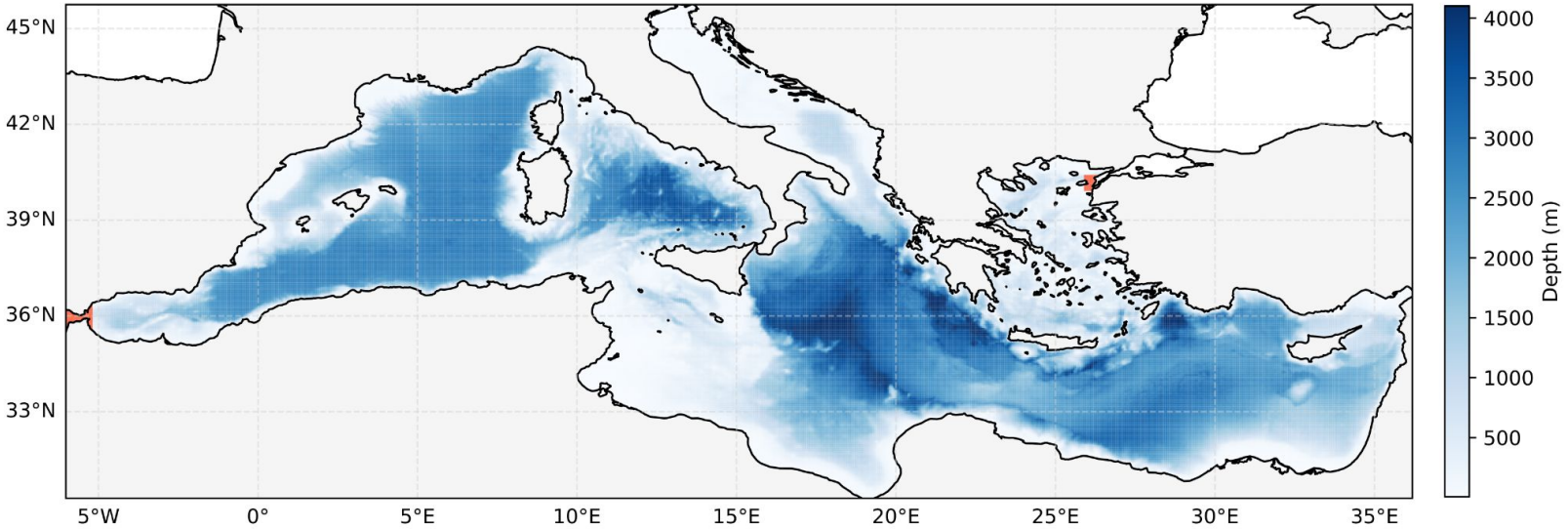
Accurate Mediterranean Sea Forecasting via Graph-based Deep Learning

Daniel Holmberg 22.04.2025

UNIVERSITY OF HELSINKI
FACULTY OF SCIENCE

Study Domain



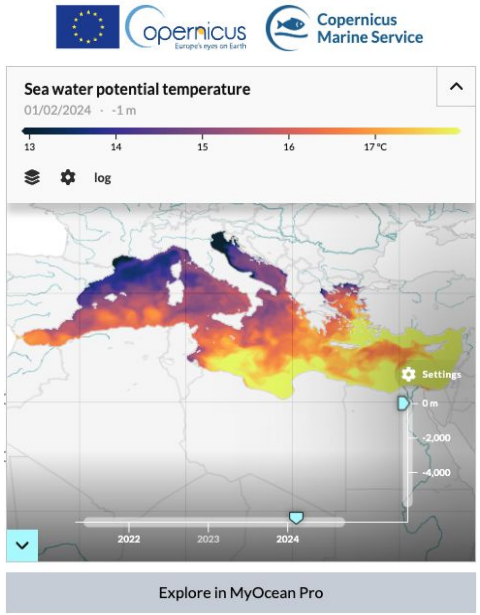
Mediterranean Physics Dataset

Variables

- Covers every other simulated depth down to 200 m (18 in total).
- 73 variables in total (18×4 + 1): temperature, salinity, meridional and zonal currents + single level SSH.
- Atmospheric forcing 10m wind stress, 2m temp, and MSLP.

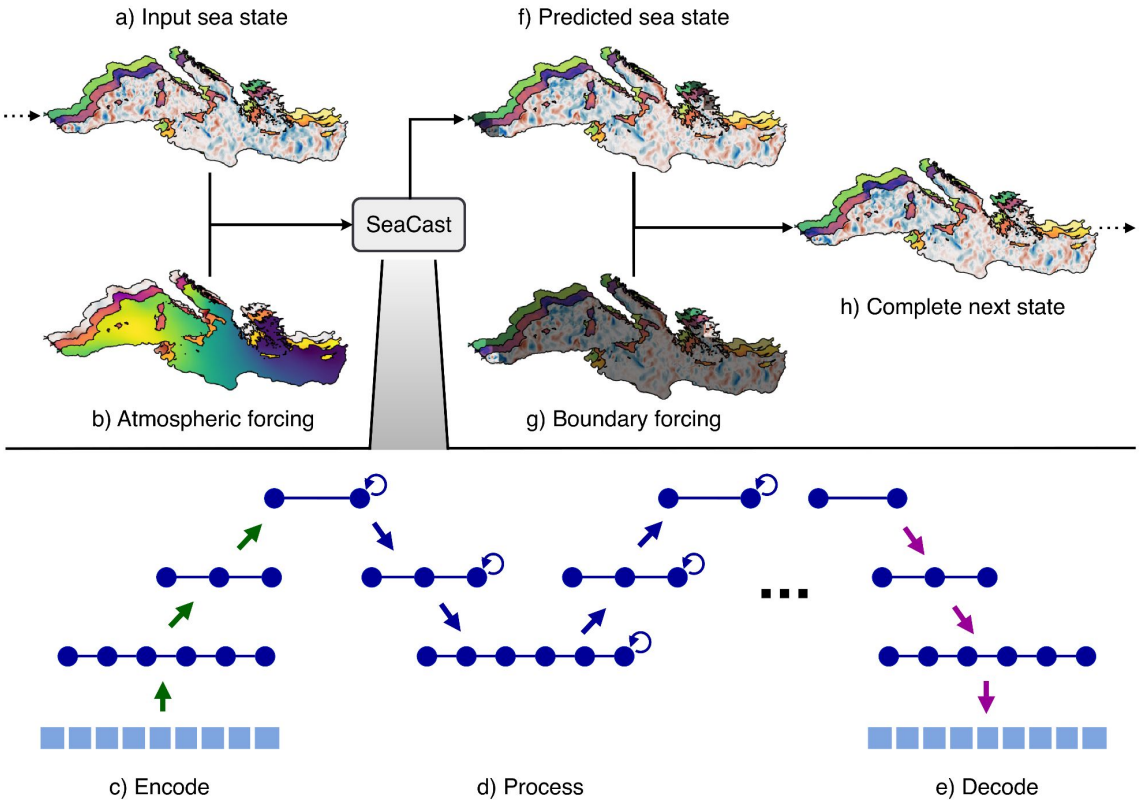
Data splits

- Train on daily 35 years of reanalysis data, finetune on 2 years of analysis data.
- Daily validation spring 2024 and test on fall 2024.
    - (test set very recent because we use the new 15-day *AIFS* weather forecast from ECMWF as forcing)



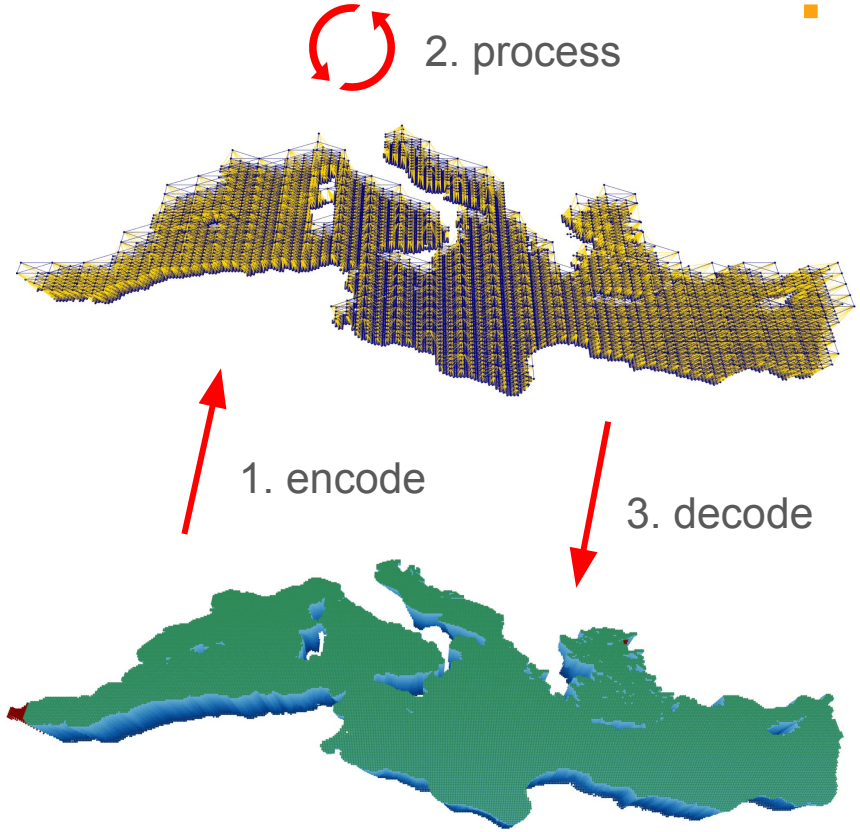
https://marine.copernicus.eu/

Model Overview

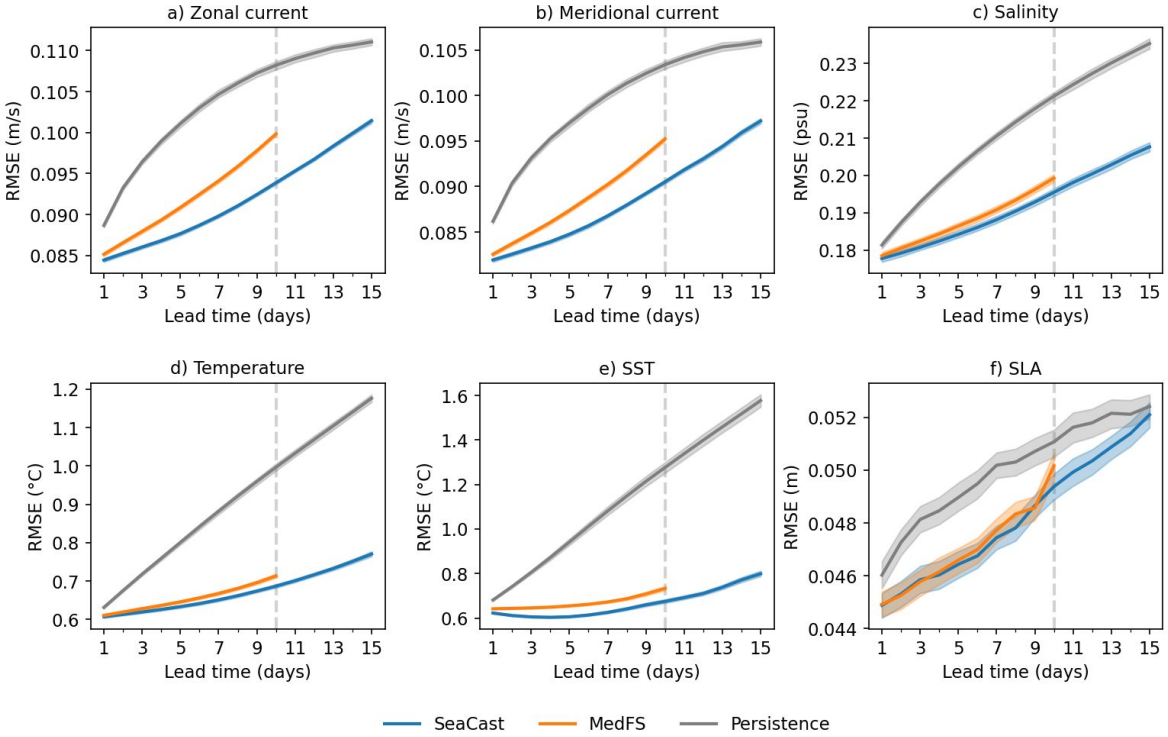


Mesh Structure

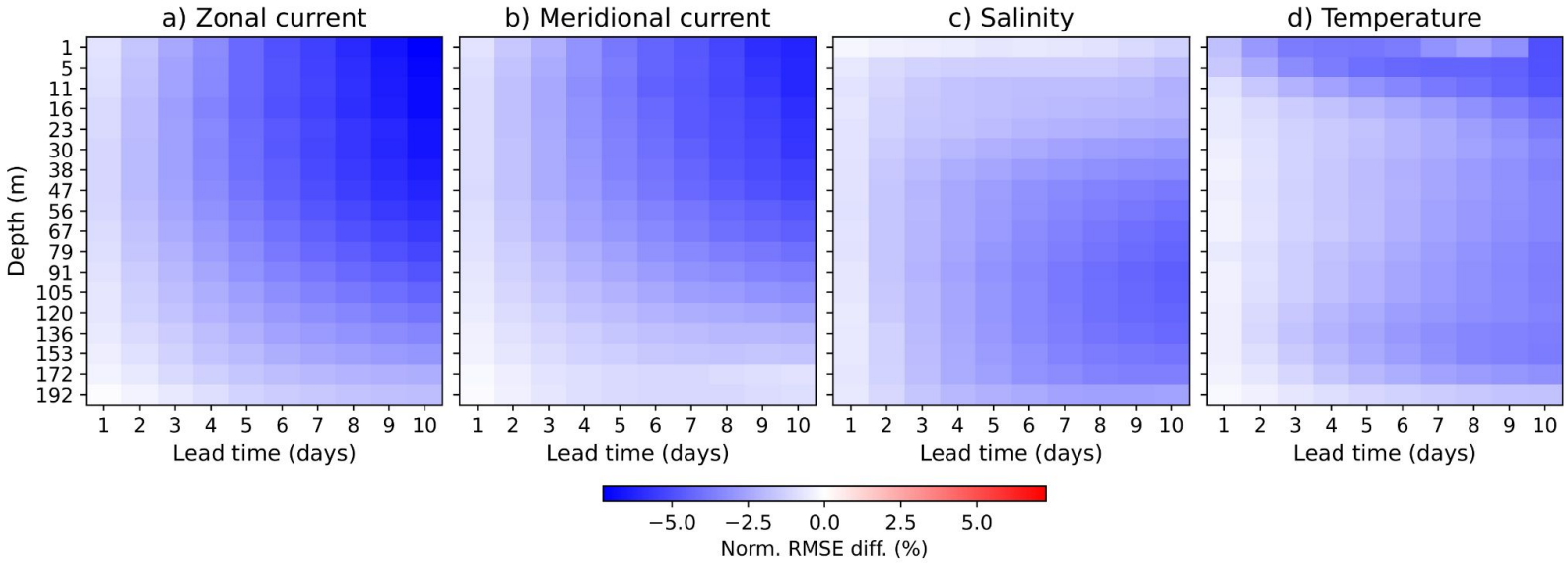
- Quadrilateral mesh used by the model:
  - Coarser than the data → efficient processing.
  - Nodes are connected with bidirectional edges to its neighbors horizontally, vertically and diagonally (repeated at 3 different resolutions tripling the distance between nodes).
- GNNs are used to:
  1. **Encode** inputs from the data grid to latent vector representations in the mesh.
  2. **Process** latent node and edge representations using yielding new latent representations.
  3. **Decode** onto the original sea grid to predict a new state.



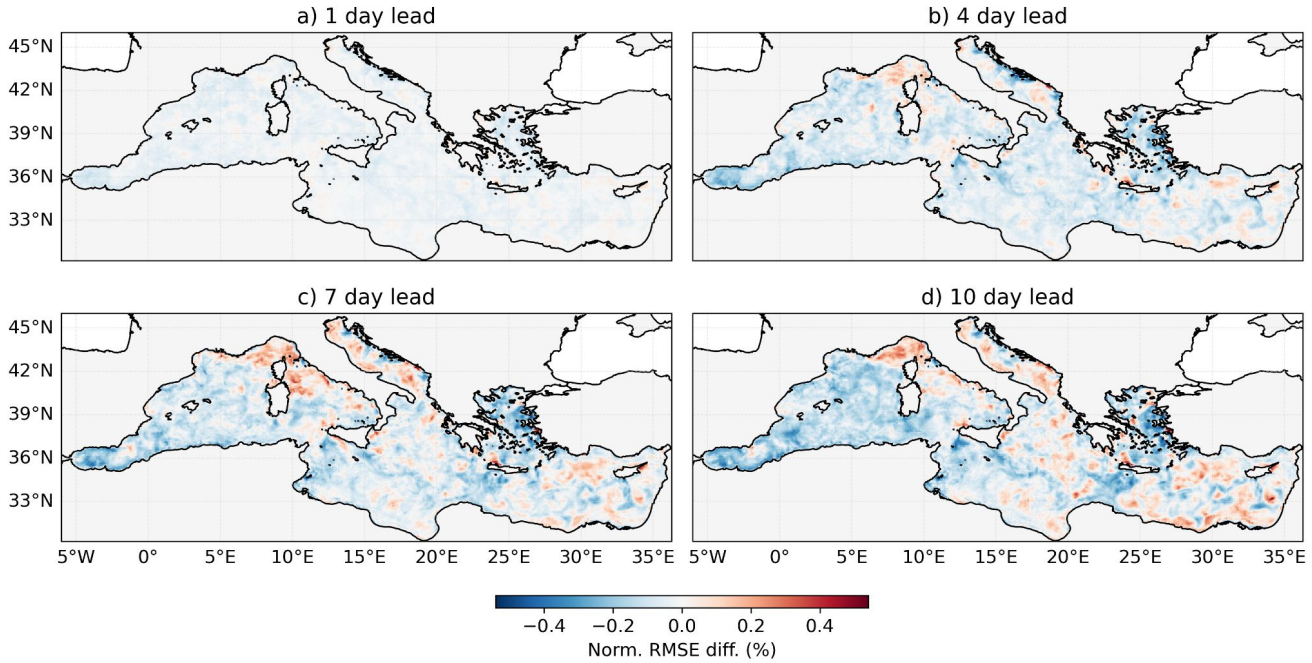
RMSE per Lead Time



RMSE Relative to MedFS

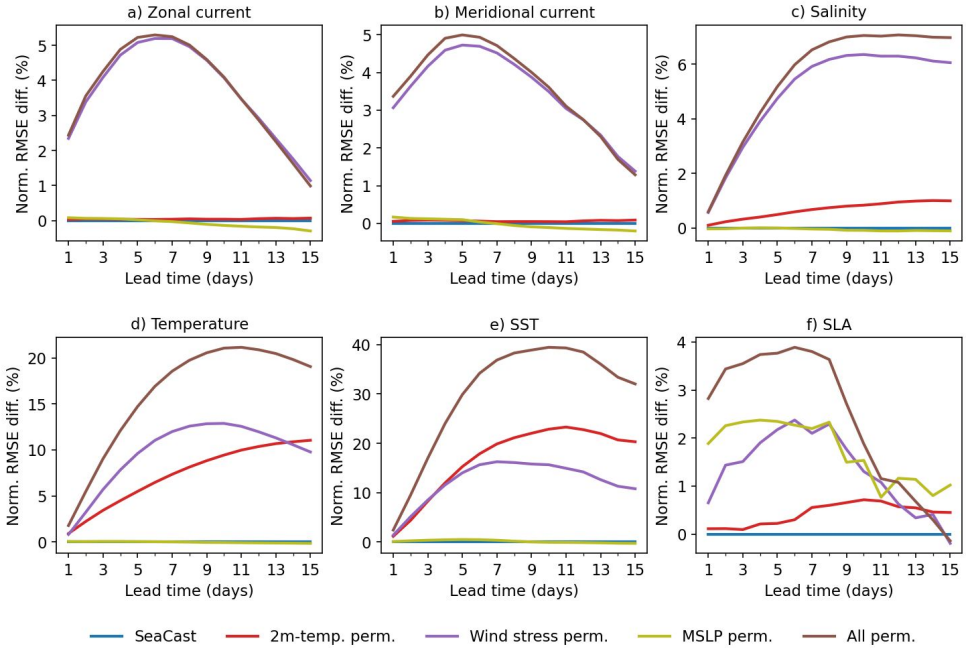


SST RMSE Relative to MedFS

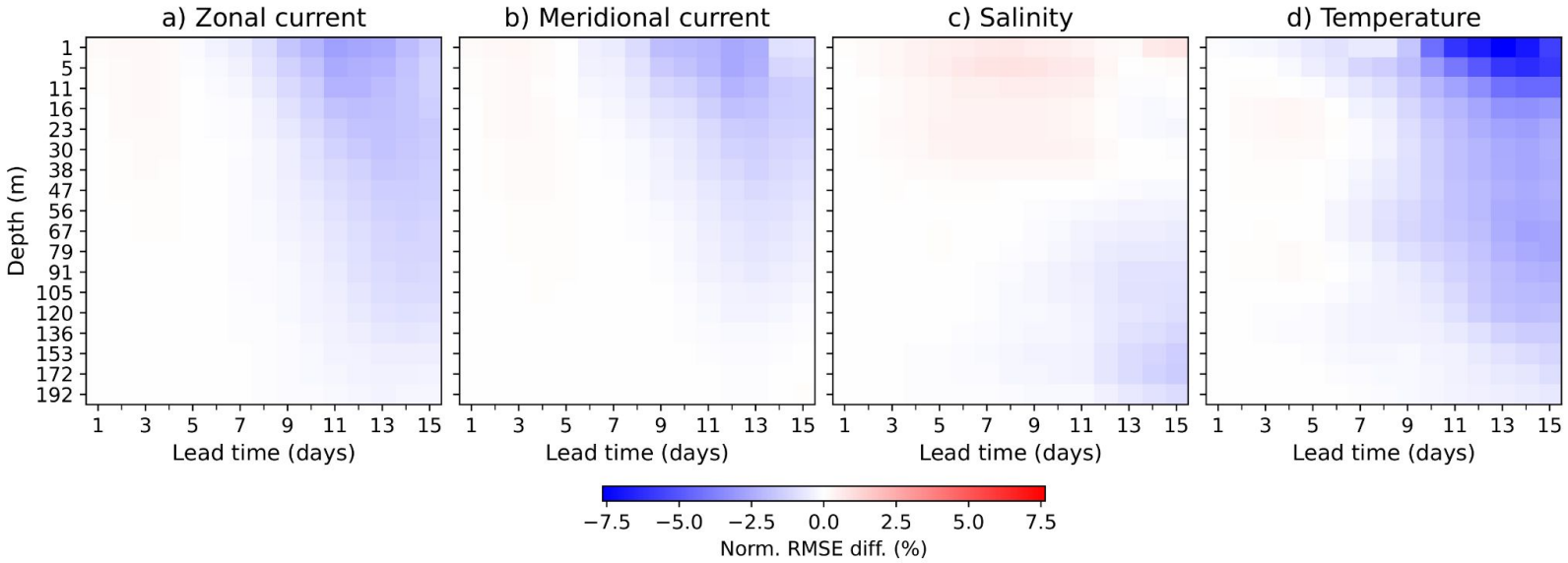


Atmospheric Feature Importance

- Sensitivity study randomly permuting atmospheric variables over grid dimension.



SeaCast-AIFS RMSE Relative to SeaCast-HRES



Summary

- Training 20.5 h on 64 AMD MI250x GPUs.
- Forecast takes 20 s on 1 GPU to generate the 15-day, 1/24° (ca. 4 km) spatial res. using ML, w.r.t. 70 minutes using 89 CPU cores for MedFS.
- Future work:
  - Couple ML forecast with waves, bio-geochemistry, and ice where applicable
  - Probabilistic ensemble forecast
  - Learned boundary forcing
- Paper: climatechange.ai/papers/neurips2024/51