

AIFS

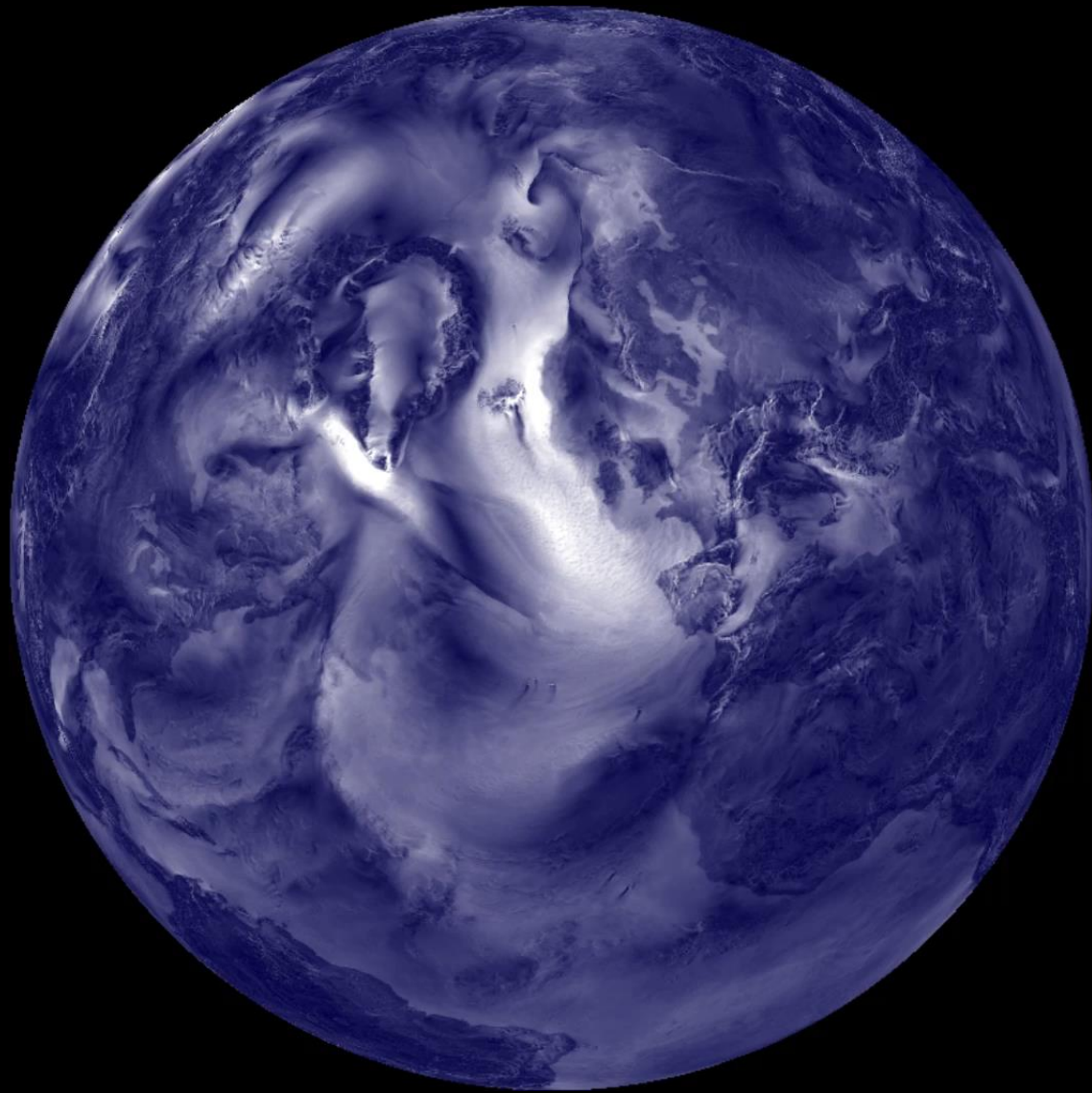
ECMWF's data driven forecast model

Simon Lang and colleagues

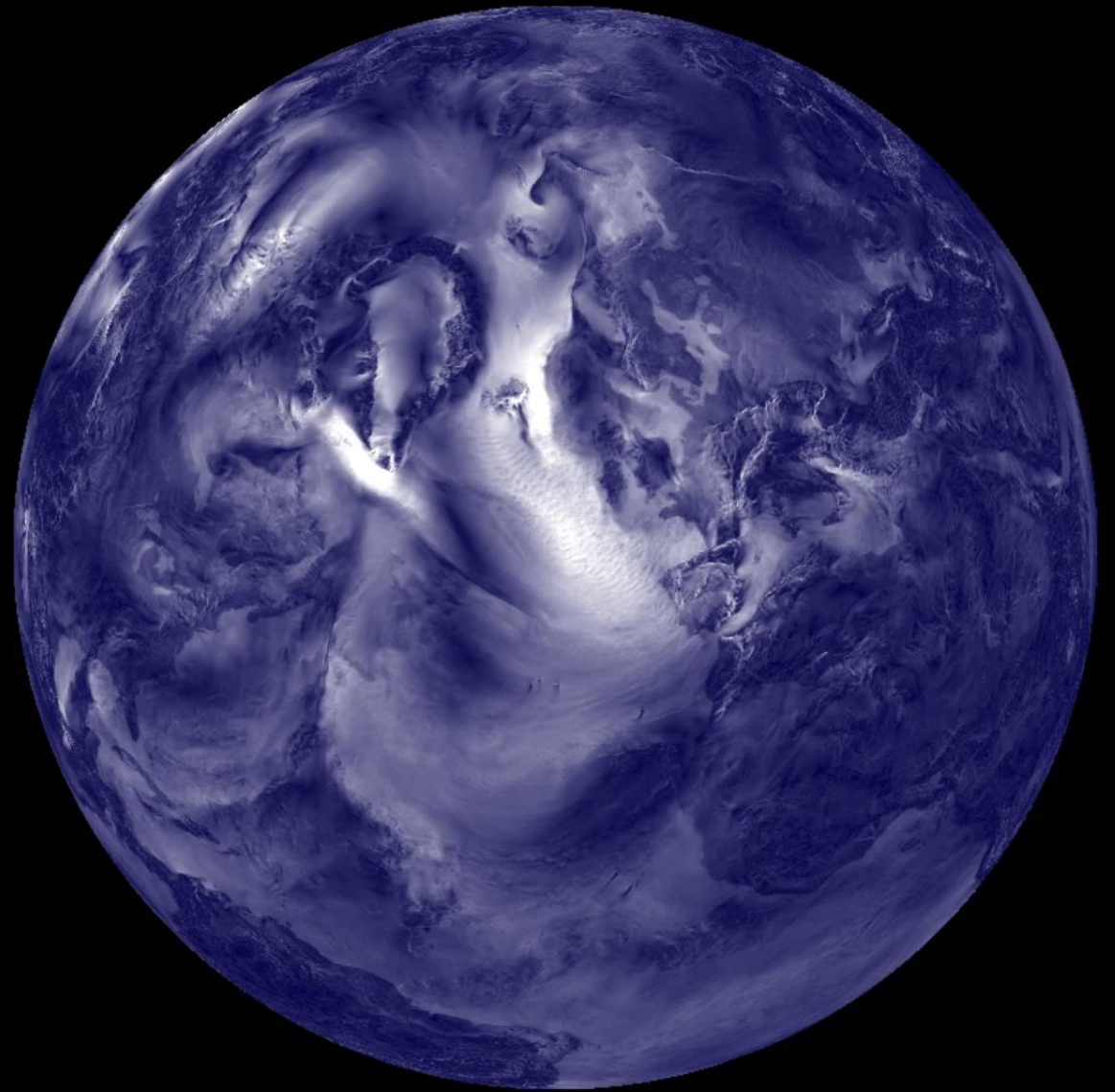
simon.lang@ecmwf.int

AIFS team: Rilwan Adewoyin, Mihai Alexe, Zied Ben Bouallègue, Matthew Chantry, Mariana Clare, Jesper Dramsch, Sara Hahner, Simon Lang, Christian Lessig, Linus Magnusson, Michael Maier-Gerber, Gert Mertes, Gabriel Moldovan, Ana Prieto Nemesio, Cathal O'Brien, Florian Pinault, Baudouin Raoult, Mario Santa Cruz, Helen Theissen, Steffen Tietsche

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



Control Member

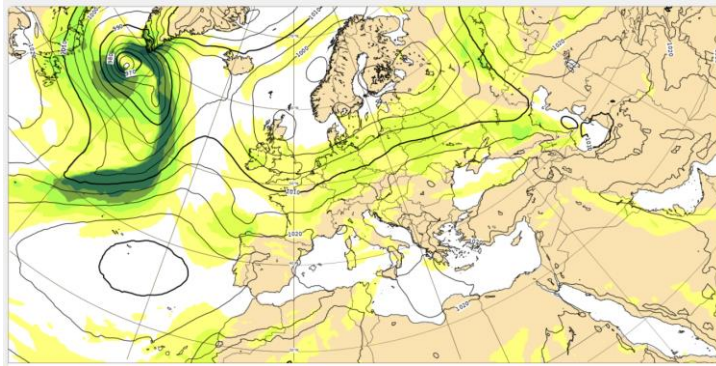


Perturbed member 1

Weather Forecasts – NWP? Data Driven?

Traditionally weather forecasts are generated by running 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 an NWP model?

Analysis

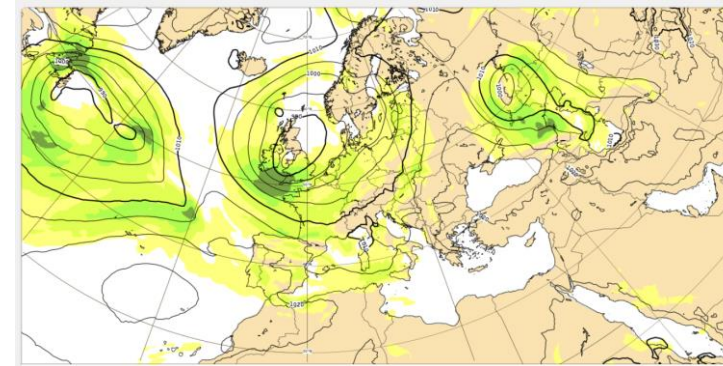


Fusion of short-range forecast with latest observations

NWP Model



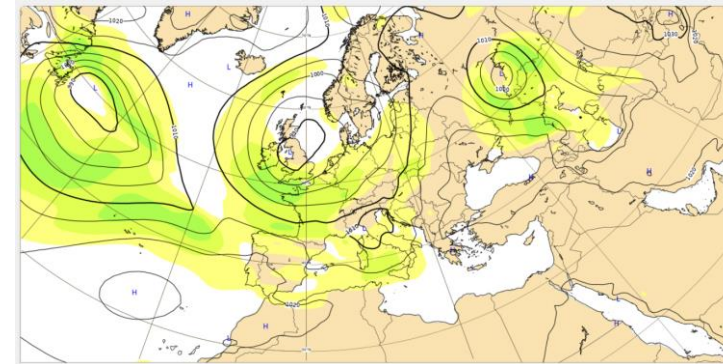
Forecast



Data Driven Model



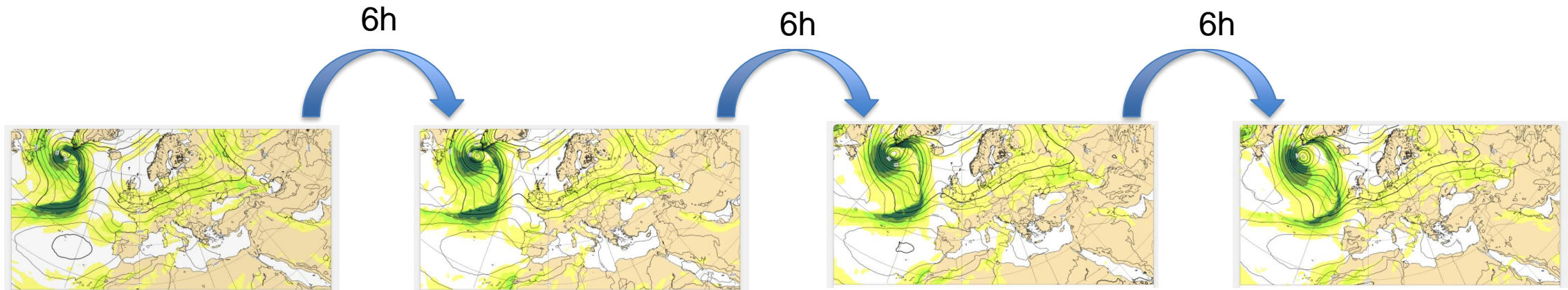
Learned from 40 years of analyses



Weather Forecasts – NWP? Data Driven?

Recent advances by individuals and tech companies show that this is possible (e.g. NVIDIA, Keisler, Google Deepmind, Huawei, ... and others)

Here, the models learn from ca. 40 years of ECMWF's ERA5 re-analysis data, stepping e.g. 6h from analysis to analysis



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

If we want to build a model, we have to choose an architecture ...

- Vision transformers
- Graph neural networks
- Convolutions
-



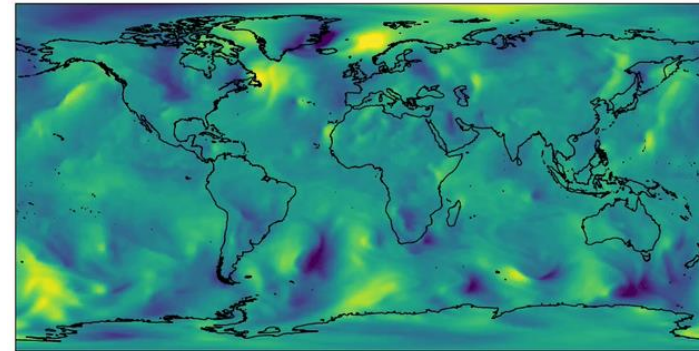
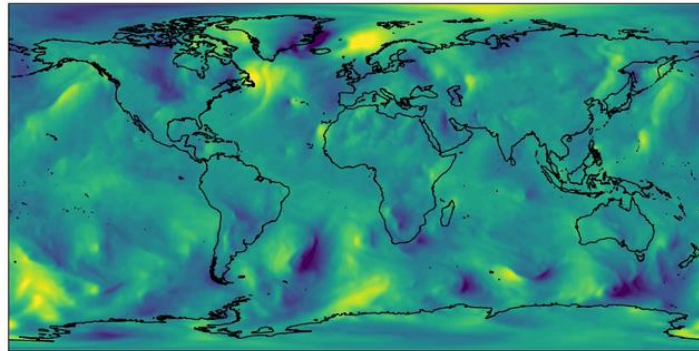
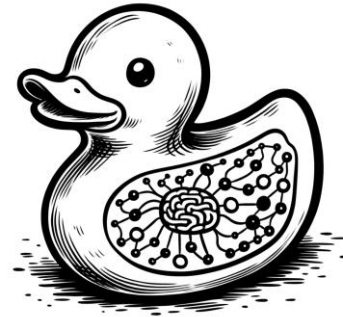
AIFS - Artificial Intelligence Forecasting System

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
- Scales to > 1000s of GPUs ; tensor parallel implementation, split model across multiple GPUs

Update beginning of 2024, update to ~ 0.25 deg:

- Attention based GNN for encoder, decoder
- Transformer backbone in processor
- Trained on 64 GPUs ~ 1 Week



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

news

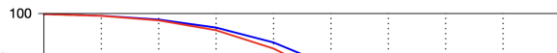
AIFS: a new ECMWF forecasting system

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

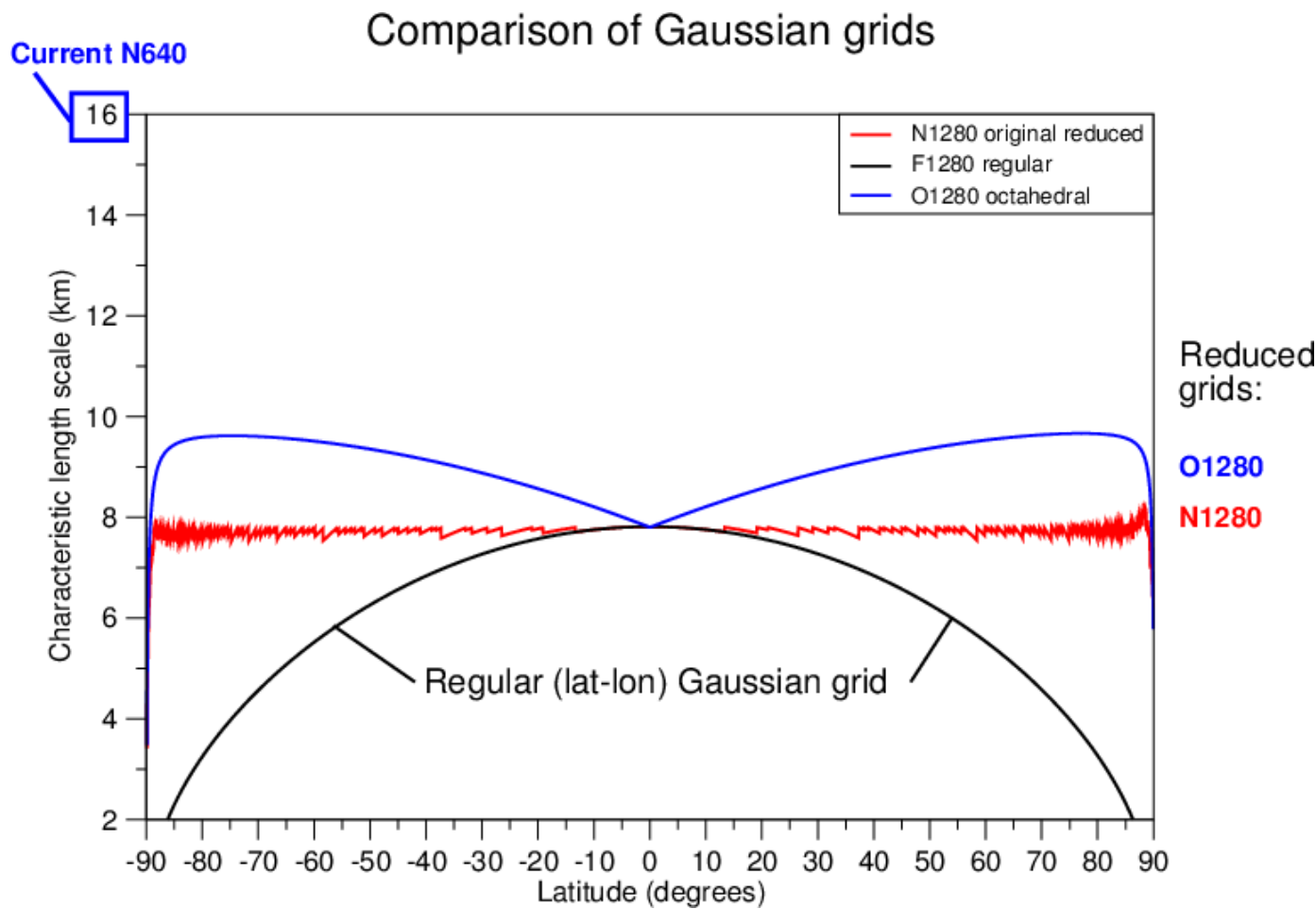
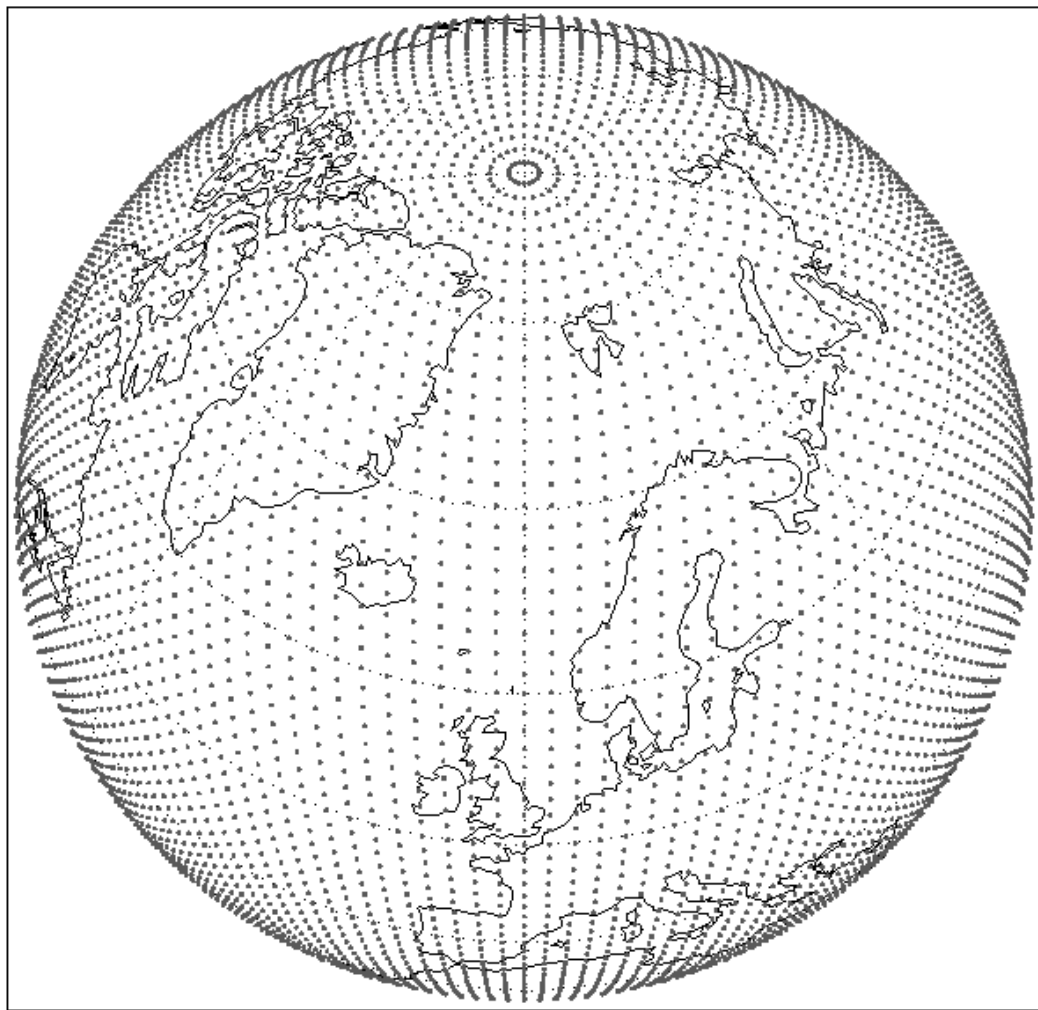
Lang S., Alexe M., Chantry M., Dramsch J., Pinault F., Raoult B., Ben Bouallègue Z., Clare M, Lessig C., Magnusson L., Prieto Nemesio A., 2024. **AIFS: a new ECMWF forecasting system.** ECMWF Newsletter number 178. p4-5. <http://dx.doi.org/10.21957/1a8466ec2f>

UTRE FOR MEDIUM-RANGE WEATHER FORECASTS

There has been substantial progress recently in the realm of data-driven weather forecasting. Disturbance

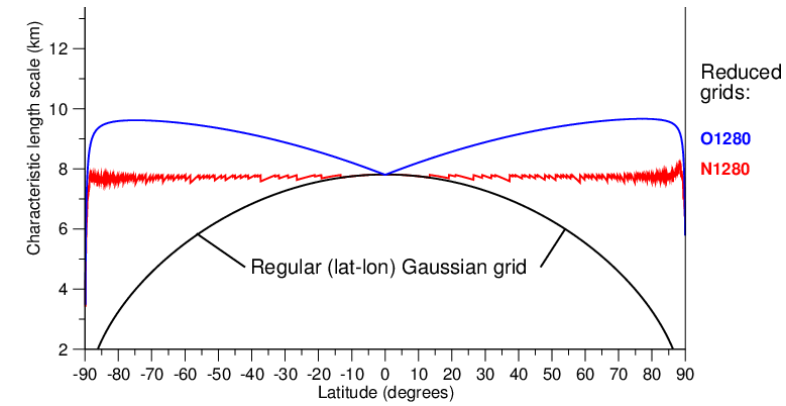


ECMWF's octahedral reduced Gaussian grid

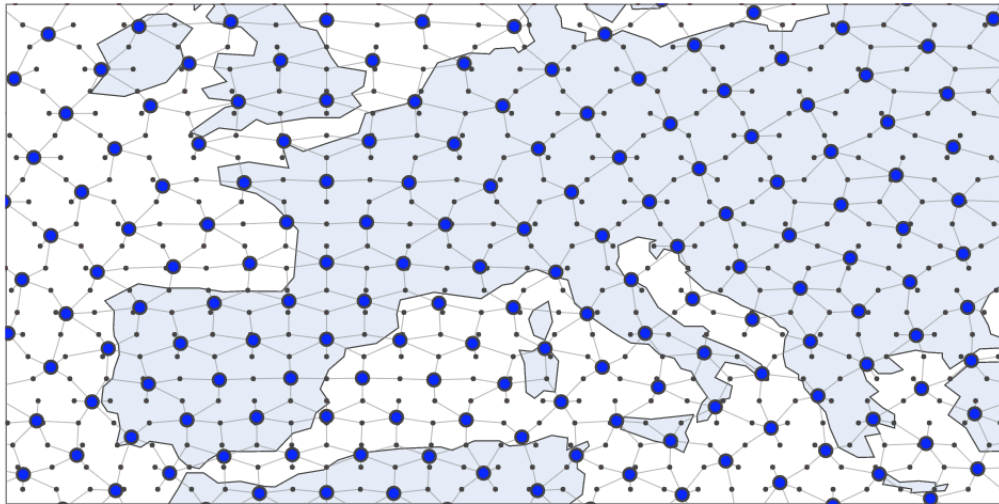


AIFS – Encoder and Decoder

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



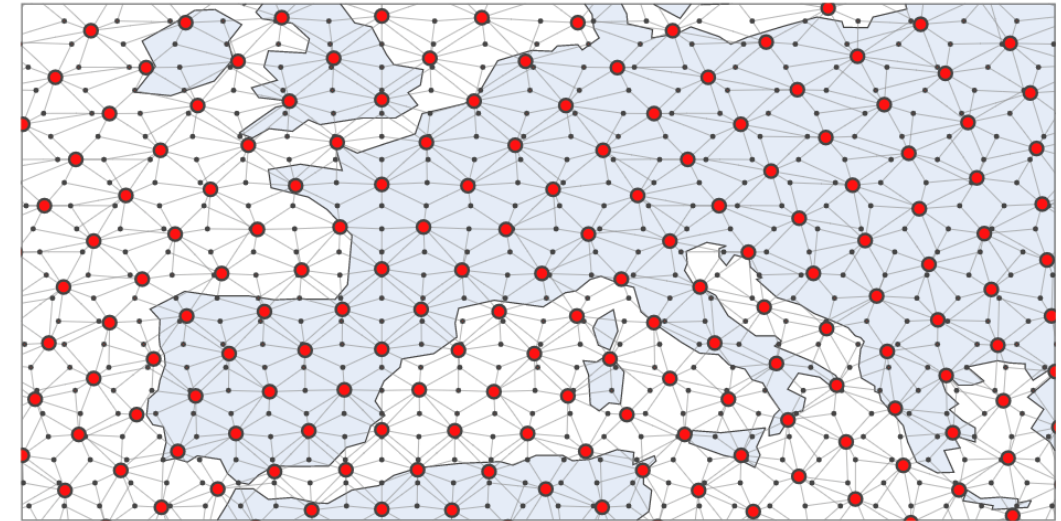
Encoder, GNN



Era5 – n320

~ 540 000 Nodes,
~ 1 Million Edges

Decoder, GNN



Era5 – n320

16 x Processor
with skip-connections

O96 ~ 40 000 Nodes

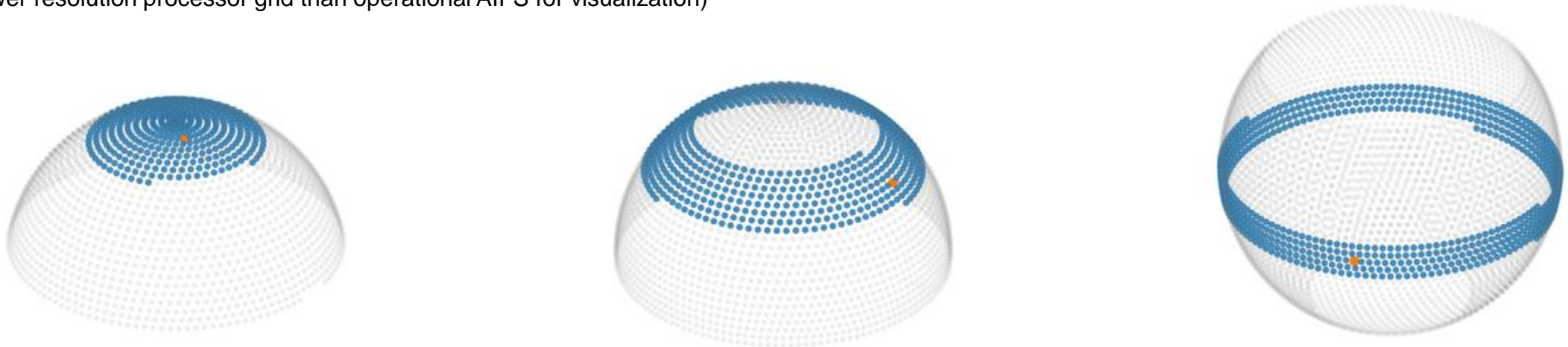
AIFS – Processor

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)

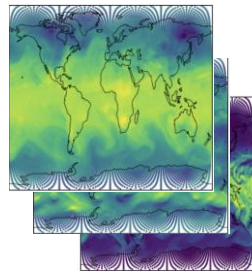
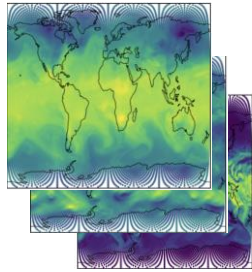


since February 2024, See AIFS blog:

Lang et al. 2024: <https://www.ecmwf.int/en/about/media-centre/aifs-blog/2024/first-update-aifs>

Chantry et al. 2024: <https://www.ecmwf.int/en/about/media-centre/aifs-blog/2024/its-raining-data>

Atmospheric state:
 $X(t), X(t-6h)$



previous
 $X(t)$

Prediction:
 $X(t+6h)$

AR predictions

$AIFS_{t+6h \rightarrow t+12h}$

$WMSE_{t+12h}$

...

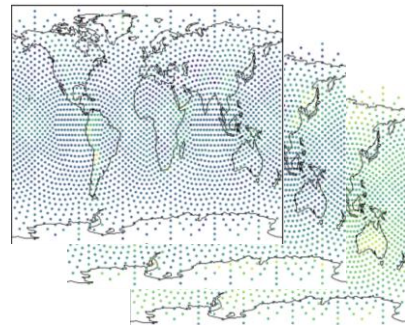
$WMSE_{t+6h}$

encoder

decoder

processor:
16 MP rounds

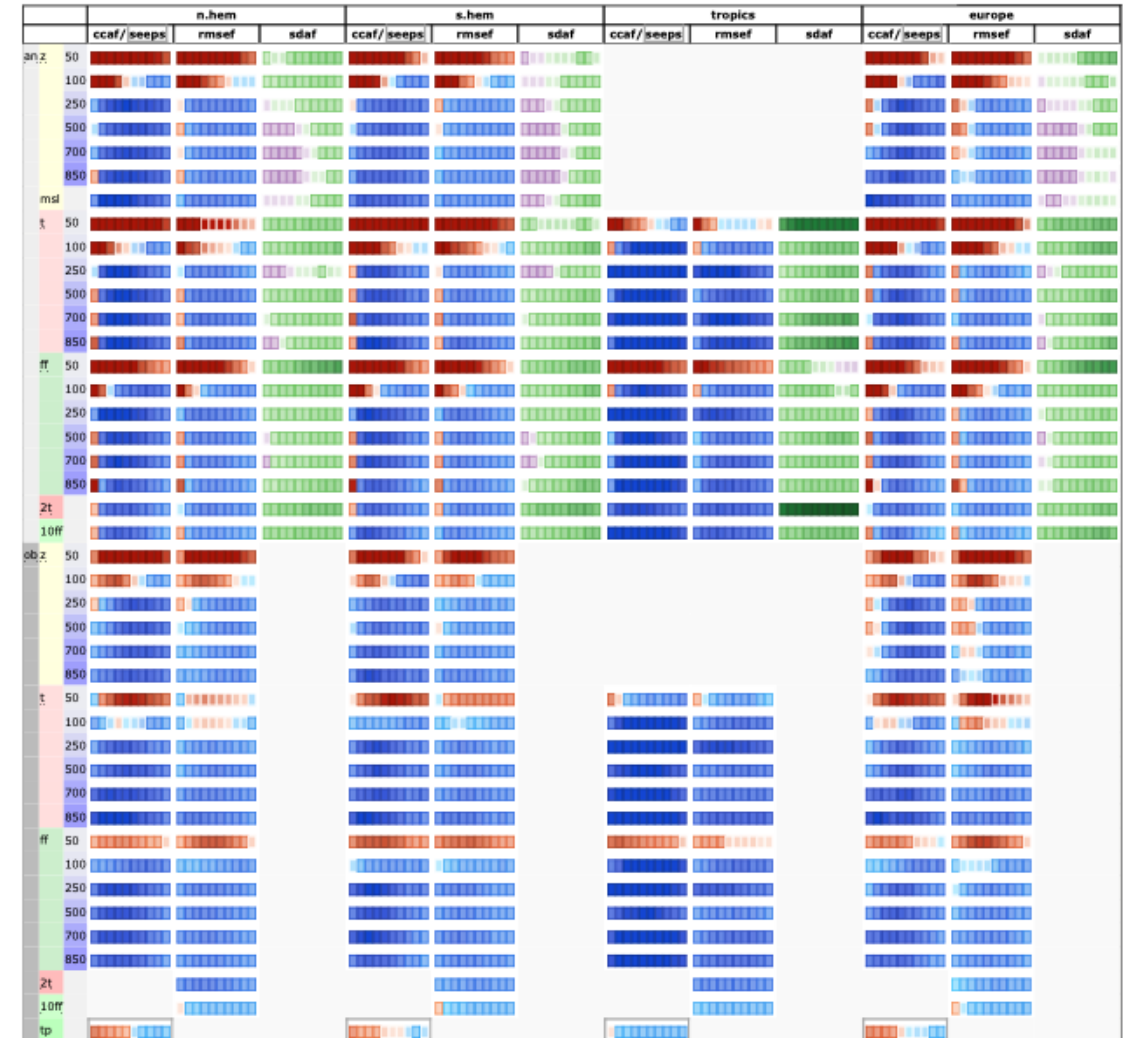
Aggregate
WMSE



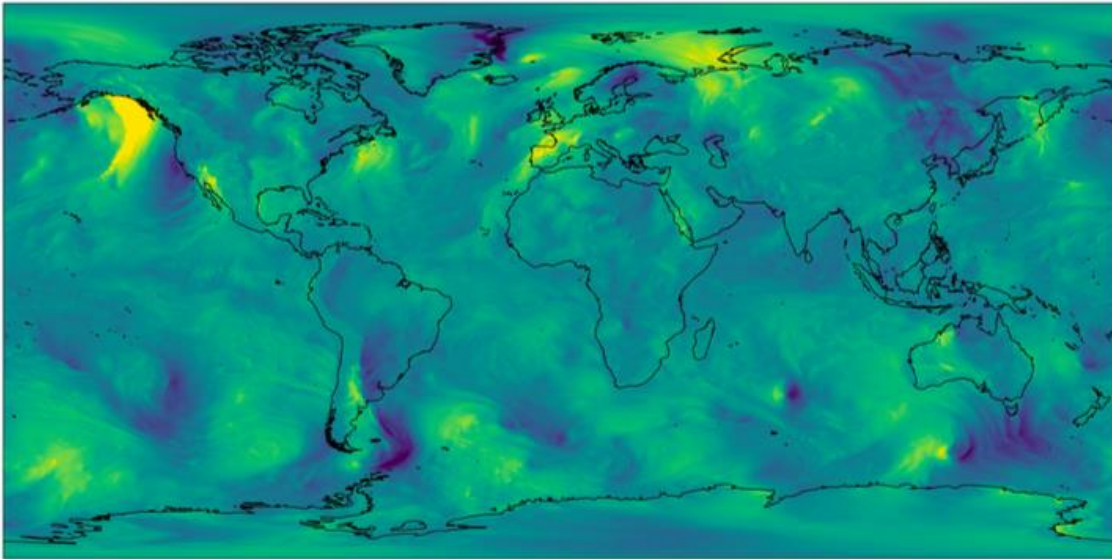
$AIFS_{t \rightarrow t+6h}$

Scorecard -> compared to IFS (2022)

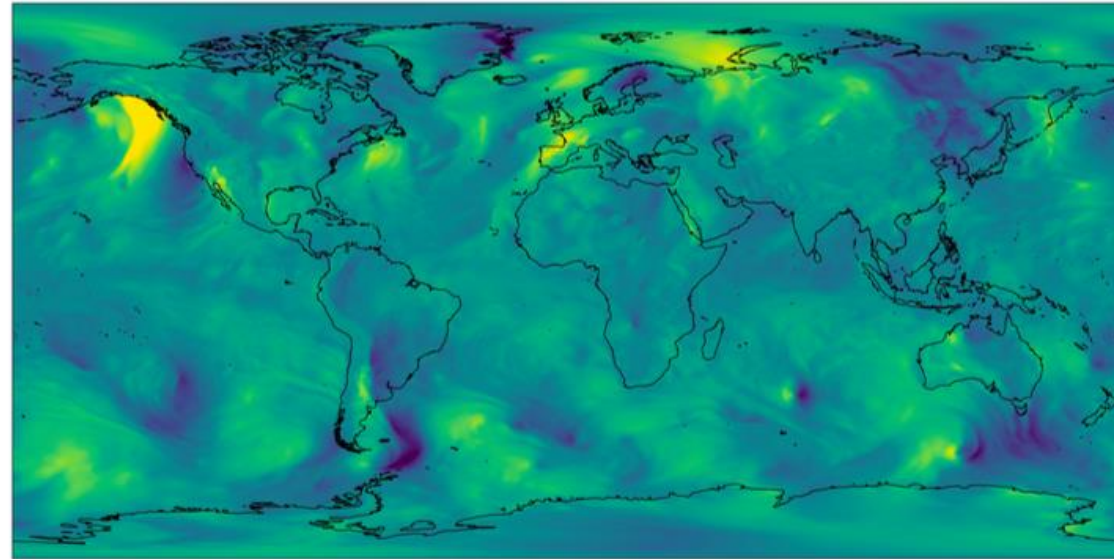
Scorecard comparing forecast scores of AIFS versus IFS (2022). Forecasts are initialised on 00 and 12 UTC. Shown are relative score changes as function of lead time (day 1 to 10) for northern extra-tropics (n.hem), southern extra-tropics (s.hem), tropics and Europe. Blue colours mark score improvements and red colours score degradations. Purple colours indicate an increased in standard deviation of forecast anomaly, while green colours indicate a reduction. Framed rectangles indicate 95% significance level. Variables are geopotential (z), temperature (t), wind speed (ff), mean sea level pressure (msl), 2 m temperature (2t), 10 m wind speed (10ff) and 24 hr total precipitation (tp). Numbers behind variable abbreviations indicate variables on pressure levels (e.g., 500 hPa), and suffix indicates verification against IFS NWP analyses (an) or radiosonde and SYNOP observations (ob). Scores shown are anomaly correlation (ccaf), SEEPS (seeps, for precipitaion), RMSE (rmsef) and standard deviation of forecast anomaly (sdaf, see text for more explanation).



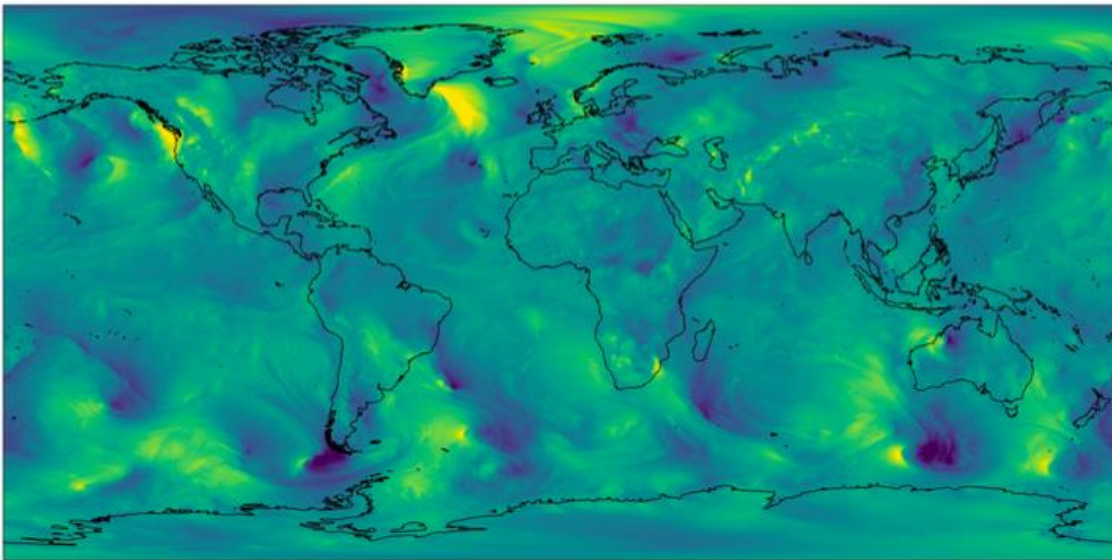
IFS +12h



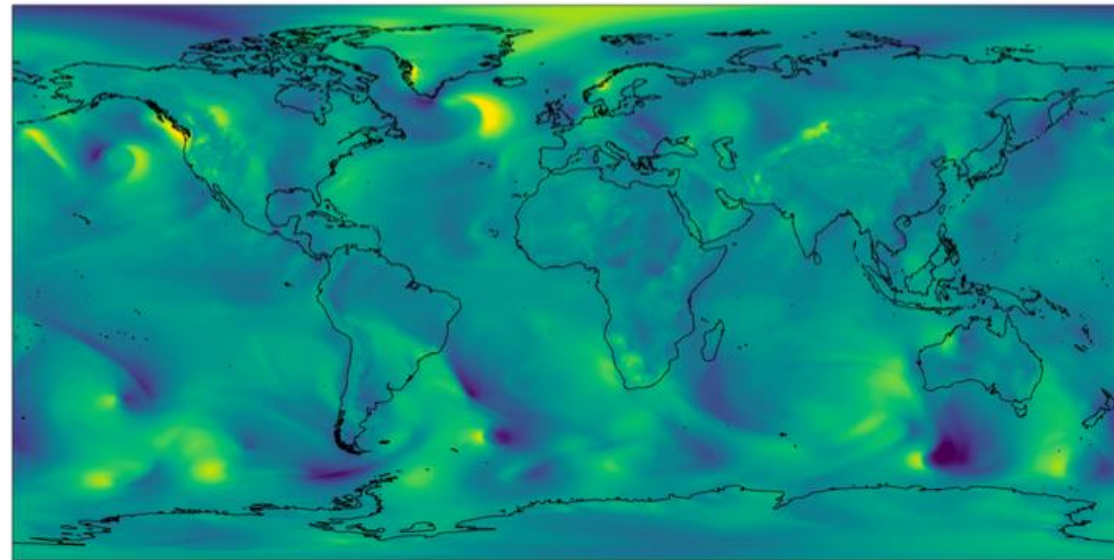
AIFS +12h



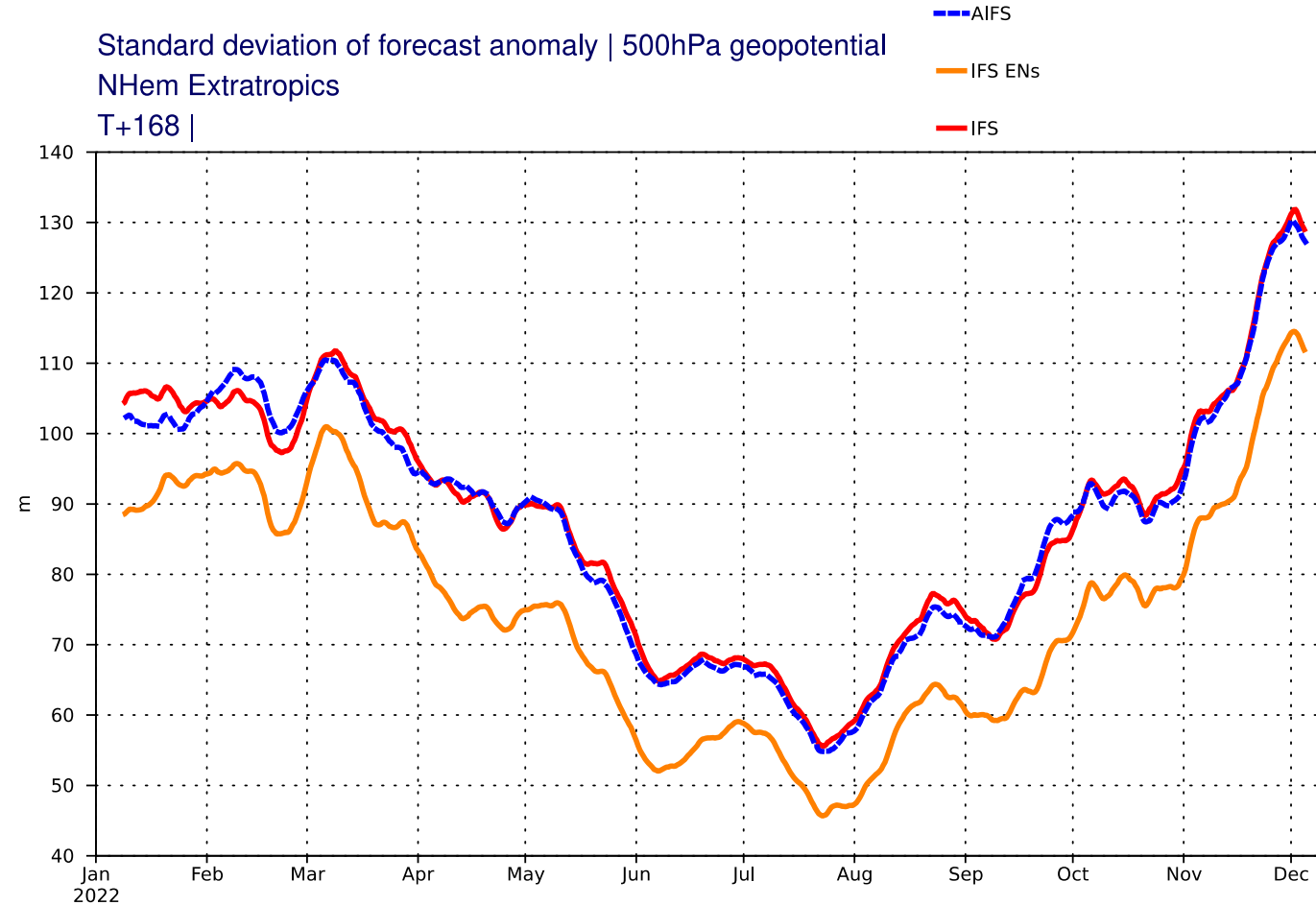
IFS +120h



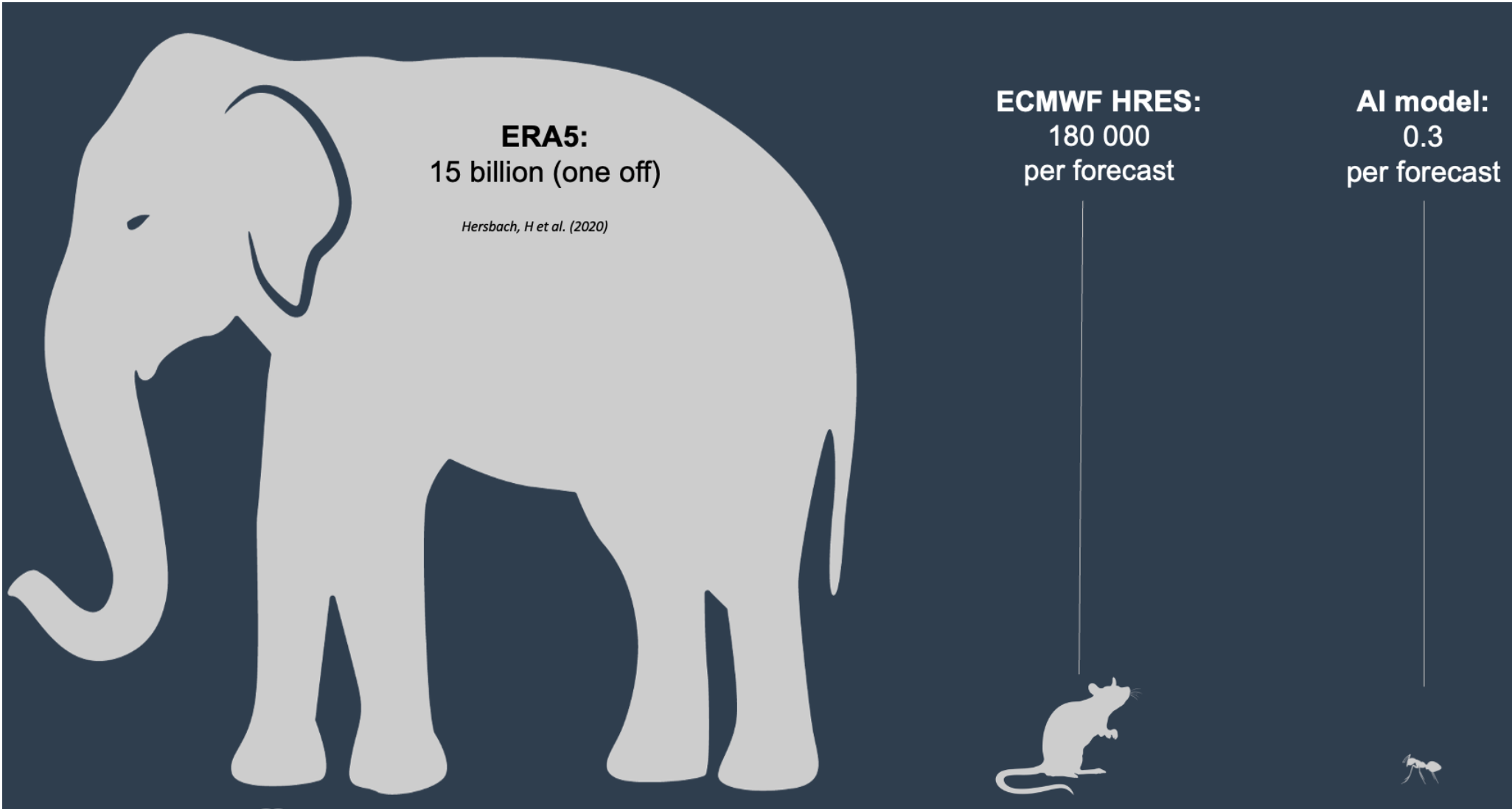
AIFS +120h



z500 forecast activity (2022)

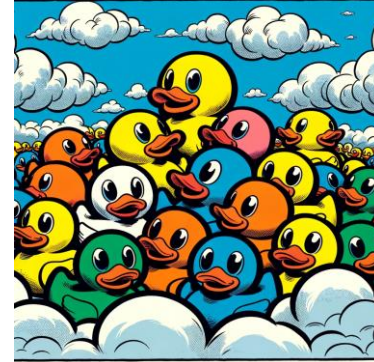
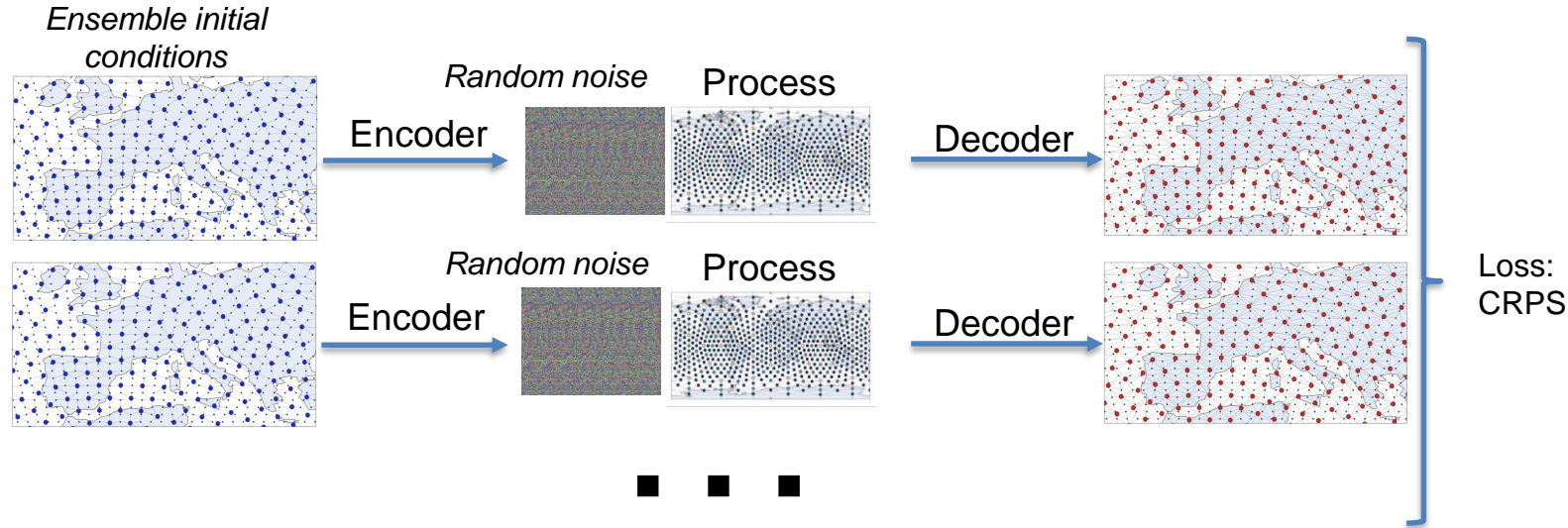


How costly?



Next? Ensemble forecasts ...

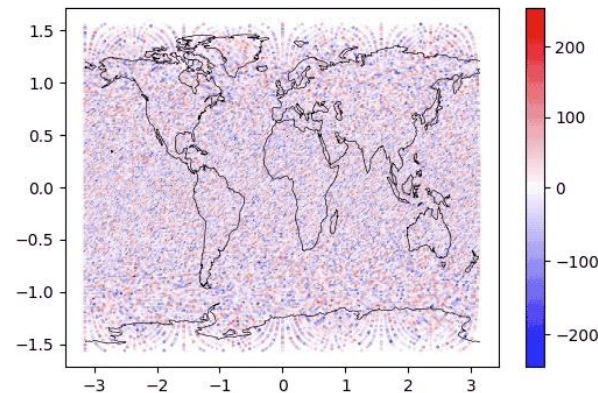
Instead of a MSE loss, learn an ensemble via optimizing probabilistic scores



“Generative AI”, e.g. create a forecast as de-noising task (diffusion probabilistic models)

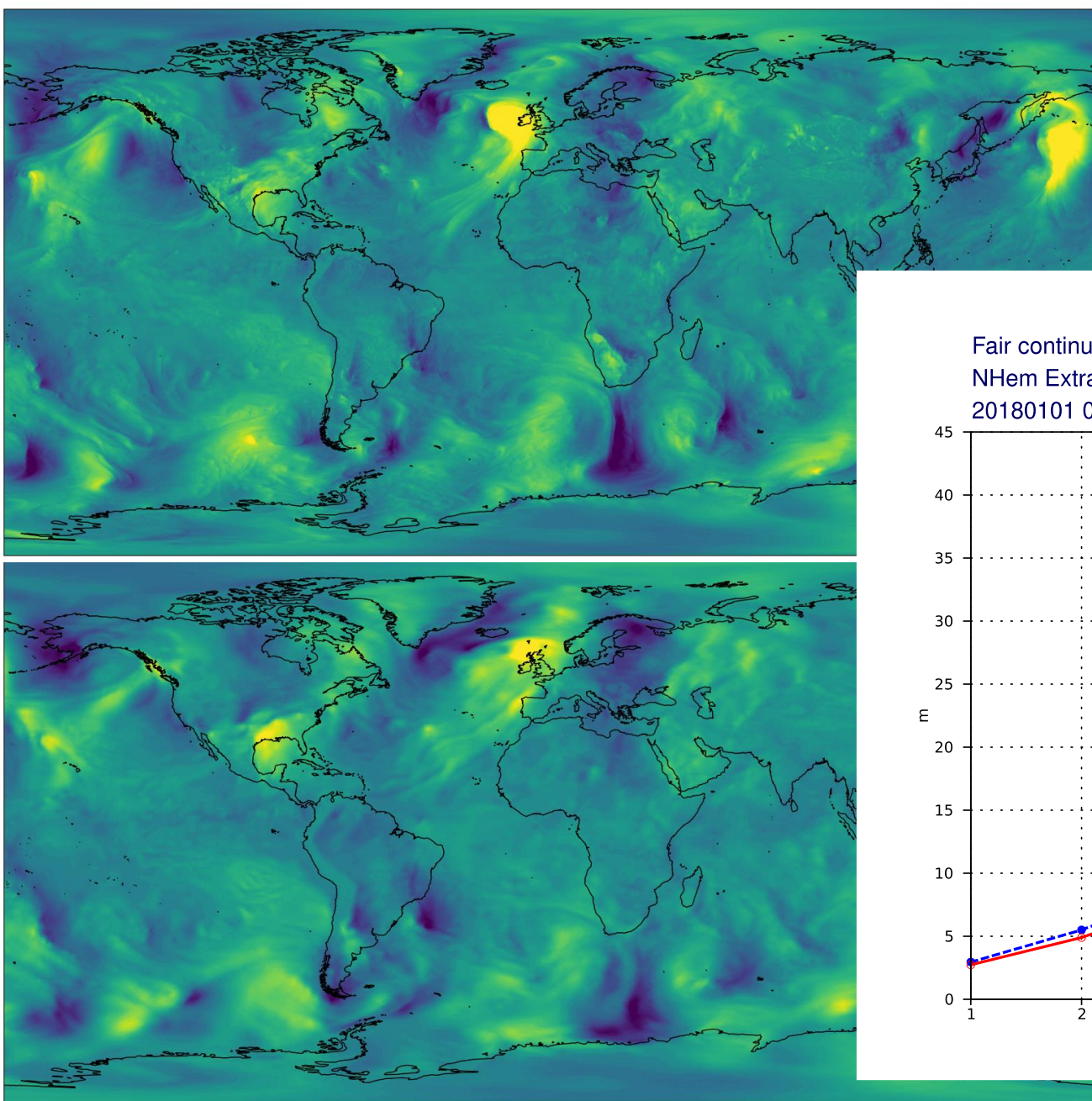
for example:

- Stable diffusion -> Images
- Sora -> Video
- Gencast -> Weather

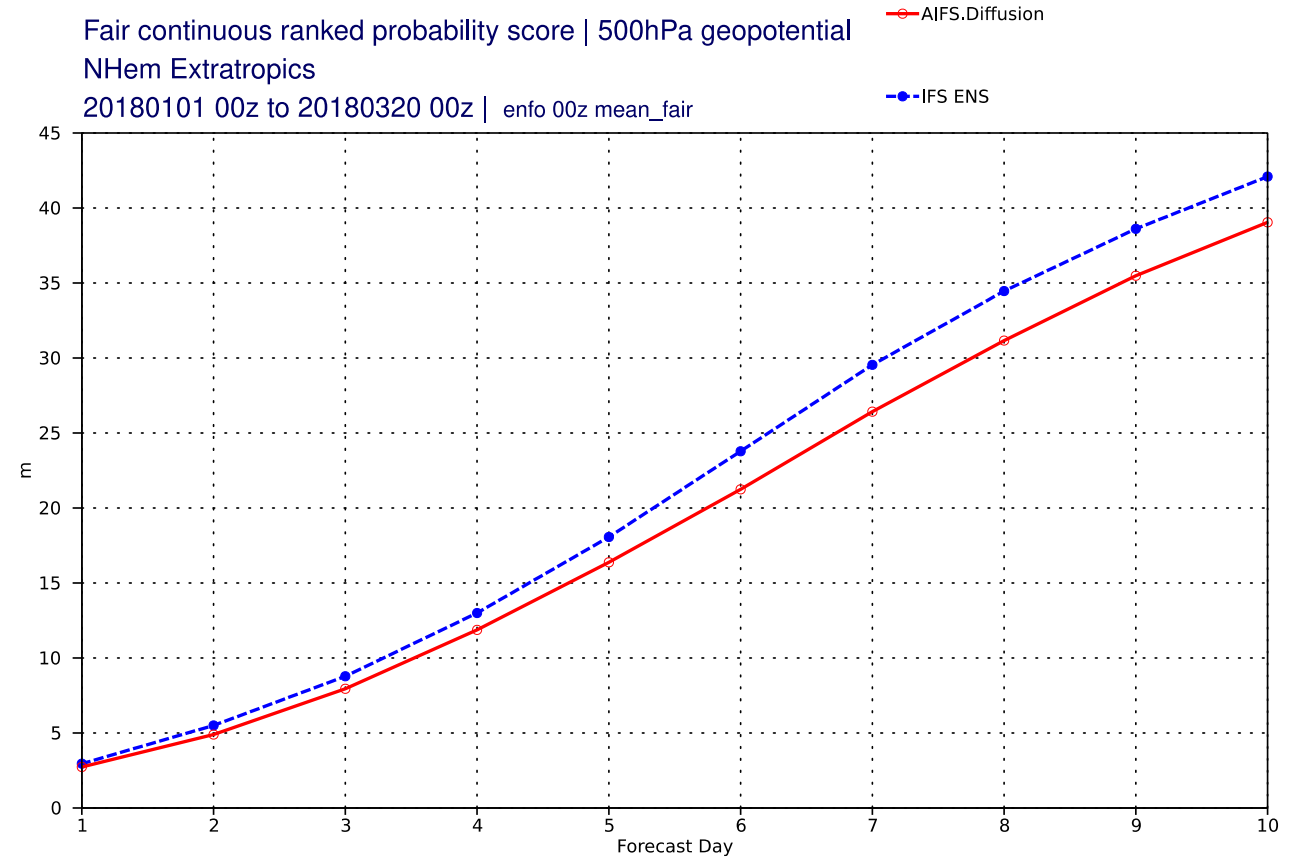


AIFS

~ 0.25 degree resolution



Fair continuous ranked probability score | 500hPa geopotential
NHem Extratropics
20180101 00z to 20180320 00z | enfo 00z mean_fair



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**

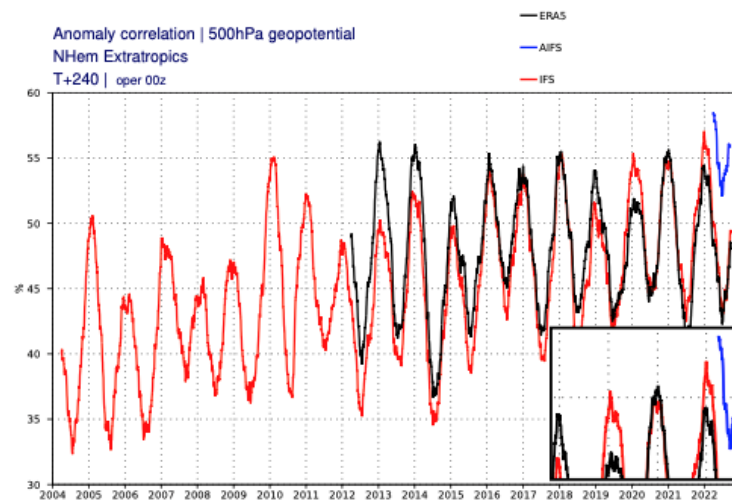
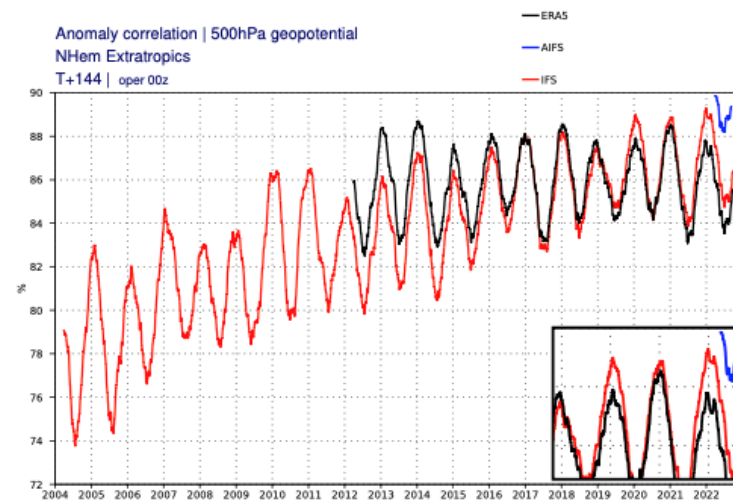
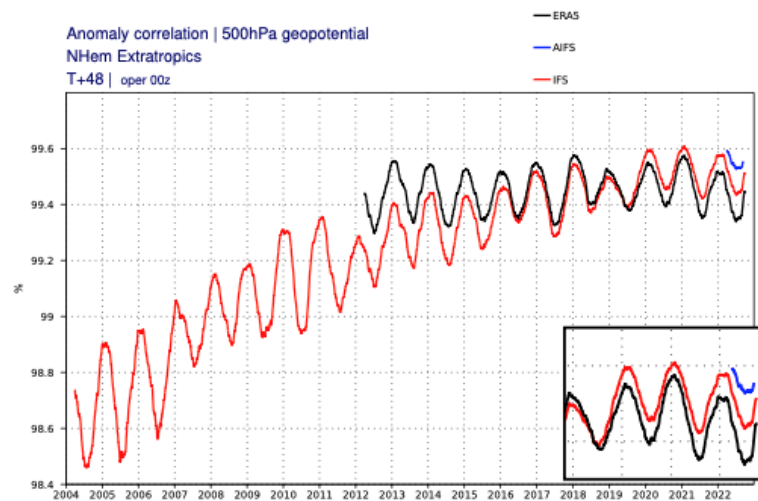
Florence Rabier

May 2024

ABSTRACT

Machine learning based weather forecasting models have quickly emerged as a plausible methodology for accurate medium range global weather forecasting. Here, we introduce the Artificial Intelligence Forecasting System (AIFS), a data driven forecast model developed by the European Centre for Medium-Range Weather Forecasts (ECMWF). AIFS is based on a graph neural network (GNN) encoder and decoder, and a sliding window transformer processor, and is trained on ECMWF's

Thank you!



Northern Hemisphere ACC (anomaly correlation) of geopotential at 500 hPa for AIFS (blue) and IFS (red) and ERA5 (black) forecasts for different years: 2 day forecasts (top left), 6 day forecasts (top right) and 10 day forecasts (bottom). Forecasts are initialised on 00 UTC each day and shown is a 30 day running mean. Insets show a zoomed-in view.

z500 anomaly correlation (2022)

