Data-driven ocean modelling at ECMWF

Rachel Furner, Lorenzo Zampieri, Mario Santa Cruz, Sara Hahner, Sarah Keeley, Kristian Mogensen

ECMWF, Reading, United Kingdom and Bonn, Germany. rachel.furner@ecmwf.int

Overview of activities

ML activities at ECMWF look to develop ocean and sea-ice models for predicting from medium range out to seasonal forecast periods. Varying modelling strategies are employed for the different use cases.

In parallel we are developing a surface only model, and a 3D model. Experiments have considered different time steps, different architectures, and different vertical structure. Prototype models exist for a variety of use cases, and an ML based surface ocean component is expected to become part of the AIFS medium range forecast system in the coming years. In terms of coupling these models to the atmosphere (and other components) we are investigating the impact of training a full ocean-atmosphere model "all at once", vs training individual components which pass fields during inference in the more traditional manner.

Ocean models have been trained on the ORAS6 reanalysis, with 1993-2022 used for testing, and 2023 for validation. We model the ocean temperature, salinity, currents, and sea surface height anomaly. For sea-ice, we model the concentration, salinity, albedo, velocities, and snow volume over ice. We allow for atmospheric forcing by providing forcing fields including 10m winds, 2m temperature and humidity, the mean sea-level pressure, long & short wave radiation, & precipitation.

Zuo, Hao, et al. "ECMWF's next ensemble reanalysis system for ocean and sea ice: ORAS6." ECMWF Newsletter 180.10.21957 (2024).

3D ocean model

Our 3D ocean model captures the full ocean depth, taking data from roughly every 5th level of ORAS6. The model takes 24 hour time steps. A graph transformer is used with attention over a local mesh.

Preliminary results looking at medium range forecasting are shown here, with comparisons to profile observations for temperature and salinity, and surface temperature from drifting buoys, all at 10 day lead time. This model will provide the most benefit in long range forecasting, such as seasonal timescales, when the sub-surface ocean becomes increasingly important, and future work will address longer term stability & skill.

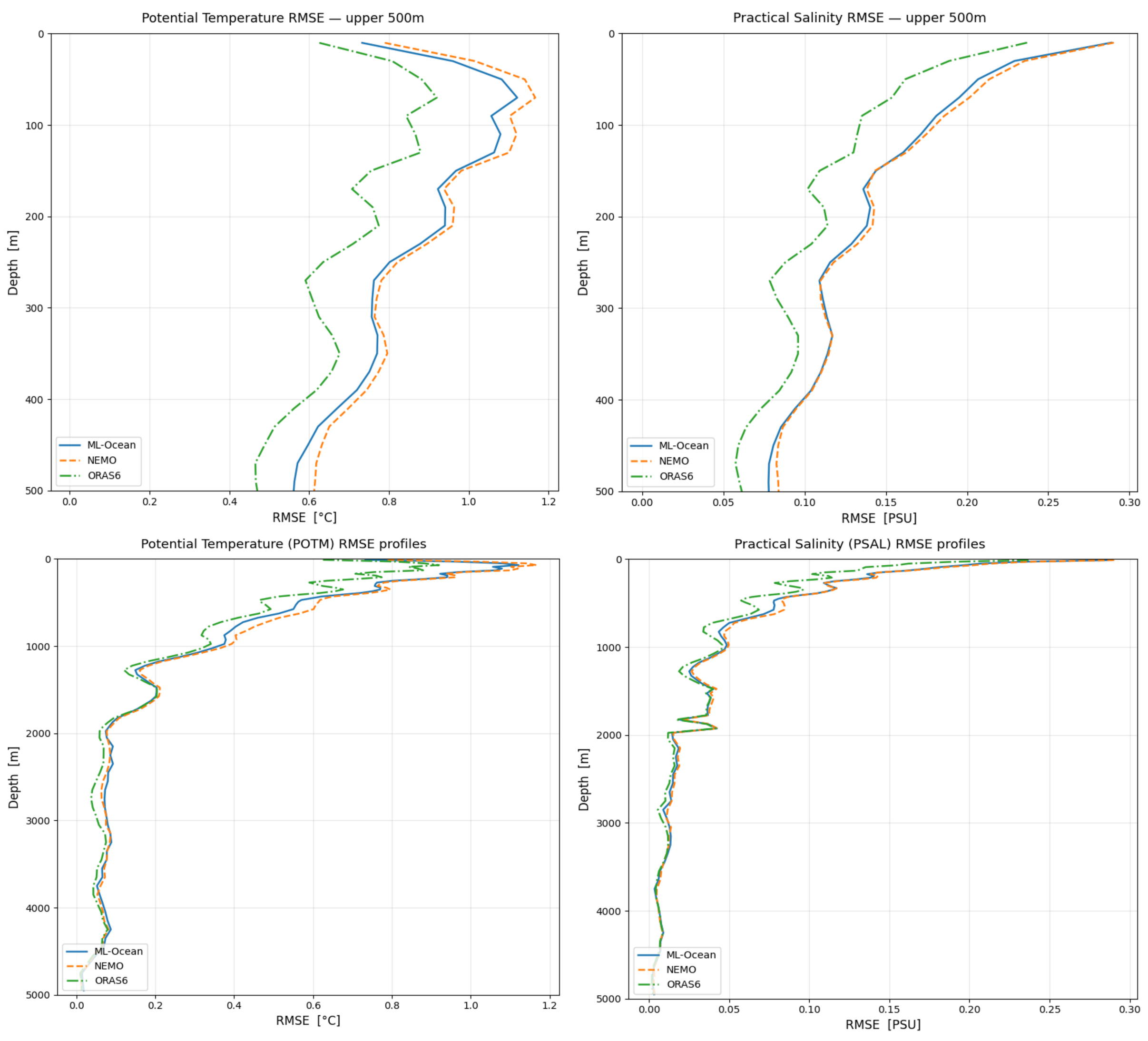


Figure: Profile RMSE for temperature (l) &salinity (r) for ORAS6, NEMO, and ML-Ocean at 10 day lead time for the top 500m (top) and 5000m (bottom)

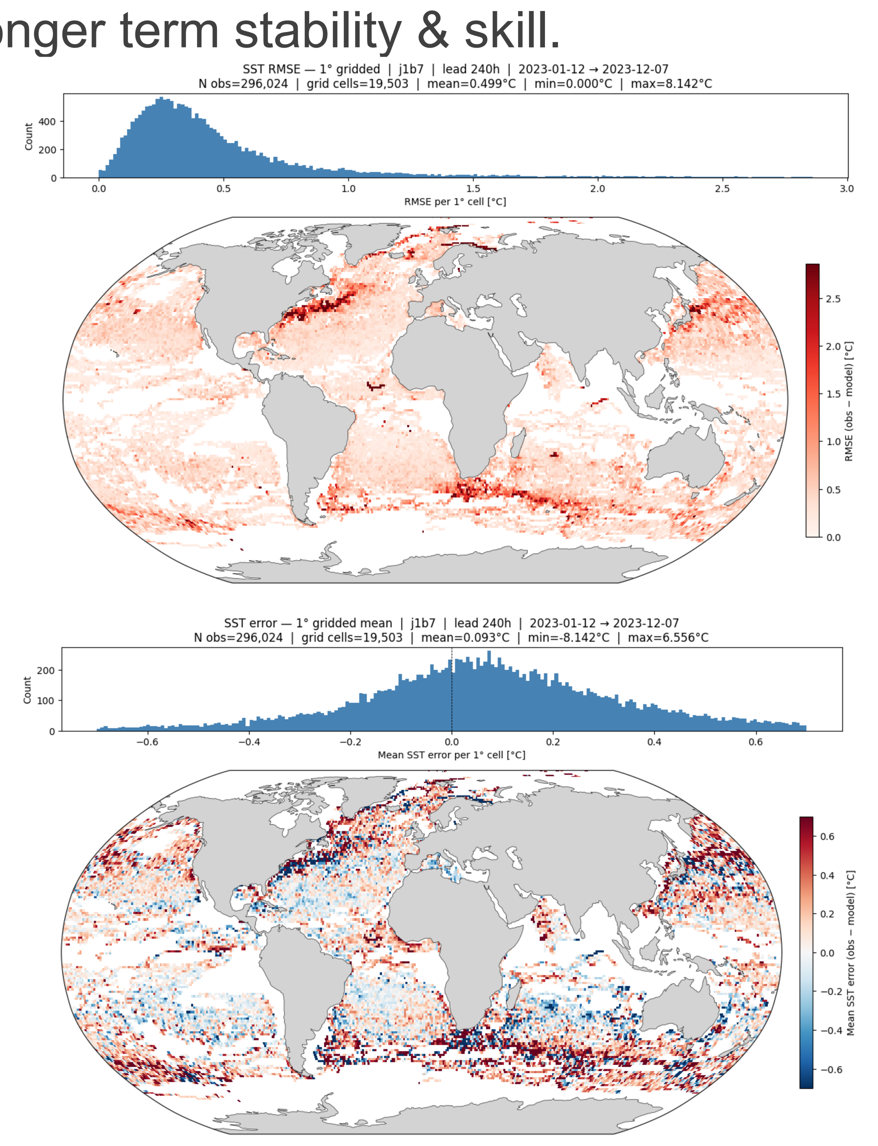


Figure: ML-Ocean RMSE (top) and mean error (bottom) for SST compared with surface drifter observations at 10 day lead time

Surface-only ocean

We model the surface ocean & sea ice as a standalone model, and also jointly with the atmosphere (AIFS-Ocean), and with the atmosphere and waves (AIFS-Marine). This configuration is suited to medium-range and sub-seasonal forecasting. These models runs with a 6 hour time step. They use a sliding window transformer, as used in the AIFS.

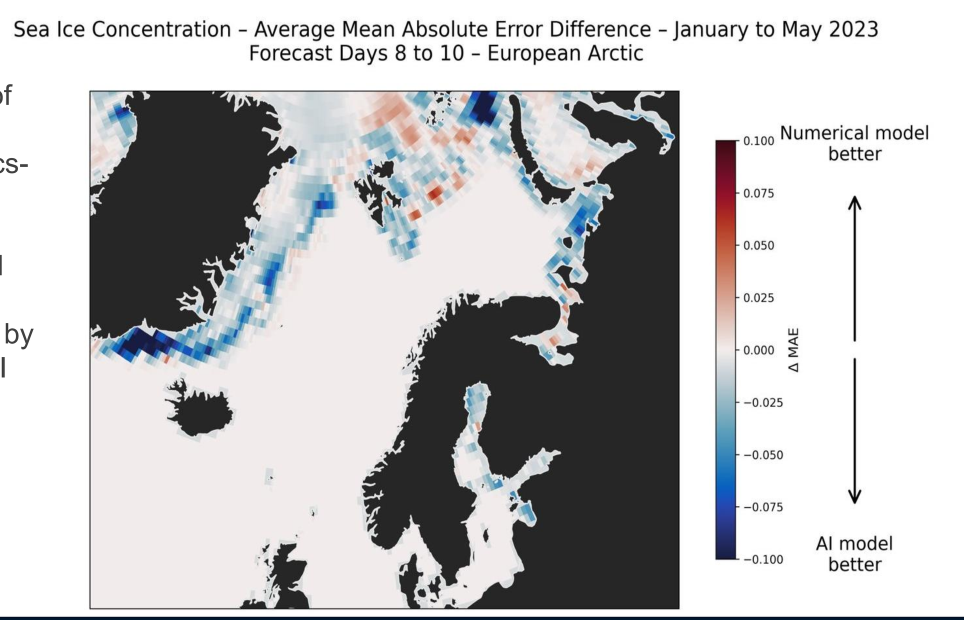


Figure: Comparison of sea ice concentration errors from the physics-based IFS-NEMO coupled system (the numerical model) and the standalone ML surface model forced by AIFS forecasts (the AI model). Results are shown over the arctic during the spring melting season.

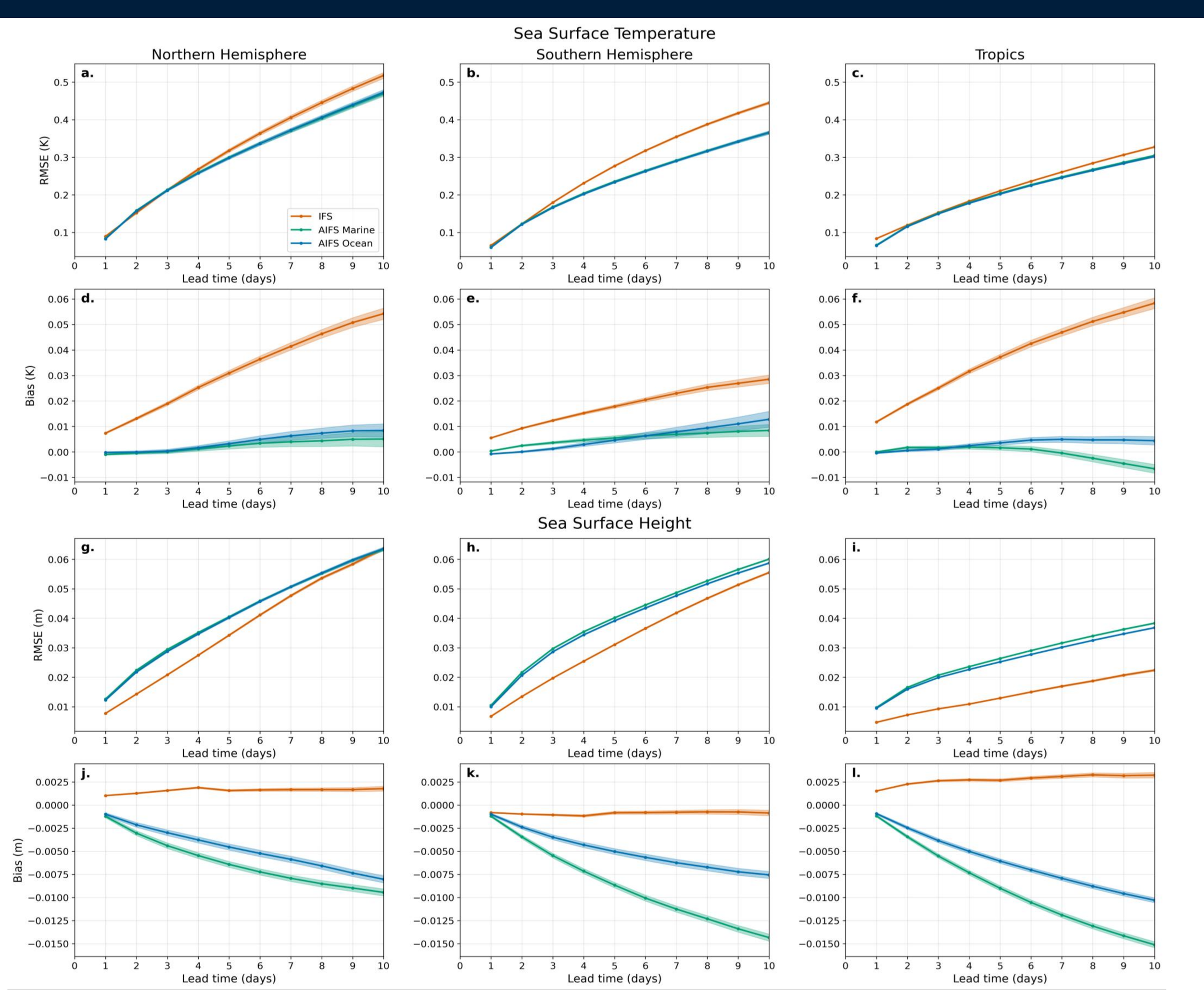


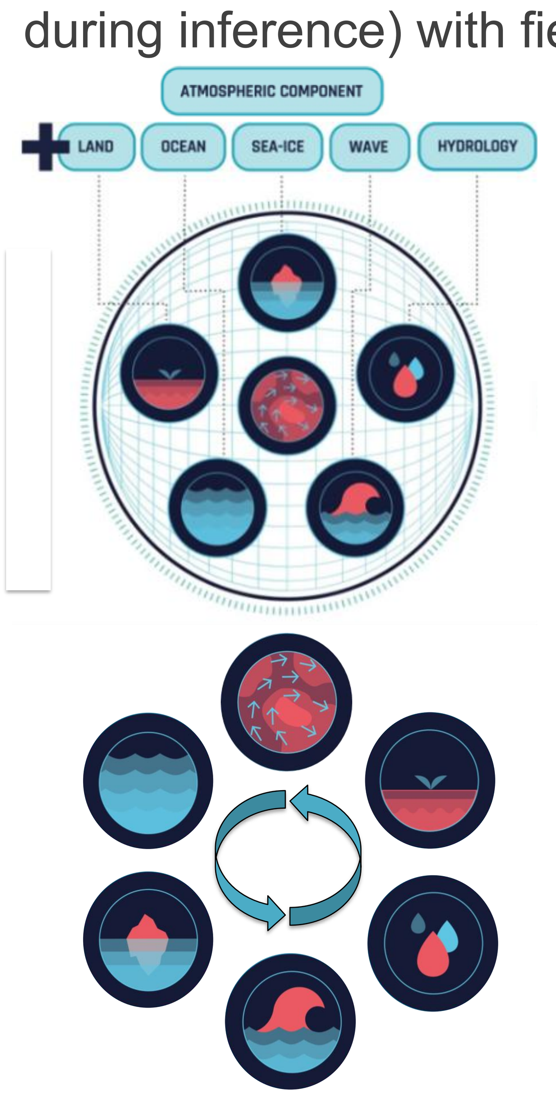
Figure: RMSE and Bias for SST and SSH, separated by region, for 3 models: the physics-based IFS-NEMO coupled system; AIFS-Ocean (joint ML model of atmosphere, surface ocean, & sea-ice); AIFS-Marine (joint ML model of atmosphere, surface ocean, sea-ice, & waves)

An ML-based earth system model

This work is part of a wider initiative to develop a full ML-based Earth system model within the Destination Earth project. As well as ocean and sea-ice models, machine learning based wave, land surface, and hydrology models are being developed at ECMWF. You can find out more about all of these in our blog series:

Coupling approaches

We are exploring two approaches to coupling. In one, we have a joint model where variables from all components are modelled together, in a component agnostic way. The ML model sees and predicts all variables at once. The second approach mimics a traditional coupled model set up, whereby models for each individual component are built using idealised forcing. These are then coupled at run-time (i.e. during inference) with fields being passed between the models.



FOR ATMOSPHERE:
Both approaches are viable and present overall similar behaviours.

FOR OCEAN:
The coupled model is a better approach for ocean fields. Joint training leads to degradation for all variables.

Experiment Configuration:

- Atmosphere (786) + ocean (512), **coupled model** — 696 forecasts
- Ocean (512) **forced model** — 696 forecasts
- Atmosphere + ocean (1024), **joint model** — 696 forecasts

Coupling timestep: 6h (also the forecast timestep)

Verification year: 2023

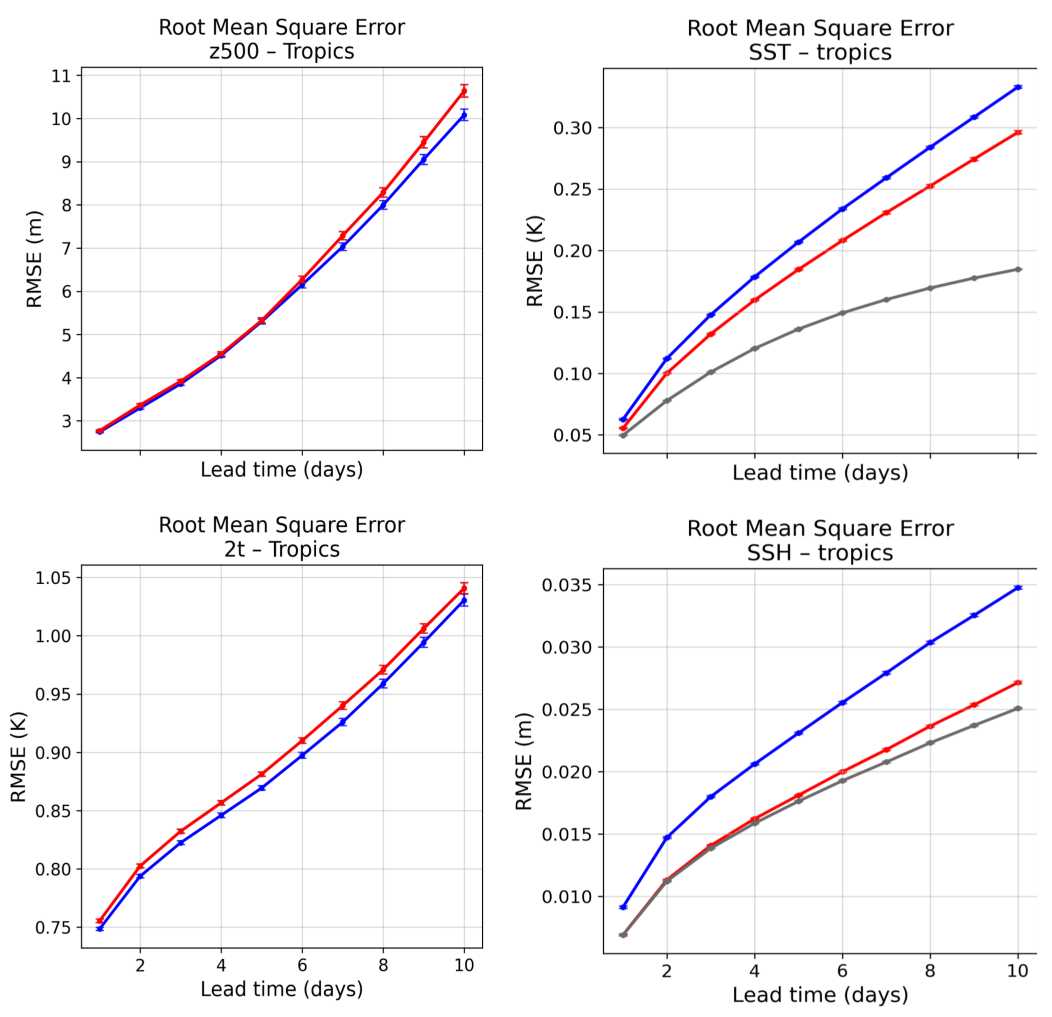


Figure: Joint vs coupled model RMSE for atmospheric (l) & ocean (r) variables

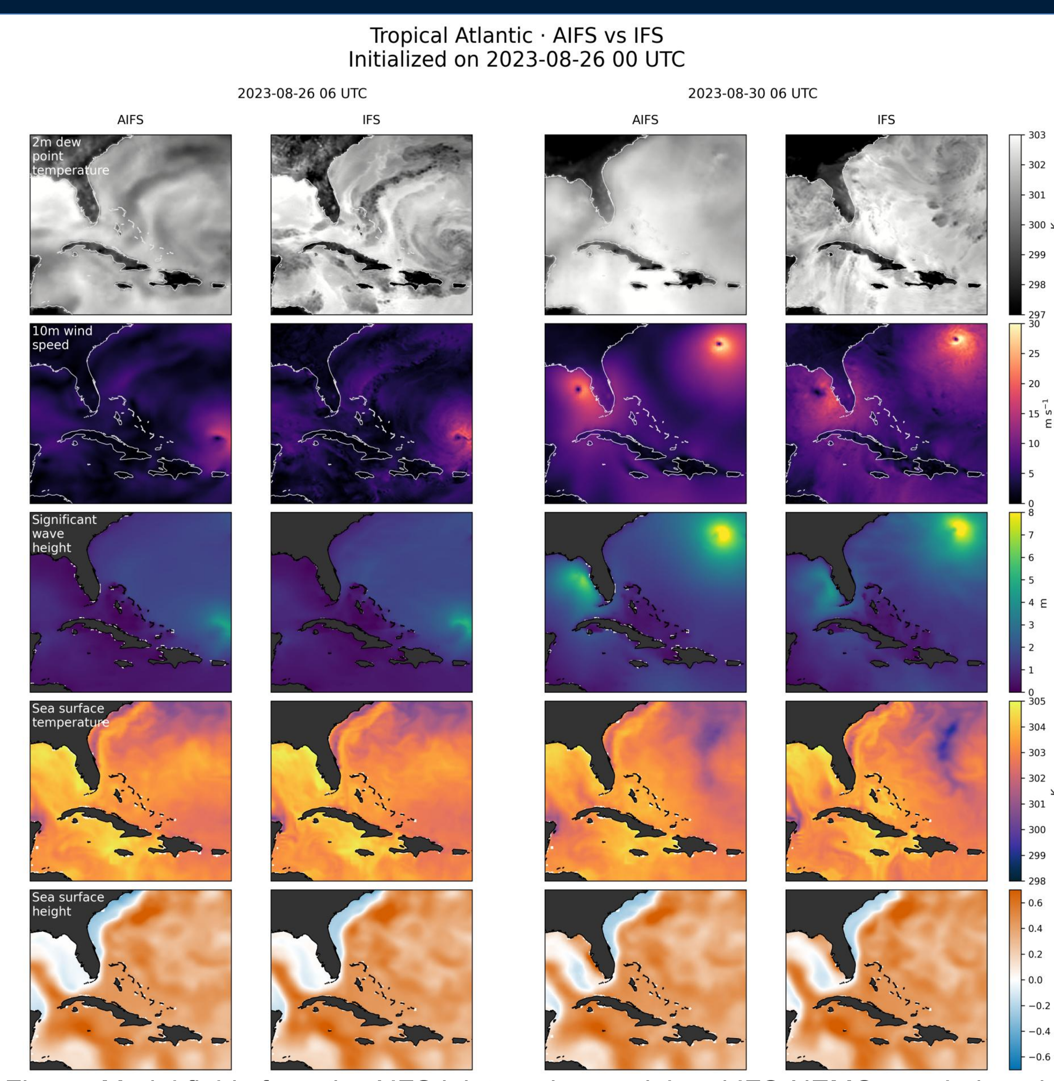


Figure: Model fields from the AIFS joint marine model and IFS-NEMO coupled model for the tropical Atlantic, looking at a hurricane case study.

We show a tropical cyclone case study from a joint ML model of the atmosphere, waves, surface-ocean and sea-ice. Results are broadly consistent with the physics-based IFS prediction, albeit with some smoothing and lower resolution. In particular:

- Cold wakes develop along the TC track.
- Storm surges are predicted in coastal regions.
- The high winds generate ocean waves.

Funded by the European Union

Destination Earth

implemented by ECMWF esa EUMETSAT