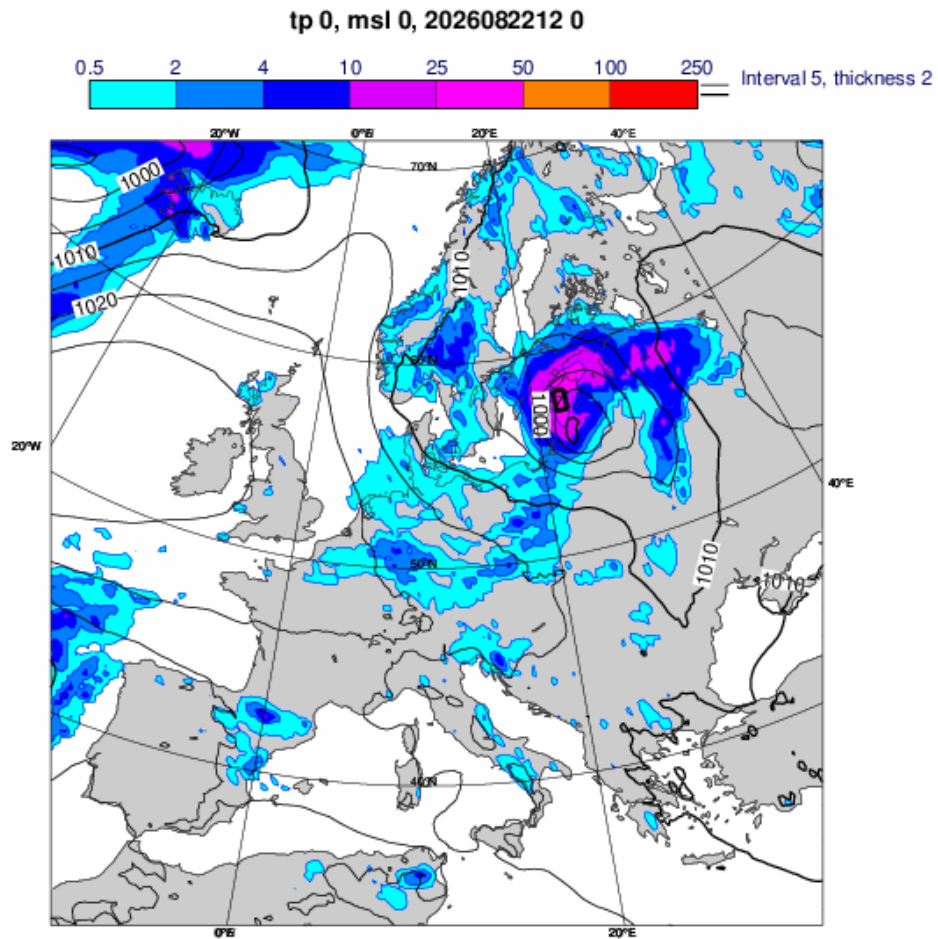
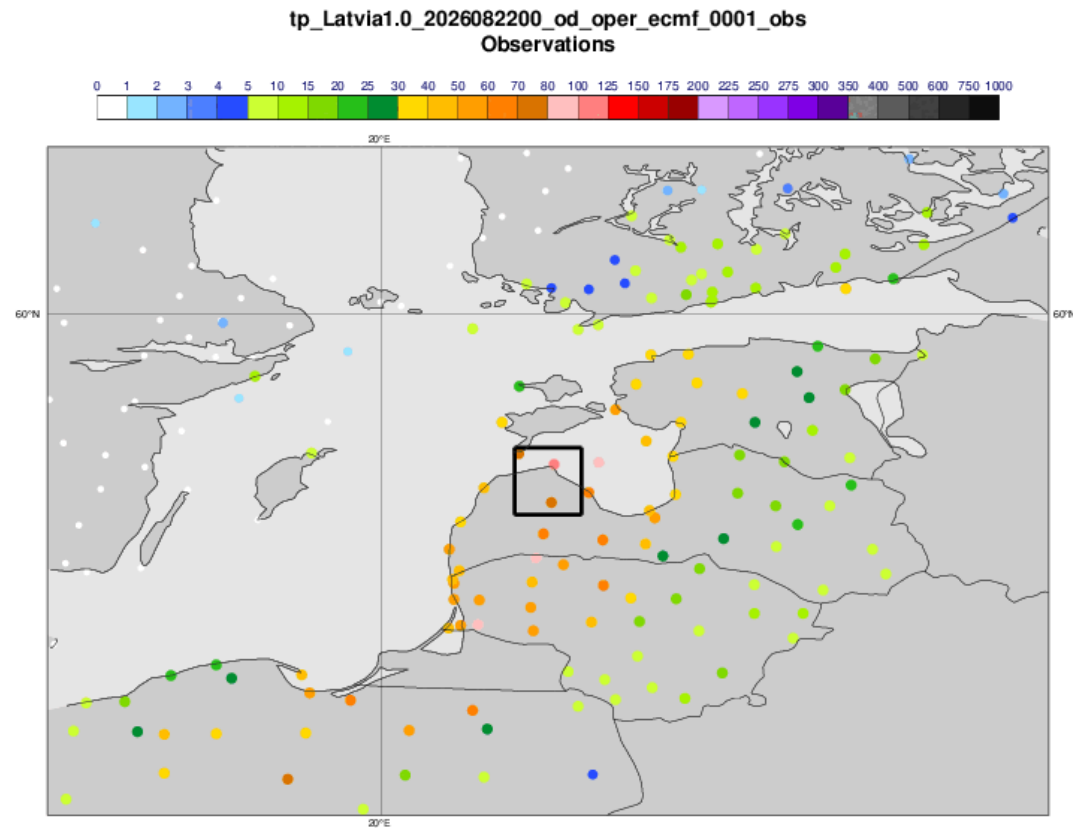
Given observations in the past, predict observations in the future



Learning a model from observational data

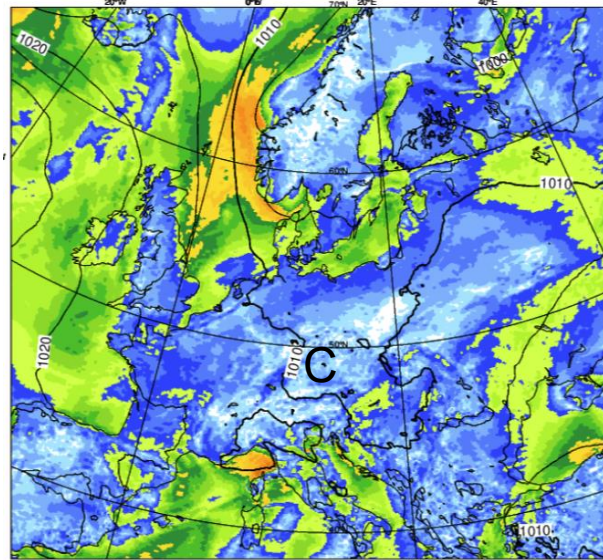
From Peter Lean (ECMWF)

Recent 5b cyclone hitting the Baltic States (August 2026)

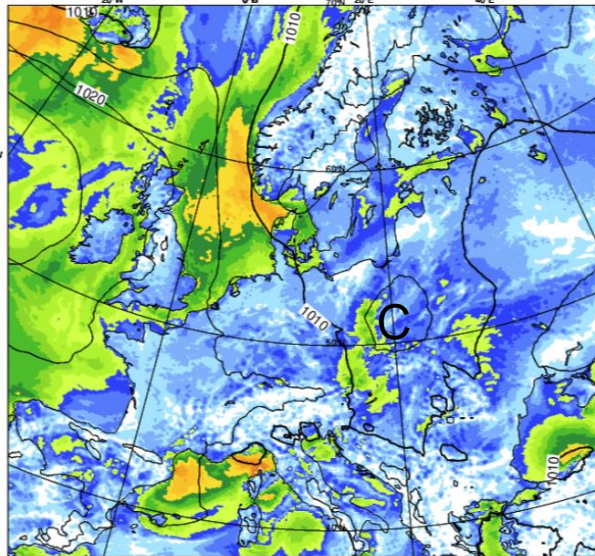


IFS-control forecast from 20 August 00UTC

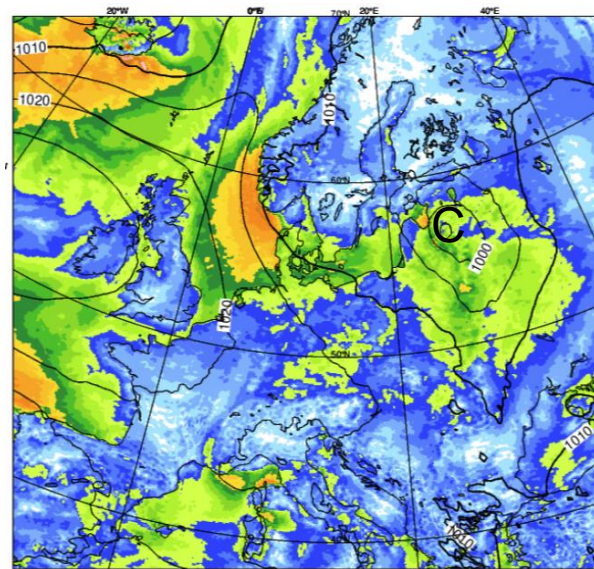
+36h



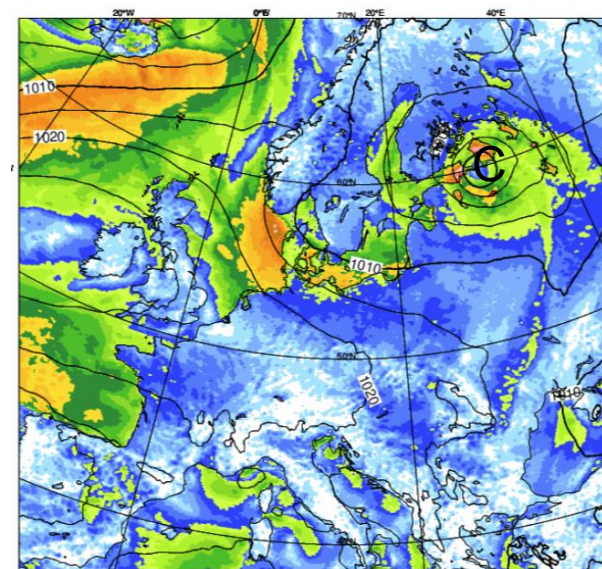
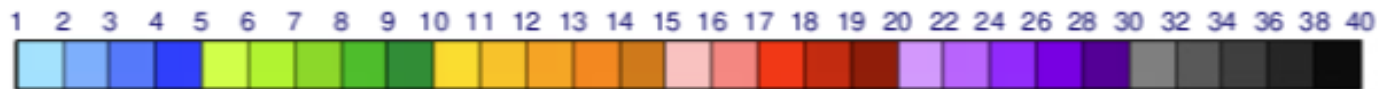
+48h



+60h

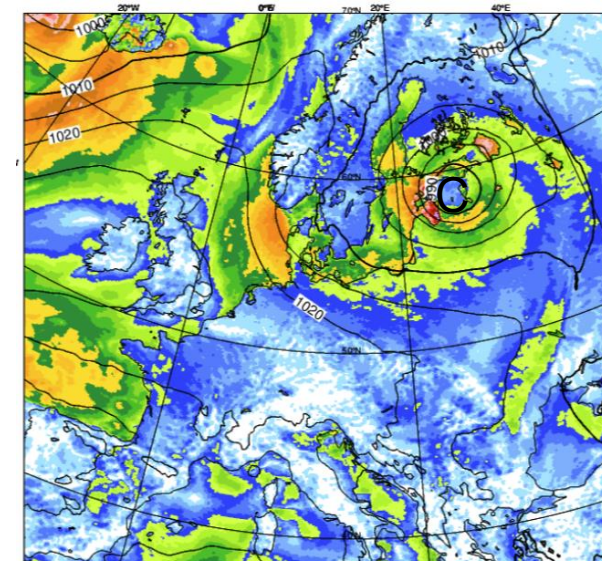


+72h



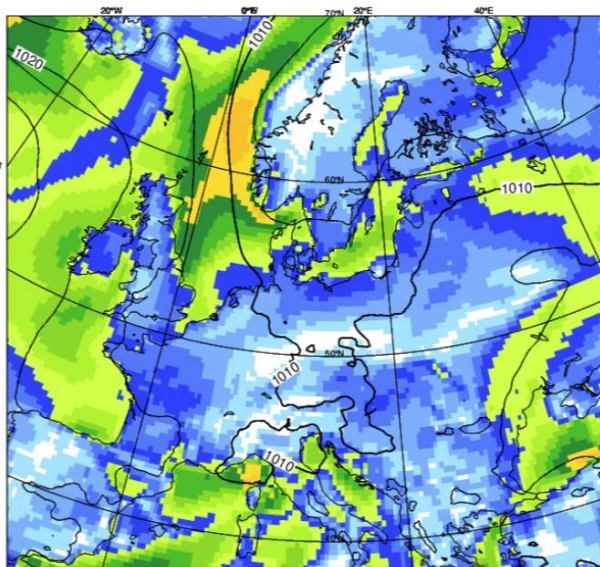
MSLP – contour
10-m mean wind - shade

IFS Analysis

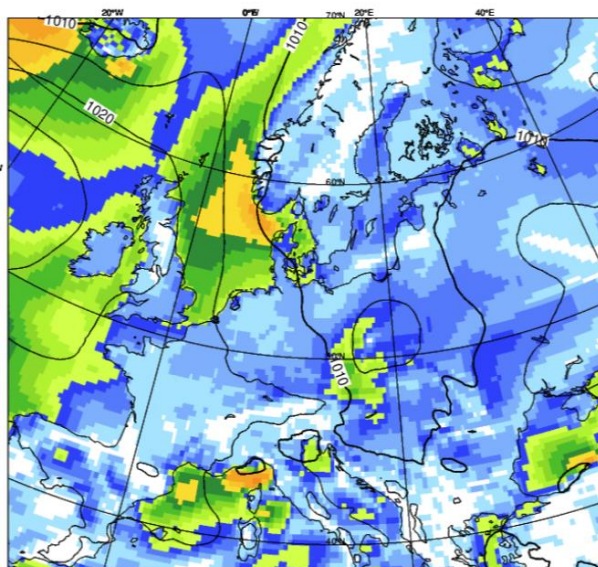


AIFS-single 20 August 00UTC

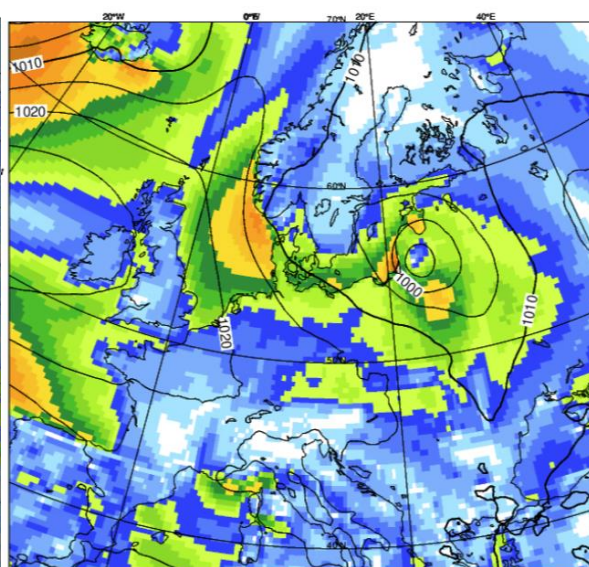
+36h



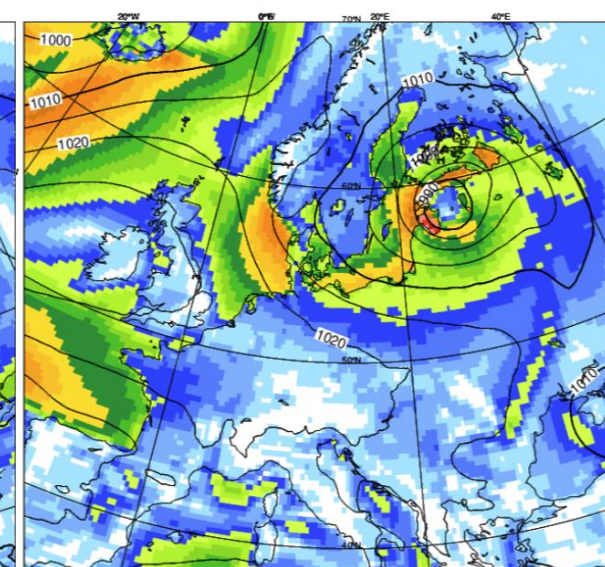
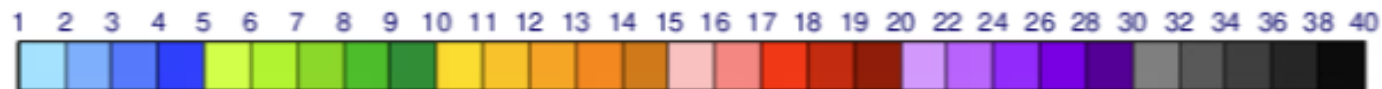
+48h



+60h

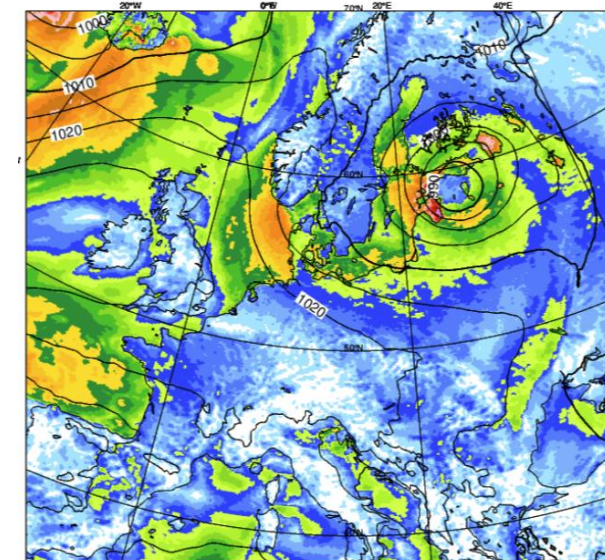


+72h



MSLP – contour
10-m mean wind - shade

IFS Analysis



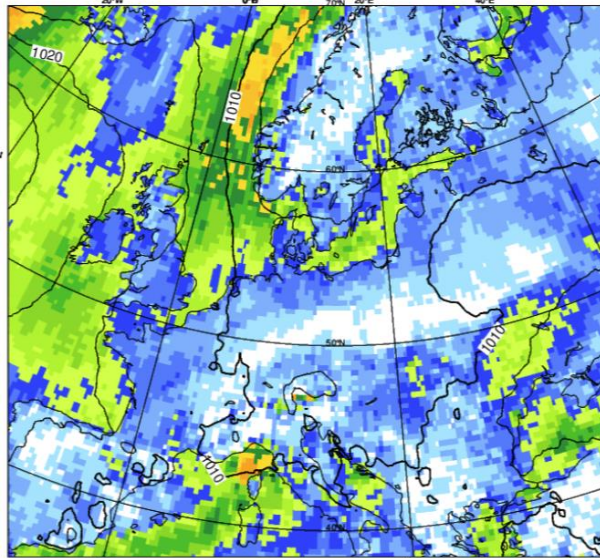
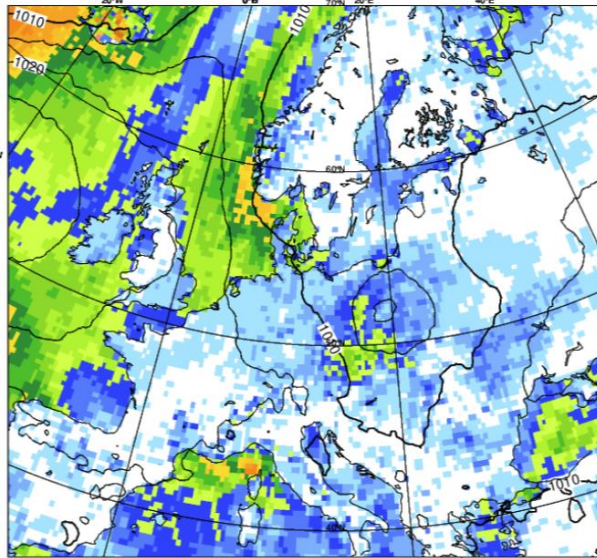
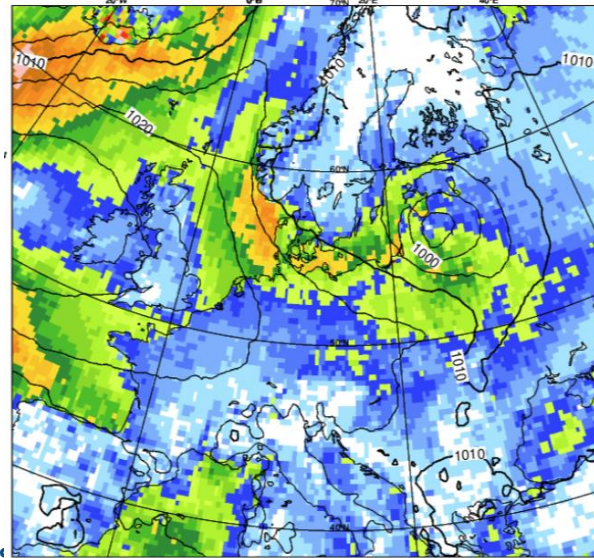
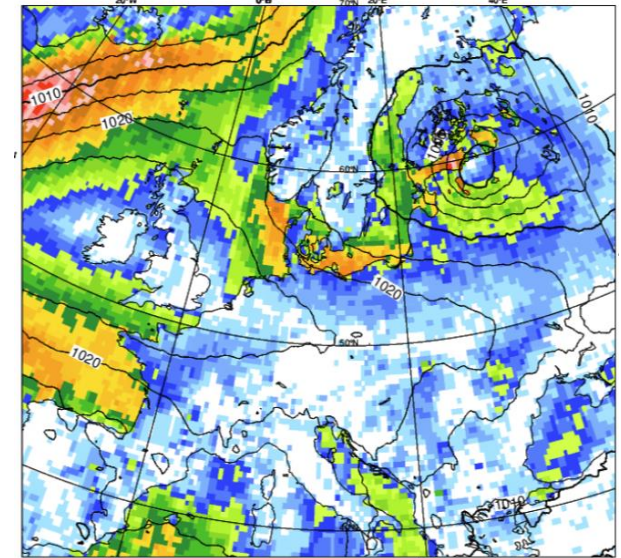
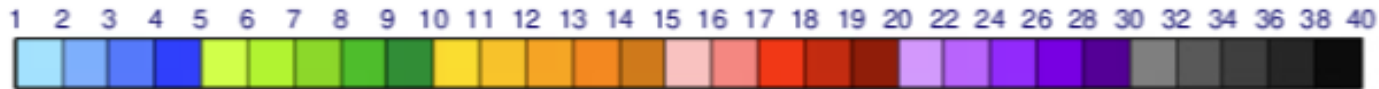
DOP 20 August 00UTC

+36h

+48h

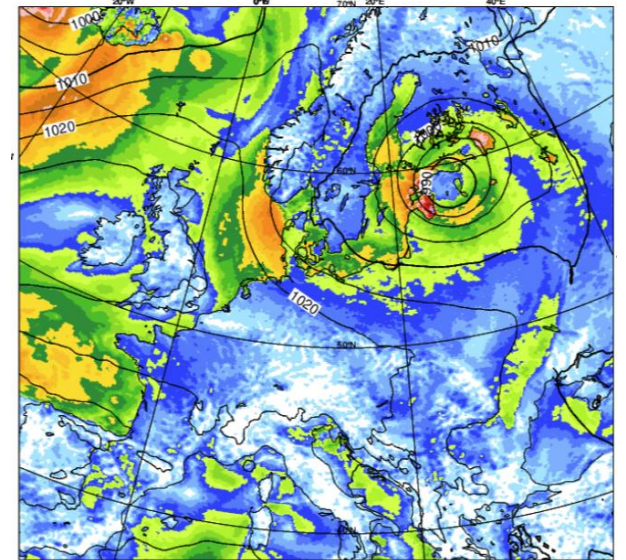
+60h

+72h



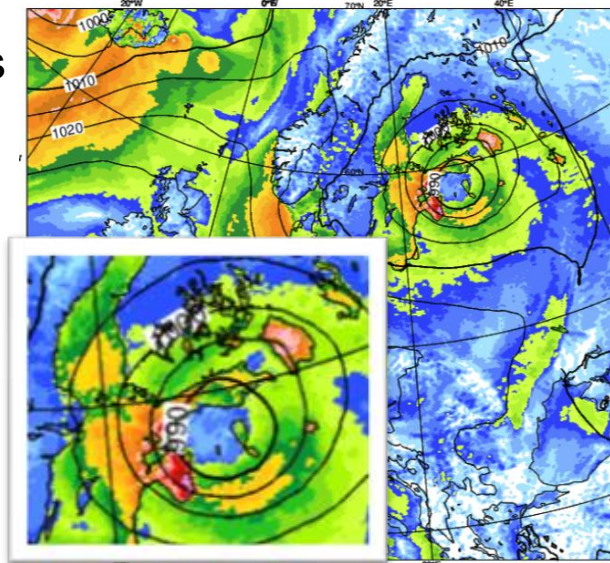
MSLP – contour
10-m mean wind - shade

IFS Analysis

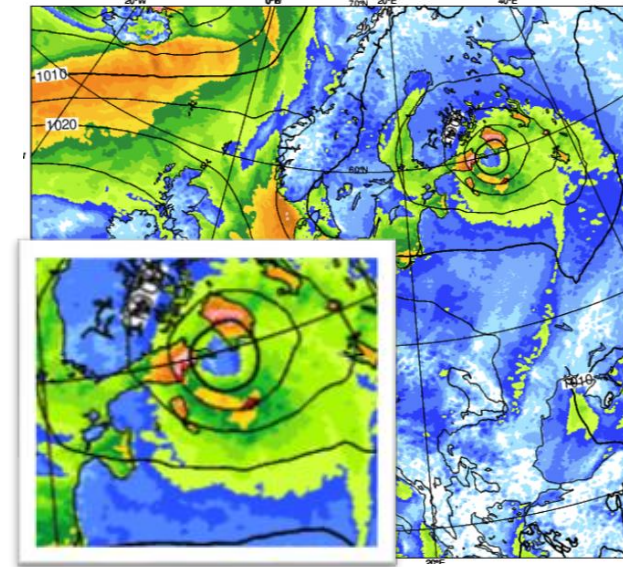


20 August 00UTC +48h

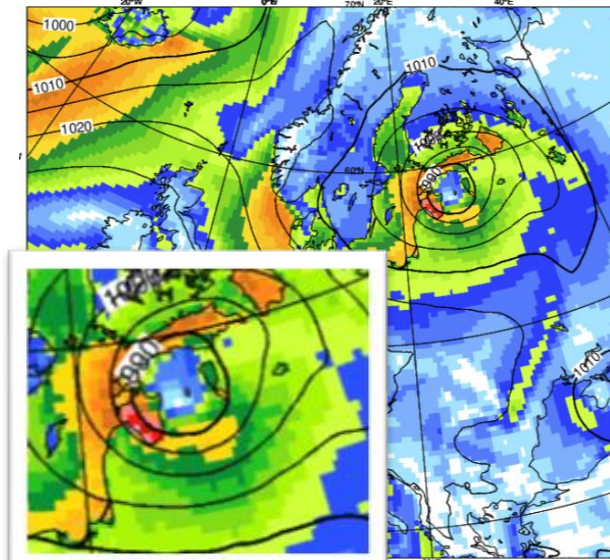
IFS analysis



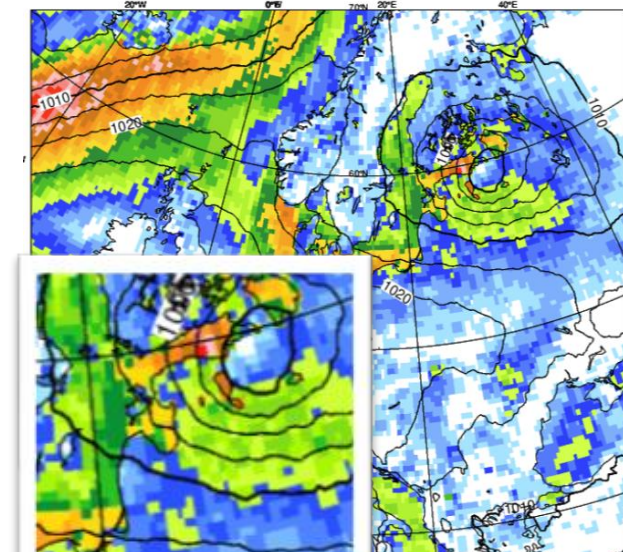
IFS-control



AIFS-single



AI-DOP



AIFS - Single v2 new variables



Forcings:

- Land-sea mask
- Orography
- Deviation of sub-grid orography
- Slope of sub-scale orography
- Insolation
- Latitude/longitude
- Time of day + day of year
- **Model bathymetry**

Prognostics:

Pressure levels at **10**, 50, 100, 150, 200, 250, 300, 400, 500, 600, 700, 850, 925, 1000 hPa:

- Geopotential
- Horizontal ~~and vertical~~ wind components
- Temperature

Pressure levels at 50, 100, 150, 200, 250, 300, 400, 500, 600, 700, 850, 925, 1000 hPa:

- Specific humidity

Surface fields:

- Surface pressure
- Mean sea-level pressure
- Sea-surface temperature
- Skin temperature
- 2m temperature
- 2m dewpoint temperature
- 10 m horizontal wind components (u,v)
- Total column water
- **Snow cover**
- **Significant wave height**
- **Significant wave height of all waves with periods**
- **Coefficient of drag with waves**
- **Mean wave period and direction**

Diagnostics:

Pressure levels at 10, 50, 100, 150, 200, 250, 300, 400, 500, 600, 700, 850, 925, 1000 hPa:

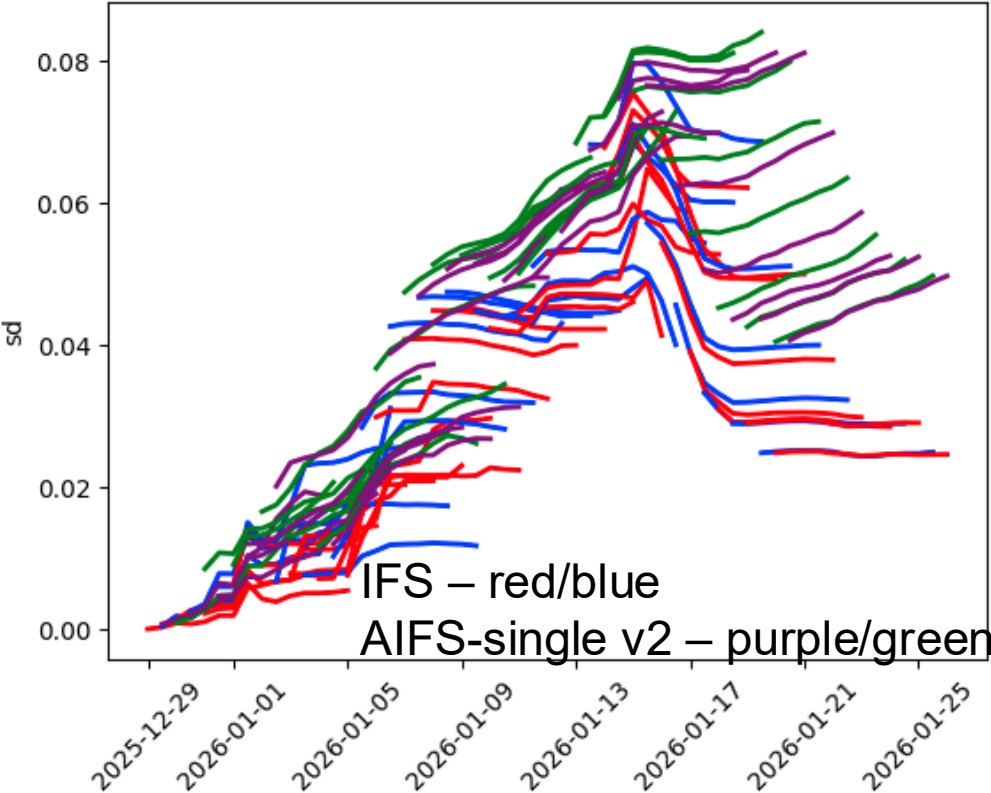
- **vertical wind components**

Surface fields:

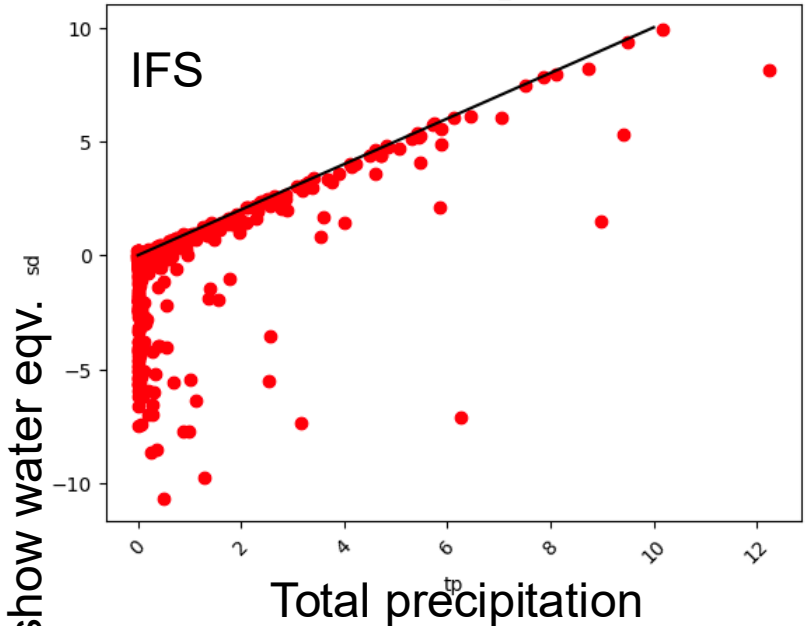
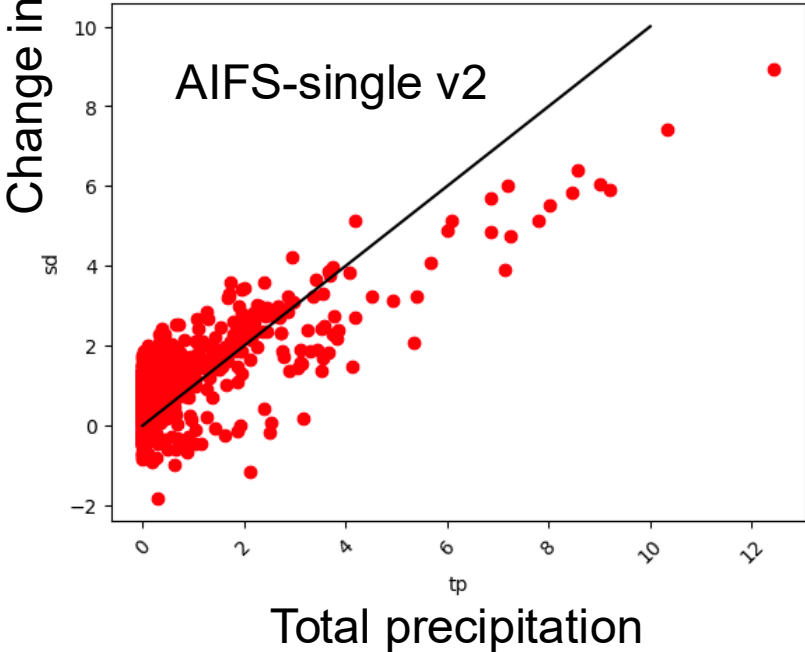
- Total and convective precipitation
- **Snow depth**

Example of physical inconsistency: Snow depth evolution

10-day forecasts for snow water equivalent

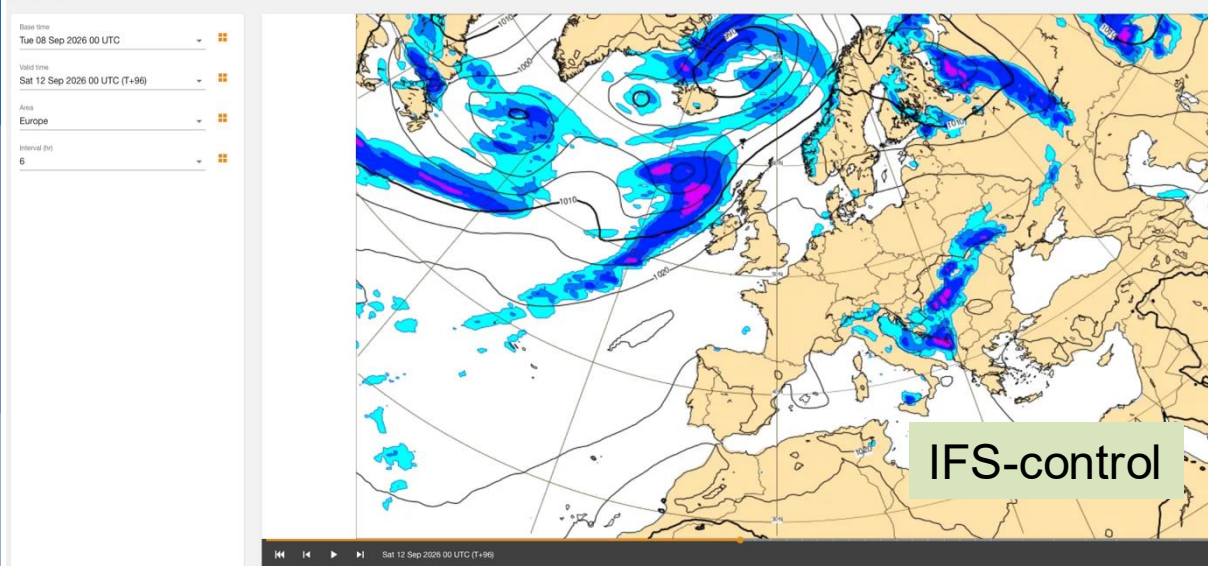


Accumulation of snow without precipitation!
Probably due to large snow increments in ERA5

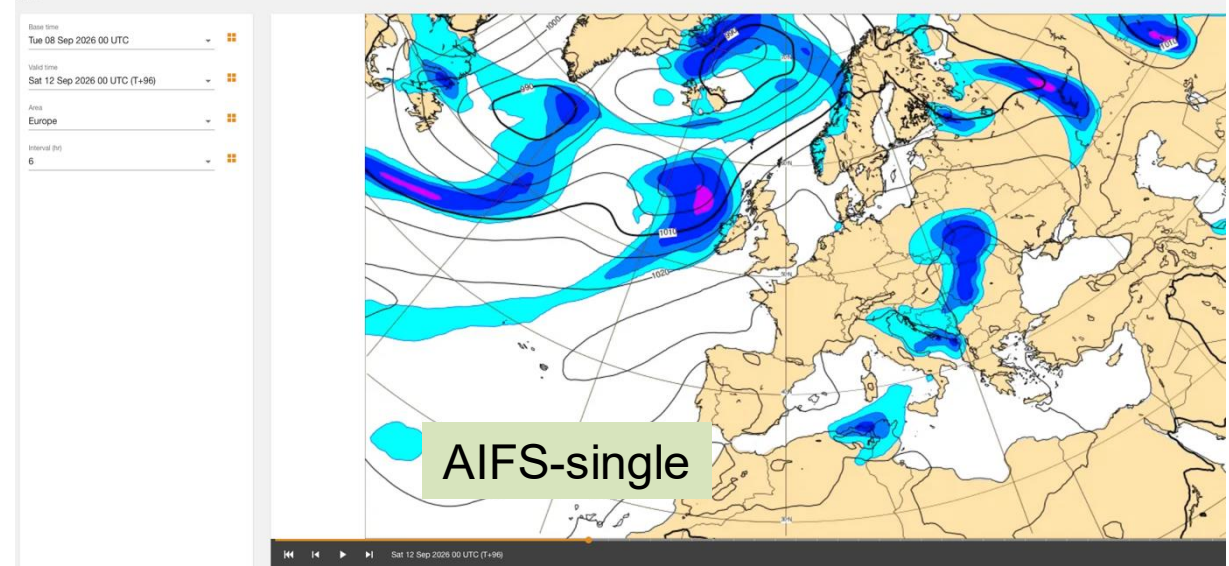


Why a hybrid model? - Pos and cons of machine-learning models

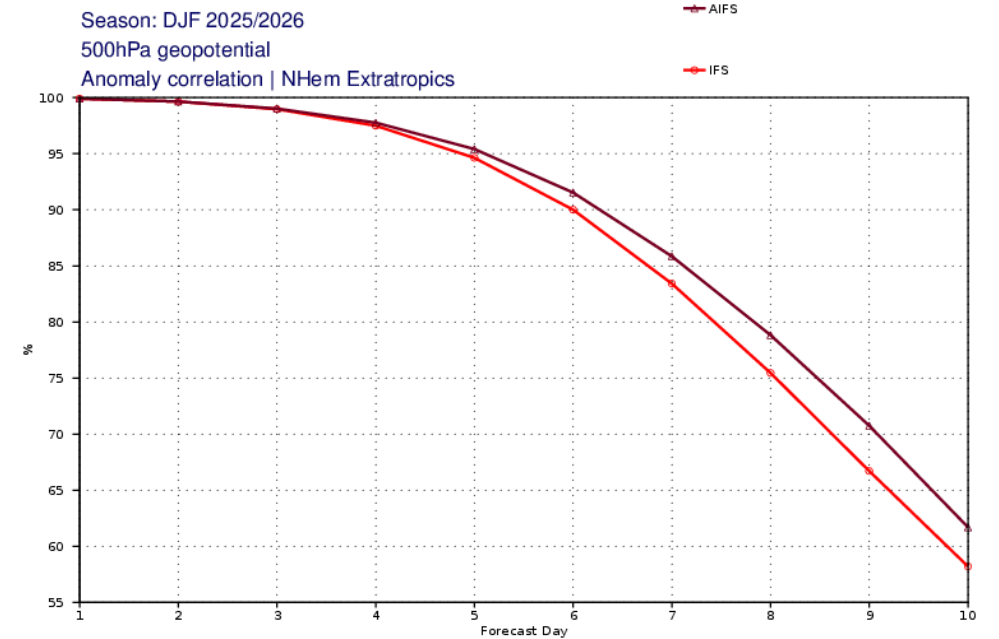
Rain and mean sea level pressure
Control forecast



AIFS Single: Rain and mean sea level pressure
AIFS



AIFS ENS Control: Rain and mean sea level pressure
AIFS ENS



What is Hybrid?

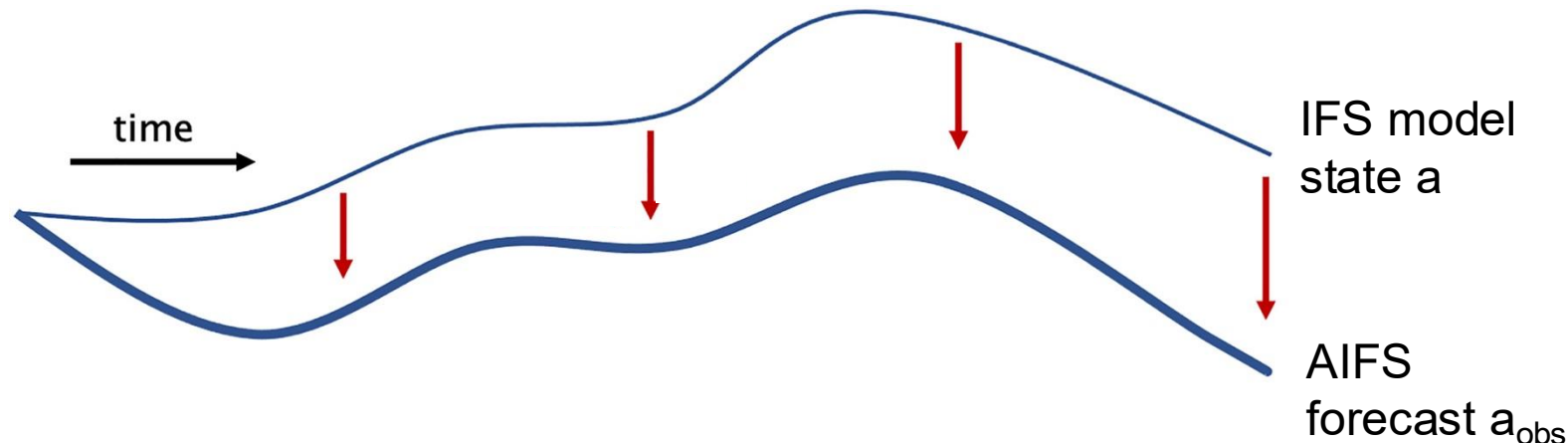
- Use spectral nudging to a bespoke AIFS models that predict fields on IFS model levels.
- Take AIFS-Single and AIFS-ENS forecasts and spectrally nudge IFS-CF and IFS-ENS to them. Tv and VOR nudged up to total wavenumber T21 in the **troposphere**.
Like Hussain et al. (2024) @ECCC
- **Pros:**
 - Benefit from the higher synoptic skill from AIFS
 - Small-scales not damped
 - Retains better physical consistency of a physics-based model
 - Produce the same output as IFS
- **Cons:**
 - Complex operational system

$$\frac{\partial a}{\partial t} = -\mathbf{v} \cdot \nabla a + Q_a^p - \frac{a - a_{obs}}{\tau},$$

$a = \{T_v, VOR\}$

$\tau = 12 \text{ hours}$

$a_{obs} = \text{AIFS (wavenumbers 0-21)}$

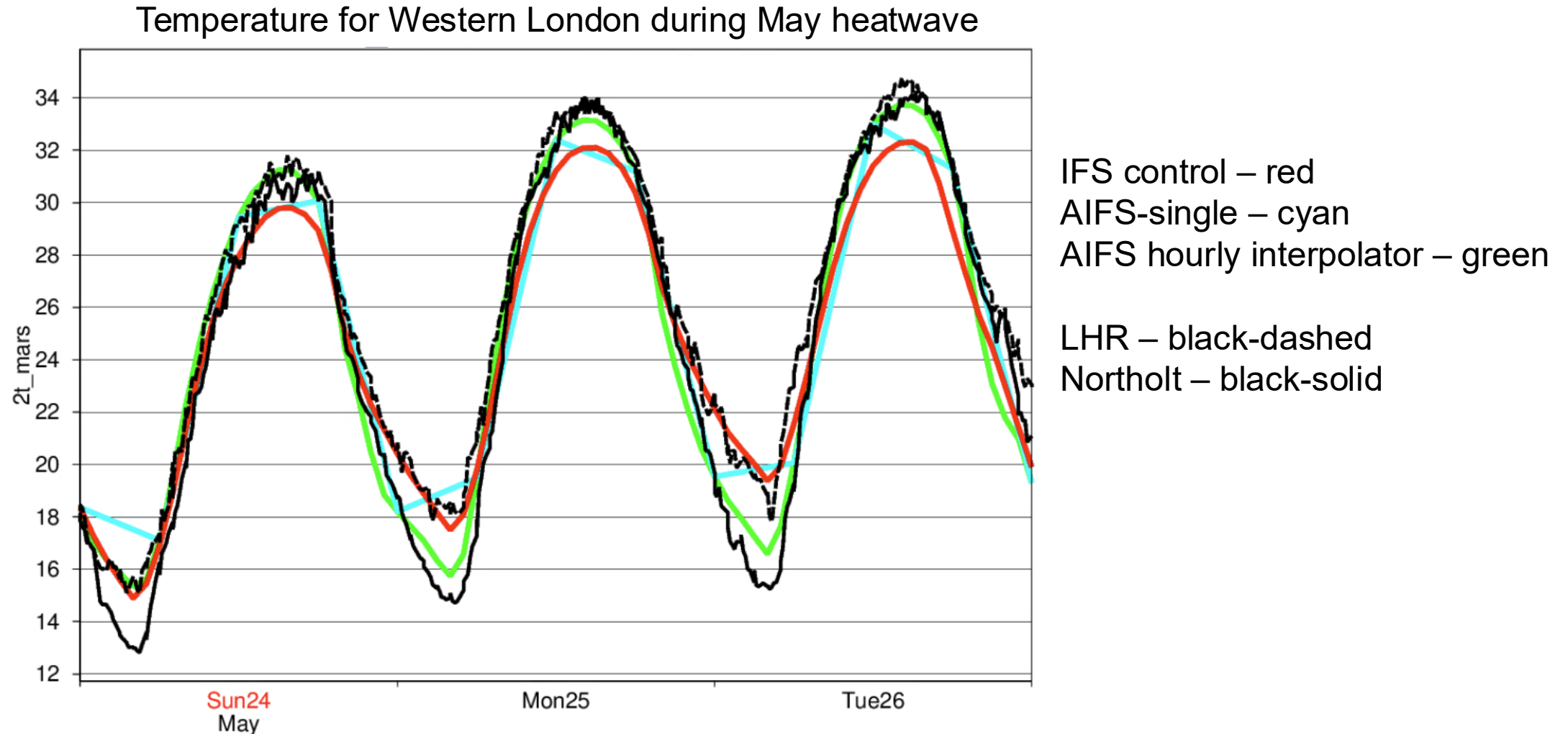


From Watt-Meyer et al. (2021)

What are the next steps?

- Reforecasts in order to calculation Extreme Forecast Index
 - Sub-seasonal forecast system
 - Hourly "interpolator"
 - Train AIFS on combination of reanalysis and observations?
-
- How far into an end-to-end forecasting system should we go?
 - How much of the traditional post-processing should be handle by the prediction model?
 - Think outside the traditional way we see model variables
 - Near-road temperatures trained on road-side observations?

Diurnal cycle of 2-metre temperature – experimentation with hourly interpolator



Forecast Comparison Dashboards

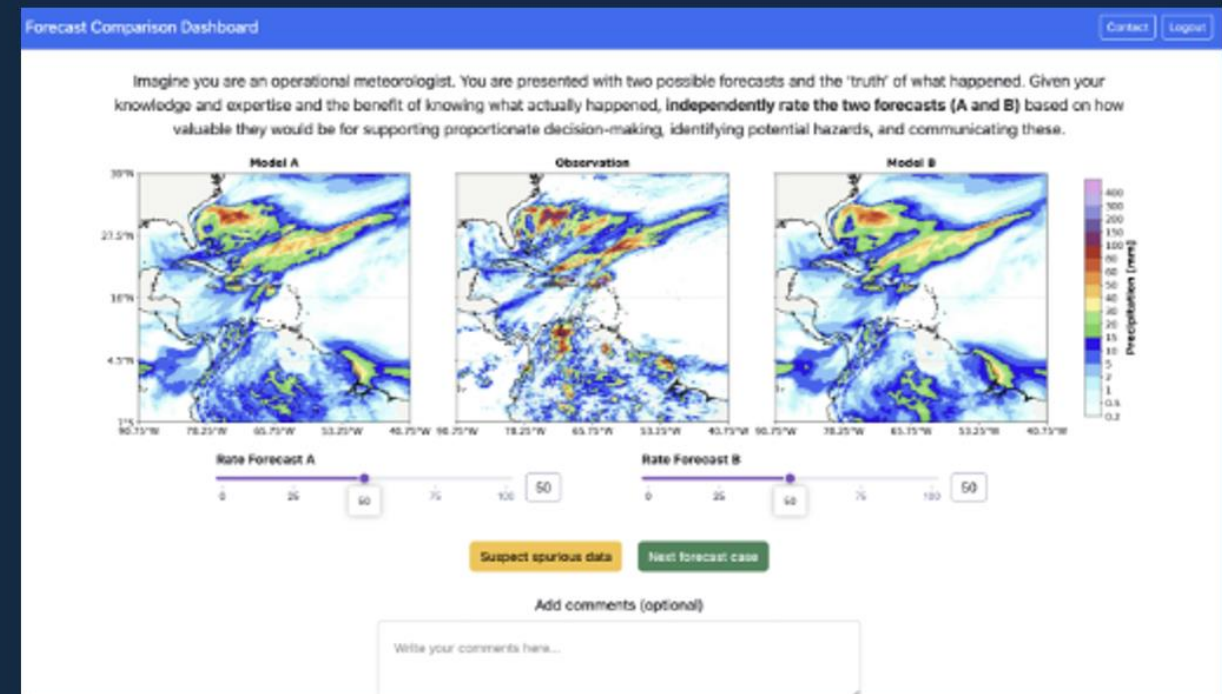
An international research activity answering how useful AI, NWP and hybrid forecasts really are to the meteorologists who use them

THE SCIENCE QUESTIONS WE WILL ANSWER

- 01 How do meteorologists rate the usefulness of AIWP, NWP and hybrid forecasts?
- 02 Which spatial verification methods best reproduce expert forecaster rankings?
- 03 What forecast characteristics do meteorologists actually prefer — intensity, displacement, sharpness?
- 04 Can we build better verification methods that capture desirable forecast behaviour?

Take part at www.forecastcomparison.org

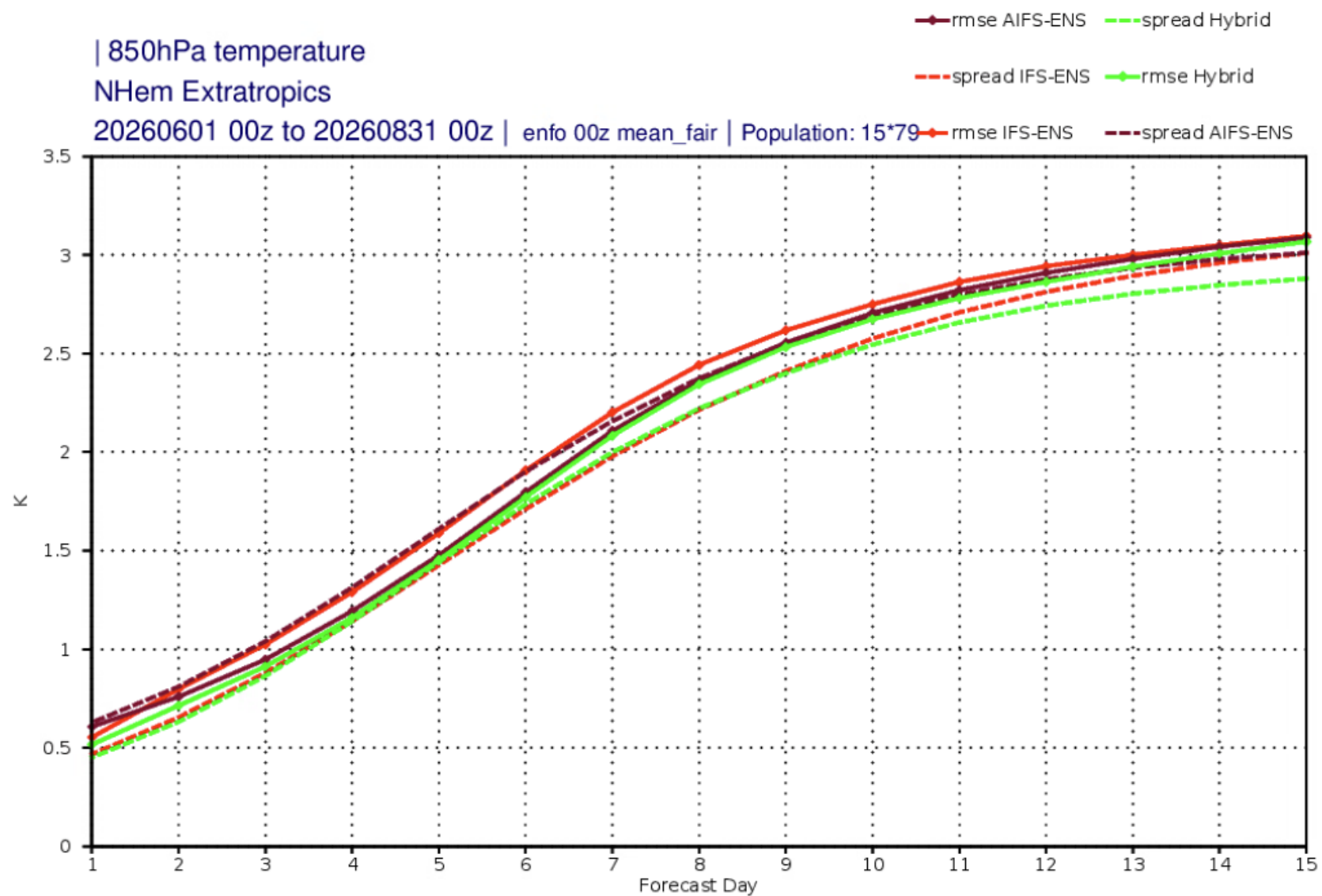
Questions? Email nicholas.loveday@bom.gov.au



Two unlabelled forecasts appear beside the observed truth for a random model pair, region, lead time and date. Experts rate usefulness blind; each rating feeds the dataset behind these questions.

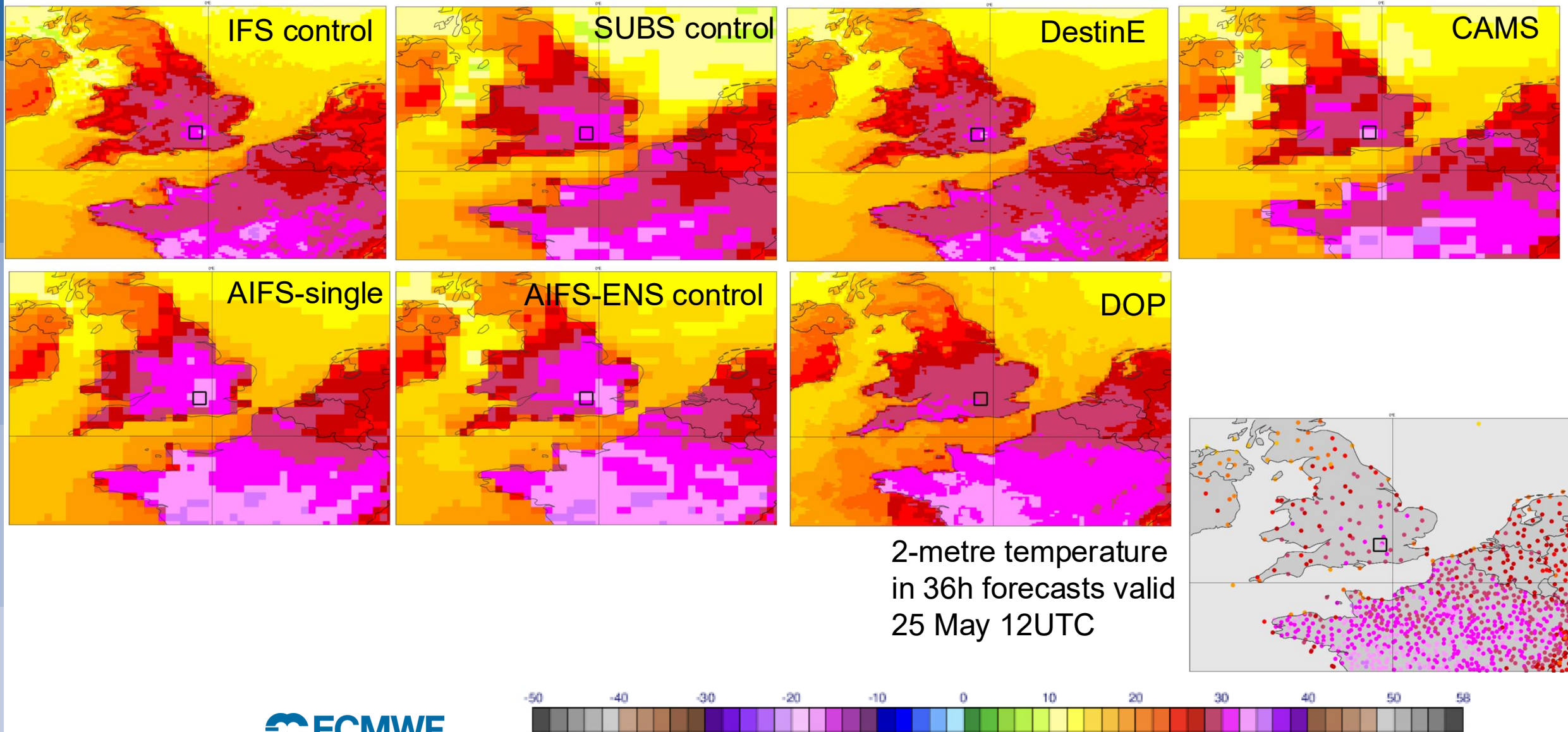


Verification of IFS-ENS, AIFS-ENS and Hybrid for JJA 2026

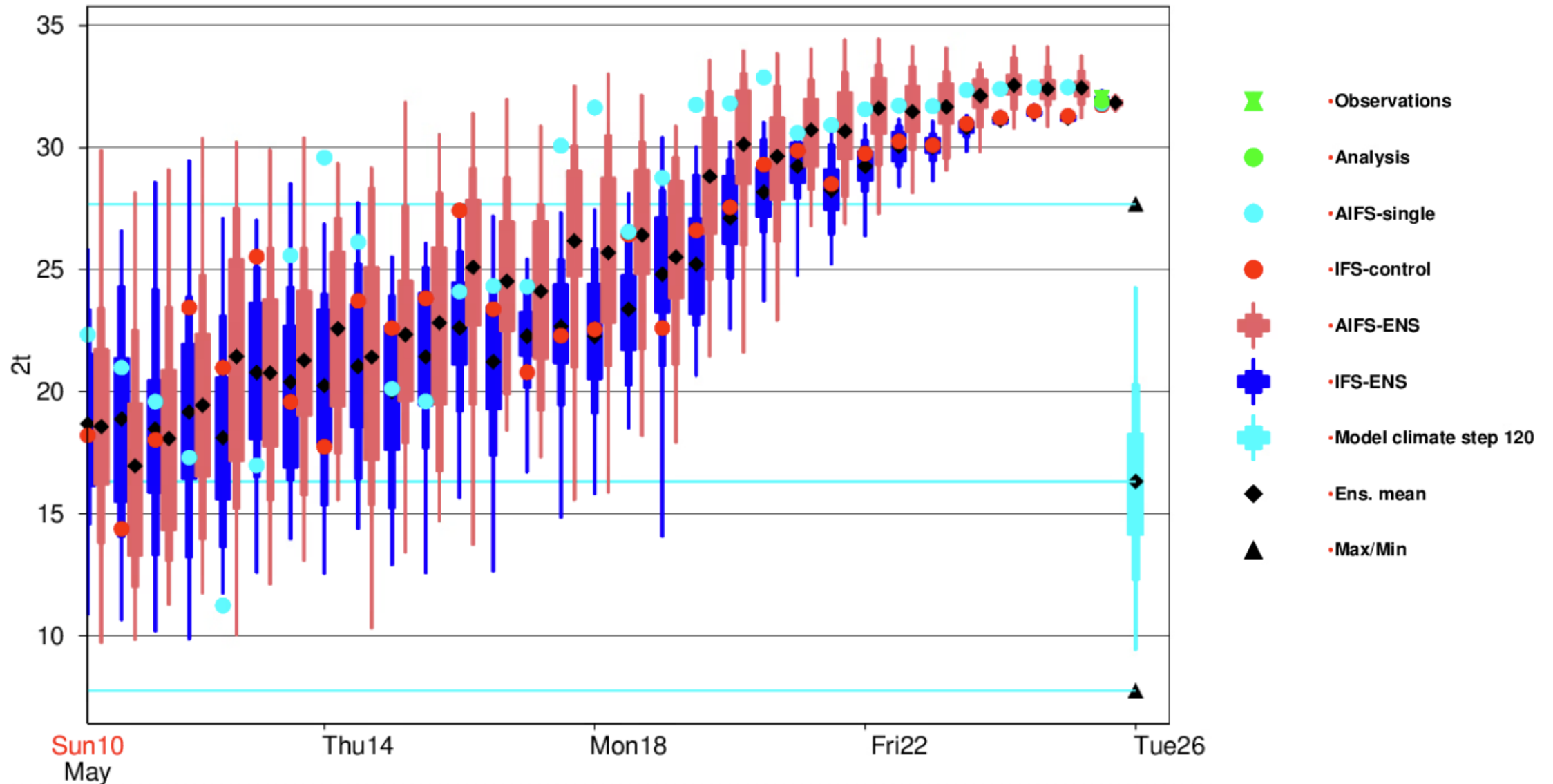


Short-range prediction of 2-metre temperature at 12 UTC on 25 May 2026

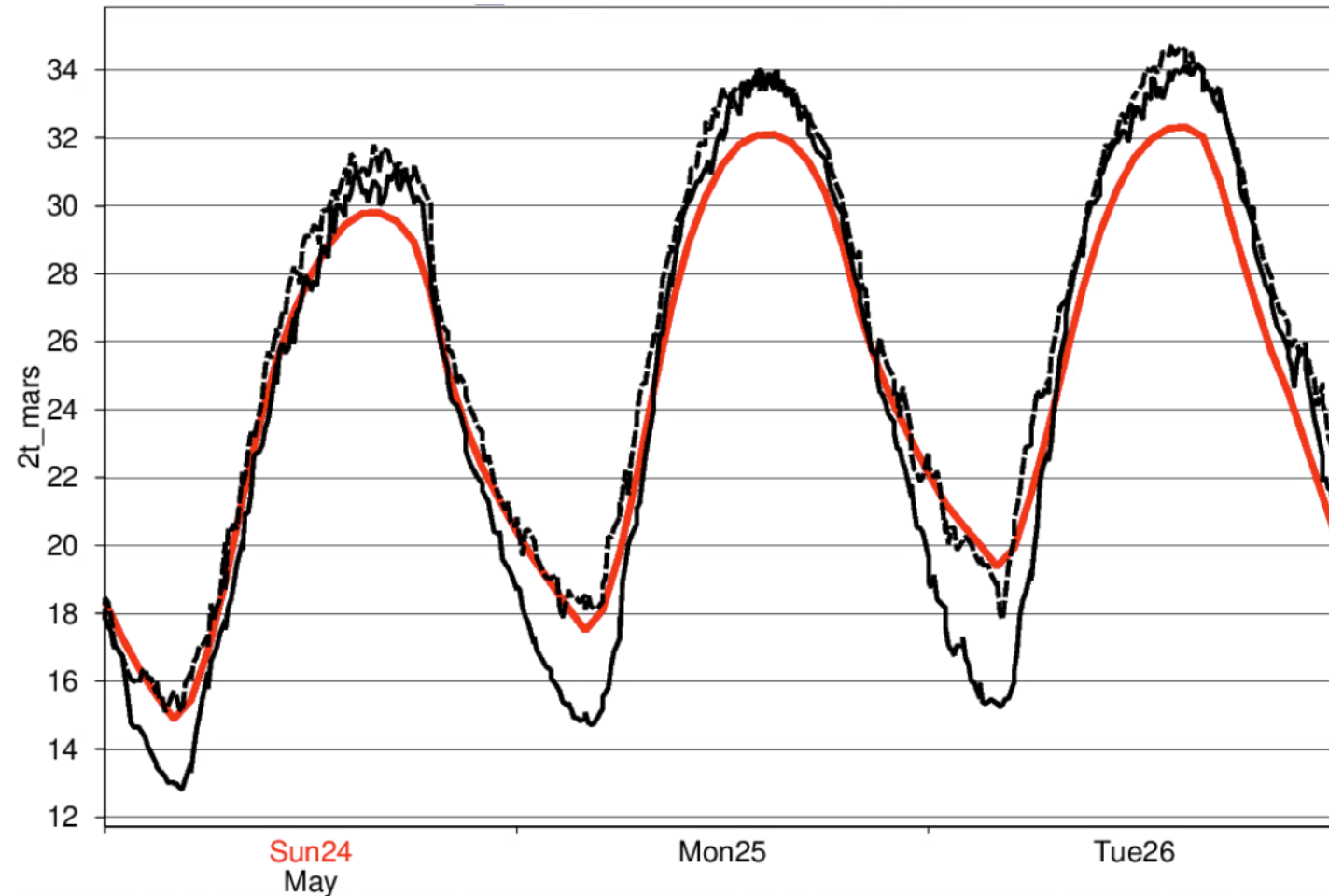
- The ECMWF Collection



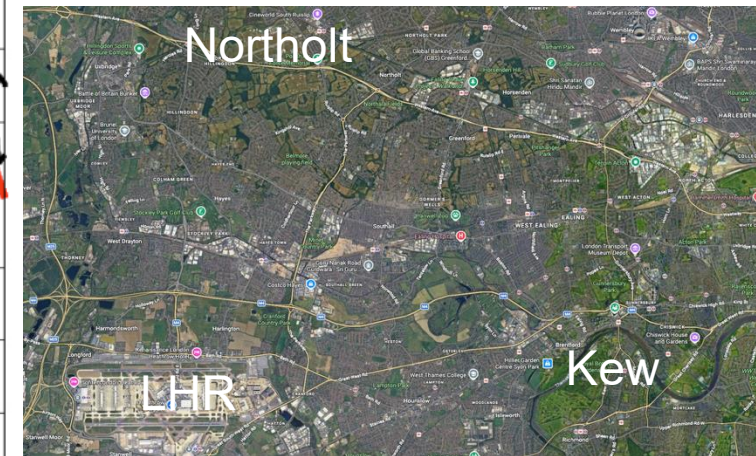
Evolution of 2-metre temperature at 12UTC on 25 May around London-Heathrow



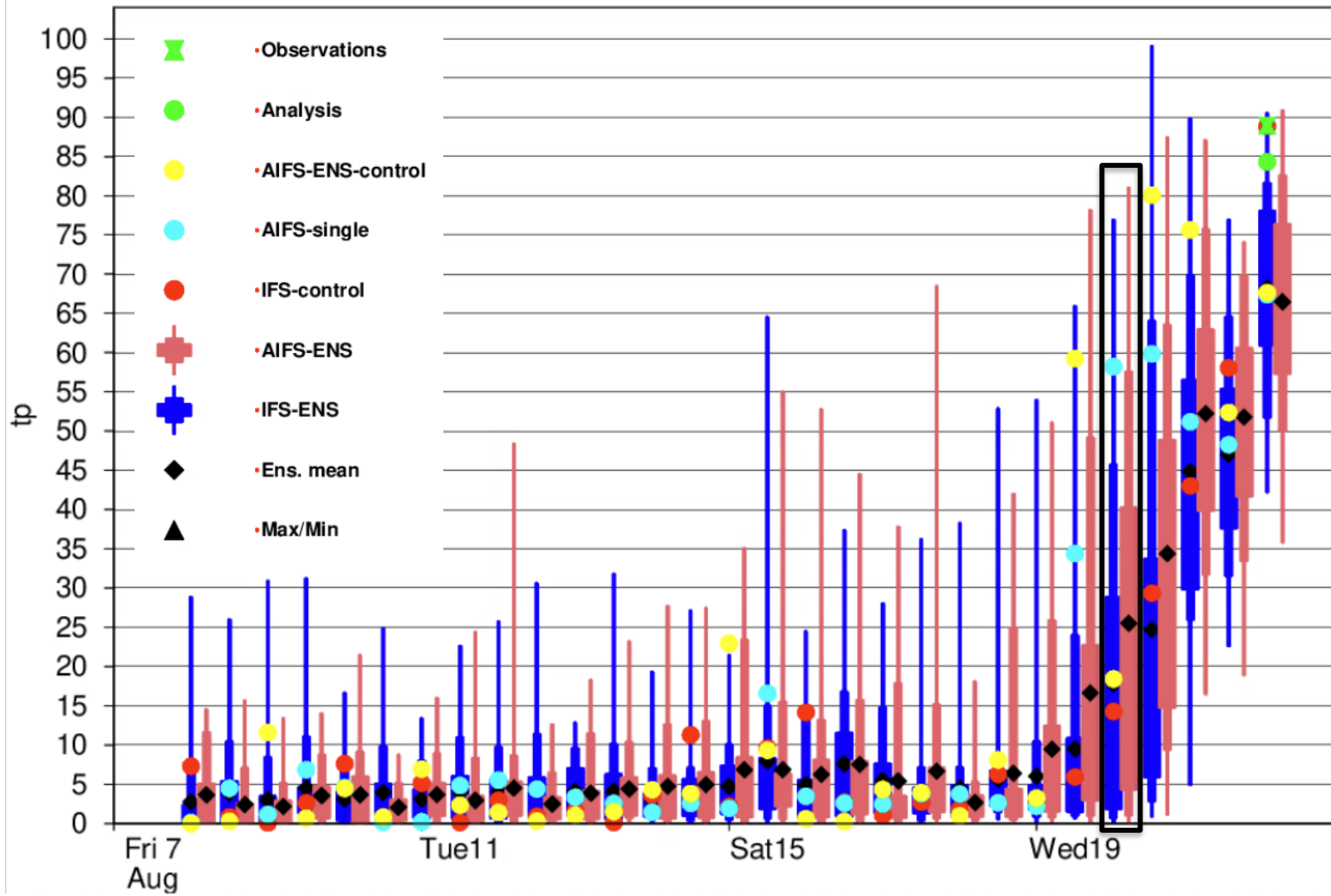
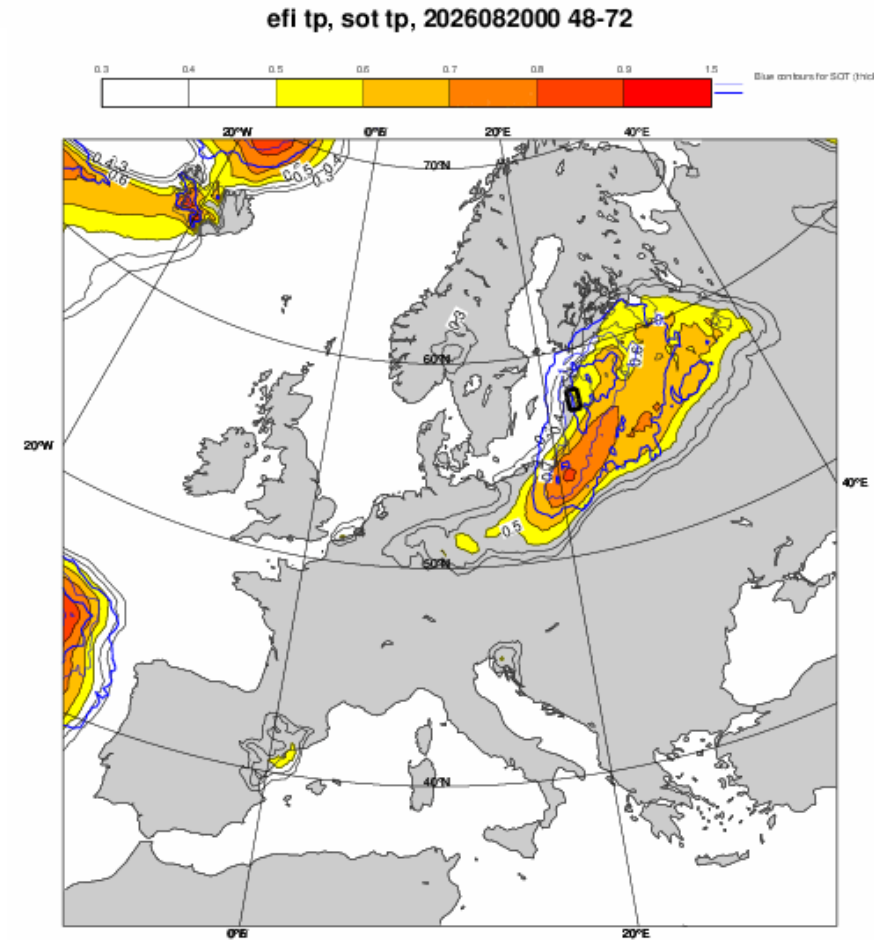
Diurnal cycle of 2-metre temperature



IFS control – red
LHR – black-dashed
Northolt – black-solid



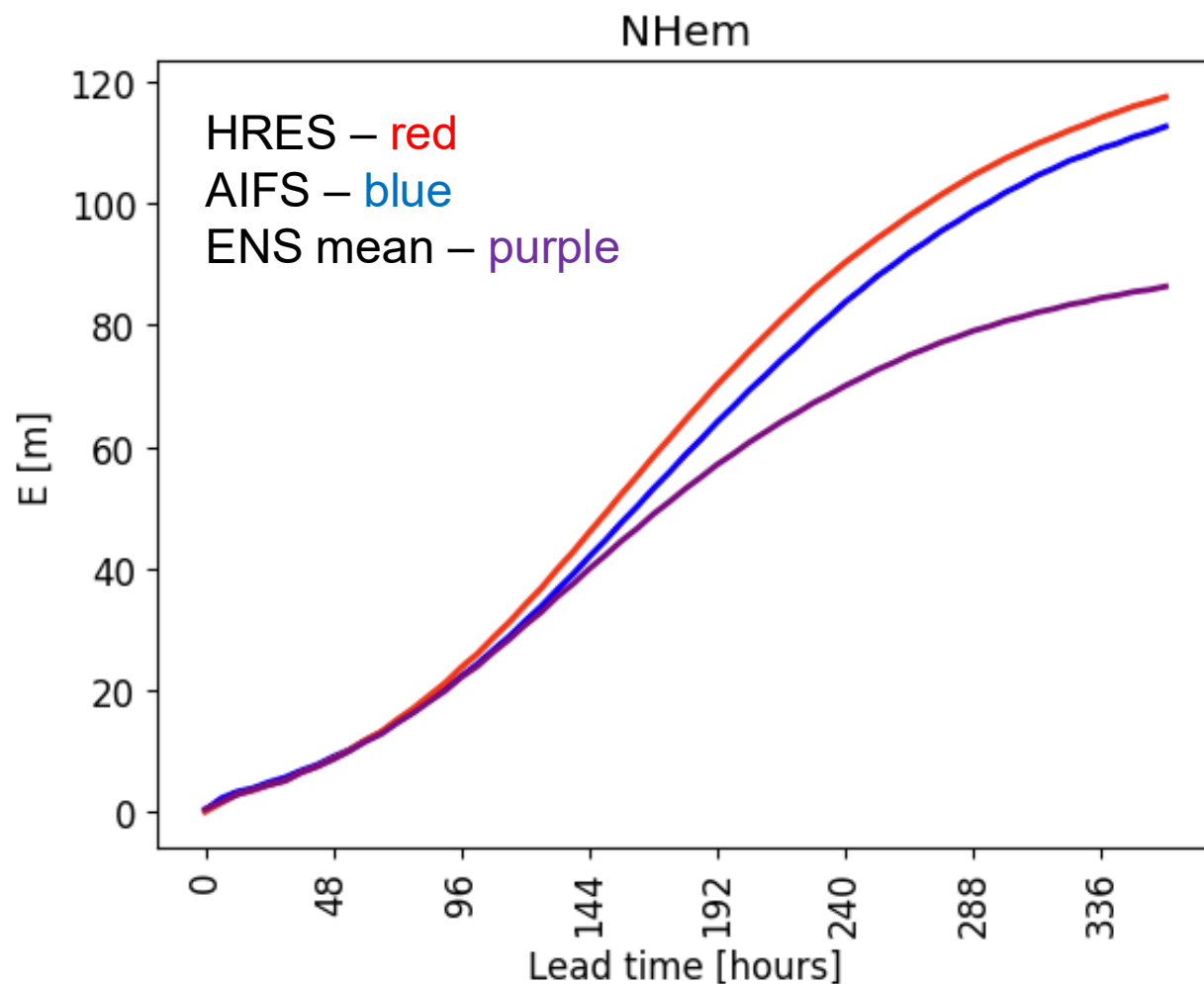
24-hour precipitation on 20 August inside the box



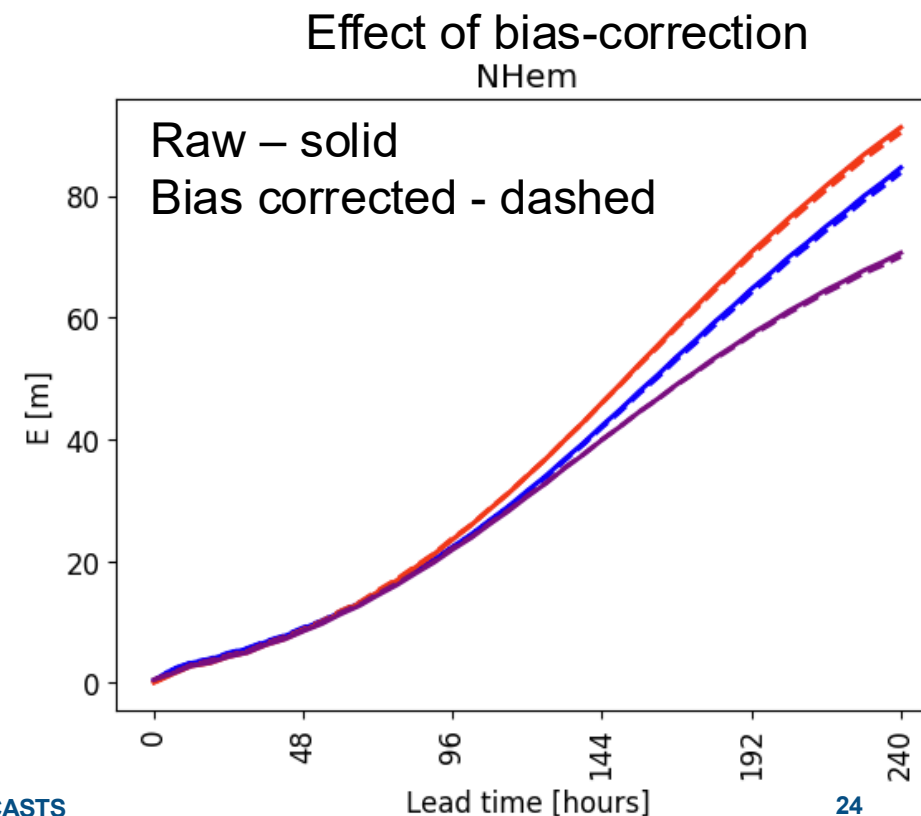
Understanding large-scale score differences between IFS and ML models

RMSE for z500, N.Hem March 2024 – mid-Feb 2025

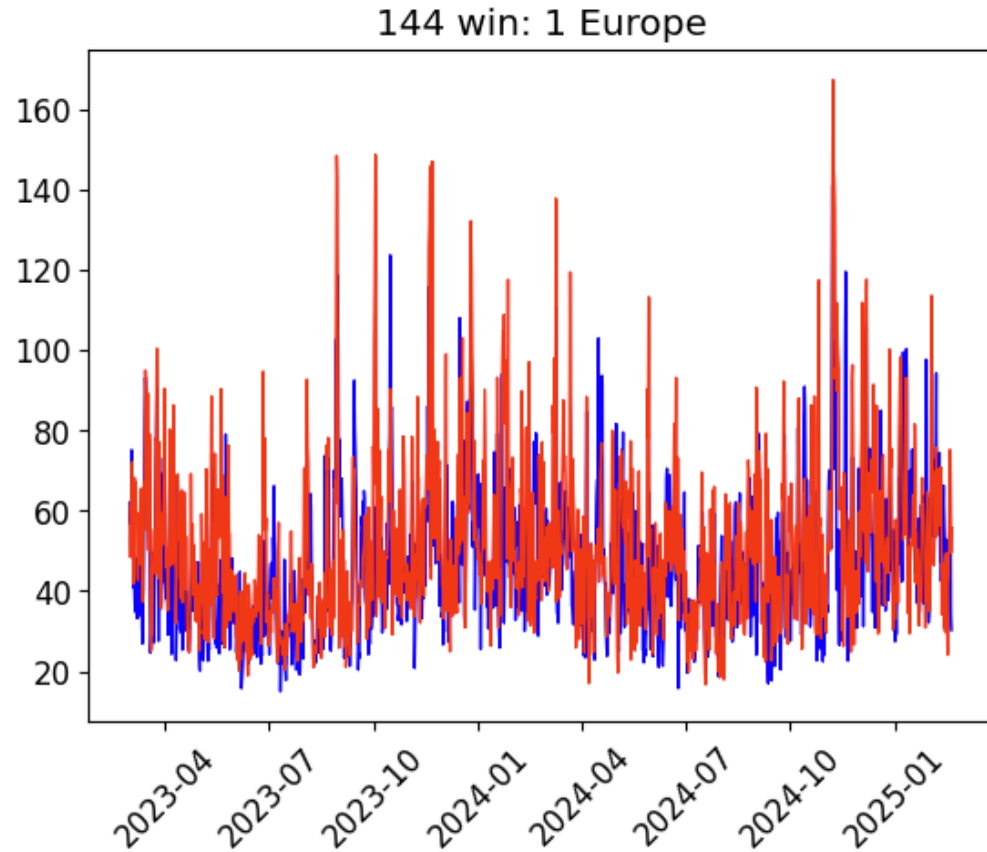
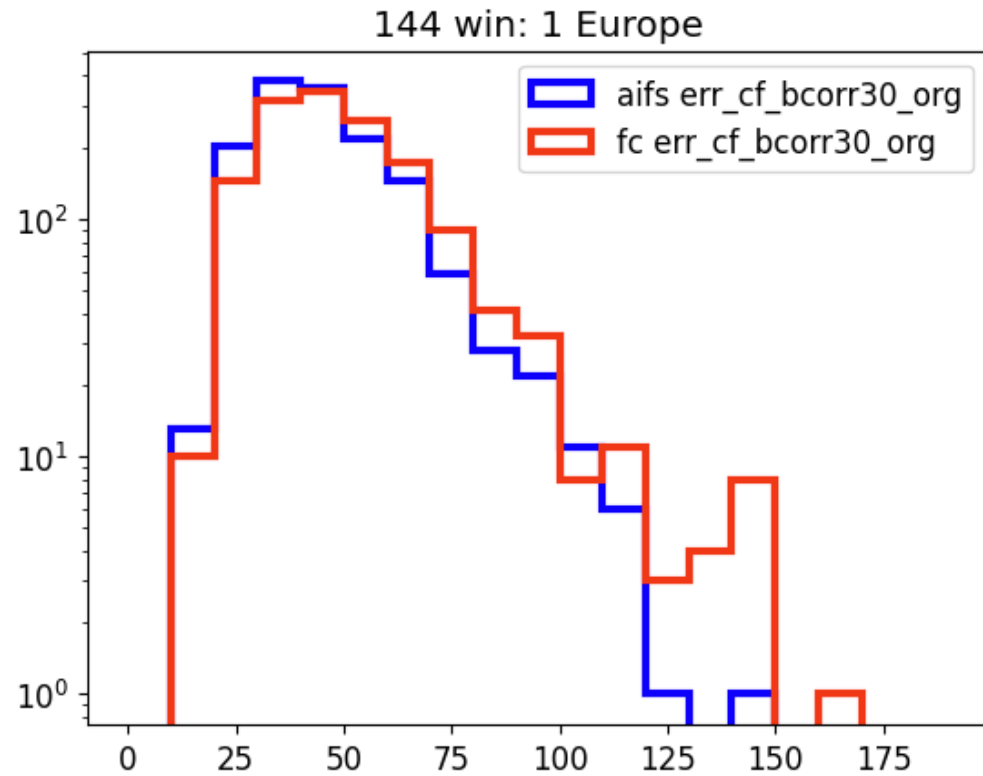
- Chaotic error growth
- Initial condition error
- Real / Model activity
- “Random” model error



- AIFS initialised from ECMWF analysis



Frequency of large errors over Europe day 6



I hope you watched
the poster from
Tobias Selz!