AN OVERVIEW OF DATA-DRIVEN FORECASTING AT ECMWF

DestinE

ML Project

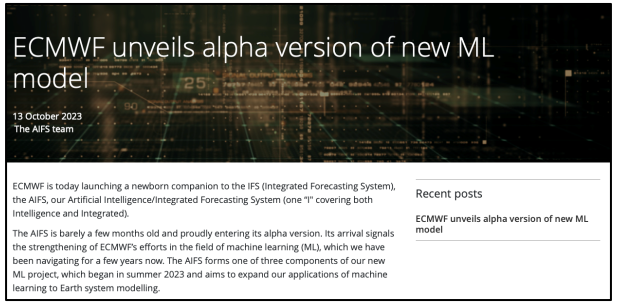
Member State Pilot Project

DestinE partners

Today's presentation

c.f. Alexe et al., Thurs 9am

Led by MET Norway & MeteoSwiss


ECMWF unveils alpha version of new ML model

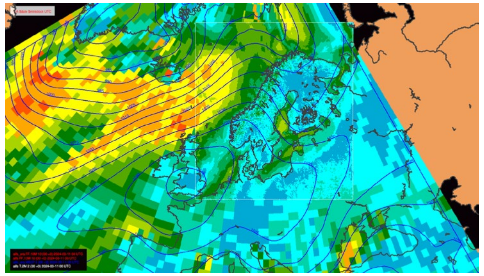
13 October 2023
The AIFS team

ECMWF is today launching a newborn companion to the IFS (Integrated Forecasting System), the AIFS, our Artificial Intelligence/Integrated Forecasting System (one "I" covering both Intelligence and Integrated).

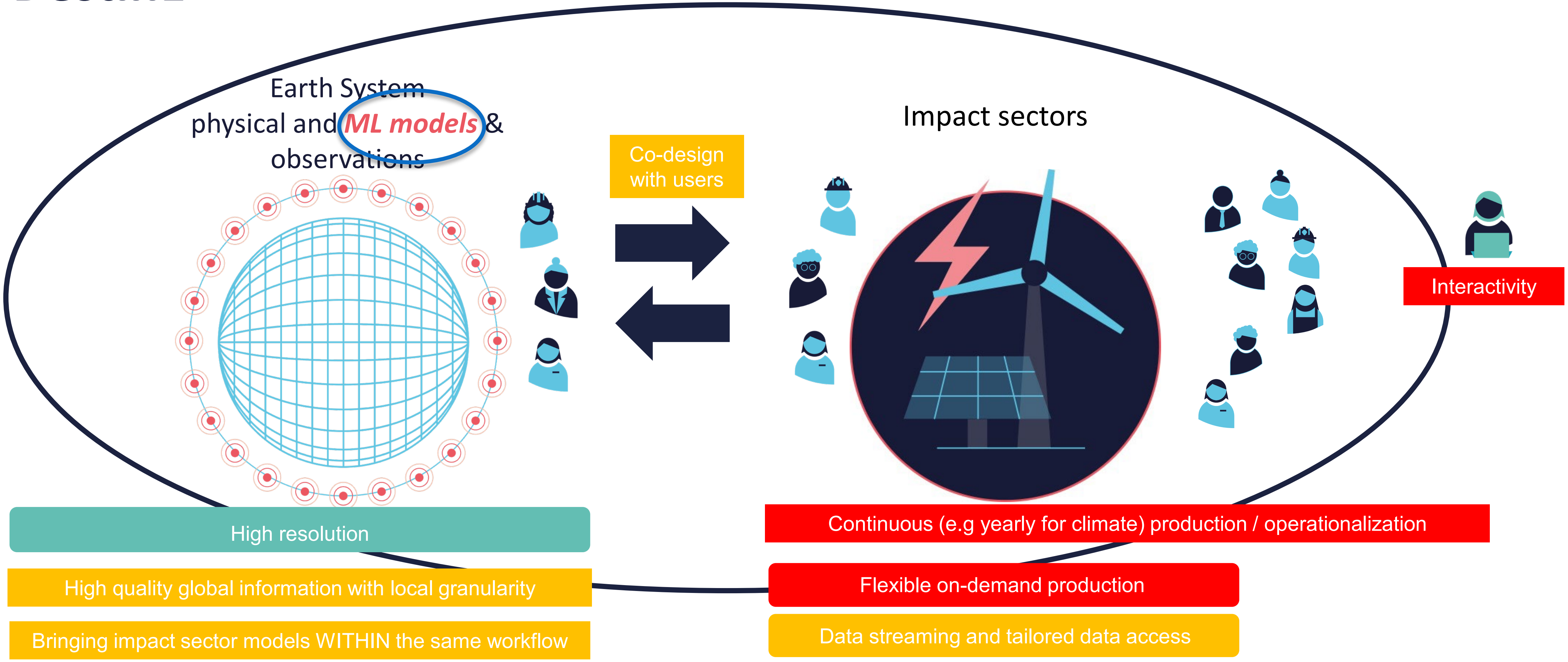
The AIFS is barely a few months old and proudly entering its alpha version. Its arrival signals the strengthening of ECMWF's efforts in the field of machine learning (ML), which we have been navigating for a few years now. The AIFS forms one of three components of our new ML project, which began in summer 2023 and aims to expand our applications of machine learning to Earth system modelling.

Recent posts

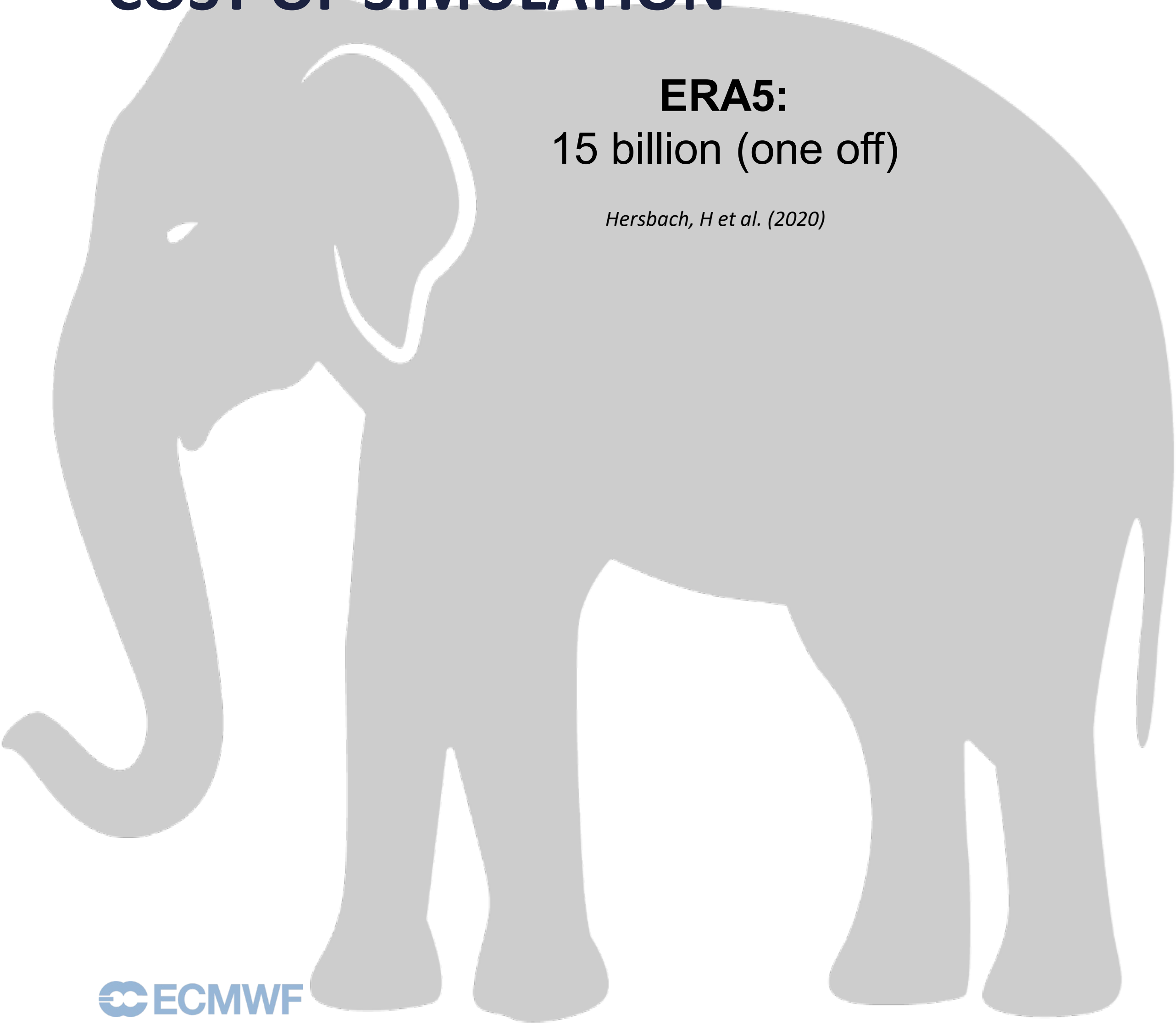
ECMWF unveils alpha version of new ML model



DestinE


Earth System
physical and ML models &
observations
Co-design
with users
Impact sectors
Interactivity
High resolution
High quality global information with local granularity
Bringing impact sector models WITHIN the same workflow
Continuous (e.g yearly for climate) production / operationalization
Flexible on-demand production
Data streaming and tailored data access

COST OF SIMULATION



ERA5:
15 billion (one off)

Hersbach, H et al. (2020)

DestinE 4.4 km:
1 600 000
per forecast

ECMWF Operational:
180 000
per forecast

AI Model:
0.3
per forecast

For ensemble forecasts, multiply this cost by number of ensemble members

UNCERTAINTY QUANTIFICATION FROM A DETERMINISTIC FORECAST

Use a Bayesian Neural Network to predict the distribution of the km-scale DestinE forecast error

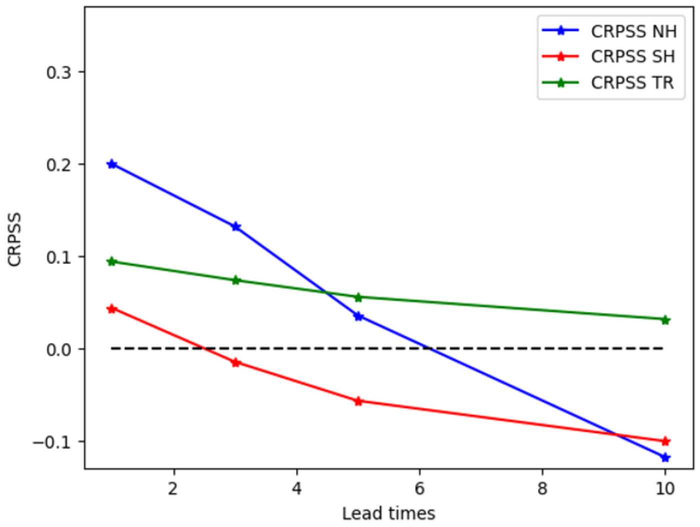
Post-processed probabilistic forecast = Deterministic km-scale forecast + Probabilistic Forecast Error

f_i

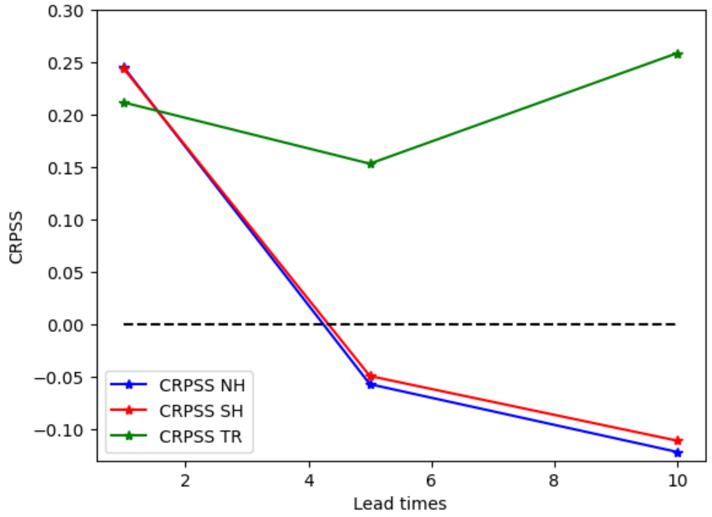
Better or comparable CRPS at short lead times but clear degradation from day 5 in extra-tropics.

At all lead times, post-processed forecasts have spread/skill ≈ 1

2m temperature



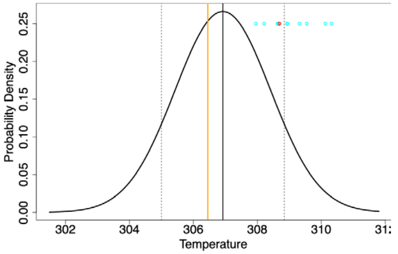
Geopotential at 500hPa

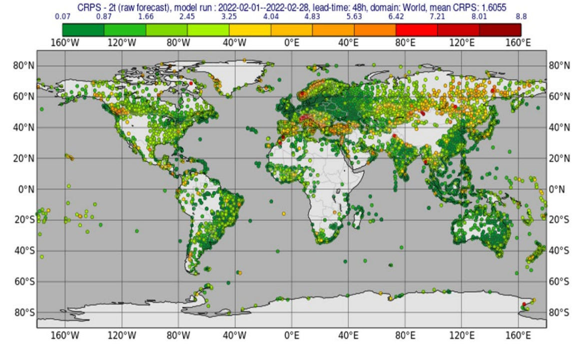


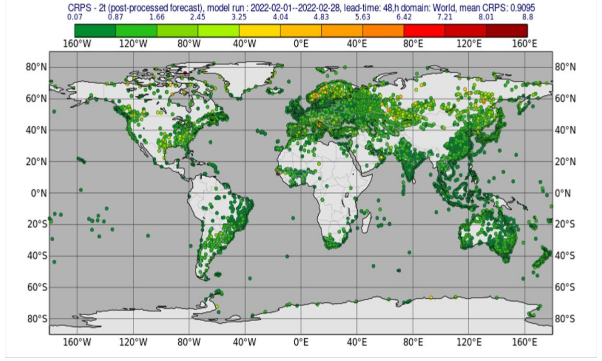
CRPSS relative to operational ensemble

UNCERTAINTY QUANTIFICATION AGAINST OBSERVATIONS

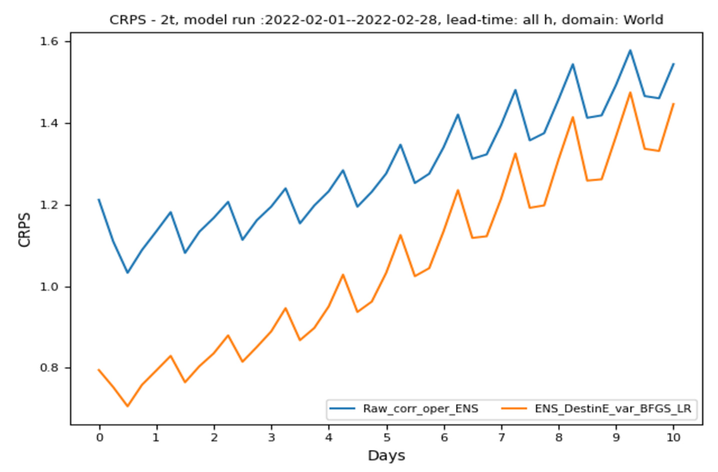
Use Ensemble Model Output Statistics (EMOS) Method to predict the distribution of the km-scale DestinE forecast





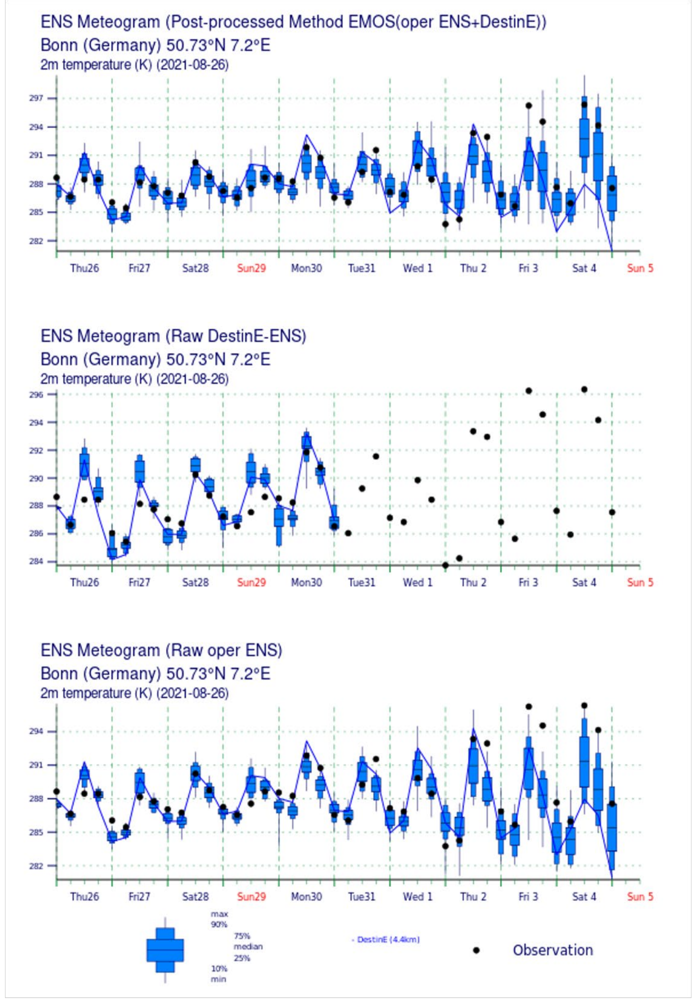


Post-processed probabilistic forecast = estimated conditional distribution using operational ensemble ENS, single DestinE forecast
+ observations



EMOS generated post-processed ensemble forecasts for 2mT (trained on 30 previous days rolling period):

- better performance vs. raw ensembles using CRPS for all lead-times;
- Meteograms (vs. raw and DestinE ENS);

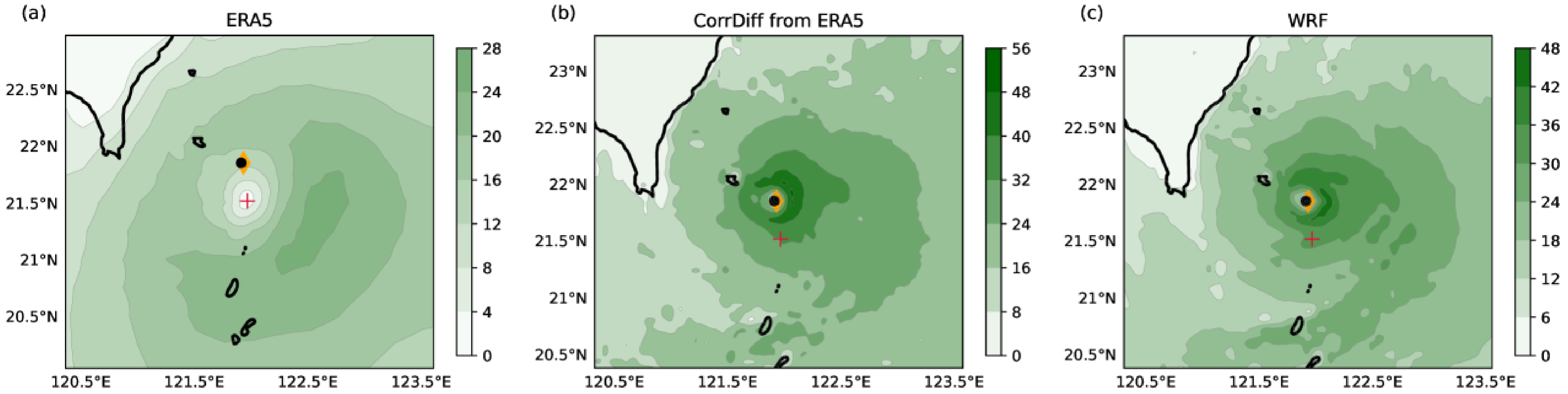


For more results including 10m wind speed and bias and spread/error see Ivana Aleksovska's Poster

DOWNSCALING ENSEMBLE MEMBERS TO KM-SCALE RESOLUTIONS (ONGOING WORK)

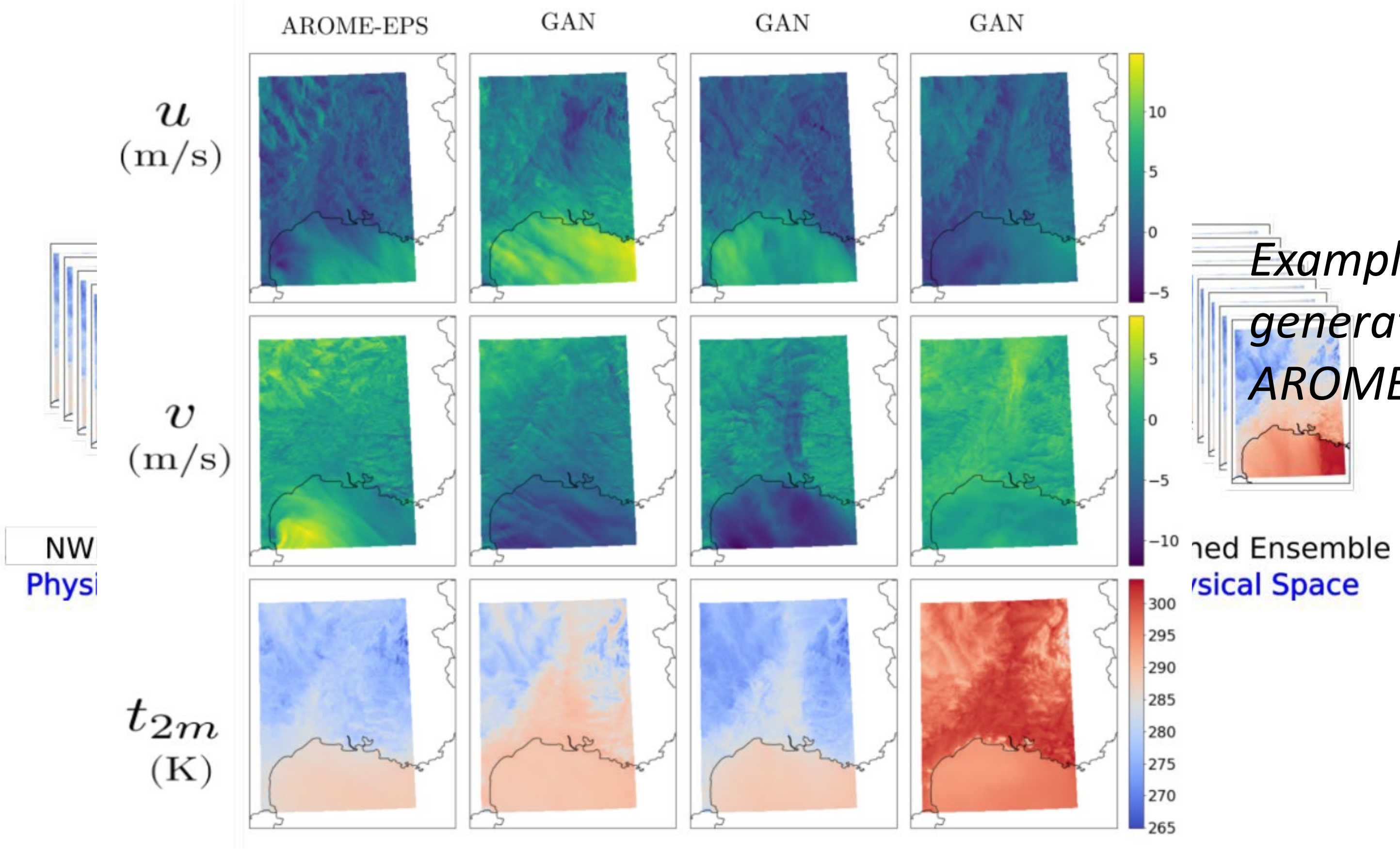
Can use diffusion models to downscale ensemble members, thus producing high-resolution ensembles much more efficiently than classical approaches

Example from the literature:
CorrDiff trained to downscale ERA5 to WRF simulations at 2km resolution over Taiwan (Mardani et al., 2023)



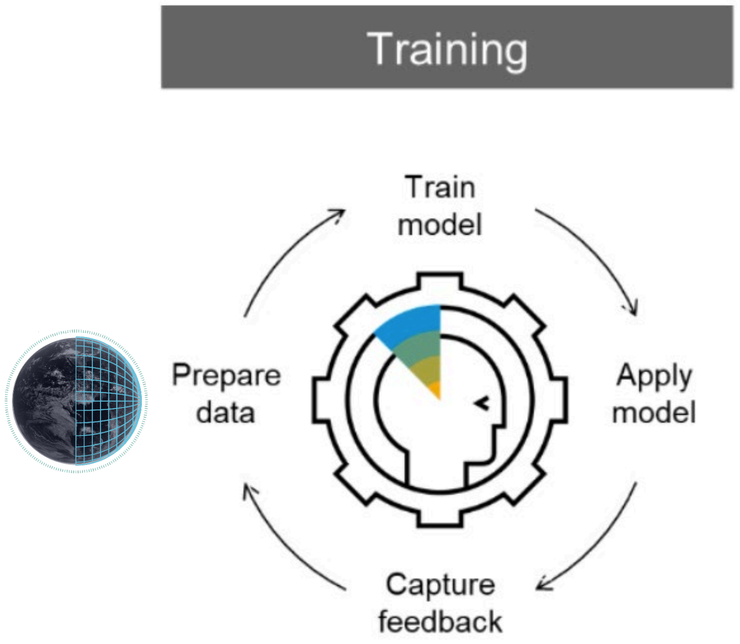
UNCERTAINTY QUANTIFICATION AND TEMPORAL INTERPOLATION (ONGOING WORK BY MET NORWAY, METEO FRANCE, SMHI)

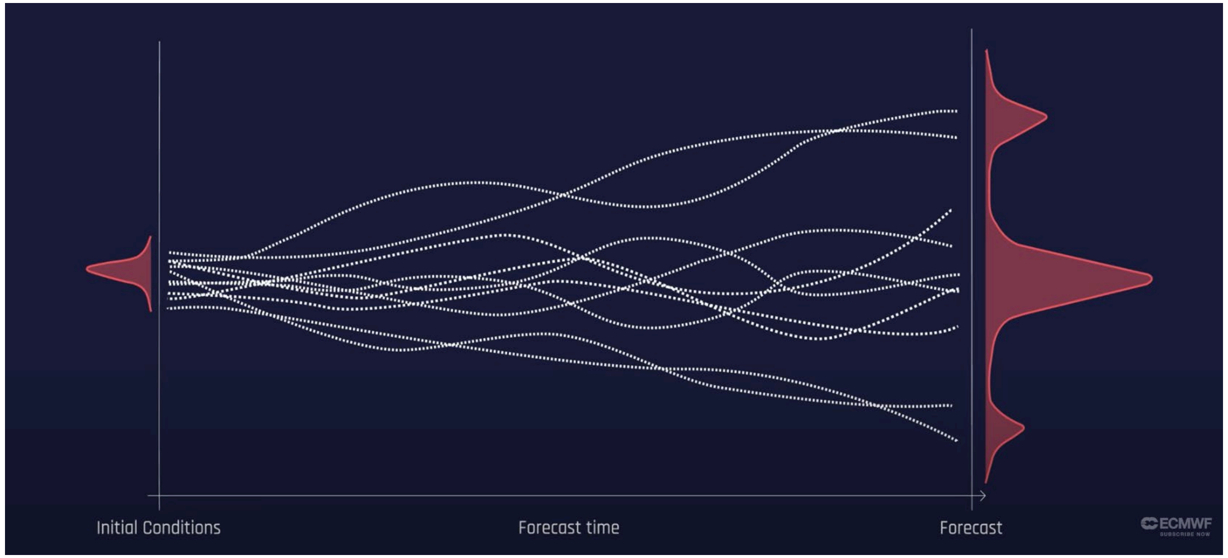
StyleGAN - Input can be ensemble or deterministic forecast



Example ensemble members generated by MeteoFrance on AROME data (Brochet et al., 2023).

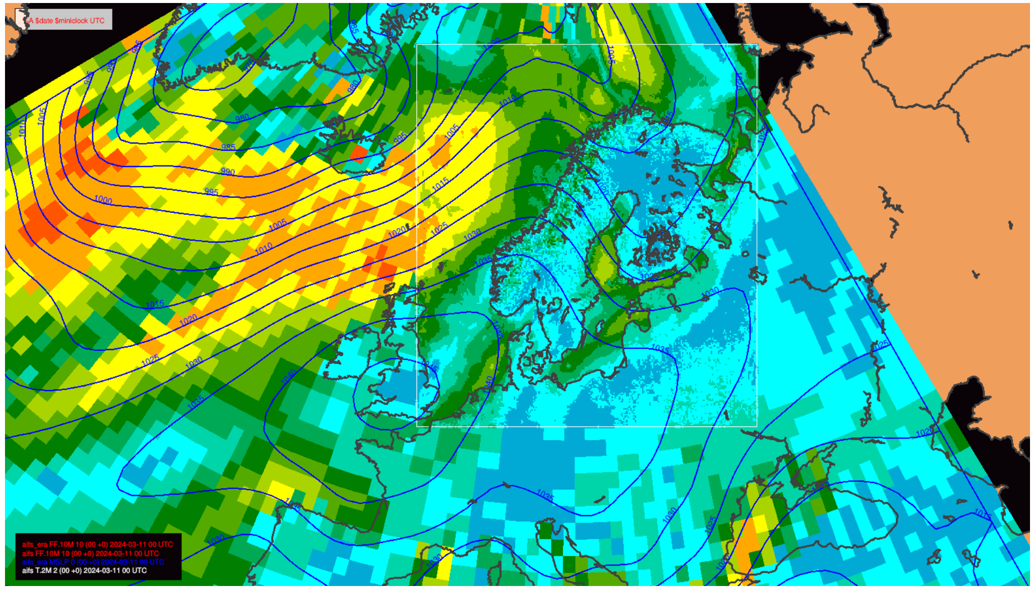
DATA-DRIVEN FORECASTS FOR UNCERTAINTY QUANTIFICATION



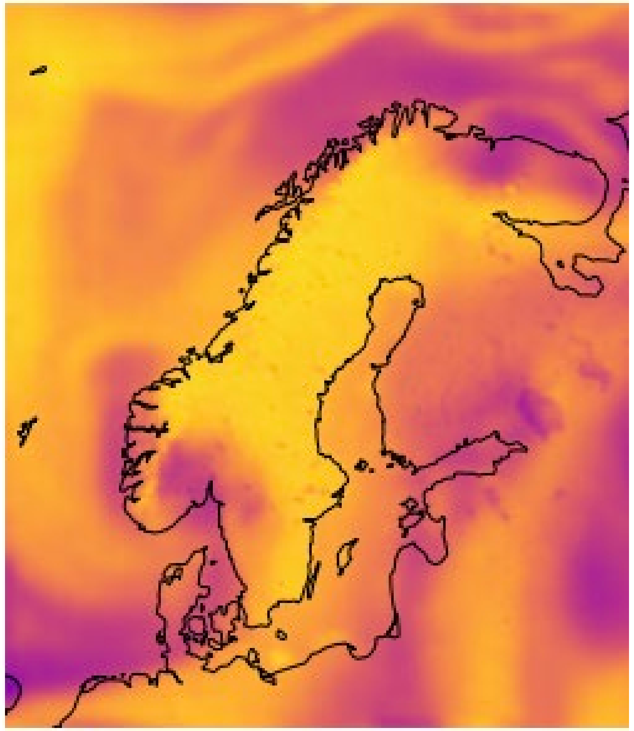


Developing & running both global & local data-driven models to create ensembles that complement DestinE simulations

EXAMPLES IN LITERATURE: LOCAL DATA-DRIVEN FORECAST MODELS



Stretched grid model (Nipen et al., 2024)



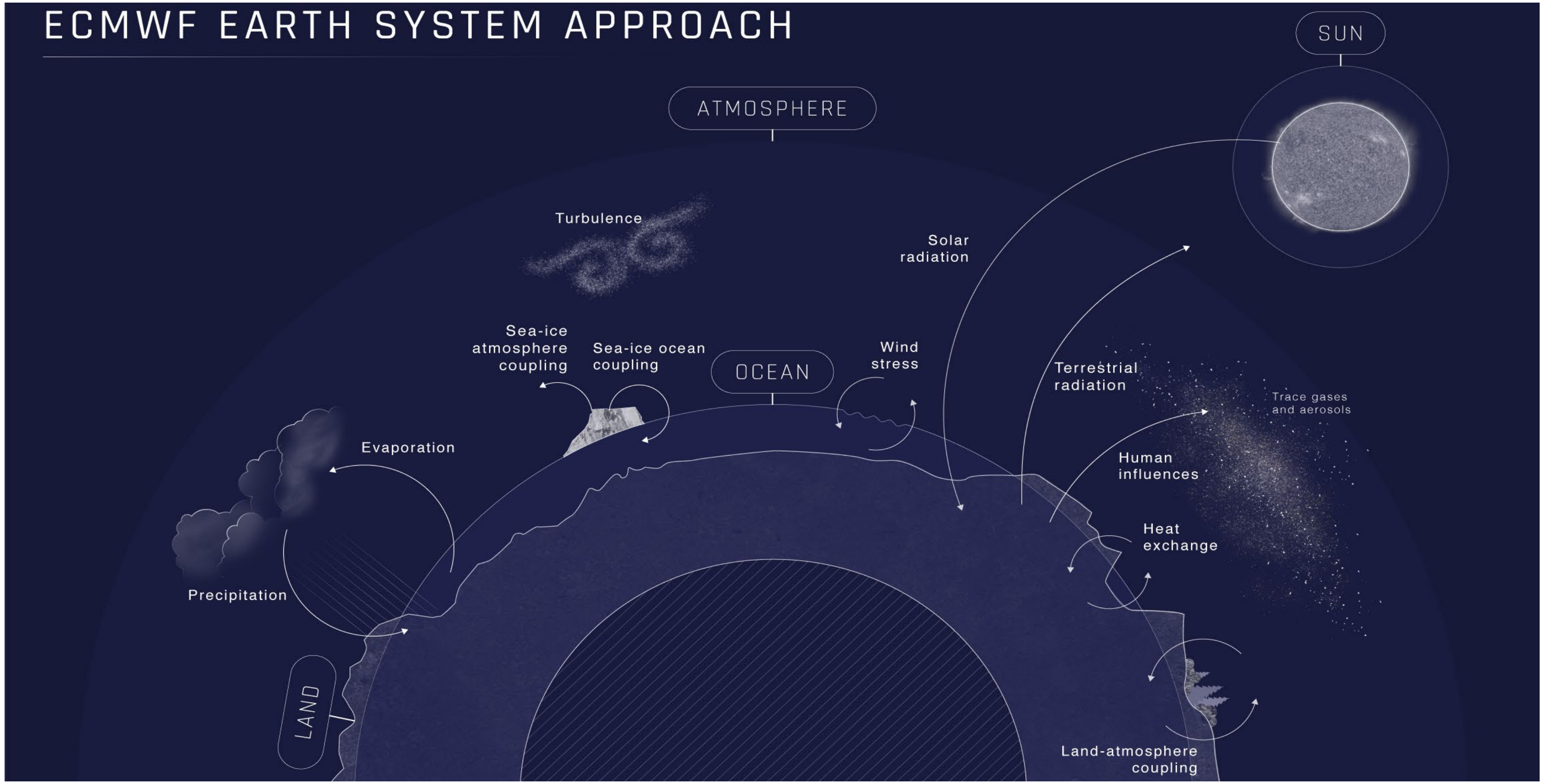
Limited Area Model (Oskarsson et al., 2023)

Cf. Oskarsson et al. Thurs 9.50am; Buurman et al. (Poster)

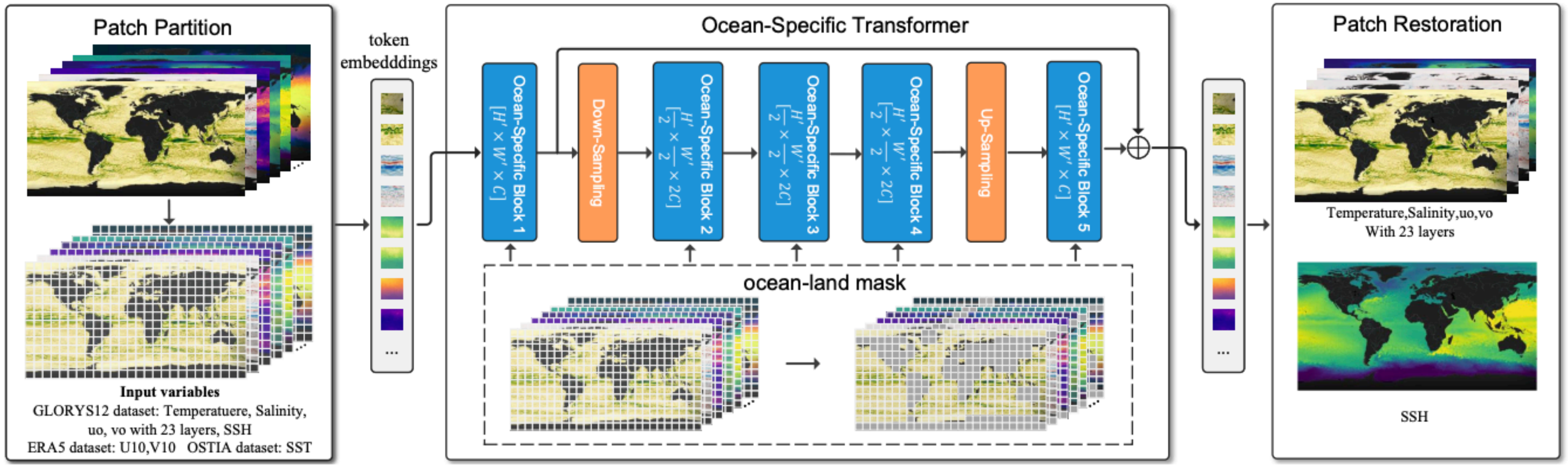
AI EARTH SYSTEM MODEL

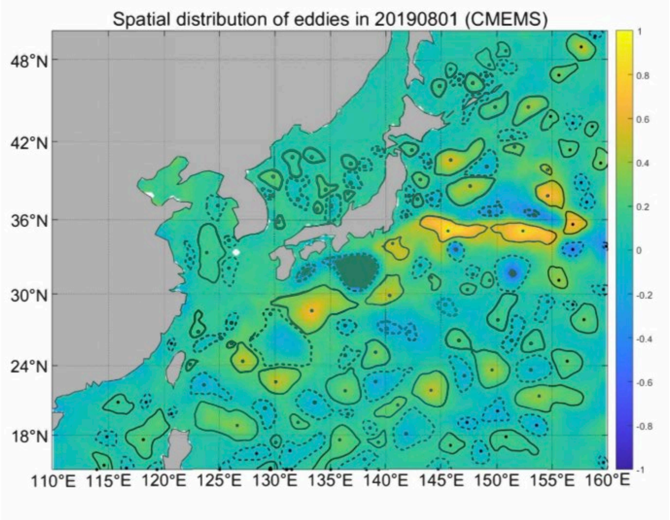
Build full Earth System model with land, ocean, sea-ice and hydrology components

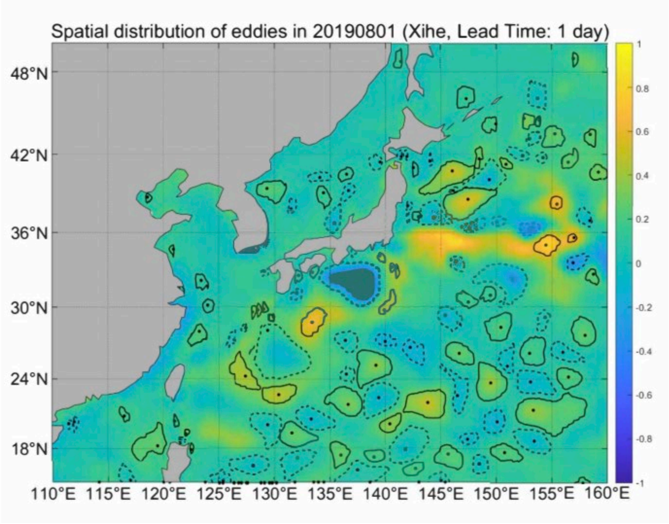
Leverage developments made in the ML project especially ensemble developments and learning from observations



EXAMPLE FROM THE LITERATURE:
XIHE: A DATA-DRIVEN MODEL FOR GLOBAL OCEAN EDDY-RESOLVING FORECASTING

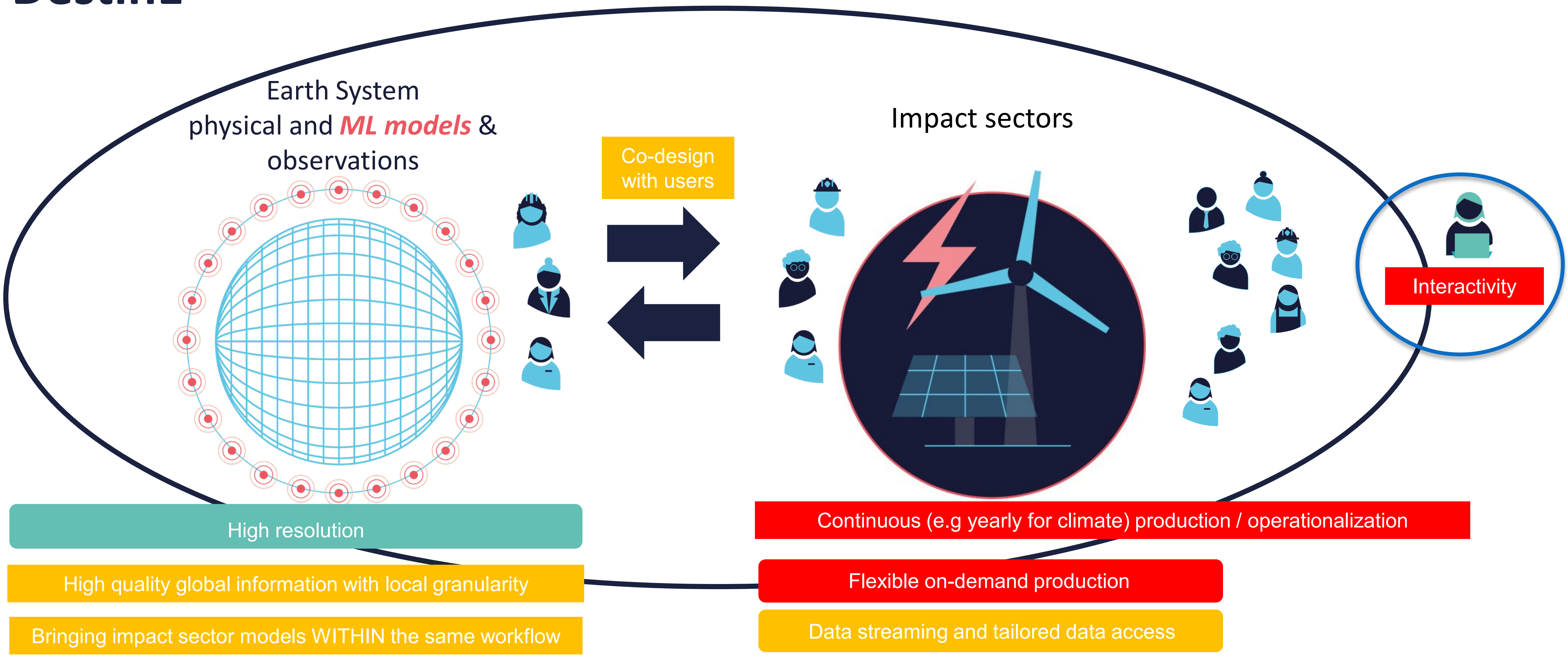






Wang et al. (2024)

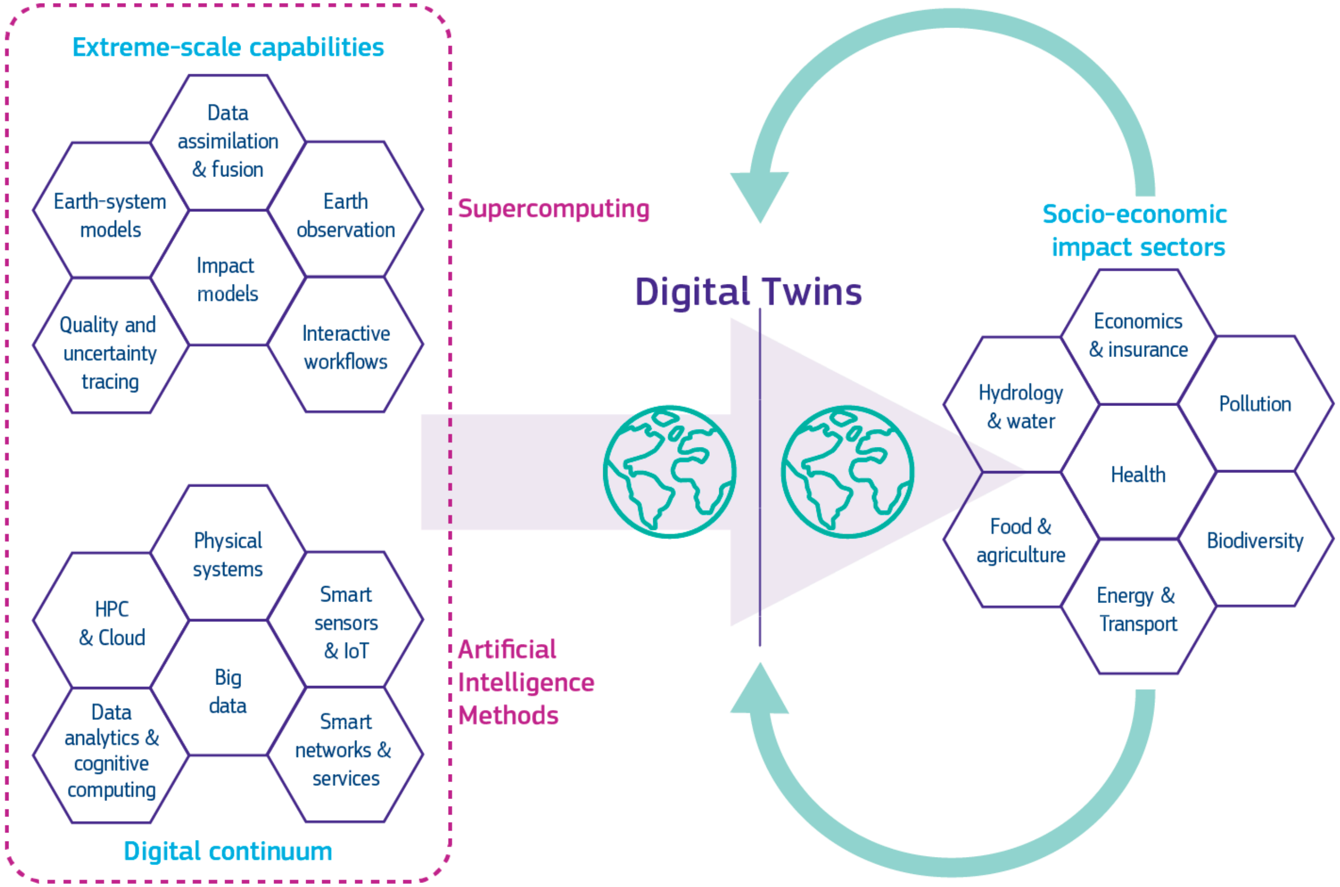
DestinE


Earth System
physical and *ML models* &
observations
Co-design
with users
Impact sectors
Interactivity
High resolution
High quality global information with local granularity
Bringing impact sector models WITHIN the same workflow
Continuous (e.g yearly for climate) production / operationalization
Flexible on-demand production
Data streaming and tailored data access

INTERACTIVITY AND ACCESSIBILITY

Machine learning will be used to help stakeholders and policy makers interact with the digital twins.

This will make the data more accessible to users

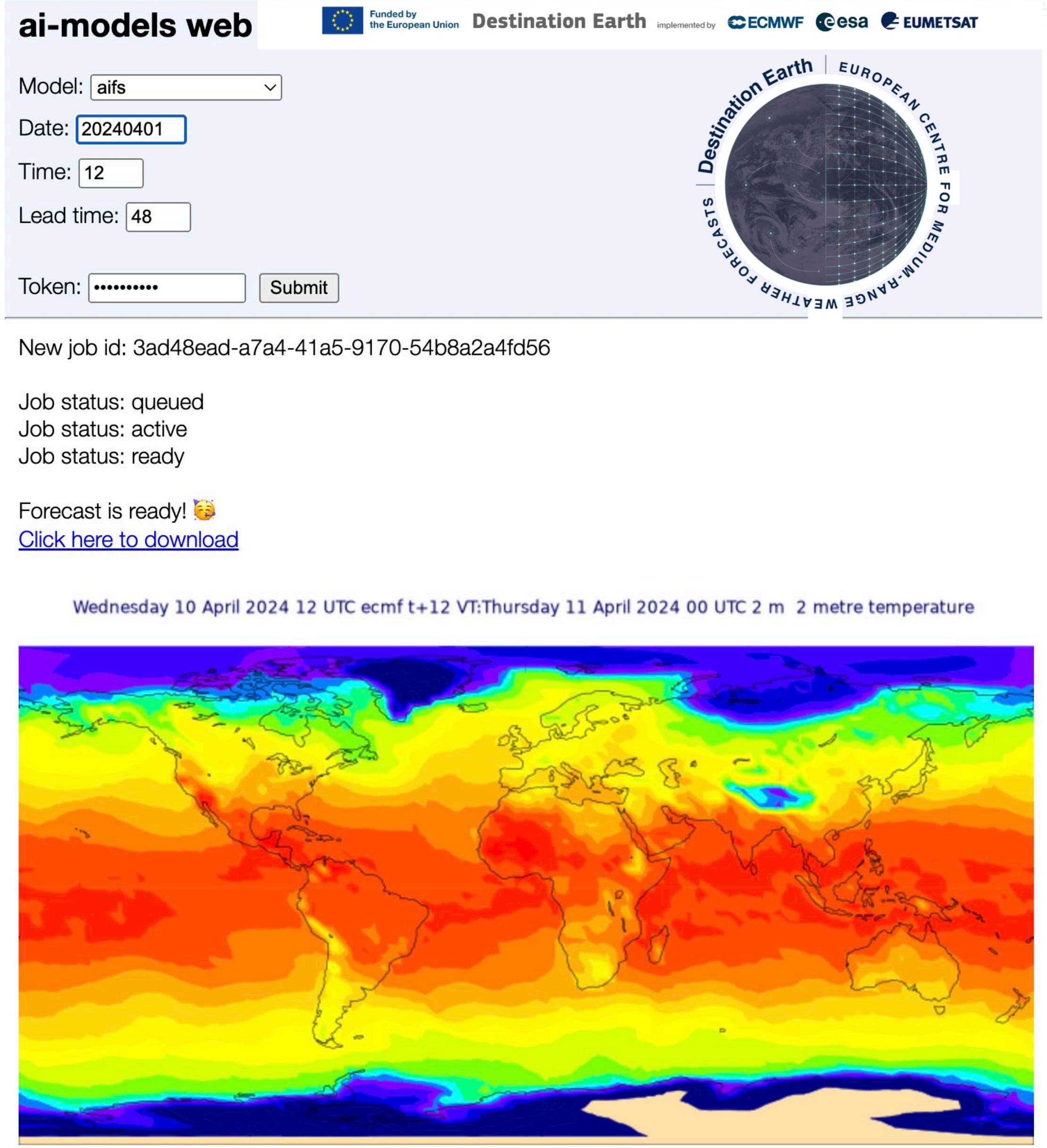


FORECAST-IN-A-BOX

Providing a packaged system with data-retrieval, forecasting & postprocessing.

This system runs on local hardware or cloud and is delivered in a matter of minutes

It is configurable for Earth-System components and user-defined outputs.



CLIMSIGHT – INSPIRATION FOR ENHANCED INTERACTIVITY

communications earth & environment

Explore content About the journal Publish with us

nature > communications earth & environment > comment > article

Comment | Open access | Published: 05 January 2024

Local climate services for all, courtesy of large language models

Nikolay Koldunov & Thomas Jung

Communications Earth & Environment **5**, Article number: 13 (2024) | Cite this article

Koldunov & Jung 2024

User request: "What is the best crop to grow here"

Location: 5.3992N, 6.5163E

Climate impact evaluation

Based on the environmental variables, the location Ogba/Egbema/Ndoni in Rivers State, Nigeria is currently experiencing a tropical climate with a mean monthly temperature ranging from 24.088 °C to 29.071 °C and high precipitation throughout the year. The soil type is Acrisols which are typically acidic and low in base cations but can be suitable for a variety of crops with appropriate management.

In the future, there is a projected increase in temperature and an overall increase in precipitation, particularly in the late part of the year. The wind direction also seems to be shifting more towards the east and south.

Crop recommendations

Taking into consideration the local climate and soil conditions, the following crops could be suitable:

- Cassava: This crop thrives in warm, tropical climates and can tolerate poor soil conditions. With the projected increase in temperature, Cassava could be a viable option.

SUMMARY OF AI ACTIVITIES

Towards an earth-system machine learning model leveraging DestinE data

Developing end-to-end workflows for ML model components like land, ocean, sea-ice, hydrology

Enhance Digital Twin Engine with ML pipelines from training to post-processing

Using data-driven methods for uncertainty quantification of Extremes and Climate Digital Twin

Climate emulator to rapidly explore ‘what-if’ scenarios

Enhanced interactivity

Developing a forecast-in-a-box concept.

Building ML demonstrators for impact-sectors (e.g., health, agriculture, urban)

Develop of a weather and climate chatbot

KEY REFERENCES

- Brochet, C., Raynaud, L., Thome, N., Plu, M., & Rambour, C. (2023). Multivariate Emulation of Kilometer-Scale Numerical Weather Predictions with Generative Adversarial Networks: A Proof of Concept. *Artificial Intelligence for the Earth Systems*, *2*(4), 230006.
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., ... & Thépaut, J. N. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999-2049.
- Koldunov, N., & Jung, T. (2024). Local climate services for all, courtesy of large language models. *Communications Earth & Environment*, *5*(1), 13.
- Lang, S., Alexe, M., Chantry, M., Dramsch, J., Pinault, F., Raoult, B., Ben Bouallègue, Z., Clare, M., Lessig, C., Magnusson, L., Prieto, A. N. (2024). AIFS: A new ECMWF forecasting system. ECMWF Newsletter, (178), 4–5. https://doi.org/10.21957/1a8466ec2f
- Mardani, M., Brenowitz, N., Cohen, Y., Pathak, J., Chen, C. Y., Liu, C. C., ... & Pritchard, M. (2023). Generative residual diffusion modeling for km-scale atmospheric downscaling. *arXiv preprint arXiv:2309.15214*.
- Nipen, T, Chantry, M. et al. (2024). Data driven regional modelling. https://www.ecmwf.int/en/about/media-centre/aifs-blog/2024/data-driven-regional-modelling
- Oskarsson, J., Landelius, T., & Lindsten, F. (2023). Graph-based neural weather prediction for limited area modeling. *arXiv preprint arXiv:2309.17370*.
- Wang, X., Wang, R., Hu, N., Wang, P., Huo, P., Wang, G., ... & Song, J. (2024). XiHe: A Data-Driven Model for Global Ocean Eddy-Resolving Forecasting. *arXiv preprint arXiv:2402.02995*.