simon.lang@ecmwf.int

Machined-learned weather forecasting with AIFS

Simon Lang and ECMWF Colleagues

AIFS - ECMWF's data-driven forecasting system

A Preprint

Simon Lang* Mihai Alexe* Matthew Chantry Jesper Dramsch Florian Pinault Baudouin Raoult Mariana C. A. Clare Christian Lessig Michael Maier-Gerber Linus Magnusson Zied Ben Bouallègue Ana Prieto Nemesio Peter D. Dueben Andrew Brown Florian Pappenberger Florence Rabier

May 2024

Abstract

Machine learning-based weather forecasting models have quickly emerged as a promising method-ology for accurate medium-range global weather forecasting. Here, we introduce the Artificial Intel-

AIFS-CRPS: Ensemble forecasting using a model trained with a loss function based on the Continuous Ranked Probability Score

A Preprint

Simon Lang Mihai Alexe Mariana C. A. Clare Christopher Roberts Rilwan Adewoyin Zied Ben Bouallègue Matthew Chantry Jesper Dramsch Peter D. Dueben Sara Hahner Pedro Maciel Ana Prieto-Nemesio Cathal O'Brien Florian Pinault Jan Polster Baudouin Raoult Steffen Tietsche Martin Leutbecher

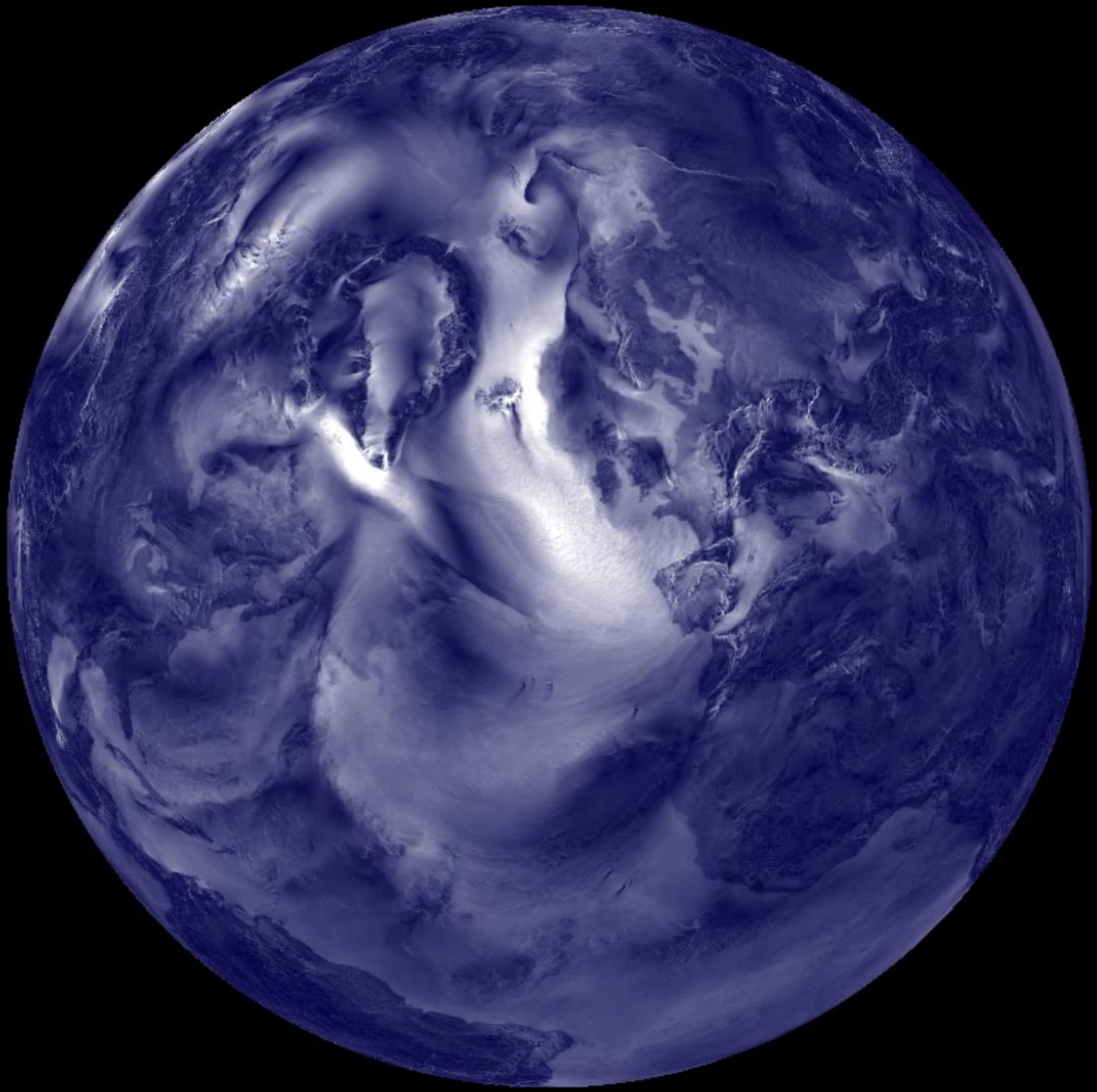
European Centre for Medium-Range Weather Forecasts (ECMWF)

December 23, 2024

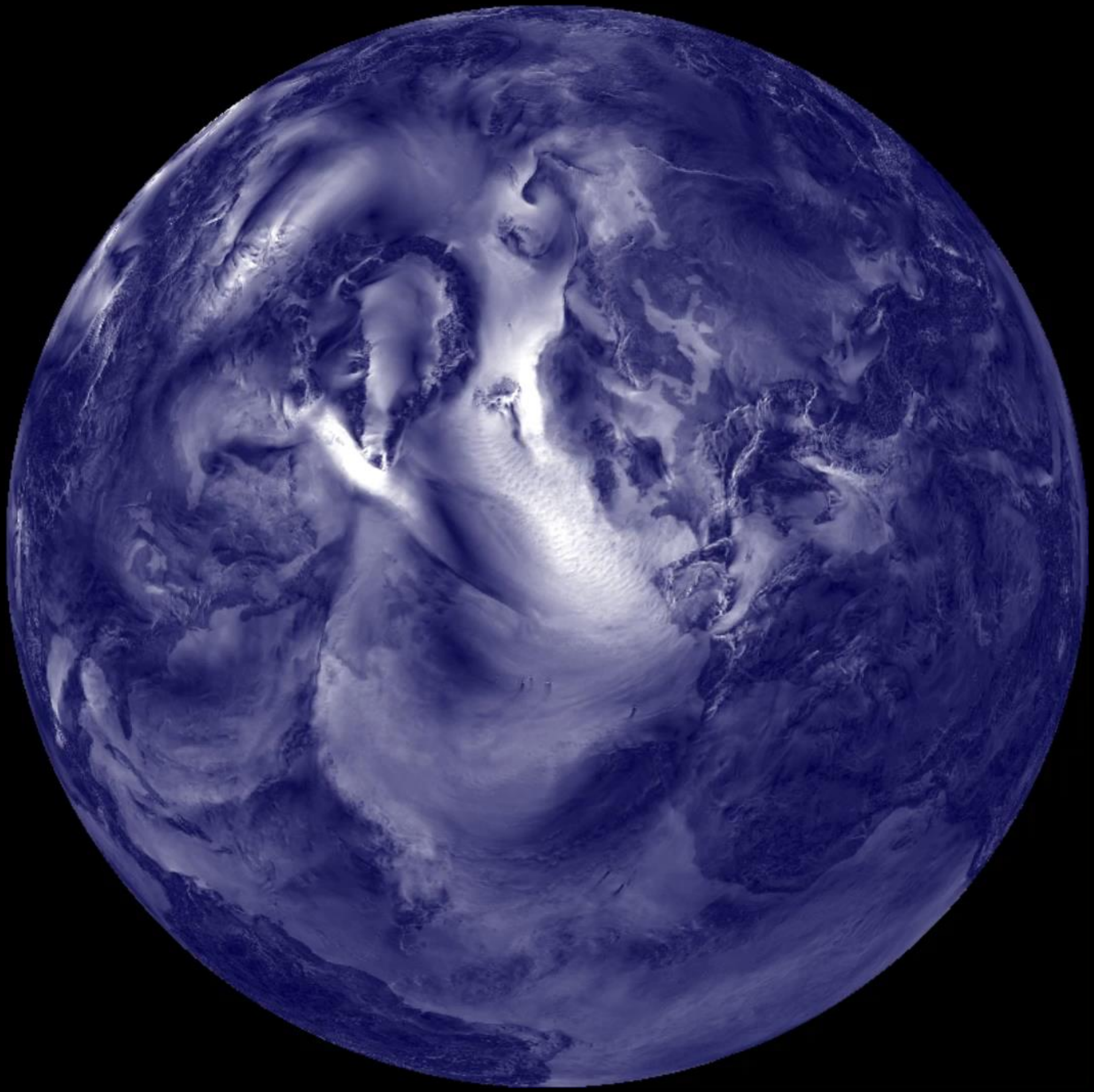
Abstract

Over the last three decades, ensemble forecasts have become an integral part of forecasting the weather. They provide users with more complete information than single forecasts as they permit to estimate the probability of weather events by representing the sources of uncertainties and accounting for the day-to-day variability of error growth in the atmosphere. This paper presents a novel approach

IFS 10m wind gusts, 2020-12-04 00 UTC 720h forecasts, 9 km spatial resolution



Control Member

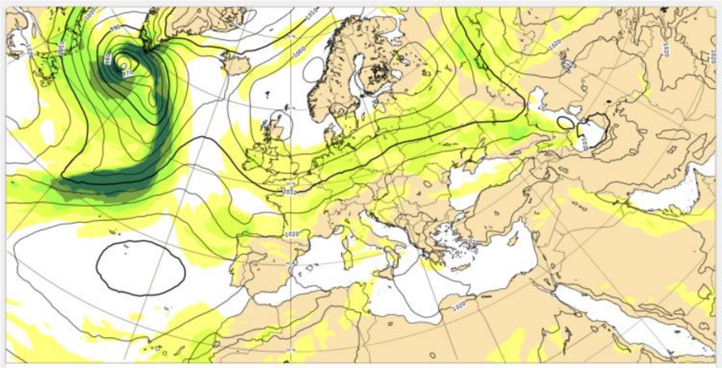


Perturbed member 1

Weather Forecasts – NWP? Data Driven?

Traditionally weather forecasts are generated by running an NWP model – computer code that has been designed to represent the physical processes governing the evolution of the atmosphere running on 1000s of CPUs. But can you produce a forecast without one?

Analysis



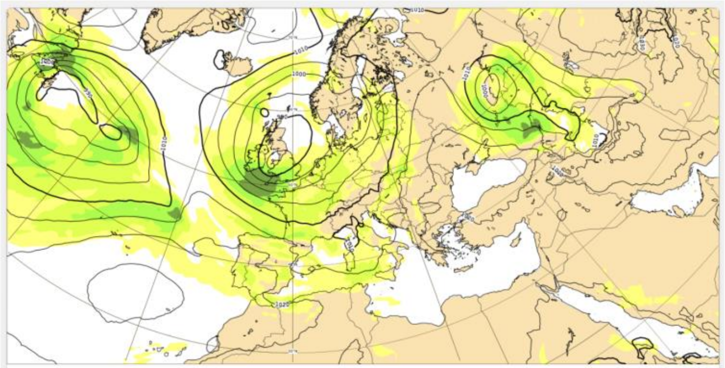
Fusion of short-range forecast with latest observations

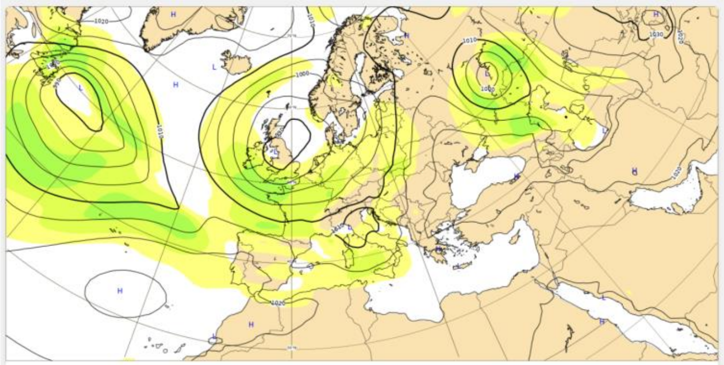
NWP Model

Data Driven Model

Learned from 40 years of analyses

Forecast

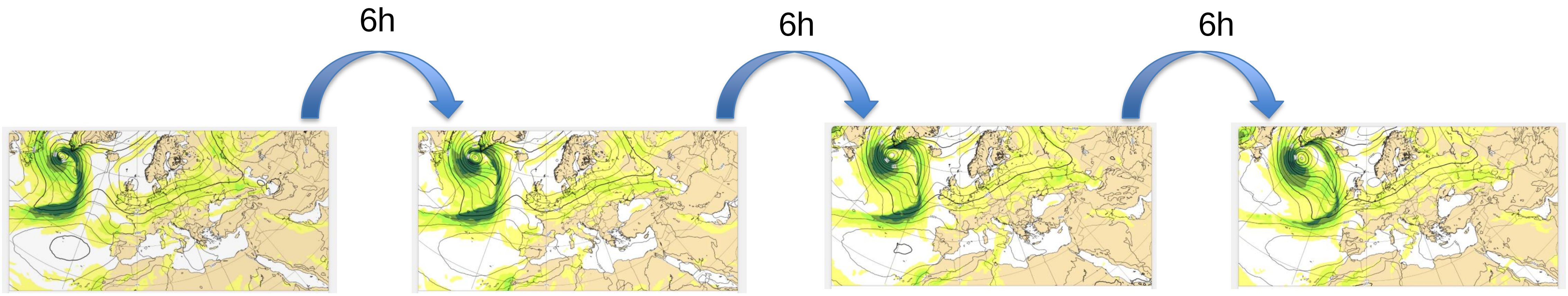




Weather Forecasts – NWP? Data Driven?

Recently it emerged, you can … (see Mat's talk for a summary of developments)

The model learns from ca. 40 years of ECMWF's ERA5 re-analysis data, stepping e.g. 6h from analysis to analysis ; then fine-tuning on oper analyses …

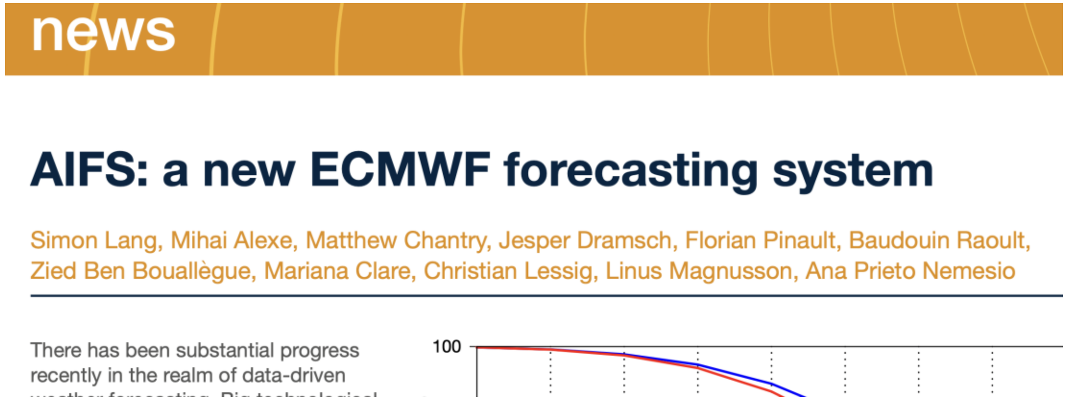


The forecast is then autoregressively stepping 6h into the future $x_n = f(x_{n-1})$ …

Can be much cheaper ~ 1000x, Skill?

AIFS - Artificial Intelligence Forecasting System…

AIFS - ECMWF's data-driven forecasting system. arXiv.
https://doi.org/10.48550/ARXIV.2406.01465


news

AIFS: a new ECMWF forecasting system

Simon Lang, Mihai Alexe, Matthew Chantry, Jesper Dramsch, Florian Pinault, Baudouin Raoult, Zied Ben Bouallègue, Mariana Clare, Christian Lessig, Linus Magnusson, Ana Prieto Nemesio

There has been substantial progress recently in the realm of data-driven

First implementation (~ 1deg resolution) in 2023, following Keisler 2022 and Lam et. al 2022:

- GNN architecture: Interaction Networks (Battaglia et. al 2016)
- Graph representation, hidden multi-scale mesh, edge features

-> decision: re-factor AIFS to create Anemoi -> make it accessible for a wider community to build models on top etc.

First update beginning of 2024, -> 0.25 deg:

- Attention / Transformer based GNN for encoder, decoder (Shi et al., 2021)
- Transformer backbone in processor (with sliding window, e.g. Child et al. 2019, Jiang et al. 2023)

June 2024, Diffusion based ensemble enters real-time mode, 1 deg resolution, 50 members

December 2024, AIFS weights are published

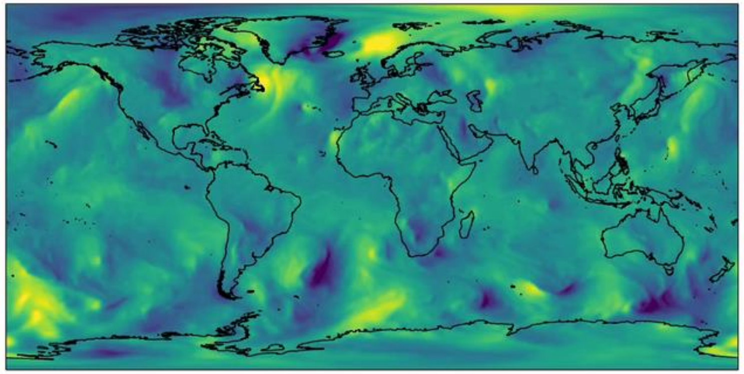
ECMWF has taken the Artificial Intelligence Forecasting System (AIFS) into operations today, 25 February 2025, to run side by side with its traditional physics-based Integrated

AIFS-Single becomes operational: 25th February 2025

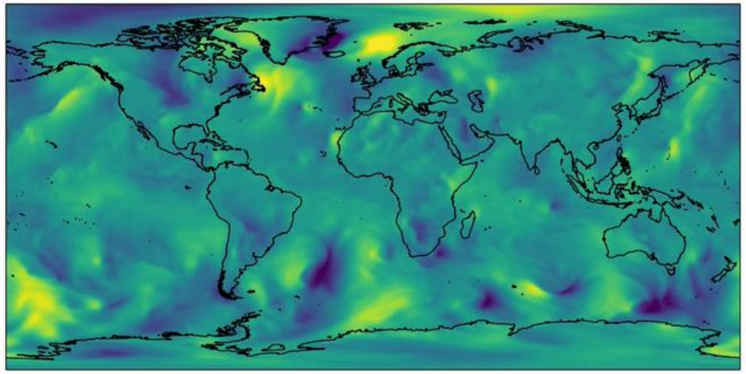
For more, pls see -> AIFS blog: https://www.ecmwf.int/en/about/media-centre/aifs-blog

AIFS - Artificial Intelligence Forecasting System…

- Attention based GNN for encoder, decoder
- Transformer backbone in processor



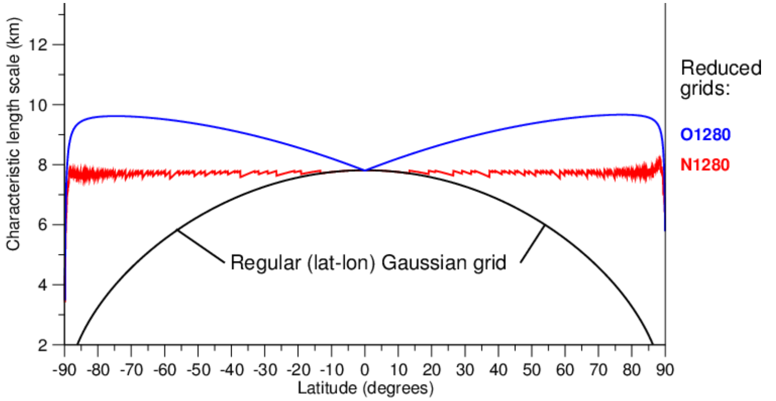
IFS



AIFS forecast v0.1

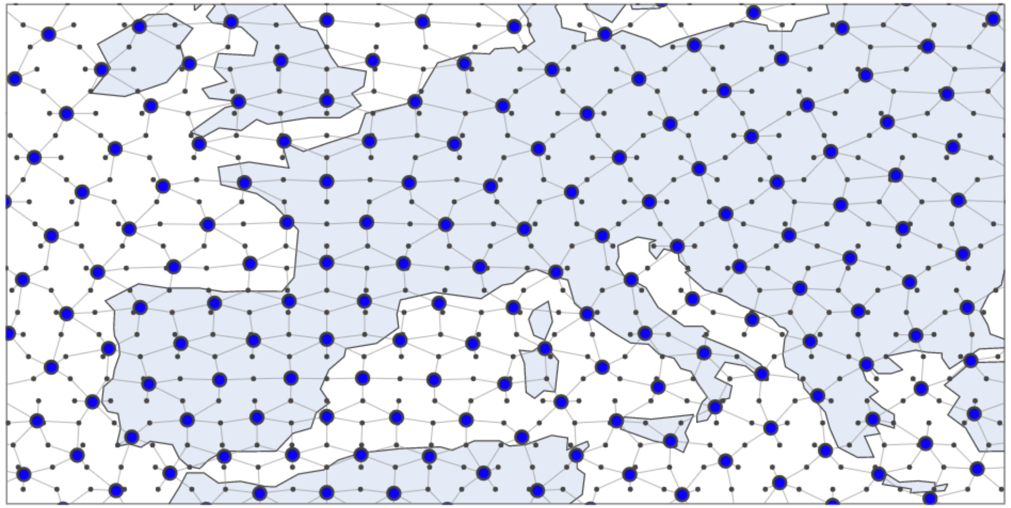
Why GNN Encoder / Decoder: can handle arbitrary input / output grids, local and ad hoc grid refinement, changing grids etc. ; attractive for use in earth system science

AIFS – Encoder and Decoder



AIFS works with the native IFS reduced gaussian grids ; possible to split model and input data across multiple GPUs to handle large memory requirements

Encoder, GNN



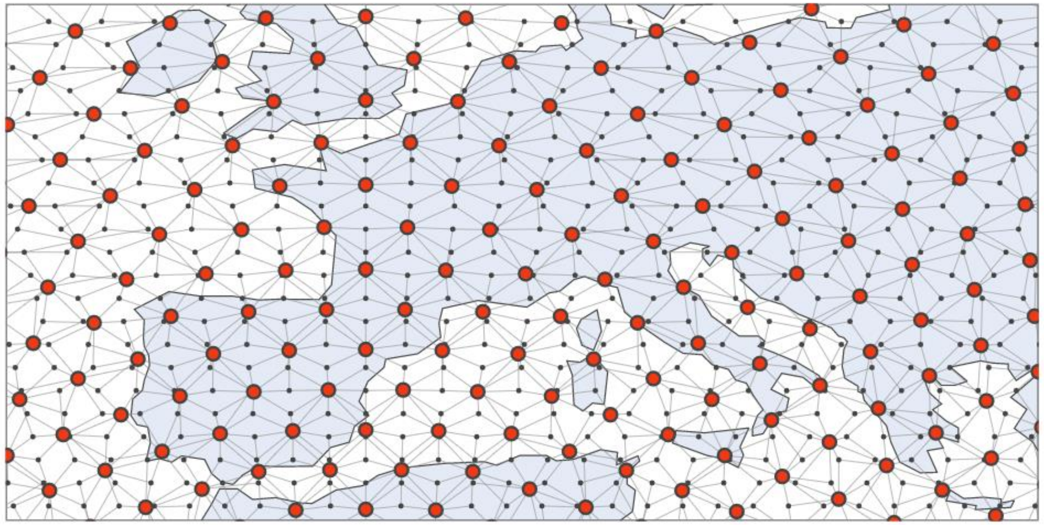
Era5 – n320

~ 540 000 Nodes,
~ 1 Million Edges

16 Processor Layers

O96 ~ 40 000 Nodes

Decoder, GNN

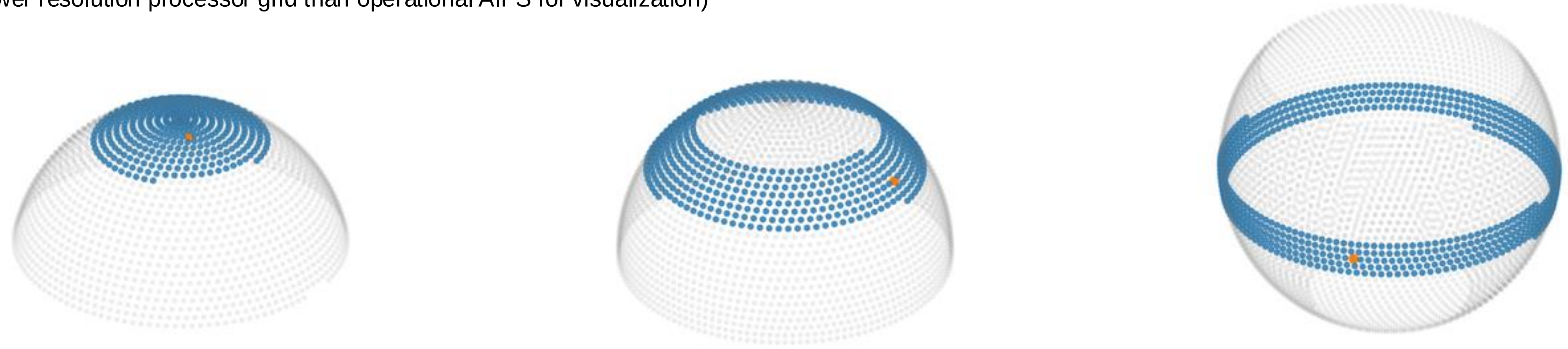


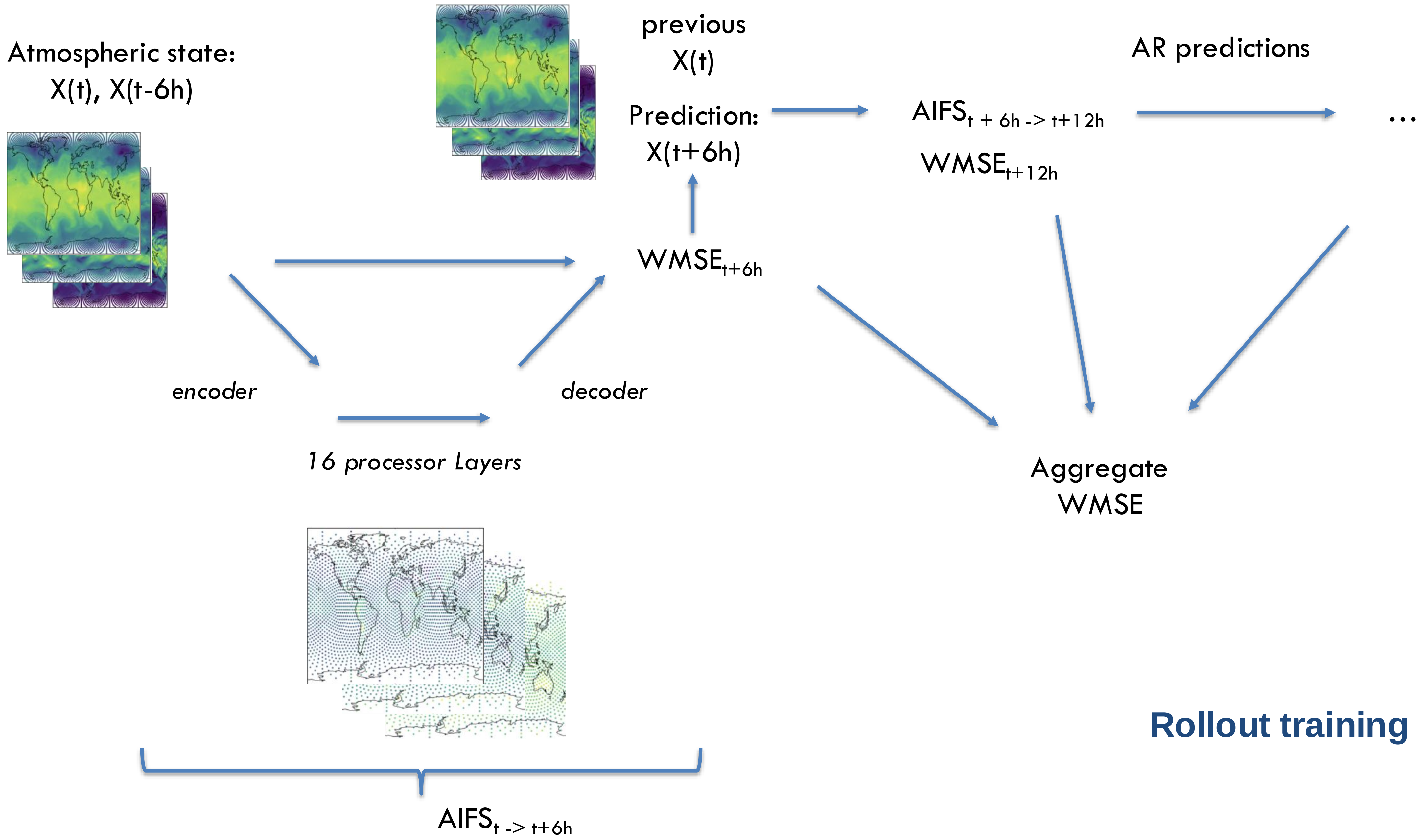
Era5 – n320

AIFS – Processor

Transformer (like LLMs) that works with a sliding attention window -> attention bands around the globe

Red: target node
Blue: Nodes target node attends to in one processor layer
Grey: How far information can travel within e.g. 6 processor layers
(here lower resolution processor grid than operational AIFS for visualization)

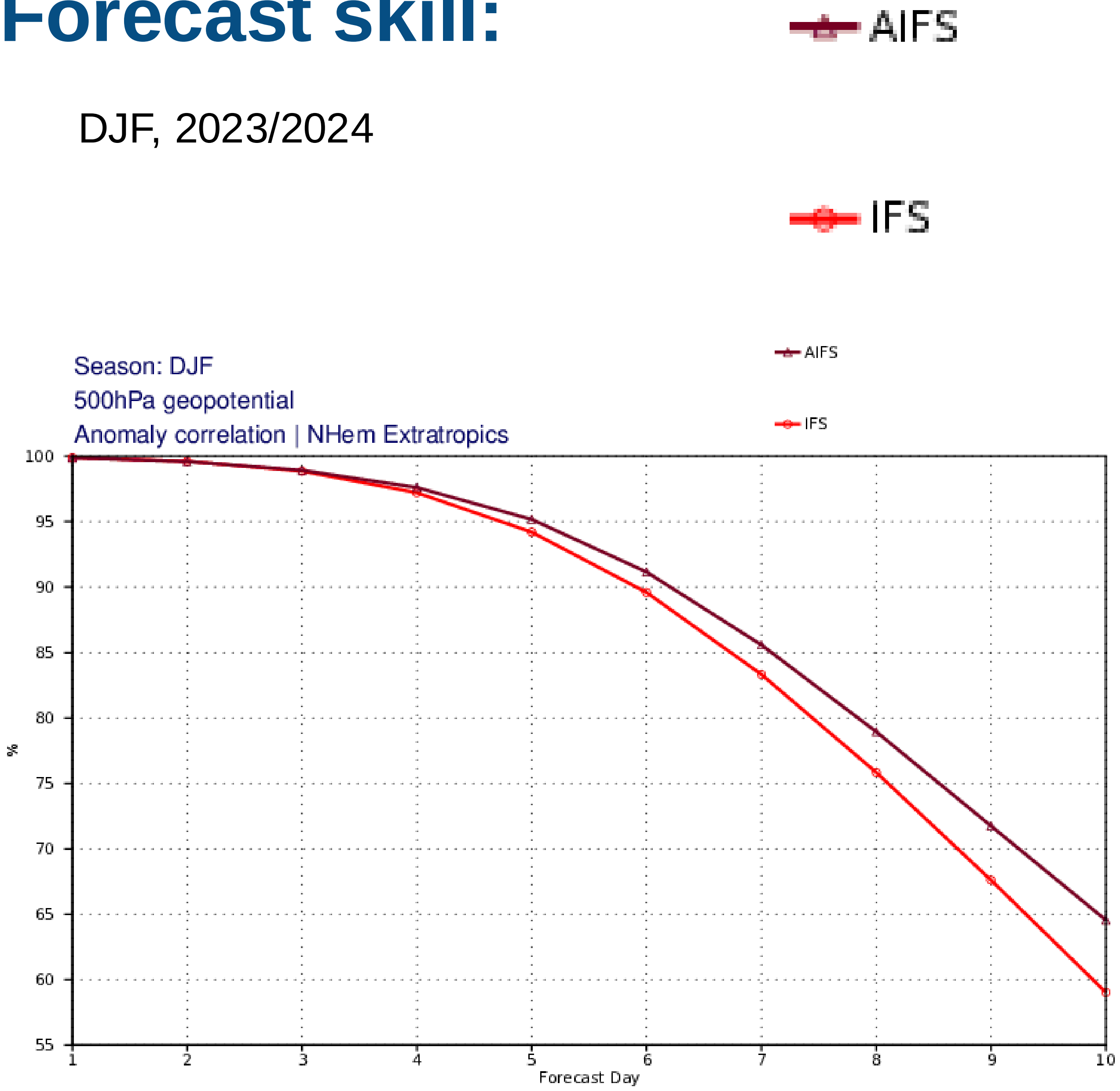


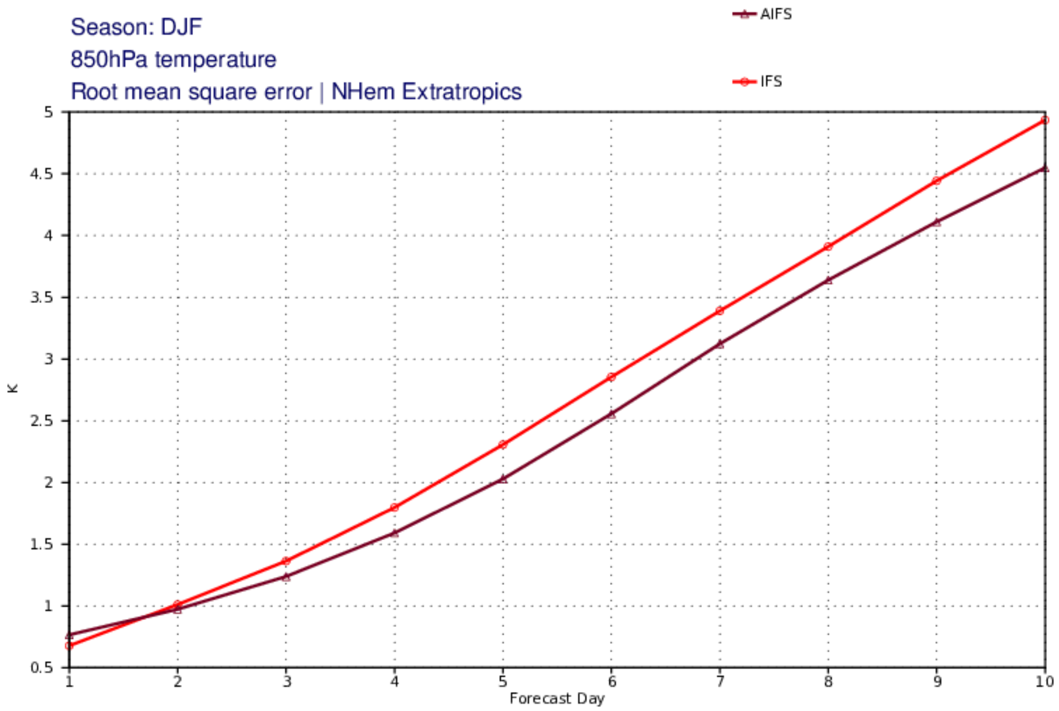


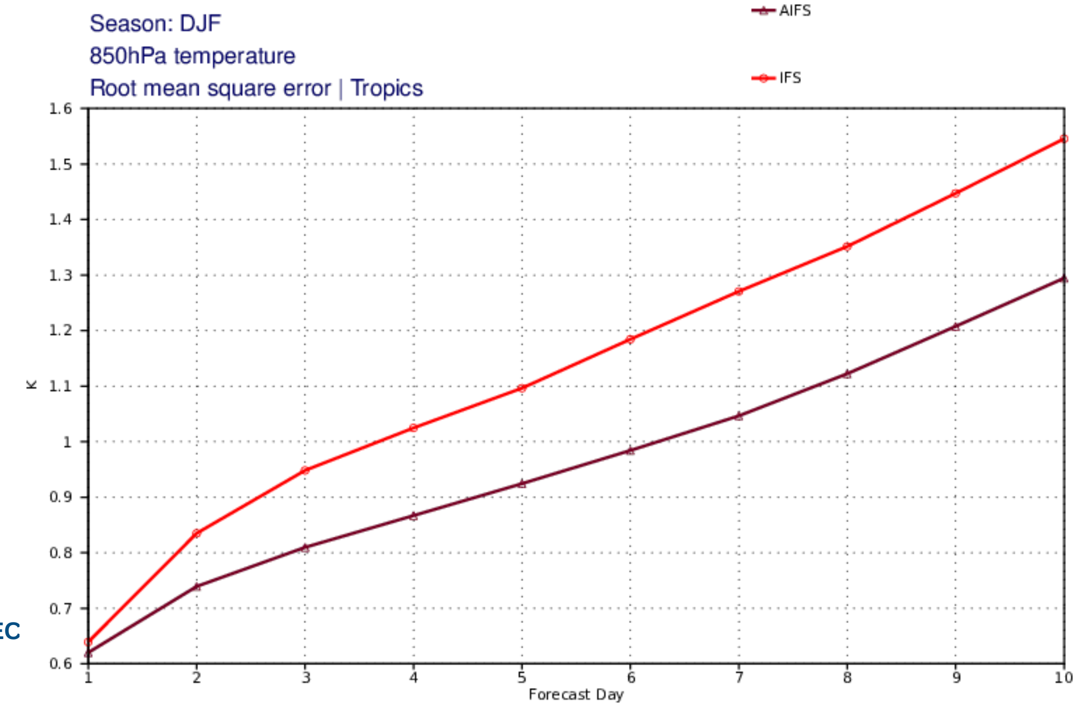
Rollout training

Forecast skill:

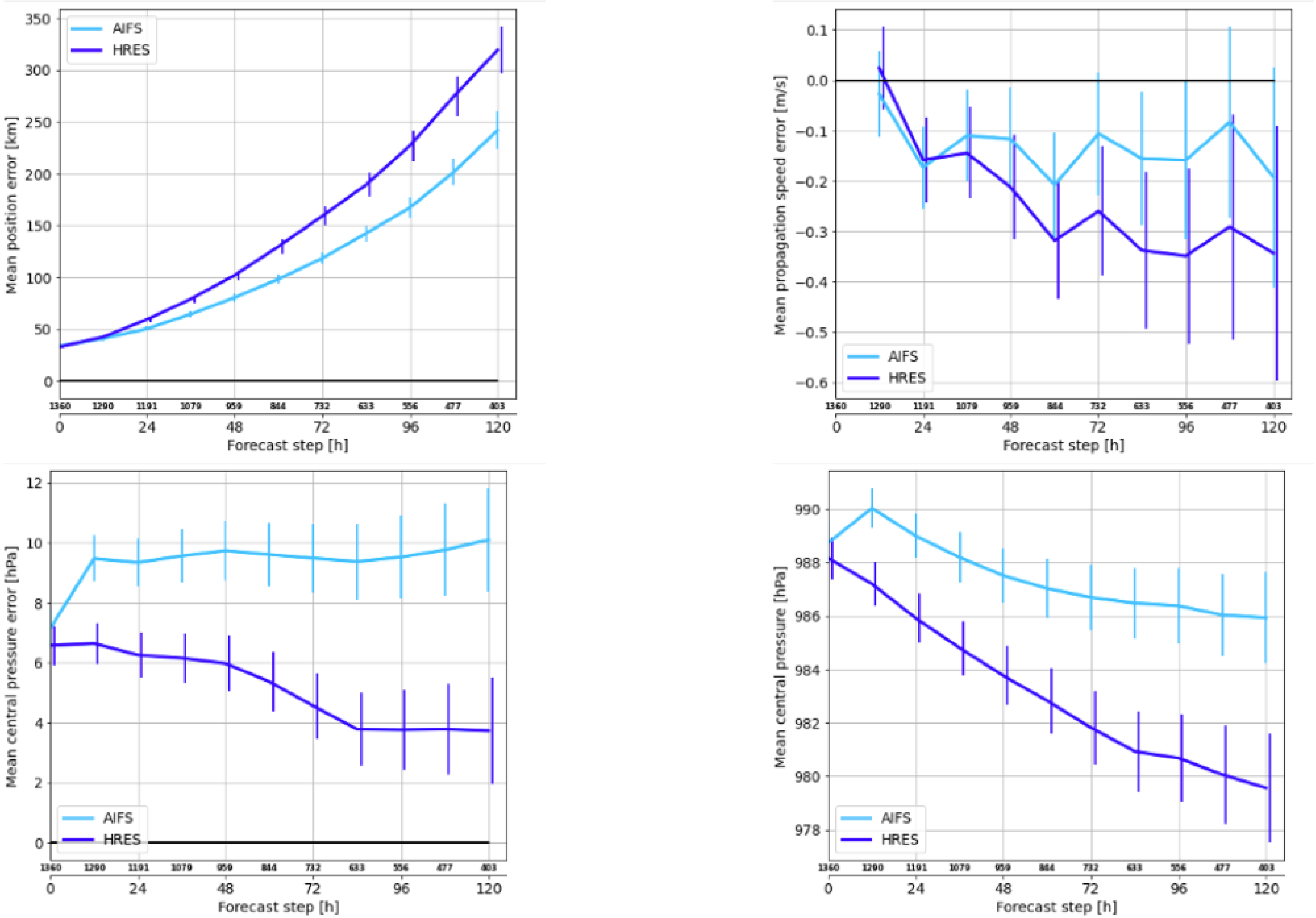
DJF, 2023/2024



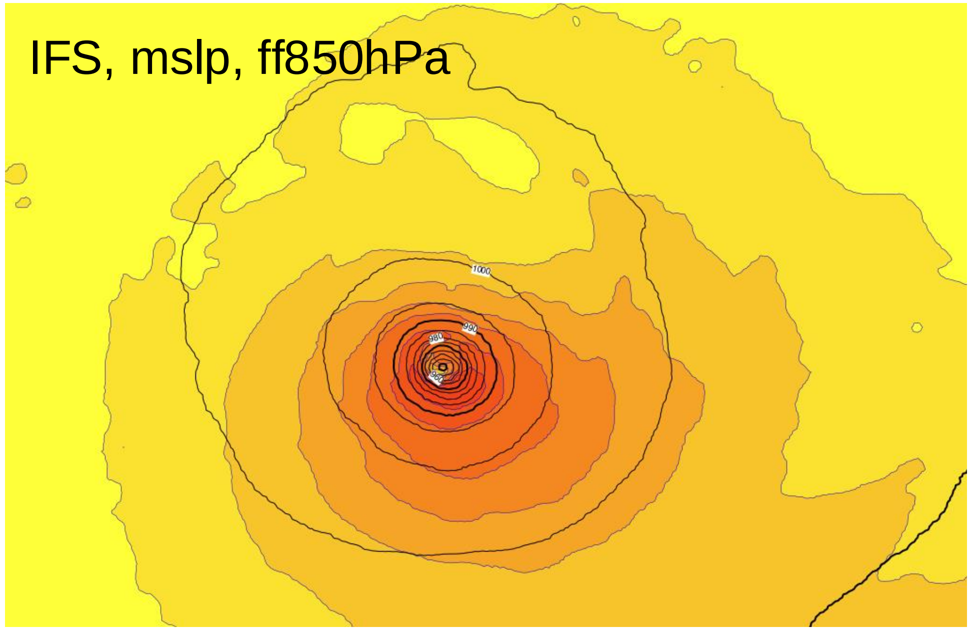


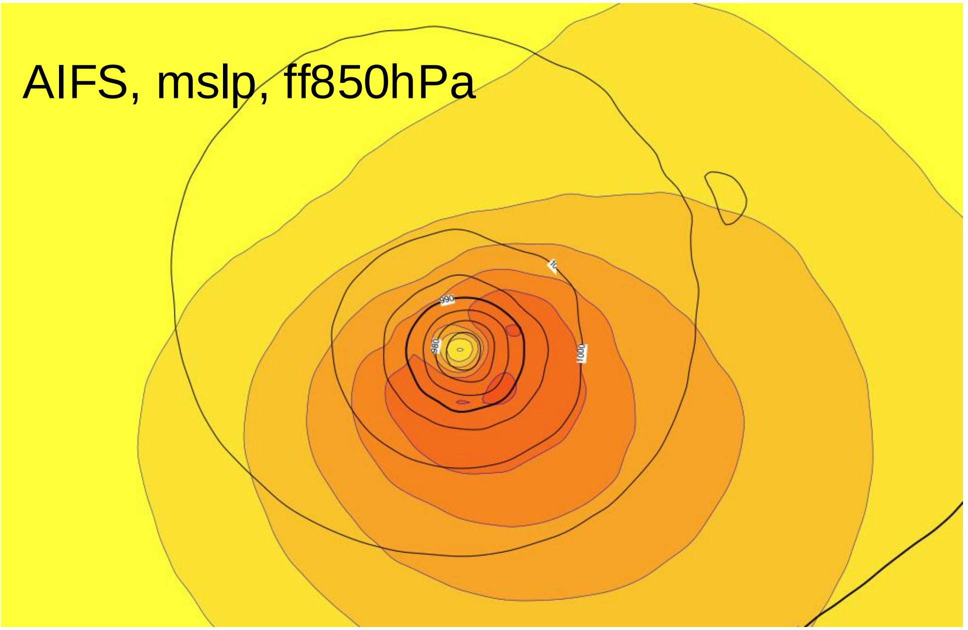


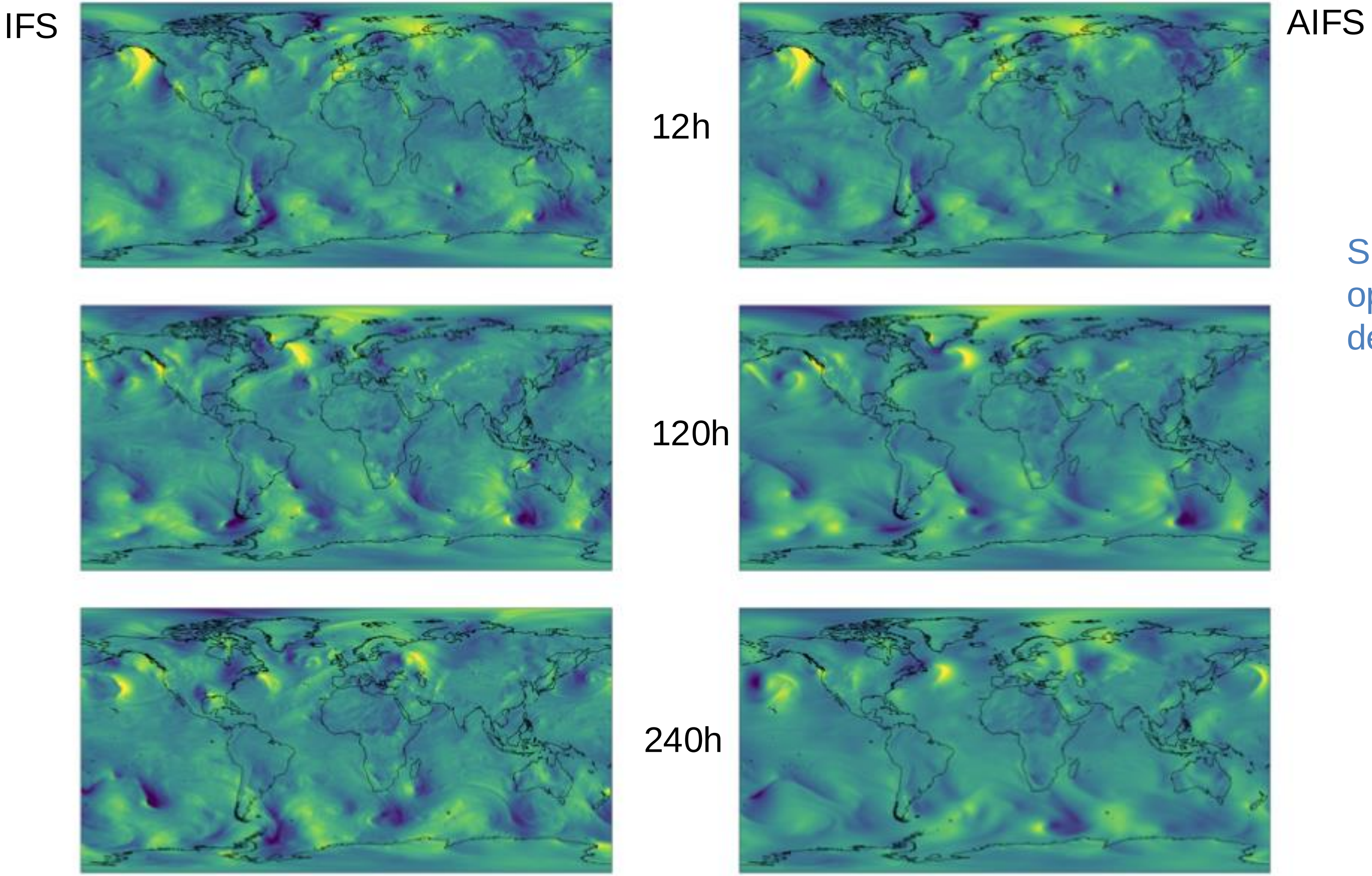
Forecast skill, TCs:



Comparison of IFS and AIFS Mean tropical cyclone position error (upper-left), mean tropical cyclone propagation speed (upper-right), tropical cyclone mean central pressure error (bottom-left) and tropical cyclone mean central pressure (bottom right) of AIFS (light blue) and IFS (dark blue), January 2022 to December 2023.





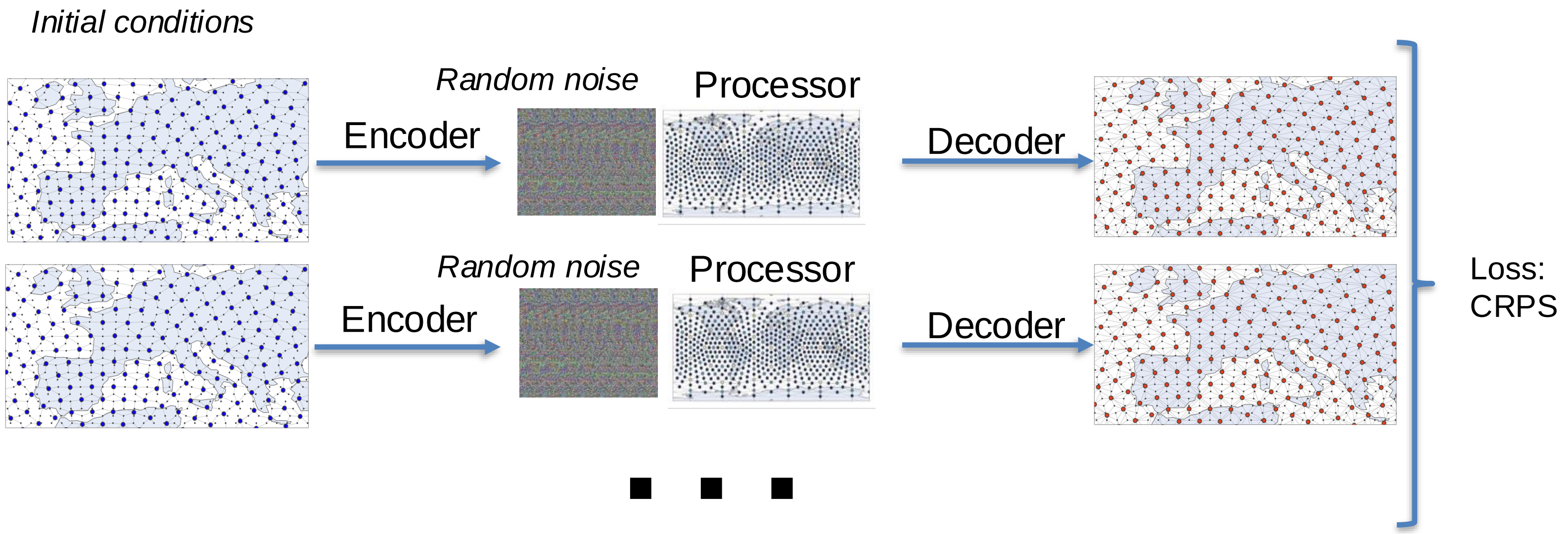


Smoothing due to MSE optimization – deterministic training

Meridional wind at 850 hPa: IFS (left) and AIFS (right) for 1 January 2023 00 UTC date, 12 h (top), 120 h (middle) and 240 h (bottom) forecasts. For plotting, IFS and AIFS forecast fields were interpolated to a 0.25° regular latitude-longitude grid. The MSE objective used in AIFS training leads to more smoothing at longer lead times.

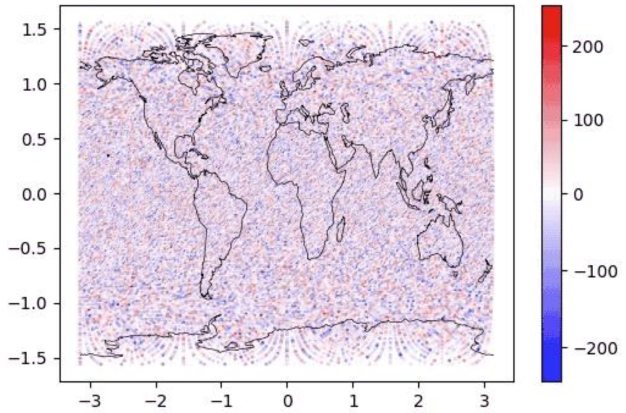
AIFS ensemble: two approaches

Instead of a MSE loss, learn an ensemble via optimizing probabilistic scores

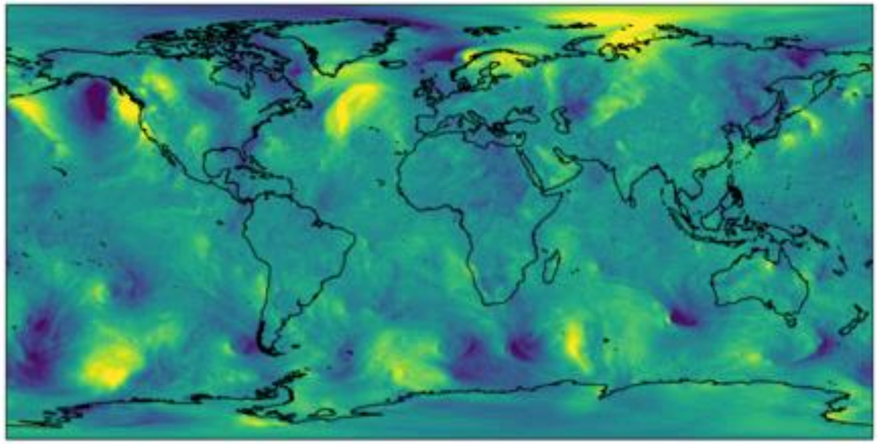


Create a forecast as de-noising task

-> diffusion models

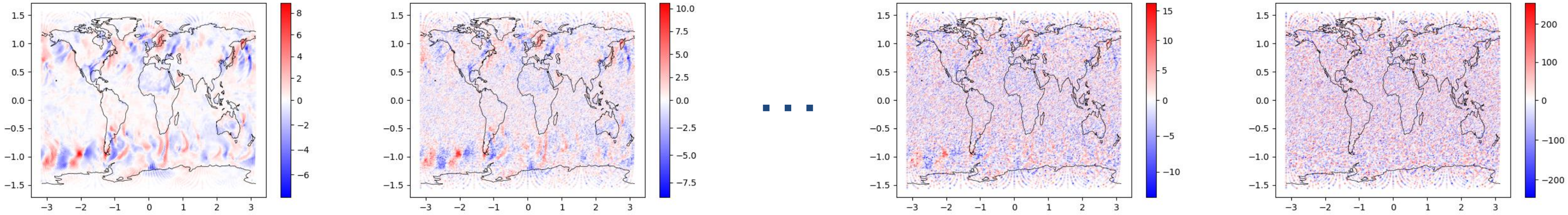


Probabilistic training encourages realistic levels of variability



Diffusion – AIFS-Diffusion:

Forward process: add noise



T_0, S_0 T_1, S_1 … T_{N-1}, S_{N-1} T_N, S_N

Reverse process (learned): remove noise
Model input: noised tendency, initial condition, noise level

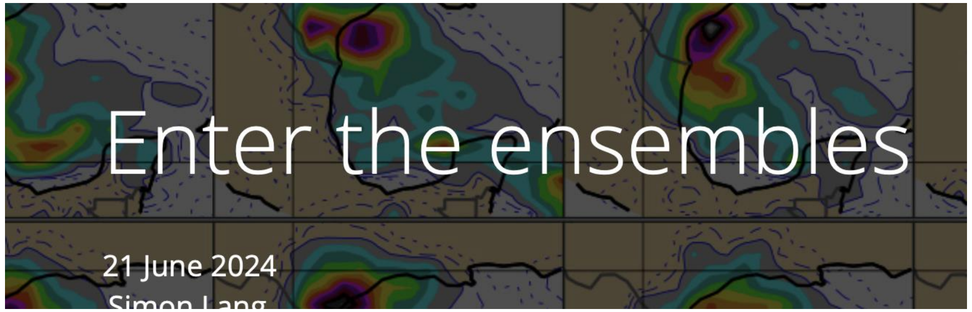
To generate a forecast (single member): start from pure noise with initial conditions, run the model e.g. 40 times, to slowly create a forecast tendency step by step ; repeat for the next forecast step etc. -> 12h step, 360h lead time, call the model 1200x

Diffusion – AIFS-Diffusion:

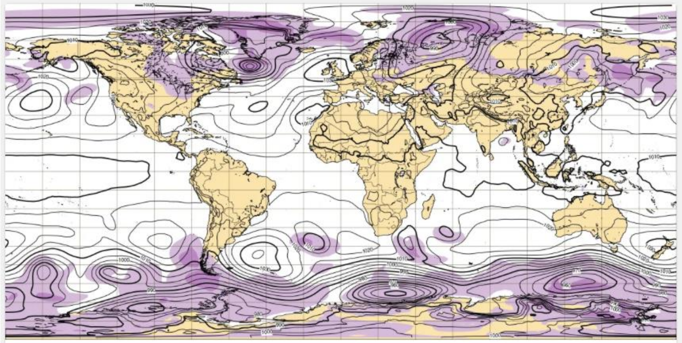
First test system running now 2x daily,

based on AIFS diffusion flavor,

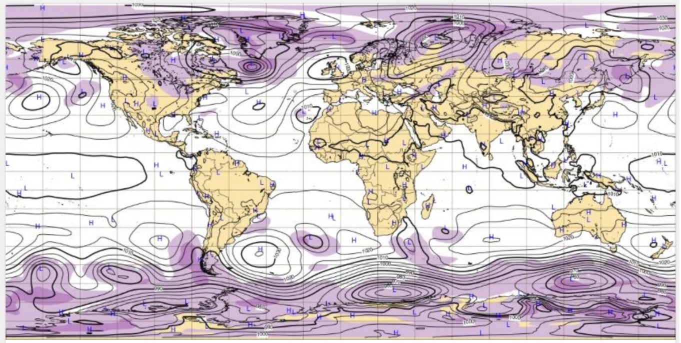
~ 1 deg resolution, 12h timestep, fine-tuned on oper analyses



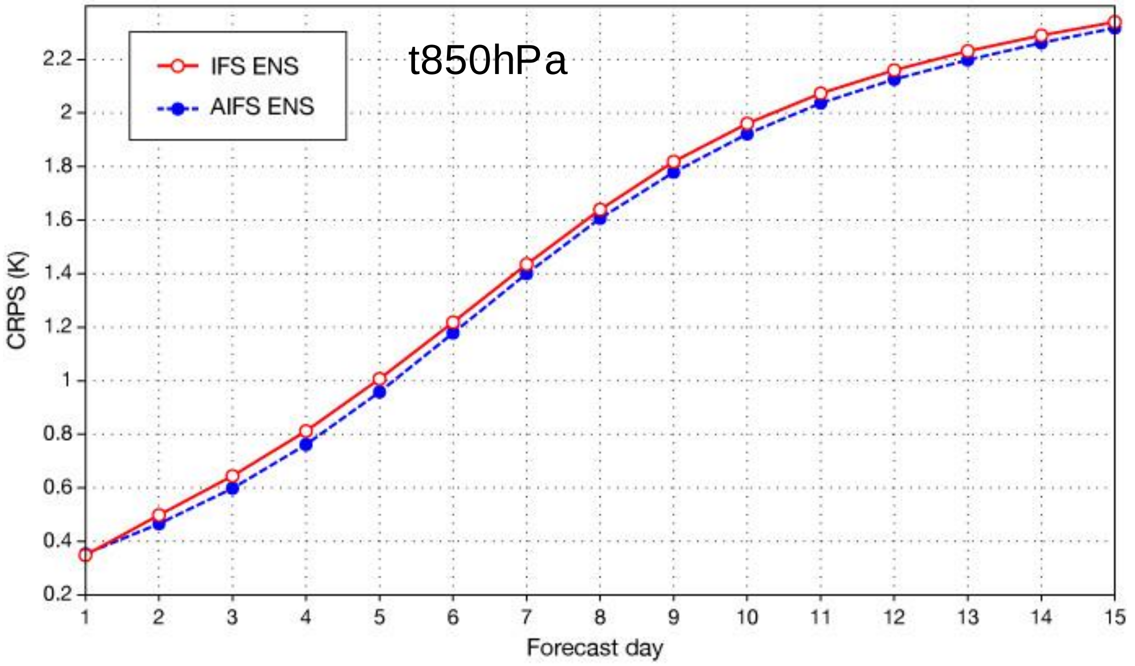
https://www.ecmwf.int/en/about/media-centre/aifs-blog/2024/enter-ensembles

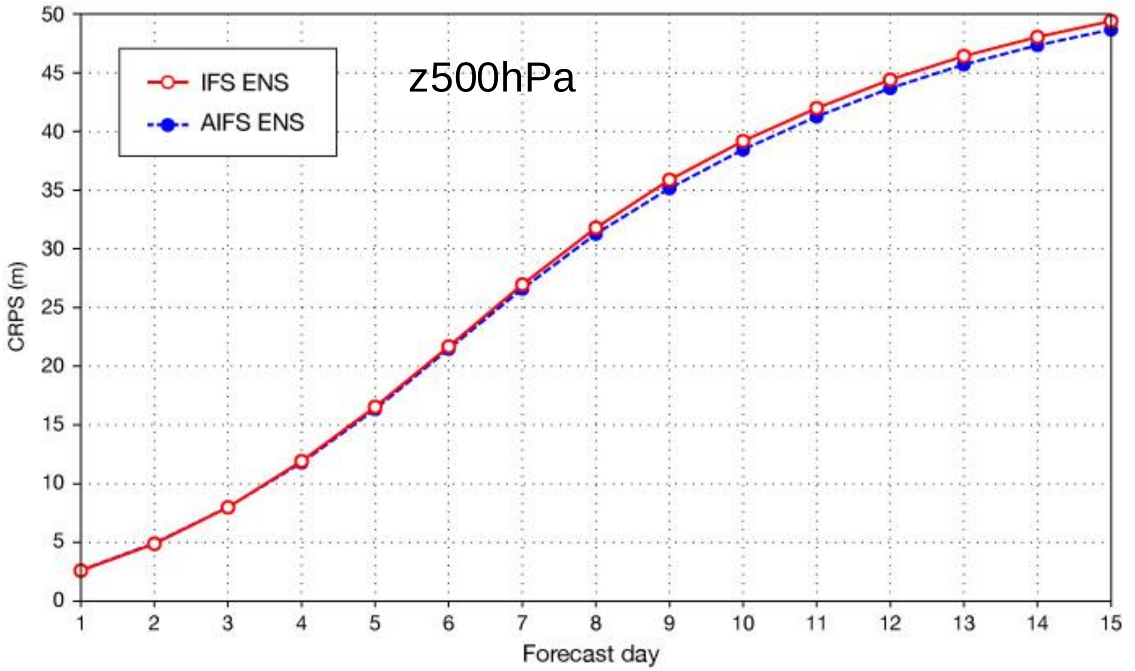


Ensemble mean and spread for mean sea level pressure



Experimental: AIFS ENS (ECMWF) ML model: Ensemble mean and spread for mean sea level pressure





Next: AIFS-CRPS:

Probabilistic training of AIFS:

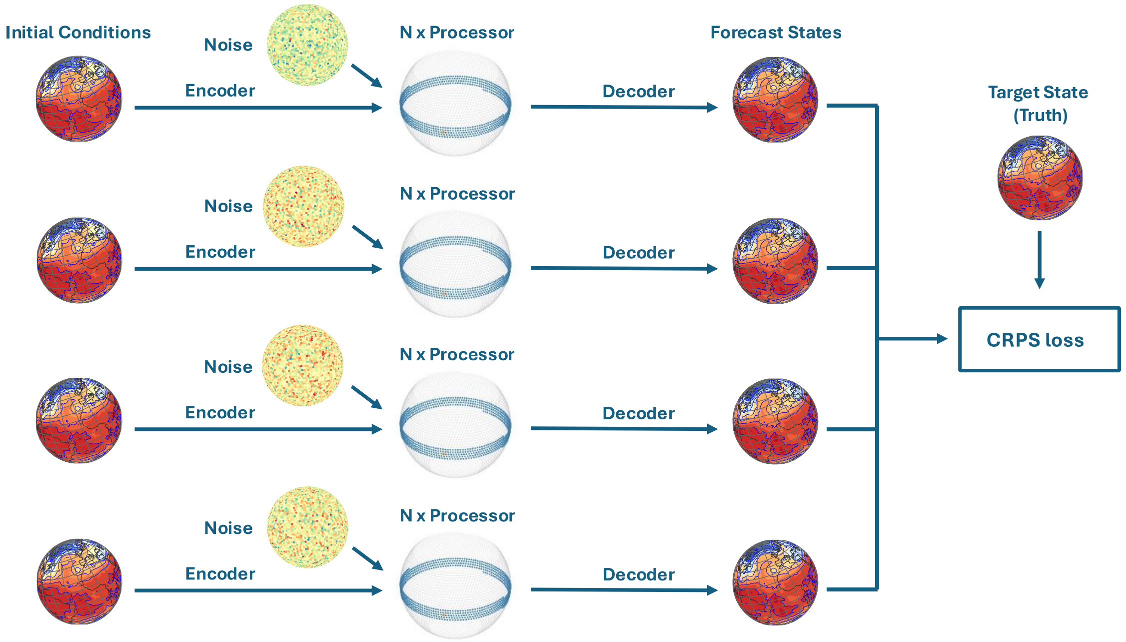
$$\mathrm{afCRPS}_\alpha := \alpha\,\mathrm{fCRPS} + (1-\alpha)\mathrm{CRPS}$$

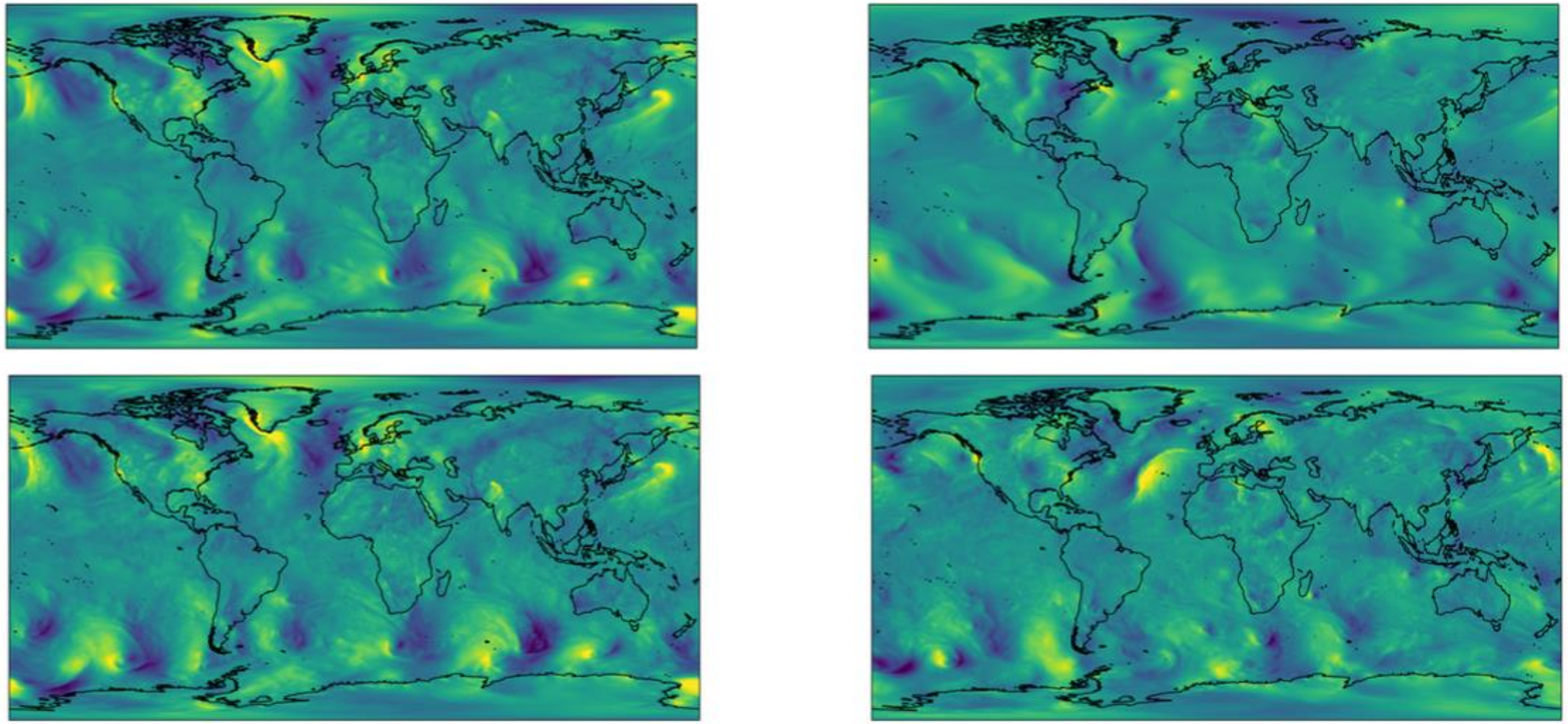
$$= \frac{1}{M}\sum_{j=1}^{M}|x_j - y| - \frac{M-1+\alpha}{2M^2(M-1)}\sum_{j=1}^{M}\sum_{k=1}^{M}|x_j - x_k|$$

$$= \frac{1}{M}\sum_{j=1}^{M}|x_j - y| - \frac{1-\epsilon}{2M(M-1)}\sum_{j=1}^{M}\sum_{k=1}^{M}|x_j - x_k|$$

Probabilistic training -> there is no such thing as an unperturbed control forecast

AIFS-CRPS runs with a 6h timestep





AIFS-single, day 1 and 10

AIFS-CRPS, day 1 and 10

https://arxiv.org/abs/2412.15832

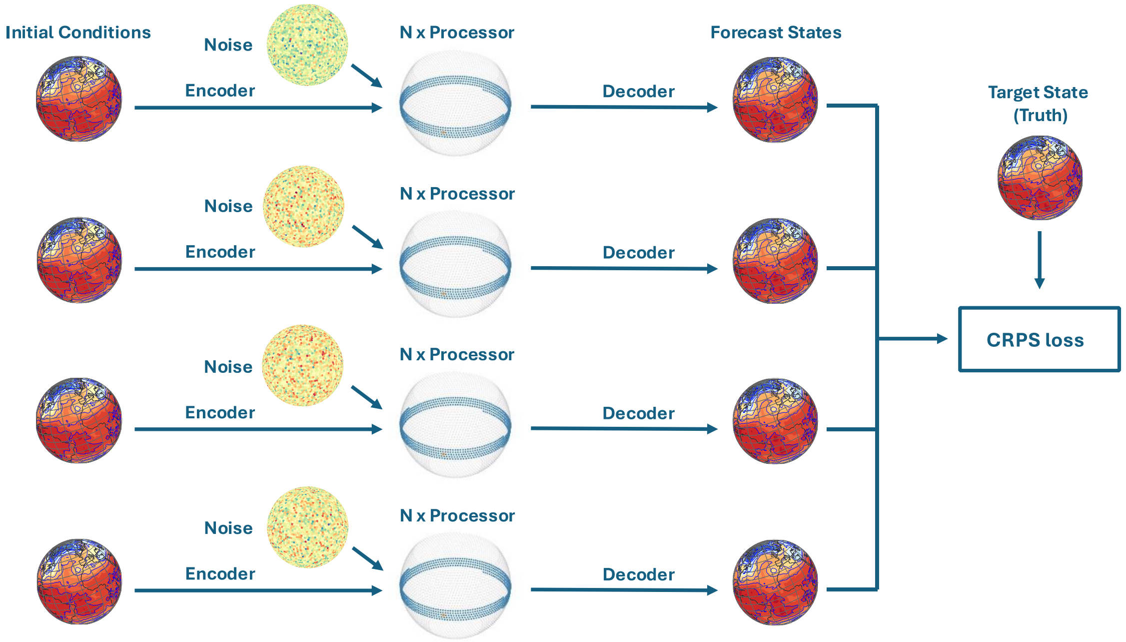
Next: AIFS-CRPS:

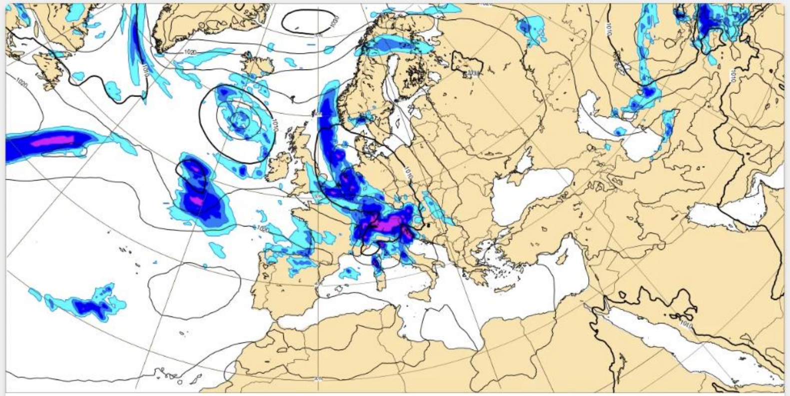
Probabilistic training of AIFS:

$$
\begin{aligned}
\mathrm{afCRPS}_\alpha &:= \alpha\,\mathrm{fCRPS} + (1-\alpha)\mathrm{CRPS} \\
&= \frac{1}{M}\sum_{j=1}^{M}|x_j - y| - \frac{M-1+\alpha}{2M^2(M-1)}\sum_{j=1}^{M}\sum_{k=1}^{M}|x_j - x_k| \\
&= \frac{1}{M}\sum_{j=1}^{M}|x_j - y| - \frac{1-\epsilon}{2M(M-1)}\sum_{j=1}^{M}\sum_{k=1}^{M}|x_j - x_k|
\end{aligned}
$$

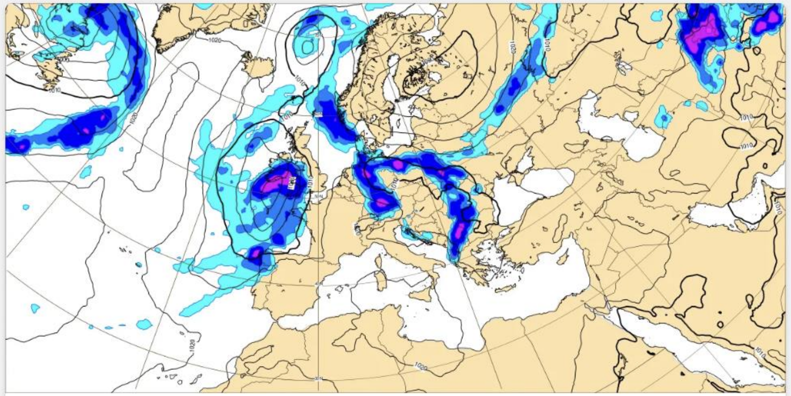
Probabilistic training -> there is no such thing as an unperturbed control forecast

AIFS-CRPS runs with a 6h timestep

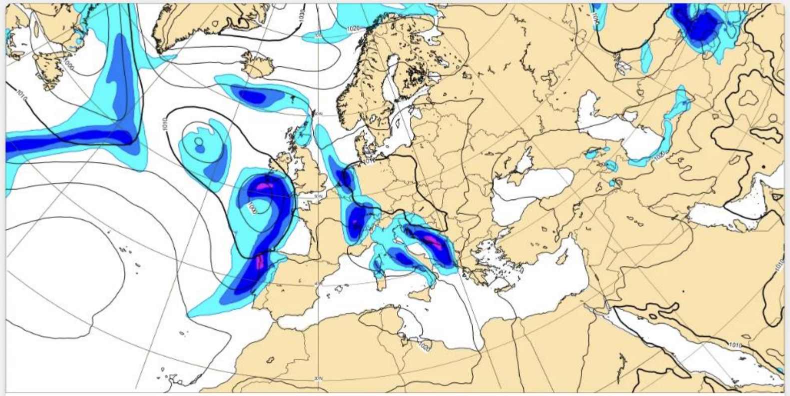




Rain and mean sea level pressure



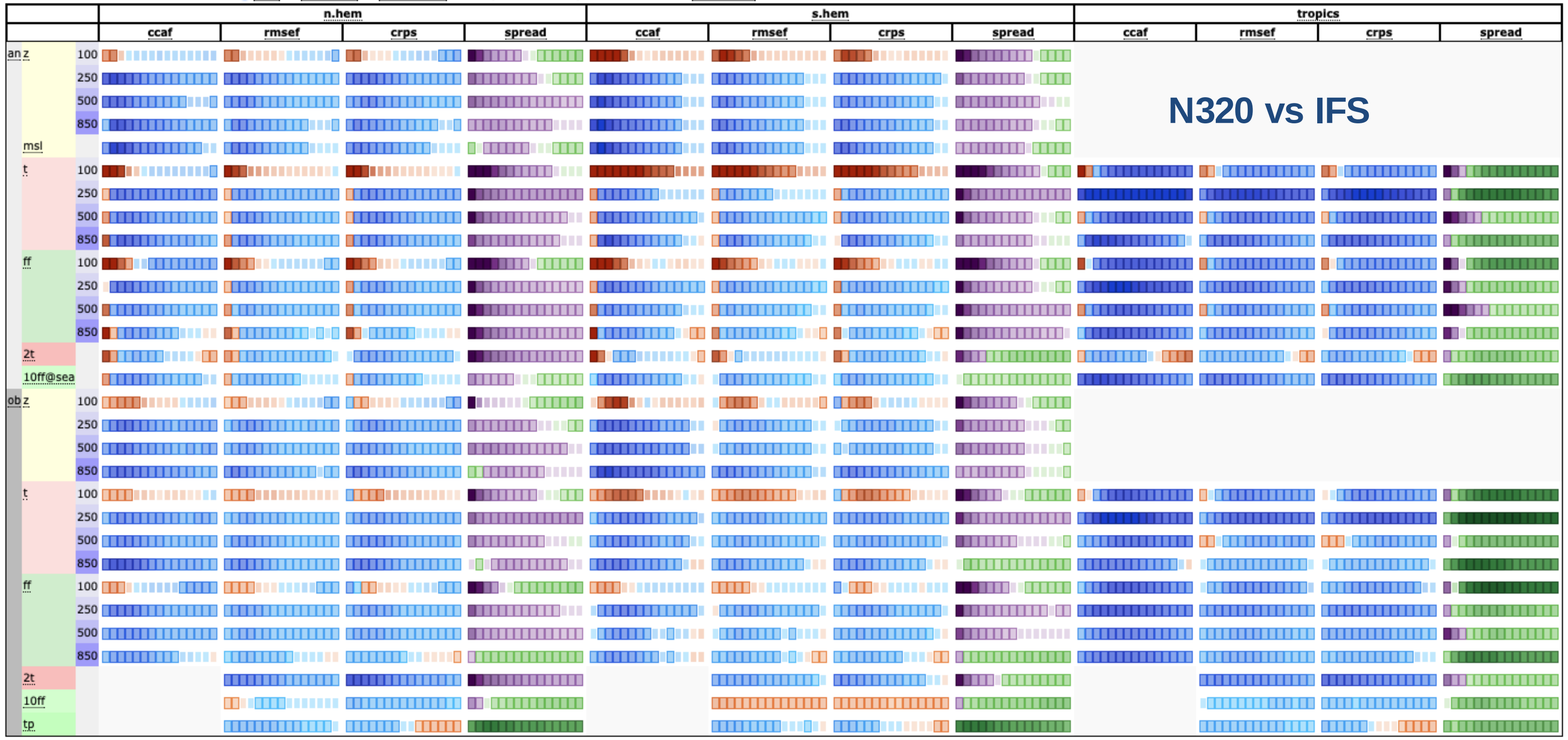
AIFS CRPS : Rain and mean sea level pressure



AIFS Single: Rain and mean sea level pressure

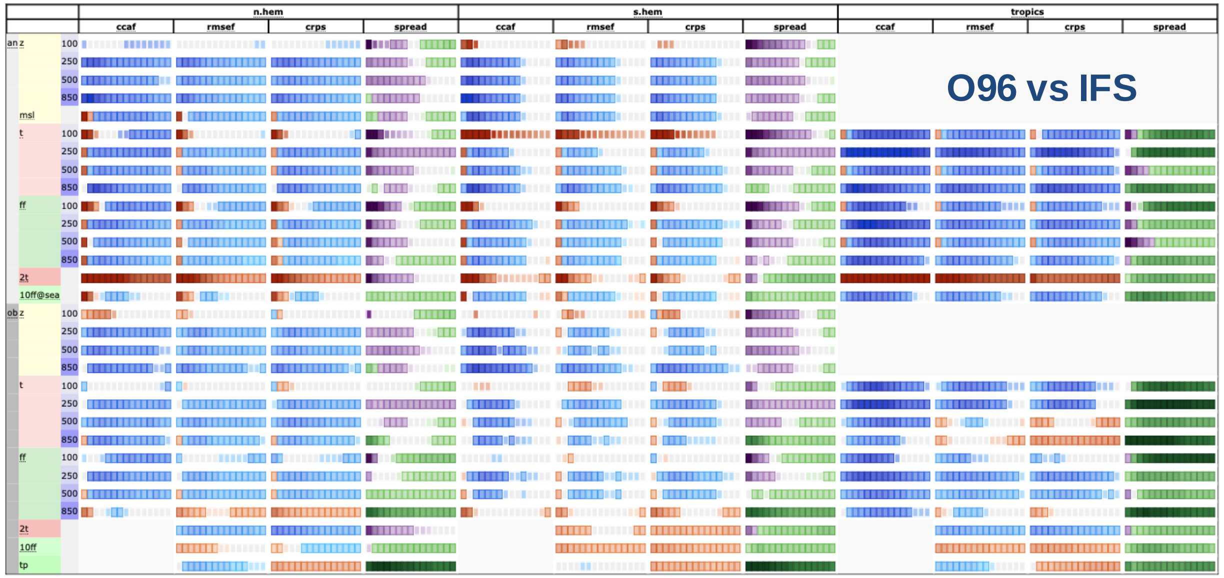
10-day forecasts

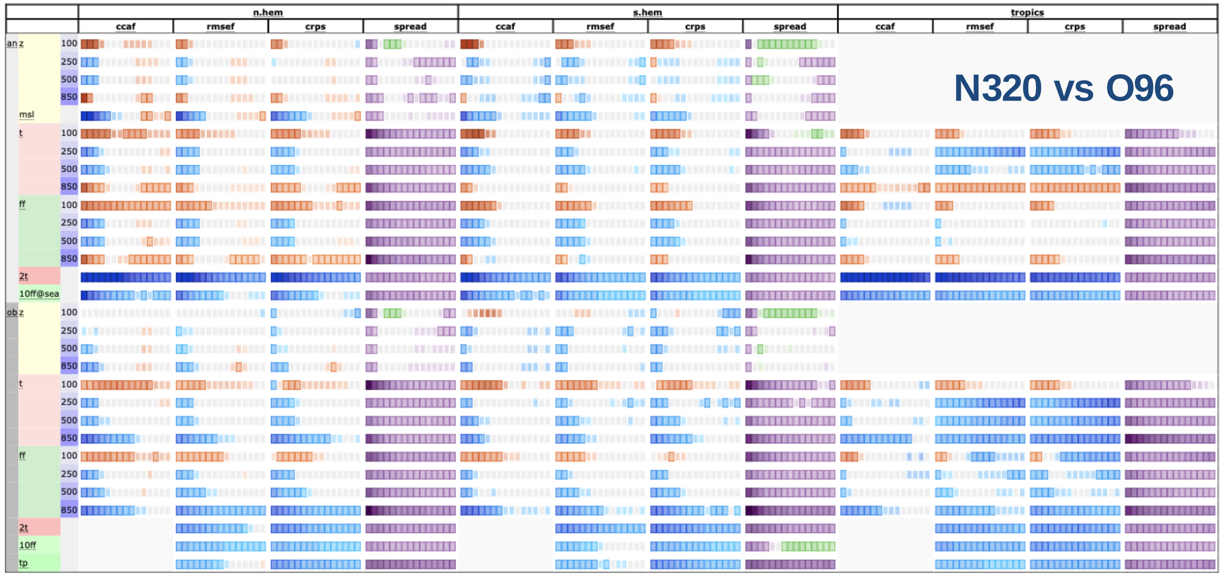
AIFS-CRPS Medium-Range evaluation, 50 member, O96, N320 vs O1280 IFS



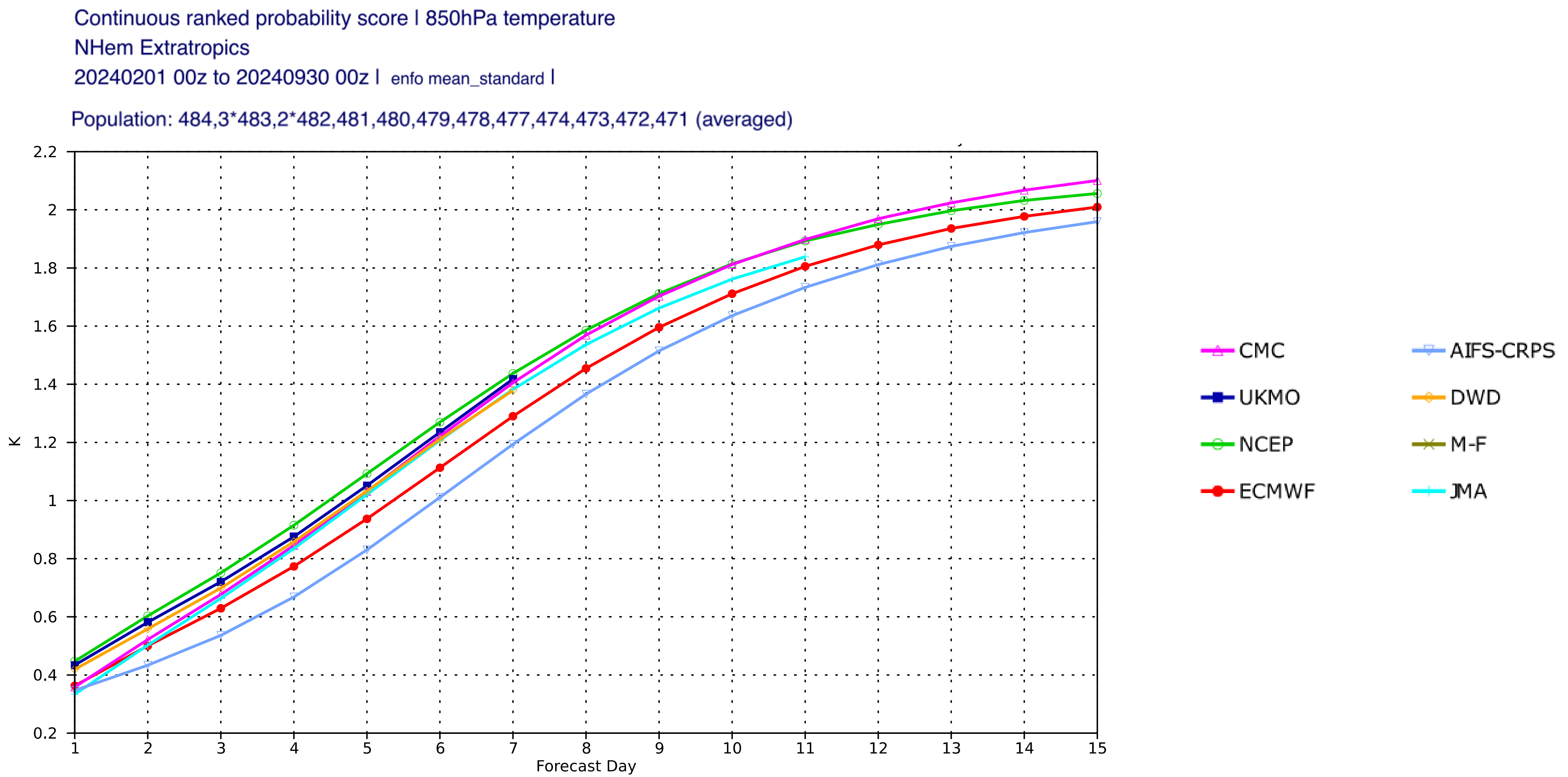
AIFS-CRPS O96 ~ 1.0deg
AIFS-CRPS N320 ~ 0.25deg
IFS 1280 ~ 0.1 deg

Verification: Oper An, Obs



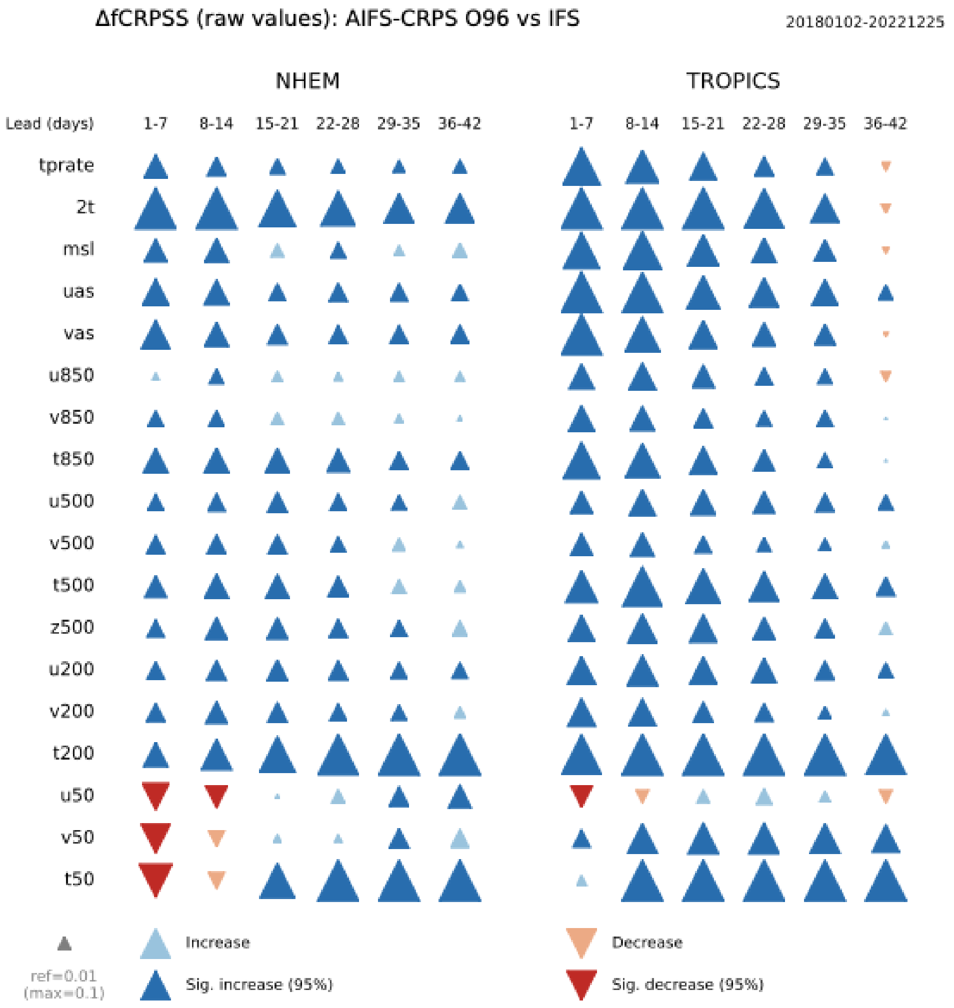


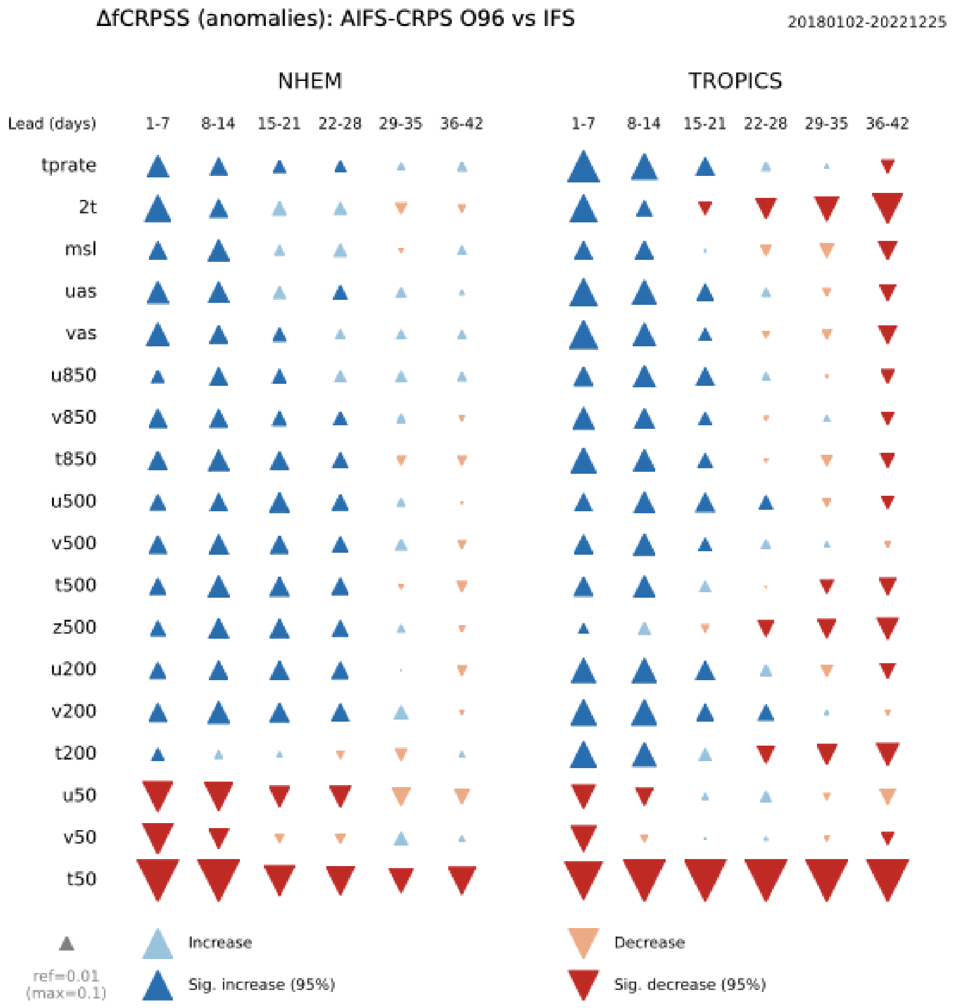
AIFS-CRPS, comparison to TIGGE ensembles



TIGGE: https://doi.org/10.1175/2010BAMS2853.1

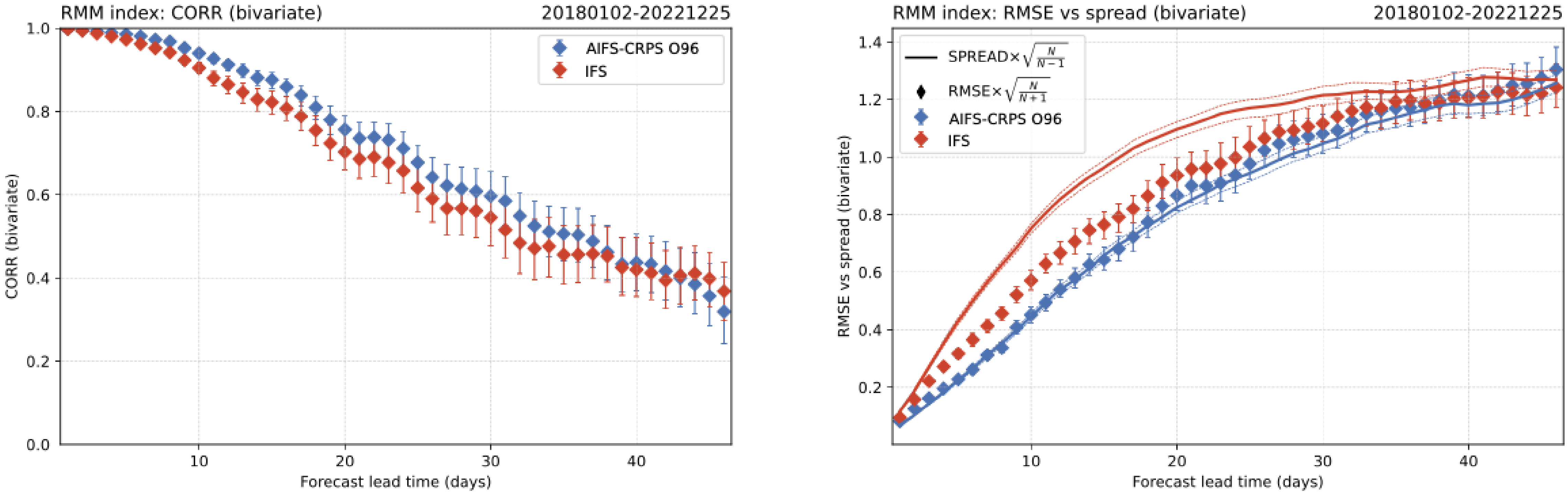
AIFS-CRPS sub-seasonal skill



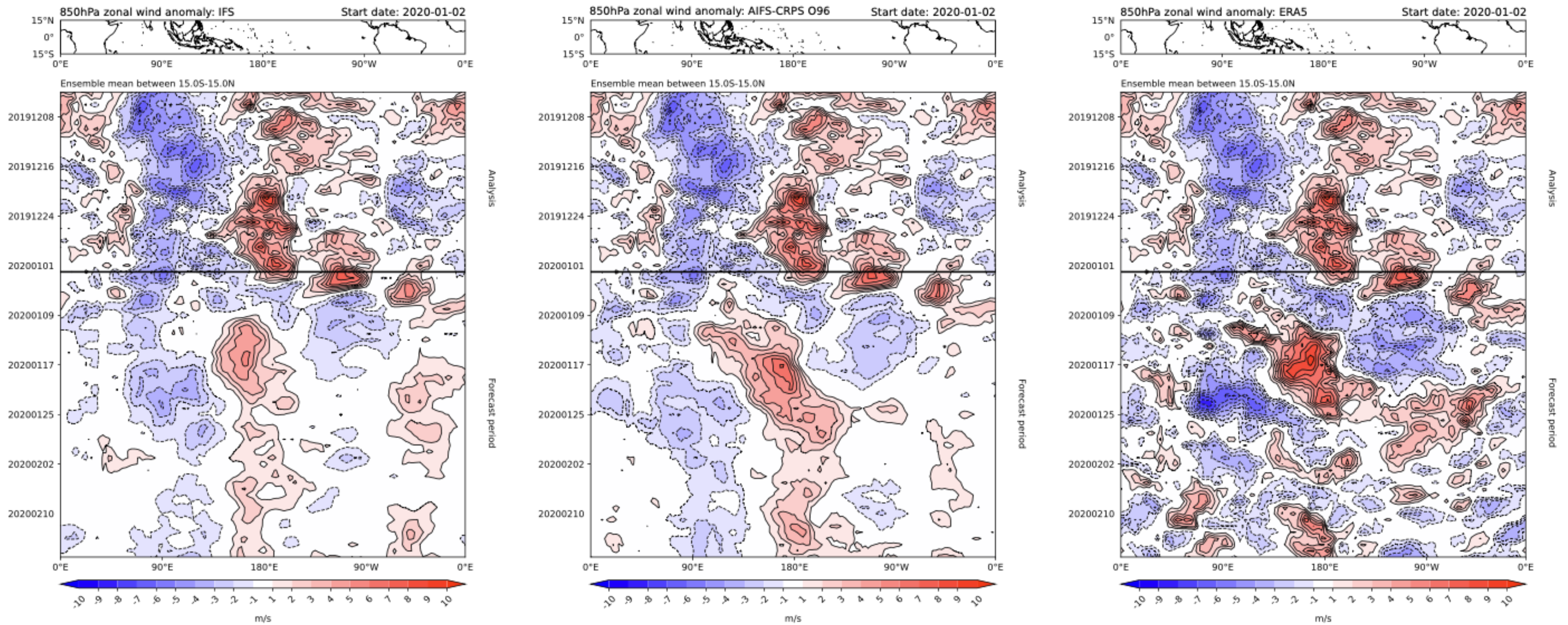


AIFS-CRPS sub-seasonal skill – Madden Julien Oscillation

IFS, AIFS-CRPS O96



(a) Bivariate correlations for an MJO index calculated from 200 hPa and 850 hPa zonal wind anomalies for AIFS-CRPS (blue) and operational IFS reforecasts run in 2023 (red). The MJO index used here is an approximation for the full Wheeler and Hendon [2004] Real-time Multivariate MJO index as it excludes contributions from outgoing longwave radiation that are not available from AIFS-CRPS. For both systems, correlations are calculated with respect to the same indices calculated from ERA5. Error bars represent the 2.5th and 97.5th percentiles of the distribution created by block-bootstrap resampling of the available start dates. (b) Estimates of root mean square error (RMSE; diamonds) and average ensemble spread (solid lines) for the MJO index described in the text. Spread and RMSE are scaled by factors of $sqrt\frac{N}{N-1}$ and $sqrt\frac{N}{N+1}$, respectively, to ensure estimates are unbiased with sample size (N) as described in Leutbecher and Palmer [2008].



IFS

(a)

AIFS-CRPS

(b)

ERA5

(c)

Hovmöller diagrams showing the evolution of zonal wind anomalies at 850 hPa meridionally averaged from 15°S-15°N. All panels show the evolution of zonal wind anomalies in ERA5 for the 30 days prior to the forecast start date (i.e. data above the grey line). Anomalies below the grey line are from (a) IFS ensemble mean forecast initialized on 2020-01-02, (b) AIFS-CRPS ensemble mean forecast initialized on 2020-01-02, and (c) ERA5.

See: https://arxiv.org/abs/2412.15832
https://arxiv.org/abs/2406.01465

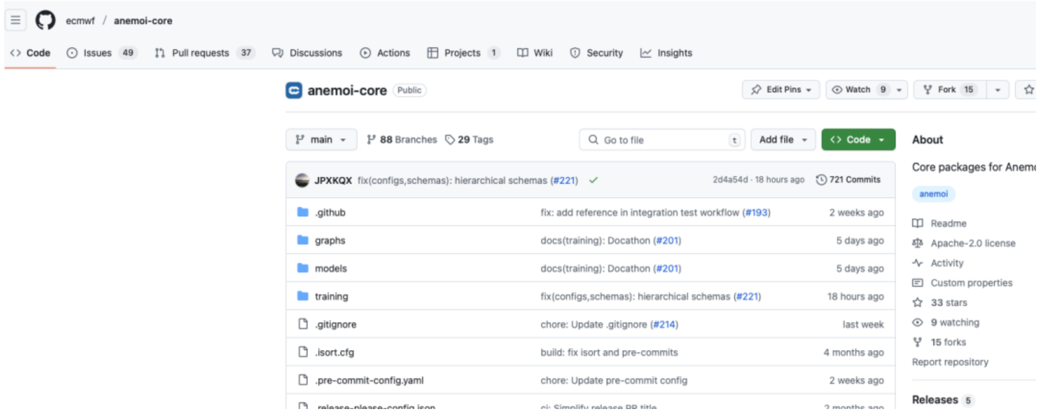
AIFS Single is now operational

➔ Forecasts available via charts.ecmwf.int and data via ECMWF's open data policy
➔ Code, training recipes and model weights open

Next:

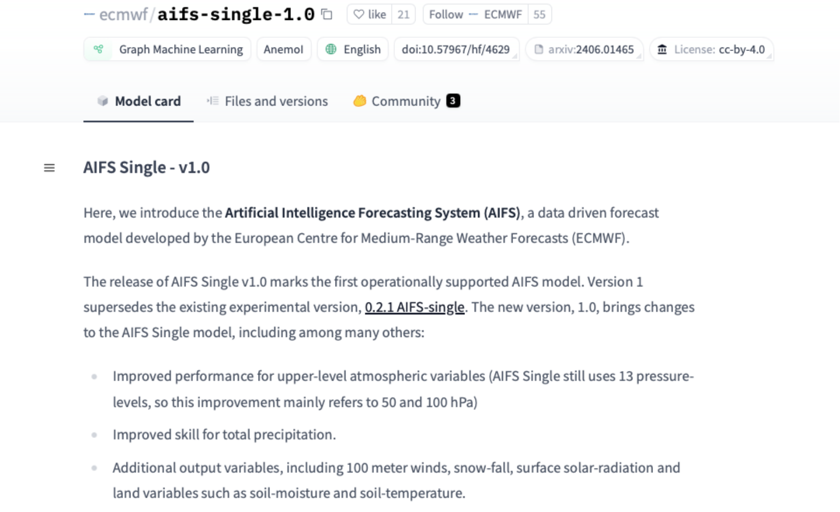
➔ AIFS-CRPS will enter experimental real-time mode soon
➔ Planned: publish weights, code and data
➔ Include more variables, increase resolution, etc.
➔ Bigger models?
➔ Training data …
➔ Guide NWP models …

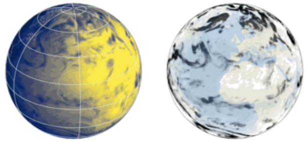
Code (including training etc.)



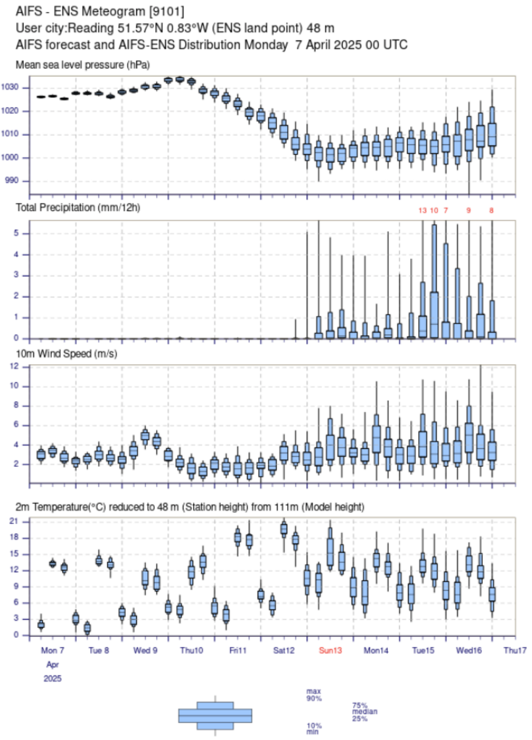
ECMWF

Model weights:



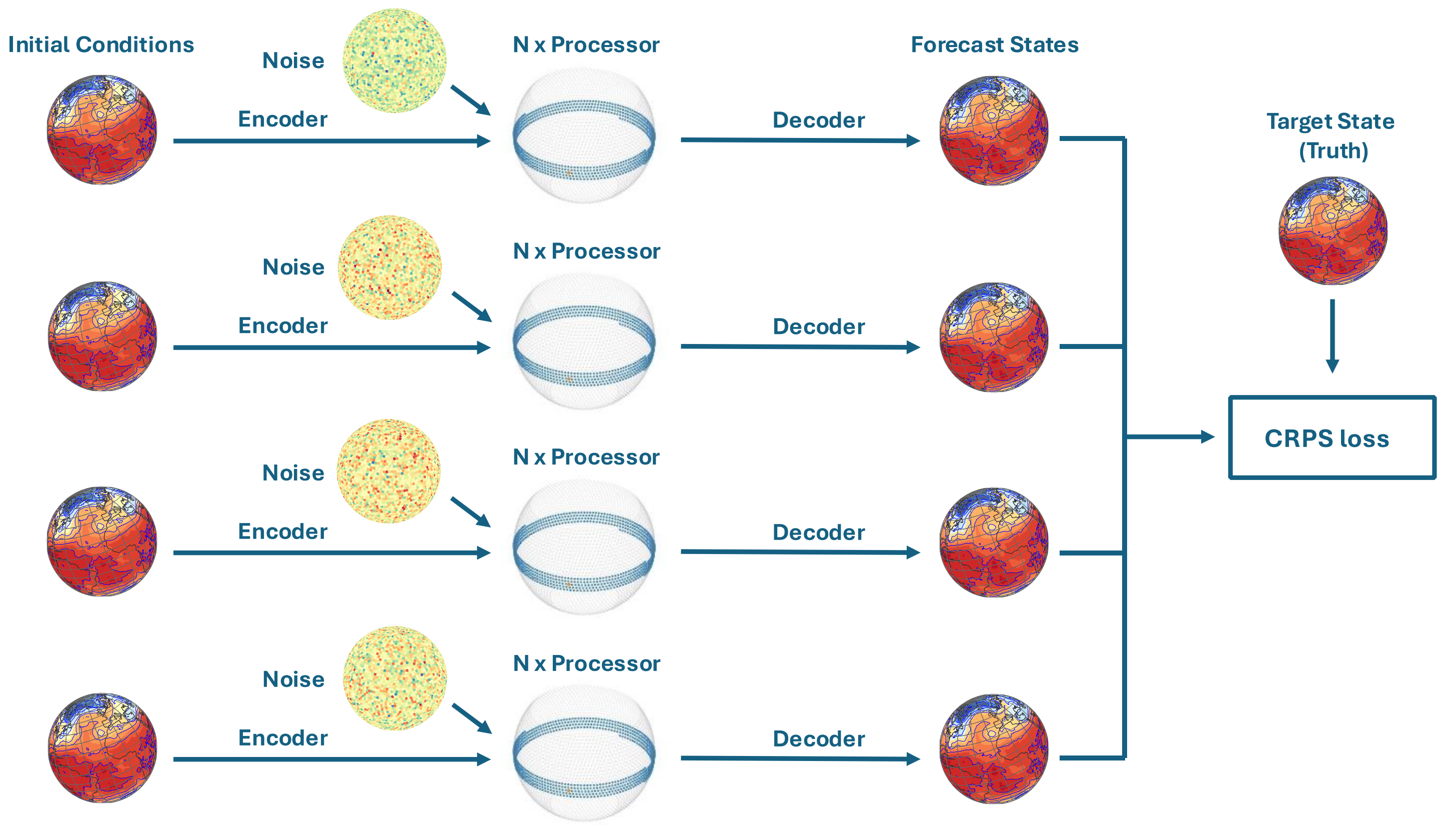


charts



Proper probabilistic score loss – AIFS-CRPS:

In training: run (small) ensemble:



To generate a forecast (single member): run model with noise realization for each forecast step

Mitigating error accumulation during rollout

$$x_{t+1} = U(D(x_t)) + f(x_t)$$

Reframe training objective:

- Predict an increment to a coarsened input field.
- Effectively predict an increment to the state for large scale.
- Predict the small scale directly on top of the smoothed field.

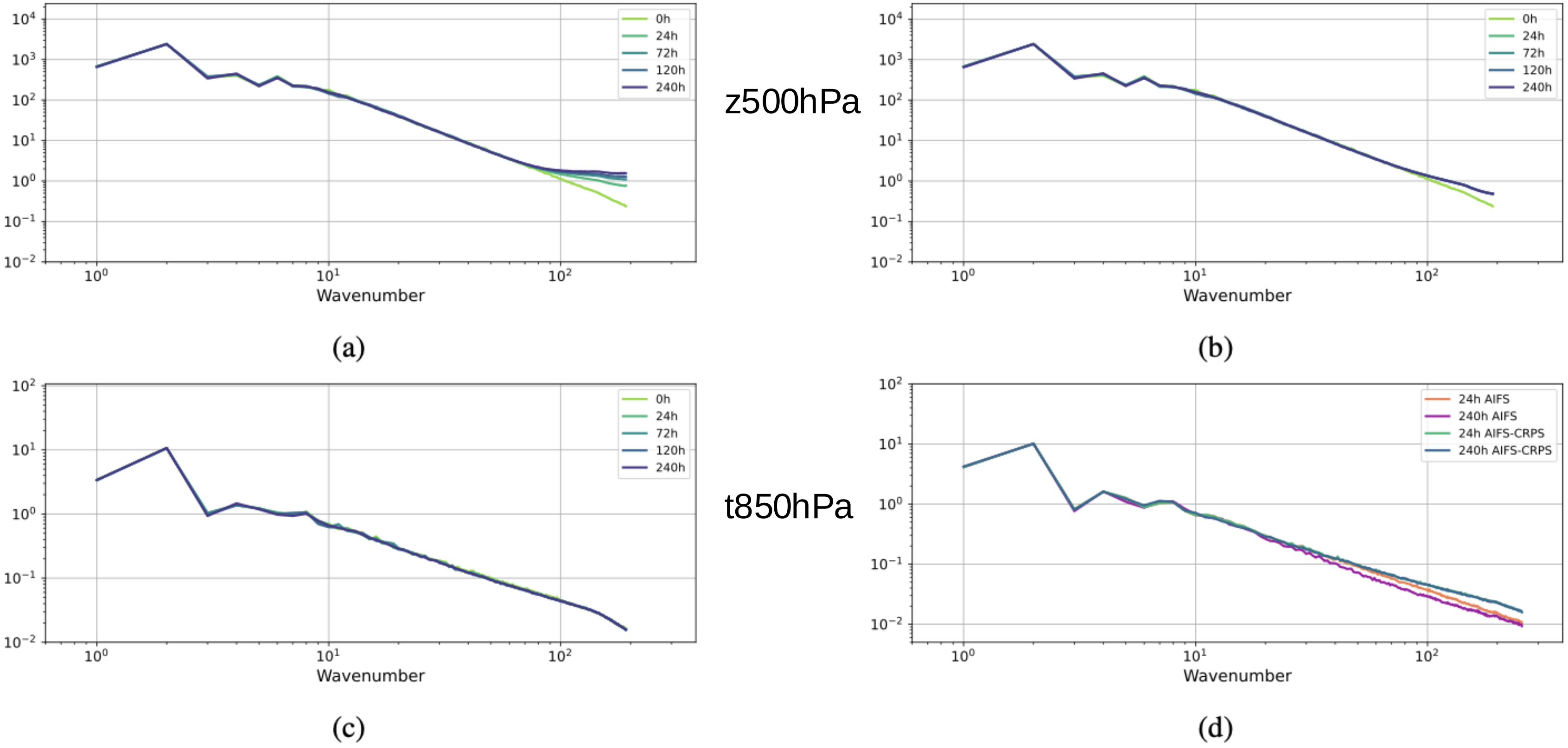
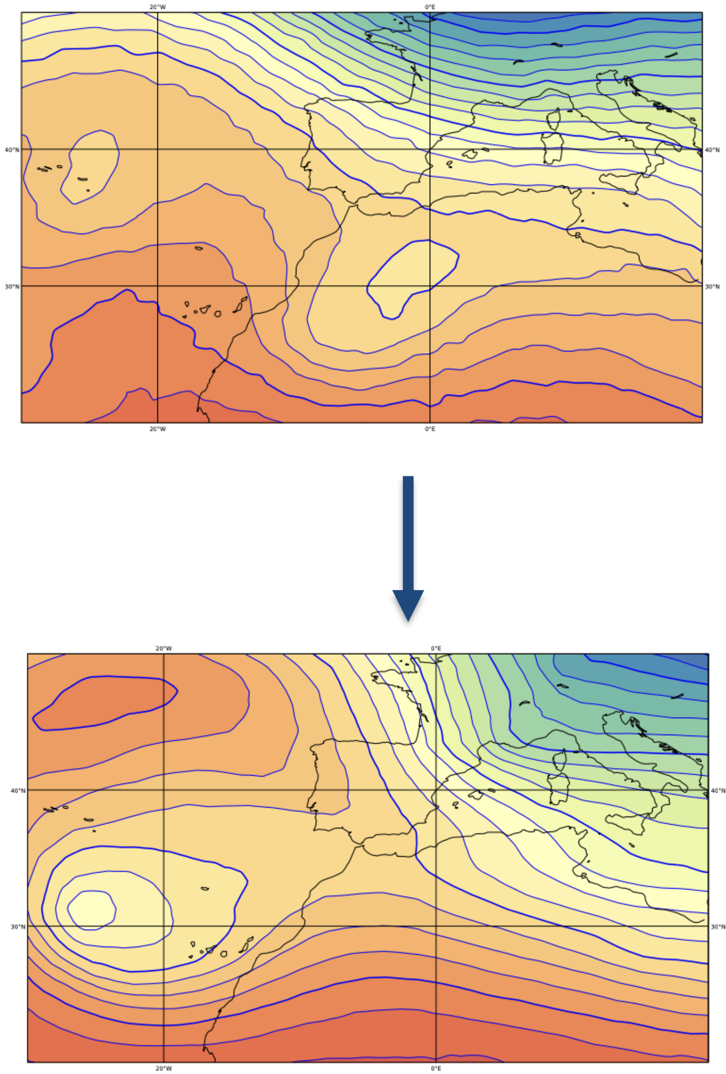


Figure 4: Spectra of geopotential at 500 hPa (a, b) and temperature at 850 hPa (c, d) for different lead times. Step 0 h refer to the initial conditions / IFS analysis. Shown are the AIFS-CRPS ensemble without (a) and with reference field truncation (b, c, d), and AIFS (d). Spectra are averaged over 12 initial dates and the first 8 ensemble members (a, b and c). For the AIFS and AIFS-CRPS comparison (d), the spectra are averaged over 12 initial dates and AIFS-CRPS perturbed member 1 only. For more explanation, please see the text.



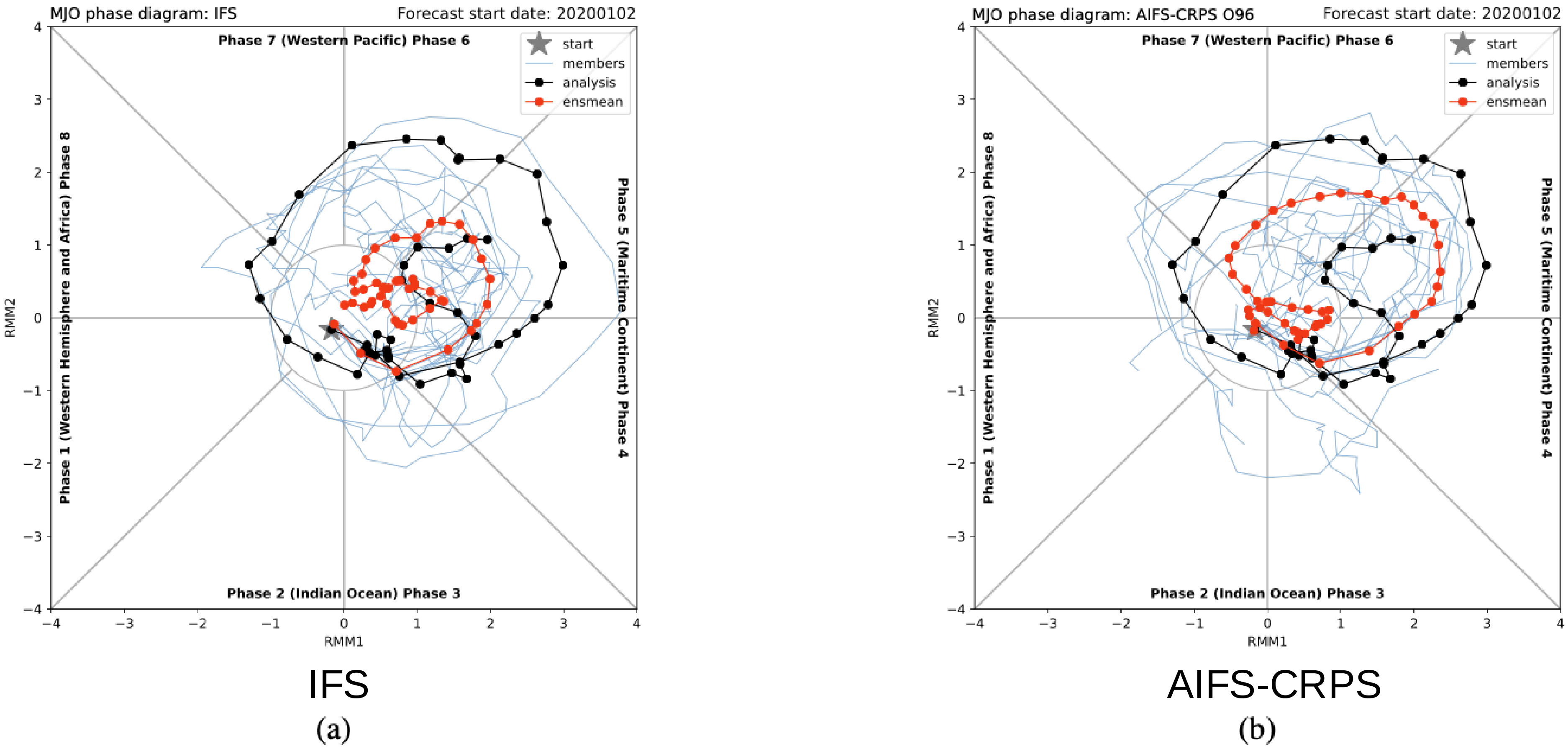
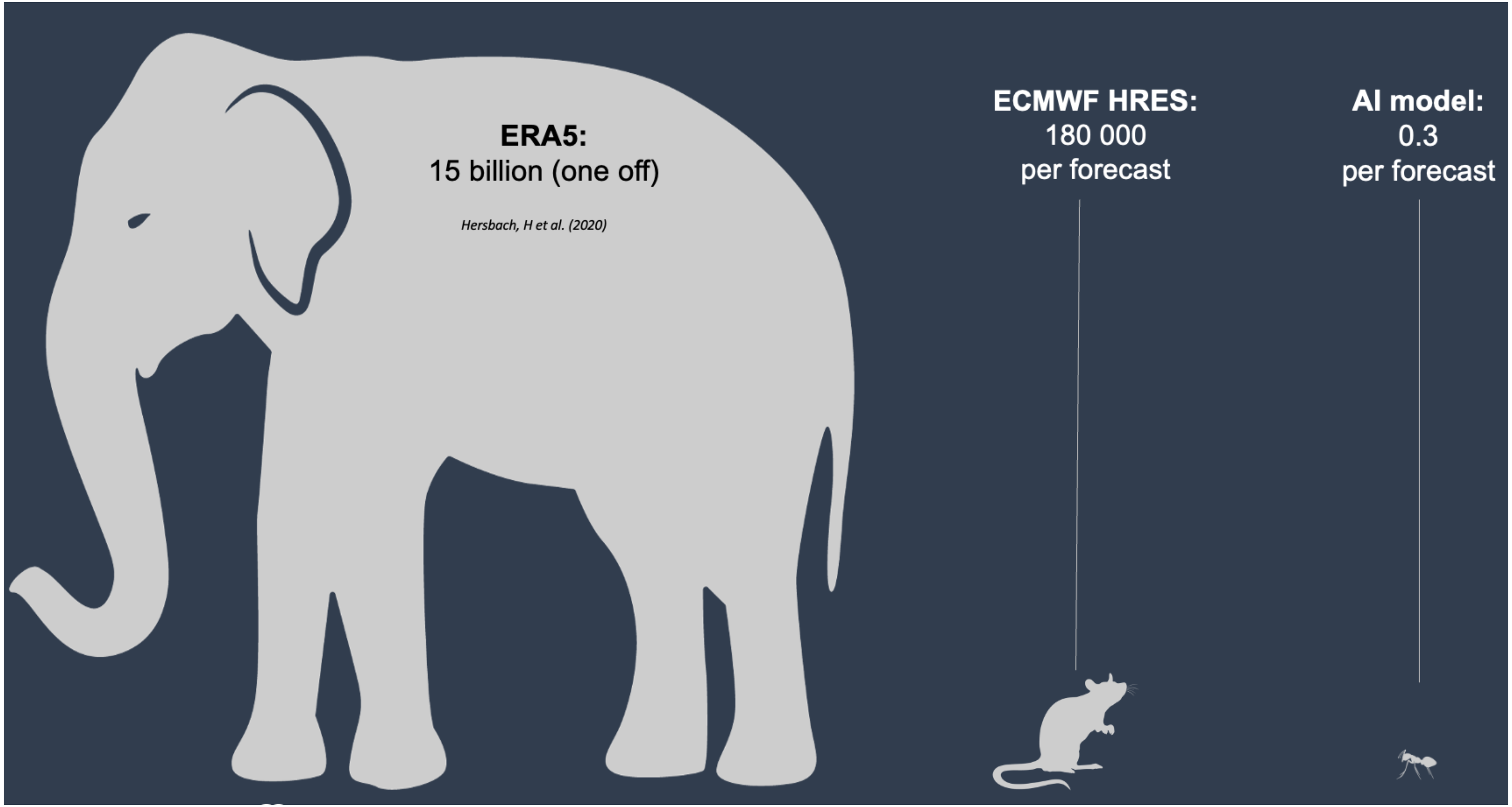


Figure 13: Phase diagrams based on the surrogate Real-time Multivariate MJO index described in the text for 46-day ensemble forecasts initialized on 2020-01-02 from (a) IFS and (b) AIFS-CRPS reforecasts.

Field / Variable	Level type	Input/Output	Normalisation
Geopotential, horizontal and vertical wind components, specific humidity, temperature	Pressure level: 50, 100, 150, 200, 250, 300, 400, 500, 600, 700, 850, 925, 1000	Both	Standardised, apart from the geopotential, which is max-scaled
Surface pressure, mean sea-level pressure, skin temperature, 2 m temperature, 2 m dewpoint temperature, 10 m horizontal wind components, total column water	Surface	Both	Standardised
Total precipitation	Surface	Output	Standard deviation changed but mean kept the same
Land-sea mask, orography, standard deviation and slope of sub-grid orography, insolation, cosine and sine of latitude and longitude, cosine and sine of the local time of day and day of year	Surface	Input	Sub-grid orography max-scaled; all others not normalised
Time step identifier	-	Input	-

Table 2: Input and output variables of AIFS-CRPS.

How costly?


ERA5:
15 billion (one off)
Hersbach, H et al. (2020)
ECMWF HRES:
180 000
per forecast
AI model:
0.3
per forecast